DUALITIES, CONES AND SPECTRAL DECOMPOSITIONS ARISING IN MATHEMATICAL POPULATION GENETICS

Martin Möhle, University of Tübingen

Workshop on Duality of Markov Processes, TU Berlin

November 7, 2014

Duality

$X = (X_t)_{t \in T}$ Markov process with state space $(E_1, \mathcal{F}_1)$

$Y = (Y_t)_{t \in T}$ Markov process with state space $(E_2, \mathcal{F}_2)$

$B(E) :=$ space of real-valued bounded measurable functions on $E := E_1 \times E_2$

Definition. (Duality, Liggett, 1985) The process X is said to be dual to Y w.r.t. $H \in B(E)$ if

$\boxed{\mathbb{E}^x H(X_t, y) = \mathbb{E}^y H(x, Y_t)}$ for all $(x, y) \in E$, $t \in T$.

Remarks.

- Dual processes occur in many applications, usually when considering some phenomena forwards and backwards in time.
- Duality is a powerful tool, for example in the physics literature on interacting particle systems and in mathematical population genetics.
- Duality can be expressed via semigroups and is essentially equivalent to duality of generators (Voss-Böhme, Schenk and Koellner (2011), Jansen and Kurt (2014)).

Duality space and cones

Definition. (Duality space) The set of all duality functions

$$U \ := \ U(X, Y) \ := \ \{H \in B(E) \ : \ X \text{ is dual to } Y \text{ w.r.t. } H\}$$

is called the duality space of X and Y.

Remark. U is a closed subspace of $B(E)$, hence, a Banach space. Depending on X and Y, U can be rather small or rather large. Typical questions: Basis of U?, Dimension of U?

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

Relation between Liggett duality and cones

Let

$$C_1 = C_1(H) = \{f : E_1 \to \mathbb{R} \,:\, f(x) = \int_{E_2} H(x,y)\, Q_2(\mathrm{d}y) \text{ for some } Q_2 \in \mathcal{M}(E_2)\}$$

$$C_2 = C_2(H) = \{g : E_2 \to \mathbb{R} \,:\, g(y) = \int_{E_1} H(x,y)\, Q_1(\mathrm{d}x) \text{ for some } Q_1 \in \mathcal{M}(E_1)\}$$

Proposition 1. If X is dual to Y w.r.t. H, then C_1 is a cone of X and C_2 is a cone of Y.

Duality in the sense of Liggett implies cone duality.

The following kind of converse of Proposition 1 holds under additional assumptions.

Relation between Liggett duality and cones (continued)

Proposition 2. Let $X = (X_t)_{t\in T}$ be a Markov process with state space $(E_1, \mathcal{F}_1)$. Assume that there exists $(E_2, \mathcal{F}_2)$, $C_1 \subseteq B(E_1)$, and $H \in B(E_1 \times E_2)$ such that

(i) $H(.,y) \in C_1$ for every $y \in E_2$.

(ii) C_1 is a cone of X.

(iii) C_1 has a unique integral representation over E_2 w.r.t. H, i.e., for every $f \in C_1$, there exists a unique probability measure Q_f on $(E_2, \mathcal{F}_2)$ such that $f = \int_{E_2} H(.,y)\, Q_f(\mathrm{d}y)$.

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 w.r.t. H. The process Y is unique in distribution.

If $(T_t)_{t\in T}$ denotes the semigroup of X, then Y has transition kernel $\mathbb{P}(Y_t \in B \,|\, Y_0 = y) = Q_{T_t H(.,y)}(B)$, $B \in \mathcal{F}_2$, $y \in E_2$.

Relation between Liggett duality and cones (continued)

Remarks.

1. If $E_2 \subseteq C_1$, then a typical duality function is $H(x,y) := y(x)$ (evaluation duality). E_2 is usually called the set of extremals. For this particular choice of H one arrives at the cone duality of Klebaner, Rösler and Sagitov (2007).

2. In general, E_2 is not necessarily a subset of C_1, not even of $B(E_1)$.

3. Typical other duality functions: $H(x,y) = x^y$ (moment duality), $H(x,y) = \exp(-xy)$ (Laplace duality)

4. Cone duality essentially deals with the question when a semigroup preserves a convex set C in some Banach space B. Related problems have been addressed in the functional analytic literature, see, for example, Brezis and Pazy (1970) and Ouhabaz (1996, 1999).

Stochastically monotone Markov chains, Sigmund 1976

Let $Y = (Y_n)_{n\in\mathbb{N}_0}$ be a Markov chain with state space $E_2 := \mathbb{N}_0$.

Proposition 3. If Y is stochastically monotone, i.e. $\mathbb{P}(Y_{n+1} \geq j \,|\, Y_n = i)$ is non-decreasing in i for every j, and if $\lim_{i\to\infty} \mathbb{P}(Y_{n+1} \geq j \,|\, Y_n = i) = 1$ for all j, then there exists a Markov chain $X = (X_n)_{n\in\mathbb{N}_0}$ such that X is dual to Y w.r.t. $H(i,j) := 1_{\{i\leq j\}}$.

Remarks.

- Duality relation: $\boxed{\mathbb{P}(X_{n+1} \leq k \,|\, X_n = i) = \mathbb{P}(Y_{n+1} \geq i \,|\, Y_n = k)}$.
- The cone C_1 consists of all non-negative, non-increasing functions f on $\mathbb{N}_0$ satisfying $f(0) = 1$ and $\lim_{i\to\infty} f(i) = 0$.
- See also Asmussen and Sigman (1996) and Sigman and Ryan (2000) for the continuous-time setting.

Brownian motion with reflection and with absorption

- Probably one of the oldest examples of duality (Lévy 1948, Breiman 1968)
- Continuous time analog of the previous example
- Brownian motion with reflection at 0 is dual to Brownian motion with absorption at 0 w.r.t. $H(x, y) := 1_{\{x \leq y\}}$.
- The cone C_1 is the set of all non-negative, non-increasing, left-continuous functions f on $[0, \infty)$ satisfying $f(0) = 1$ and $\lim_{x \to \infty} f(x) = 0$.
- The cone C_2 coincides with the set of all distribution functions $g : [0, \infty) \to \mathbb{R}$.
- C_1 and C_2 have both distance 1 from the origin and diameter 1.

Forward and backward process of Cannings models

- Discrete time population model with non-overlapping generations, fixed population size $N \in \mathbb{N}$ and exchangeable offspring mechanism. Examples: Wright–Fisher model, discrete Moran model
- Forward process $X = (X_n)_{n\in\mathbb{N}_0}$ counts the number of descendants forwards in time. Transition matrix $\Pi = (\pi_{ij})_{0\le i,j\le N}$
- Backward process $Y = (Y_n)_{n\in\mathbb{N}_0}$ counts the number of ancestors backward in time. Transition matrix $P = (p_{ij})_{0\le i,j\le N}$ is triangular.
- State space $S := \{0, \ldots, N\}$.
- X is dual to Y, for example w.r.t. $\boxed{H(i,j) := \binom{i}{j}/\binom{N}{j}}$, $i, j \in S$.
- Matrix notation: $\Pi H = HP'$, where P' denotes the transpose of P.
- Algebraic interpretation: Π can be transformed into a triangular matrix $P' = H^{-1}\Pi H$.

Forward and backward process of Cannings models (continued)

- 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 $i \in S$.
- The cone C_2 of Y consists of all functions $g : S \to \mathbb{R}$ satisfying $g(0) = 1$ such that $j \mapsto g(N-j)$ is absolutely monotone.
- Both, C_1 and C_2, have distance 1 from the origin and diameter 1.
- Duality space $U = U(X, Y)$ typically has dimension $N + 3$.

Wright–Fisher diffusion and block counting process of the Kingman coalescent

The Wright–Fisher diffusion $X = (X_t)_{t\geq 0}$ is a Markov process with state space $[0,1]$ 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 the block counting process of the Kingman coalescent (Markov chain, state space $\mathbb{N}_0$, infinitesimal rates $g_{i,i-1} := -g_{ii} := i(i-1)/2$ and $g_{ij} := 0$ otherwise).

Then $\boxed{X \text{ is dual to } Y}$ w.r.t. $H : [0,1] \times \mathbb{N}_0 \to \mathbb{R}$, $H(x,n) := x^n$ (moment duality).

The cone C_1 of X consists of all series $f : [0,1] \to \mathbb{R}$ of the form $f(x) = \sum_{n=0}^{\infty} a_n x^n$, where $a_n \geq 0$ and $\sum_n a_n = 1$ (isomorphic to the set of probability measures on $\mathbb{N}_0$).

The cone C_2 of Y consists of all functions $g : \mathbb{N}_0 \to \mathbb{R}$ of the form $g(n) = \int_{[0,1]} x^n Q(\mathrm{d}x)$ for some probability measure Q on $[0,1]$ (isomorphic to the set of all Hausdorff moment sequences $(g(n))_{n\in\mathbb{N}_0}$).

Again, C_1 and C_2 have distance 1 from the origin and diameter 1.

Fleming–Viot process (1979) and Kingman coalescent (1982)

Fleming–Viot process: measure valued Markov process $F = (F_t)_{t\geq 0}$, state space $\mathcal{M}(E) :=$ set of probability measures on some compact Polish space E.

The generator L of F acts on test functions $G_f(\mu) := \int_{E^n} f(x)\mu^n(\mathrm{d}x)$ via

$$LG_f(\mu) = \sum_{1\leq i<j\leq n} \int_{E^n} (f(x(i,j)) - f(x))\,\mu^n(\mathrm{d}x), \qquad f \in B(E^n), \mu \in \mathcal{M}(E),$$

where $x(i,j) \in E^n$ is obtained from $x = (x_1, \ldots, x_n) \in E^n$ by replacing x_j by x_i.

Kingman coalescent: partition valued Markov process $\Pi = (\Pi_t)_{t\geq 0}$, state space $\mathcal{P}$, the set of partitions of $\mathbb{N}$. If the process is in a state with b blocks than two blocks, chosen at random, merge together at rate 1.

$\mathcal{P}_n :=$ set of partitions of $\{1, \ldots, n\}$

$\varrho_n : \mathcal{P} \to \mathcal{P}_n$ (natural restriction from $\mathcal{P}$ to $\mathcal{P}_n$)

Often the (restricted) Kingman n-coalescent $(\varrho_n \circ \Pi_t)_{t\geq 0}$ is considered.

Fleming–Viot process and Kingman coalescent (continued)

For $h \in B(E^n)$ define the duality function $H_n : \mathcal{M}(E) \times \mathcal{P}_n \to \mathbb{R}$ via

$$\boxed{H_n(\mu, \xi) \ := \ \int_{E^n} h(x[\xi])\, \mu^n(\mathrm{d}x)},$$

where, for $\xi \in \mathcal{P}_n$ with blocks $B_1, \ldots, B_k$ and $x \in E^n$, $x[\xi] \in E^n$ has by definition entries $(x[\xi])_i := x_{\min B_j}$ if $i \in B_j$, $i \in \{1, \ldots, n\}$.

The Fleming–Viot process F is dual to the Kingman n-coalescent $(\varrho_n \circ \Pi_t)_{t \geq 0}$ w.r.t. H_n, i.e.

$$\boxed{\mathbb{E}^\mu H_n(F_t, \xi) \ = \ \mathbb{E}^\xi H_n(\mu, \varrho_n \circ \Pi_t)}, \qquad t \geq 0, \mu \in \mathcal{M}(E), \xi \in \mathcal{P}_n.$$

Fleming–Viot process and Kingman coalescent (continued)

The cone $C_1 = C_1(H_n)$ of F consists of all functions $f : \mathcal{M}(E) \to \mathbb{R}$ of the form

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

The cone $C_2 = C_2(H_n)$ of the Kingman n-coalescent consists of all functions $g : \mathcal{P}_n \to \mathbb{R}$ of the form

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

Questions. Simple representation for C_1 and C_2? Underlying Lie algebra?

Remark. The same duality relation holds between the Ξ-Fleming–Viot process F and the Ξ-coalescent $\Pi = (\Pi_t)_{t\geq 0}$ (exchangeable coalescent with simultaneous multiple collisions). The cones C_1 and C_2 are as above, since the duality function H_n does not depend on the measure Ξ.

From duality to spectral decompositions

- In the discrete (and haploid) setting duality usually transforms a forward transition matrix into a similar triangular backward transition matrix.
- In the algebraic language this corresponds to triangularisability.
- It is natural to go one step further and to ask for diagonalisability.
- This leads to spectral decompositions.
- Typical questions: Existence?, Formulas for the eigenvalues and for the eigenvectors (similarity transformation)?
- Existence is usually clear, for example, if the eigenvalues are distinct.
- For triangular matrices the eigenvectors can be obtained recursively (see, for example, Gladstien), explicit solutions are however not known in general.

Spectral decomposition of the block counting process of the Kingman coalescent

$Q = (q_{ij})_{i,j\in\mathbb{N}} :=$ generator of the block counting process of the Kingman coalescent.

$q_{i,i-1} = i(i-1)/2 = -q_{ii} =: q_i$ (total rate)

Proposition 4. (Spectral decomposition) $\boxed{Q = RDL}$, where $D = (d_{ij})_{i,j\in\mathbb{N}}$ is the diagonal matrix with entries $d_{ii} := -(i(i-1)/2$ and $R = (r_{ij})_{i,j\in\mathbb{N}}$ and $l = (l_{ij})_{i,j\in\mathbb{N}}$ are lower left triangular matrices with entries

$$r_{ij} = \prod_{k=j+1}^{i} \frac{q_k}{q_k - q_j} = \frac{i!(i-1)!(2j-1)!}{j!(j-1)!(i-j)!(i+j-1)!}, \qquad i \geq j,$$

and

$$l_{ij} = \prod_{k=j}^{i-1} \frac{q_{k+1}}{q_k - q_i} = (-1)^{i-j} \frac{(i-1)!i!(i+j-2)!}{(j-1)!j!(i-j)!(2i-2)!}, \qquad i \geq j.$$

Result known at least since 1984 (Tavaré).

Spectral decomposition of the Bolthausen–Sznitman coalescent

$Q = (q_{ij})_{i,j\in\mathbb{N}} :=$ generator of the block counting process of the B.–S. coalescent.
$q_{ij} = i/((i-j)(i-j+1))$ for $j < i$, $q_{ii} = 1-i$, $q_{ij} = 0$ otherwise.

Let $s(i,j)$ and $S(i,j)$ denote the Stirling numbers of the first and second kind respectively.

Theorem 1. (Spectral decomposition, M. and Pitters, 2014)
$\boxed{Q = RDL}$, where $D = (d_{ij})_{i,j\in\mathbb{N}}$ is the diagonal matrix with entries $d_{ii} := 1-i$ and $R = (r_{ij})_{i,j\in\mathbb{N}}$ and $l = (l_{ij})_{i,j\in\mathbb{N}}$ are lower left triangular matrices with entries

$$\boxed{r_{ij} = \frac{(j-1)!}{(i-1)!}|s(i,j)|} \quad \text{and} \quad \boxed{l_{ij} = \frac{(j-1)!}{(i-1)!}S(i,j)} \qquad i,j\in\mathbb{N}.$$

Remark. Proof based on generating functions.

Remark. Analog formulas for the spectral decomposition of the generator of the (partition valued) Kingman n-coalescent and for the Bolthausen–Sznitman n-coalescent are available (Kukla and Pitters, work in progress).

Hitting probabilities and absorption time of the Bolthausen–Sznitman coalescent

Corollary 1. (Hitting probability) The hitting probability $h(i,j)$ that the block counting process ever hits state j when started from state i is given by $h(i,1)=1$ and

$$h(i,j) \;=\; (j-1)(-1)^{i+j}\frac{(j-1)!}{(i-1)!}\sum_{k=j}^{i}\frac{s(i,k)S(k,j)}{k-1}, \qquad 2\le j\le i.$$

Corollary 2. (Absorption time)

The absorption time τ_n of the Bolthausen–Sznitman n-coalescent has distribution function

$$\mathbb{P}(\tau_n \le t) \;=\; \frac{\Gamma(n-e^{-t})}{\Gamma(n)\Gamma(1-e^{-t})}, \qquad t\in(0,\infty).$$

In particular, $\boxed{\tau_n - \log\log n \to \tau}$ in distribution as $n\to\infty$, where τ is Gumbel distributed.

Remark. The convergence to the Gumbel distribution can be alternatively verified via random recursive trees (Goldschmidt and Martin (2005)) or via the Chinese restaurant process.

The Mittag–Leffler process

Let η_t be Mittag–Leffler distributed with parameter e^{-t}, i.e. $\boxed{\mathbb{E}(\eta_t^m) = \frac{\Gamma(1+m)}{\Gamma(1+me^{-t})}}$.

$C_0(E) :=$ space of real-valued continuous functions on $E := [0, \infty)$ vanishing at infinity

Proposition 5. (M., 2014) $\boxed{T_t f(x) := \mathbb{E}(f(x^{e^{-t}}\eta_t))}$ defines a conservative, positive, strongly continuous contraction semigroup $(T_t)_{t\geq 0}$ on $C_0(E)$. The associated cadlag Markov process $X = (X_t)_{t\geq 0}$ is called the Mittag–Leffler process.

Remarks.

- Initial state $X_0 = 1$
- X_t is Mittag–Leffler distributed with parameter e^{-t}.
- Stationary distribution is the standard exponential distribution.

A scaling limit for the block counting process

$N_t^{(n)}$:= number of blocks of the Bolthausen–Sznitman n-coalescent at time t.

Define $\boxed{X_t^{(n)} := \frac{N_t^{(n)}}{n^{e^{-t}}}}$ and $\boxed{X^{(n)} := (X_t^{(n)})_{t\geq 0}}$ (scaled block counting process)

Note that $X^{(n)}$ is Markovian but time-inhomogeneous.

Theorem 2. (M., 2014) As $n \to \infty$, the scaled block counting process $X^{(n)}$ converges in $D_E[0,\infty)$ to the Mittag–Leffler process X.

Idea of Proof. Use spectral decomposition to show that

$$\mathbb{E}(N_t^{(n)}) = \frac{\Gamma(n+e^{-t})}{\Gamma(n)\Gamma(1+e^{-t})} \sim n^{e^{-t}}\mathbb{E}(X_t), \qquad n \to \infty.$$

Higher/joint moments are treated similarly. $\Rightarrow$ Convergence of the finite-dim. distributions.
Proof of the convergence in $D_E[0,\infty)$ is more tricky.

Remark. Baur and Bertoin (2014) independently obtained similar results via recursive trees.

The dual of the Mittag–Leffler process

Proposition 6. (Duality, M., 2014)

The Markov process $Y = (Y_t)_{t\geq 0}$ with state space $E = [0,\infty)$ and transition mechanism $\boxed{\mathbb{P}(Y_t \geq x \,|\, Y_0 = y) = \mathbb{P}(x^{e^{-t}}\eta_t \leq y)}$ is dual to the Mittag–Leffler process X w.r.t. $H(x,y) := 1_{\{x\leq y\}}$.

Proof 1. X is stochastically monotone. Result follows via Sigmund duality $\mathbb{P}(Y_t \geq x \,|\, Y_0 = y) = \mathbb{P}(X_t \leq y \,|\, X_0 = x)$.

Proof 2. Let C_1 denote the set of non-negative, non-increasing, left-continuous functions $f : E \to \mathbb{R}$ satisfying $f(0) = 1$ and $\lim_{x\to\infty} f(x) = 0$. Check that C_1 is a cone of X and that every $f \in C_1$ has a unique integral representation over E w.r.t. H. [Hint: Define the probability measure Q_f via $Q_f([x,\infty)) := f(x)$ for all $x \in E$.] Now apply Proposition 2.

Questions. Interpretation of the dual process Y? Properties of Y?

A dual formulation of Theorem 2

If η is Mittag–Leffler distributed with parameter $\alpha \in (0,1]$, then $\xi := \eta^{-1/\alpha}$ is α-stable with Laplace transform $\lambda \mapsto e^{-\lambda^\alpha}$.

Define $\boxed{Y_t^{(n)} := (X_t^{(n)})^{-e^t} = \frac{n}{(N_t^{(n)})^{e^t}}}$ and $\boxed{Y^{(n)} := (Y_t^{(n)})_{t\geq 0}}$.

Theorem 3. (M., 2014) As $n \to \infty$, the process $Y^{(n)}$ converges in $D_E[0,\infty)$ to the dual process Y. Note that Y_t is α-stable with Laplace transform $\lambda \mapsto e^{-\lambda^\alpha}$, where $\alpha := e^{-t}$.

Thank you very much for your attention!

References I

BAUR, E. AND BERTOIN, J. (2014) The fragmentation process of an infinite recursive tree and Ornstein-Uhlenbeck type processes. HAL preprint.

BOLTHAUSEN, E. AND SZNITMAN, A.-S. (1998) On Ruelle's probability cascades and an abstract cavity method. *Commun. Math. Phys.* **197**, 247–276.

BREZIS H. AND PAZY, A. (1970) Semigroups of nonlinear contractions on convex sets. *J. Func. Anal.* **6**, 237–281.

JANSEN, S. AND KURT, N. (2014) On the notion(s) of duality for Markov processes. *Probab. Surv.* **11**, 59–120.

KLEBANER, F.C., RÖSLER, U. AND SAGITOV, S. (2007) Transformations of Galton–Watson processes and linear fractional reproduction. *Adv. Appl. Probab.* **39**, 1036–1053.

LIGGETT, T. M. (1985) *Interacting Particles Systems*. Springer, Berlin.

MÖHLE, M. (1999) The concept of duality and applications to Markov processes arising in neutral population genetics modes. *Bernoulli* **5**, 761–777.

References II

MÖHLE, M. (2013) Duality and cones of Markov processes and their semigroups. *Markov Process. Related Fields* **19**, 149–162.

MÖHLE, M. (2014) The Mittag–Leffler process and a scaling limit for the block counting process of the Bolthausen–Sznitman coalescent. Preprint.

MÖHLE, M. AND PITTERS, H. (2014) A spectral decomposition for the block counting process of the Bolthausen–Sznitman coalescent. *Electron. Comm. Probab.* **19**, paper 47, 1–11.

OUHABAZ, E. (1996) Invariance of closed convex sets and domination criteria for semigroups. *Potential Anal.* **5**, 611–625.

OUHABAZ, E. (1999) L^p contraction semigroups for vector valued functions. *Positivity* **3**, 83–93.

VOSS-BÖHME, A., SCHENK, W. AND KOELLNER, A.-K. (2011) On the equivalence between Liggett duality of Markov processes and the duality relation between their generators. *Markov Process. Related Fields* **17**, 315–346.