A MEASURE-VALUED DIFFUSION PROCESS DESCRIBING THE STEPPING STONE MODEL WITH INFINITELY MANY ALLELES

Kenji HANDA

Department of Mathematics, School of Science, Nagoya University, Chikusa-ku, Nagoya 464-01, Japan

Received 4 April 1989
Revised 7 February 1990

In this paper we formulate the stepping stone model in population genetics as a measure-valued diffusion process. In order to formulate this model we introduce an appropriate martingale problem and show that it is well-posed. In the selectively neutral case an ergodic property of the process corresponding to the solution of the martingale problem is proved, under a suitable assumption on the mechanism of mutation. If in addition the mutation mechanism is of jump type, then simple calculations involving the generator associated with our martingale problem give us equations for the probabilities of identity at equilibrium.

stepping stone model * measure-valued diffusion * martingale problem * duality * equilibrium probability of identity

1. Introduction

In population genetics theory many stochastic models have been studied. The stepping stone model which we discuss in this paper is one of these models and is concerned with geographical structure. Migration has been taken into consideration by Kimura (1953) and genetic systems with geographical structure have also been studied from a mathematical point of view (cf. Cox and Griffeath, 1987).

The model we propose here is intuitively illustrated as follows. Let $S$ be a countable set of colonies and $\{X_1, \ldots, X_d\}$ be the set of all possible alleles or types of individuals. For each colony $k \in S$, the frequencies of types are described by the $d$-dimensional vector $\vec{p}_k = (p_{1k}, \ldots, p_{dk})$, where $p_{ik}$ represents the frequency of the individuals of the type $X_i$ in colony $k$ and so $p_{1k} + \cdots + p_{dk} = 1$.

Set

$$\Sigma = \left\{ P = \{p_{ik}\}_{k \in S, 1 \leq i \leq d} ; 0 \leq p_{ik} \leq 1, \sum_{i=1}^{d} p_{ik} = 1, k \in S \right\}. \quad (1.1)$$
The stepping stone model is originally a discrete time $\Sigma$-valued Markov chain, which evolves under the mechanisms of mutation, selection and migration.

Sato (1983) has given diffusion approximations for stepping stone models and Shiga (1982) discussed the properties (e.g. stationary distributions) of the limiting diffusion processes.

For genetic systems without geographical structures, measure-valued diffusion processes are introduced by Fleming and Viot (1979) and developed by Ethier and Kurtz (1987), etc. Behind the formulation of these diffusion processes is the fact that the number of alleles is very large, and they describe the models with infinitely many alleles. Indeed the set of types $E$ is allowed to be a compact metric space, and any possible distribution of types in the population is described by a probability measure $\mu$ on $E$ (cf. Hochberg, 1986).

The first aim of this paper is to formulate the stepping stone model with infinitely many alleles as a measure-valued diffusion process solving a certain martingale problem. Roughly speaking, it is a continuum limit for measure-valued processes obtained from the original stepping stone models by identifying the vector $p^{(k)}$ with the measure $\sum_{i=1}^d p^{(k)}_i \delta_{x_i}$, where $\delta_x$ denotes the $\delta$-distribution at $x$. The generator associated with the martingale problem is a generalization of the generator discussed by Ethier and Kurtz (1987), and it will be defined in the next section.

In Section 3 we shall prove the well-posedness of the martingale problem. The method of proof is similar to that used by Ethier and Kurtz (1987). Our situation, however, needs some lemmas concerned with the geographical structure.

Section 4 is devoted to the proof of Theorem 3.5 already used in Section 3 to give the uniqueness of the solution of the martingale problem. The theorem asserts the duality relationship between the measure-valued process and another stochastic process. It can be regarded as a generalization of the duality shown by Ethier and Kurtz (1987).

The model described in our formulation should of course approximate the discrete model and be tractable for certain purposes. The former will be shown in the proof of the existence theorem. The latter can be found in the final section, where we consider the following special case. The mutation mechanism is of jump type and is assumed to satisfy a suitable convergence condition to the equilibrium. It is then shown in the selectively neutral case that the measure-valued process has a unique equilibrium state.

Another aim of this paper is to investigate the equilibrium state. We are interested in a certain quantity of the model. It is called the 'equilibrium probability of identity' and has been discussed by Maruyama (1977), Nagylaki (1983), and by other authors. Let $I_{kl}$ denote the probability at equilibrium that two distinct individuals chosen at random, one from colony $k$ and the other from colony $l$, are of the same type. Equations for $\{I_{kl}\}$ will easily be obtained by calculations involving the generator associated with our martingale problem. Moreover in the case where the array of colonies is the integer lattice $\mathbb{Z}^d$ and where migration is spatially homogeneous, we shall prove that the system $\{I_{kl}\}$ is characterized as a unique solution of the equations.
in a certain class, and the effect of distance between colonies on the decrease of $I_{kl}$ is investigated.

2. Martingale problem

To find an appropriate formulation, we begin by reviewing the finite (say $d$) alleles stepping stone model of the type discussed by Shiga (1982). Let $S$ be the countable set of the colonies and define $\Sigma$ as in (1.1). Consider a $\Sigma$-valued diffusion processes \( \{P(t); t \geq 0\} \) corresponding to the solution of the martingale problem for the generator

\[
G = \sum_{k \in S} \sum_{i=1}^{d} \left\{ L_i(\bar{p}^{(k)}) + H_i(\bar{p}^{(k)}) + \sum_{k' \in S} m_{k'k} p_i^{(k)} \right\} \frac{\partial}{\partial p_i^{(k)}} \\
+ \frac{1}{2} \sum_{k \in S} \sum_{i,j=1}^{d} p_i^{(k)}(\delta_{ij} - p_i^{(k)}) \frac{\partial^2}{\partial p_i^{(k)} \partial p_j^{(k)}}
\]

where $\bar{p}^{(k)} = (p_1^{(k)}, \ldots, p_d^{(k)})$,

\[
L_i(\bar{p}) = \sum_{j=1}^{d} \theta_{ij} p_j,
\]

\[
H_i(\bar{p}) = p_i \left( \sum_{j=1}^{d} \sigma_{ij} p_j - \sum_{j,l=1}^{d} \sigma_{il} p_j p_l \right) \quad \text{for} \quad \bar{p} = (p_1, \ldots, p_d)
\]

and the coefficients above satisfy

\[
\begin{align*}
\theta_{ij} &\geq 0 \quad \text{for} \quad i \neq j, \quad \theta_{ii} = -\sum_{j \neq i} \theta_{ij}, \\
\sigma_{ij} &= \sigma_{ji} \quad \text{for} \quad i, j = 1, \ldots, d, \\
m_{k'k} &\geq 0 \quad \text{for} \quad k' \neq k,
\end{align*}
\]

and

\[
\sup_{k \in S} |m_{kk}| < \infty \quad \text{with} \quad m_{kk} = - \sum_{k' \neq k} m_{k'k}. \tag{2.2b}
\]

Let $E = \{X_1, \ldots, X_d\}$ be the set of types. Then (2.1) has the following population genetics interpretation. For $i \neq j$, $\theta_{ij}$ represents the mutation rate of $X_i$ to $X_j$. For each $i$ and $j$, $\sigma_{ij}$ is the selection intensity of the pair of $(X_i, X_j)$. For $k \neq k'$, $m_{k'k}$ represents the migration rate from colony $k'$ to colony $k$. The terms involving the second derivatives come from the effect of random mating.

Let $\mathcal{P}(E)$ be the set of probability measures on $E$ and set

\[
\tilde{\mathcal{P}} = \mathcal{P}(E)^\circ = \{\tilde{\mu} = \{\mu_k\}_{k \in S}; \mu_k \in \mathcal{P}(E), k \in S\}.
\]

Considering the one to one mapping $\xi : \Sigma \rightarrow \tilde{\mathcal{P}}$ such that

\[
\xi P = \left\{ \sum_{i=1}^{d} p_i^{(k)} \delta_{X_i} \right\}_{k \in S} \quad \text{for} \quad P = \{p_i^{(k)}\} \in \Sigma,
\]
we can identify \( \{P(t); t \geq 0\} \) with the \( \mathcal{P} \)-valued process

\[
\{\tilde{\mu}(t) = \{\mu_k(t)\}_{k \in S}; \ t \geq 0\}
\]

given by

\[
\mu_k(t) = \sum_{i=1}^{d} p^{(k)}_i(t) \delta_X.
\]

Define the function \( \Phi \) on \( \tilde{\mathcal{P}} \) by

\[
\Phi(\tilde{\mu}) = \prod_{i=1}^{m} \langle f_i, \mu_k \rangle
\]

where \( m \in \mathbb{N} \), \( f_i : E \to \mathbb{R} \), and \( k_i \in S \) are fixed and where \( \langle f, \mu \rangle \) denotes the integral of \( f \) with respect to the measure \( \mu \). Define the function \( g \) on \( \Sigma \) by \( g = \Phi \circ \xi \). Straightforward calculation then shows

\[
Gg(P) = \sum_{i=1}^{m} \left( \langle Lf_i, \mu_{k_i} \rangle + \langle \sigma \cdot f_i, \mu_{k_i} \rangle - \langle \sigma, \mu_{k_i} \times \mu_{k_i} \rangle \langle f_i, \mu_{k_i} \rangle \right)
\]

\[
+ \sum_{k' \in S} m_{k,k'} \langle f_i, \mu_{k} \rangle \prod_{j \neq i} \langle f_j, \mu_{k} \rangle
\]

\[
+ \sum_{i < j < m \atop k_i = k_j} (\langle f_i f_j, \mu_{k_i} \rangle - \langle f_i, \mu_{k_i} \rangle \langle f_j, \mu_{k_j} \rangle) \prod_{i \neq i, j} \langle f_i, \mu_{k_i} \rangle
\]

(2.3)

where

\[
Lf(X_i) = \sum_{j=1}^{d} \theta_{ij} f(X_j) \quad \text{for} \ f : E \to \mathbb{R},
\]

\[
\sigma(X_i, X_j) = \sigma_{ij}, \quad (\sigma \cdot f)(X, Y) = \sigma(X, Y) f(Y) \quad \text{for} \ X, Y \in E,
\]

and \( \{\mu_k\}_{k \in S} = \xi P \).

Note that the right hand side of (2.3) could make sense for more general sets \( E \).

In the rest of this paper, we allow the set \( E \) of possible types to be a compact metric space. \( \mathcal{P}(E) \) is then the set of Borel probability measures on \( E \) with the \( w^* \)-topology and \( \tilde{\mathcal{P}} = \mathcal{P}(E)^S \) is equipped with the product topology. Set

\[
B(E) = \{\text{real-valued bounded Borel functions on } E\}
\]

and

\[
C(E) = \{\text{continuous functions on } E\}.
\]

For \( m \in \mathbb{N} \), \( E^m \) denotes the \( m \)-fold product of \( E \). \( B(E^m) \), \( C(E^m) \) and \( \mathcal{P}(E^m) \) are defined similarly. We use \( \| \cdot \| \) and \( \langle \cdot , \cdot \rangle \) as the notation for the sup norm and the integration on any space respectively.

Following Ethier and Kurtz (1987), to describe the mechanisms of mutation and selection we consider a linear operator \( L \) on \( C(E) \) (‘mutation operator’) which generates a Feller semigroup \( \{T(t)\}_{t \geq 0} \) on \( C(E) \) and a symmetric function \( \sigma \) in
For migration rates \( \{m_{k,k'}\} \), we impose (2.2a, b) and

\[
M^* = \sup_{k \in S} \left| \sum_{k' \neq k} m_{k,k'} \right| < \infty. \tag{2.4}
\]

Before defining the martingale problem, we observe the following.

**Lemma 2.1.** The subspace of \( C(\mathcal{P}) \),

\[
\{ \Phi; \Phi(\tilde{\mu}) = F((f_1, \mu_{k_1}), \ldots, (f_m, \mu_{k_m})), m \in \mathbb{N}, F \text{ is a polynomial with } m \text{ variables}, f_i \in \mathcal{D}(L), k_i \in S \},
\]

is dense in \( C(\mathcal{P}) \) in the sense of sup norm, where \( \mathcal{D}(L) \) is the domain of \( L \).

**Proof.** The proof is essentially the same as that of Lemma 2 of Fleming and Viot (1979) and is omitted. \( \square \)

For \( m \in \mathbb{N}, f_i \in \mathcal{D}(L), \) and \( k_i \in S \) \((i = 1, \ldots, m)\), set

\[
f = (f_1, \ldots, f_m) \in \mathcal{D}(L)^m, \quad k = (k_1, \ldots, k_m) \in S^m,
\]

and define a function \( \Phi_{f,k} \) on \( \mathcal{P} \) by

\[
\Phi_{f,k}(\tilde{\mu}) = \prod_{i=1}^{m} (f_i, \mu_{k_i}), \quad \tilde{\mu} = \{\mu_k\}_{k \in S} \in \mathcal{P}.
\]

Let \( \Psi_{f,k} \) be the function of \( \tilde{\mu} \) having the same form as the right hand side of (2.3). Define \( A \subset C(\mathcal{P}) \times C(\mathcal{P}) \) by

\[
A = \{ (\Phi_{f,k}, \Psi_{f,k}); m \in \mathbb{N}, f \in \mathcal{D}(L)^m, k \in S^m \}.
\]

\( C([0, \infty), \mathcal{P}) \) denotes the space of continuous functions \( \omega : [0, \infty) \rightarrow \mathcal{P} \) with the topology of uniform convergence on compact subsets of \([0, \infty)\). \( D([0, \infty), \mathcal{P}) \) denotes the space of right continuous functions \( \omega : [0, \infty) \rightarrow \mathcal{P} \) with left limits with the Skorokhod topology. The coordinate process \( \{\tilde{\mu}(t) = (\mu_k(t))_{k \in S}; t > 0\} \) on \( \Omega = C([0, \infty), \mathcal{P}) \) (or \( D([0, \infty), \mathcal{P}) \)) is defined by

\[
\tilde{\mu}(t)(\omega) = \omega(t), \quad \omega \in \Omega,
\]

and we set the \( \sigma \)-fields

\[
\mathcal{F} = \sigma(\tilde{\mu}(s); s > 0) \quad \text{and} \quad \mathcal{F}_t = \sigma(\tilde{\mu}(s); 0 < s \leq t) \text{ for } t > 0.
\]

We define the martingale problem in the form due to Ethier and Kurtz (1986).

**Definition 2.2.** Given \( \tilde{\mu}^0 \in \mathcal{P} \), we call a probability measure \( P \) on \( (C([0, \infty), \mathcal{P}), \mathcal{F}) \) (or \( (D([0, \infty), \mathcal{P}), \mathcal{F}) \)) a solution of the \( C([0, \infty), \mathcal{P}) \) (respectively \( D([0, \infty), \mathcal{P}) \)) martingale problem for \( (A, \tilde{\mu}^0) \) if

\[
P(\tilde{\mu}(0) = \tilde{\mu}^0) = 1,
\]
and if
\[
\Phi(\mu(t)) - \int_0^t \Psi(\mu(s)) \, ds
\]
is a \((P, \{\mathcal{F}_t\})\)-martingale for all \((\Phi, \Psi) \in A\).

3. Existence and uniqueness theorems

In this section we prove that the martingale problem defined in the previous section is well-posed. The idea in Ethier and Kurtz (1987) plays an important role in the following argument.

3.1. Existence

First we discuss the existence of solutions of the martingale problem for \(A\). Solutions will be obtained as limits in distribution of a sequence of measure-valued stepping stone models.

From the assumption on the mutation operator \(L\), there exists a sequence of transition functions \(\{p_N(x, dy)\}\) on \(E \times \mathcal{P}(E)\) such that

for all \(f \in \mathcal{D}(L)\), where \(Q_Nf(x) = \int_E f(y)p_N(x, dy)\). (See Ethier and Kurtz (1986), Lemma 4.5.3.)

For each integer \(N > N_0 = \max\{M^*, \|\sigma\|\}\), we can construct a Markov chain \(X^N(\tau) = \{(X^{N,k}_k(\tau), \ldots, X^{N,k}_K(\tau))\}_{k \in S}\), \(\tau = 0, 1, 2, \ldots\), in \((E^N)^S\) as follows. Given \(X^N(\tau)\), the components of \(X^N(\tau + 1)\) are mutually conditionally independent, that is

\[
P(X^N(\tau + 1) \in dy | X^N(\tau) = x) = \prod_{k \in S} \sum_{\sigma_{k,\alpha}^{(N)}} P(X_{k,\alpha}^{N}(\tau + 1) \in dy_{k,\alpha} | X^N(\tau) = x)
\]

for \(x, y = \{(y_{k,1}, \ldots, y_{k,N})\} \in (E^N)^S\). For each \(k \in S\), each of \(X^N_{k,1}(\tau + 1), \ldots, X^N_{k,N}(\tau + 1)\) has an identical conditional distribution given \(X^N(\tau) = x\), which is given by

\[
P(X^N_{k,1}(\tau + 1) \in dy | X^N(\tau) = x) = \sum_{k \in S} \sum_{\alpha = 1}^N \sigma^{(N)}_{k,\alpha}(x) p_N(x_{k,\alpha}, dy)
\]

where

\[
n^{(N)}_{kk'} = \frac{m^{(N)}_{kk'}}{\sum_{l \in S} m^{(N)}_{lk}}, \quad m^{(N)}_{lk} = \begin{cases} N^{-1} m_{lk} & \text{if } l \neq k, \\ 1 - N^{-1} \sum_{k' \neq k} m_{kk'} & \text{if } l = k, \end{cases}
\]
and where
\[ \sigma^{(N)}_{k,\alpha}(x) = \sum_{\beta=1}^{N} \frac{W_{N}(x_{k,\alpha}, x_{k,\beta})}{N} \sum_{\rho, \gamma=1}^{N} W_{N}(x_{k,\beta}, x_{k,\gamma}) \]
for \( x \in (E^N)^S \)
with the notation \( W_{N}(x, y) = 1 + N^{-1}\sigma(x, y) \) for \( x, y \in E \).

These Markov chains are considered as stepping stone models involving selection and mutation. Each component \( X_{k,\alpha}(T) \) is viewed as the type of the \( \alpha \)th individual belonging to colony \( k \) in generation \( T \). Each individual in generation \( T+1 \) independently selects a parent from generation \( T \). For each of the \( N \) individuals belonging to colony \( k \), the probability that its parent has migrated from the colony \( k' \) to colony \( k \) is \( n_{kk'}(N) \), and in this case the \( \alpha \)th individual in colony \( k' \) in generation \( T \) is selected as the parent with probability \( \sigma^{(N)}_{k,\alpha}(x) \) where \( x \) is the vector of types of the individuals in generation \( T \). If the parent is of type \( x \), then the offspring's type belongs to \( A \) with probability \( p_{N}(x, \Lambda) \).

Define a mapping \( \tilde{\rho}_{N} : (E^N)^S \rightarrow \tilde{\varnothing} \) by
\[ \tilde{\rho}_{N}(x) = \left\{ N^{-1} \sum_{\alpha=1}^{N} \delta_{s_{k,\alpha}} \right\}_{k \in S}, \quad x = \{(x_{k,1}, \ldots, x_{k,N})\}_{k}. \]
The range of \( \tilde{\rho}_{N} \) is denoted by \( \varnothing_{N} \). Set \( \tilde{v}^{N}(\tau) = \tilde{\rho}_{N}(X^{N}(\tau)) \). Then \( \tilde{v}^{N}(\tau), \tau = 0, 1, 2, \ldots, \) is a Markov chain in \( \varnothing_{N} \). The measure-valued process \( \{\tilde{\mu}(t); t \geq 0\} \) which we will obtain as a solution of the martingale problem for \( A \) can be shown to approximate these Markov chains in the sense that a subsequence of \( \{\tilde{\mu}^{N}(t); t \geq 0\}_{N > N_0} \) defined by \( \tilde{\mu}^{N}(t) = \tilde{v}^{N}(\lfloor Nt \rfloor) \) converges to it. To see this, we first verify the tightness of the \( \varnothing_{N} \)-valued processes \( \{\tilde{\mu}^{N}(t); t \geq 0\}_{N > N_0} \).

**Lemma 3.1.** For \( f, g \in \mathcal{D}(L) \) and \( k, k' \in S \),
\[ \mathcal{N} \cdot \{E[f(X_{k,1}(1))]|X^{N}(0) = x] - \langle f, \mu_{k} \rangle \}
= \langle Lf, \mu_{k} \rangle + \langle \sigma \cdot f, \mu_{k} \times \mu_{k} \rangle - \langle \sigma, \mu_{k} \times \mu_{k} \rangle(f, \mu_{k}) + \sum_{k' \in S} m_{k,k'}(f, \mu_{k}) + o(1), \]
\[ \mathcal{N} \cdot E[f(X_{k,1}(1)) \cdot g(X_{k',1}(1))]|X^{N}(0) = x]
= \begin{cases} \langle f, \mu_{k} \rangle \langle g, \mu_{k} \rangle + o(1) & \text{if } k \neq k', \\ \langle fg, \mu_{k} \rangle + o(1) & \text{if } k = k', \end{cases} \]
as \( N \to \infty \), uniformly in \( \tilde{\mu} = \{\mu_{k}\}_{k \in S} = \tilde{\rho}_{N}(x) \in \varnothing_{N} \).

**Proof.** From the definition of \( X^{N}(\tau) \) and (3.1), one can easily obtain these estimates by direct calculations. \( \square \)

Using the above estimates and arguing as in the proof of Theorem 10.4.1 of Ethier and Kurtz (1986), we have
\[ \mathcal{N} \cdot E[\Phi(\tilde{v}^{N}(1)) - \Phi(\tilde{v}^{N}(0))|\tilde{v}^{N}(0) = \tilde{\mu}] = \Psi(\tilde{\mu}) + o(1) \] (3.2)
as \( N \rightarrow \infty \), uniformly in \( \mu \in \mathcal{F}_N \) for all \((\Phi, \Psi) \in A\). Applying Theorem 3.9.1 and Theorem 3.8.6 of Ethier and Kurtz (1986) and using Lemma 2.1 and (3.2), we can verify that the family of processes \( \{\mu^N(t); t \geq 0\}_{N \geq N_0} \) is tight in \( D([0, \infty), \mathcal{F}) \). Furthermore arguing about the limit point, we get an existence theorem as follows.

**Theorem 3.2.** For any \( \mu^0 \in \mathcal{F} \), there exists a solution of the \( C([0, \infty), \mathcal{F}) \) martingale problem for \((A, \mu^0)\).

**Proof.** Choose a sequence \( \{\mu^N_0\}_N \) such that \( \mu^N_0 \in \mathcal{F}_N \) and that \( \mu^N_0 \rightarrow \mu^0 \) in \( \mathcal{F} \) as \( N \rightarrow \infty \). We consider the processes \( \{\mu^N(t); t \geq 0\}_{N \geq N_0} \) with \( \mu^N(0) = \mu^N_0 \). From the tightness we can take a process \( \{\tilde{\mu}(t); t \geq 0\} \) to which a subsequence of \( \{\mu^N(t); t \geq 0\}_{N \geq N_0} \) converges weakly in \( D([0, \infty), \mathcal{F}) \). Since (3.2) implies that for \( 0 \leq s \leq t \) and \((\Phi, \Psi) \in A\), the sequence

\[
\left| E \left[ \Phi(\tilde{\mu}^N(t)) - \Phi(\tilde{\mu}^N(s)) - \int_s^t \Psi(\tilde{\mu}^N(u)) \, du \right] \right|, \quad N > N_0,
\]

can be dominated by a deterministic sequence converging to 0, where \( \mathcal{F} = \sigma(\tilde{\mu}^N(u); 0 \leq u \leq s) \), we conclude that \( P \), the law of \( \{\tilde{\mu}(t); t \geq 0\} \) on \( D([0, \infty), \mathcal{F}) \), solves the martingale problem for \((A, \mu^0)\). (See Ethier and Kurtz (1986), Lemma 4.5.1.)

It remains only to show that \( P \) has a support in \( C([0, \infty), \mathcal{F}) \). For this purpose we prove

\[
P(\{(f, \mu_k(\cdot)) \in C([0, \infty), \mathbb{R})\}) = 1 \tag{3.3}
\]

for \( f \in \mathfrak{D}(L) \) and \( k \in S \). Consider \( \mathbb{R} \)-valued processes

\[
X(t) = \langle f, \mu_k(t) \rangle, \quad U(t) = \frac{1}{2}(\langle f^2, \mu_k(t) \rangle - \langle f, \mu_k(t) \rangle^2)
\]

and

\[
V(t) = \langle Lf, \mu_k(t) \rangle + \langle \sigma \cdot f, \mu_k(t) \times \mu_k(t) \rangle - \langle \sigma, \mu_k(t) \times \mu_k(t) \rangle \langle f, \mu_k(t) \rangle + \sum_{k' \in S} m_{k'k}(f, \mu_k(t)).
\]

Then for each \( m \in \mathbb{N} \),

\[
X(t)^m - \int_0^t \{m(m-1)U(s)X(s)^{m-2} + mV(s)X(s)^{m-1}\} \, ds
\]

is a \((P, \{\mathcal{F}_t\})\)-martingale. These facts lead us to an estimate of the form

\[
E^P [X(t) - X(s)]^4 \leq C |t - s|^2, \quad 0 \leq s \leq t < \infty,
\]

which implies (3.3). (See Ikeda and Watanade (1981), Corollary to Theorem 1.4.3.) \( \Box \)
The method used to show (3.3) is due to Lemma 2.1 of Ethier and Kurtz (1987).

As was seen above, every solution of the $D([0, \infty), \mathcal{F})$ martingale problem for $A$ has a support in $C([0, \infty), \mathcal{F})$. From now on we are only concerned with the $C([0, \infty), \mathcal{F})$ martingale problem and call it simply the martingale problem for $A$.

### 3.2. Dual process

To prove uniqueness we introduce a dual process. We start with a rephrasing of the definition of the function $\Psi_{f,k}$ (the right hand side of (2.3)), which suggests how to define the dual process and is expressed in the form

$$\Psi_{f,k}(\hat{\mu}) = \langle L^{(m)} f, \hat{\mu} \rangle + \hat{\sigma} \sum_{i=1}^{m} \langle K_{im} f, \hat{\mu}_{\alpha,k} \rangle$$

$$+ \sum_{i=1}^{m} \sum_{k' \in S} m_{k,k'} \langle f, \hat{\mu}_{\gamma,(k')k} \rangle$$

$$+ \sum_{1 \leq i < j \leq m} ((\Phi_{ij} f, \hat{\mu}_{\beta,k}) - \langle f, \hat{\mu}_{k} \rangle)$$

(3.4)

for $f = (f_1, \ldots, f_m) \in \mathcal{D}(L)^m$, $k = (k_1, \ldots, k_m) \in S^m$, and $\hat{\mu} = (\mu_k)_{k \in S} \in \mathcal{F}$, where we identify the vector $f$ with the function $f(x_1, \ldots, x_m) = \prod_{i=1}^{m} f_i(x_i)$ in $C(E^m)$ to define the functions and the measures in the right hand side of (3.4) in the following manner. Let $L^{(m)}$ be a linear operator on $C(E^m)$ which generates the Feller semigroup $\{T(t)\}$ corresponding to $m$ independent copies of the processes corresponding to $\{T_n(t)\}$. In particular,

$$L^{(m)} f(x_1, \ldots, x_m) = \sum_{i=1}^{m} L f_i(x_i) \prod_{j \neq i} f_j(x_j)$$

for $f(x_1, \ldots, x_m) = \prod_{i=1}^{m} f_i(x_i)$ with $f_i \in \mathcal{D}(L)$, $i = 1, \ldots, m$. For $f \in C(E^m)$ and $1 \leq i < j \leq m$, let $\Phi_{ij} f$ be the function in $C(E^{m-1})$ obtained from $f$ by replacing $x_j$ by $x_i$ and renumbering the variables, more explicitly

$$\Phi_{ij} f(x_1, \ldots, x_{m-1}) = f(x_1, \ldots, x_j, x_i, x_j, \ldots, x_{m-1}).$$

Define $K_{im} : C(E^m) \to C(E^{m+2})$ for $1 \leq i < j \leq m$ by

$$K_{im} f(x_1, \ldots, x_{m+2}) = \hat{\sigma}^{-1} \{ \sigma(x_i, x_{m+1}) - \sigma(x_{m+1}, x_{m+2}) \} f(x_1, \ldots, x_m)$$

where $\hat{\sigma}$ is a positive constant satisfying

$$\hat{\sigma} \geq \sup \{ |\sigma(x, y) - \sigma(y, z)|; x, y, z \in E \}.$$

For $k = (k_1, \ldots, k_m) \in S^m$, define

$$\alpha_i k = (k_1, \ldots, k_{i-1}, k_i, k_{i+1}, \ldots, k_m) \in S^{m+2} \quad \text{for} \ 1 \leq i \leq m,$$

$$\beta_j k = (k_1, \ldots, k_{j-1}, k_{j+1}, \ldots, k_m) \in S^{m-1} \quad \text{for} \ 2 \leq j \leq m,$$

$$\gamma_{i}(k') k = (k_1, \ldots, k_{i-1}, k', k_{i+1}, \ldots, k_m) \in S^m \quad \text{for} \ k' \in S, \ 1 \leq i \leq m,$$
Having established the relation (3.4), we are now ready to construct the dual process \( \{(Y(t), k(t)); \ t \geq 0\} \) in \( \bigcup_{m=1}^{\infty} (C(E^m) \times S^m) \). First let \( K = \{k(t); \ t \geq 0\} \) be a Markov chain in \( S^* = \bigcup_{m=1}^{\infty} S^m \) whose generator \( \mathcal{L} \) is of the form

\[
\mathcal{L} h(k) = \bar{\sigma} \sum_{i=1}^{[k]} \{h(\alpha_i k) - h(k)\} + \sum_{j=2}^{[k]} \#\{i; 1 \leq i < j, k_i = k_j\} \cdot \{h(\beta_j k) - h(k)\}
\]

\[+ \sum_{i=1}^{[k]} \sum_{k' : k_i = k'_{k_i}} m_{k_i k_i} \{h(\gamma_i(k') k) - h(k)\},\]

where \( k_i \) is the \( i \)th component of \( k \) and \( |k| \) is a positive integer \( m \) such that \( k \in S^m \). In this connection, \( \{\tau_n\} \) denotes the sequence of jump times of \( K \) (take \( \tau_0 = 0 \)) and we set \( M(t) = |k(t)| \) for convenience. For each \( k \in S^* \), let \( \lambda(k) \) be the transition rate from \( k \), that is

\[
\lambda(k) = \bar{\sigma} |k| + \#\{(i, j); 1 \leq i < j \leq |k|, k_i = k_j\} + \sum_{i=1}^{[k]} \sum_{k' : k_i = k'_{k_i}} m_{k_i k_i} \cdot \quad (3.5)
\]

Next define random operators \( \{\Gamma_n\}_{n \geq 1} \) which are conditionally independent given \( K \) and satisfy

\[
\mathcal{L} h(k) = \bar{\sigma} |k| + \#\{(i, j); 1 \leq i < j \leq |k|, k_i = k_j\} + \sum_{i=1}^{[k]} \sum_{k' : k_i = k'_{k_i}} m_{k_i k_i} \cdot \quad (3.5)
\]

for \( 1 \leq i \leq m \),

\[
P(\Gamma_n = \Phi_{ij} | K) = \#\{l; 1 \leq l < j, k_l(\tau_n-) = k_j(\tau_n-)\}^{-1}
\]

\[
\cdot 1_{\{k(\tau_n-) = \beta_j k(\tau_n-) \text{ and } k_i(\tau_n-) = k_j(\tau_n-), k_i(\tau_n-) = k_j(\tau_n-)\}}
\]

for \( 1 \leq i < j \), and

\[
P(\Gamma_n = \text{id} | K) = 1_{\{|k(\tau_n-) - |k(\tau_n-)|\}},
\]

where \( 1_A \) denotes the indicator function of an event \( A \).

Finally we define a \( \bigcup_{m=1}^{\infty} C(E^m) \)-valued process \( \{Y(t); \ t \geq 0\} \) by

\[
Y(0) \in C(E^{M(0)}),
\]

\[
Y(t) = T_{M(\tau_n)}(t - \tau_n) Y(\tau_n) \text{ for } t \in [\tau_n, \tau_{n+1}),
\]

\[
(3.6)
\]

and

\[
Y(\tau_{n+1}) = \Gamma_{n+1} T_{M(\tau_n)}(\tau_{n+1} - \tau_n) Y(\tau_n)
\]

\[
(3.7)
\]

for \( n = 0, 1, 2, \ldots \).

**Remark 3.3.** From the definition of the processes \( Y(t) \) and \( k(t) \), we have \( (Y(t), k(t)) \in \bigcup_{m=1}^{\infty} (C(E^m) \times S^m) \) and \( \|Y(t)\| \leq \|Y(0)\| \) for each \( t \geq 0 \).

Recall one of the assumptions on the migration rates \( \{m_{k', k}\} \):

\[
M = \sup_{k \in S} |m_{k', k}| < \infty \quad (3.8)
\]
where $m_{kk} = -\sum_{k' \neq k} m_{k'k}$. This condition ensures the conservativeness of the Markov process $K$, that is

**Lemma 3.4.** For each $k \in S^*$, it holds

$$P_k(\lim_{n \to \infty} \tau_n = \infty) = 1,$$

where $P_k$ denotes the law of the process $\{k(t); t \geq 0\}$ starting from $k$.

**Proof.** For each positive integer $l$, define a stopping time $\rho_l$ by

$$\rho_l = \inf\{t > 0; |k(t)| \geq l\}.$$

Noting that by (3.5) and (3.8) there exists a positive constant $C_1$ such that

$$\lambda(k) \leq C_1 |k|^2 \quad \text{for } k \in S^*, \quad (3.9)$$

we can easily show

$$P_k(\lim_{n \to \infty} \tau_n < \infty \text{ and } \rho_l = \infty) = 0$$

for $l = 1, 2, \ldots$, and hence

$$P_k(\lim_{n \to \infty} \tau_n < \infty \text{ and } \sup_l \rho_l = \infty) = 0.$$

Therefore it is enough to prove

$$P_k(\sup_l \rho_l = \infty) = 1. \quad (3.10)$$

Apply Dynkin's formula to the function $h(k) = \log(|k| \wedge (l + 1))$ (where we used the notation $a \wedge b = \min\{a, b\}$) and observe that $\mathcal{L}h(k) \leq 2\tilde{\sigma}$ holds for all $k$. We then obtain

$$E_k[\log|k(t \wedge \rho_l)|] \leq \log|k| + 2\tilde{\sigma}t$$

for all $t \geq 0$ and $l = 1, 2, \ldots$. Together with

$$E_k[\log|k(t \wedge \rho_l)|] \geq (\log l) P_k(\rho_l \leq t),$$

we have

$$P_k(\sup_l \rho_l \leq t) = \lim_{l \to \infty} P_k(\rho_l \leq t)$$

$$\leq \lim_{l \to \infty} (\log l)^{-1}(\log|k| + 2\tilde{\sigma}t) = 0,$$

which implies (3.10) since $t$ is arbitrary. \qed

**3.3. Uniqueness**

Uniqueness of the solution of the martingale problem for $A$ follows from the duality relationship described in the following theorem, which will be proved in the next section.
Theorem 3.5. Let \((f, k)\) be in \(\bigcup_{m=1}^{\infty} (C(E^m) \times S^m)\) and suppose that there exists \(t_0 > 0\) satisfying
\[
E_k \left[ \exp \left( \sum_0^t \left| k(u) \right| du \right) \right] < \infty. \tag{3.11}
\]
Then for each \(\mu^0 \in \hat{\mathcal{P}}\) every solution \(P\) of the martingale problem for \((A, \mu^0)\) satisfies
\[
P^P \left[ \langle f, \mu(t, k) \rangle \right] = E_{(f, k)} \left[ \langle Y(t), \mu^0_{(k,t)} \rangle \exp \left( \sum_0^t \left| k(u) \right| du \right) \right] \tag{3.12}
\]
for \(0 \leq t \leq t_0\), where \(E_{(f, k)}\) represents the expectation with respect to the process \(\{(Y(t), k(t)); t \geq 0\}\) with \((Y(0), k(0)) = (f, k)\).

Existence of \(t_0\) as above is verified in the next lemma. In its proof we compare \(K = \{k(t); t \geq 0\}\) with a certain birth process in \(\mathbb{N} = \{1, 2, 3, \ldots\}\) by using the coupling technique.

Lemma 3.6. Let \(\{N(t); t \geq 0\}\) be a Markov process in \(\mathbb{N}\) with transition rates \(\hat{\sigma}\) from \(n\) to \(n+2\) for \(n = 1, 2, 3, \ldots\). Then for \(k \in S^*\) and \(t \geq 0\) we have
\[
E_k \left[ \exp \left( \sum_0^t \left| k(u) \right| du \right) \right] \leq \bar{E}_n \left[ \exp \left( \sum_0^t N(u) du \right) \right], \tag{3.13}
\]
where \(\bar{E}_n\) denotes the expectation with respect to the process \(\{N(t); t \geq 0\}\) starting from \(n\). Moreover the right hand side of (3.13) is finite for \(t = t_0 = (4\hat{\sigma})^{-1}\).

Proof. Define a coupled Markov process \(\{(\bar{k}(t), \bar{N}(t)); t \geq 0\}\) in \(X = S^* \times \mathbb{N}\) with generator \(\bar{\mathcal{P}}\) given by
\[
\bar{\mathcal{P}} h(k, n) = \hat{\sigma} \sum_{i=1}^{\left| k \right| \land n} \{h(\alpha, k, n+2) - h(k, n)\}
+ \hat{\sigma} \sum_{i=\left| k \right| \land n+1}^{\left| k \right|} \{h(\alpha, k, n) - h(k, n)\}
+ \sum_{j=2}^{\left| k \right|} \# \{i; 1 \leq i < j, k_i = k_j\} \cdot \{h(\beta, k, n) - h(k, n)\}
+ \sum_{i=1}^{\left| k \right|} \sum_{k' = k_i} \{h(\gamma_i(k'), k, n) - h(k, n)\}
+ \hat{\sigma} (n - \left| k \right| \land n) \{h(k, n+2) - h(k, n)\}.
\]
The conservativeness of this process follows by the same argument as in the proof of Lemma 3.4.

Let \(h_1\) and \(h_2\) be functions on \(S^*\) and on \(\mathbb{N}\) respectively. If we set \(\bar{h}_1(k, n) = h_1(k)\) and \(\bar{h}_2(k, n) = h_2(n)\), then \(\bar{\mathcal{P}} \bar{h}_1(k, n) = \mathcal{L} h_1(k)\) and \(\bar{\mathcal{P}} \bar{h}_2(k, n) = \mathcal{L} h_2(n)\) hold for \((k, n) \in X\), where \(\mathcal{L}_2\) is the generator of \(\{N(t); t \geq 0\}\). This observation implies that
the marginal distributions of \{((\bar{k}(t), \bar{N}(t)); t \geq 0\} with respect to \{k(t); t \geq 0\} and 
\{\bar{N}(t); t \geq 0\} coincide with the distributions of \{k(t); t \geq 0\} and 
\{N(t); t \geq 0\} respectively. So we shall prove

\[ P_{(k,n)}(|\bar{k}(t)| \leq \bar{N}(t) \text{ for all } t \geq 0) = 1 \]  
(3.14)

for \((k, n) \in X\) with \(|k| \leq n\), which gives (3.13) immediately.

Apply Dynkin's formula to the function \(h(k, n) = |k| - |k| \wedge n\) and observe that
\(\mathcal{L}h(k, n) \leq 2\bar{\sigma} h(k, n)\) holds for all \((k, n) \in X\). We then obtain

\[ E_{(k,n)}[h(\bar{k}(t \wedge \bar{\rho}_l), \bar{N}(t \wedge \bar{\rho}_l))] \]

\[ \leq h(k, n) + 2\bar{\sigma} \int_0^t E_{(k,n)}[h(s \wedge \bar{\rho}_l), \bar{N}(s \wedge \bar{\rho}_l))] \, ds \quad \text{for } t \geq 0, \]

where \(\bar{\rho}_l = \inf\{t > 0; \bar{k}(t) \in \bigcup_{m=1}^{\infty} S^m \text{ or } \bar{N}(t) = l\}, l = 1, 2, \ldots\).

If \(|k| \leq n\), then the above inequality implies

\[ E_{(k,n)}[h(\bar{k}(t \wedge \bar{\rho}_l), \bar{N}(t \wedge \bar{\rho}_l))] = 0 \]

since \(h(k, n) = 0\) and \(h\) is a nonnegative function. Consequently

\[ P_{(k,n)}(|\bar{k}(t \wedge \bar{\rho}_l)| \leq \bar{N}(t \wedge \bar{\rho}_l)) = 1 \]

holds for all \(t \geq 0\) and \(l = 1, 2, \ldots\). Noting that \(\sup_l \bar{\rho}_l = \infty\) a.s. (this can be shown as was (3.10)) and letting \(l \to \infty\), we obtain

\[ P_{(k,n)}(|\bar{k}(t)| \leq \bar{N}(t)) = 1 \]

for \(t \geq 0\). Therefore (3.14) follows from the right continuity of the paths.

To show the second assertion, observe that

\[ \bar{E}_n \left[ \exp \left( \bar{\sigma} \int_0^t N(u) \, du \right) \right] \leq \bar{E}_n[\exp(\bar{\sigma}tN(t))] \]

\[ = \sum_{r=0}^{\infty} e^{\bar{\sigma}(n+2r)} \bar{P}_n(N(t) = n+2r) = \sum_{r=0}^{\infty} a_r(n, t). \]  
(3.15)

Using the identity

\[ \bar{P}_n(N(t) = n+2r) = \frac{n(n+1) \cdots (n+2(r-1))}{r!2^r} e^{-\bar{\sigma}t}(1-e^{-2\bar{\sigma}t})^r, \]

we have \(\lim_{r \to \infty} a_{r-1}(n, t_0)/a_r(n, t_0) = \exp(2\bar{\sigma}t_0) - 1 = \exp(\frac{1}{2}) - 1 < 1\) and so the series in (3.15) converges for \(t = t_0\). 

**Theorem 3.7.** For each \(\bar{\mu}^0 \in \bar{\mathcal{P}}\), the solution of the martingale problem for \((A, \bar{\mu}^0)\) is unique.

**Proof.** Let \(t_0\) be as in Lemma 3.6 and \(P\) be any solution of the martingale problem for \((A, \bar{\mu}^0)\). Then by Theorem 3.5, (3.12) holds for all \((f, k) \in \bigcup_{m=1}^{\infty} (C(E^m) \times S^m)\)

\[ \]
and $0 \leq t \leq t_0$. Therefore for each $t \in [0, t_0]$ the distribution $P(\tilde{\mu}(t) \in \cdot)$ on $\tilde{H}$ depends only on $t$ and $\tilde{\mu}^0$, which is written by $Q_t^0\tilde{\mu}^0$.

Noting that the regular conditional distribution of $\theta_{t_0} \tilde{\mu}(\cdot)$ given $\tilde{\mu}(t_0) = \tilde{\nu} \in \tilde{H}$ under $P$ (where $\theta_{t_0}$ is the shift operator on $C([0, \infty), \tilde{H})$) is also a solution of the martingale problem for $(A, \tilde{\nu})$ for $Q_t^0\tilde{\mu}^0$-almost all $\tilde{\nu}$, we have again by Theorem 3.5,

$$E^P[(f, \tilde{\mu}(t_0 + t)_k)]$$

$$= \int_{\tilde{H}} E_{(f,k)}[\langle Y(t), \tilde{\nu}_{k(t)} \rangle \exp\left( \tilde{\sigma} \int_0^t |k(u)| \, du \right)] Q_t^0\tilde{\mu}^0(\,d\tilde{\nu})$$

for all $(f, k) \in \bigcup_{m=1}^{\infty} (C(E^m) \times S^m)$ and $0 \leq t \leq t_0$. Consequently for each $t \in [t_0, 2t_0]$ the distribution $P(\tilde{\mu}(t) \in \cdot)$ is uniquely determined by $t$ and $\tilde{\mu}^0$.

Repeating this procedure shows that for each $\tilde{\mu}^0 \in \tilde{H}$ any two solutions of the martingale problem for $(A, \tilde{\mu}^0)$ have the same one-dimensional distributions and we conclude uniqueness. (See Theorem 4.4.2 of Ethier and Kurtz (1986).) $\square$

**Remark 3.8.** Using the same method as in Ethier and Kurtz (1987), one can show that it is possible to allow $\sigma$ to be a symmetric function in $B(E^2)$. That is, if we define $A \subset B(\tilde{H}) \times B(\tilde{H})$ for such $\sigma$ in the same way as in Section 2, then well-posedness of the martingale problem for $A$ still holds.

### 4. Proof of Theorem 3.5

This section is devoted to the proof of Theorem 3.5.

Take $(f, k) \in \bigcup_{m=1}^{\infty} (C(E^m) \times S^m)$, $\tilde{\mu}^0 \in \tilde{H}$ and a solution $P$ of the martingale problem for $(A, \tilde{\mu}^0)$. What we have to prove is the identity

$$E^P[(f, \tilde{\mu}(t)_k)] = E_{(f,k)}[\langle Y(t), \tilde{\mu}^0_{k(t)} \rangle W(t)]$$

(4.1)

for $0 \leq t \leq t_0 = (4\tilde{\sigma})^{-1}$, where $W(t) = \exp(\tilde{\sigma} \int_0^t |k(u)| \, du)$.

Before we come to the proof of (4.1), we note the following. We can construct the $\tilde{H}$-valued process $\{\tilde{\mu}(t); t \geq 0\}$ whose distribution is $P$ and the dual process $\{(Y(t), k(t)); t \geq 0\}$ (defined in the previous section) starting from $(f, k)$ on the same probability space so that these two processes are mutually independent. Fix $t \in (0, t_0]$ and set

$$F(s) = E[\langle Y(s), \tilde{\mu}(t-s)_{k(s)} \rangle W(s)]$$

for $0 \leq s \leq t$. Then (4.1) is equivalent to the identity $F(0) = F(t)$.

The proof of this equality is divided into three steps. First we will show that the proof of (4.1) is reduced to estimations of certain expectations.
Proposition 4.1. For an integer \( l > |k| \), let \( \rho_l \) be the stopping time related to \( \{ k(t); t \geq 0 \} \) introduced in the proof of Lemma 3.4. If

\[
E[(Y(s+h), \tilde{\mu}(t-s-h)_{k(s+h)}) W(s+h); \rho_l > s] - E[(Y(s), \tilde{\mu}(t-s)_{k(s)}) W(s); \rho_l > s] \leq ||f|| \cdot E_k[W(t); s < \rho_l \leq s + h] + O(h^2)
\]

holds as \( h \downarrow 0 \) uniformly in \( s \in [0, t) \), then (4.1) is valid.

Proof. For \( s \in [0, t] \), set

\[
F_i(s) = E[(Y(s \wedge \rho_l), \tilde{\mu}(t-s)_{k(s)}) W(s \wedge \rho_l)].
\]

Then for \( s \in [0, t] \) and \( h > 0 \) with \( s + h \leq t \), we have

\[
F_i(s + h) - F_i(s)
\]

\[
= E[(Y(\rho_l), \tilde{\mu}(t-s-h)_{k(\rho_l)}) W(\rho_l); s < \rho_l \leq s + h] + E[(Y(s+h), \tilde{\mu}(t-s-h)_{k(s+h)}) W(s+h); \rho_l > s] - E[(Y(s), \tilde{\mu}(t-s)_{k(s)}) W(s); \rho_l > s] + E[\langle Y(\rho_l), \tilde{\mu}(t-s-h)_{k(\rho_l)} \rangle - \langle Y(\rho_l), \tilde{\mu}(t-s)_{k(\rho_l)} \rangle] W(\rho_l); \rho_l \leq s].
\]

Set \( t_i = (i/N)t \), \( i = 0, 1, 2, \ldots \). Assume that (4.2) is true. Since the right hand side of (4.2) majorizes the sum of the second and the fourth terms, we obtain

\[
|F_i(t) - F_i(0)| = \lim_{N \to \infty} \left| \sum_{i=1}^{N} (F_i(t_i) - F_i(t_{i-1})) \right|
\]

\[
\leq \limsup_{N \to \infty} \sum_{i=1}^{N} \left| E[(Y(\rho_i), \tilde{\mu}(t-t_i)_{k(\rho_i)}) W(\rho_i); t_{i-1} < \rho_i \leq t_i] \right|
\]

\[
+ \limsup_{N \to \infty} \sum_{i=1}^{N} \left| E[(Y(t_i), \tilde{\mu}(t-t_i)_{k(t_i)}) W(t_i); t_{i-1} < \rho_i \leq t_i] \right|
\]

\[
+ \limsup_{N \to \infty} \left| \sum_{i=1}^{N} E[\langle Y(\rho_i), \tilde{\mu}(t_i)_{k(\rho_i)} \rangle \cdot W(\rho_i); \rho_i \leq t_{i-1}] \right|
\]

\[
+ ||f|| \cdot E_k[W(t); \rho_i \leq t] \leq 3||f|| \cdot E_k[W(t); \rho_i \leq t]
\]

\[
+ \limsup_{N \to \infty} \left| F[(Y(\rho_i), \tilde{\mu}(0)_{k(\rho_i)}) W(\rho_i); \rho_i \leq t] \right|
\]

\[
- \sum_{i=1}^{N} E[(Y(\rho_i), \tilde{\mu}(t-t_i)_{k(\rho_i)}) W(\rho_i); t_{i-1} < \rho_i \leq t_i] \right| \leq 5||f|| \cdot E_k[W(t); \rho_i \leq t].
\]
Consequently
\[ |E^{P}[\langle f, \tilde{\mu}(t)k \rangle] - E_{\{\nu\}}[\langle Y(t \wedge \rho_l, \tilde{\mu}_{k(t, \rho_l)}^0, W(t \wedge \rho_l) \rangle]| \leq 5\|f\| \cdot E_k[W(t); \rho_l \leq t]. \]

Here note (3.10) and that \(E_k[W(t)] < \infty\) by Lemma 3.6. Letting \(l \to \infty\), we have (4.1). \(\square\)

We shall show in what follows that the estimate (4.2) is true. Now let \(l > |k|\) and \(s \in [0, t)\) be fixed. Set \(M(u) = |k(u)|\) and
\[ I(h) = E[\langle Y(s + h), \tilde{\mu}((t - s - h)_{k(s + h)}, W(s + h); \rho_l > s \rangle] \]
for \(0 \leq h \leq t - s\). Then we can rephrase (4.2) in the form
\[ |I(h) - I(0)| \leq \|f\| \cdot E_k[W(t); s < \rho_l \leq s + h] + O(h^2). \]  
Set
\[ I(h) = I_1(h) + I_2(h) + I_3(h) + I_4(h) + R(h) \]  
where

\[ I_1(h) = E[\langle T_{M(s)}(h) Y(s), \tilde{\mu}((t - s - h)_{k(s)}), W(s); \rho_l > s \rangle], \]
\[ I_2(h) = E \left[ \sum_{i=1}^{M(s)} \sum_{i \leq j \leq M(s)} \int_0^h \left( T_{M(s) - 1}(h - r) \Phi\alpha T_{M(s)}(r) Y(s), \tilde{\mu}((t - s - h)_{\alpha(k(s))}) \right) \right. \]
\[ \cdot W(s); \rho_l > s \]  
\[ - \langle T_{M(s)}(h) Y(s)/\tilde{\mu}((t - s - h)_{k(s)}) \rangle \cdot W(s); \rho_l > s \],
\[ I_4(h) = E \left[ \sum_{i=1}^{M(s)} \sum_{k \in S} m_k T_{M(s)}(h) Y(s), \tilde{\mu}((t - s - h)_{\gamma(k)}) \right]  
\cdot W(s); \rho_l > s \] \( \cdot h, \)

and \(R(h)\) is the remainder term.

**Proposition 4.2.** With the notation established above, we have the following inequality:
\[ |R(h)| \leq \|f\| \cdot E_k[W(t); s < \rho_l \leq s + h] + O(h^2) \]  
as \(h \downarrow 0\) uniformly in \(s\).
Proof. Since \{\tilde{\mu}(t); t \geq 0\} and \{(Y(t), k(t)); t \geq 0\} are independent, \(I(h)\) is expressible as

\[
I(h) = \int_{\mathcal{F}} E_{(f, k)}[\langle Y(s+h), \tilde{\nu}_{k(s+h)} \rangle W(s+h); \rho_i > s] P(\tilde{\mu}(t-s-h) \in d\tilde{\nu}).
\]  

(4.6)

For a moment \(\tilde{\nu} \in \mathcal{F}\) is fixed and \(E_{(f, k)}\) is denoted simply by \(E\). Set

\[
X(h) = \langle Y(s+h), \tilde{\nu}_{k(s+h)} \rangle \exp\left(\tilde{\sigma} \int_s^{s+h} M(u) \, du\right).
\]

Then clearly it holds that

\[
E[\langle Y(s+h), \tilde{\nu}_{k(s+h)} \rangle W(s+h); \rho_i > s] = E[E[X(h) | \mathcal{F}_x^k] \cdot W(s); \rho_i > s]  
\]

(4.7)

where \(\mathcal{F}_x^k = \sigma(k(u); 0 \leq u \leq s)\). To compute the conditional expectation in (4.7), we decompose the total event into three parts according to the numbers of jumps of \(\{k(t); t \geq 0\}\) in the time interval \((s, s+h]\) as follows:

\[
J_0 = \{ \text{no jump occurs in } (s, s+h] \},
\]

\[
J_1 = \{ \text{only one jump occurs in } (s, s+h] \},
\]

\[
J_2 = \{ \text{two or more jumps occur in } (s, s+h] \}.
\]

First from the definition of \(Y\) and \(K\), it holds that

\[
E[X(h)1_{J_0} | \mathcal{F}_x^k] = E[\langle T_{M(s)}(h) Y(s), \tilde{\nu}_{k(s)} \rangle | \mathcal{F}_x^k] \cdot \exp(\tilde{\sigma} M(s) h) \cdot P_k(s)(\tau_1 > h)
\]

\[
= E[\langle T_{M(s)}(h) Y(s), \tilde{\nu}_{k(s)} \rangle | \mathcal{F}_x^k] \cdot \exp((\tilde{\sigma} M(s) - \lambda(k(s))) h).
\]

(See (3.5) for the definition of \(\lambda(\cdot)\) and note that \(E[\langle T_{M(s)}(h) Y(s), \tilde{\nu}_{k(s)} \rangle | \mathcal{F}_x^k]\) is \(\mathcal{F}_x^k\)-measurable.) Therefore

\[
E[\langle Y(s+h), \tilde{\nu}_{k(s+h)} \rangle W(s+h); \rho_i > s] \cap J_0]
\]

\[
= E[\langle T_{M(s)}(h) Y(s), \tilde{\nu}_{k(s)} \rangle \exp((\tilde{\sigma} M(s) - \lambda(k(s))) h) \cdot W(s); \rho_i > s]
\]

\[
= E[\langle T_{M(s)}(h) Y(s), \tilde{\nu}_{k(s)} \rangle W(s); \rho_i > s]
\]

\[
+ E[\langle T_{M(s)}(h) Y(s), \tilde{\nu}_{k(s)} \rangle | \mathcal{F}_x^k] \exp((\tilde{\sigma} M(s) - \lambda(k(s))) h) \cdot W(s); \rho_i > s] + O(h^2)
\]

(4.8)

uniformly in \(s\) and \(\tilde{\nu}\), where the last equality follows from the inequality \(\tilde{\sigma} M(s) \leq \lambda(k(s))\).

Next we decompose \(J_1\) depending on the values of \(M(s+h) - M(s)\), in such a way that

\[
J_1 = J_\alpha \cup J_\beta \cup J_\gamma,
\]
where $J_\alpha = J_1 \cap \{M(s+h)-M(s) = 2\}$, $J_\beta = J_1 \cap \{M(s+h)-M(s) = -1\}$ and $J_\gamma = J_1 \cap \{M(s+h)-M(s) = 0\}$. Corresponding to these events the following equalities hold:

$$E[X(h)1_{J_\alpha} | \mathcal{F}_s^K]$$

$$= \sum_{i=1}^{M(s)} \int_0^h E[(T_{M(s)+2}(h-r)K_{iM(s)}T_{M(s)}(r) Y(s), \tilde{\nu}_{\alpha,k_i(s)}|K]$$

$$\cdot \exp\{\tilde{\sigma}(M(s)+2)(h-r) + \tilde{\sigma}M(s)r\} \cdot P_{\alpha,k_i(s)}(\tau_i > h-r)$$

$$\cdot P_{k_i(s)}(\tau_i \in dr, k(\tau_i) = k')_{k'_i < \alpha, k_i(s)}$$

$$= \tilde{\sigma} \sum_{i=1}^{M(s)} \int_0^h E[(T_{M(s)+2}(h-r)K_{iM(s)}T_{M(s)}(r) Y(s), \tilde{\nu}_{\alpha,k_i(s)}|K]$$

$$\cdot \exp\{\tilde{\sigma}M(s)h + 2\tilde{\sigma}(h-r) - \lambda(\alpha,k_i(s))(h-r) - \lambda(k_i(s))r\} \cdot P_{k_i(s)}(\tau_i > h-r)$$

$$\cdot \exp\{\tilde{\sigma}(M(s)-1)(h-r) + \tilde{\sigma}M(s)r\} \cdot P_{\beta,k_i(s)}(\tau_i > h-r)$$

$$\cdot P_{k_i(s)}(\tau_i \in dr, k(\tau_i) = k')_{k'_i < \beta, k_i(s)}$$

$$= \sum_{1 \leq i < j < M(s)} \int_0^h E[(T_{M(s)-1}(h-r)\Phi_i T_{M(s)}(r) Y(s), \tilde{\nu}_{\beta,k_i(s)}|K]$$

$$\cdot \exp\{\tilde{\sigma}M(s)h - \tilde{\sigma}(h-r) - \lambda(\beta,k_i(s))(h-r)$$

$$- \lambda(k_i(s))r\} \cdot P_{k_i(s)}(\tau_i > h-r)$$

$$\cdot \exp\{\tilde{\sigma}(M(s)-1)(h-r) + \tilde{\sigma}M(s)r\} \cdot P_{\beta,k_i(s)}(\tau_i > h-r)$$

$$\cdot P_{k_i(s)}(\tau_i \in dr, k(\tau_i) = k')_{k'_i < \beta, k_i(s)}$$

$$E[X(h)1_{J_\beta} | \mathcal{F}_s^K] = \sum_{j=2}^{M(s)} \sum_{1 \leq i < j \leq k(s)} \#\{l; 1 \leq l < j, k_s(s) = k_j(s)\}^{-1}$$

$$\cdot \int_0^h E[(T_{M(s)-1}(h-r)\Phi_i T_{M(s)}(r) Y(s), \tilde{\nu}_{\beta,k_i(s)}|K]$$

$$\cdot \exp\{\tilde{\sigma}(M(s)-1)(h-r) + \tilde{\sigma}M(s)r\} \cdot P_{\beta,k_i(s)}(\tau_i > h-r)$$

$$\cdot P_{k_i(s)}(\tau_i \in dr, k(\tau_i) = k')_{k'_i < \beta, k_i(s)}$$

$$= \sum_{1 \leq i < j < M(s)} \int_0^h E[(T_{M(s)-1}(h-r)\Phi_i T_{M(s)}(r) Y(s), \tilde{\nu}_{\beta,k_i(s)}|K]$$

$$\cdot \exp\{\tilde{\sigma}M(s)h - \tilde{\sigma}(h-r) - \lambda(\beta,k_i(s))(h-r)$$

$$- \lambda(k_i(s))r\} \cdot P_{k_i(s)}(\tau_i > h-r)$$

$$\cdot \exp\{\tilde{\sigma}(M(s)-1)(h-r) + \tilde{\sigma}M(s)r\} \cdot P_{\beta,k_i(s)}(\tau_i > h-r)$$

$$\cdot P_{k_i(s)}(\tau_i \in dr, k(\tau_i) = k')_{k'_i < \beta, k_i(s)}$$

and

$$E[X(h)1_{J_\gamma} | \mathcal{F}_s^K] = \sum_{i=1}^{M(s)} \sum_{k_i(s) > k(s)} m_{k_i(s)} E[(T_{M(s)}(h) Y(s), \tilde{\nu}_{\gamma(k')}(k_i(s))|K]$$

$$\cdot \int_0^h \exp\{\tilde{\sigma}M(s)h - \lambda(\gamma_i(k'))(h-r)$$

$$- \lambda(k_i(s))r\} \cdot P_{k_i(s)}(\tau_i > h-r)$$

$$\cdot \exp\{\tilde{\sigma}(M(s)-1)(h-r) + \tilde{\sigma}M(s)r\} \cdot P_{\gamma(k')}(\tau_i > h-r)$$

From (4.7) and (4.9a) we have

$$E[(Y(s+h), \tilde{\nu}_{k(s+h)} W(s+h); \{\rho_t > s \cap J_\alpha\}]$$
\[ \begin{align*}
= E \left[ \sigma \sum_{i=1}^{M(s)} \int_0^h \langle T_{M(s) + 2}(h-r) K_{iM(s)} T_{M(s)}(r) Y(s), \tilde{\nu}_{\alpha, k(s)} \rangle \cdot \exp \{ \sigma M(s) h + 2 \sigma (h-r) - \lambda (\alpha, k(s))(h-r) - \lambda (k(s)) r \} dr \cdot W(s); \rho_l > s \right] \\
= E \left[ \sigma \sum_{i=1}^{M(s)} \int_0^h \langle T_{M(s) + 2}(h-r) K_{iM(s)} T_{M(s)}(r) Y(s), \tilde{\nu}_{\alpha, k(s)} \rangle dr \cdot W(s); \rho_l > s \right] + O(h^2)
\end{align*} \] (4.10a)

uniformly in \( s \) and \( \tilde{\nu} \), where the last equality follows from the inequality
\[ \sigma M(s) h + 2 \sigma (h-r) - \lambda (\alpha, k(s))(h-r) - \lambda (k(s)) r \leq 0. \]

Similar observations on (4.9b) and (4.9c) show that
\[ \begin{align*}
E[\langle Y(s+h), \tilde{\nu}_{k(s+h)} \rangle W(s+h); \{ \rho_l > s \} \cap J_\beta] \\
= E \left[ \sum_{1 \leq i < j \leq M(s)} \int_0^h \langle T_{M(s)-1}(h-r) \Phi_i T_{M(s)}(r) Y(s), \tilde{\nu}_{\beta, k(s)} \rangle dr \cdot W(s); \rho_l > s \right] + O(h^2)
\end{align*} \] (4.10b)

and
\[ \begin{align*}
E[\langle Y(s+h), \tilde{\nu}_{k(s+h)} \rangle W(s+h); \{ \rho_l > s \} \cap J_\gamma] \\
= E \left[ \sum_{i=1}^{M(s)} \sum_{k' \neq k(s)} m_{k', k(s)} \langle T_{M(s)}(h) Y(s), \tilde{\nu}_{\gamma, k'(k(s))} \rangle \cdot W(s); \rho_l > s \right] + h + O(h^2)
\] (4.10c)

hold as \( h \to 0 \) uniformly in \( s \) and \( \tilde{\nu} \).

It remains to estimate the integral over \( J_2 \). It is easy to see that there exists a positive constant \( C_2 \) such that
\[ P_k(\tau_2 \leq h) \leq C_2 h^2 |k|^4 \quad \text{for all } h > 0 \text{ and } k \in S^*. \]

From this it follows that if we set
\[ R^*(h) = E[\langle Y(s+h), \tilde{\nu}_{k(s+h)} \rangle W(s+h); \{ \rho_l > s \} \cap J_2], \] (4.11)

then
\[ |R^*(h)| \leq \| f \| \cdot E_k[ W(s+h); \{ s < \rho_l \leq s+h \} \cap J_2] + \| f \| \cdot \exp(\tilde{\sigma} l) P_k(\{ \rho_l > s+h \} \cap J_2) \leq \| f \| \cdot E_k[ W(t); s < \rho_l \leq s+h ] + O(h^2) \]

uniformly in \( s \) and \( \tilde{\nu} \).
Summing (4.8), (4.10) and (4.11), integrating the sum by $P(\mu(t-s-h) \in d\nu)$, and using the relation

$$\sigma M(s) - A(\mathbf{k}(s)) = -\#\{(i, j); 1 \leq i < j \leq M(s), k_i(s) = k_j(s)\} + \sum_{i=1}^{M(s)} m_{k_i(s)k_i(s)},$$

we finally come to the desired estimate (4.5). \(\square\)

The following lemma is needed for the final step.

**Lemma 4.3.** Let $P$ be a solution of the martingale problem for $A$. Set

$$R_i(f, \mathbf{k}; u) = \sum_{i=1}^m \langle K_{im} T_m(u) f, \mu(t-u)_{\alpha_i} \rangle$$

$$+ \sum_{1 \leq i < j \leq m, k_i = k_j} \left( \langle \Phi_j T_m(u) f, \mu(t-u)_{\beta_j} \rangle - \langle T_m(u) f, \mu(t-u)_{\beta_j} \rangle \right)$$

$$+ \sum_{i=1}^m \sum_{k' \in S} m_{k'k_i} \langle T_m(u) f, \mu(t-u)_{\gamma_i(k')} \rangle$$

(4.12)

for $f \in C(E^m)$, $\mathbf{k} = (k_1, \ldots, k_m) \in S^m$ and $0 \leq u \leq t$.

Then for each $s \in [0, t]$,

$$|R_i(f, \mathbf{k}; u)| \leq 2\lambda(\mathbf{k}) \|f\|$$

(4.13)

and

$$E^P[\langle f, \mu(t) \rangle] = E^P[\langle T_m(s) f, \mu(t-s) \mu \rangle] + E^P \left[ \int_0^s R_i(f, \mathbf{k}; u) \, du \right].$$

(4.14)

**Proof.** (4.13) follows from the definitions of the operators in the right hand side of (4.12).

By Lemma 4.3.4 of Ethier and Kurtz (1986),

$$Z(s) = \langle T_m(t-s) f, \mu(s) \rangle - \int_0^s R_i(f, \mathbf{k}; t-u) \, du$$

is a $(P, \{\mathcal{F}_t\})$-martingale for $s \in [0, t]$. In particular it holds that $E^P[Z(t)] = E^P[Z(t-s)]$, which is equivalent to (4.14). \(\square\)

We are now ready to prove (4.3). Noting that the processes $\{\mu(t); t \geq 0\}$ and $\{(Y(t), \mathbf{k}(t)); t \geq 0\}$ are independent, we have the following equalities as an immediate consequence of Lemma 4.3:

$$I(0) = E[\langle I_{M(s)}(s) Y(s), \mu(t-s-h)_{k(s)} \rangle W(s); \rho_i > s]$$

$$+ E \left[ \int_0^h R_{t-h} (Y(s), \mathbf{k}(s); r) \, dr \cdot W(s); \rho_i > s \right]$$

$$= I_1^\infty(h) + I_2^\infty(h) + I_3^\infty(h) + I_4^\infty(h),$$

(4.15)
where

\[ I_1^*(h) = I_1(h), \]

\[ I_2^*(h) = E \left[ \hat{\sigma} \sum_{i=1}^{M(s)} \int_0^h \langle K_{iM(s)} T_{M(s)}(r) Y(s), \tilde{\mu}(t-s-r) \alpha_{i,k(s)} \rangle \, dr \cdot W(s); \rho_i > s \right], \]

\[ I_3^*(h) = E \left[ \sum_{\substack{1 \leq i < j \leq M(s) \ q_{i(s)}, k_{j(s)} = k_{j(s)} \}} \int_0^h \langle \Phi_{ij} T_{M(s)}(r) Y(s), \tilde{\mu}(t-s-r) \beta_{i,k(s)} \rangle 
\quad - \langle T_{M(s)}(r) Y(s), \tilde{\mu}(t-s-r) \beta_{i,k(s)} \rangle \, dr \cdot W(s); \rho_i > s \right] \]

and

\[ I_4^*(h) = E \left[ \sum_{i=1}^{M(s)} \sum_{k' \in S} m_{k,k}(s) \int_0^h \langle T_{M(s)}(r) Y(s), \tilde{\mu}(t-s-r) \gamma_{i,k(s)} \rangle \, dr \cdot W(s); \rho_i > s \right]. \]

Comparing (4.15) with (4.4) and using Proposition 4.2, we see that (4.3) is proved by verifying that

\[ I_i(h) - Z_i(h) = O(h^i) \quad (4.16) \]

holds as \( h \downarrow 0 \) uniformly in \( s \in [0, t) \) for \( i = 2, 3 \) and 4.

Observing that again by Lemma 4.3,

\[ E[\langle K_{iM(s)} T_{M(s)}(r) Y(s), \tilde{\mu}(t-s-r) \alpha_{i,k(s)} \rangle | K] \]

\[ = E[\langle T_{M(s)+2} h-r K_{iM(s)} T_{M(s)}(r) Y(s), \tilde{\mu}(t-s-h) \alpha_{i,k(s)} \rangle | K] \]

\[ + E \left[ \int_0^{h-r} R_{i-s-r} \langle K_{iM(s)} T_{M(s)}(r) Y(s), \alpha_{i,k(s)}(u) \rangle \, du | K \right] \]

holds and using (4.13) and (3.9), we obtain

\[ |I_2(h) - I_2^*(h)| = \left| E \left[ \hat{\sigma} \sum_{i=1}^{M(s)} \int_0^h \int_0^{h-r} du R_{i-s-r} \times \langle K_{iM(s)} T_{M(s)}(r) Y(s), \alpha_{i,k(s)}(u) \rangle \cdot W(s); \rho_i > s \right] \right| \]

\[ \leq \exp(\hat{\sigma} t l) E_{f,k} \left[ 2\hat{\sigma} \sum_{i=1}^{M(s)} \lambda(\alpha_{i,k(s)}) \int_0^h dr \cdot \| K_{iM(s)} T_{M(s)}(r) Y(s) \| (h-r); \rho_i > s \right] \]

\[ \leq \exp(\hat{\sigma} t l) \hat{\sigma} l \cdot C_1 (1+1)^2 \| f \| h^2, \]

and hence (4.16) holds for \( i = 2 \).

Similar calculations give (4.16) for \( i = 3 \) and 4. We now obtain (4.3), which completes the proof of Theorem 3.5 with the help of Proposition 4.1.
5. The equilibrium state and probabilities of identity

In this section, we shall study the equilibrium state of our stepping stone model which is defined as the solution of the martingale problem for $A$. Our attention will be focussed on the neutral case where no selection acts, in other words we take 'the selection intensity function' $\sigma = 0$. Thus $A$ has the form

$$A = \{(\Phi_{f,k}, \Psi_{f,k}); m \in \mathbb{N}, f \in \mathcal{D}(L)^m, k \in S^m\}$$

where

$$\Phi_{f,k}(\tilde{\mu}) = (f, \tilde{\mu}_k)$$

and where

$$\Psi_{f,k}(\tilde{\mu}) = \langle L^m f, \tilde{\mu}_k \rangle + \sum_{i=1}^{m} \sum_{k \in S} m_{k'} k \langle f, \tilde{\mu}_{\gamma(k')k} \rangle$$

$$+ \sum_{1 \leq i < j \leq m, k_i = k_j} (\langle \Phi_{f,j} f, \tilde{\mu}_{\beta,k} \rangle - \langle f, \tilde{\mu}_k \rangle)$$

(5.1)

(cf. (3.4)).

Note that as in Section 3.2 we can construct the dual process $\{(Y(t), k(t)); t \geq 0\}$ associated with the solution of the martingale problem for $A$. Indeed the generator $\mathcal{L}$ of the Markov process $\{k(t); t \geq 0\}$ in $\bigcup_{m=1}^{\infty} S^m$ is given by

$$\mathcal{L}h(k) = \sum_{j=2}^{[k]} \sum_{i \neq j} m_{k_i k_j} \{h(\gamma_i(k')k) - h(k)\}$$

$$+ \sum_{i=1}^{[k]} \sum_{k \neq k_i} m_{k_i k_i} \{h(\gamma_i(k')k) - h(k)\},$$

and $\{Y(t); t \geq 0\}$ is defined by (3.6) and (3.7). In this case the duality relationship (Theorem 3.5) is described as

$$E_{\tilde{\nu}}[\langle f, \tilde{\mu}(t) \rangle] = E_{\tilde{\nu}_k}(\langle Y(t), \tilde{\nu}_k(t) \rangle)$$

(5.2)

for $t \geq 0$, $\tilde{\nu} \in \tilde{\mathcal{P}}$, and $(f, k) \in \bigcup_{m=1}^{\infty} \{(C(E^m) \times S^m)\}$, where $E_{\tilde{\nu}}$ denotes the expectation with respect to $P_{\tilde{\nu}}$, the solution of the martingale problem for $(A, \tilde{\nu})$.

We first discuss the convergence of $\tilde{\mu}(t)$ to equilibrium as $t \to \infty$ by using the equality (5.2).

**Theorem 5.1.** Assume that there exists $Q \in \mathcal{P}(E)$ such that for all $f \in C(E)$,

$$T(t)f \to \langle f \rangle = \int_E f(x)Q(dx)$$

(5.3)

as $t \to \infty$ uniformly on $E$, where $\{T(t)\}_{t \geq 0}$ is the Feller semigroup generated by the mutation operator $L$. Then the process $\{\tilde{\mu}(t); t \geq 0\}$ corresponding to the solution of
the martingale problem for A has a unique stationary distribution \( \hat{Q} \in \mathcal{P}(\mathcal{F}) \) and

\[
\int_\mathcal{F} \langle \hat{f}, \mu_k \rangle \hat{Q}(d\hat{\mu}) = \langle f \rangle
\]  

for all \( k \in S \) and \( f \in B(E) \).

**Proof.** Existence of a stationary distribution follows from compactness of \( \mathcal{F} = \mathcal{P}(E)^S \). By Lemma 2.1, to show uniqueness it is sufficient to prove that for all \( (f, k) \in \bigcup_{m=1}^\infty (C(E^m) \times S^m) \) and all \( \hat{\nu} \in \hat{\mathcal{F}} \) the left hand side of (5.2) converges as \( t \to \infty \) and that the limit does not depend on \( \nu \). We therefore consider the right hand side of (5.2).

First note that \( |k(t)| \) is non-increasing in \( t \). Set

\[
M(\infty) = \lim_{t \to \infty} |k(t)| \quad \text{and} \quad \rho_i = \inf\{t > 0; |k(t)| \leq l\}
\]

for \( l \in \mathbb{N} \). For \( k \in S^m \), clearly

\[
0 = \rho_m' < \rho_{m-1}' \leq \cdots \leq \rho_1' \leq \infty, \quad P_k - \text{a.s.}
\]

It is easy to show from (5.3) that for \( g \in C(E^l) \)

\[
T_l(t)g \to \langle g \rangle_l = \int_{E^l} g(x_1, \ldots, x_l)Q(dx_1) \times \cdots \times Q(dx_l)
\]

as \( t \to \infty \) uniformly on \( E^l \). These observations and (5.2) immediately give

\[
\lim_{t \to \infty} E_\mathcal{F}[\langle f, \hat{\mu}(t) \rangle] = \sum_{l=1}^{[k]} E_{\phi_k}[\langle Y(\rho_i) \rangle_l; M(\infty) = l]
\]

for all \( \hat{\nu} \in \hat{\mathcal{F}} \), so that we obtain a unique stationary distribution \( \hat{Q} \), which satisfies

\[
\int_\mathcal{F} \langle f, \hat{\mu}_k \rangle \hat{Q}(d\hat{\mu}) = \sum_{l=1}^{[k]} E_{\phi_k}[\langle Y(\rho_i) \rangle_l; M(\infty) = l]
\]  

for all \( (f, k) \in \bigcup_{m=1}^\infty (C(E^m) \times S^m) \). The formula (5.4) follows from (5.5) with \( m = 1 \).

We discuss the stationary state \( \hat{Q} \) for a special model, which is given by

\[
L_f(x) = \frac{1}{2} \theta \int_E (f(y) - f(x)) \pi(x, dy)
\]

where \( \theta \) is a positive constant and where \( \pi(x, dy) \) is a Feller transition function. We assume that \( L \) satisfies the condition in Theorem 5.1 and that \( \pi(x, \cdot) \) has no atoms for each \( x \in E \). (Typically consider the case \( E = [0, 1] \) and \( \pi(x, dy) \) is Lebesgue measure on \([0, 1]\) for each \( x \in [0, 1] \). We call the corresponding measure-valued process the infinitely many alleles stepping stone model with uniform mutation.)
Regarding $L$ as an operator on $B(E)$, we can extend the operator $L^{(m)}$ defined in Section 3.2 to that on $B(E^m)$ by setting

$$L^{(m)}f(x_1, \ldots, x_m) = \sum_{i=1}^{m} \{Lf(x_1, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_m)\}(x_i)$$

$$= \frac{1}{2} \theta \sum_{i=1}^{m} \{ \pi_i f(x_1, \ldots, x_m) - f(x_1, \ldots, x_m) \}$$

where

$$\pi_i f(x_1, \ldots, x_m) = \int_{E} f(x_1, \ldots, x_i, y, x_{i+1}, \ldots, x_m) \pi(x_i, dy).$$

We can also define $\Psi_{f,k}$ for $(f, k) \in \bigcup_{m=1}^{\infty} (B(E^m) \times S^m)$ as the function on $\tilde{P}$ by the same expression as (5.1).

The next lemma gives us some more information on $\tilde{Q}$.

**Lemma 5.2.** For all $(f, k) \in \bigcup_{m=1}^{\infty} (B(E^m) \times S^m)$,

$$\int_{\tilde{P}} \Psi_{f,k}(\tilde{\mu}) \tilde{Q}(d\tilde{\mu}) = 0. \tag{5.7}$$

**Proof.** For $f$ such that $f(x_1, \ldots, x_m) = f_1(x_1) \cdots f_m(x_m), f_i \in C(E)$, (5.7) is obvious from the definition of the martingale problem and from the time stationarity property of the process $\{\tilde{\mu}(t); t \geq 0\}$ with initial distribution $\tilde{Q}$. Every $g$ in $C(E^m)$ can be approximated uniformly by finite linear combinations of the above $f_i$'s, and so (5.7) holds for each continuous function $f$. Furthermore the extension of (5.7) to bounded measurable functions $f$ can be done by approximation argument using Proposition 3.4.2 of Ethier and Kurtz (1986).

As is mentioned in Section 1, we are interested in a certain quantity related to $\tilde{Q}$. In biological terminology, it is the probability that two distinct individuals chosen at random, one from colony $k$ and the other from colony $l$, are of the same type. The probabilities of such an identity in geographically structured populations have been discussed in many contexts, see, e.g., Maruyama (1977) and Nagylaki (1983).

With our setting the quantity in question is given by

$$I_{ki} = \int_{\tilde{P}} \langle 1_D, \mu_k \times \mu_l \rangle \tilde{Q}(d\tilde{\mu})$$

for $k, l \in S$, where $D = \{(x, x); x \in E\}$ and $1_D$ is the indicator function of $D$.

Here note that there is a possibility that $I_{ki}$ is positive. For by a slight modification of the proof of Theorem 10.4.5 of Ethier and Kurtz (1986) (i.e. by replacing $\varphi_\gamma$ in the proof by $\varphi_\gamma(\tilde{\mu}) \equiv \sum_{x \in E} \mu_k(\{x\})^\gamma$ with $k \in S$ fixed) one can prove that every solution $P$ of the martingale problem for $A$ satisfies

$$P(\mu_k(t) \in \mathcal{P}_d(E) \text{ for all } t > 0 \text{ and } k \in S) = 1,$$
where \( \mathcal{P}_a(E) \) denotes the set of purely atomic Borel probability measures on \( E \), and hence
\[
\tilde{Q}(\mu_k \in \mathcal{P}_a(E) \text{ for all } k \in S) = 1.
\]

Letting \( m = 2, f = 1_D \) and \( k = (k, l) \) in (5.1) and (5.6), we have
\[
\Psi_{1_D,(k,l)}(\vec{\mu}) = -\theta(1_D, \mu_k \times \mu_l) + \sum_{k \in S} m_{k,k'}(1_D, \mu_k \times \mu_{k'})
+ \sum_{k' \in S} m_{k'1}(1_D, \mu_k \times \mu_{k'}) + \delta_{kl}(1 - (1_D, \mu_k \times \mu_l)).
\]

Integrating both sides by \( \tilde{Q}(d\vec{\mu}) \), Lemma 5.2 gives
\[
(\theta - m_{kk} - m_{ll} + \delta_{kl}) I_{kl} = \sum_{k' \neq k} m_{k,k'} I_{k'1} + \sum_{k' \neq l} m_{k'1} I_{kk'} + \delta_{kl}.
\]

In what follows we consider the case \( S = \mathbb{Z}^d \), the \( d \)-dimensional integer lattice, and assume that migration is homogeneous, i.e., \( m_{kl} \) depends only on the displacement \( k - l \). In this case we can write
\[
m_{kl} = m_{k-l}, \quad \sum_{k \neq 0} m_k = m.
\]

It follows from (5.5) with \( m = 2 \) that \( I_{kl} \) also depends only on \( k - l \):
\[
I_{kl} = I_{k-l}.
\]

Then (5.8) can be rewritten as
\[
(\theta + 2m + \delta_{k0}) I_k = \sum_{l \neq 0} \bar{m}_l I_{k+l} + \delta_{k0},
\]
where \( \bar{m}_l = m_l + m_{-l} \).

Consider the Banach space
\[
\mathcal{A} = \{a = \{a_k\}_{k \in \mathbb{Z}^d}; \sup_k |a_k| < \infty\}
\]
with sup norm.

**Lemma 5.3.** \( I = \{I_k\} \) is characterized as the unique element in \( \mathcal{A} \) satisfying (5.9).

**Proof.** Define operators \( \Delta \) and \( M \) on \( \mathcal{A} \) by
\[
(\Delta a)_k = \delta_{k0} a_k \quad \text{and} \quad (Ma)_k = \sum_{l \neq 0} \bar{m}_l a_{k+l}
\]
respectively, and write (5.9) as
\[
(\theta + 2m) I + \Delta I = MI + d
\]
where \( d = \{\delta_{k0}\} \). Multiplying by \( \Delta \) on both sides of (5.9'), we obtain
\[
(\theta + 2m + 1) \Delta I = \Delta MI + d.
\]
Combining (5.9') with (5.10), we have

\[
(\theta + 2m)I - \left(1 - \frac{1}{\theta + 2m + 1}\Delta\right)M + \frac{\theta + 2m}{\theta + 2m + 1}d,
\]

where 1 denotes the identity operator. Here we note

\[
\left\|1 - \frac{1}{\theta + 2m + 1}\Delta\right\|_{\mathscr{A} \rightarrow \mathscr{A}} = 1 \quad \text{and} \quad \|M\|_{\mathscr{A} \rightarrow \mathscr{A}} \leq \sum_{l=0}^{\infty} \bar{m}_l = 2m.
\]

Therefore if we set

\[H = \frac{1}{\theta + 2m}\left(1 - \frac{1}{\theta + 2m + 1}\Delta\right)^*M,\]

then the inverse of \(1 - H\) on \(\mathscr{A}\) exists and by (5.11)

\[I = \frac{1}{\theta + 2m + 1}(1 - H)^{-1}d. \quad \square\]

Having proved this lemma, we can also show \(\sum_k I_k < \infty\). Indeed following the argument of the proof of the lemma, where \(\mathscr{A}\) is replaced by

\[\mathscr{A}_1 = \left\{a = \{a_k\}; \sum_k |a_k| < \infty\right\},\]

we obtain uniqueness among solutions of the equation

\[(\theta + 2m)a + \Delta a = Ma + d\]

in \(\mathscr{A}_1 (\subset \mathscr{A})\). So by Lemma 5.3 the solution is just 1.

Our final concern is the rate of decrease of \(I_k\) as \(|k| \to \infty\), where \(|k| = |k_1| + \cdots + |k_d|, k = (k_1, \ldots, k_d) \in \mathbb{Z}^d\).

**Theorem 5.4.** Assume that there exists \(\lambda > 1\) such that

\[
\sum_{k \neq 0} m_k \lambda^{|k|} < \infty.
\]

Then

\[
\limsup_{n \to \infty} \frac{1}{n} \log(\max_{k:|k|=n} I_k) \leq -\log \lambda^* \quad \text{(5.13)}
\]

where

\[
\lambda^* = \sup\left\{\lambda > 1; \sum_{k \neq 0} m_k \lambda^{|k|} < \frac{1}{2} \theta + m\right\}.
\]
Remark 5.5. (i) (5.12) and the dominated convergence theorem imply
\[
\lim_{\lambda \uparrow 1} \sum_{k \neq 0} m_k \lambda^{|k|} = \sum_{k \neq 0} m_k = m < \frac{1}{2} \theta + m,
\]
and hence \( \lambda^* > 1 \).

(ii) If \( \{m_k\} \) is of finite range, i.e., there exists a positive integer \( R \) such that
\[
m_k = 0 \quad \text{whenever } |k| > R,
\]
then (5.12) obviously holds for all \( \lambda > 1 \) and therefore \( \lambda^* \) is given as the positive root of the equation in \( \lambda \),
\[
\sum_{k \neq 0} m_k \lambda^{|k|} = \frac{1}{2} \theta + m.
\]

Proof of Theorem 5.4. For \( \lambda > 1 \), let \( \mathcal{A}_\lambda \) be the Banach space
\[
\left\{ a = \{a_k\}; \sum_k |a_k| \lambda^{|k|} < \infty \right\}
\]
with norm \( \|a\|_\lambda = \sum_k |a_k| \lambda^{|k|} \). Consider the operators \( \Delta, M \) and \( H \) introduced in the proof of Lemma 5.3. It is easy to verify
\[
\left\| 1 - \frac{1}{\theta + 2m + 1} \Delta \right\|_{\mathcal{A}_\lambda \rightarrow \mathcal{A}_\lambda} = 1 \quad \text{and} \quad \|M\|_{\mathcal{A}_\lambda \rightarrow \mathcal{A}_\lambda} \leq \sum_{k \neq 0} \tilde{m}_k \lambda^{|k|}.
\]
Therefore for \( \lambda \) satisfying (5.12), \( H \) is a bounded operator on \( \mathcal{A}_\lambda \) with
\[
\|H\|_{\mathcal{A}_\lambda \rightarrow \mathcal{A}_\lambda} \leq \frac{1}{\theta + 2m + 1} \sum_{k \neq 0} \tilde{m}_k \lambda^{|k|}.
\]
Furthermore arguing as in the proof of Lemma 5.3, it follows that if \( \lambda \) satisfies
\[
\sum_{k \neq 0} \tilde{m}_k \lambda^{|k|} < \theta + 2m, \quad \text{or equivalently} \quad \sum_{k \neq 0} m_k \lambda^{|k|} < \frac{1}{2} \theta + m,
\]
then
\[
I = \frac{1}{\theta + 2m + 1} (1 - H)^{-1} d \quad \text{in } \mathcal{A}_\lambda,
\]
in particular
\[
\limsup_{n \to \infty} \frac{1}{n} \log \left( \max_{|k| = n} I_k \right) \leq - \log \lambda.
\]
Letting \( \lambda \uparrow \lambda^* \), we have (5.13). \( \square \)

This result shows the effect of distance on genetic differences between colonies.

Acknowledgements

The author wishes to thank Professors T. Hida and T. Funaki for helpful comments. Particular thanks are due to Professor A. Shimizu for very useful discussions and suggestions. The author would also like to thank the unknown referee for careful reading of the manuscript and for useful suggestions.
References