Clustering of Linearly Interacting Diffusions
and Universality of their Long-Time Distribution

J. M. Swart

Report No. 9923 (May 1999)
Clustering of Linearly Interacting Diffusions and Universality of their Long-Time Distribution

J. M. Swart*

Abstract

We study infinite systems of diffusions indexed by an Abelian group $\Lambda$ and taking values in a compact convex set $K \subset \mathbb{R}^d$ ($d \geq 1$). Each diffusion is subject to: (1) a linear drift towards diffusions at surrounding positions, weighted with an interaction kernel $a(\cdot)$ on $\Lambda$, and (2) a diffusion with local rate $\sigma(\cdot)$ on $K$.

For one-dimensional $K$, it is known that the system clusters (that is, becomes locally flat) if and only if the random walk on $\Lambda$ with symmetrized kernel $a_S(\cdot) := a(\cdot) + a(-\cdot)$ is recurrent. We investigate the generalization of this statement to higher-dimensional $K$, focusing on a comparison argument that has been used in the one-dimensional case. We show that this argument is linked to the universality of the long-time distribution of the system, within the class of recurrent interaction kernels $a_S$, and this universality is in turn shown to follow from a condition involving the harmonic functions of the system. Under this condition we prove that the system clusters and we determine its long-time distribution. We give a general formula for certain special diffusion matrices that have previously appeared in the renormalization of the system, and we argue that universality properties found in this renormalization analysis find their origin in the same condition on the harmonic functions that we use.

AMS 1991 subject classifications. 60H10, 60J60, 60K35.

Key words and phrases. Interacting diffusions, clustering, universality, population genetics.

Acknowledgment. Research supported by the Dutch Association for Mathematical Physics. We thank Prof. dr. F. den Hollander and Prof. dr. A. van Rooij for valuable comments during the preparation of the manuscript.

*Mathematical Institute, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands, E-mail: janswart@sci.kun.nl.
1 Introduction and main results

1.1 Definitions

We consider models of linearly interacting diffusion processes. Models of this type were introduced in population biology and have been the subject of a considerable amount of mathematical work. We consider a family

\[ X = (X_i)_{i \in \Lambda} = (X_i(t))_{t \geq 0, i \in \Lambda} \]  

(1.1)

of stochastic processes, solving a system of stochastic differential equations of the following type:

\[ dX_i(t) = \sum_{j \in \Lambda} a(j - i)(X_j(t) - X_i(t))dt + \sigma(X_i(t))dB_i(t) \quad (i \in \Lambda, t \geq 0). \]  

(1.2)

Here the following definitions apply.

- The \((B_i)_{i \in \Lambda}\) are standard Brownian motions, independent of each other and of the initial condition \(X(0)\).
- The index set \(\Lambda\) is a finite or countable Abelian group, with
  
  \[ \text{group product} \quad i + j \]
  
  \[ \text{inverse} \quad -i \]
  
  \[ \text{unit element} \quad 0. \]  

(1.3)

For example, \(\Lambda\) may be the \(n\)-dimensional integer lattice \(\mathbb{Z}^n\) or the \(N\)-dimensional hierarchical group \(\Omega_N\) (as in [1], [7] and [11]). We sometimes refer to \(i + j\) as addition and to \(0\) as the origin.

- The interaction kernel \(a : \Lambda \to \mathbb{R}\) satisfies
  
  \[ a(i) \geq 0 \quad (i \in \Lambda) \]
  
  \[ \sum_{i \in \Lambda} a(i) < \infty. \]  

(1.4)

It is the kernel of a continuous-time random walk on \(\Lambda\) that jumps from a point \(i\) to a point \(j\) with rate \(a(j - i)\). We assume that this random walk is irreducible.

- Each single component \(X_i(t)\) takes values in a state space \(K\) that is a non-empty convex compact subset of \(\mathbb{R}^d\). Thus, each component \(X_i(t)\) itself exists of \(d\) components:
  
  \[ X_i(t) = (X^1_i(t), \ldots, X^d_i(t)). \]  

(1.5)

Equation (1.2) componentwise reads

\[ dX^\alpha_i(t) = \sum_{j} a(j - i)(X^j_i(t) - X^\alpha_i(t))dt + \sum_{\beta} \sigma_{\alpha\beta}(X_i(t))dB^\beta_i(t) \quad (i \in \Lambda, \alpha = 1, \ldots, d, t \geq 0). \]  

(1.6)

We adopt the convention that sums over Roman indices \(i, j, k, \ldots\) range over \(\Lambda\), while sums over Greek indices \(\alpha, \beta, \gamma, \ldots\) range from 1 to \(d\).
• The function $\sigma$ is a continuous function from $K$ into $\mathbb{R}^d \otimes \mathbb{R}^d$, the space of $d \times d$ real matrices. It is a root of the diffusion matrix $w : K \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$:

$$w_{\alpha\beta}(x) := \frac{1}{2} \sum_{\gamma} \sigma_{\alpha\gamma}(x)\sigma_{\beta\gamma}(x).$$  \hspace{1cm} (1.7)

We assume that $w$ satisfies

$$\sum_{\alpha,\beta} \varepsilon^{\alpha} w_{\alpha\beta}(x) z^{\beta} = 0 \quad \forall x \in K, \ z \in I_x^\perp,$$  \hspace{1cm} (1.8)

where $I_x^\perp$ is the space of vectors perpendicular to

$$I_x := \{ y \in \mathbb{R}^d : \exists \varepsilon > 0 \text{ such that } x + \lambda y \in K \ \forall |\lambda| \leq \varepsilon \}.$$  \hspace{1cm} (1.9)

$I_x$ is the space of directions in which the boundary of $K$ at $x$ is flat. In terms of the process $X$, condition (1.8) guarantees that the components $X_i(t)$ cannot leave the state space $K$.

We equip the space $K^\Lambda$ with the product topology. In this topology $K^\Lambda$ is a compact separable metrizable space. $C(K^\Lambda)$ is the Banach space of continuous real-valued functions on $K^\Lambda$, equipped with the supremum norm $\| \cdot \|_\infty$.

Solutions to (1.2), whenever they exist, are continuous $K^\Lambda$-valued processes that solve the martingale problem for a linear operator $A$ on $C(K^\Lambda)$ given by

$$(Af)(x) := \left( \sum_{ij} \sum_{\alpha} a(j-i)(x_j^\alpha - x_i^\alpha) \frac{\partial}{\partial x_i^\alpha} + \sum_{i} \sum_{\alpha,\beta} w_{\alpha\beta}(x_i) \frac{\partial}{\partial x_i^\alpha} \frac{\partial}{\partial x_i^\beta} \right)f(x).$$  \hspace{1cm} (1.10)

Here $f$ is a real function on $K^\Lambda$, and a typical element $x \in K^\Lambda$ is written as

$$x = (x_i)_{i \in \Lambda} = (x_i^\alpha)^{\alpha = 1, \ldots, d}.$$  \hspace{1cm} (1.11)

The operator $A$ in (1.10) has domain

$$D(A) := C^2_{\text{int}}(K^\Lambda),$$  \hspace{1cm} (1.12)

the space of all $C^2$-functions depending on finitely many coordinates only. For such functions the infinite sums of derivatives in (1.10) reduce to finite sums. Condition (1.8) guarantees that the operator $A$ in (1.10) satisfies the maximum principle.

### 1.2 Existence and uniqueness: Theorems 1.1 and 1.2

We focuss our attention on shift-invariant solutions to (1.2). For $j \in \Lambda$, let the shift operator $T_j : K^\Lambda \rightarrow K^\Lambda$ be defined as

$$(T_j x)_i := x_{i-j}.$$

We say that a solution $X$ to (1.2) is shift-invariant if for each $j \in \Lambda$ the processes $(X(t))_{t \geq 0}$ and $(T_j X(t))_{t \geq 0}$ have the same finite-dimensional distributions. We say that a probability measure $\mu$ on $K^\Lambda$ (equipped with the product-$\sigma$-field) is shift-invariant if $\mu = \mu \circ T_j^{-1}$ for all $j \in \Lambda$. 
Theorem 1.1 For each probability measure $\mu$ on $K^\Lambda$, there exists a solution $(X(t))_{t \geq 0}$ to (1.2) with initial condition $\mathcal{L}(X(0)) = \mu$ and sample paths in the continuous functions from $[0, \infty)$ to $K^\Lambda$. If the $\mu$ is shift-invariant, then (1.2) has a shift-invariant solution with the same properties.

If solutions to (1.2) are weakly unique, then any solution with a shift-invariant initial condition must be shift-invariant. Unfortunately, it is at present not very well understood when weak uniqueness holds for (1.2). Standard techniques give:

Theorem 1.2 Assume that the function $\sigma : K \to \mathbb{R}^d \otimes \mathbb{R}^d$ is Lipschitz continuous. Then, for each $K^\Lambda$-valued initial condition $X(0)$, strong uniqueness holds for equation (1.2).

Strong uniqueness for (1.2) implies weak uniqueness, just as in the case of finite-dimensional stochastic differential equations [22]. Theorem 1.2 does not cover many interesting cases. For example, for the Wright-Fisher diffusion matrix (see (1.15)), no root $\sigma$ of $w$ exists that satisfies the conditions of Theorem 1.2. For uniqueness results in this and a few more special cases, see [19, 20, 22]. In what follows, we avoid problems of uniqueness by assuming only the existence of solutions to (1.2).

1.3 Biological background

In population biology, models of the form (1.2) are used to describe the genetic composition of a population of individuals as a function of time. It is supposed that the population is divided into colonies, each containing a large number of individuals. A component $X_i(t) \in K$ describes the genetic composition of the population in colony $i$ at time $t$. A typical choice for $K$ is

$$K = S^{p-1} := \{ (x^1, \ldots, x^{p-1}) : x^a \geq 0 \ \forall a, \sum_a x^a \leq 1 \}. \quad (1.14)$$

For $p = 2$ we have a ‘2-type model’. In such a model a gene comes in two types (‘alleles’), say type I and II. $X_i(t) \in S^1 = [0, 1]$ is interpreted as the relative frequency of type I in colony $i$ at time $t$, the relative frequency of type II being $1 - X_i(t)$. More generally, in a ‘$p$-type model’ one considers the relative frequency of $p$ types. $X_i(t)$ is the relative frequency of the $\alpha$-th type ($\alpha = 1, \ldots, p - 1$), the relative frequency of the remaining $p$-th type being $1 - \sum_{\alpha} X_i^p(t)$.

The genetic compositions $X_i(t)$ change in time due to migration and resampling. Individuals in the population migrate between colonies according to a continuous-time random walk, jumping from $j$ to $i$ with rate $a(j - i)$. This migration causes an attractive interaction between components, expressed by the drift term $\sum_j a(j - i)(X_j(t) - X_i(t))dt$ in equation (1.2). At each colony individuals are after an exponential waiting time replaced by individuals of a type chosen at random from the colony. This resampling is expressed by the diffusion term $\sigma(X_i(t))dB_i(t)$ in equation (1.2), where $\sigma$ is any continuous root of the diffusion matrix $w$. A typical choice for $w$ is the Wright-Fisher diffusion matrix $w_{\alpha\beta}(x) = x^\alpha(\delta_{\alpha\beta} - x^\beta)$. \quad (1.15)
For other choices, see the models listed in [3] and [21]. Often one's aim is to prove statements for as wide a class of diffusion matrices as possible. The following examples are found in the literature.

**2-type models** References [1, 7, 12] are concerned with diffusion functions \( w : [0,1] \to [0,\infty) \), Lipschitz continuous, satisfying
\[
w(x) = 0 \Leftrightarrow x \in \{0,1\}.
\]

**Isotropic models** Reference [11] is concerned with diffusion matrices of the form
\[
w_{\alpha\beta}(x) = \delta_{\alpha\beta} g(x),
\]
where \( g : K \to [0,\infty) \) is a nice function satisfying
\[
g(x) = 0 \Leftrightarrow x \in \partial K,
\]
with \( \partial K \) the (topological) boundary of \( K \).

Work on the non-compact state space \( K = [0,\infty) \) can be found in [2, 8]. Reference [6] is concerned with an isotropic model on \( K = [0,\infty)^2 \). A generalization of the \( p \)-type model to infinitely many types is studied in [9].

### 1.4 The non-interacting model

Later on, we will make a comparison between the model in (1.2) and a model without interaction. For this, we consider the case that the Abelian group \( \Lambda \) consists of only one element. Equation (1.2) now reduces to
\[
dX(t) = \sigma(X(t))dB(t) \quad (t \geq 0),
\]
where \( (X(t))_{t \geq 0} \) is a \( K \)-valued stochastic process. Uniqueness of solutions to (1.19) can be proved under considerably weaker conditions than those needed for equation (1.2) (see the examples in section 1.10). We therefore prefer, when possible, to assume only existence of solutions to (1.2) and uniqueness of solutions to (1.19).

Solutions to (1.19) are bounded martingales, and hence they have a last element:
\[
X^x(t) \to X^x(\infty) \quad \text{a.s. as } t \to \infty \quad (x \in K),
\]
where \( X^x \) is the solution of (1.19) starting in \( x \). \( X^x(\infty) \) takes values in the set
\[
\partial w K := \{ x \in K : w_{\alpha\beta}(x) = 0 \ \forall \alpha, \beta \}.
\]
In typical examples, \( \partial w K \) is a subset of the (topological) boundary of \( K \). We call \( \partial w K \) the effective boundary of \( K \). We denote the law of \( X^x(\infty) \) by
\[
\Gamma_x := \mathcal{L}(X^x(\infty)).
\]
The collection \((\Gamma_x)_{x \in K}\) we call the boundary distribution associated with the diffusion matrix \( w \). We note that different diffusion matrices \( w \) may share the same boundary
distribution. For example, by the martingale property of solutions to (1.19), diffusions on $[0,1]$ with $w$ as in (1.16) all have

$$\Gamma_x = (1-x)\delta_0 + x\delta_1. \tag{1.23}$$

For diffusions with isotropic $w$ as in (1.17), solutions to (1.19) are time-transformed Brownian motions, and therefore

$$\Gamma_x = \mathcal{L}(B^x_t), \tag{1.24}$$

where $(B^x_t)_{t \geq 0}$ is Brownian motion starting in $x$ and

$$\tau := \inf\{t \geq 0 : B_t \in \partial K\}. \tag{1.25}$$

We try to answer two questions. When does the distribution of components $X_i(t)$ of the interacting system in (1.2) converge to a distribution on the effective boundary $\partial wK$? And when is this limiting distribution actually the same as in the non-interacting system?

### 1.5 Clustering: Theorem 1.3

In order to state our first result, we introduce the symmetrized kernel

$$a_s(i) := a(i) + a(-i) \quad (i \in \Lambda). \tag{1.26}$$

By the random walk with kernel $a_s$ we mean a continuous-time random walk on $\Lambda$ that jumps from a point $j$ to a point $i$ with rate $a_s(j-i)$. By $\Rightarrow$ we denote weak convergence of probability measures on $K^\Lambda$.

**Theorem 1.3** Let $X$ be a shift-invariant solution to (1.2) and assume that there exists a $K^\Lambda$-valued random variable $X(\infty)$ such that

$$X(t) \Rightarrow X(\infty) \quad \text{as } t \to \infty. \tag{1.27}$$

If the random walk with kernel $a_s$ is recurrent, then

(i) $P[X_i(\infty) \in \partial_w K \quad \forall i \in \Lambda] = 1$

(ii) $P[X_i(\infty) = X_j(\infty) \quad \forall i,j \in \Lambda] = 1. \tag{1.28}$

If the random walk with kernel $a_s$ is transient, $E[X_0(0)] \not\in \partial_w K$ and $\mathcal{L}(X(0))$ is spatially ergodic, then

(i) $P[X_i(\infty) \in \partial_w K] < 1 \quad \forall i \in \Lambda$

(ii) $P[X_i(\infty) = X_j(\infty)] < 1 \quad \forall i \neq j \in \Lambda. \tag{1.29}$

Note that Theorem 1.3 makes a statement about the possible properties of a limiting distribution $X(\infty)$, but that it does not answer the question whether such a limiting distribution actually exists. Provided we know in some way that $X(t)$ converges weakly to a limit, Theorem 1.3 says the following.
In the \textit{recurrent} case, the configuration in any finite window \( \Delta \subset \Lambda \) after a sufficiently long time becomes almost flat. At large but finite time there are in the system regions, called ‘clusters’, of typical sizes that grow with time, in which all components are almost equal. This behavior is called ‘clustering’. The behavior is similar to that of the voter model in low \( (d \leq 2) \) dimension. In fact, 2-type models as in (1.16) are believed to be asymptotically equivalent, in some sense, to the voter model on the same lattice. See \cite{5} for some pictures of simulations of the (clustering) voter model on \( \mathbb{Z}^2 \).

In the \textit{transient} case, such clustering behavior cannot occur. Instead, the system converges to a ‘true’ equilibrium \( X(\infty) \). We refer to this as ‘stable’ behavior.

Although it seems hard to imagine a shift-invariant solution to (1.2) that does not converge as \( t \to \infty \), the convergence in (1.27) is in general hard to prove. For finite \( \Lambda \), one may exploit the fact that \( \sum_i X_i(t) \) is a bounded martingale to get the convergence in (1.27), not only in the sense of weak convergence, but also in \( L^2 \)-norm. For infinite \( \Lambda \), convergence in \( L^2 \)-norm in general does not hold.

For the 2-type model, the convergence in (1.27) has been proved in \cite{4, 16}. In \cite{4}, this is achieved for transient \( a_S \) by a coupling technique, and for recurrent \( a_S \) by a ‘duality comparison argument’. This argument, as well as Theorem 1.3, are based on calculations involving covariances between components.

1.6 Covariance calculations: Lemma 1.4

In this section we explain the relation between covariance calculations, the random walk with kernel \( a_S \), and clustering properties of the system in (1.2).

For any two \( K \)-valued random variables \( X \) and \( Y \) the covariance of \( X \) and \( Y \) is the quantity

\[
\text{Cov}(X, Y) = E[X \cdot Y] - E[X] \cdot E[Y],
\]

where \( \cdot \) denotes the inner product \( x \cdot y = \sum_{a} x^a y^a \). By \( \text{tr}(w) \) we denote the trace

\[
\text{tr}(w)(x) = \sum_{a=1}^{d} w_{aa}(x) \quad (x \in K)
\]

of the diffusion matrix \( w \). The following lemma follows from a little calculation involving Itô’s formula and a bit of continuity.

\textbf{Lemma 1.4} Let \( X \) be a shift-invariant solution of (1.2). Then there exists a \( \theta \in K \) such that

\[
E[X_i(t)] = \theta \quad (t \geq 0, \ i \in \Lambda)
\]

and there exists a function \( C : [0, \infty) \times \Lambda \to \mathbb{R} \) such that

\[
\text{Cov}(X_i(t), X_j(t)) = C_t(j - i) \quad (t \geq 0, \ i, j \in \Lambda).
\]

For each \( i \), the function \( t \mapsto C_t(i) \) is continuously differentiable and satisfies

\[
\frac{\partial}{\partial t} C_t(i) = \sum_j a_S(j - i)(C_t(j) - C_t(i)) + 2\delta_{ij} E[\text{tr}(w)(X_0(t))].
\]
The right-hand side of (1.34) contains the operator

\[(Gf)(i) := \sum_j a_S(j - i)(f(j) - f(i)), \tag{1.35}\]

acting on bounded functions \( f : \Lambda \to \mathbb{R} \). \( G \) is the generator of the random walk with kernel \( a_S \). For solutions to (1.34) we have the representation

\[ C_t(i) = \sum_j P_t(j - i)C_0(j) + 2\int_0^t P_s(0 - i)E[tr(w)(X_0(t - s))]ds, \tag{1.36} \]

where \( P_t(j - i) \) is the probability that the random walk with kernel \( a_S \) starting from a point \( i \), is in \( j \) at time \( t \).

In view of the biological background of the model, the representation in (1.36) can be understood in terms of a 'historic process' tracing back where ancestors of two individuals from colonies at 0 and \( i \) lived at previous times. The time the symmetrized random walk spends at the origin is the time the ancestors lived in the same colony, and hence had a chance of descending from a common ancestor.

This sort of reasoning works best when \( w \) is the Wright-Fisher diffusion matrix. In that case the system (1.2) is in duality with a system of delayed coalescing random walks (see formula (4.1) in [12] or Lemma 2.3 in [18]) and all mixed moments of the type \( E[X_i(t)X_j(t)], E[X_i(t)X_j(t)X_k(t)], \ldots \) may be expressed in terms of the dual model. This duality has been exploited in [18] to show the dichotomy between clustering and stable behavior for the Wright-Fisher diffusion on \([0, 1]\).

For arbitrary \( w \), the representation (1.36) is sufficient to derive Theorem 1.3, but not to derive the convergence in (1.27). For 2-type models as in (1.16), this shortcoming can be overcome by using a 'duality comparison argument' as in [4] (see also [3]), which makes a comparison between models with arbitrary \( w \) and the special model with Wright-Fisher diffusion, for which clustering is known by duality.

### 1.7 Universality of the long-time distribution: Theorem 1.5

We give sufficient conditions for the convergence in (1.27) and for the uniqueness in distribution of the limit \( X(\infty) \). For this we need to look at the differential equation

\[ dY(t) = (\theta - Y(t))dt \quad (t \geq 0), \tag{1.37} \]

where \( \theta \in K \) is a fixed parameter. By the convexity of \( K \), the solution of (1.37) starting from a point \( x \in K \):

\[ Y^x(t) = \theta + (x - \theta)e^{-t} \quad (t \geq 0), \tag{1.38} \]

stays in \( K \) for all time. Solutions to (1.37) are associated with a semigroup \((T_{\theta,t})_{t \geq 0}\) on the space \( B(K) \) of bounded measurable real functions on \( K \), given by

\[ (T_{\theta,t}f)(x) := E[f(Y^x(t))] = f(\theta + (x - \theta)e^{-t}) \quad (x, \theta \in K, \ t \geq 0). \tag{1.39} \]
We are going to compare equation (1.37) (non-zero drift, zero diffusion) with the non-interacting equation (1.19) (zero drift, non-zero diffusion).

Let us assume that for each initial condition \( x \in K \), the non-interacting equation (1.19) has a unique weak solution \((X^x(t))_{t \geq 0}\), and let us denote the associated semigroup on \( B(K) \) by

\[
(S_t f)(x) := E[f(X^x(t))] \quad (x \in K, \ t \geq 0).
\]

We add a ‘last element’ \( S_\infty \) to this semigroup by defining

\[
(S_\infty f)(x) := \int_K \Gamma_x(dy)f(y) \quad (x \in K, \ f \in B(K)),
\]

where \((\Gamma_x)_{x \in K}\) is the boundary distribution associated with \( w \), introduced in (1.22).

With this notation, we formulate a condition that will guarantee that the long-time behavior of the non-interacting model is not changed by the introduction of a linear drift.

**Definition 1.1** Let \( w \) be a diffusion matrix on \( K \) such that weak uniqueness holds for (1.19), and let \((\Gamma_x)_{x \in K}\) be the associated boundary distribution. We say that \((\Gamma_x)_{x \in K}\) is stable against a linear drift if

\[
S_\propto T_{\theta \cdot t} S_\infty f = T_{\theta \cdot t} S_\infty f \quad \forall \theta \in K, \ t \geq 0, \ f \in B(K).
\]

Since \( S_\infty S_\infty = S_\infty \), we can read equation (1.42) as: \( S_\infty \) and \( T_{\theta \cdot t} \) commute on functions of the form \( S_\infty f \).

For technical reasons, we will restrict ourselves to the case that

\[
S_\infty(C(K)) \subset C(K).
\]

This condition guarantees that \( S_\infty f \) is a \( w \)-harmonic function for all \( f \in C(K) \), where the space of \( w \)-harmonic functions is defined as

\[
H := \{ f \in D(G) : Gf = 0 \},
\]

with \( G \) the full generator of the process in (1.19) and \( D(G) \) its domain. In particular, \( C^2 \)-functions are \( w \)-harmonic if and only if they solve the equation

\[
\sum_{\alpha \beta} w_{\alpha \beta}(x) \frac{\partial^2}{\partial x^\alpha \partial x^\beta} f(x) = 0 \quad (x \in K).
\]

It turns out that condition (1.42) is equivalent to

\[
T_{\theta \cdot t}(H) \subset H \quad \forall \theta \in K, \ t \geq 0.
\]

That is, for each \( \theta \) the space of \( w \)-harmonic functions is invariant under the semigroup \((T_{\theta \cdot t})_{t \geq 0}\).

With these definitions, our main result reads as follows.
Theorem 1.5 Let $X$ be a shift-invariant solution to (1.2) such that $\mathcal{L}(X(0))$ is spatially ergodic and
\[
E[X_i(0)] = \theta \quad (i \in \Lambda) \tag{1.47}
\]
for some $\theta \in K$. Assume that weak uniqueness holds for the non-interacting equation (1.19), that the associated boundary distribution is stable against a linear drift, that $S_\infty(\mathcal{C}(K)) \subset \mathcal{C}(K)$ and that $H$ is contained in the bp-closure of $\mathcal{C}^2(K) \cap H$. If the random walk with kernel $a_S$ is recurrent, then there exists a $K^\Lambda$-valued random variable $X(\infty)$ such that
\[
X(t) \Rightarrow X(\infty) \quad \text{as } t \to \infty, \tag{1.48}
\]
where
\[
\mathcal{L}(X_i(\infty)) = \Gamma_{\theta} \quad (i \in \Lambda). \tag{1.49}
\]
The bp-closure of a set is the smallest set containing it that is closed under bounded pointwise limits.

Note that by Theorem 1.3, $P[X_i(\infty) = X_j(\infty) \quad \forall i, j \in \Lambda] = 1$. Thus, the fact that the boundary distribution is stable against a linear drift not only allows us to conclude that $X(t)$ converges to a limit $X(\infty)$, it also allows us to completely specify its distribution. This distribution turns out to be universal in all recurrent random walk kernels $a_S$ and Abelian groups $\Lambda$, and in all diffusion matrices $w$ sharing the same boundary distribution $(\mathcal{C}(K) \cap H)$.

1.8 Harmonic functions: Lemma 1.6

To see what goes into proving Theorem 1.5, we mention the following:

Lemma 1.6 Let $X$ be a solution to (1.2). Assume that weak uniqueness holds for the non-interacting equation (1.19), that the associated boundary distribution is stable against a linear drift, that $S_\infty(\mathcal{C}(K)) \subset \mathcal{C}(K)$ and that $H$ is contained in the bp-closure of $\mathcal{C}^2(K) \cap H$. Then
\[
E[f(X_i(t))] = E\left[f\left(\sum_j P_t(j-i)X_j(0)\right)\right] \quad \forall f \in H, \ i \in \Lambda, \ t \geq 0, \tag{1.50}
\]
where $P_t(j-i)$ is the probability that the random walk with kernel $a$ starting from $i$, is in $j$ at time $t$.

The situation is particularly simple when $X_i(0) = \theta$ for all $i \in \Lambda$. In that case
\[
E[f(X_i(t))] = f(\theta) \quad \forall f \in H, \ i \in \Lambda, \ t \geq 0. \tag{1.51}
\]

For a 2-type model with diffusion matrix $w$ as in (1.16), the class $H$ contains only affine functions $x \mapsto a + bx \ (a, b \in \mathbb{R})$, and (1.51) says no more than that the mean of the components is conserved. Since there is only one distribution on $\{0,1\}$ with a given mean, it is then immediately clear (for recurrent $a_S$) that there is only one possible long-time distribution for the process in (1.2). In the general higher-dimensional case, we need to specify a distribution on the effective boundary $\partial_w K$, and for this we need the expectation of sufficiently many harmonic functions. We may describe (1.51) by saying that the `$w$-harmonic mean' of the components is conserved.
1.9 Special models: Corollary 1.7

The proof of Theorem 1.5 is by a comparison argument, in the spirit of the ‘duality comparison argument’ in [4]. In our comparison argument we use objects related to the special diffusion matrix

\[ w_{\alpha,\beta}(x) := \int_K \Gamma_x(dy)(y^\alpha - x^\alpha)(y^\beta - x^\beta) \quad (x \in K, \alpha, \beta = 1, \ldots, d). \]  

(1.52)

We do not have a duality for the model with \( w^* \), but we can find an expression for second moments, which is enough for our purposes. For the special model with \( w = \lambda w^* \ (\lambda > 0) \) the proof of Theorem 1.5 yields the following corollary.

**Corollary 1.7** In addition to the assumptions in Theorem 1.5, assume that for some \( \lambda \in (0, \infty) \)

\[ w(x) = \lambda w^*(x) \quad (x \in K). \]  

(1.53)

Then for each \( t \geq 0, \ i, j \in \Lambda, \ \alpha, \beta = 1, \ldots, d \)

\[ E[(X^\alpha_i(t) - \theta^\alpha)(X^\beta_j(t) - \theta^\beta)] = w_{\alpha,\beta}^*(\theta)K^\lambda_{\alpha,\beta}(i,j), \]  

(1.54)

where \( K^\lambda_{\alpha,\beta}(i,j) \) denotes the probability that two delayed coalescing random walks, each with kernel \( a \), starting in points \( i \) respectively \( j \) and coalescing with rate \( 2\lambda \), have coalesced before time \( t \).

1.10 Examples

We close this introduction by giving two examples of classes of diffusion matrices \( w \) satisfying the assumptions in Theorem 1.5.

The first example arises when we generalize the 2-type models mentioned in (1.16) to \( p \)-type models in the following way.

**Example 1.8 (p-type models)** Assume that \( K \) is the \((p - 1)\)-dimensional simplex \( S^{p-1} \), and that \( x \mapsto w(x) \) is Lipschitz continuous and satisfies (compare (1.8))

\[ \sum_{\alpha,\beta} z^\alpha w_{\alpha,\beta}(x)z^\beta = 0 \iff z \in I_2^+ \quad (x \in K). \]  

(1.55)

Then:

(a) Weak uniqueness holds for the non-interacting equation (1.19). The boundary distribution is stable against a linear drift, \( \mathcal{S}_\infty(\mathcal{C}(K)) \subset \mathcal{C}(K), \) and \( H \) is contained in the hp-closure of \( \mathcal{C}^2(H) \cap H \).

(b) The class of \( w \)-harmonic functions consists of all affine functions

\[ x \mapsto a + \sum_a b^a x^a \quad (a, b_1, \ldots, b_d \in \mathbb{R}). \]  

(1.56)

(c) The associated special diffusion matrix is the Wright-Fisher diffusion matrix

\[ w_{\alpha,\beta}^*(x) = x^\alpha(\delta_{\alpha,\beta} - x^\beta) \quad (x \in K, \ \alpha, \beta = 1, \ldots, d). \]  

(1.57)
The second example is formed by the class of isotropic diffusion matrices (compare (1.17)).

**Example 1.9 (isotropic models)** Assume that $K$ has non-empty interior $K^\circ$, and let $\partial K := K \setminus K^\circ$ denote its topological boundary. Assume that

$$w_{\alpha\beta}(x) = \delta_{\alpha\beta}g(x) \quad (x \in K, \alpha, \beta = 1, \ldots, d)$$

(1.58)

for some Lipschitz continuous function $g : K \to [0, \infty)$ satisfying

$$g(x) = 0 \iff x \in \partial K.$$  

(1.59)

Then:

(a) Weak uniqueness holds for the non-interacting equation (1.19). The boundary distribution is stable against a linear drift, $S_\infty(C(K)) \subset C(K)$, and $H$ is contained in the bp-closure of $C^2(H) \cap H$.

(b) The class of $w$-harmonic functions is given by

$$H = \left\{ f \in C(K) \cap C^2(K^\circ) : \sum_\alpha \frac{\partial^2}{\partial x^2} f(x) = 0 \text{ on } K^\circ \right\}. \quad (1.60)$$

(c) The associated special diffusion matrix is given by

$$w^*_{\alpha\beta}(x) = \delta_{\alpha\beta}g^*(x) \quad (x \in K, \alpha, \beta = 1, \ldots, d),$$

(1.61)

where $g^* \in C(K) \cap C^2(K^\circ)$ is the unique solution of

$$-\frac{1}{2} \sum_\alpha \frac{\partial^2}{\partial x^2} g^*(x) = 1 \quad (x \in K^\circ)$$

$$g^*(x) = 0 \quad (x \in \partial K).$$

(1.62)

One can find a few more examples of diffusion matrices satisfying the assumptions in Theorem 1.5, but it turns out that these are mainly trivial variations on the two examples mentioned above. The message of Theorem 1.5 is that all these examples fall into the same framework. The common property that unites them is the stability of the boundary distribution against a linear drift.

In fact, we conjecture that this property is a necessary condition for the universality of the long-time distribution. If the boundary distribution is not stable against a linear drift, it seems likely that still $X(t) \to \infty$, although we do not know how to prove this for infinite $\Lambda$. But we believe that in this case the law of $X(\infty)$ will depend on the choice of the recurrent kernel $a_S$ and the Abelian group $\Lambda$. However, we have at present very little knowledge about the nature of this dependence.

In conclusion, we have found that the ‘duality comparison argument’ developed in [4] is linked to universality of the long-time distribution of solutions to (1.2). A similar relation between comparison arguments and universality has been found for models on the hierarchical group $\Omega_N$ with $N$ large in [1, 7, 11]. There, the system in (1.2) is studied by means of a renormalization transformation acting on diffusion matrices.
Under iteration of the transformation, the renormalized diffusion matrices converge to a limit. In the clustering case a comparison argument shows that this limit is universal within a large ‘universality class’ of matrices. This has been worked out for 2-type models in [1] and for isotropic models in [11]. The universal limit that is found is exactly the $w^*$ in formula (1.52). The conclusion we can draw from Theorem 1.5 is that the correct ‘universality classes’ of diffusion matrices one should look at are formed by all diffusion matrices $w$ that share the same boundary distribution $(\Gamma_x)_{x \in K}$. Furthermore, universal behavior can be expected only if this boundary distribution is stable against a linear drift.

2 Proofs of Theorems 1.1 and 1.2

2.1 Proof of Theorem 1.1

If $\Delta \subset \Lambda$ is finite then $C^2(K^\Delta)$ is the space of real functions on $K^\Delta$ that have a $C^2$-extension to all of $(\mathbb{R}^d)^\Delta$. $C^0_{\text{fin}}(K^\Lambda)$ consists of all functions that are the lifting to the larger space $K^\Lambda$ of a function in $C^2(K^\Delta)$ for some finite $\Delta \subset \Lambda$.

**Lemma 2.1** The operator $A$ in (1.10) with domain $D(A)$ in (1.12) is a densely defined linear operator on the Banach space $C(K^\Lambda)$, and satisfies the maximum principle.

**Proof of Lemma 2.1:** By the Stone-Weierstrass theorem, $C^2(K^\Delta)$ is dense in $C(K^\Lambda)$ for each finite $\Delta \subset \Lambda$. Pick a bijection between $\Lambda$ and the positive integers and fix a point $z \in K$. Define, for $x \in K^\Lambda$,

$$\pi_n(x) := (x_1, x_2, \ldots, x_n, z, z, \ldots).$$  

(2.1)

The sets $\pi(K^\Lambda)$ are uniformly dense in $K^\Lambda$, and since each $f \in C(K^\Lambda)$ is uniformly continuous, it is the uniform limit of functions

$$f_n(x) := f(\pi_n(x))$$  

(2.2)

depending on finitely many coordinates. Hence $C^0_{\text{fin}}(K^\Lambda)$ is dense in $C(K^\Lambda)$.

To see that $A$ satisfies the maximum principle, fix $f \in C^0_{\text{fin}}(K^\Lambda)$ and suppose that $f$ assumes its maximum in a point $x$. Fix an $i \in \Lambda$. Keeping all $(x_j)_{j \neq i}$ fixed, $f$ assumes its maximum as a function of the remaining variable in the point $x_i$. By the convexity of $K$ it is easily checked that

$$\sum_j \sum_\alpha a(j - i)(x_j^\alpha - x_i^\alpha) \frac{\partial}{\partial x_i^\alpha} f(x) \leq 0.$$

(2.3)

Condition (1.8) ensures that

$$\sum_{\alpha, \beta} w_{\alpha, \beta}(x_i) \frac{\partial^2}{\partial x_i^\alpha \partial x_i^\beta} f(x) \leq 0,$$

(2.4)

as can be seen by writing the matrix $w(x)$ in diagonal form:

$$\sum_{\alpha, \beta} w_{\alpha, \beta}(x_i) \frac{\partial^2}{\partial x_i^\alpha \partial x_i^\beta} f(x) = \sum_\alpha \lambda_\alpha(x) \frac{\partial^2}{\partial x_i^\alpha} f(x)$$

(2.5)

14
for an appropriate orthonormal basis \((x^a)\) of \(\mathbb{R}^d\). By condition (1.8), the only non-zero terms in (2.5) occur for directions that lie in the space \(I_x\), and for such directions the second derivative is non-positive.

We equip \(K^A\) with the Borel \(\sigma\)-field generated by the open sets. We write \(D_{K^A}[0,\infty)\) for the cadlag functions from \([0,\infty)\) to \(K^A\), equipped with the metric \(d\) from chapter 3, section 5 of [10], which generates the Skorohod topology. By \(C_{K^A}[0,\infty)\) we denote the continuous functions from \([0,\infty)\) to \(K^A\). On \(D_{K^A}[0,\infty)\) we choose the Borel \(\sigma\)-field generated by the open sets of this topology. We equip the probability measures on \(D_{K^A}[0,\infty)\) with the topology of weak convergence and we denote weak convergence of processes with sample paths in \(D_{K^A}[0,\infty)\) by \(\Rightarrow\). Thus \(X_n \Rightarrow X\) means that

\[
E[f(X_n)] \to E[f(X)] \quad \text{as } n \to \infty
\]

for all bounded continuous real functions \(f\) on \(D_{K^A}[0,\infty)\). By a solution to the martingale problem we always mean a solution with sample paths in \(D_{K^A}[0,\infty)\).

**Lemma 2.2** For each probability measure on \(K^A\) there exists a solution to the martingale problem for \(A\) with initial condition \(\mu\). Each solution to the martingale problem for \(A\) has sample paths in \(C_{K^A}[0,\infty)\). The space of solutions to the martingale problem for \(A\) is compact in the topology of weak convergence. If \(X_n, X\) solve the martingale problem for \(A\), then \(X_n \Rightarrow X\) implies \(X_n(t) \Rightarrow X(t)\) for all \(t \geq 0\).

**Proof of Lemma 2.2:** Existence of solutions to the martingale problem for \(A\) follows from Lemma 2.1 in combination with Theorem 5.4 and Remark 5.5 from chapter 4 of [10].

The continuity of sample paths can be shown by Problem 19 from the same chapter: for this one needs to find for every \(x \in K^A\) a function \(f_x \in \mathcal{D}(A)\) such that for every \(\varepsilon > 0\)

\[
\inf\{f_x(y) - f_x(x) : x, y \in K^A, \quad d(x, y) \geq \varepsilon\} > 0 \tag{2.7}
\]

and such that \(\lim_{x \to y} A f_x(y) = A f_y(y) = 0\) for all \(x \in K^A\). Instead of working with \(A\), one may also use the closure of \(A\). Applying Lemma 4.5 below and defining \((\gamma_i)_{i \in A}\) as in (2.14), it is not hard to check that the functions

\[
f_x(y) := \sum_i \gamma_i |x_i - y_i|^q \tag{2.8}
\]

satisfy the requirements.

Compactness of the space of solutions follows from Lemma 5.1 and Remark 5.2 from chapter 4 of [10]. Finally, weak convergence in path space of solutions \(X_n\) to the martingale problem for \(A\) implies convergence of finite-dimensional distributions by Theorem 7.8 from chapter 3 of [10] and the continuity of sample paths.

**Proof of Theorem 1.1:** Solutions to the martingale problem for \(A\) are guaranteed by Lemma 2.2. Corollary 3.4 from chapter 5 of [10] generalizes in a straightforward way to the infinite-dimensional case, and so for each solution to the martingale problem for \(A\) we can find a weak solution to the stochastic differential equation (1.2).
We next show that for each shift-invariant initial condition \( \mu \), there exists a shift-invariant solution to (1.2). It suffices to construct a shift-invariant solution to the martingale problem for \( A \). We define a shift operation on \( \mathcal{D}_{K,\Lambda}[0,\infty) \) in the obvious way, by putting
\[
(T_j x)(t) := x_{i-j}(t) \quad (i,j \in \Lambda, \ t \geq 0).
\]
Let \( X \) be a solution to the martingale problem for \( A \) with initial condition \( \mathcal{L}(X(0)) = \mu \). By Lemma 3.3 below, there exists a sequence of functions \( p_n : \Lambda \to [0,\infty) \) such that \( \sum p_n(i) = 1 \) for each \( n \) and
\[
\lim_{n \to \infty} \sum_k |p_n(i - k) - p_n(j - k)| = 0 \quad \forall i, j \in \Lambda.
\]
Let \( (X_n) \) be a sequence of processes with sample paths in \( \mathcal{D}_{K,\Lambda}[0,\infty) \) with law
\[
\mathcal{L}(X_n) = \sum_k p_n(k)\mathcal{L}(T_k X).
\]
Then each \( X_n \) solves the martingale problem for \( A \) with initial condition \( \sum_k p_n(k) (\mu \circ T_k^{-1}) = \mu \), where we use that \( \mu \) is shift-invariant. By Lemma 2.2 we can find a subsequence \( (X_{n_k}) \) and a solution \( X^\infty \) to the martingale problem for \( A \) such that \( X_{n(k)} \Rightarrow X^\infty \). Clearly \( X^\infty \) has initial condition \( \mathcal{L}(X^\infty(0)) = \mu \) and for any bounded continuous real function \( f \) on \( \mathcal{D}_{K,\Lambda}[0,\infty) \) we have
\[
\begin{align*}
|E[f(T_j X_{n(k)})] - E[f(X_{n(k)})]| & = \left| \sum_k p_n(k) E[f(T_j T_k X)] - \sum_i p_n(k) E[f(T_k X)] \right| \\
& = \left| \sum_k p_n(k) E[f(T_j T_k X)] - \sum_i p_n(k) E[f(T_k X)] \right| \\
& \leq \sum_i |p_n(k - j) - p_n(k)| \|f\|_\infty.
\end{align*}
\]
By (2.10) it follows that \( T_j X^\infty \) and \( X^\infty \) have the same distribution as a probability measure on \( \mathcal{D}_{K,\Lambda}[0,\infty) \), which implies that their finite-dimensional distributions agree. Hence \( X^\infty \) is shift-invariant.

\[\text{2.2 Proof of Theorem 1.2}\]
Define a normalized interaction kernel \( \bar{a} \) and a normalizing constant \( Z \) by
\[
Z := \sum_i a(i) \quad \bar{a}(i) := Z^{-1} a(i).
\]
For each \( M > 1 \) there exist \([21]\) strictly positive numbers \( (\gamma_i)_{i \in \Lambda} \) such that \( \sum_i \gamma_i < \infty \) and
\[
\sum_i \bar{a}(j - i) \gamma_i \leq M \gamma_j \quad (j \in \Lambda).
\]
Let $L^2(\gamma)$ be the Hilbert space
\[ L^2(\gamma) := \{ x \in (\mathbb{R}^d)^\Lambda : \sum_i \gamma_i |x_i|^2 < \infty \} \quad (2.15) \]
with inner product
\[ \langle x, y \rangle_\gamma := \sum_i \gamma_i x_i \cdot y_i, \quad (2.16) \]
where $\cdot$ denotes the standard inner product on $\mathbb{R}^d$. Clearly, $K^\Lambda \subset L^2(\gamma)$ and the topology on $K^\Lambda$ coincides with the topology on $L^2(\gamma)$. We write $\|x\|_\gamma := \sqrt{\langle x, x \rangle_\gamma}$ for the Hilbertian norm on $L^2(\gamma)$.

Set $\Delta(t) := X(t) - \bar{X}(t)$, where $X$ and $\bar{X}$ are solutions to (1.2), starting in $X(0) = \bar{X}(0)$ and adapted to the same set of Brownian motions. Then
\[ d\Delta^\alpha_i(t) = \sum_j \bar{a}(j - i) (\Delta^\alpha_j(t) - \Delta^\alpha_j(\bar{t})) dt \]
\[ + \sum_{\beta} (\sigma_{\alpha\beta}(X_i(t)) - \sigma_{\alpha\beta}(\bar{X}_i(t))) dB^\beta_i(t). \quad (2.17) \]
By Itô’s formula we see that
\[ E|\Delta(T)|^2_\gamma \leq \int_0^T \left\{ 2 \sum_{i,\alpha} \sum_j \gamma_i \Delta^\alpha_j(t) Z \sum_j \bar{a}(j - i) (\Delta^\alpha_j(t) - \Delta^\alpha_j(\bar{t})) \right. \]
\[ + \left. \sum_i \gamma_i \sum_{\alpha\beta} (\sigma_{\alpha\beta}(X_i(t)) - \sigma_{\alpha\beta}(\bar{X}_i(t)))^2 \right\} dt. \quad (2.18) \]
By the Lipschitz property of $\sigma$ we have
\[ \left( \sum_{\alpha\beta} (\sigma_{\alpha\beta}(x) - \sigma_{\alpha\beta}(y))^2 \right)^{\frac{1}{2}} \leq L |x - y| \quad (x, y \in K) \quad (2.19) \]
for some $L < \infty$. With $(\bar{a}\Delta(t))^\alpha := \sum_j \bar{a}(j - i) \Delta^\alpha_j(t)$ it follows that
\[ E|\Delta(T)|^2_\gamma \leq \int_0^T \left\{ 2 Z (\Delta(t), \bar{a}\Delta(t) - \Delta(t))_\gamma + L^2 (\Delta(t))^2_\gamma \right\} dt \]
\[ \leq \int_0^T \left\{ 2 Z (|\Delta(t)|^2_\gamma, |\Delta(t)|^2_\gamma) + L^2 (\Delta(t))^2_\gamma \right\} dt \]
\[ \leq \int_0^T (2 Z (M^\frac{1}{2} - 1) + L^2) E|\Delta(t)|^2_\gamma dt, \quad (2.20) \]
where we used Cauchy-Schwarz and the fact that, by Jensen’s inequality and (2.14),
\[ \|\bar{a}x\|_\gamma^2 = \sum_i \gamma_i \left| \sum_j \bar{a}(j - i) x_j \right|^2 \leq \sum_i \gamma_i \bar{a}(j - i) |x_j|^2 \]
\[ \leq \sum_j M \gamma_j |x_j|^2 = M \|x\|_\gamma^2, \quad (2.21) \]
The result now follows from Gronwall’s lemma. \hfill \blacksquare
3 Proofs of Theorem 1.3 and Lemma 1.4

3.1 Proof of Lemma 1.4

Note that, since any solution \( X \) to (1.2) solves the martingale problem for the operator \( A \) in (1.10), we have for any \( f \in C^2_{\text{fin}}(K^\Lambda) \)

\[
E[f(X(t))] - E[f(X(0))] = \int_0^t E[Af(X(s))] ds. \tag{3.1}
\]

Using the continuity of \( Af \), the continuity of the sample paths of \( X \), and bounded convergence, we see that the function \( t \mapsto E[Af(X(t))] \) is continuous. It follows that the function \( t \mapsto E[f(X(t))] \) is continuously differentiable and satisfies

\[
\frac{\partial}{\partial t} E[f(X(t))] = E[Af(X(t))]. \tag{3.2}
\]

Applying the remarks above to the function \( f(x) = x_\alpha \) and using bounded convergence to interchange an infinite sum and expectation, we see that

\[
\frac{\partial}{\partial t} E[X_\alpha(t)] = \sum_j a(j - i)(E[X_\alpha^j] - E[X_\alpha^i]). \tag{3.3}
\]

When \( X \) is shift-invariant, there clearly exist functions \( \theta : [0, \infty) \to K \) and \( C : [0, \infty) \times \Lambda \to \mathbb{R} \) such that

\[
\begin{align*}
E[X_\alpha^i(t)] &= \theta_\alpha(t) \\
\text{Cov}(X_i(t), X_j(t)) &= C_{ij}(j - i) \\
\end{align*}
\tag{3.4}
\]

Applying this to (3.3), we see that \( \frac{\partial}{\partial t} \theta(t) = 0 \) and hence

\[
E[X_i(t)] = \theta \quad (t \geq 0, \ i \in \Lambda) \tag{3.5}
\]

for some \( \theta \in K \).

Let us put \( \bar{X}_i := X_i - \theta \). Applying (3.2) to the function \( f(x) = \sum_\alpha (x_\alpha^i - \theta_\alpha)(x_\alpha^j - \theta_\alpha) \), using bounded convergence to interchange an infinite sum and expectation, we get

\[
\frac{\partial}{\partial t} \text{Cov}(X_i(t), X_j(t)) = \sum_{k,l} a(k - l)E \left[ \sum_\alpha (\bar{X}_k^\alpha(t) - \bar{X}_i^\alpha(t))(\delta_{kl}X_k^\alpha(t) + \delta_{jl}\bar{X}_i^\alpha(t)) \right] \tag{3.6}
\]

Inserting (3.4) we get

\[
\frac{\partial}{\partial t} C_{ij}(j - i) = \sum_k a(k - i)(C_{ij}(j - k) - C_{ij}(j - i)) + \sum_k a(k - j)(C_{ij}(k - i) - C_{ij}(j - i)) + 2\delta_{ij}E[tr(w)(X(t))]. \tag{3.7}
\]
Substituting $i := j - i$, $j := k - i$ and $k := j - k$ and reordering the summations, we find that

$$
\frac{\partial}{\partial t} C_t(i) = \sum_j a(j)(C_t(i - j) - C_t(j)) + \sum_k a(-k)(C_t(i - k) - C_t(i)) + 2\delta_{i0} E[tr(w)(X(t))].
$$

(3.8)

This shows that formula (1.34) holds.

### 3.2 Random walk representations

Let $B(\Lambda)$ be the Banach space of bounded real functions on $\Lambda$, equipped with the supremum norm. The operator $G$ in (1.35) is a bounded linear operator on $B(\Lambda)$. We define a Feller semigroup on $B(\Lambda)$ by

$$
P_t f := e^{tG} f,
$$

(3.9)

where $e^{tG} := \sum_{n=0}^{\infty} \frac{1}{n!} (tG)^n$. This semigroup corresponds to a continuous-time random walk $(I_t)_{t \geq 0}$ on $\Lambda$ that jumps from $i$ to $j$ with rate $a_S(j - i)$. By shift-invariance there exists a function $P : [0, \infty) \times \Lambda \rightarrow \mathbb{R}$ such that

$$
P_t(j - i) = P^t[I_t = j].
$$

(3.10)

We can consider $P_t(j - i)$ as the $(i, j)$-th element of the matrix of the operator $P_t$ in (3.9), in the following sense

$$
(P_t f)(i) = \sum_j P_t(j - i) f(j).
$$

(3.11)

**Lemma 3.1** Assume that $f, g : [0, \infty) \rightarrow B(\Lambda)$ are continuous functions, where $t \mapsto f_t(i)$ is continuously differentiable for each $i \in \Lambda$ and

$$
\frac{\partial}{\partial t} f_t(i) = \sum_j a_S(j - i)(f_t(j) - f_t(i)) + g_t(i) \quad (t > 0, \ i \in \Lambda).
$$

(3.12)

Then

$$
f_t(i) = \sum_j P_t(j - i) f_0(j) + \int_0^t \sum_j P_s(j - i) g_{t-s}(j) ds \quad (t > 0, \ i \in \Lambda).
$$

(3.13)

**Proof of Lemma 3.1:** We define derivatives and Riemann integrals of $B(\Lambda)$-valued functions as in [10], chapter 1. In that language, we would like to rewrite (3.12) as

$$
\frac{\partial}{\partial t} f_t = G f_t + g_t \quad (t \geq 0).
$$

(3.14)

However, care is needed because it is not immediately clear that the derivative $\frac{\partial}{\partial t} f_t := \lim_{\varepsilon \to 0} \varepsilon^{-1} (f_{t+\varepsilon} - f_t)$ exists in the topology on $B(\Lambda)$. To see that this is all right, we note that the function

$$
t \mapsto G f_t + g_t
$$

(3.15)
is continuous in $t$ and therefore

$$t \mapsto \int_0^t (Gf_s + g_s)ds$$

(3.16)

exists and is a continuously differentiable $B(\Lambda)$-valued function. Formula (3.12) implies that

$$f_t = \int_0^t (Gf_s + g_s)ds$$

(3.17)

and it follows that $t \mapsto f_t$ is continuously differentiable and (3.14) holds. Let $(I_t)_{t \geq 0}$ be the continuous-time random walk with kernel $a_s$. This process solves the martingale problem for $G$, and therefore

$$E^\mu[I_t] = E^\mu[I_0] - \int_0^t E^\mu[(Gf_s + G)f_{t-s}](I_s)ds$$

(3.18)

$$= E^\mu[I_0] + \int_0^t E^\mu[g_{t-s}(I_s)]ds.$$ 

This is formula (3.13).

3.3 Spatially ergodic measures

The $\sigma$-field of shift-invariant events is

$$\mathcal{S} := \{ A \in B(K^\Lambda) : T_i^{-1}(A) = A \ \forall i \in \Lambda \}.$$  

(3.19)

A probability measure $\mu$ on $K^\Lambda$ is spatially ergodic if for every $A \in \mathcal{S}$ either $\mu(A) = 1$ or $\mu(A) = 0$. We state the following standard ergodic theorem in $L^2$ without proof (see [13]).

**Lemma 3.2** For $n = 1, 2, \ldots$, let $p_n : \Lambda \to [0, \infty)$ be functions satisfying $\sum_i p_n(i) = 1$ and

$$\lim_{n \to \infty} \sum_k |p_n(i - k) - p_n(j - k)| = 0 \quad \forall i, j \in \Lambda.$$  

(3.20)

Let $X = (X_i)_{i \in \Lambda}$ be a family of $K$-valued random variables with shift-invariant ergodic law $\mathcal{L}(X)$. If $E[X_0] = \theta$, then

$$\lim_{n \to \infty} E \left[ \left| \theta - \sum_i p_n(i)X_i \right|^2 \right] = 0.$$  

(3.21)

In our case, probability distributions $p_n$ satisfying (3.20) will arise in the following way.

**Lemma 3.3** Let $P : [0, \infty) \times \Lambda \to \mathbb{R}$ be as in (3.10). Then for any $i, j \in \Lambda$:

$$\lim_{t \to \infty} \sum_k |P_t(i - k) - P_t(j - k)| = 0.$$  

(3.22)
**Proof of Lemma 3.3:** We use the Ornstein coupling [15]. To see how this works for random walks on arbitrary Abelian groups, let $\Delta \subset \Lambda$ be a set such that $a_S(k) > 0$ for each $k \in \Delta$ and such that of each $k \in \Lambda$ with $a_S(k) > 0$, either $k$ or $-k$ (but not both) is in $\Delta$. By irreducibility, we can decompose $j - i$ as

$$j - i = \sum_{k \in \Delta} n(k)k,$$

where $n(k) \in \mathbb{Z}$ and only a finite number of $n(k)$'s are non-zero. We may couple two random walks starting in points $i$ and $j$ in such a way that they always make a jump of size $k$ or $-k$ at the same time. They choose $k$ or $-k$ independently of each other, until the walk starting in $j$ has made $n(k)$ more of these jumps than the walk starting in $i$. After that, they choose either both $k$ or both $-k$. This coupling is obviously successful and Lemma 3.3 now follows easily. \[\blacksquare\]

**3.4 Proof of Theorem 1.3**

The proof consists of several steps.

**$X(\infty)$ is an invariant law.** By this we mean that there exists a shift-invariant solution $X^{\infty}$ to the martingale problem for the operator $A$ in (1.10) such that

$$\mathcal{L}(X^{\infty}(t)) = X(\infty) \quad \forall t \geq 0.$$  

(3.24)

To see this, define solutions to the martingale problem for $A$ by

$$X_n(t) := X(t_n + t),$$  

(3.25)

where $(t_n)$ is some sequence tending to infinity. By Lemma 2.2 we can find a subsequence $(X_{n(k)})$ that converges weakly to some solution $X^{\infty}$ to the martingale problem for $A$. Now

$$\mathcal{L}(X^{\infty}(t)) = \lim_{n \to \infty} \mathcal{L}(X(t_n + t)) = \mathcal{L}(X(\infty)) \quad \forall t \geq 0,$$

(3.26)

where the limit denotes weak convergence of probability measures on $K^\Lambda$. It is easy to see that $X^{\infty}$ is shift-invariant.

**Recurrent $a_S$, $P[X_i(\infty) \in \partial_\omega K \quad \forall i \in \Lambda] = 1.$** Let us write

$$\text{Cov}(X_i^{\infty}(t), X_j^{\infty}(t)) = C_i^{\infty}(j - i)$$

(3.27)

for covariances belonging to the process $X^{\infty}$ constructed above. We can apply Lemma 1.4 to this process. Lemma 3.1 now leads to the representation

$$C_i^{\infty}(i) - \sum_j P_t(j - i)C_0^{\infty}(j) = 2 \int_0^t P_s(0 - i)E[\text{tr}(w)(X_0^{\infty}(t - s))]ds.$$  

(3.28)

21
By the compactness of the state space $K$, the left-hand side of (3.28) is bounded. The right-hand side is equal to

$$2E[tr(w)(X_0(\infty))] \int_0^t P_s(0-i)ds. \quad (3.29)$$

By the recurrence of the random walk with kernel $a_S$, the integral in (3.29) diverges as $t$ tends to infinity, and therefore (3.28) can only hold if

$$E[tr(w)(X_0(\infty))] = 0. \quad (3.30)$$

This proves that $P[X_0(\infty) \in \partial_\omega K] = 1$ and by shift-invariance

$$P[X_i(\infty) \in \partial_\omega K \ \forall i \in \Lambda] = 1. \quad (3.31)$$

**Recurrent $a_S$,** $P[X_i(\infty) = X_j(\infty) \ \forall i, j \in \Lambda] = 1$. Applying Lemma 1.4 to the process $X^\infty$, we see that

$$\frac{\partial}{\partial t} C^\infty_t(i) = \sum_j a_S(j-i)(C^\infty_t(j) - C^\infty_t(i)) + 2\delta_0 E[tr(w)(X^\infty_0(t))]. \quad (3.32)$$

Here $C^\infty_t(i) = C^\infty(i)$, where we use the notation

$$\text{Cov}(X_i(\infty), X_j(\infty)) = C^\infty(j-i). \quad (3.33)$$

Note that $\frac{\partial}{\partial t} C^\infty_t(i) = 0$, while $E[tr(w)(X^\infty_0(t))] = 0$ by (3.30). Inserting this into (3.32), we get

$$\sum_j a_S(j-i)(C^\infty(j) - C^\infty(i)) = 0. \quad (3.34)$$

This means that $C^\infty$ is a bounded $a_S$-harmonic function. By the Choquet-Deny theorem (which follows easily from Lemma 3.3 – see [15], section II) it follows that $C^\infty$ is constant. We write $\check{X}_i(t) := X_i(t) - \theta$ with $\theta$ as in Lemma 1.4 and note that by Cauchy-Schwarz

$$C^\infty(j-i) = E[\check{X}_i(\infty) \cdot \check{X}_j(\infty)] \leq E[|\check{X}_i(\infty)|^2]^{\frac{1}{2}} E[|\check{X}_j(\infty)|^2]^{\frac{1}{2}} = E[|\check{X}_0(\infty)|^2] = C^\infty(0), \quad (3.35)$$

where equality holds if and only if $P[X_i(\infty) = X_j(\infty)] = 1$. This proves that

$$P[X_i(\infty) = X_j(\infty) \ \forall i, j \in \Lambda] = 1. \quad (3.36)$$

**Transient $a_S$,** $P[X_i(\infty) \in \partial_\omega K] < 1 \ \forall i \in \Lambda$. We start by noting that the ergodicity of $L(X(0))$ implies that for each $i \in \Lambda$

$$\lim_{t \to \infty} \sum_j P_t(j-i)C_0(j) = 0. \quad (3.37)$$
To see this, write $X_j(0) := X_i(0) - \theta$ as before and note that by Lemma 3.2 and 3.3

$$\lim_{t \to \infty} E \left[ \sum_j P_t(j)X_j(0) \right] = 0. \tag{3.38}$$

Here

$$E \left[ \sum_j P_t(j)X_j(0) \right]^2 = \sum_{j,k} P_t(j)P_t(k)E[X_j(0)X_k(0)]$$

$$\quad = \sum_{j,k} P_t(j)P_t(k)C_0(k-j)$$

$$\quad = \sum_{i,j} P_t(i)P_t(i+j)C_0(i)$$

$$\quad = \sum_i \left( \sum_j P_t(j)P_t(i-j) \right)C_0(i)$$

$$\quad = \sum_i P_{2t}(i)C_0(i), \tag{3.39}$$

where all infinite sums are absolutely convergent and we have used that, by the symmetry of $a_S$, $P_t(i) = P_t(-i)$. Formula (3.38) and (3.39) show that (3.37) holds for $i = 0$. Using Lemma 3.3 we can easily generalize this to arbitrary $i \in \Lambda$.

By Lemma 1.4 and Lemma 3.1 we have the representation

$$C_t(i) = \sum_j P_t(j-i)C_0(j) + 2 \int_0^t P_s(0-i)E[tr(w)(X_0(t-s))]ds. \tag{3.40}$$

Taking the limit $t \to \infty$ we get with the help of (3.37) that

$$C^\infty(i) = \lim_{t \to \infty} 2 \int_0^t P_s(0-i)E[tr(w)(X_0(t-s))]ds$$

$$\quad = 2E[tr(w)(X_0(\infty))] \int_0^\infty P_t(0-i)dt, \tag{3.41}$$

where we use the notation in (3.33). Let us assume for the moment that $E[tr(w)(X_0(\infty))] = 0$. Then $P[X_0(\infty) \in \partial w K] = 1$. On the other hand, (3.41) gives $C^\infty(0) = 0$ and hence $P[X_0(\infty) = \theta] = 1$. This contradicts our assumption that $\theta \not\in \partial w K$ and we conclude that $E[tr(w)(X_0(\infty))] > 0$. Therefore $P[X_0(\infty) \in \partial w K] < 1$ and the claim follows from shift-invariance.

**Transient $a_S$, $P[X_i(\infty) = X_j(\infty)] < 1 \ \forall i \neq j \in \Lambda$.** Let $(I_t)_{t \geq 0}$ be the random walk with kernel $a_S$. Let $\tau_i$ be the stopping time

$$\tau_i := \inf \{ t \geq 0 : I_t = i \} \quad (i \in \Lambda). \tag{3.42}$$

It is easy to see that for all $i \in \Lambda$

$$\int_0^\infty P_t(0-i)dt = P^{\tau_0 \leq \infty} \int_0^\infty P_t(0)dt. \tag{3.43}$$

23
Let us assume that for some $i \neq 0$ we have $P^i[\tau_0 < \infty] = 1$. Then by the symmetry of the random walk, also $P^0[\tau_1 < \infty] = 1$. But this implies that the random walk starting in 0 visits 0 infinitely often, which contradicts our assumption that it is transient. It follows that $P^i[\tau_j < \infty] < 1$ for all $i \neq j$. Combining (3.43) and (3.41) we can conclude that

$$C^\infty(i) < C^\infty(0) \quad \forall i \neq 0. \quad (3.44)$$

Now Cauchy-Schwarz in (3.35) implies that $P[X_i(\infty) = X_j(\infty)] < 1$ for all $i \neq j$. ■

4 Proofs of Theorem 1.5, Lemma 1.6 and Corollary 1.7

4.1 Potential theory

In this section we collect some elementary facts about $w$-harmonic functions from potential theory. We assume that for each $x \in K$, the non-interacting equation (1.19) has a unique weak solution $X^x$ with initial condition $X^x(0) = x$. We denote its last element by $X^x(\infty)$. We denote the semigroup on $B(K)$ associated with (1.19) by $(S_t)_{t \geq 0}$, we write $G$ for its full generator and we add a last element $S_\infty$ to the semigroup as in (1.41).

Lemma 4.1 For each solution $X$ to (1.19)

$$P[X(\infty) \in \partial_w K] = 1. \quad (4.1)$$

Proof of Lemma 4.1: Since $X$ is a bounded martingale, it converges. Now the lemma is just a special case of Theorem 1.3. ■

Lemma 4.2 Assume that $S_\infty(\mathcal{C}(K)) \subset \mathcal{C}(K)$. Consider sets $H, H', H'', H'''$ defined as

$$H := \{ f \in \mathcal{D}(G) : Gf = 0 \}$$
$$H' := \{ f \in \mathcal{C}(K) : S_t f = f \ \forall t \in [0, \infty] \}$$
$$H'' := \{ f \in \mathcal{C}(K) : S_\infty f = f \}$$
$$H''' := \{ S_\infty \phi : \phi \in \mathcal{C}(K) \} \quad (4.2)$$

Then $H = H' = H'' = H'''$. For each $\phi \in \mathcal{C}(K)$ there exists a unique $f \in H$ such that

$$f(x) = \phi(x) \quad (x \in \partial_w K) \quad (4.3)$$

and this $f$ is given by

$$f = S_\infty \phi. \quad (4.4)$$

Proof of Lemma 4.2: It is easy to see that $H \subset H' \subset H'' \subset H'''$. To see that $H''' \subset H$, note that $\phi(X(t)) \to \phi(X(\infty))$ almost surely, so bounded convergence implies that $E^x[\phi(X(t))] \to E^x[\phi(X(\infty))]$ for each $x \in K$. Since $|E^x[\phi(X(t))]| \leq ||\phi||_\infty < \infty$, it
follows that $S_t \phi \to S_\infty \phi$ as $t \to \infty$ in the sense of bounded pointwise convergence. Therefore

$$
(S_t S_\infty \phi)(x) = \lim_{s \to \infty} (S_{s+t} S_\infty \phi)(x) = (S_\infty \phi)(x) \quad (x \in K).
$$

(4.5)

It follows that $t^{-1}(S_t - 1)S_\infty \phi = 0$ for all $t$, so

$$
\lim_{t \to 0} t^{-1}(S_t - 1)S_\infty \phi = 0
$$

(4.6)

in the topology on $\mathcal{C}(K)$ and this proves that $H'' \subset H$.

By (4.2), $S_\infty \phi \in H$ for each $\phi \in \mathcal{C}(K)$. To see that $f := S_\infty \phi$ solves (4.3) it suffices to note that for each $x \in \partial \omega K$ the process

$$
X(t) := x
$$

(4.7)

solves (1.19). To see that $f$ is the unique $\omega$-harmonic function satisfying (4.3), suppose that $\tilde{f} := \mathcal{S}_\infty \tilde{\phi} \in H$ is another one. Then by (4.2) and by Lemma 4.1

$$
\tilde{f} = S_\infty \tilde{f} = S_\infty \phi = f.
$$

(4.8)


In the proof of Theorem 1.5 we will make use of the function $v^*$, given by

$$
v^*(x) = \text{tr}(w^*),
$$

(4.9)

where $w^*$ is the special diffusion matrix mentioned in (1.52). The following lemma collects some elementary facts about $v^*$.

**Lemma 4.3** Assume that $S_\infty(\mathcal{C}(K)) \subset \mathcal{C}(K)$. Then there exists a unique function $v^* \in \mathcal{D}(G)$ such that

$$
-\frac{1}{2}(Gv^*)(x) = \text{tr}(w)(x) \quad (x \in K)
$$

$$
v^*(x) = 0 \quad (x \in \partial \omega K).
$$

(4.10)

This function $v^*$ satisfies

$$
v^*(x) \geq 0 \quad x \in K
$$

$$
v^*(x) = 0 \iff x \in \partial \omega K,
$$

(4.11)

and is given by the formula

$$
v^*(x) = \text{Var}(X^x(\infty)).
$$

(4.12)

**Proof of Lemma 4.3:** We write $x^2$ for the function $x \mapsto |x|^2$. Then

$$
Gx^2 = 2\text{tr}(w).
$$

(4.13)

Thus, (4.10) can be rewritten as

$$
G(v^* + x^2) = 0 \quad \text{on } K
$$

$$
(v^* + x^2) = x^2 \quad \text{on } \partial \omega K.
$$

(4.14)
Lemma 4.2 shows that $v^* + x^2$ can be uniquely solved from these equations. Using the fact that $X^x$ solves the martingale problem for $G$ we see with the help of (4.10) and (4.13) that

$$E \left| X^x(\infty) - x \right|^2 = 2 \int_0^\infty E[tr(w)(X(t))]dt$$

$$= -E[\nu^*(X^x(\infty))] + \nu^*(x).$$

(4.15)

Lemma 4.1 implies that $E[\nu^*(X^x(\infty))] = 0$ and so we see that (4.12) holds. Formula (4.12) immediately implies that $\nu^*(x) \geq 0$. Finally $\nu^*(x) = 0$ implies $\text{Var}(X^x(\infty)) = 0$ so that $X^x(\infty) = x$, and by Lemma 4.1 this in turn implies that $x \in \partial_w K$.

4.2 Infinite-dimensional differentiation

We will need to extend the domain of the operator $A$ in (1.10) to include functions depending on infinitely many coordinates. In order to do this properly, we introduce the space $C^2_{\text{sum}}(K^\Lambda)$ of functions with summable continuous second derivatives. For a function $f : (\mathbb{R}^d)^\Lambda \to \mathbb{R}$ we define $\frac{\partial}{\partial x_i}f$ in the usual way. $C^2((\mathbb{R}^d)^\Lambda)$ is the class of functions for which all zeroth, first and second order derivatives are continuous functions on $(\mathbb{R}^d)^\Lambda$. $C^2(K^\Lambda)$ is the set of functions on $K^\Lambda$ that can be extended to functions in $C^2(\mathbb{R}^d)$. $C^2_{\text{sum}}(K^\Lambda)$, finally, is the space of functions in $C^2(K^\Lambda)$ for which

$$x \mapsto \left( \frac{\partial}{\partial x_i^j}f(x) \right)_{i,j=1,...,d},$$

are continuous functions from $K^\Lambda$ into the spaces $l^1(\{1,\ldots,d\} \times \Lambda)$ and $l^1(\{1,\ldots,d\}^2 \times \Lambda^2)$ of absolutely summable sequences, equipped with the $l^1$-norm.

**Lemma 4.4** For $i,j \in \Lambda$ and $\alpha,\beta = 1,\ldots,d$, let $b_{i,\alpha}$ and $a_{ij,\alpha,\beta}$ be functions in $C(K^\Lambda)$ satisfying uniform bounds

$$\|b_{i,\alpha}\|_\infty \leq M_1 \quad \forall i \in \Lambda, \quad \alpha = 1,\ldots,d$$

$$\|a_{ij,\alpha,\beta}\|_\infty \leq M_2 \quad \forall i,j \in \Lambda, \quad \alpha,\beta = 1,\ldots,d.$$  

(4.17)

Then for each $f \in C^2_{\text{sum}}(K^\Lambda)$ and for all finite $\Delta_n \subset \Lambda$ with $\Delta_n \uparrow \Lambda$, the limit

$$\lim_{n \to \infty} \left( \sum_{i \in \Delta_n, \alpha} b_{i,\alpha}(x) \frac{\partial}{\partial x_i} f(x) + \sum_{ij \in \Delta_n, \alpha,\beta} a_{ij,\alpha,\beta}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) \right)$$

(4.18)

exists in the topology on $C(K^\Lambda)$ and does not depend on the choice of the $\Delta_n$.

**Proof of Lemma 4.4:** We treat only the convergence of the first order derivatives; the argument for second order derivatives is then the same. Define operators $(A_n)$ by

$$(A_n f)(x) := \sum_{i \in \Delta_n, \alpha} b_{i,\alpha}(x) \frac{\partial}{\partial x_i} f(x).$$

(4.19)
For each \( n \leq m \) and for each \( f \in C^2_{\text{sum}}(K^\Lambda) \) we have
\[
\|A_n f - A_m f\|_\infty = \sup_{x \in K^\Lambda} \left| \sum_{i \in A_n \setminus A_m} b_{i,\alpha}(x) \frac{\partial}{\partial x_i} f(x) \right|
\leq M_1 \sup_{x \in K^\Lambda} \sum_{i \in A \setminus A_n} \left| \frac{\partial}{\partial x_i} f(x) \right|.
\] (4.20)

The functions \( g_n : K \rightarrow l^1(\{1, \ldots, d\} \times \Lambda) \) given by
\[
g_n(x) := \sum_{i,\alpha} \left| \frac{\partial}{\partial x_i} f(x) \right|
\] (4.21)
are continuous functions decreasing to zero as \( n \to \infty \), and hence by Dini’s theorem \( g_n \to 0 \) uniformly. Thus we see by (4.20) that the sequence \((A_n f)\) is a Cauchy sequence in the Banach space \( C(K^\Lambda) \).

To see that the limit does not depend on the summation order, observe that for each \( x \in K^\Lambda \)
\[
\sum_{i,\alpha} \left| b_{i,\alpha}(x) \frac{\partial}{\partial x_i} f(x) \right| \leq M_1 \sum_{i,\alpha} \left| \frac{\partial}{\partial x_i} f(x) \right| < \infty.
\] (4.22)
This means that we can write
\[
\lim_{n \to \infty} (A_n f)(x) = \sum_{i,\alpha} b_{i,\alpha}(x) \frac{\partial}{\partial x_i} f(x),
\] (4.23)
where the sums are pointwise absolutely convergent so that the result does not depend on the summation order.

For each \( i \in \Lambda \) and \( \alpha, \beta \in \{1, \ldots, d\} \) the maps
\[
\begin{align*}
x &\mapsto \sum_{j} a(j-i)(x_j^\alpha - x_i^\beta) \\
x &\mapsto w_{\alpha,\beta}(x_i)
\end{align*}
\] (4.24)
are continuous on \( K^\Lambda \) and satisfy the uniform bounds
\[
\sup_{x \in K^\Lambda} \left| \sum_{j} a(j-i)(x_j^\alpha - x_i^\beta) \right| \leq \left( \sum_{k} a(k) \right) \sup_{\alpha', \beta'} \sup_{y, z \in K} \left| y^{\alpha'} - z^{\alpha'} \right|
\] (4.25)
\[
\sup_{x \in K^\Lambda} \left| w_{\alpha,\beta}(x_i) \right| \leq \sup_{\alpha', \beta'} \| w_{\alpha',\beta'} \|_\infty.
\] (4.26)
Therefore, by Lemma 4.4, the operator
\[
(A' f)(x) := \left( \sum_{i,\alpha} \sum_{j} a(j-i)(x_j^\alpha - x_i^\beta) \frac{\partial}{\partial x_i} + \sum_{i,\alpha,\beta} w_{\alpha,\beta}(x_i) \frac{\partial^2}{\partial x_i \partial x_j} \right) f(x)
\] (4.26)
is well-defined on \( C^2_{\text{sum}}(K^\Lambda) \), where the infinite sums are convergent in \( C(K^\Lambda) \) and the result is independent of the summation order. We now show that \( A \) is a core for \( A' \).
Lemma 4.5 Let $\overline{A}$ be the closure of the operator $A$ in (1.10), and let $\mathcal{D}(\overline{A})$ be its domain. Then $\mathcal{C}^2_{\text{sum}}(K^A) \subset \mathcal{D}(\overline{A})$ and

$$(\overline{A}f)(x) = \left( \sum_{i,\alpha} \sum_{j} a(i-j)(x_j^i - x_j^i) \frac{\partial}{\partial x_j^i} + \sum_{i,\alpha,\beta} w_{\alpha\beta}(x_i) \frac{\partial^2}{\partial x_j^i \partial x_k^i} \right) f(x)$$

(4.27)

for each $f \in \mathcal{C}^2_{\text{sum}}(K^A)$.

Proof of Lemma 4.5: We have to show that for each $f \in \mathcal{C}^2_{\text{sum}}(K^A)$ there exist $f_n \in \mathcal{D}(A)$ such that $f_n \to f$ and $Af_n \to A'f$, with $A'$ as in (4.26). Fix $f \in \mathcal{C}^2_{\text{sum}}(K^A)$ and define $\pi_n$, $f_n$ as in (2.1) and (2.2), so that $f_n \to f$. It is easy to see that there exists a constant $C$ such that

$$\|Af_n(x) - A'f(x)\|_\infty \leq C \sup_{x \in K^A} \left( \sum_{i,\alpha} \left| \frac{\partial}{\partial x_j^i} f_n(x) - \frac{\partial}{\partial x_j^i} f(x) \right| 
+ \sum_{i,\alpha,\beta} \left| \frac{\partial^2}{\partial x_j^i \partial x_k^i} f_n(x) - \frac{\partial^2}{\partial x_j^i \partial x_k^i} f(x) \right| \right)$$

$$\leq C \sup_{x \in K^A} \left( \sum_{i \leq n, \alpha} \left| \frac{\partial}{\partial x_j^i} f(x) \right| \right) \pi_n(x) - \frac{\partial}{\partial x_j^i} f(x) \right| 
+ \sum_{i \leq n, \alpha} \left| \frac{\partial^2}{\partial x_j^i \partial x_k^i} f(x) \right| \pi_n(x) - \frac{\partial^2}{\partial x_j^i \partial x_k^i} f(x) \right| 
+ \sum_{i > n, \alpha} \left| \frac{\partial}{\partial x_j^i} f(x) \right| \right)$$

(4.28)

By the compactness of $K^A$, the maps

$$x \mapsto \left( \frac{\partial}{\partial x_j^i} f(x) \right)_{i=1, \ldots, d}$$

$$x \mapsto \left( \frac{\partial^2}{\partial x_j^i \partial x_k^i} f(x) \right)_{i, j \in \Lambda}$$

(4.29)

are uniformly continuous with respect to the norm on the spaces $l^1(\{1, \ldots, d\} \times \Lambda)$ and $l^1(\{1, \ldots, d\}^2 \times \Lambda^2)$. This implies that

$$\lim_{n \to \infty} \sup_{x \in K^A} \sum_{i,\alpha} \left| \frac{\partial}{\partial x_j^i} f(x) \pi_n(x) - \frac{\partial}{\partial x_j^i} f(x) \right| = 0,$$

(4.30)

and similarly for second derivatives. Finally, by Dini’s theorem (see (4.21))

$$\lim_{n \to \infty} \sup_{x \in K^A} \sum_{i > n, \alpha} \left| \frac{\partial}{\partial x_j^i} f(x) \right| = 0,$$

(4.31)

and similarly for second derivatives, so $Af_n \to A'f$.  

4.3 Proof of Lemma 1.6

The model with zero diffusion In the special case that \( w = 0 \), the system of stochastic differential equations (1.2) reduces to

\[
dX_i(t) = \sum_{j \in \Lambda} a(j - i)(X_j(t) - X_i(t))dt \quad (i \in \Lambda, \ t \geq 0).
\]

By Theorems 1.1 and 1.2 this system of equations has a unique solution. We can write down the solution of (4.32) explicitly in terms of the random walk on \( \Lambda \) that jumps from \( i \) to \( j \) with rate \( a(j - i) \). Let \( P_t(j - i) \) denote the probability that this random walk, starting in \( i \) at time 0, is in \( j \) at time \( t \). Then the unique solution of (4.32) is given by (see Lemma 3.1)

\[
X_i(t) = \sum_j P_t(j - i)X_j(0).
\]

Let \( (P_t)_{t \geq 0} \) be the semigroup on \( B(\Lambda) \) associated with the random walk with kernel \( a \) (see section 3.2). Let us denote by \( (R_t)_{t \geq 0} \) the Feller semigroup on \( C(\Lambda) \) associated with the process in (4.32):

\[
(R_t f)(x) := E[f(X^x(t))] = f(P_t x).
\]

Applying Lemma 4.5 to the case \( w = 0 \), we see that the generator of \( (R_t)_{t \geq 0} \) is an extension of the operator

\[
(Bf)(x) := \sum_{i \in \Lambda} \sum_j a(j - i)(x_j - x_i) \frac{\partial}{\partial x_i} f(x)
\]

with domain \( D(B) := C^2_{\text{sum}}(\Lambda) \).

Evolution of harmonic functions We now set out to prove Lemma 1.6. We start with the case \( f \in C^2(\Lambda) \cap H \). Fix \( i \in \Lambda \), and let \( h \in C^2_{\text{fin}}(\Lambda) \) be given by

\[
h(x) := f(x_i).
\]

In the language above, we want to show that

\[
E[h(X(t))] = E[(R_th)(X(0))].
\]

It is not hard to see that \( R_th \in C^2_{\text{sum}}(\Lambda) \) for each \( t \geq 0 \), where by (4.34)

\[
\frac{\partial}{\partial x_i} (R_th)(x) = P_t(k - i) \left( \sum_j P_t(j - i) x_j \right)
\]

and

\[
\frac{\partial^2}{\partial x_i^2} (R_th)(x) = P_t(k - i)P_t(l - i) \left( \sum_j P_t(j - i) x_j \right).
\]

General theory (see [10], chapter 1) now tells us that \( t \mapsto R_th \) is continuously differentiable in \( C(\Lambda) \) and

\[
\frac{\partial}{\partial t} R_th = BR_th,
\]
with $B$ the operator in (4.35). By Lemma 4.5, $X$ solves the martingale problem for the operator
\begin{equation}
A' := B + C,
\end{equation}
where
\begin{equation}
(Cf)(x) := \sum_{\alpha,\beta} w_{\alpha,\beta}(x_1) \frac{\partial^2}{\partial x_1^\alpha \partial x_1^\beta} f(x)
\end{equation}
and $\mathcal{D}(A') = C^2_{\text{sum}}(K^\Lambda)$. It follows that
\begin{equation}
E[(R_0h)(X(T))] - E[(R_T h)(X(0))] = E \int_0^T (B + C + \frac{\partial}{\partial t})(R_{T-t}h)(X(t))dt
\end{equation}
\begin{equation}
= E \int_0^T (C R_{T-t}h)(X(t))dt.
\end{equation}
By (4.38) we have, for any $x \in K^\Lambda$
\begin{equation}
(C R_{T-t}h)(x) = \sum_i P_{T-t}(j-i)^2 \sum_{\alpha,\beta} w_{\alpha,\beta}(x_1) \left( \frac{\partial^2}{\partial x_1^\alpha \partial x_1^\beta} f \right) \left( \sum_j P_t(j-i)x_j \right).
\end{equation}
Using Lemma 4.2 it is not hard to see that the stability of the boundary distribution against a linear drift is equivalent to formula (1.46). The semigroup $(T_{\theta,t})_{\theta \geq 0}$ maps differentiable functions into differentiable functions, and hence
\begin{equation}
T_{\theta,t}(C^2(K) \cap H) \subset C^2(K) \cap H \quad \forall \theta \in K, \ t \geq 0.
\end{equation}
This means that $G T_{\theta,t} f = 0$ for all $f \in C^2(K) \cap H$ and $\theta \in K$, $t \geq 0$, which says that for any $x \in K$
\begin{equation}
\sum_{\alpha,\beta} w_{\alpha,\beta}(x_1) \frac{\partial^2}{\partial x_1^\alpha \partial x_1^\beta} (f(\theta + (x - \theta)e^{-t}))
\end{equation}
\begin{equation}
e^{-2t} \sum_{\alpha,\beta} w_{\alpha,\beta}(x_1) \left( \frac{\partial^2}{\partial x_1^\alpha \partial x_1^\beta} f \right)(e^{-t}x + (1 - e^{-t})\theta) = 0
\end{equation}
\begin{equation}
\forall f \in C^2(K) \cap H, \ \theta \in K, \ t \geq 0.
\end{equation}
For the $x$ here we insert the $x_i$ in (4.43) and we fit $\theta$ and $t$ such that $e^{-t} = P_1(0)$ and $(1 - e^{-t})\theta = \sum_{j \neq i} P_t(j-i)x_j$. Inserting this into (4.43) we see that each term in the sum over $i$ there is zero, and therefore (4.42) gives
\begin{equation}
E[f(X_i(T))] = E[f \left( \sum_j P_t(j-i)X_j(0) \right)].
\end{equation}
To generalize this to arbitrary $f \in H$ it suffices to note that the set of functions $f \in B(K)$ for which (4.46) holds is bp-closed. 

30
4.4  Proof of Theorem 1.5

Comparison argument  The function \( x \mapsto tr(w)(x) \) is continuous, takes only non-negative values, and satisfies

\[
tr(w)(x) = 0 \Leftrightarrow x \in \partial_w K. \tag{4.47}
\]

The same is true for the function \( x \mapsto v^*(x) \) (see Lemma 4.3) and therefore for each \( \varepsilon > 0 \) we can find a \( \lambda > 0 \) such that

\[
tr(w)(x) \geq \lambda (v^*(x) - \varepsilon) \quad (x \in K). \tag{4.48}
\]

When we insert this inequality into formula (1.34) in Lemma 1.4 we see that for all \( i \in I, \ t \geq 0 \)

\[
\frac{\partial}{\partial t} C_t(i) \geq \sum_j a_S(j - i)(C_t(j) - C_t(i)) + 2\lambda \delta_{i0}(E[v^*(X_0(t))] - \varepsilon). \tag{4.49}
\]

We apply Lemma 4.3 to see that the function

\[
x \mapsto v^*(x) + |x - \theta|^2
\]

is \( w \)-harmonic. Lemma 1.6 therefore tells us that for all \( t \geq 0 \)

\[
E[v^*(X_0(t)) + \text{Var}(X_0(t)) = E\left[v^* \left( \sum_j P_t(j)X_j(0) \right) \right] + \text{Var}\left( \sum_j P_t(j)X_j(0) \right). \tag{4.51}
\]

By Lemma 3.3, Lemma 3.2 and the spatial ergodicity of \( \mathcal{L}(X(0)) \) this implies that

\[
\lim_{t \to \infty} E[v^*(X_0(t))] + C_t(0) = v^*(\theta). \tag{4.52}
\]

Combining this with (4.49) we see there exists a \( T \) such that for all \( t \geq T \)

\[
\frac{\partial}{\partial t} C_t(i) \geq \sum_j a_S(j - i)(C_t(j) - C_t(i)) + 2\lambda \delta_{i0}(v^*(\theta) - C_t(0) - 2\varepsilon). \tag{4.53}
\]

Random walk representation  Let us define

\[
D_t(i) := v^*(\theta) - C_t(i) - 2\varepsilon \quad (i \in I, \ t \geq 0). \tag{4.54}
\]

Then (4.53) can be rewritten as

\[
\frac{\partial}{\partial t} D_t(i) \leq \sum_j a_S(j - i)(D_t(j) - D_t(i)) - 2\lambda \delta_{i0} D_t(0) \quad (t \geq T). \tag{4.55}
\]

We note that since \( t \mapsto C_t \) is continuously differentiable in \( B(\Lambda) \), so is \( t \mapsto D_t \). Arguing as in the proof of Lemma 3.1, we can represent solutions of the differential inequality (4.55) in terms of a contracting semigroup \( (P^\Lambda_t)_{t \geq 0} \) on \( B(\Lambda) \), with generator

\[
(Gf)(i) := \sum_j a_S(j - i)(f(j) - f(i)) - 2\lambda \delta_{i0} f(0) \quad (f \in B(\Lambda)). \tag{4.56}
\]
This semigroup is related to a random walk on $\Lambda$ that jumps from $i$ to $j$ with rate $a_\Lambda(j)$ and that is killed at the origin with rate $2\Lambda$. When $P_t^\Lambda(j-i)$ denotes the probability that this random walk, starting from a point $i$, is in $j$ at time $t$, then

$$(P_t^\Lambda f)(i) = \sum_j P_t^\Lambda(j-i)f(j) \quad (f \in B(\Lambda)), \quad (4.57)$$

and for solutions of (4.55) we have the representation

$$D_{T+t}(i) \leq \sum_j P_t^\Lambda(j-i)D_T(j) \quad (t \geq 0). \quad (4.58)$$

**Convergence of the covariance function** If $a_\Lambda$ is recurrent, then the random walk is killed with probability one. This means that for each $i \in \Lambda$

$$\lim_{t \to \infty} \sum_j P_t^\Lambda(j-i) = 0. \quad (4.59)$$

Combining this with (4.58) and using the boundedness of $K$ we see that for each $i \in \Lambda$ there exists a $T'$ such that for all $t \geq T'$

$$C_t(i) \geq v^*(\theta) - 3\varepsilon. \quad (4.60)$$

We have thus shown that for every $i \in \Lambda$

$$\liminf_{t \to \infty} C_t(i) = v^*(\theta). \quad (4.61)$$

On the other hand, with the help of formula (4.52) it is easy to see that

$$\limsup_{t \to \infty} C_t(0) = v^*(\theta). \quad (4.62)$$

By Cauchy-Schwarz we have $C_t(i) \leq C_t(0)$ for all $i \in \Lambda$ (compare (3.35)), and hence

$$\lim_{t \to \infty} C_t(i) = v^*(\theta) \quad \forall i \in \Lambda. \quad (4.63)$$

**Convergence of $X_0(t)$** Let $(S_t)_{t \geq 0}$ be the semigroup in (1.40). Pick any function $\phi \in C(K)$. By Lemma 4.2

$$\phi(x) - (S_\infty \phi)(x) = 0 \quad (x \in \partial_\omega K). \quad (4.64)$$

Formulas (4.52) and (4.63) imply that

$$\lim_{t \to \infty} E[v^*(X_0(t))] = 0. \quad (4.65)$$

Since $v^*$ is continuous, non-negative, and zero only at $\partial_\omega K$ (see Lemma 4.3) formulas (4.64) and (4.65) imply that

$$\lim_{t \to \infty} \left( E[\phi(X_0(t))] - E[(S_\infty \phi)(X_0(t))] \right) = 0. \quad (4.66)$$
Now $S_{\infty} f \in H$ (Lemma 4.2) and therefore Lemmas 1.6, 3.2 and 3.3 imply that
\[ \lim_{t \to \infty} E[(S_{\infty}\phi)(X_0(t))] = (S_{\infty}\phi)(\theta) = \int_K \Gamma_\phi(dx)\phi(x). \] (4.67)
Thus we see that
\[ X_0(t) \Rightarrow X_0(\infty) \quad \text{as } t \to \infty, \] (4.68)
where the law of $X_0(\infty)$ is given by
\[ \mathcal{L}(X_0(\infty)) = \Gamma_\theta. \] (4.69)

**Convergence of $X(t)$** Formula (4.61) and Cauchy-Schwarz imply that for all $i, j \in \Lambda$
\[ \lim_{t \to \infty} E[|X_i(t) - X_j(t)|^2] = 0. \] (4.70)
Combining this with (4.68) we easily see that for each finite $\Delta \subset \Lambda$ the collection $(X_i(t))_{i \in \Delta}$ converges weakly to a limit $(X_i(\infty))_{i \in \Delta}$. By the fact that continuous functions depending on finitely many coordinates only are dense in $C(K^\Lambda)$ (see the proof of Lemma 2.1) this implies weak convergence of $X(t)$. □

### 4.5 Proof of Corollary 1.7

Under the condition $w = \lambda w^*$, most inequalities in the proof of Theorem 1.5 can be replaced by equalities. In fact, under the weaker condition (recall that $w^* = tr(w^*)$)
\[ tr(w) = \lambda w^*, \] (4.71)
we have equality in (4.48) with $\varepsilon = 0$. Since we are working with the initial condition $X_i(0) = \theta$ for all $i \in \Lambda$, formula (4.52) strengthens to
\[ E[w^*(X_0(t))] + C_t(0) = w^*(\theta) \quad (t \geq 0). \] (4.72)
Formula (4.58) with $T = 0 = \varepsilon$ and an equality sign reads
\[ (w^*(\theta) - C_t(i)) = \sum_j P^\lambda_t(j - i)(w^*(\theta) - C_0(j)), \] (4.73)
where $C_0(j) = 0$ for all $j$. In this way we find that
\[ C_t(i) = w^*(\theta) \left( 1 - \sum_j P^\lambda_t(j - i) \right). \] (4.74)
Here $1 - \sum_j P^\lambda_t(j - i)$ is the same as the probability $K^\lambda(i)$ appearing in (1.54). One can derive formula (1.54) in a similar way as formula (4.74). For that, one needs to replace the covariance function $C_t(i)$ by a covariance matrix function
\[ C_t(j - i)_{\alpha\beta} := E[(X^\alpha_t(t) - \theta^\alpha)(X^\beta_j(t) - \theta^\beta)]. \] (4.75)
Generalizing Lemma 4.3 one then finds that, for each $\alpha, \beta$, the function

$$x \mapsto w_{\alpha,\beta}^*(x)$$

is the unique function in $D(G)$ solving

$$-\frac{1}{2}Gw_{\alpha,\beta}^*(x) = w_{\alpha,\beta}(x) \quad (x \in K)$$

$$w_{\alpha,\beta}^*(x) = 0 \quad (x \in \partial_w K).$$

The rest of the proof is now in complete analogy with the proof of formula (4.74).

5 Proofs of the examples

5.1 Proof of Example 1.8

Weak uniqueness for (1.19) The uniqueness proof in section 4 of [17], although stated there only for diffusion matrices of a special form, carries over to our situation. For this, the main fact one has to check is the following.

**Lemma 5.1** For $\alpha = 1, \ldots, p$, let

$$F_{\alpha} := \{(x^1, \ldots, x^{p-1}) \in S^{p-1} : x^\alpha = 0\}$$

$$F_{\alpha} := \{(x^1, \ldots, x^{p-1}) \in S^{p-1} : \sum_{\beta} x^\beta = 1\}$$

be the $\alpha$-face of the $(p-1)$-dimensional simplex $S^{p-1}$. Then for any solution $X$ to (1.19) with $X(0) \in F_{\alpha}$

$$P[X(t) \in F_{\alpha} \quad \forall t \geq 0] = 1.$$  

**Proof of Lemma 5.1:** immediate by the martingale property of solutions to (1.19).

In order to show weak uniqueness for (1.19) we prove strong uniqueness for the special case that $\sigma$ is the unique positive symmetric root of $w$ (recall (1.7)). By (1.55), this $\sigma$ is Lipschitz continuous on the interior of $\mathbb{S}^{p-1}$ and therefore a standard argument gives uniqueness of solutions to (1.19) up to the first hitting of a face $F_{\alpha}$.

By Lemma 5.1, the process stays in this face after hitting it. Each face is isomorphic to $S^{p-2}$ and therefore strong uniqueness can be proved by induction. For details we refer to [17].

(a), (b) and (c) By (1.55), the effective boundary of $K$ consists of the extremal points of $S^{p-1}$:

$$\partial_w K = \{e_1, \ldots, e_p\},$$

where $e_1 = (1,0,\ldots,0)$, $e_{p-1} = (0,\ldots,0,1)$ and $e_p = (1,\ldots,1)$. It follows that for any $f \in \mathcal{C}(K)$

$$(S_\infty f)(x) = \sum_{\alpha=1}^p P[X^\sigma(\infty) = e_\alpha]f(e_\alpha),$$

where the probabilities $P[X^\sigma(\infty) = e_p]$ follow from the martingale property of solutions to (1.19). The rest of the assertions are now trivial.
5.2 Proof of Example 1.9

Uniqueness of solutions to (1.19) is proved in the same way as in Example 1.8, where this time one needs to check that any solution \( X \) starting in \( x \in \partial K \) is constant with probability one. By the convexity of \( K \) we can without lack of generality assume that \( x = 0 \) and \( y^t \geq 0 \) for all \( y \in K \). Then by the martingale property of solutions to (1.19) we have

\[
P[X^t(t) = 0 \ \forall t \geq 0] = 1
\]  

(5.5)

and this implies that almost surely \( X(t) \in \partial K \) for all \( t \geq 0 \) and hence

\[
E[X(t)]^2 = \int_0^t E[2g(X(s))]ds = 0 \ \forall t \geq 0.
\]  

(5.6)

Thus we see that almost surely \( X(t) = x \) for all \( t \geq 0 \).

To see that \( S_\infty(C(K)) \subset C(K) \) and that the class of harmonic functions is given by formula (1.60), we can use [14], Proposition 4.2.7 and Theorems 4.2.12 and 4.2.19, where by (1.24) the harmonic functions of the process in (1.19) are the same as the harmonic functions for Brownian motion. The same references show that (1.62) has a unique solution. It follows from (1.60) that \( T_{\alpha,\beta}(H) \subset H \) and this implies that the boundary distribution is stable against a linear drift. The other assertions in Example 1.9 are now readily checked.

References


