

This is the peer reviewed version of the following article:

Dualities in population genetics: A fresh look with new dualities / Carinci, Gioia; Giardina', Cristian; Giberti, Claudio; Frank, Redig. - In: STOCHASTIC PROCESSES AND THEIR APPLICATIONS. - ISSN 0304-4149. - STAMPA. - 125:(2015), pp. 941-969. [10.1016/j.spa.2014.10.009]

Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

25/04/2024 07:55

(Article begins on next page)

Dualities in population genetics: a fresh look with new dualities.

Gioia Carinci^(a), Cristian Giardinà^(a),
Claudio Giberti^(b), Frank Redig^(c).

^(a) Department of Mathematics, University of Modena and Reggio Emilia
via G. Campi 213/b, 41125 Modena, Italy

^(b) Department of Sciences and Methods for Engineering,
University of Modena and Reggio Emilia
via Giovanni Amendola 2, 42122 Reggio Emilia, Italy

^(c) Delft Institute of Applied Mathematics, Technische Universiteit Delft
Mekelweg 4, 2628 CD Delft, The Netherlands

October 19, 2014

Abstract

We apply our general method of duality, introduced in [15], to models of population dynamics. The classical dualities between forward and ancestral processes can be viewed as a change of representation in the classical creation and annihilation operators, both for diffusions dual to coalescents of Kingman's type, as well as for models with finite population size.

Next, using $SU(1, 1)$ raising and lowering operators, we find new dualities between the Wright-Fisher diffusion with d types and the Moran model, both in presence and absence of mutations. These new dualities relates two forward evolutions. From our general scheme we also identify self-duality of the Moran model.

1 Introduction

Duality is one of the most important techniques in interacting particle systems [26], models of population dynamics [18, 27, 21], mutually catalytic branching [29] and general Markov process theory [10]. See [24] for a recent review paper containing an extensive list of references, going even back to the work of Lévy.

In all the interacting particle systems models (e.g. symmetric exclusion, voter model, contact process, etc.) about which we know fine details, such as complete ergodic theorems or explicit formulas for time-dependent correlation functions, a non-trivial duality or self-duality relation plays a crucial role. This fact also becomes more and more apparent in recent exact formulas for the transition probabilities of the asymmetric exclusion process [36], [3], where often the starting point is a duality of the type first revealed in this context by Schütz in [32]. In [33] the notion of “stochastic integrability” is coined, and related to duality. It is therefore important to gain deeper understanding of “what is behind dualities”, i.e., why some processes admit nice dual processes and others not, and where the duality functions come from. This effort is not so much a quest of creating a general abstract framework of duality. It is rather a quest to create a workable constructive approach towards duality, and using this creating both new dualities in known contexts as well as new Markov processes with nice duality properties.

In the works [15, 16] duality between two stochastic processes, in the context of interacting particle systems and non-equilibrium statistical mechanics, has been related to a change of representation of an underlying Lie algebra. More precisely, if the generator of a Markov process is built from lowering and raising operators (in physics language creation and annihilation operators) associated to a Lie algebra, then different representations of these operators give rise to processes related to each other by duality. The intertwiner between the different representations is exactly the duality function. Furthermore, self-dualities [17] can be found using symmetries related to the underlying Lie algebra (see also [31], [34]).

The fact that generators of Markov processes can be built from raising and lowering operators is a quite natural assumption. In interacting particle systems, the dynamics consists of removing particles at certain places and putting them at other places. If the rates of these transitions are appropriately chosen, then the operators of which the effect is to remove or to add a particle (with appropriate coefficients), together with their commutators, generate a Lie algebra. For diffusion processes the generator is built from a combination of multiplication operators and (partial) derivatives. For specific choices, these correspond to differential operator representations of a Lie algebra. If this Lie algebra also possesses a discrete representation, then this can lead to a duality between a diffusion process and a process of jump

type, such as the well-known duality between the Wright-Fisher diffusion and the Kingman's coalescent.

It is the aim of this paper to show that this scheme of finding dual processes via a change of representation can be applied in the context of mathematical population genetics. First, we give a fresh look at the classical dualities between processes of Wright-Fisher type and their dual coalescents. These dualities correspond to a change of representation in the creation and annihilation operators generating the Heisenberg algebra. For population models in the diffusion limit (infinite population size limit) the duality comes from the standard representation of the Heisenberg algebra in terms of the multiplicative and derivative operators $(x, d/dx)$, and another discrete representation, known as the Doi-Peliti representation. The intertwiner is in this case simply the function $D(x, n) = x^n$. In the case of finite population size, dualities arise from going from a finite-dimensional representation (finite dimensional creation and annihilation operators satisfying the canonical commutation relations) to the Doi-Peliti representation. The intertwiner is exactly the hypergeometric polynomial found e.g. in [21], [20], and gives duality between the Moran model with finite population size and the Kingman's coalescent.

Next we use the $SU(1, 1)$ algebra to find previously unrevealed dualities between the discrete Moran model and the Wright-Fisher diffusion, as well as self-duality of the discrete Moran model.

These are in fact applications of the previously found dualities between the Brownian energy process and the symmetric inclusion process, as well as the self-duality of the symmetric inclusion process, which we have studied in another context in [15], [16], [17]. Put into the context of population dynamics, these dualities give new results for the multi-type Moran model, as well as the multi-type Wright-Fisher model.

The rest of our paper is organized as follows. In section 2 we give the general setting and view on duality. Though many elements of this formalism are already present in previous papers, we find it useful to put these together here in a unifying, more transparent and widely applicable framework. In section 3 we discuss dualities in the context of the Heisenberg algebra. This leads to dualities between diffusions and discrete processes, as well as between different diffusion processes, and finally between different discrete processes. In section 4 we show how this gives the dualities between forward (in time) population processes and their ancestral dual coalescents. In section 5 we apply the $SU(1, 1)$ algebra techniques in the context of population dynamics, finding new dualities, this time between two forward (in time) processes: duality between the Wright-Fisher diffusion and the Moran model, and duality of the Moran model with itself. We give three concrete computations using these new dualities as an illustration.

2 Abstract Setting

2.1 Functions and operators.

Let $\Omega, \widehat{\Omega}$ be metric spaces. We denote $\mathcal{F}(\Omega), \mathcal{F}(\widehat{\Omega}), \mathcal{F}(\Omega \times \widehat{\Omega})$ a space of real-valued functions from Ω (resp. $\widehat{\Omega}, \Omega \times \widehat{\Omega}$). Typical examples to have in mind are $\mathcal{C}(\Omega), \mathcal{C}_c(\Omega), \mathcal{C}_0(\Omega)$ the sets of continuous real-valued functions on Ω , resp. continuous real-valued functions with compact support on Ω , and continuous real-valued functions on Ω going to zero at infinity. In what follows, this choice of function space is not so stringent, and can be replaced if necessary by other function spaces such as L^p spaces or Sobolev spaces. For a function $\psi : \Omega \times \widehat{\Omega} \rightarrow \mathbb{R}$ and linear operators $K : \mathcal{D}(K) \subset \mathcal{F}(\Omega) \rightarrow \mathcal{F}(\Omega), \widehat{K} : \mathcal{D}(\widehat{K}) \subset \mathcal{F}(\widehat{\Omega}) \rightarrow \mathcal{F}(\widehat{\Omega})$ we define the left action of K on ψ and the right action of \widehat{K} on ψ via

$$(K_l\psi)(x, y) = (K\psi(\cdot, y))(x), \quad (\widehat{K}_r\psi)(x, y) = (\widehat{K}\psi(x, \cdot))(y), \quad (1)$$

where we assume that ψ is such that these expressions are well-defined, i.e. $\psi(\cdot, y) \in \mathcal{D}(K), \psi(x, \cdot) \in \mathcal{D}(\widehat{K})$. An important special case to keep in mind is $\Omega = \{1, \dots, n\}$ and $\widehat{\Omega} = \{1, \dots, m\}$ finite sets, in which case we identify functions on Ω or $\widehat{\Omega}$ with column vectors and functions on $\Omega \times \widehat{\Omega}$ with $n \times m$ matrices. In that case, an operator K (resp. \widehat{K}) on functions on Ω (resp. $\widehat{\Omega}$) coincides with a $n \times n$ (resp. $m \times m$) matrix. Denoting by \mathbb{K} (resp. $\widehat{\mathbb{K}}$) such matrices, one has

$$K_l\psi(x, y) = \sum_{z \in \Omega} \mathbb{K}(x, z)\psi(z, y) = (\mathbb{K}\psi)(x, y), \quad (2)$$

$$\widehat{K}_r\psi(x, y) = \sum_{u \in \widehat{\Omega}} \widehat{\mathbb{K}}(y, u)\psi(x, u) = (\psi\widehat{\mathbb{K}}^T)(x, y). \quad (3)$$

Namely, left action of K corresponds to left matrix multiplication and right action of \widehat{K} corresponds to right multiplication with the transposed matrix. The same picture arises when $\Omega, \widehat{\Omega}$ are countable sets.

For two operators K_1, K_2 working on the same domain we denote, as usual, K_1K_2 their product or composition, i.e.

$$(K_1K_2f)(x) = (K_1(K_2f))(x)$$

and

$$[K_1, K_2] = K_1K_2 - K_2K_1$$

the commutator of K_1 and K_2 . In order to be well-defined, in particular in the case of unbounded operators, we assume that K_1, K_2 are working on a common domain of functions \mathbf{D} , such that $K_i\mathbf{D} \subset \mathbf{D}$ for $i = 1, 2$.

More precisely, we abbreviate $\mathcal{B}(\Omega)$ for algebras of linear operators working on a common domain \mathbf{D} , i.e. $\mathbf{D} \subset \mathcal{D}(K)$ for all $K \in \mathcal{B}(\Omega)$. This common domain \mathbf{D} is left invariant by the operators, i.e. $K(\mathbf{D}) \subset \mathbf{D}$ for all

$K \in \mathcal{B}(\Omega)$. So in such a context expressions like $\prod_{i=1}^n K_i^{n_i}$ are well-defined ($n_i \in \mathbb{N}$). The most important contexts in which we naturally have such a common domain and an algebra of operators working on it are the following.

1. Formal differential operators working on smooth functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ with compact support (or smooth functions from a bounded subset of \mathbb{R}^d). Both differential operators, and multiplication with a polynomial keep this domain invariant. Moreover, in most cases, if an operator belonging to this class generates a semigroup of contractions (defined via the Hille-Yosida theorem), then this common domain is a core, i.e., the graph closure of the operator working on \mathbf{D} coincides with the generator. See e.g. [10],[9] for details on generators and semigroups, and corresponding diffusion processes. See also [26] for standard arguments to extend duality from generators to contraction semigroups. Subtleties can arise for domains with boundary: in that case is important to specify boundary conditions to fix the closure of the operators. These issues will not be dealt with in general here, but on a case-to-case basis later when we work with diffusions on domains.
2. Context of finite set: in this case all functions belong to \mathbf{D} , and operators can be identified with matrices, i.e., in that case $\mathcal{B}(\Omega)$ is just a subalgebra of matrices.
3. Context of a countable set $\Omega = \mathbb{N}_0$, where \mathbb{N}_0 denotes the set of positive integers, with zero included. In that case, the operators we have in mind are finite difference operators, and multiplication operators, working on e.g. a common domain of functions $f : \mathbb{N}_0 \rightarrow \mathbb{R}$ going to zero faster than any polynomial as $n \rightarrow \infty$.

These algebras are often representations of an abstract Lie algebra \mathcal{H} (such as e.g. the Heisenberg algebra, see Sec. 3 and [19]). For a general algebra \mathcal{H} , we define the dual algebra \mathcal{H}^* as the algebra with the same elements as in \mathcal{H} but with product “*” defined by $a * b = b \cdot a$, where \cdot is the product in \mathcal{H} . A typical example in the finite dimensional setting is the algebra of $n \times n$ matrices, where the map $A \rightarrow A^T$ maps the algebra into the dual algebra ($(AB)^T = B^T A^T$).

DEFINITION 2.1. *For a function $\psi : \Omega \times \widehat{\Omega} \rightarrow \mathbb{R}$ we say that ψ is left exhaustive if the relation $K_l \psi = 0$ implies $K = 0$, and correspondingly we call ψ right exhaustive if the relation $\widehat{K}_r \psi = 0$ implies $\widehat{K} = 0$.*

Notice that in the context of finite sets $\Omega, \widehat{\Omega}$ being right or left exhaustive just means that the matrix associated to ψ is invertible. In particular we must then require $|\Omega| = |\widehat{\Omega}|$.

2.2 Duality

We begin this section with the standard definition of duality [10, 26].

DEFINITION 2.2. *Suppose $\{X_t\}_{t \geq 0}$, $\{\widehat{X}_t\}_{t \geq 0}$ are Markov processes with state spaces Ω and $\widehat{\Omega}$ and $D : \Omega \times \widehat{\Omega} \rightarrow \mathbb{R}$ a bounded measurable function. The processes $\{X_t\}_{t \geq 0}$, $\{\widehat{X}_t\}_{t \geq 0}$ are said to be dual with respect to D if*

$$\mathbb{E}_x D(X_t, \widehat{x}) = \widehat{\mathbb{E}}_{\widehat{x}} D(x, \widehat{X}_t), \quad (4)$$

for all $x \in \Omega, \widehat{x} \in \widehat{\Omega}$ and $t > 0$. In (4) \mathbb{E}_x is the expectation with respect to the law of the $\{X_t\}_{t \geq 0}$ process started at x , while $\widehat{\mathbb{E}}_{\widehat{x}}$ denotes expectation with respect to the law of the $\{\widehat{X}_t\}_{t \geq 0}$ process initialized at \widehat{x} .

Since in most of the examples of duality between two processes, this property is equivalent with duality of the corresponding generators, we focus here on dualities between linear operators.¹

Indeed, in most examples, duality of the generators in turn follows from duality between more elementary “building blocks” (such as derivatives and multiplication operators). So in the next section, we focus on duality between general operators and show how from that notion, which is conserved under sums and products, it is natural to consider duality between algebras of operators.

DEFINITION 2.3. *Let $K \in \mathcal{B}(\Omega)$, $\widehat{K} \in \mathcal{B}(\widehat{\Omega})$ and $D : \Omega \times \widehat{\Omega} \rightarrow \mathbb{R}$. Then we say that K and \widehat{K} are dual to each other with duality function D if*

$$K_l D = \widehat{K}_r D \quad (5)$$

where we assume that both sides are well defined, i.e. $D(\cdot, \widehat{x}) \in \mathcal{D}(K)$ for all $\widehat{x} \in \widehat{\Omega}$ and $D(x, \cdot) \in \mathcal{D}(\widehat{K})$ for all $x \in \Omega$. We denote this property by $K \xrightarrow{D} \widehat{K}$.

In the following we collect elementary but important properties of the relation \xrightarrow{D} .

THEOREM 2.1. *Let $K_1, K_2 \in \mathcal{B}(\Omega)$, $\widehat{K}_1, \widehat{K}_2 \in \mathcal{B}(\widehat{\Omega})$. Suppose that $K_1 \xrightarrow{D} \widehat{K}_1$, $K_2 \xrightarrow{D} \widehat{K}_2$, and further $c_1, c_2 \in \mathbb{R}$ then we have*

1. $\widehat{K}_1 \xrightarrow{\tilde{D}} K_1$, with $\tilde{D}(\widehat{x}, x) = D(x, \widehat{x})$.
2. $c_1 K_1 + c_2 K_2 \xrightarrow{D} c_1 \widehat{K}_1 + c_2 \widehat{K}_2$.
3. $K_1 K_2 \xrightarrow{D} \widehat{K}_2 \widehat{K}_1$, in particular $K_1^n \xrightarrow{D} \widehat{K}_1^n$, $n \in \mathbb{N}$.

¹An exception is the duality between Brownian motion with reflection and Brownian motion with absorption, because in that case the duality function $D(x, y) = I(x \leq y)$ (with I denoting the indicator function) is not in the domain of the generator. See [24] for details on dualities of this type.

4. $[K_1, K_2] \rightarrow^D [\widehat{K}_2, \widehat{K}_1] = -[\widehat{K}_1, \widehat{K}_2]$, i.e., commutators of dual operators are dual.
5. If $S \in \mathcal{B}(\Omega)$ commutes with K_1 , then $K_1 \rightarrow^{S_i D} \widehat{K}_1$, and if $\widehat{S} \in \mathcal{B}(\widehat{\Omega})$ commutes with \widehat{K}_1 , then $K_1 \rightarrow^{\widehat{S}_r D} \widehat{K}_1$.
6. If for a collection $\{K_i, i \in I\} \subset \mathcal{B}(\Omega)$, and $\{\widehat{K}_i : i \in I\} \subset \mathcal{B}(\widehat{\Omega})$ we have $K_i \rightarrow^D \widehat{K}_i$ then every element of the algebra generated by $\{K_i, i \in I\}$ is dual to an element of the algebra generated by $\{\widehat{K}_i : i \in I\}$. More precisely:

$$K_{i_1}^{n_1} \dots K_{i_k}^{n_k} \rightarrow^D \widehat{K}_{i_k}^{n_k} \dots \widehat{K}_{i_1}^{n_1} \quad (6)$$

for all $n_1, \dots, n_k \in \mathbb{N}$, and for constants $\{c_i : i \in I\}$

$$\sum_{i \in I} c_i K_i \rightarrow^D \sum_{i \in I} c_i \widehat{K}_i .$$

7. Suppose (only in this item) that $K_1 \in \mathcal{B}(\Omega_1)$, $\widehat{K}_1 \in \mathcal{B}(\widehat{\Omega}_1)$, $K_2 \in \mathcal{B}(\Omega_2)$, $\widehat{K}_2 \in \mathcal{B}(\widehat{\Omega}_2)$. If $K_1 \rightarrow^{D_1} \widehat{K}_1$ and $K_2 \rightarrow^{D_2} \widehat{K}_2$, then $K_1 \otimes K_2 \rightarrow^{D_1 \otimes D_2} \widehat{K}_1 \otimes \widehat{K}_2$, where

$$D_1 \otimes D_2(x_1, x_2; \widehat{x}_1, \widehat{x}_2) = D_1(x_1, \widehat{x}_1) D_2(x_2, \widehat{x}_2).$$

8. If K and \widehat{K} generate Markov semigroups $S_t = e^{tK}$ and $\widehat{S}_t = e^{t\widehat{K}}$, and if $D : \Omega \times \widehat{\Omega} \rightarrow \mathbb{R}$ are functions such that $D(\cdot, \widehat{x}), \widehat{S}_t D(\cdot, \widehat{x}) \in \mathcal{D}(K)$ for all $\widehat{x} \in \widehat{\Omega}$ and $D(x, \cdot), S_t D(x, \cdot) \in \mathcal{D}(\widehat{K})$ for all $x \in \Omega$, then $K \rightarrow^D \widehat{K}$ implies $S_t \rightarrow^D \widehat{S}_t$. If moreover these semigroups correspond to Markov processes $\{X_t, t \geq 0\}, \{\widehat{X}_t, t \geq 0\}$ on $\Omega, \widehat{\Omega}$, the relation $S_t \rightarrow^D \widehat{S}_t$ reads in terms of these processes:

$$\mathbb{E}_x D(X_t, \widehat{x}) = \widehat{\mathbb{E}}_{\widehat{x}} D(x, \widehat{X}_t) , \quad (7)$$

for all $x \in \Omega, \widehat{x} \in \widehat{\Omega}$ and $t > 0$.

PROOF. The properties listed in the theorem are elementary and their proof is left to the reader. The only technical issue is item 8, i.e., passing from duality on the level of generator to duality of the corresponding semigroups and processes. This result is obtained by using the uniqueness of the semigroup. More precisely, if K (or the closure of K) generates a semigroup S_t (formally denoted by e^{tK}) then for all functions $D(\cdot, \widehat{x})$ which are in the domain of K for all $\widehat{x} \in \widehat{\Omega}$, the unique solution of the equation

$$\frac{d}{dt} f_t(x, \widehat{x}) = K_t f_t(x, \widehat{x}) \quad (8)$$

with initial condition $f_0(x, \hat{x}) = D(x, \hat{x})$ is given by $f_t(x, \hat{x}) = (S_t)_l D(x, \hat{x})$.

On the other hand, the relation $K \rightarrow^D \widehat{K}$ implies that also $f_t(x, \hat{x}) = (\widehat{S}_t)_r D(x, \hat{x})$ solves the same equation. Indeed, since \widehat{S}_t has generator \widehat{K} , it follows that

$$\frac{d}{dt}(\widehat{S}_t)_r D(x, \hat{x}) = (\widehat{S}_t)_r \widehat{K}_r D(x, \hat{x}) = (\widehat{S}_t)_r K_l D(x, \hat{x}) = K_l (\widehat{S}_t)_r D(x, \hat{x})$$

where in the last equality we used that $\widehat{S}_t D(\cdot, \hat{x}) \in \mathcal{D}(K)$.

As the equation (8) has a unique solution it follows that $S_t \rightarrow^D \widehat{S}_t$. For more details see [26] theorem 4.13, Chapter 3, pag. 161 in the context of spin systems and also [9, 24] for more general cases. \square

REMARK 2.1. *Item 6 of theorem 2.1 is useful in particular if the collection $\{K_i, i \in I\} \subset \mathcal{B}(\Omega)$ is a generating set for the algebra. Then every element of the algebra has a dual by (6), and it suffices to know dual operators for the generating set to infer dual operator for a general element of the algebra. In practice, one starts from such a generating set and the commutation relations between its elements (defining the algebra) and associates to it by a single duality function a set of dual operators with the same commutation relations up to a change of sign (cfr. item 4). In other words, one moves via the duality function from a representation of the algebra to a representation of the dual algebra.*

REMARK 2.2. *The relation (7) is the form in which one usually formulates duality between two Markov processes (cfr. Def. 2.2). Remark however that the relation $S_t \rightarrow^D \widehat{S}_t$ between the semigroups is more general. It may happen that S_t is a Markov semigroup, whereas \widehat{S}_t is not. E.g., mass can get lost in the evolution according to the dual semigroup \widehat{S}_t , which means $\widehat{S}_t 1 \neq 1$, or it can happen that \widehat{S}_t is not a positive operator (see e.g. Remark 4.2 of [5] where the duality between the generator of the symmetric exclusion process and a non-positive differential operator is exhibited, and e.g. [26] chapter III, section 4, for duality with Feynman Kac factors in the context of spin systems).*

In [27], the author studies the so-called duality space associated to two operators. In our notation, this is the set

$$\mathcal{D}(K, \widehat{K}) = \{D : \Omega \times \widehat{\Omega} \rightarrow \mathbb{R} \mid K \rightarrow^D \widehat{K}\}. \quad (9)$$

Our point of view here is *instead to consider for a fixed duality function the set of pairs of operators (K, \widehat{K}) such that $K \rightarrow^D \widehat{K}$* . These operators then usually form a representation and a dual representation of a given algebra, with D as intertwiner (see, however, Eq. (12)).

In theorem 2.1 item 5, we see that we can produce new duality functions via “symmetries”, i.e., operators commuting with K or \widehat{K} . In the context

of finite sets $\Omega = \widehat{\Omega} = \{1, \dots, n\}$, we have more: if there exists an invertible duality function D , then all other duality functions are obtained via symmetries acting on D . We formulate this more precisely in the following proposition. We will give examples in the subsequent sections.

PROPOSITION 2.1.

1. Let $\Omega = \widehat{\Omega} = \{1, \dots, n\}$, and let $K \rightarrow^D \widehat{K}$. Suppose furthermore that the associated $n \times n$ matrix \mathbb{D} is invertible. Then, if $K \rightarrow^{D'} \widehat{K}$, we have that there exists S commuting with K such that $D' = SD$.
2. For general $\Omega, \widehat{\Omega}$ we have the following. Suppose D and D' are duality functions for the duality between K and \widehat{K} . Suppose furthermore that $D' = S_l D$, for some operator S . Then we have

$$(KS - SK)_l D = 0 . \quad (10)$$

In particular if D is left exhaustive, then we conclude $[S, K] = 0$. Similarly, if $D' = \widehat{S}_r D$ then

$$(\widehat{K}\widehat{S} - \widehat{S}\widehat{K})_r D = 0 . \quad (11)$$

and if D is right exhaustive, then we conclude $[\widehat{S}, \widehat{K}] = 0$.

PROOF. In the proof of the first item, with slight abuse of notation, we use the notation K, \widehat{K}, D, S both for the operators and for their associated matrices. We have, by assumption

$$KD = D\widehat{K}^T .$$

By invertibility of D , $S = D'D^{-1}$ is well-defined and we have

$$SK = D'D^{-1}K = D'\widehat{K}^T D^{-1} ,$$

and

$$KS = KD'D^{-1} = D'\widehat{K}^T D^{-1} ,$$

hence $[K, S] = 0$.

For the second item, use $K \rightarrow^D \widehat{K}, K \rightarrow^{D'} \widehat{K}$ to conclude

$$(KS)_l D = K_l(S_l D) = K_l D' = \widehat{K}_r D' ,$$

as well as

$$(SK)_l D = S_l \widehat{K}_r(D) = \widehat{K}_r(S_l D) = \widehat{K}_r D' ,$$

Hence

$$([S, K])_l D = 0 .$$

The remaining part of the proof (for the right action case) is similar. \square

2.3 Self duality

Self-duality is duality between an operator and itself, i.e., referring to definition 2.3: $\Omega = \widehat{\Omega}$ and $K = \widehat{K}$. The corresponding duality function such that $K \xrightarrow{D} K$ is then a function $D : \Omega \times \Omega \rightarrow \mathbb{R}$ and we call it a self-duality function. For self-duality, of course, all the properties listed in theorem 2.1 hold.

In the finite case $\Omega = \{1, \dots, n\}$, a self-duality function is a $n \times n$ matrix and self-duality reads, in matrix form,

$$KD = DK^T .$$

Therefore, in this setting such a matrix D can always be found because every matrix is similar to its transposed [35], i.e., self-duality always holds with an invertible D .

Other self-duality functions can then be found by acting on a given self-duality function with symmetries of K (i.e. operators S commuting with K), as we derived in item 5 of theorem 2.1 and in proposition 2.1. In particular we have in this finite context, in the notation (9):

$$\mathcal{D}(K, K) = \{SD : [S, K] = 0\} , \tag{12}$$

with D an arbitrary invertible self-duality function. So this means that the correspondence between self-duality functions and symmetries of K is one-to-one. This characterization of the duality space has the advantage that the set of operators commuting with a given operator is easier to identify. In the finite setting, if K is the generator (resp. transition operator) of a continuous-time (resp. discrete-time) Markov chain, then if this Markov chain has a reversible probability measure $\mu : \Omega \rightarrow [0, 1]$, a duality function for self-duality is given by the diagonal matrix

$$D(x, y) = \delta_{x,y} \frac{1}{\mu(x)} .$$

This is easily verified from the detailed balance relation $\mu(x)\mathbb{K}(x, y) = \mathbb{K}(y, x)\mu(y)$. This “cheap” duality function is usually not very useful since it is diagonal, but it can be turned in a more “useful” one by acting with symmetries. All the known self-duality functions in discrete interacting particle systems such as the exclusion process, independent random walkers, the inclusion process, etc. can be obtained by this procedure [15].

3 Dualities in the context of the Heisenberg algebra.

The (abstract) Heisenberg algebra $\mathcal{H}(m)$ [19] is an algebra generated by $2m$ elements $\mathbf{K}_i, \mathbf{K}_i^\dagger, i = 1, \dots, m$, satisfying the following commutation relations:

$$[\mathbf{K}_i, \mathbf{K}_j] = 0, \quad [\mathbf{K}_i^\dagger, \mathbf{K}_j^\dagger] = 0, \quad [\mathbf{K}_i, \mathbf{K}_j^\dagger] = \delta_{i,j}I, \quad i, j = 1, \dots, m \quad (13)$$

where I is the unit element of $\mathcal{H}(m)$. Relations (13) are called *canonical commutation relations*.

In this section we focus on representations of Heisenberg algebra and its dual algebra, and the corresponding duality functions that connect these different representations.

3.1 Standard creation and annihilation operators

As a first example of a representation of $\mathcal{H}(1)$, let us start with the operators A^\dagger, A working on smooth functions $f : \mathbb{R} \rightarrow \mathbb{R}$ with compact support, defined as

$$Af(x) = f'(x), \quad A^\dagger f(x) = xf(x), \quad (14)$$

in physical jargon: the annihilation and creation operators. These operators satisfy the canonical commutation relations (13) for $m = 1$ with $\mathbf{K}_1 = A, \mathbf{K}_1^\dagger = A^\dagger$. Indeed $[A, A^\dagger] = I$, where I is the identity operator, while the remaining relations are trivially satisfied.

The same commutation relations (13), up to a negative sign, can be achieved using operators working on discrete functions. Considering

$$af(n) = nf(n-1), \quad a^\dagger f(n) = f(n+1), \quad (15)$$

acting on functions $f : \mathbb{N}_0 \rightarrow \mathbb{R}$, we have $[a, a^\dagger] = -I$. Therefore, in view of the item 4 of theorem 2.1 and Remark 2.1, the operators a, a^\dagger are natural candidates for duality with A, A^\dagger .

To find D such that $A \xrightarrow{D} a$, we use the definition 2.3:

$$A_l D(x, n) = D'(x, n) = a_r D(x, n) = nD(x, n-1),$$

which yields

$$D(x, n) = \sum_{k=0}^n \binom{n}{k} c_{n-k} x^k$$

with $\{c_i : i \in \mathbb{N}\}$ a sequence of constants. In the same way, the duality condition $A^\dagger \xrightarrow{D} a^\dagger$, produces

$$A_l^\dagger D(x, n) = xD(x, n) = a_r^\dagger D(x, n) = D(x, n+1),$$

which gives

$$D(x, n) = x^n D(x, 0),$$

with $D(x, 0)$ an arbitrary function. Therefore, if we want both dualities to hold with the same duality function, then we are restricted to the choice

$$D(x, n) = c_0 x^n .$$

Without loss of generality we can choose $c_0 = 1$.

As a consequence, by using item 6 of theorem 2.1, we obtain the following result.

THEOREM 3.1. *For $0 \leq n \leq m$, let $\alpha_n : \mathbb{R} \rightarrow \mathbb{R}$ be a finite sequence of polynomials. The differential operator K defined on smooth functions with compact support, $f \in \mathcal{C}_0^\infty(\mathbb{R})$, of the form*

$$K = \sum_{n=0}^m \alpha_n(x) \frac{d^n}{dx^n} = \sum_{n=0}^m \alpha_n(A^\dagger) A^n$$

is dual with duality function $D(x, n) = x^n$ to the operator \hat{K} acting on the space of real valued functions $f : \mathbb{N}_0 \rightarrow \mathbb{R}$, $f = \{f_n\}_{n \in \mathbb{N}_0}$:

$$\hat{K} = \sum_{n=0}^m a^n \alpha_n(a^\dagger)$$

where the operators A, A^\dagger are defined in (14) and a, a^\dagger are defined in (15).

REMARK 3.1. *Using the Doi-Peliti method, in [30], some results of the type of theorem 3.1 are obtained for reaction diffusion systems.*

We close this section with a representation of the Heisenberg algebra $\mathcal{H}(m)$ with $m > 1$, which generalizes the previous one to functions of several variables and that will be used in the next sections. The generators of this representation, working on the smooth functions $f : \mathbb{R}^m \rightarrow \mathbb{R}$ with compact support, are

$$A_i f(x) = \frac{\partial}{\partial x_i} f(x), \quad A_i^\dagger f(x) = x_i f(x), \quad i = 1, \dots, m, \quad (16)$$

also in this case they are called annihilation and creation operators. Clearly A_i and A_i^\dagger satisfy the canonical commutation relations (13), thus they generate a representation of the Heisenberg algebra $\mathcal{H}(m)$.

In analogy with (15) we introduce a discrete algebra generated by the operators acting on functions $f : \mathbb{N}_0^m \rightarrow \mathbb{R}$ via

$$a_i f(n) = n_i f(n - e_i), \quad a_i^\dagger f(n) = f(n + e_i), \quad i = 1, \dots, m \quad (17)$$

where $n \in \mathbb{N}_0^m$ and $e_i \in \mathbb{N}_0^m$ is the i -th canonical unit vector defined via $(e_i)_j = \delta_{i,j}$. On the basis of the previous discussion we have that, for each i

the dualities $A_i \xrightarrow{D_i} a_i, A_i^\dagger \xrightarrow{D_i} a_i^\dagger$ hold with duality function $D_i(x_i, n_i) = x_i^{n_i}$. Thus by Theorem 2.1, item 7, we have duality between the tensor products of the generators of the continuous representation ($\otimes_{i=1}^m K_i$ with $K_i \in \{A_i, A_i^\dagger\}$) and the tensor products of the generators of the discrete one ($\otimes_{i=1}^m \widehat{K}_i$ with $\widehat{K}_i \in \{a_i, a_i^\dagger\}$). The duality function is given by

$$D(n, x) = \prod_{i=1}^m D_i(x_i, n_i) = \prod_{i=1}^m x_i^{n_i} \quad (18)$$

that is $\otimes_{i=1}^m K_i \xrightarrow{D} \otimes_{i=1}^m \widehat{K}_i$.

3.2 Generalization

In the following proposition we show how to generate the duality functions for more general generators $\mathbf{A}, \mathbf{A}^\dagger$ of a representation of the Heisenberg algebra, namely by repetitive action of the creation operator on the “vacuum” which is annihilated by the operator \mathbf{A} .

PROPOSITION 3.1. *Suppose $[\mathbf{A}, \mathbf{A}^\dagger] = I$, and let $D(x, n)$ be functions such that*

$$\begin{aligned} (\mathbf{A}_l^\dagger)^n D(x, 0) &= D(x, n), \\ \mathbf{A}_l D(x, 0) &= 0, \end{aligned} \quad (19)$$

then $\mathbf{A} \xrightarrow{D} a$ and $\mathbf{A}^\dagger \xrightarrow{D} a^\dagger$, where a, a^\dagger are the discrete representation defined in (15). As a consequence, for a finite sequence of polynomials α_n , with $0 \leq n \leq m$, we have the analogue of theorem 3.1:

$$\sum_{n=0}^m \alpha_n(\mathbf{A}^\dagger) \mathbf{A}^n \xrightarrow{D} \sum_{n=0}^m a^n \alpha_n(a^\dagger).$$

PROOF. We have $\mathbf{A}^\dagger \xrightarrow{D} a^\dagger$ by the assumption on \mathbf{A}^\dagger in (19) and the definition of a^\dagger in (15). We therefore have to prove $\mathbf{A} \xrightarrow{D} a$. Start from the commutation relation $[\mathbf{A}, \mathbf{A}^\dagger] = I$ to write

$$\mathbf{A}(\mathbf{A}^\dagger)^n = (\mathbf{A}^\dagger)^n \mathbf{A} + [\mathbf{A}, (\mathbf{A}^\dagger)^n] = (\mathbf{A}^\dagger)^n \mathbf{A} + n(\mathbf{A}^\dagger)^{n-1}.$$

Then use the assumptions (19) to deduce

$$\begin{aligned} \mathbf{A}_l D(x, n) &= \mathbf{A}_l (a_r^\dagger)^n D(x, 0) = \mathbf{A}_l (\mathbf{A}_l^\dagger)^n D(x, 0) \\ &= (\mathbf{A}_l^\dagger)^n \mathbf{A}_l D(x, 0) + n(\mathbf{A}_l^\dagger)^{n-1} D(x, 0) = n(\mathbf{A}_l^\dagger)^{n-1} D(x, 0) \\ &= n(a_r^\dagger)^{n-1} D(x, 0) \\ &= nD(x, n-1) = a_r D(x, n). \end{aligned}$$

□

As an application, we can choose linear combinations of multiplication and derivative

$$\mathbf{A} = c_1x + c_2\frac{d}{dx}, \quad \mathbf{A}^\dagger = c_3x + c_4\frac{d}{dx}, \quad (20)$$

with the real constants satisfying $c_2c_3 - c_1c_4 = 1$, then we satisfy the commutation relation $[\mathbf{A}, \mathbf{A}^\dagger] = I$. To find the corresponding duality function that “switches” from $\mathbf{A}, \mathbf{A}^\dagger$ to a, a^\dagger , we start with

$$\mathbf{A}_l D(x, 0) = c_1xD(x, 0) + c_2D'(x, 0) = a_r D(x, 0) = 0$$

which gives as a choice

$$D(x, 0) = \exp\left(-\frac{c_1}{c_2}\frac{x^2}{2}\right)$$

and next,

$$D(x, n) = (\mathbf{A}_l^\dagger)^n D(x, 0) = \left(c_3x + c_4\frac{d}{dx}\right)^n D(x, 0).$$

An important particular case (related to the harmonic oscillator in quantum mechanics and the Ornstein-Uhlenbeck process), is when $c_1 = c_2 = 1/2$ and $c_3 = -c_4 = 1$. With this choice one finds that the duality function is $D(x, n) = e^{-x^2/2}H_n(x)$, where H_n is the Hermite polynomial of order n .

3.3 Dualities with two continuous variables

Within the scheme described in section 2 we can also find dualities between two operators both working on continuous variables, as the following example shows.

Consider again the operators A, A^\dagger in (14). A “dual” commutation relation (in the sense of item 4 of theorem 2.1) can be obtained by considering a copy of those operators and exchanging their role. Namely, we look for dualities $d/dx \xrightarrow{D} y, x \xrightarrow{D} d/dy$. Imposing that the left action of d/dx (resp. x) does coincide with the right action of y (resp. d/dy) one finds the duality function $D(x, y) = e^{xy}$. As a consequence one immediately has the following

THEOREM 3.2. *For $0 \leq n \leq m$, let $\alpha_n : \mathbb{R} \rightarrow \mathbb{R}$ be a finite sequence of polynomials. A differential operator working on smooth functions of the real variable x and with the generic form*

$$K = \sum_{n=0}^m \alpha_n(x) \frac{d^n}{dx^n}$$

is dual, with duality function $D(x, y) = e^{xy}$, to the operator working on smooth functions of the real variable y given by

$$\widehat{K} = \sum_{n=0}^m y^n \alpha_n\left(\frac{d}{dy}\right).$$

As a first simple illustration, consider the operator $\frac{1}{2}\frac{d}{dx^2}$, which is dual to the multiplication operator $\frac{y^2}{2}$ (with duality function e^{xy}). The semigroup with generator $\frac{1}{2}\frac{d}{dx^2}$ is Brownian motion, and the “semigroup” generated by $\frac{y^2}{2}$ is of course multiplication with $e^{ty^2/2}$. As a consequence, denoting by $(W_t)_{t \geq 0}$ the standard Brownian motion, we have for the corresponding semigroups:

$$\mathbb{E}_x [e^{yX_t}] = \mathbb{E} [e^{y(x+W_t)}] = \widehat{\mathbb{E}}_y [e^{Y_t x}] = e^{\frac{ty^2}{2}} e^{xy}$$

which in this case can of course be directly verified from the equality $\mathbb{E}[e^{yW_t}] = e^{\frac{ty^2}{2}}$.

If we specify that the operators work on functions $f : [0, \infty) \rightarrow \mathbb{R}$, we can use the duality $d/dx \rightarrow^D -y, x \rightarrow^D -d/dy$ with duality function $D(x, y) = e^{-xy}$ and when the operators can be interpreted as pregenerators of diffusions one has the following

COROLLARY 3.1. *The diffusion pregenerator*

$$\mathcal{L} = (c_1 x^2 + c_2 x) \frac{d^2}{dx^2} + (c_3 x) \frac{d}{dx}$$

with $c_1 > 0, c_2 \geq 0$ on the domain

$$\mathcal{D}(\mathcal{L}) = \{f : [0, \infty) \rightarrow \mathbb{R} : f, f', f'' \in \mathcal{C}([0, \infty)), \mathcal{L}f(0) = 0\}$$

is dual to

$$\widehat{\mathcal{L}} = c_1 y^2 \frac{d^2}{dy^2} + (-c_2 y^2 + c_3 y) \frac{d}{dy}$$

on the same domain, with duality function $D(x, y) = e^{-xy}$. For the corresponding diffusion processes $\{X_t : t \geq 0\}$, $\{Y_t : t \geq 0\}$ we thus have

$$\mathbb{E}_x e^{-yX_t} = \widehat{\mathbb{E}}_y e^{-xY_t} . \quad (21)$$

The particular case $c_2 = 0$ gives that the diffusion pregenerator $c_1 x^2 d^2/dx^2 + c_3 x d/dx$ is self-dual.

REMARK 3.2. Notice that naively applying $d/dx \rightarrow^D y, x \rightarrow^D d/dy$ with duality function $D(x, y) = e^{xy}$ in the previous context yields that

$$\mathcal{L} = (c_1 x^2 + c_2 x) \frac{d^2}{dx^2} + (c_3 x) \frac{d}{dx}$$

is dual to

$$\widehat{\mathcal{L}} = c_1 y^2 \frac{d^2}{dy^2} + (c_2 y^2 + c_3 y) \frac{d}{dy}$$

with duality function $D(x, y) = e^{xy}$ but it might be that the corresponding relation on the level of the semigroup

$$\mathbb{E}_x e^{yX_t} = \widehat{\mathbb{E}}_y e^{xY_t} . \quad (22)$$

does not yield useful information because the corresponding processes do not possess exponential moments, i.e., the relation reads “ $\infty = \infty$ ”.

3.4 Discrete creation and annihilation operators

The following example starts from a finite dimensional representation of the Heisenberg algebra $\mathcal{H}(1)$ (in the spirit of [7, 20]). We consider $\Omega = \Omega_N = \{0, \dots, N\}$ and $\widehat{\Omega} = \mathbb{N}$. For functions $f : \Omega_N \rightarrow \mathbb{R}$ we define the operators

$$\begin{aligned} a_N f(k) &= (N - k) f(k + 1) + (2k - N) f(k) - k f(k - 1) , \\ a_N^\dagger f(k) &= \sum_{r=0}^{k-1} (-1)^{k-1-r} \frac{\binom{N}{r}}{\binom{N}{k}} f(r) , \end{aligned} \quad (23)$$

with the convention $f(-1) = f(N + 1) = 0$. Consider

$$D_N(k, n) = \frac{\binom{k}{n}}{\binom{N}{n}} = \frac{k(k-1) \cdots (k-(n-1))}{N(N-1) \cdots (N-(n-1))} \quad (24)$$

with the convention $D_N(k, 0) = 1$, $D_N(k, N + 1) = 0$. Let us denote by \mathcal{W}_N the vector space generated by the functions $k \mapsto D_N(k, n)$, $0 \leq n \leq N$.

PROPOSITION 3.2.

$$\begin{aligned} (a_N)_l D_N(k, n) &= n D_N(k, n - 1), \quad \forall 1 \leq n, \forall k \geq n - 1 , \\ (a_N)_l D_N(k, 0) &= 0 \quad \forall 0 \leq k \leq N , \\ (a_N^\dagger)_l D_N(k, n) &= D_N(k, n + 1) \quad \forall 0 \leq n \leq N, k \geq n . \end{aligned} \quad (25)$$

As a consequence, as operators on \mathcal{W}_N we have

$$[a_N, a_N^\dagger] = I , \quad (26)$$

i.e., a_N, a_N^\dagger form a finite dimensional representation of the canonical commutation relations.

PROOF. Straightforward computation. \square

REMARK 3.3. Notice that in the limit $N \rightarrow \infty$, putting $k/N = x$, and $f(k) = \phi(x) = \phi(k/N)$, $a_N f(k)$ converges to $d\phi/dx$. Next, notice that $D_N(k, n) = \phi_N^{(n)}(x)$, where

$$\phi_N^{(n)}(x) = x \left(\frac{x - \frac{1}{N}}{1 - \frac{1}{N}} \right) \cdots \left(\frac{x - \frac{n-1}{N}}{1 - \frac{n-1}{N}} \right) ,$$

converges to x^n . The effect of the operator a_N^\dagger on $D_N(k, n)$ is to raise the index n by one. Since $D_N(k, n) = \phi_N^{(n)}(x) \rightarrow x^n$ for $N \rightarrow \infty$ and $a_N^\dagger D_N(k, n) = A^\dagger \phi_N^{(n)}(x) \rightarrow x^{n+1}$ for $N \rightarrow \infty$, we conclude that in the limit $N \rightarrow \infty$, the operator a_N^\dagger coincides with the multiplication operator A^\dagger defined in (14).

The discrete finite dimensional representation of the Heisenberg algebra given in proposition 3.2 will be used at the end of section 4 to fit within the scheme of a change of representation the classical duality between the Moran model and the block-counting process of the Kingman's coalescent. Since we only use the block-counting process of the Kingman's coalescent (rather than the full partition-valued process) we will use the name "Kingman's coalescent" for that block-counting process here and also later on. We end this section with a comment on the relation between the discrete representation in proposition 3.2 and the Binomial distribution. This also offers an alternative simple way to see the commutation relation (26).

3.5 Relation with invariant measures

In many models where there is duality or self-duality (see e.g. Exclusion, Inclusion and Brownian Energy processes), there exists a one-parameter family of invariant measures ν_ρ , (see e.g. Section 3.1 of [5] for more details) and integrating the duality function w.r.t. these measures usually gives a simple expression of the parameter ρ . In the context of diffusion processes with discrete dual, this relation is usually that the duality function with n dual particles integrated over the distribution ν_ρ equals ρ^n . A similar relation connects the polynomials $D_N(k, n)$ in (24) to the binomial distribution. This general relation between a natural one-parameter family of measures and the duality functions cannot be a coincidence and requires further investigation.

The polynomials $D_N(k, n)$ are (as a function of k) indeed naturally associated to the binomial distribution. Denoting by

$$\nu_{N,\rho}(k) = \binom{N}{k} \rho^k (1-\rho)^{N-k}$$

the binomial distribution with success probability $\rho \in [0, 1]$, we have

$$\sum_{k=0}^N D_N(k, n) \nu_{N,\rho}(k) = \rho^n. \quad (27)$$

For a function $f : \Omega_N \rightarrow \mathbb{R}$ we define its binomial transform $\mathcal{T}f : [0, 1] \rightarrow \mathbb{R}$ by

$$(\mathcal{T}f)(\rho) = \sum_{k=0}^N f(k) \nu_{N,\rho}(k). \quad (28)$$

If for such f , we write its expansion

$$f(k) = \sum_{r=0}^N c_r D_N(k, r)$$

we say that f is of degree l if $c_l \neq 0$ and all higher coefficients $c_k, k > l$ are zero. We then have, using (27) and (28),

$$(\mathcal{T}f)(\rho) = \sum_{r=0}^N c_r \rho^r.$$

The functions f and $\mathcal{T}f$ have therefore the same components with respect to two different bases: one given by $\{D_N(k, r), r = 0, \dots, N\}$ which is a base of \mathbb{R}^{N+1} and the other given by $\{\rho^r, r = 0, \dots, N\}$ which is a base of the space of polynomials on $[0, 1]$ of degree at most equal to N . We then have, for all f :

$$(\mathcal{T}a_N f)(\rho) = (\mathcal{T}f)'(\rho)$$

and for all f with degree less than or equal to $N - 1$:

$$(\mathcal{T}a_N^\dagger f)(\rho) = \rho \cdot (\mathcal{T}f)(\rho).$$

This relation shows that the operators a_N, a_N^\dagger after binomial transformation turn into the standard creation and annihilation operators $(\rho, d/d\rho)$ for a restricted set of functions (polynomials of degree at most N).

4 Classical dualities of population dynamics

The scheme developed in the section 2, together with the change of representation discussed in section 3, allows to recover many of the well-know dualities of classical models of population genetics [1, 12, 25]. We first consider diffusion processes of the Wright-Fisher diffusion type and then discrete processes for a finite population of N individuals of the Moran type. In this section, as well as in section 5 when we consider generators L of diffusion processes on an interval or on a multidimensional simplex Ω , we will always define them with absorbing boundary conditions, i.e., the pregenerator (of which the generator is the graph closure) is defined on the domain of smooth functions f with compact support such that Lf vanishes on the boundary. In the case that the boundary is not attainable (such as Wright-Fisher diffusion with mutation, depending on the mutation rate) the domain of the pregenerator consists of smooth functions f with compact support contained in the interior of Ω . See [10] chapter 8, section 1, Theorem 1.4 for more details on generators with absorbing boundary conditions.

Diffusions and coalescents.

Consider smooth functions $f : [0, 1] \rightarrow \mathbb{R}$ vanishing at the boundaries 0 and 1. A diffusion process on $[0, 1]$ has generator of the form

$$\mathcal{L} = \alpha(x) \frac{d^2}{dx^2} + \beta(x) \frac{d}{dx} = \alpha(A^\dagger)A^2 + \beta(A^\dagger)A, \quad (29)$$

with A and A^\dagger defined in (14). More precisely, we choose

$$\begin{aligned} \alpha(x) &= \sum_{k=1}^{\infty} \alpha_k x^k, \\ \beta(x) &= \sum_{k=0}^{\infty} \beta_k x^k, \end{aligned} \quad (30)$$

where the coefficients α_k, β_k satisfy the following

$$\begin{aligned} \alpha_2 &= - \sum_{k \neq 2, k=1}^{\infty} \alpha_k, & \alpha_k &\geq 0 \quad \forall k \neq 2, \\ \beta_1 &= - \sum_{k \neq 1, k=0}^{\infty} \beta_k, & \beta_k &\geq 0 \quad \forall k \neq 1. \end{aligned} \quad (31)$$

Typical choices are $\alpha(x) = x - x^2$, $\beta(x) = (1 - x)$. By the duality $A \rightarrow^D a$, $A^\dagger \rightarrow^D a^\dagger$ and by theorem 3.1 we find that \mathcal{L} is dual to

$$\begin{aligned} \hat{\mathcal{L}}f(n) &= \left(a^2 \alpha(a^\dagger) + a \beta(a^\dagger) \right) f(n) \\ &= n(n-1) \sum_{k=1}^{\infty} \alpha_k (f(n+k-2) - f(n)) \\ &\quad + n \sum_{k=0}^{\infty} \beta_k (f(n+k-1) - f(n)) \end{aligned} \quad (32)$$

with duality function $D(x, n) = x^n$. By the conditions (31) on the coefficients, this corresponds to a Markov chain on the natural numbers.

We can then list a few examples.

1. Wright Fisher neutral diffusion.

$$\mathcal{L} = x(1-x) \frac{d^2}{dx^2} = A^\dagger(1 - A^\dagger)A^2.$$

This corresponds to $\beta = 0$ and $-\alpha_2 = \alpha_1 = 1$ and gives the dual

$$\begin{aligned} \hat{\mathcal{L}}f(n) &= \left(a^2(a^\dagger(1 - a^\dagger)) \right) f(n) \\ &= n(n-1)(f(n-1) - f(n)), \end{aligned}$$

which is the well-known Kingman's coalescent block-counting process.

2. **Wright Fisher diffusion with mutation.**

$$\mathcal{L} = x(1-x)\frac{d^2}{dx^2} + \theta(1-x)\frac{d}{dx} = A^\dagger(1-A^\dagger)A^2 + \theta(1-A^\dagger)A.$$

This corresponds to $\alpha_1 = -\alpha_2 = 1$, $\beta_0 = -\beta_1 = \theta$. This gives the dual

$$\begin{aligned}\hat{\mathcal{L}}f(n) &= \left(a^2(a^\dagger(1-a^\dagger)) + \theta a(1-a^\dagger)\right) f(n) \\ &= n(n-1)(f(n-1) - f(n)) + \theta n(f(n-1) - f(n)),\end{aligned}$$

which corresponds to Kingman's coalescent with extra rate θn to go down from n to $n-1$, due to mutation.

3. **Wright Fisher diffusion with “negative” selection.**

$$\mathcal{L} = x(1-x)\frac{d^2}{dx^2} - \sigma x(1-x)\frac{d}{dx} = A^\dagger(1-A^\dagger)(A^2 - \sigma A) \quad (33)$$

with $\sigma > 0$, which corresponds to $\alpha_1 = -\alpha_2 = 1$, $\beta_2 = -\beta_1 = \sigma$. The dual is

$$\begin{aligned}\hat{\mathcal{L}}f(n) &= \left((a^2 - \sigma a)a^\dagger(1-a^\dagger)\right) f(n) \\ &= n(n-1)(f(n-1) - f(n)) + \sigma n(f(n+1) - f(n))\end{aligned} \quad (34)$$

Notice that

$$\mathcal{L} = x(1-x)\frac{d^2}{dx^2} + \sigma x(1-x)\frac{d}{dx} = A^\dagger(1-A^\dagger)(A^2 - \sigma A) \quad (35)$$

with $\sigma > 0$ (i.e., “positive selection”) can also be dealt with. Indeed it is dual to the same process (34), but now with duality function $(1-x)^n$, coming from the representation $(1-x)$, $-d/dx = d/d(1-x)$ of the generators of the Heisenberg algebra. This in turn corresponds to the transformation $x \mapsto 1-x$.

4. **Stepping stone model.** This is an extension of the Wright-Fisher diffusion, modelling subpopulations of which the individuals have two types, and which evolve within each subpopulation as in the neutral Wright Fisher diffusion, and additionally, after reproduction a fraction of each subpopulation is exchanged with other subpopulations. These subpopulations are indexed by a countable set S . The variables $x_i \in [0, 1]$, $i \in S$ then represent the fraction of type 1 in the i^{th} subpopulation. The generator of this model is defined on smooth local functions (i.e., depending on a finite number of variables) on the set $\Omega = [0, 1]^S$ and given by

$$\mathcal{L} = \sum_{i,j \in S} p(i,j)(x_j - x_i) \left(\frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_j} \right) + \sum_{i \in S} x_i(1-x_i) \frac{\partial^2}{\partial x_i^2}. \quad (36)$$

Here $p(i, j) = p(j, i)$, with positive entries outside of the diagonal and with $\sum_{j \in S} p(i, j) = 1$, is an irreducible symmetric Markov transition kernel on the set S .

In terms of the standard creation and annihilation operators $A_i^\dagger = x_i$, $A_i = \frac{\partial}{\partial x_i}$, introduced (in the case of $m = |S|$ finite) at the end of Section 3.1, this generator reads

$$\mathcal{L} = \sum_{i, j \in S} p(i, j) (A_j^\dagger - A_i^\dagger) (A_i - A_j) + \sum_{i \in S} A_i^\dagger (1 - A_i^\dagger) A_i^2. \quad (37)$$

The dual operators a_i and a_i^\dagger have also been introduced in (17) and the duality function between tensor products of operators is (18). As a consequence the generator \mathcal{L} in (37) is dual to $\hat{\mathcal{L}}$ given by

$$\hat{\mathcal{L}} = \sum_{i, j \in S} p(i, j) (a_i - a_j) (a_j^\dagger - a_i^\dagger) + \sum_{i \in S} a_i^2 a_i^\dagger (1 - a_i^\dagger) \quad (38)$$

or equivalently the generator \mathcal{L} in (36) is dual to $\hat{\mathcal{L}}$ given by

$$\begin{aligned} \hat{\mathcal{L}} f(n) &= \sum_{i, j \in S} p(i, j) n_i (f(n - e_i + e_j) - f(n)) \\ &+ \sum_{i, j \in S} p(j, i) n_j (f(n + e_i - e_j) - f(n)) \\ &+ \sum_{i \in S} n_i (n_i - 1) (f(n_i - e_i) - f(n)), \end{aligned} \quad (39)$$

which is the generator of a Markov process on \mathbb{N}_0^S with transitions $n \rightarrow n - e_i + e_j$ (resp. $n \rightarrow n - e_j + e_i$) at rate $n_i p(i, j)$ (resp. $n_j p(j, i)$) and $n \rightarrow n - e_i$ at rate $n_i (n_i - 1)$. Here e_i denotes the vector with components $(e_i)_k = \delta_{i, k}$. The first type of transitions are of random walk type and correspond to the exchange of subpopulations, whereas the second type are the transitions corresponding to the Kingmans' coalescent in each subpopulation.

Finite-size populations [4, 14] and coalescents.

As a final example, we illustrate the use of the discrete creation and annihilation operators a_N^\dagger, a_N , corresponding to population models with N individuals in the discrete Moran model. This is the discrete analogue of the neutral Wright-Fisher diffusion

$$\mathcal{L}_N f(k) = \frac{N^2}{2} \frac{k}{N} \left(1 - \frac{k}{N}\right) (f(k+1) + f(k-1) - 2f(k)). \quad (40)$$

In terms of the discrete creation and annihilation operators a_N, a_N^\dagger defined in (23), this generator reads

$$\mathcal{L}_N = a_N^\dagger (1 - a_N^\dagger) a_N^2. \quad (41)$$

By theorem 2.1 and proposition 3.2, we therefore obtain immediately that this generator is dual to the generator of the Kingman's coalescent with duality function (24).

5 $SU(1, 1)$ algebra and corresponding dualities

In this section we show new dualities for models of population dynamics, using dualities between well-chosen differential operators and discrete operators. These operators have been used in the context of particle systems and models of heat conduction [15]. Interpreted here in terms of population models, they yield in that context new dualities.

In the whole of this section, the common domain \mathbf{D} of the operators that will appear (as section 2.1) will be the set of multivariate polynomials. This set is closed under the action of the operators, and forms a core of the Markov generators that will appear.

The results of this section are obtained applying the $SU(1, 1)$ algebra [23], which is an (abstract) algebra generated by a set of elements $\{\mathbf{K}_i^+, \mathbf{K}_i^-, \mathbf{K}_i^o\}$, $i = 1, \dots, n$ that satisfy the following commutation relations:

$$[\mathbf{K}_i^o, \mathbf{K}_i^\pm] = \pm \mathbf{K}_i^\pm, \quad [\mathbf{K}_i^-, \mathbf{K}_i^+] = 2\mathbf{K}_i^o. \quad (42)$$

We start with the following two families (labeled by m) of infinite dimensional representations of the algebra $SU(1, 1)$. The first family of operators acts on smooth functions $f : [0, \infty) \rightarrow \mathbb{R}$, whereas the second family acts on functions $f : \mathbb{N}_0 \rightarrow \mathbb{R}$.

$$\begin{aligned} \mathcal{H}^+ &= z, \\ \mathcal{H}^- &= z \frac{d^2}{dz^2} + \frac{m}{2} \frac{d}{dz}, \\ \mathcal{H}^0 &= z \frac{d}{dz} + \frac{m}{4}, \end{aligned} \quad (43)$$

and

$$\begin{aligned} K^+ f(n) &= \left(\frac{m}{2} + n\right) f(n+1), \\ K^- f(n) &= n f(n-1), \\ K^0 f(n) &= \left(\frac{m}{4} + n\right) f(n). \end{aligned} \quad (44)$$

The \mathcal{H} operators satisfy the $SU(1, 1)$ commutation relations (42) whereas the K operators satisfy the dual commutation relations (i.e., with opposite sign). Therefore, the operators are candidates for a duality relation (see item 4 of Theorem 2.1 and Remark 2.1).

In order to find corresponding duality functions, we now first give the analogue of Proposition 3.1 in the context of the $SU(1, 1)$ algebra. This

tells us that if \mathcal{K}^+ and \mathcal{K}^0 are dual to their discrete analogues given in (44) with duality function D , and $D(z, 0)$ is “annihilated” by \mathcal{K}^- (i.e., $\mathcal{K}^-D(z, 0) = 0$), then, using the commutation relations (42), we obtain that \mathcal{K}^- and K^- are also dual with the same duality function, and hence, by item 6 of Theorem 2.1, the whole algebra spanned by \mathcal{K}^α is dual to the algebra spanned by the K^α , $\alpha \in \{+, -, 0\}$.

PROPOSITION 5.1. *Suppose $D(z, n)$ are functions such that*

$$\begin{aligned}\mathcal{K}_1^+ D(z, n) &= \left(\frac{m}{2} + n\right) D(z, n+1) \\ \mathcal{K}_1^0 D(z, n) &= \left(n + \frac{m}{4}\right) D(z, n) \\ \mathcal{K}_1^- D(z, 0) &= 0\end{aligned}\tag{45}$$

where the \mathcal{K}^α , for $\alpha \in \{+, -, 0\}$, are working on the z -variable. Then we have $\mathcal{K}^\alpha \rightarrow^D K^\alpha$, where K^α are the discrete operators defined in (44).

PROOF. By assumption (45) we have $\mathcal{K}^\alpha \rightarrow^D K^\alpha$ for $\alpha \in \{+, 0\}$. Therefore we have to prove that $\mathcal{K}^- \rightarrow^D K^-$. In this proof we abuse notation and denote $\mathcal{K}^\alpha D(z, n) = \mathcal{K}_1^\alpha D(z, n)$. We start by proving that

$$\mathcal{K}^- D(z, 1) = D(z, 0)\tag{46}$$

Using (45) with $n = 1$

$$\begin{aligned}\mathcal{K}^- D(z, 1) &= \mathcal{K}^- \left(\frac{\mathcal{K}^+ D(z, 0)}{m/2}\right) \\ &= \frac{2}{m} \mathcal{K}^- \mathcal{K}^+ D(z, 0) \\ &= \frac{2}{m} (\mathcal{K}^+ \mathcal{K}^- + [\mathcal{K}^-, \mathcal{K}^+]) D(z, 0) \\ &= \frac{2}{m} (2\mathcal{K}^0) D(z, 0) \\ &= \frac{2}{m} \left(\frac{m}{2} D(z, 0)\right) = D(z, 0)\end{aligned}\tag{47}$$

Then, we proceed by induction. Assume $\mathcal{K}^- D(z, n-1) = (n-1)D(z, n-2)$.

Then

$$\begin{aligned}
\mathcal{K}^- D(z, n) &= \mathcal{K}^- \left(\frac{\mathcal{K}^+ D(z, n-1)}{\frac{m}{2} + n - 1} \right) \\
&= \frac{1}{\frac{m}{2} + n - 1} (\mathcal{K}^+ \mathcal{K}^- + [\mathcal{K}^-, \mathcal{K}^+]) D(z, n-1) \\
&= \frac{1}{\frac{m}{2} + n - 1} (\mathcal{K}^+ \mathcal{K}^- + 2\mathcal{K}^0) D(z, n-1) \\
&= \frac{1}{\frac{m}{2} + n - 1} \mathcal{K}^+ ((n-1)D(z, n-2)) + \\
&\quad + \frac{1}{\frac{m}{2} + n - 1} \left(2n - 2 + \frac{m}{2} \right) D(z, n-1) \\
&= \frac{D(z, n-1)}{\frac{m}{2} + n - 1} \left(\left(\frac{m}{2} + n - 2 \right) (n-1) + 2n - 2 + \frac{m}{2} \right) \\
&= nD(z, n-1) \tag{48}
\end{aligned}$$

Here in the third step we used the commutation relations, in the fourth step the induction hypothesis and (45), and in the fifth step (45). \square

To find the duality function $d : [0, \infty) \times \mathbb{N} \rightarrow \mathbb{R}$ relating the discrete and continuous representations (43) and (44) we use the previous proposition: first

$$\mathcal{K}_l^- d(z, 0) = \left(z \frac{d^2}{dz^2} + \frac{m}{2} \frac{d}{dz} \right) d(z, 0) = K_r^- d(z, 0) = 0$$

which gives as a possible choice $d(z, 0) = 1$. Then, we can act with \mathcal{K}^+ :

$$(\mathcal{K}_l^+)^n d(z, 0) = z^n = (K_r^+)^n d(z, 0) = \frac{m}{2} \left(\frac{m}{2} + 1 \right) \dots \left(\frac{m}{2} + n - 1 \right) d(z, n),$$

and we find

$$d(z, n) = \frac{z^n}{\frac{m}{2} \left(\frac{m}{2} + 1 \right) \dots \left(\frac{m}{2} + n - 1 \right)} = \frac{z^n \Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m}{2} + n\right)}. \tag{49}$$

Since $d(z, n)$ is of the form $c_n z^n$ we also see that

$$\mathcal{K}_l^0 d(z, n) = \left(z \frac{d}{dz} + \frac{m}{4} \right) d(z, n) = \left(n + \frac{m}{4} \right) d(z, n)$$

Then, by proposition 5.1 $\mathcal{K}_l^- d(z, n) = K_r^- d(z, n)$. We can then summarize these findings in the following result.

PROPOSITION 5.2. *The family of operators given by (43) and the family of operators given by (44) are dual with duality function given by (49). As a consequence, every element of the algebra generated by the operators (43) is dual to an element of the algebra generated by (44), obtained by replacing the operators by their duals and reverting the order of products.*

5.1 Markov generators constructed from $SU(1, 1)$ raising and lowering operators.

The relevance of the K^\pm, \mathcal{K}^\pm lies in the fact that some natural generators of diffusion processes of population dynamics can be rewritten in terms of them. As mentioned before (see beginning of section 4), these will be generators of processes on a multidimensional simplex with *absorbing boundary conditions*, i.e., the domain of the pregenerator L consists of smooth functions on the simplex such that Lf vanishes at the boundary. We start now with defining these generators.

DEFINITION 5.1 ([12], p. 55). *The d -types Wright-Fisher model with symmetric parent-independent mutation at rate $\theta \in \mathbb{R}$ is a diffusion process on the simplex $\sum_{i=1}^d x_i = 1$ defined by the generator*

$$\begin{aligned} \mathcal{L}_{d,\theta}^{WF} g(x) &= \sum_{i=1}^{d-1} \frac{1}{2} x_i (1 - x_i) \frac{\partial^2 g(x)}{\partial x_i^2} - \sum_{1 \leq i < j \leq d-1} x_i x_j \frac{\partial^2 g(x)}{\partial x_i \partial x_j} \\ &+ \frac{\theta}{d-1} \sum_{i=1}^{d-1} (1 - dx_i) \frac{\partial g(x)}{\partial x_i}. \end{aligned} \quad (50)$$

DEFINITION 5.2. *The Brownian Energy process with parameter $m \in \mathbb{R}$ on the complete graph with d vertices ($BEP(m)$) is a diffusion on \mathbb{R}_+^d with generator*

$$\begin{aligned} \mathcal{L}_d^{BEP(m)} f(y) &= \frac{1}{2} \sum_{1 \leq i < j \leq d} y_i y_j \left(\frac{\partial}{\partial y_i} - \frac{\partial}{\partial y_j} \right)^2 f(y) \\ &- \frac{m}{4} \sum_{1 \leq i < j \leq d} (y_i - y_j) \left(\frac{\partial}{\partial y_i} - \frac{\partial}{\partial y_j} \right) f(y). \end{aligned} \quad (51)$$

PROPOSITION 5.3. *The Brownian Energy process with parameter $m \in \mathbb{R}$ on the complete graph with d vertices and with initial condition $\sum_{i=1}^d x_i = 1$ coincides with the d -types Wright-Fisher model with symmetric parent-independent mutation at rate $\theta = \frac{m}{4}(d-1)$, i.e.*

$$\mathcal{L}_d^{BEP(m)} f(x_1, \dots, x_{d-1}, x_d) = \mathcal{L}_{d, \frac{m}{4}(d-1)}^{WF} g(x_1, \dots, x_{d-1})$$

with

$$g(x_1, \dots, x_{d-1}) = f(x_1, \dots, x_{d-1}, 1 - \sum_{j=1}^{d-1} x_j).$$

PROOF. The statement of the proposition is a consequence of the fact the BEP evolution conserves the quantity $x_1 + \dots + x_d$. Consider the initial condition $\sum_{i=1}^d x_i = 1$ and define the function $\phi : \mathbb{R}^{d-1} \rightarrow \mathbb{R}^d$ such that

$$(x_1, \dots, x_{d-1}) = x \mapsto \phi(x) = (x_1, \dots, x_{d-1}, 1 - \sum_{j=1}^{d-1} x_j).$$

Then $g(x) = f(\phi(x))$ and, for all $i = 1, \dots, d-1$, using the chain rule gives

$$\frac{\partial g(x)}{\partial x_i} = \frac{\partial f(\phi(x))}{\partial y_i} - \frac{\partial f(\phi(x))}{\partial y_d}.$$

A computation shows that

$$\mathcal{L}_d^{BEP(m)} f(x_1, \dots, x_{d-1}, x_d) = \mathcal{L}_{d, \frac{m}{4}(d-1)}^{WF} g(x_1, \dots, x_{d-1}).$$

□

DEFINITION 5.3 ([13], Eq. (12)). *In the d -types Moran model with population size N and with symmetric parent-independent mutation at rate θ , a pair of individuals of types i and j are sampled uniformly at random, one dies with probability $1/2$ and the other reproduces. In between reproduction events each individual accumulates mutations at a constant rate θ and his type mutates to any of the others with the same probability. Therefore, denoting types occurrences by $k = (k_1, \dots, k_{d-1})$, where k_i is the number of individuals of type i , the process has generator*

$$\begin{aligned} \mathcal{L}_{N,d,\theta}^{Mor} g(k) = & \frac{1}{2} \sum_{1 \leq i < j \leq d-1} \left[k_i \left(k_j + \frac{2\theta}{d-1} \right) (g(k - e_i + e_j) - g(k)) \right. \\ & \left. k_j \left(k_i + \frac{2\theta}{d-1} \right) (g(k + e_i - e_j) - g(k)) \right] \\ & + \frac{1}{2} \sum_{i=1}^{d-1} \left[\left(N - \sum_{j=1}^{d-1} k_j \right) \left(k_i + \frac{2\theta}{d-1} \right) (g(k + e_i) - g(k)) \right. \\ & \left. k_i \left(N - \sum_{j=1}^{d-1} k_j + \frac{2\theta}{d-1} \right) (g(k - e_i) - g(k)) \right]. \end{aligned} \quad (52)$$

DEFINITION 5.4. *The Symmetric Inclusion process with parameter $m \in \mathbb{R}$ on the complete graph with d vertices ($SIP(m)$) is a Markov process on \mathbb{N}_0^d with generator*

$$\begin{aligned} \mathcal{L}_d^{SIP(m)} f(k) = & \frac{1}{2} \sum_{1 \leq i < j \leq d} \left[k_i \left(k_j + \frac{m}{2} \right) (f(k - e_i + e_j) - f(k)) \right. \\ & \left. + k_j \left(k_i + \frac{m}{2} \right) (f(k + e_i - e_j) - f(k)) \right]. \end{aligned} \quad (53)$$

PROPOSITION 5.4. *The generator of the Symmetric Inclusion process with parameter $m \in \mathbb{R}$ on the complete graph with d vertices and with initial condition $\sum_{i=1}^d n_i = N$ coincides with the generator of the d -types Moran model with population size N and with symmetric parent-independent mutation at rate $\theta = \frac{m}{4}(d-1)$.*

PROOF. One verifies that

$$\mathcal{L}_d^{SIP(m)} f(k_1, \dots, k_{d-1}, k_d) = \mathcal{L}_{N,d,\frac{m}{4}(d-1)}^{Mor} g(k_1, \dots, k_{d-1})$$

with

$$g(k_1, \dots, k_{d-1}) = f(k_1, \dots, k_{d-1}, N - \sum_{j=1}^{d-1} k_j) .$$

□

We can now state our duality result.

THEOREM 5.1. *In the presence of symmetric parent-independent mutation at rate θ , the d -types Wright-Fisher diffusion process with generator (50) and the d -types Moran model with N individuals and with generator (52) are dual with duality function*

$$\tilde{D}_N(x, k) = \prod_{i=1}^d \frac{x_i^{k_i}}{\Gamma(\frac{2\theta}{d-1} + k_i)} , \quad (54)$$

with

$$x_d = 1 - \sum_{j=1}^{d-1} x_j , \quad k_d = N - \sum_{j=1}^{d-1} k_j .$$

PROOF. The statement of the theorem is a consequence of the duality between BEP(m) and SIP(m), which we now recall. We consider the two families of operators representing the $SU(1, 1)$ and dual $SU(1, 1)$ commutation relations, now rewritten in d coordinates:

$$\begin{cases} \mathcal{H}_{m,i}^+ = x_i \\ \mathcal{H}_{m,i}^- = x_i \frac{\partial^2}{\partial x_i^2} + \frac{m}{2} \frac{\partial}{\partial x_i} \\ \mathcal{H}_{m,i}^0 = x_i \frac{\partial}{\partial x_i} + \frac{m}{4} \end{cases} , \quad (55)$$

and the corresponding discrete operators

$$\begin{cases} K_{m,i}^+ f(k_i) = (k_i + \frac{m}{2} - 1) f(k_i - 1) \\ K_{m,i}^- f(k_i) = (k_i + 1) f(k_i + 1) \\ K_{m,i}^o f(k_i) = (k_i + \frac{m}{4}) f(k_i) . \end{cases} \quad (56)$$

The generator of the BEP(m) then reads

$$\mathcal{L}_m = \frac{1}{2} \sum_{1 \leq i < j \leq d} \left(\mathcal{K}_{m,i}^+ \mathcal{K}_{m,j}^- + \mathcal{K}_{m,i}^- \mathcal{K}_{m,j}^+ - 2\mathcal{K}_{m,i}^o \mathcal{K}_{m,j}^o + \frac{m^2}{8} \right), \quad (57)$$

By proposition 5.2, combined with theorem 2.1, we find that this operator is dual to the operator

$$L_m = \frac{1}{2} \sum_{1 \leq i < j \leq d} \left(K_{m,i}^+ K_{m,j}^- + K_{m,i}^- K_{m,j}^+ - 2K_{m,i}^o K_{m,j}^o + \frac{m^2}{8} \right), \quad (58)$$

This operator is exactly the generator of the SIP(m). The duality function is given by, using once more theorem 2.1, item 7:

$$D_N(x, k) = \prod_{i=1}^d d(x_i, k_i).$$

where $d(z, k)$ is given in (49). The multiplicative constant $\Gamma(m/2)$ in (49) can be dropped, and the result of the theorem thus follows from combining the duality between BEP(m) and SIP(m) with proposition 5.3 and proposition 5.4. \square

5.2 Limiting duality between d -types Wright-Fisher diffusion and d -types Moran model

We can now also let $m \rightarrow 0$, or correspondingly $\theta \rightarrow 0$ to obtain a duality result between the neutral Wright-Fisher diffusion and the standard Moran model.

Let $d \geq 2$ be an integer denoting the number of types (or alleles) in a population.

DEFINITION 5.5. *The d -types Wright-Fisher model is a diffusion process on the simplex $\sum_{i=1}^d x_i = 1$ defined by the generator*

$$\mathcal{L}_d^{WF} g(x) = \sum_{i=1}^{d-1} \frac{1}{2} x_i (1 - x_i) \frac{\partial^2 g(x)}{\partial x_i^2} - \sum_{1 \leq i < j \leq d-1} x_i x_j \frac{\partial^2 g(x)}{\partial x_i \partial x_j}. \quad (59)$$

DEFINITION 5.6. *The Brownian Energy process with $m = 0$ on the complete graph with d vertices is a diffusion on \mathbb{R}_+^d given by the generator*

$$\mathcal{L}_d^{BEP(0)} f(y) = \frac{1}{2} \sum_{1 \leq i < j \leq d} y_i y_j \left(\frac{\partial}{\partial y_i} - \frac{\partial}{\partial y_j} \right)^2 f(y). \quad (60)$$

PROPOSITION 5.5. *The generator of the Brownian Energy process with $m = 0$ on the complete graph with d vertices and with initial condition $\sum_{i=1}^d x_i = 1$ does coincide with the generator of the d -types Wright-Fisher diffusion.*

PROOF. Similar to the proof of proposition 5.3 \square

DEFINITION 5.7. *In the d -types Moran model with population size N a pair of individuals of types i and j are sampled uniformly at random, one dies with probability $1/2$ and the other reproduces. Therefore, denoting type occurrences by $k = (k_1, \dots, k_{d-1})$, where k_i is the number of individuals of type i , the process has generator*

$$\begin{aligned} \mathcal{L}_{N,d}^{Mor} g(k) &= \frac{1}{2} \sum_{1 \leq i < j \leq d-1} k_i k_j (g(k + e_i - e_j) + g(k - e_i + e_j) - 2g(k)) \\ &+ \frac{1}{2} \sum_{i=1}^{d-1} k_i \left(N - \sum_{j=1}^{d-1} k_j \right) (g(k + e_i) + g(k - e_i) - 2g(k)) . \end{aligned} \tag{61}$$

DEFINITION 5.8. *The Symmetric Inclusion process with $m = 0$ on the complete graph with d vertices is a Markov process on \mathbb{N}_0^d with generator*

$$\mathcal{L}_d^{SIP(0)} f(k) = \frac{1}{2} \sum_{1 \leq i < j \leq d} k_i k_j (f(k + e_i - e_j) + f(k - e_i + e_j) - 2f(k)) . \tag{62}$$

PROPOSITION 5.6. *The generator of the Symmetric Inclusion process with $m = 0$ on the complete graph with d vertices and with initial condition $\sum_{i=1}^d n_i = N$ does coincide with the generator of the d -types Moran model with population size N , i.e.*

$$\mathcal{L}_d^{SIP(0)} f(k_1, \dots, k_{d-1}, k_d) = \mathcal{L}_{N,d}^{Mor} g(k_1, \dots, k_{d-1})$$

with

$$g(k_1, \dots, k_{d-1}) = f(k_1, \dots, k_{d-1}, N - \sum_{j=1}^{d-1} k_j) .$$

PROOF. Similarly to the proof of proposition 5.4, the result follows from the conservation law, namely the fact that the SIP evolution conserves the total number of particles $k_1 + \dots + k_d$. \square

In the duality result of theorem 5.1 we cannot directly substitute $m = 0$ because there would be problems when some $k_i = 0$. To state a duality

result for $\theta = 0$, i.e., between the Wright Fisher diffusion and the Moran model without mutation, we start again from the duality between Brownian Energy process and Symmetric Inclusion process:

$$\begin{aligned} & \mathbb{E}_x^{BEP(m)} \left(\prod_{i=1}^d \frac{x_i(t)^{\xi_i}}{\frac{m}{2} \cdots (\frac{m}{2} + \xi_i - 1)} \right) \\ &= \mathbb{E}_\xi^{SIP(m)} \left(\prod_{i=1}^d \frac{x_i^{\xi_i(t)}}{\frac{m}{2} \cdots (\frac{m}{2} + \xi_i(t) - 1)} \right). \end{aligned} \quad (63)$$

Here the products in lhs and rhs are by definition equal to 1 when $\xi_i = 0, \xi_i(t) = 0$.

For $\xi \in \mathbb{N}_0^d$, denote $\mathcal{R}(\xi) = \#\{i \in \{1, \dots, d\} : \xi_i \geq 1\}$. Then we can rewrite (63) and obtain

$$\begin{aligned} & \mathbb{E}_x^{BEP(m)} \left(\prod_{i=1, \xi_i \geq 1}^d \frac{x_i(t)^{\xi_i}}{(\frac{m}{2} + 1) \cdots (\frac{m}{2} + \xi_i - 1)} \right) \\ &= \mathbb{E}_\xi^{SIP(m)} \left(\left(\frac{m}{2}\right)^{(\mathcal{R}(\xi) - \mathcal{R}(\xi(t)))} \prod_{i=1, \xi_i(t) \geq 1}^d \frac{x_i^{\xi_i(t)}}{(\frac{m}{2} + 1) \cdots (\frac{m}{2} + \xi_i(t) - 1)} \right), \end{aligned} \quad (64)$$

where now the denominators in the products in lhs and rhs are by definition equal to 1 when $\xi_i = 1, \xi_i(t) = 1$.

Now we are in the position to take the limit $m \rightarrow 0$ and we find

$$\begin{aligned} & \mathbb{E}_x^{BEP(0)} \left(\prod_{i=1, \xi_i \geq 1}^d \frac{x_i(t)^{\xi_i}}{(\xi_i - 1)!} \right) \\ &= \lim_{m \rightarrow 0} \mathbb{E}_\xi^{SIP(m)} \left(\left(\frac{m}{2}\right)^{(\mathcal{R}(\xi) - \mathcal{R}(\xi(t)))} \prod_{i: \xi_i(t) \geq 1} \frac{x_i^{\xi_i(t)}}{(\xi_i(t) - 1)!} \right). \end{aligned} \quad (65)$$

Notice that the lhs becomes zero as soon as one of the x_i is zero, which corresponds to the fact that for all i , $x_i = 0$ is an absorbing set in the diffusion. Corresponding to this, the rhs becomes zero as soon as one of the species disappears, i.e., as soon as $\mathcal{R}(\xi)$ decreases by one unit. Notice however that in the rhs we can not simply substitute $m = 0$ as we did in the lhs, since $\left(\frac{m}{2}\right)^{(\mathcal{R}(\xi) - \mathcal{R}(\xi(t)))}$ can be of order $(1/m)^k$ with $k > 0$ with correspondingly small probability. Therefore, in the rhs we do not exactly recover the $SIP(0)$, but have to keep m positive and take the limit after the expectation. We call (65) “a limiting duality relation with duality function”

$$D(\xi, x) = \left(\prod_{i=1, \xi_i \geq 1}^d \frac{x_i^{\xi_i}}{(\xi_i - 1)!} \right). \quad (66)$$

By the correspondence of $SIP(m)$ with the Moran model, and $BEP(m)$ with the Wright-Fisher diffusion, the limiting duality relation (65) can also be read as a limiting duality between Wright-Fisher without mutation and the Moran model in the limit of zero mutation.

REMARK 5.1. *We remark that the duality results in subsections 5.1 and 5.2 are of a different nature than the usual dualities between forward process and coalescent. Indeed, we have here duality between two “forward processes” (the Wright-Fisher diffusion and the Moran model), which cannot be obtained from “looking backwards in time”, the method by which moment-dualities with the coalescent are usually obtained. In our framework, the dualities with the coalescent correspond to a change of representation in the Heisenberg algebra, whereas the dualities between e.g. Wright-Fisher and Moran model arise from a change of representation in the $SU(1, 1)$ algebra.*

5.3 Limiting self-duality of the d -types Moran model

We can push further the $SU(1, 1)$ structure behind the Moran model and deduce self-duality of the process.

THEOREM 5.2. *The d -types Moran model with N individuals and with generator (52) is self-dual with duality function*

$$\bar{D}_N(k, \xi) = \prod_{i=1}^d \frac{k_i!}{(k_i - \xi_i)!} \frac{\Gamma\left(\frac{2\theta}{d-1}\right)}{\Gamma\left(\xi_i + \frac{2\theta}{d-1}\right)}, \quad (67)$$

where $k_d = N - \sum_{i=1}^{d-1} k_i$ and $\xi_d = N - \sum_{i=1}^{d-1} \xi_i$.

PROOF. The result follows from the self-duality property of the $SIP(m)$ process [17] and from proposition 5.4. \square

The limit $m \rightarrow 0$, or equivalently $\theta \rightarrow 0$, leads to a limiting self-duality relation, i.e., the $SIP(0)$ is dual to $SIP(m)$ in the limit $m \rightarrow 0$, and correspondingly, the Moran model with zero mutation has a limiting self-duality relation with the Moran model in the limit of zero mutation

$$\begin{aligned} & \mathbb{E}_\eta^{SIP(0)} \left(\prod_{i=1, \xi_i \geq 1}^d \frac{\eta_i(t)!}{(\eta_i(t) - \xi_i)! (\xi_i - 1)!} \right) \\ &= \lim_{m \rightarrow 0} \mathbb{E}_\xi^{SIP(m)} \left(\left(\frac{m}{2} \right)^{\mathcal{R}(\xi) - \mathcal{R}(\xi(t))} \prod_{i=1, \xi_i(t) \geq 1}^d \frac{\eta_i!}{(\eta_i - \xi_i(t))! (\xi_i(t) - 1)!} \right). \end{aligned} \quad (68)$$

5.4 Examples

Here we give a few simple illustrations of concrete computations using the dualities of the present section. Notice that these computations can also be performed using the coalescent in the context of d types. One can for example use the function-valued dual process of the Fleming-Viot process, discussed e.g. in Sect. 1.12 of [11] or in Sect. 2.8 of [6], or alternatively the embedding of the Kingman coalescent into the look-down construction [8, 2].

Our approach using the duality between Wright-Fisher and Moran is an alternative, and in our opinion slightly simpler one. In general, remark that the duality between Wright-Fisher and Moran implies that if we want to compute an expectation of a polynomial of degree k at time t in the multi-type Wright Fisher model, we have to consider a Moran model with k individuals. Also, if we want to compute the expectation of a polynomial of degree k in the number of individuals of different types in a Moran model with N individuals, we can do it by using only a Moran model with k individuals. So the main simplification coming from these dualities is the fact that we can go from “many” (N) to “few” (k) individuals (which can be useful in particular in simulations). The concrete computations that follow below are chosen somewhat arbitrarily as an illustration of those simplifications.

Before we start these computations, we remark that if in (68) or (65), we start with $\mathcal{R}(\xi) = d$ equal to its maximal value, then the non-zero contributions in the limit $m \rightarrow 0$ only come from configuration $\mathcal{R}(\xi_t) = d$ (since automatically $\mathcal{R}(\xi_t) \leq d$, so in that case there are no contributions for which $\mathcal{R}(\xi_t) > \mathcal{R}(\xi)$, i.e., with a negative exponent of m).

1. Heterozygosity of two-types Wright-Fisher diffusion. This is defined as the probability that two randomly chosen individuals are of different types ([12], pag. 48). To compute this quantity we can use the limiting duality between the $BEP(0)$ process $(x(t), y(t))$ on two sites, with initial condition (x, y) such that $x + y = 1$ and the process $SIP(m)$, $m \rightarrow 0$.

$$\begin{aligned} \mathbb{E}_{x,y}^{BEP(0)}(x(t)y(t)) &= \lim_{m \rightarrow 0} \mathbb{E}_{1,1}^{SIP(m)}(xyI(n_1(t) = 1, n_2(t) = 1)) \\ &= \mathbb{E}_{1,1}^{SIP(0)}(xyI(n_1(t) = 1, n_2(t) = 1)) \\ &= xy\mathbb{P}_{1,1}(n_1(t) = 1, n_2(t) = 1) = xye^{-t}, \end{aligned}$$

where $\mathbb{P}_{1,1}$ denotes the law of the $SIP(0)$ process initialized with one particle per site.

2. Higher moments of two-types Wright-Fisher diffusion. We use the same notation of the previous item and consider for instance x^2y . Further, we notice that if we start the $SIP(0)$ from initial configuration

$(n_1, n_2) = (2, 1)$, then configurations $(3, 0)$ and $(0, 3)$ are absorbing and the only transitions before absorption are of the type $(2, 1) \rightarrow (1, 2)$ and vice versa, and both transitions occur at rate 2, whereas from any of these states, the rate to go to the absorbing states is also equal to 2. Therefore,

$$\begin{aligned} & \mathbb{E}_{xy}^{BEP(0)}(x^2(t)y(t)) \\ &= x^2y\mathbb{P}_{2,1}^{SIP(0)}((n_1(t), n_2(t)) = (2, 1)) + \\ & \quad + xy^2\mathbb{P}_{2,1}^{SIP(0)}((n_1(t), n_2(t)) = (1, 2)) \\ &= \frac{e^{-2t}}{2}(x^2y(1 + e^{-2t}) + xy^2(1 - e^{-2t})). \end{aligned}$$

3. Analogue of heterozygosity for d -types Wright Fisher diffusion. Notice that for multitype Wright Fisher, there is no simple analogue of the Kingman's coalescent, as for the two-types case. This means that we have the $BEP(0)$ started from x_1, \dots, x_d

$$\begin{aligned} & \mathbb{E}_{x_1, \dots, x_d}^{BEP(0)}(x_1(t) \dots x_d(t)) \\ &= x_1 \dots x_d \mathbb{E}_{(1,1, \dots, 1)}^{SIP(0)}(I(n_i(t) \neq 0 \ \forall i \in \{1, \dots, d\})) \\ &= x_1 \dots x_d e^{-\frac{d(d-1)t}{2}}. \end{aligned}$$

4. Analogue of x^2y for the multi-type case.

$$\begin{aligned} & \mathbb{E}_{x_1, \dots, x_d}^{BEP(0)}(x_1^2(t)x_2 \dots x_d(t)) \\ &= \mathbb{E}_{(2,1, \dots, 1)}^{SIP(0)}\left(\left(\prod_i x_i^{n_i(t)}\right) I(n_i(t) \neq 0 \ \forall i \in \{1, \dots, d\})\right) \\ &= \sum_{i=1}^d \left(\prod_{j \neq i} x_j\right) x_i^2 \mathbb{P}_{(2,1, \dots, 1)}^{SIP(0)}(n_1(t) = 1, \dots, n_i(t) = 2, \dots, n_d(t) = 1). \end{aligned}$$

To compute the latter probability, we remark that starting from the configuration $(2, 1, \dots, 1)$, the $SIP(0)$ will reach an absorbing state as soon as one of the particles on the sites with a single occupation makes a jump, which happens at rate $(d-1)(d-2) + (d-1)2 = d(d-1)$. Further, as long as absorption did not occur, the site with two particles moves as a continuous-time random walk X_t^d on the complete graph of d vertices, moving at rate 2 and starting at site 1. Therefore

$$\begin{aligned} & \mathbb{P}_{(2,1, \dots, 1)}^{SIP(0)}(n_1(t) = 1, \dots, n_i(t) = 2, \dots, n_d(t) = 1) \\ &= e^{-d(d-1)t} \mathbb{P}(X_t^d = i) \\ &= e^{-2dt} + \frac{1}{d}(1 - e^{-2dt})\delta_{i,1} + (1 - \delta_{i,1})\frac{1}{d}(1 - e^{-2dt}). \end{aligned}$$

As these examples illustrate, computations of (appropriately chosen) moments in the multi-type Wright Fisher diffusion reduce to finite dimensional Markov chain computations, associated to inclusion walkers on the complete graph until absorption, which occurs as soon as a site becomes empty. The same can be done for the multi-type Moran model, using its self-duality.

Acknowledgments. We acknowledge financial support from the Italian Research Funding Agency (MIUR) through FIRB project “Stochastic processes in interacting particle systems: duality, metastability and their applications”, grant n. RBFR10N90W and the Fondazione Cassa di Risparmio Modena through the International Research 2010 project.

References

- [1] S. Athreya, J. Swart, Systems of branching, annihilating, and coalescing particles, preprint *arXiv:1203.6477* (2012).
- [2] J. Bertoin, J.F. Le Gall, Stochastic flows associated to coalescent processes, *Probab. Theor. Rel. Fields* **126**, 261–288 (2003).
- [3] A. Borodin, I. Corwin, T. Sasamoto, From duality to determinants for q -TASEP and ASEP preprint *arXiv:1207.5035* (2012).
- [4] C. Cannings, The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models, *Advances in Applied Probability* **6**, 2, 260–290 (1974).
- [5] G. Carinci, C. Giardinà, C. Giberti, F. Redig, Duality for stochastic models of transport. *Journal of Statistical Physics* **152**, 657–697 (2013).
- [6] D. Dawson, *Measure-value Markov processes*, proceedings of the 21st St Flour summer school, 1991, Springer LNM 1541, (1993).
- [7] A. Dimakis, F. Müller-Hoisse, T. Striker, Umbral Calculus, Discretization, and Quantum Mechanics on a Lattice, *Journal of Physics A* **29**, 6861-6876, (1996).
- [8] P. Donnelly and T.G. Kurtz, A countable representation of the Fleming-Viot measure-valued diffusion, *Ann. Probab.* **24**, 698–742 (1996); Particle representations for measure-valued population models, *Ann. Probab.* **27**, 166–205 (1999).
- [9] K. Engel, R. Nagel, *A Short Course on Operator Semigroups*, Universitext, Springer-Verlag, New York, Berlin, Heidelberg, 2006.
- [10] S.N. Ethier, T.M. Kurtz, *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics, (1986).

- [11] A.E. Etheridge, *An introduction to superprocesses*, University Lecture Series, 20, AMS, (2000)
- [12] A. Etheridge, *Some Mathematical Models from Population Genetics: École D'été de Probabilités De Saint-flour XXXIX-2009*, Springer Verlag (2011).
- [13] A.M. Etheridge, R.C. Griffiths, A coalescent dual process in a Moran model with genic selection, *Theor. Pop. Biol.* **75**, 320–330 (2009).
- [14] K. Gladstien, The characteristic values and vectors for a class of stochastic matrices arising in genetics, *SIAM Journal on Applied Mathematics*, **34**, 4, 630–642 (1978).
- [15] C. Giardinà, J. Kurchan, F. Redig, Duality and exact correlations for a model of heat conduction, *Journal of Mathematical Physics* **48**, 033301 (2007).
- [16] C. Giardinà, J. Kurchan, F. Redig, K. Vafayi, Duality and hidden symmetries in interacting particle systems, *Journal of Statistical Physics* **135**, 1, 25-55 (2009).
- [17] C. Giardinà, F. Redig, K. Vafayi, Correlation inequalities for interacting particle systems with duality, *Journal of Statistical Physics* **141**, 2, 242-263 (2009).
- [18] D. A. Dawson and A. Greven, Duality for spatially interacting Fleming-Viot processes with mutation and selection, preprint: arXiv:1104.1099 (2011).
- [19] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Springer, Graduate Texts in Mathematics 222, 2003.
- [20] E. Hillion, O. Johnson, Y. Yu, A natural derivative on $[0, n]$ and a binomial Poincaré inequality, preprint arXiv:1107.0127 (2011).
- [21] T. Huillet, A Duality Approach to the Genealogies of Discrete Nonneutral Wright-Fisher Models *Journal of Probability and Statistics*, 714701 (2009)
- [22] T. Huillet, On the Karlin-Kimura approaches to the Wright-Fisher diffusion with fluctuating selection *J. Stat. Mech.* P02016 (2011)
- [23] A. Inomata, H. Kuratsuji, C. Gerry, *Path integrals and coherent states of SU (2) and SU (1, 1)*, Singapore: World Scientific, 1992.
- [24] S. Jansen, N. Kurt, On the notion(s) of duality for Markov processes, preprint arXiv:1210.7193 (2012).

- [25] S. Jansen, N. Kurt, Pathwise construction of certain moment dualities and application to population models with balancing selection, preprint *arXiv:1207.6056* (2012).
- [26] Th. M. Liggett *Interacting particle systems*, Classic in Mathematics. Springer-Verlag, Berlin 2005.
- [27] M. Möhle, The concept of duality and applications to Markov processes arising in neutral population genetics models, *Bernoulli*, **5**, 5, 761–777, (1999).
- [28] M. Möhle, Total variation distances and rates of convergence for ancestral coalescent processes in exchangeable population models, *Adv. Appl. Probab.* **32**, 983993 (2000).
- [29] L. Mytnik, Uniqueness for a mutually catalytic branching model. *Probab. Theory Related Fields* **112** 245 – 253 (1998).
- [30] J. Ohkubo, Duality in interacting particle systems and boson representation, *J. Stat. Phys.* **139**, 454–465 (2010).
- [31] G. Schütz, S. Sandow, Non-Abelian symmetries of stochastic processes: Derivation of correlation functions for random-vertex models and disordered-interacting-particle systems, *Physical Review E* **49**, 2726 – 2741 (1994).
- [32] G. Schütz, Duality relations for asymmetric exclusion processes, *J. Stat. Phys.* **86**, 1265-1287, (1997).
- [33] H. Spohn, Stochastic integrability and the KPZ equation, preprint *arXiv:1204.2657v1* (2012)
- [34] A. Sudbury, P. Lloyd, Quantum operators in classical probability theory: II. The concept of duality in interacting particle systems, *Annals of Probability*, 1816–1830 (1995).
- [35] O. Taussky, H. Zassenhaus, On the similarity transformation between a matrix and its transpose. *Pacific J. Math.* **9**, Number 3, 893-896 (1959).
- [36] C. A. Tracy, H. Widom, Integral Formulas for the Asymmetric Simple Exclusion Process *Commun. Math. Phys.* **279**, 815–844 (2008). Erratum: *Commun. Math. Phys.* **304**, 875-878 (2011).