# DUALITY AND CONES OF MARKOV PROCESSES AND THEIR SEMIGROUPS

M. Möhle[1] Mathematisches Institut, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany

March 22, 2012

**Abstract**

We show that the *cone duality* essentially coincides with Liggett's definition of duality of Markov processes. Several examples, mainly motivated from mathematical population genetics, of dual Markov processes and their corresponding convex cones are provided, including Fleming–Viot measure valued processes and their dual coalescents with simultaneous multiple collisions of ancestral lineages.

Keywords: coalescent; cone; duality; Fleming–Viot process; stochastic monotone Markov chain

AMS 2000 Mathematics Subject Classification: Primary 60K35 Secondary 92D10; 92D25

## 1 Introduction

The concept of duality is a common tool in the analysis of Markov processes. Duality is for example used in the physics literature on interacting particle systems [26, 31, 32, 41, 42] and in the literature on mathematical population genetics [2, 3, 5, 7, 17, 20, 21, 33, 34]. In fact, interacting particle systems and mathematical population genetics are closely related ([18, 19, 27]), and duality relations appear in many other situations ([12, 13, 14, 15, 16]). In Section 1.5 of [30] the *cone dual* is considered. It is based on a stochastic monotonicity property [4, 11, 39, 40] for Markov processes. In this paper the *cone duality* is slightly extended and it is shown that it coincides with the usual definition of duality given in Definition 2.1 below. Several examples of dual Markov processes are presented and their cones are characterized.

## 2 Duality and cones

Assume that $X = (X_t)_{t\in T}$ and $Y = (Y_t)_{t\in T}$ are two homogeneous-time Markov processes with state spaces $(E_1, \mathcal{F}_1)$ and $(E_2, \mathcal{F}_2)$ respectively. Typical time sets $T$ we are thinking of are finite sets $T = \{0, \ldots, n\}$ for some $n \in \mathbb{N}$, countable sets $T = \mathbb{N}_0 := \{0, 1, 2, \ldots\}$ or as well continuous time sets such as the unit interval $T = [0, 1]$ or $T = [0, \infty)$. Let $B(E)$ denote the set of all real-valued bounded measurable functions on $E := E_1 \times E_2$. We recall the following definition of duality of Markov processes (Liggett [31]).

**Definition 2.1 (duality)** *The process $X$ is said to be dual to $Y$ with respect to $H \in B(E)$ if*

$$\mathrm{E}^x H(X_t, y) \;=\; \mathrm{E}^y H(x, Y_t) \qquad (1)$$

[1] E-mail address: martin.moehle@uni-tuebingen.de

*for all* $(x,y) \in E$ *and* $t \in T$*, where* $\mathrm{E}^x$ *denotes the expectation given that the process* $X$ *starts in* $X_0 = x$ *and* $\mathrm{E}^y$ *denotes the expectation given that the process* $Y$ *starts in* $Y_0 = y$*. The function* $H$ *is called a duality function.*

Dual processes occur in many applications, usually when considering some phenomena forwards and backwards in time. For typical examples in the physics literature on interacting particle systems we refer to Liggett [31, 32]. Other important examples occur in the context of mathematical population genetics [2, 3, 5, 7, 17, 20, 21, 33, 34]. Some of these examples are presented and discussed in Section 3.

**Definition 2.2 (duality space)** *For two given Markov processes* $X = (X_t)_{t\in T}$ *and* $Y = (Y_t)_{t\in T}$ *the set of all duality functions*

$$U \ := \ U(X,Y) \ := \ \{H \in B(E) : X \text{ is dual to } Y \text{ with respect to } H\} \tag{2}$$

*is called the duality space of* $X$ *and* $Y$*.*

Clearly, $U$ is a subspace of the Banach space $B(E)$, equipped with the supremum norm $\|H\| := \sup\{|H(x,y)| : (x,y) \in E\}$, $H \in B(E)$. In order to verify that $U$ is a closed subset of $B(E)$, let $(H_n)_{n\in\mathbb{N}} \subset U$ be a convergent sequence with limit $H \in B(E)$. The sequence $(H_n)_{n\in\mathbb{N}}$ is convergent and, hence, bounded. Thus there exists a constant $C > 0$ such that $\|H_n\| \le C$ for all $n \in \mathbb{N}$ or, equivalently, $H_n(x,y) \le C$ for all $(x,y) \in E$ and all $n \in \mathbb{N}$. Moreover, the uniform convergence ($\|H_n - H\| \to 0$) implies the pointwise convergence, i.e. $H_n(x,y) \to H(x,y)$ as $n \to \infty$ for all $(x,y) \in E$. For $(x,y) \in E$ and $t \in T$, it follows by dominated convergence that

$$\begin{aligned} \mathrm{E}^x H(X_t,y) &= \mathrm{E}^x \lim_{n\to\infty} H_n(X_t,y) \ = \ \lim_{n\to\infty} \mathrm{E}^x H_n(X_t,y) \\ &= \lim_{n\to\infty} \mathrm{E}^y H_n(x,Y_t) \ = \ \mathrm{E}^y \lim_{n\to\infty} H_n(x,Y_t) \ = \ \mathrm{E}^y H(x,Y_t), \end{aligned}$$

i.e. $H \in U$. Thus, $U$ is a closed subset of the Banach space $B(E)$, and, therefore, $U$ is itself a Banach space. Depending on the structure of the Markov processes $X$ and $Y$, the duality space $U$ might be quite small or quite large. Typical problems are to determine the dimension of $U$ or to find a basis of $U$. For some results for particular Markov processes arising in mathematical population genetics we refer the reader to [33]. We will not further analyze the duality space $U$ here, and we will also not exploit the structure of the duality space in our further considerations.

Let $R = \{R_t\}_{t\in T}$ and $S = \{S_t\}_{t\in T}$ denote the semigroups of $X$ and $Y$ respectively, i.e., $R_t f(x) := \mathrm{E}^x f(X_t)$ and $S_t g(y) := \mathrm{E}^y g(Y_t)$ for all functions $f$ on $E_1$ and $g$ on $E_2$ such that the expectations exist. From (1) it follows that $X$ is dual to $Y$ with respect to $H \in B(E)$ if and only if

$$(R_t H(.,y))(x) \ = \ (S_t H(x,.))(y) \tag{3}$$

for all $x \in E_1$, $y \in E_2$ and $t \in T$. Note that (3) is a reformulation of the property that $X$ is dual to $Y$ with respect to $H$ in terms of the semigroups $R$ and $S$. It is therefore natural to call any semigroup $R = \{R_t\}_{t\in T}$ on $B(E_1)$ dual to another semigroup $S = \{S_t\}_{t\in T}$ on $B(E_2)$ with respect to a given duality function $H \in B(E)$, if (3) holds for all $x \in E_1$, $y \in E_2$ and $t \in T$. Duality can hence be defined for semigroups without the notion of Markov processes.

From a functional analytic view the definition (3) of duality in terms of semigroups seems to be more natural. However, as we will see soon, it is quite useful to work with the Markov processes $X$ and $Y$ instead of the semigroups $R$ and $S$ alone, and to exploit the relations $R_t f(x) = \mathrm{E}^x f(X_t)$ and $S_t g(y) = \mathrm{E}^y(g(Y_t))$ between the Markov processes $X$ and $Y$ and their corresponding semigroups $R$ and $S$.

**Definition 2.3 (cone)** *A set $C \subseteq B(E_1)$ is called a cone of the Markov process $X = (X_t)_{t\in T}$ if $R_t C \subseteq C$ for all $t \in T$, where $\{R_t\}_{t\in T}$ denotes the semigroup of $X$.*

**Remarks.** 1. The name *cone* appears in [30]. This name is reasonable because Definition 2.3 reminds of the standard algebraic definition of a cone $C$ as a nonempty subset of a vector space $V$ over an ordered field $F$ with the property that $\lambda C \subseteq C$ for all $\lambda \in F$ with $\lambda \geq 0$. On the other hand the name *cone* is somewhat misleading. For example, a cone in the sense of Definition 2.3 does not necessarily contain the neutral element, whereas $0 \in C$ for any cone $C$ in the standard algebraic sense. For more information on (variants of) such cones we refer the reader to [1].
2. Being a subset of $B(E_1)$, $C$ is equipped with the norm $\|f\| := \sup\{|f(x)| : x \in E_1\}$, $f \in C \subseteq B(E_1)$. For example, we can speak of the diameter $d(C) := \sup\{\|f_1 - f_2\| : f_1, f_2 \in C\} \in [0,\infty]$ of $C$, the distance $d_0(C) := \inf\{\|f\| : f \in C\} \in [0,\infty]$ of $C$ from the origin or more general the distance $d_h(C) := \inf\{\|f - h\| : f \in C\}$ of $C$ from some given $h \in B(E_1)$.

For a given function $H \in B(E)$ we are interested in the two sets $C_1 \subseteq B(E_1)$ and $C_2 \subseteq B(E_2)$ defined via

$$C_1 \; := \; C_1(H) \; := \; \{f : E_1 \to \mathbb{R} \, : \, f(x) = \textstyle\int_{E_2} H(x,y)\, Q_2(dy) \text{ for some } Q_2 \in \mathcal{M}_1(E_2)\} \quad (4)$$

and

$$C_2 \; := \; C_2(H) \; := \; \{g : E_2 \to \mathbb{R} \, : \, g(y) = \textstyle\int_{E_1} H(x,y)\, Q_1(dx) \text{ for some } Q_1 \in \mathcal{M}_1(E_1)\}, \quad (5)$$

where $\mathcal{M}_1(E_k)$ denotes the set of probability measures on $(E_k, \mathcal{F}_k)$, $k \in \{1,2\}$. Note that, for arbitrary but fixed $y \in E_2$, the set $C_1$ contains the function $H(.,y) : E_1 \to \mathbb{R}$. This property follows simply by choosing $Q_2 = \delta_y$ (Dirac measure at $y$) in (4). In the same manner, for every $x \in E_1$, the set $C_2$ contains the function $H(x,.) : E_2 \to \mathbb{R}$. The set $C_1$ (and also $C_2$) is convex, i.e., if $f_1, f_2 \in C_1$, then $qf_1 + (1-q)f_2 \in C_1$ for all $q \in [0,1]$. Moreover, $d_0(C_1), d_0(C_2) \leq \|H\| := \sup\{|H(x,y)| : (x,y) \in E\} < \infty$ and $d(C_1), d(C_2) \leq 2\|H\| < \infty$, since $\|f\|, \|g\| \leq \|H\|$ for $f \in C_1$, $g \in C_2$. If $H$ is non-negative, then $0 \leq f(x) \leq \|H\|$ and $0 \leq g(y) \leq \|H\|$ for all $f \in C_1$, $g \in C_2$, $x \in E_1$ and $y \in E_2$, leading to the better bound $d(C_1), d(C_2) \leq \|H\|$.

**Proposition 2.4** *If $X$ is dual to $Y$ with respect to $H$, then $C_1$ is a cone of $X$ and $C_2$ is a cone of $Y$.*

**Proof.** We show that $C_1$ is a cone of $X$. The proof that $C_2$ is a cone of $Y$ works analogously. Fix $t \in T$, $f \in C_1$, and $x \in E_1$. By the definition of $C_1$ there exists $Q_2 \in \mathcal{M}_1(E_2)$ such that $f(x') = \int_{E_2} H(x',y)\, Q_2(dy)$ for all $x' \in E_1$. Thus,

$$(R_t f)(x) \;=\; \mathrm{E}^x f(X_t) \;=\; \int_{E_1} f(x') P^x_{X_t}(dx') \;=\; \int_{E_1} \int_{E_2} H(x',y)\, Q_2(dy) P^x_{X_t}(dx').$$

An application of Fubini's theorem leads to

$$(R_t f)(x) \;=\; \int_{E_2}\int_{E_1} H(x',y) P^x_{X_t}(dx')\, Q_2(dy) \;=\; \int_{E_2} \mathrm{E}^x H(X_t,y)\, Q_2(dy).$$

As $X$ is dual to $Y$ with respect to $H$, we have $\mathrm{E}^x H(X_t,y) = \mathrm{E}^y H(x,Y_t)$ and, therefore, we arrive at

$$(R_t f)(x) \;=\; \int_{E_2} \mathrm{E}^y H(x,Y_t)\, Q_2(dy) \;=\; \int_{E_2}\int_{E_2} H(x,y') P^y_{Y_t}(dy')\, Q_2(dy).$$

Now apply again Fubini's theorem to conclude that

$$(R_t f)(x) \;=\; \int_{E_2} H(x,y')\, Q_2'(dy'), \tag{6}$$

where $Q_2' \in \mathcal{M}_1(E_2)$ is defined via $Q_2'(B) := \int_{E_2} P^y_{Y_t}(B)\, Q_2(dy)$ for $B \in \mathcal{F}_2$. Eq. (6) implies that $R_t f \in C_1$. Thus, it is shown that $R_t C_1 \subseteq C_1$ for all $t \in T$, i.e., $C_1$ is a cone of $X$. □

Proposition 2.4 shows that duality in the sense of Definition 2.1 implies cone duality, i.e. that $C_1$ is a cone of $X$ and that $C_2$ is a cone of $Y$. We show the following kind of converse of Proposition 2.4.

**Proposition 2.5** *Let $X = (X_t)_{t\in T}$ be a Markov process with state space $(E_1,\mathcal{F}_1)$. Assume that there exist a space $(E_2,\mathcal{F}_2)$, a subset $C_1 \subseteq B(E_1)$, and a function $H \in B(E_1 \times E_2)$ such that the following properties hold.*

*(i) For each $y \in E_2$ the function $H(.,y) : E_1 \to \mathbb{R}$ belongs to $C_1$.*

*(ii) The set $C_1$ is a cone of $X$.*

*(iii) The set $C_1$ has a unique integral representation over $E_2$ with respect to $H$, i.e., for each function $f \in C_1$, there exists a unique probability measure $Q_f$ on $(E_2,\mathcal{F}_2)$ with*

$$f(x) \;=\; \int_{E_2} H(x,y)\, Q_f(dy), \qquad x \in E_1.$$

*Then, there exists a Markov process $Y = (Y_t)_{t\in T}$ with state space $(E_2,\mathcal{F}_2)$ such that $X$ is dual to $Y$ with respect to $H$. The process $Y$ is unique in distribution. If $\{R_t\}_{t\in T}$ denotes the semigroup of $X$, then $Y$ has transition kernel $P(Y_t \in B \,|\, Y_0 = y) = Q_{R_t H(.,y)}(B)$, $B \in \mathcal{F}_2$, $y \in E_2$.*

**Proof.** Let $\{R_t\}_{t\in T}$ denote the semigroup of $X$. Fix $t \in T$ and $y \in E_2$. From $H(.,y) \in C_1$ and $R_t C_1 \subseteq C_1$ it follows that $R_t H(.,y) \in C_1$. By (iii) there exists a unique probability measure $Q_{R_t H(.,y)}$ on $(E_2,\mathcal{F}_2)$ such that

$$(R_t H(.,y))(x) \;=\; \int_{E_2} H(x,y') Q_{R_t H(.,y)}(dy'), \qquad x \in E_1. \tag{7}$$

Let $Y=(Y_t)_{t\in T}$ be a Markov process with state space $(E_2,\mathcal{F}_2)$ and transition kernel $P(Y_t\in B\,|\,Y_0=y):=P^y_{Y_t}(B):=Q_{R_tH(.,y)}(B)$ for $y\in E_2$ and $B\in\mathcal{F}_2$. Then, for $x\in E_1$ and $y\in E_2$,

$$\begin{aligned}\mathrm{E}^y H(x,Y_t) &= \int_{E_2} H(x,y')P^y_{Y_t}(dy') \;=\; \int_{E_2} H(x,y')\,Q_{R_tH(.,y)}(dy')\\ &\overset{(7)}{=} (R_tH(.,y))(x) \;=\; \int_{E_1} H(x',y)P^x_{X_t}(dx') \;=\; \mathrm{E}^x H(X_t,y).\end{aligned}$$

Thus, $X$ is dual to $Y$ with respect to $H$. The uniqueness of $Y$ in distribution follows from the fact that the integral representation in (iii) is unique. □

**Remarks.** 1. In some cases the state space $E_2$ itself is a subset of $C_1$. In this situation each $y\in E_2$ is a bounded measurable function $y:E_1\to\mathbb{R}$, and a typical duality function is

$$H(x,y) \;:=\; y(x), \tag{8}$$

the evaluation of $y$ at the point $x$. For this particular choice of $H$ one arrives at the *cone duality* of [30], which was the starting point for this article. In this context the set $E_2$ is sometimes called the set of extremals. The duality in general allows for other duality functions $H$ than that defined in (8). Moreover, in general, $E_2$ is not necessarily a subset of $C_1$ (not even of $B(E_1)$). If both, $E_1$ and $E_2$, are subsets of $[0,\infty)$, then typical other duality functions (see, for example, [2]) are $H(x,y)=x^y$ (moment duality) or $H(x,y)=\exp(-\lambda xy)$ for some $\lambda>0$ (Laplace duality).
2. Propositions 2.4 and 2.5 show that duality of Markov processes in the sense of Liggett and cone duality are closely related to each other. It is a matter of taste, once the duality function $H$ is guessed, wether it is easier to establish duality in the sense of Liggett or cone duality. Cone duality essentially deals with the question when a semigroup $\{T_t\}_{t\in T}$ preserves a convex set $C$ in some Banach space $B$. Questions of this form have been addressed in the functional analytic literature, see for example Brezis and Pazy [8] for the case when $B$ is a Hilbert space. Ouhabaz [36, 37] studies similar problems. However, these results seem to be rather abstract and not applicable directly. Our results show that such preserving properties are (at least in some cases) related to Liggett's duality of Markov processes. The following examples show that establishing duality in the sense of Liggett is often quite straightforward, usually by viewing some appropriate stochastic system forwards and backwards in time. As a consequence of Proposition 2.4, cone duality can be established in that way which is the main contribution of this article.

## 3 Examples

We start with an example involving stochastic monotone Markov processes well known from the literature (Sigmund [40], Klebaner, Rösler and Sagitov [30]).

**Stochastic monotone Markov chains**
Let $Y=(Y_n)_{n\in\mathbb{N}_0}$ be a Markov chain with state space $E_2:=\mathbb{N}_0$ and transition matrix $P=(p_{ij})_{i,j\in\mathbb{N}_0}$. Suppose that $Y$ is stochastic monotone, i.e., $P(Y_{n+1}\ge j\,|\,Y_n=i)$ is monotone increasing in $i\in\mathbb{N}_0$ for every $j\in\mathbb{N}_0$. For further information on stochastically monotone Markov chains we refer to Sigmund [40] and to the related papers of Asmussen and Sigman [4] and Sigman and Ryan [39].

We furthermore assume that $\lim_{i\to\infty} P(Y_{n+1} \geq j \,|\, Y_n = i) = 1$ for all $j \in \mathbb{N}_0$. As in [30], let $X = (X_n)_{n\in\mathbb{N}_0}$ be a Markov chain with transition probabilities $\pi_{ij} := P(X_{n+1} = j \,|\, X_n = i)$, $i, j \in \mathbb{N}_0$, defined via $\pi_{i0} := P(Y_{n+1} \geq i \,|\, Y_n = 0)$, $i \in \mathbb{N}_0$, and

$$\pi_{ij} \;:=\; P(Y_{n+1} \geq i \,|\, Y_n = j) - P(Y_{n+1} \geq i \,|\, Y_n = j-1), \quad i \in \mathbb{N}_0, j \in \mathbb{N}.$$

Note that, for $i, k \in \mathbb{N}_0$,

$$\begin{aligned} P(X_{n+1} \leq k \,|\, X_n = i) \;&=\; \sum_{j=0}^{k} \pi_{ij} \;=\; \pi_{i0} + \sum_{j=1}^{k} \pi_{ij} \\ &=\; P(Y_{n+1} \geq i \,|\, Y_n = 0) + \sum_{j=1}^{k} P(Y_{n+1} \geq i \,|\, Y_n = j) - P(Y_{n+1} \geq i \,|\, Y_n = j-1) \\ &=\; P(Y_{n+1} \geq i \,|\, Y_n = k). \end{aligned} \tag{9}$$

In particular, $\sum_{j=0}^{\infty} \pi_{ij} = \lim_{k\to\infty} \sum_{j=0}^{k} \pi_{ij} = \lim_{k\to\infty} P(Y_{n+1} \geq i \,|\, Y_n = k) = 1$. If we define the matrix $H = (h_{ij})_{i,j\in\mathbb{N}_0}$ via $h_{ij} := 1$ for $i \leq j$ and $h_{ij} := 0$ otherwise, then (9) takes the form $\sum_{j\in\mathbb{N}_0} \pi_{ij} h_{jk} = \sum_{j\in\mathbb{N}_0} h_{ij} p_{kj}$. Thus, in matrix notation we have $\Pi H = HP'$, where $P'$ denotes the transpose of $P$. By induction on $n$ it follows that $\Pi^n H = H(P')^n$ for all $n \in \mathbb{N}_0$. Thus, $X$ is dual to $Y$ with respect to $H : \mathbb{N}_0^2 \to \mathbb{R}$, $H(i,j) := h_{ij}$. By its definition (see (4)), the cone $C_1$ of $X$ consists of all functions $f : \mathbb{N}_0 \to \mathbb{R}$ of the form

$$f(i) \;=\; \sum_{j\in\mathbb{N}_0} h_{ij} Q_2(\{j\}) \;=\; \sum_{j\geq i} Q_2(\{j\}) \;=\; Q_2(\{i, i+1, \ldots\})$$

for some probability measure $Q_2$ on $\mathbb{N}_0$. Thus, $C_1$ coincides with the set of all non-negative, non-increasing functions $f$ on $\mathbb{N}_0$ satisfying $f(0) = 1$. In the same manner it follows that the cone $C_2$ of $Y$ is the set of all non-negative non-decreasing functions $g$ on $\mathbb{N}_0$ satisfying $\lim_{j\to\infty} g(j) = 1$. It is easily seen that $C_1$ and $C_2$ have both distance $d_0(C_1) = d_0(C_2) = 1$ from the origin and both diameter $d(C_1) = d(C_2) = 1$. Note that the cone $C_2$ essentially occurs in Section 1.5 on p. 1042 in [30]. A concrete example which fits into this context is given on p. 763 in [33]. In this example, $X$ is a random walk on $\mathbb{N}_0$ with absorbtion at 0 and $Y$ a random walk with reflection at 0.

**Brownian motion with reflection and Brownian motion with absorption**
There is an obvious continuous analog (with $T = E_1 = E_2 = [0, \infty)$) of the just mentioned random walk example, namely Brownian motion with reflection at 0 and Brownian motion with absorption at 0, which is well described on pp. 84–85 of [31] and which is probably one of the oldest examples for dual processes. Since $H(x, y) = 1$ for $x \leq y$ and $H(x, y) = 0$ otherwise, it is easily seen that the cone $C_1$ is the set of all non-negative, non-increasing, left-continuous functions $f : [0, \infty) \to \mathbb{R}$ satisfying $f(0) = 1$ and that the cone $C_2$ coincides with the set of all distribution functions $g : [0, \infty) \to \mathbb{R}$. Again, $C_1$ and $C_2$ both have distance $d_0(C_1) = d_0(C_2) = 1$ from the origin and both diameter $d(C_1) = d(C_2) = 1$.

The following examples, mainly motivated from mathematical population genetics, provide some further insight into the typical structure of cones. The examples in particular nicely illustrate that cones usually consist of functions having properties which are preserved under convex combination.

**Forward and backward process of Cannings models**
In population genetics, Cannings models [9, 10] describe the evolution of a population of a fixed size of $N \in \mathbb{N}$ individuals evolving in non-overlapping generations. The forward process $X$ counts the number of descendants forwards in time, whereas the backward process $Y$ counts the number of ancestors backwards in time. Note that both processes have state space $S := \{0, \ldots, N\}$. It is well-known [23, 24, 25, 33] that $X$ is dual to $Y$, for example with respect to the duality function $H : S^2 \to \mathbb{R}$ defined via $H(i,j) := \binom{i}{j}/\binom{N}{j}$, $i,j \in S$. The function $H$, considered as a $(N+1)\times(N+1)$-matrix, is non-singular and its inverse $H^{-1}$ has entries $H^{-1}(i,j) = (-1)^{i-j}\binom{i}{j}\binom{N}{i}$, $i,j \in S$. By (4), each $f \in C_1$ has the representation

$$f(i) \ = \ \sum_{j\in S} H(i,j)\, Q_2(\{j\}) \ = \ \sum_{j=0}^{i} \frac{\binom{i}{j}}{\binom{N}{j}} Q_2(\{j\}), \qquad i \in S,$$

for some probability distribution $Q_2$ on $S$. In particular $f(N) = Q_2(S) = 1$. Together with the inversion formula

$$\binom{N}{i}\sum_{j=0}^{i}(-1)^{i-j}\binom{i}{j} f(j) \ = \ \sum_{j\in S} H^{-1}(i,j)\, f(j) \ = \ Q_2(\{i\}) \ \geq \ 0, \qquad i \in S,$$

it follows that the cone $C_1$ of $X$ consists of all functions $f : S \to \mathbb{R}$ satisfying $f(N) = 1$ and being absolutely monotone, i.e. $\sum_{j=0}^{i}(-1)^{i-j}\binom{i}{j}f(j) \geq 0$ for all $i \in S$. In particular, $0 \leq f(0) \leq f(1) \leq \cdots \leq f(N-1) \leq f(N) = 1$ for all $f \in C_1$. Note that $\|f\| = \sup_{i\in S}|f(i)| = f(N) = 1$ for all $f \in C_1$. In particular, $C_1$ has distance $d_0(C_1) = 1$ from the origin. For $f_k \in C_1$, $k \in \{1,2\}$, we have $f_k(i) \in [0,1]$ for $i \in S$ and, therefore, $\|f_1 - f_2\| = \sup_{i\in S}|f_1(i) - f_2(i)| \leq 1$. The maximal value $\|f_1 - f_2\| = 1$ is for example obtained for $f_1, f_2 \in C_1$ defined via $f_1(i) := 1$ and $f_2(i) := i$, $i \in S$. Thus, $C_1$ has diameter $d(C_1) = \sup\{\|f_1 - f_2\| : f_1, f_2 \in C_1\} = 1$.
Similarly it follows from (5) that the cone $C_2$ of $Y$ consists of all functions $g : S \to \mathbb{R}$ satisfying $g(0) = 1$ and $\sum_{j=i}^{N}\binom{N-i}{j-i}(-1)^{j-i}g(j) \geq 0$ for all $i \in S$, so the map $j \mapsto g(N-j)$ is absolutely monotone. In particular, $1 = g(0) \geq g(1) \geq \cdots \geq g(N-1) \geq g(N) \geq 0$ for all $g \in C_2$. The cone $C_2$ has distance $d_0(C_2) = 1$ from the origin and diameter $d(C_2) = \sup\{\|g_1 - g_2\| : g_1, g_2 \in C_2\} = 1$.

**Wright–Fisher diffusion and block counting process of the Kingman coalescent**
The Wright–Fisher diffusion $X = (X_t)_{t\geq 0}$ (without mutation and selection) is a Markov process with state space $E_1 := [0,1]$ and generator $Af(x) := \frac{1}{2}x(1-x)f''(x)$, $f \in C^2([0,1])$, $x \in [0,1]$.
Let $Y = (Y_t)_{t\geq 0}$ be a decreasing Markov chain with state space $E_2 := \mathbb{N}_0$ and infinitesimal rates $g_{ij} := \lim_{t\searrow 0} t^{-1}(P(Y_t = j \,|\, Y_0 = i) - \delta_{ij})$, $i,j \in \mathbb{N}_0$, given by $g_{i,i-1} := -g_{ii} := i(i-1)/2$ and $g_{ij} := 0$ otherwise. Note that $Y$ is the block counting process of the Kingman coalescent [28, 29].
It is well known (see, for example, [34]) that $X$ is dual to $Y$ with respect to the duality function $H : [0,1] \times \mathbb{N}_0 \to \mathbb{R}$, $H(x,n) := x^n$. By (4), the cone $C_1$ of $X$ consists of all series

$f:[0,1]\to\mathbb{R}$ of the form $f(x)=\sum_{n=0}^\infty a_n x^n$, with non-negative coefficients $a_0,a_1,\ldots$ such that $\sum_{n=0}^\infty a_n=1$. Note that $\|f\|=\sup_{x\in[0,1]}|f(x)|=f(1)=1$ for all $f\in C_1$. In particular, $C_1$ has distance $d_0(C_1)=1$ from the origin. For $f_i\in C_1$, $i\in\{1,2\}$, we have $f_i(x)\in[0,1]$ for $x\in[0,1]$ and, therefore, $\|f_1-f_2\|=\sup_{x\in[0,1]}|f_1(x)-f_2(x)|\le 1$. The maximal value $\|f_1-f_2\|=1$ is for example obtained for $f_1,f_2\in C_1$ defined via $f_1(x):=1$ and $f_2(x):=x$, $x\in[0,1]$. Thus, $C_1$ has diameter $d(C_1)=\sup\{\|f_1-f_2\|:f_1,f_2\in C_1\}=1$.
The cone $C_2$ of $Y$ consists of all functions $g:\mathbb{N}_0\to\mathbb{R}$ of the form $g(n)=\int_{[0,1]}x^nQ(dx)$ for some probability measure $Q$ on $[0,1]$. The cone $C_2$ therefore can be identified with the set of all Hausdorff moment sequences $(g(n))_{n\in\mathbb{N}_0}$. From $1=g(0)\ge g(1)\ge g(2)\ge\cdots$ for $g\in C_2$ it follows that $\|g\|=g(0)=1$ for $g\in C_2$. In particular, $C_2$ has distance $d_0(C_2)=1$ from the origin. The same argument as above for $C_1$ shows that $\|g_1-g_2\|\le 1$ for $g_1,g_2\in C_2$. The maximal value $\|g_1-g_2\|=1$ is for example obtained for $g_1,g_2\in C_2$ defined via $g_1(n):=\int_{[0,1]}x^n\,\delta_1(dx)=1$ and $g_2(n):=\int_{[0,1]}x^n\,\delta_0(dx)=\delta_{0n}$, $n\in\mathbb{N}_0$. Thus, $C_2$ has diameter $d(C_2)=\sup\{\|g_1-g_2\|:g_1,g_2\in C_2\}=1$. This example can be extended as follows.

**Fleming–Viot process and Kingman coalescent**
The classical Fleming–Viot process [22] (without mutation) is a Markov process $F=(F_t)_{t\ge0}$ with state space $E_1=\mathcal{M}_1(E)$, the set of all probability measures on a given compact Polish space $E$. In biological applications the space $E$ represents the set of possible types. Typical examples of type spaces $E$ are finite sets or the unit interval $[0,1]$. The generator of $F$, denoted by $L$, acts on test functions $G_f$ of the form

$$G_f(\mu)\ :=\ \int_{E^n}f(x)\,\mu^n(dx),\qquad f\in B(E^n),\mu\in\mathcal{M}_1(E),$$

where $\mu^n$ denotes the $n$-fold product measure of $\mu$, via

$$LG_f(\mu)\ =\ \sum_{1\le i<j\le n}\int_{E^n}(f(x(i,j))-f(x))\,\mu^n(dx),$$

where $x(i,j)\in E^n$ is obtained from $x=(x_1,\ldots,x_n)\in E^n$ by replacing the entry $x_j$ by $x_i$.
The Kingman coalescent [28, 29] is a Markov process $\Pi=(\Pi_t)_{t\ge0}$ with state space $E_2=\mathcal{E}$, the set of equivalence relations on $\mathbb{N}$. Fix $n\in\mathbb{N}$. Let $\varrho_n:\mathcal{E}\to\mathcal{E}_n$ denote the natural restriction to the set $\mathcal{E}_n$ of all equivalence relations on $\{1,\ldots,n\}$. In the restricted coalescent process $(\varrho_nR_t)_{t\ge0}$ during each transition exactly two equivalence classes (blocks) merge together at rate 1 to form a single block.
For an arbitrary but fixed function $h\in B(E^n)$, define $H_n:\mathcal{M}_1(E)\times\mathcal{E}_n\to\mathbb{R}$ via

$$H_n(\mu,\xi)\ :=\ \int_{E^n}h(x[\xi])\,\mu^n(dx),\tag{10}$$

where, for any equivalence relation $\xi=\{B_1,\ldots,B_k\}$ with equivalence classes (blocks) $B_1,\ldots,B_k$ and $x=(x_1,\ldots,x_n)\in E^n$, the element $x[\xi]\in E^n$ has by definition entries $(x[\xi])_i:=x_{\min B_j}$ if $i\in B_j$, $i\in\{1,\ldots,n\}$. It is well known that for each $n\in\mathbb{N}$ the Fleming–Viot process $F$ is dual to the restricted coalescent process $(\varrho_n\Pi_t)_{t\ge0}$ with respect to $H_n$, i.e.,

$$\mathrm{E}^\mu H_n(F_t,\xi)\ =\ \mathrm{E}^\xi H_n(\mu,\varrho_n\Pi_t)\tag{11}$$

for $t \geq 0$, $\mu \in \mathcal{M}_1(E)$ and $\xi \in \mathcal{E}_n$. An elementary proof of (11) is provided in the following example in an even more general setting. The cone $C_1 = C_1(H_n)$ of the Fleming–Viot process $F$ consists of all functions $f : \mathcal{M}_1(E) \to \mathbb{R}$ of the form

$$f(\mu) \;=\; \sum_{\xi \in \mathcal{E}_n} H_n(\mu, \xi)\, Q_2(\{\xi\}) \;=\; \int_{E^n} \sum_{\xi \in \mathcal{E}_n} Q_2(\{\xi\})\, h(x[\xi])\, \mu^n(dx)$$

for some $Q_2 \in \mathcal{M}_1(\mathcal{E}_n)$. The cone $C_2 = C_2(H_n)$ of the restricted coalescent process $(\varrho_n \Pi_t)_{t \geq 0}$ contains all functions $g : \mathcal{E}_n \to \mathbb{R}$ of the form

$$g(\xi) \;=\; \int_{\mathcal{M}_1(E)} H_n(\mu, \xi)\, Q_1(d\mu) \;=\; \int_{\mathcal{M}_1(E)} \int_{E^n} h(x[\xi])\, \mu^n(dx)\, Q_1(d\mu)$$

for some $Q_1 \in \mathcal{M}_1(\mathcal{M}_1(E))$. Clearly, $d_0(C_1), d_0(C_2) \leq \|H_n\| \leq \|h\| < \infty$ and $d(C_1), d(C_2) \leq 2\|H_n\| \leq 2\|h\| < \infty$. It does not seem to be straightforward to provide simpler representations of the cones $C_1$ and $C_2$ for arbitrary $h \in B(E^n)$.

**$\Xi$-Fleming–Viot process and $\Xi$-coalescent**

Coalescent processes with simultaneous multiple collisions have been introduced by Möhle and Sagitov [35] and Schweinsberg [38]. Schweinsberg characterizes these processes via a finite measure $\Xi$ on the infinite simplex $\Delta := \{x = (x_1, x_2, \ldots) : x_1 \geq x_2 \geq \cdots \geq 0, \sum_{i=1}^{\infty} x_i \leq 1\}$. These processes are therefore also called $\Xi$-coalescents. The Kingman coalescent corresponds to the case when $\Xi$ is the Dirac measure at $(0, 0, \ldots) \in \Delta$. On the other hand, there is also a natural extension of the classical Fleming–Viot process, called the $\Xi$-Fleming–Viot process. The duality relation (11) still holds (see, for example, [6, Lemma 5.1]) for the $\Xi$-Fleming–Viot process $F$ and its dual $\Xi$-coalescent $\Pi$. Since the function $H_n$ in (11) does not depend on the measure $\Xi$, the cones $C_1$ and $C_2$ do not depend on $\Xi$ and are hence identical to those already discussed in the previous example for the classical Fleming–Viot process and the dual Kingman coalescent. We now present an elementary proof of the duality relation (11). For $\xi \in \mathcal{E}_n$ define $G_\xi : \mathcal{M}_1(E) \to \mathbb{R}$ via $G_\xi(\mu) := H_n(\mu, \xi)$ with $H_n$ as defined in (10). The proof of the duality relation (11) relies on the fact that the generator $L$ of the $\Xi$-Fleming–Viot process $F$ acts on $G_\xi$ via

$$LG_\xi(\mu) \;=\; \sum_{\eta \in \mathcal{E}_n} q_{\xi\eta} G_\eta(\mu), \tag{12}$$

where the $q_{\xi\eta}$, $\xi, \eta \in \mathcal{E}_n$, are the infinitesimal rates of the restricted coalescent process $(\varrho_n \Pi_t)_{t \geq 0}$. The semigroup operator $R_t$ of the Fleming–Viot process $F$, i.e., $R_t G(\mu) := \mathrm{E}^\mu G(F_t)$, satisfies $\frac{d}{dt} R_t = R_t L$. Thus,

$$\frac{d}{dt} R_t G_\xi(\mu) = R_t L G_\xi(\mu) \;=\; R_t\Big(\sum_{\eta \in \mathcal{E}_n} q_{\xi\eta} G_\eta\Big)(\mu) \;=\; \sum_{\eta \in \mathcal{E}_n} q_{\xi\eta} R_t G_\eta(\mu).$$

As $R_t G_\xi(\mu) = \mathrm{E}^\mu G_\xi(F_t) = \mathrm{E}^\mu H_n(F_t, \xi)$, this equation is equivalent to

$$\frac{d}{dt} \mathrm{E}^\mu H_n(F_t, \xi) \;=\; \sum_{\eta \in \mathcal{E}_n} q_{\xi\eta} \mathrm{E}^\mu H_n(F_t, \eta).$$

On the other hand,

$$
\begin{aligned}
\frac{d}{dt}\mathrm{E}^{\xi} H_n(\mu, \varrho_n \Pi_t) &= \sum_{\tau \in \mathcal{E}_n} H_n(\mu, \tau) \frac{d}{dt} P(\varrho_n \Pi_t = \tau \mid \varrho_n \Pi_0 = \xi) \\
&= \sum_{\tau \in \mathcal{E}_n} H_n(\mu, \tau) \sum_{\eta \in \mathcal{E}_n} q_{\xi\eta} P(\varrho_n \Pi_t = \tau \mid \varrho_n \Pi_0 = \eta) \\
&= \sum_{\eta \in \mathcal{E}_n} q_{\xi\eta} \sum_{\tau \in \mathcal{E}_n} H_n(\mu, \tau) P(\varrho_n \Pi_t = \tau \mid \varrho_n \Pi_0 = \eta) \\
&= \sum_{\eta \in \mathcal{E}_n} q_{\xi\eta} \mathrm{E}^{\eta} H_n(\mu, \varrho_n \Pi_t).
\end{aligned}
$$

Thus, the functions $t \mapsto \mathrm{E}^{\mu} H_n(F_t, \xi)$ and $t \mapsto \mathrm{E}^{\xi} H_n(\mu, \varrho_n \Pi_t)$ satisfy the same differential equation. As these functions have for $t = 0$ the same initial value $H_n(\mu, \xi)$, these functions have to be equal, and (11) is established.

**Acknowledgement.** Many thanks go to Uwe Rösler for pointing the author to Section 1.5 of [30], which was the starting point for this work.

# References

[1] Aliprantis, C. D. and Tourky, R. (2007) *Cones and Duality.* Graduate Studies in Mathematics **84**, AMS, Providence.

[2] Alkemper, R. and Hutzenthaler, M. (2007) Graphical representation of some duality relations in stochastic population models. *Electron. Commun. Probab.* **12**, 206–220.

[3] Athreya, S. R. and Swart, J. M. (2005) Branching-coalescing particle systems. *Probab. Theory Relat. Fields* **131**, 376–414.

[4] Asmussen, S. and Sigman, K. (1996) Monotone stochastic recursions and their duals. *Prob. Eng. Inf. Sci.* **10**, 1–20.

[5] Bertoin, J. and Le Gall, J.-F. (2003) Stochastic flows associated to coalescent processes. *Probab. Theory Relat. Fields* **126**, 261–288.

[6] Birkner, M., Blath, J., Möhle, M., Steinrücken, M. and Tams, J. (2008) A modified lookdown construction for the Ξ-Fleming-Viot process with mutation and populations with recurrent bottlenecks. *Latin American J. of Prob. and Math. Stat. (ALEA)* **6**, 25–61.

[7] Blythe, R. A. (2007) The propagation of a cultural or biological trait by neutral genetic drift in a subdivided population. *Theor. Popul. Biol.* **71**, 454–472.

[8] Brezis, H. and Pazy, A. (1970) Semigroups of nonlinear contractions on convex sets. *J. Func. Anal.* **6**, 237–281.

[9] Cannings, C. (1974) The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models. *Adv. Appl. Probab.* **6**, 260–290.

[10] Cannings, C. (1975) The latent roots of certain Markov chains arising in genetics: a new approach, II. Further haploid models. *Adv. Appl. Probab.* **7**, 264–282.

[11] Clifford, P. and Sudbury, A. (1985) A sample path proof of the duality for stochastically monotone Markov processes. *Ann. Probab.* **13**, 558–565.

[12] Cox, J. T. and Rösler, U. (1984) A duality relation for entrance an exit laws for Markov processes. *Stoch. Process. Appl.* **16**, 141–156.

[13] Dette, H., Fill, J. A., Pitman, J. and Studden, W. J. (1997) Wall and Siegmund duality relations for birth and death chains with reflecting barrier. *J. Theor. Probab.* **10**, 349–374.

[14] Diaconis, P. and Fill, J. A. (1990) Strong stationary times via a new form of duality. *Ann. Probab.* **18**, 1483–1522.

[15] Dong, R., Goldschmidt, C. and Martin, J. B. (2006) Coagulation-fragmentation duality, Poisson-Dirichlet distributions and random recursive trees. *Ann. Appl. Probab.* **16**, 1733–1750.

[16] Donnelly, P. (1985) Dual processes and an invariance result for exchangeable models in population genetics. *J. Math. Biol.* **23**, 103–118.

[17] Donnelly, P. (1986) Dual processes in population genetics. In: *Stochastic Spatial Processes* (Lecture Notes in Mathematics 1212), Springer, Berlin, pp. 95–105.

[18] Donnelly, P. and Kurtz, T. G. (1996) A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.* **24**, 698–742.

[19] Donnelly, P. and Kurtz, T. G. (1999) Particle representations for measure-valued population models. *Ann. Probab.* **27**, 166–205.

[20] Ethier, S. N. and Krone, S. M. (1995) Comparing Fleming-Viot and Dawson-Watanabe processes. *Stoch. Process. Appl.* **60**, 401–421.

[21] Ethier, S. N. and Kurtz, T. G. (1986) *Markov Processes. Characterization and Convergence*. Wiley, New York.

[22] Fleming, W. H. and Viot, M. (1979) Some measure-valued Markov processes in population genetics theory. *Indiana University Mathematics Journal* **28**, 817–843.

[23] Gladstien, K. (1976) Loss of alleles in a haploid population with varying environment. *Theor. Popul. Biol.* **10**, 383–394.

[24] Gladstien, K. (1977) Haploid populations subject to varying environment: the characteristic values and the rate of loss of alleles. *SIAM J. Appl. Math.* **32**, 778–783.

[25] Gladstien, K. (1978) The characteristic values and vectors for a class of stochastic matrices arising in genetics. *SIAM J. Appl. Math.* **34**, 630–642.

[26] Griffeath, D. (1993) Frank Spitzer's pioneering work on interacting particle systems. *Ann. Probab.* **21**, 608–621.

[27] Hobson, T. and Tribe, R. (2005) On the duality between coalescing Brownian particles and the heat equation driven by Fisher-Wright noise. *Electron. Commun. Probab.* **10**, 136–145.

[28] Kingman, J. F. C. (1982) On the genealogy of large populations. *J. Appl. Probab.* **19A**, 27–43.

[29] Kingman, J. F. C. (1982) The coalescent. *Stoch. Process. Appl.* **13**, 235–248.

[30] Klebaner, F. C., Rösler, U., and Sagitov, S. (2007) Transformations of Galton-Watson processes and linear fractional reproduction. *Adv. Appl. Probab.* **39**, 1036–1053.

[31] Liggett, T. M. (1985) *Interacting Particle Systems*. Springer, Berlin.

[32] Liggett, T. M. (1999) *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer, Berlin.

[33] Möhle, M. (1999) The concept of duality and applications to Markov processes arising in neutral population genetics models. *Bernoulli* **5**, 761–777.

[34] Möhle, M. (2001) Forward and backward diffusion approximations for haploid exchangeable population models, *Stoch. Process. Appl.* **95**, 133–149.

[35] Möhle, M. and Sagitov, S. (2001) A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29**, 1547–1562.

[36] Ouhabaz, E. (1996) Invariance of closed convex sets and domination criteria for semigroups. *Potential Analysis* **5**, 611–625.

[37] Ouhabaz, E. (1999) $L^p$ contraction semigroups for vector valued functions. *Positivity* **3**, 83–93.

[38] Schweinsberg, J. (2000) Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5**, 1–50.

[39] Sigman, K. and Ryan, R. (2000) Continuous-time monotone stochastic recursions and duality. *Adv. Appl. Probab.* **32**, 426–445.

[40] Sigmund, D. (1976) The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probab.* **4**, 914–924.

[41] Spitzer, F. (1970) Interaction of Markov processes. *Adv. in Math.* **5**, 246–290.

[42] Sudbury, A. and Lloyd, P. (1995) Quantum operators in classical probability theory: II. The concept of duality in interacting particle systems. *Ann. Probab.* **23**, 1816–1830.