J Stat Phys (2008) 132: 561–567
DOI 10.1007/s10955-008-9571-6

Global Spectral Gap for Dirichlet-Kac Random Motions

Gaston Giroux · René Ferland

Received: 4 March 2008 / Accepted: 13 May 2008 / Published online: 28 May 2008

Abstract We prove that the global spectral gap, for any Dirichlet-Kac random motion, is equal to the convergence rate of the limit motion.

Keywords Spectral gap · Dirichlet distribution · Boltzmann-like equations · Kac's model

1 Introduction

In this paper, we consider random motions on N-dimensional simplexes generated by Dirichlet interaction schemes on randomly selected coordinate pairs. We use [2, 3] to compute the spectral gaps for the underlying interacting kernels. We show that the global spectral gap is equal to the convergence rate of the limit motion coming from the random motions as N goes to infinity.

A special case of these random motions can be obtained from Kac's original model [9] by a simple change of variable. The spheres are then replaced by simplexes, and the uniform distribution by a Dirichlet distribution. In this way, Kac's model can be extended to a quite larger class, including bioinformatics models related to genetics [10, 11].

G. Giroux
410 Vimy, Suite 1, Sherbrooke, QC J1K 3M9, Canada
e-mail: gastongiroux2000@yahoo.ca

R. Ferland (✉)
Department of Mathematics, University of Quebec in Montreal, P.O. Box 8888, Downtown Station, Montreal, QC H3C 3P8, Canada
e-mail: ferland.rene@uqam.ca

2 The Dirichlet-Kac Random Motions

2.1 From Velocity to Energy

In a pioneer work, Mark Kac proposed a model for the time evolution of N particle velocities $v_1, v_2, \dots, v_N$ undergoing a random collision mechanism [9]. He was mainly interested in chaos propagation [13], which provides a formal derivation for Boltzmann-like equations.

In Kac's model, when a pair collision takes place, the two involved velocities v_i and v_j are transformed to

$$v_i^* = v_i \cos(\theta) + v_j \sin(\theta) \quad \text{and} \quad v_j^* = -v_i \sin(\theta) + v_j \cos(\theta), \qquad \theta \in (-\pi, \pi],$$

so the total energy is preserved locally (thus globally). Using ideas from [1, 7], one can see how this model leads to Dirichlet random motions. Assume for a moment that v_i and v_j are independent standard Gaussian variables. For any θ, v_i^* and v_j^* are also independent standard Gaussian. Consider the change of variable $x = v^2/2$ giving the energy. Then x_i and x_j are independent random variables with a Gamma distribution $\Gamma(\frac{1}{2}, 1)$. This is true for x_i^* and x_j^* as well and energy conservation gives $x_i^* + x_j^* = x_i + x_j$. Using the latter fact we may write

$$(x_i^*, x_j^*) = \left(\frac{x_i^*}{x_i^* + x_j^*}(x_i + x_j), \frac{x_j^*}{x_i^* + x_j^*}(x_i + x_j) \right) = (H(x_i + x_j), (1 - H)(x_i + x_j))$$

where the couple $(H, 1 - H)$ has a Dirichlet distribution $\mathrm{Dir}(\frac{1}{2}, \frac{1}{2})$. This is an example of what we call a *Dirichlet interaction scheme*.

Remark The (non-negative) random vector $(H_1, \dots, H_n)$ has a Dirichlet distribution $\mathrm{Dir}(\alpha_1, \dots, \alpha_n)$ if

(a) $H_n = 1 - \sum_{j=1}^{n-1} H_j$,
(b) the sub-vector $(H_1, \dots, H_{n-1})$ has a probability density function given by

$$p(y_1, \dots, y_{n-1}) = \frac{\Gamma(\sum_{j=1}^n \alpha_j)}{\prod_{j=1}^n \Gamma(\alpha_j)} \prod_{j=1}^{n-1} y_j^{\alpha_j - 1} \left(1 - \sum_{j=1}^{n-1} y_j \right)^{\alpha_n - 1}$$

on the set $D = \{(y_1, \dots, y_{n-1}) \in \mathbf{R}_+^{n-1} : \sum_{j=1}^{n-1} y_j \leq 1\}$.

See [8], Chapter 40, Sect. 5, for more details on the Dirichlet distribution and its connection with the univariate Gamma distribution.

2.2 The N-Particle System

Our purpose in this paper is to study random motions similar to Kac's model but for non negative vectors and a Dirichlet interaction scheme. So let us consider a system of N particles represented by a state vector $x = (x_1, \dots, x_N) \in \mathbf{R}_+^N$ and which evolves as a Markov chain similar to the one introduced by Kac. At each step of this Markov chain, x is updated due to the effect of a binary interaction between particles. These are randomly chosen, two at a time, at the random jump times of a Poisson process with intensity $N\lambda/2$, and are transformed by the Dirichlet scheme:

$$(x_i, x_j) \to (H(x_i + x_j), (1 - H)(x_i + x_j))$$

where $(H, 1 - H)$ has a Dirichlet distribution $\mathrm{Dir}(\alpha, \alpha)$ with $\alpha > 0$.

We write $X_k^N(n)$ for the value of the k-th component of the state vector after the n-th jump of the Poisson process, and set $X^N(n) = (X_1^N(n), \dots, X_N^N(n))$. The process $\{X^N(n), n \geq 0\}$ is a Markov chain with a transition kernel Q^N given by

$$\begin{aligned} &Q^N(x_1, \dots, x_N; A_1 \times \cdots \times A_N) \\ &\quad = \Pr\{X^N(n+1) \in A_1 \times \cdots \times A_N \big| X^N(n) = (x_1, \dots, x_N)\} \\ &\quad = \binom{N}{2}^{-1} \sum_{i<j} Q(x_i, x_j; A_i \times A_j)\delta^{i,j} \end{aligned}$$

where

$$Q(x, y; A \times B) = \Pr\{H(x+y) \in A, (1-H)(x+y) \in B\}$$

and

$$\delta^{i,j} = \begin{cases} 1, & \text{if } x_k \in A_k \text{ for all } k \neq i, j; \\ 0, & \text{otherwise.} \end{cases}$$

The *continuous-time* version Y^N of the Markov chain X^N is simply defined by

$$Y^N(t) = X^N(n), \quad T_n \leq t < T_{n+1}$$

with $0 = T_0 < T_1 < T_2 < \cdots$, the successive jump times of the Poisson process. The (transition) semi-group $\{G_t^N, t \geq 0\}$ of Y^N is given by

$$G_t^N = e^{(\frac{N\lambda}{2})(Q^N - I)t}.$$

The Markov chains X^N and Y^N are not irreducible but any simplex

$$B^N(a) = \{(x_1, \dots, x_N) \in \mathbf{R}_+^N : x_1 + x_2 + \cdots + x_N = a\}$$

is an irreducible class.

2.3 Stationary Distributions

A simple computation shows that, when Z_1 and Z_2 are independent random variables with common Gamma distribution $\Gamma(\alpha, \theta)$, the random variables $H(Z_1 + Z_2)$ and $(1 - H)(Z_1 + Z_2)$ are also independent with the same distribution. This means that the N-product of that Gamma distribution, say γ^N, is a stationary distribution for the Markov chain X^N (or Y^N). In addition, since Q^N preserves the sum, the conditional distribution $\gamma^N(\cdot \mid B^N(a))$ is stationary on $B^N(a)$. It is not difficult to find this conditional distribution. Indeed we have

$$\begin{aligned} (X_1^N(1), \dots, X_N^N(1)) &= \sum_{k=1}^N X_k^N(1) \left(\frac{X_1^N(1)}{\sum_{k=1}^N X_k^N(1)}, \dots, \frac{X_N^N(1)}{\sum_{k=1}^N X_k^N(1)} \right) \\ &= \sum_{k=1}^N X_k^N(0)(H_1, H_2, \dots, H_N) \end{aligned}$$

and the random vector $(H_1, H_2, \dots, H_N)$ has a multivariate Dirichlet distribution $\mathrm{Dir}(\alpha, \dots, \alpha)$. Therefore, conditioning on $X^N(0)$ being in $B^N(a)$, we see that $\gamma^N(\cdot \mid B^N(a))$ is the Dirichlet distribution on $B^N(a)$.

2.4 The Limit Motion

In studying the global spectral gap we have to look at Q^N as N goes to infinity. But it turns out that, as $N \to \infty$, the random motions have a limit motion. Ferland [4] showed in particular that if the $Y_k^N(0)$ are IID with a common distribution μ_0, having a finite first moment, then for all $t > 0$, the distribution of any component $Y_k^N(t)$ converges weakly to a probability measure μ_t, and $t \mapsto \mu_t$ is the unique solution of the Boltzmann-like equation

$$\dot{\mu}_t = \lambda(\mu_t \circ \mu_t - \mu_t)$$

where

$$\mu_t \circ \mu_t(C) = \int_0^\infty \int_0^\infty \Pr\{H(x+y) \in C\}\, \mu_t(dx)\mu_t(dy),$$

and $(H, 1-H)$ has a Dirichlet distribution $\mathrm{Dir}(\alpha, \alpha)$.

It is possible to define a non homogenous continuous-time Markov chain Y such that the distribution of $Y(t)$ is precisely μ_t. This chain is what we call the *limit motion*. Moreover, one can show (see [5]) that μ_t converges weakly, as $t \to \infty$, to the Gamma distribution $\Gamma(\alpha, \theta)$ with θ chosen to match the first moment of μ_0.

Theorem 1 *The limit motion converges to a Gamma distribution at geometric speed and the rate is* $\lambda\alpha/(2\alpha+1)$.

The proof is in Sect. 3.

2.5 Spectral Gaps

Observe that when we start the N-th motion with an arbitrary distribution u_0 on $B^N(a)$, symmetric or not, we have a geometric speed of convergence of $u_t = G_t^N u_0$ towards the stationary Dirichlet distribution (see [12], Chapter 1, Theorem 7.1). The *spectral gap* for $\lambda N(Q^N - I)/2$ is

$$\Delta_N = \left(\frac{N\lambda}{2}\right) \rho_1(N)$$

and gives the L^2 rate of convergence of u_t (see [2], page 5). It is proportional to the gap $\rho_1(N)$ that exists in the spectrum of Q^N when the latter is viewed as a self-contraction operator acting on the square-integrable functions on $B^N(a)$. It does not depend on a since Q^N commutes with the change of scale that relates the different simplexes. But more can be said: the limit

$$\Delta = \lim_N \left(\frac{N\lambda}{2}\right) \rho_1(N)$$

is strictly greater than 0, and the sequence of motions have a *global spectral gap*.

Theorem 2 *The global spectral gap is equal to the convergence rate of the limit motion*:

$$\Delta = \lambda\left(\frac{\alpha}{2\alpha+1}\right).$$

3 Proofs

Proof of Theorem 1 Theorem 1 is proved in [6] for the case $\lambda = 1$ and an interaction scheme of the form

$$(x, y) \to ((1 - Z_1)x + H(Z_1x + Z_2y), (1 - Z_2)y + (1 - H)(Z_1x + Z_2y))$$

where $(H, 1 - H)$ has a Dir(α, α) distribution and Z_1, Z_2 are independent random variables with a common Beta(α, β) distribution. The proof uses an explicit formula for μ_t known as Wild's sum [14]:

$$\mu_t = \sum_{n\geq 1} e^{-t}(1 - e^{-t})^{n-1}\mu^{(n)}$$

where $\{\mu^{(n)}, n \geq 1\}$ is given by the recursive formula:

$$\mu^{(n+1)} = \frac{1}{n}\sum_{j=1}^{n} \mu^{(j)} \circ \mu^{(n+1-j)}, \quad \mu^{(1)} = \mu_0.$$

The basic idea is to build a μ_t-distributed random variable using $\mu^{(n)}$-distributed ordered binary trees, and then apply an inductive argument on the left and right subtrees. The rate of convergence obtained is

$$\eta = (1 - (E[H_1^2] + E[H_2^2]))$$

where $(H_1, H_2) = ((1 - Z_1) + Z_1H, Z_2H)$. When $\lambda > 0$, the Wild's sum becomes

$$\mu_t = \sum_{n\geq 1} e^{-\lambda t}(1 - e^{-\lambda t})^{n-1}\mu^{(n)}$$

and the rate is just multiplied by λ. But in fact, looking at the proof more closely, one realizes it applies to the limit case $Z_1 = Z_2 = 1$ as well, yielding the convergence rate

$$\eta = \lambda(1 - 2E[H^2]) = \lambda\left(1 - 2\left(\frac{\alpha + 1}{2(2\alpha + 1)}\right)\right) = \lambda\left(\frac{\alpha}{2\alpha + 1}\right).$$ □

Proof of Theorem 2 The Dirichlet-Kac random motion is an example of what is called a *Kac system* in [2]. The general results of [2] (Theorem 2.1, Theorem 2.2 and Corollary 2.3) are available at hand. However to compute the spectral gaps, we need to prove an analogue of Theorem 3.1. For this we have to find the spectrum of an operator acting on the single-particle space.

Let μ^N be the Dirichlet distribution on the simplex $B^N(1)$, and ν^N be its 1-marginal; ν^N is a beta distribution with parameters $(\alpha, (N - 1)\alpha)$. For $g \in L^2([0, 1], \nu^{N-1})$ we define

$$K^N g(y) = \int_0^1 g((1 - y)w)\nu^{N-1}(dw).$$

If g is a polynomial of the form

$$g(z) = a_0 + a_1z + \cdots + a_kz^k$$

then

$$K^N g(y) = a_0 + a_1(1-y)m_1^{N-1} + \cdots + a_k(1-y)^k m_k^{N-1}$$

with m_j^{N-1} the j-th moment of ν^{N-1}. This implies that the eigenvectors of K^N are polynomials. There is exactly one such eigenvector for each degree k, and the corresponding eigenvalue is $\alpha_k = (-1)^k m_k^{N-1}$. The moment m_k^{N-1} is easy to compute:

$$\begin{aligned} m_k^{N-1} &= \int_0^1 w^k \frac{\Gamma((N-1)\alpha)}{\Gamma(\alpha)\Gamma((N-2)\alpha)} w^{\alpha-1}(1-w)^{(N-2)\alpha-1}\,dw \\ &= \frac{\Gamma(\alpha+k)\Gamma((N-1)\alpha)}{\Gamma(\alpha)\Gamma((N-1)\alpha+k)} \\ &= \frac{(\alpha+k-1)(\alpha+k-2)\cdots\alpha}{((N-1)\alpha+k-1)((N-1)\alpha+k-2)\cdots(N-1)\alpha}. \end{aligned}$$

Since, for $N \geq 3$, we have

$$m_2^{N-1} = \frac{\alpha+1}{((N-1)\alpha+1)(N-1)} > \frac{1}{(N-1)^2} = \frac{1}{N-1} m_1^{N-1},$$

and the m_k^{N-1}'s are strictly decreasing in k, we may conclude, exactly as in [2], that

$$\begin{aligned} \Delta_N &= \prod_{j=3}^{N} (1 - m_2^{j-1})\Delta_2 \\ &= \prod_{j=3}^{N} \left(\frac{(j-2)(j\alpha+1)}{(j-1)((j-1)\alpha+1)} \right) \Delta_2 \\ &= \left(\frac{1}{2\alpha+1} \right) \left(\frac{N\alpha+1}{N-1} \right) \Delta_2. \end{aligned}$$

Therefore

$$\Delta = \lim_N \left(\frac{1}{2\alpha+1} \right) \left(\frac{N\alpha+1}{N-1} \right) \Delta_2 = \lambda \left(\frac{\alpha}{2\alpha+1} \right),$$

because, trivially, $\Delta_2 = \lambda$. Theorem 2 is proved. □

Remarks (1) Recently in [3], more elaborate results were obtained for Kac's 3 dimensional models. For our model we are in the happy circumstance where the interacting kernel Q^N has a single gap eigenfunction f which is also the gap eigenfunction of the projection operator P^N for all N, giving an identity relating Δ_N and Δ_{N-1}.

(2) In our first deduction we use Feature 4 of [2]; but now it is a little quicker to use Lemma 3.1 of [3] which incorporates in its proof Feature 4. The form of the global rate Δ indicates that there is an identity relating Δ_N and Δ_{N-1}.

References

1. Barnsley, M., Turchetti, G.: A study of Boltzmann energy equations. Ann. Phys. **159**, 1–61 (1985). doi:10.1016/0003-4916(85)90191-5
2. Carlen, E., Carvalho, M.C., Loss, M.: Determination of the spectral gap in Kac's master equation and related stochastic evolutions. Acta Math. **191**, 1–54 (2003). doi:10.1007/BF02392695

3. Carlen, E., Geronimo, J., Loss, M.: Determination of the spectral gap in the Kac model for physical momentum and energy conserving collisions. Available at http://arxiv.org/abs/0705.3729v3 (2007)
4. Ferland, R.: Law of large numbers for pairwise interacting particle systems. Math. Models Methods Appl. Sci. **4**, 1–15 (1994). doi:10.1142/S0218202594000029
5. Ferland, R., Giroux, G.: Cutoff-type Boltzmann equations: convergence of the solution. Adv. Appl. Math. **8**, 98–107 (1987). doi:10.1016/0196-8858(87)90007-8
6. Ferland, R., Giroux, G.: An exponential rate of convergence for a class of Boltzmann processes. Stoch. Stoch. Rep. **35**, 79–91 (1991). doi:10.1080/17442509108833691
7. Hoare, M.H.: Quadratic transport and soluble Boltzmann equation. Adv. Chem. Phys. **56**, 1–140 (1984)
8. Johnson, N.L., Kotz, S.: Distributions in Statistics: Continuous Multivariate Distributions. Wiley, New York (1972)
9. Kac, M.: Foundations of kinetic theory. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III, pp. 171–197. University of California Press, Berkeley (1956)
10. Niu, T., Qin, Z.S., Xu, X., Liu, J.S.: Bayesian haplotype inference for multiple linked single-nucleotide polymorphisms. Am. J. Hum. Genet. **70**, 157–169 (2002). doi:10.1086/338446
11. Matsui, T., Motoki, M., Kamatani, N.: Polynomial time approximation sampler for discretized Dirichlet distribution. In: Ibaraki, T., Katoh, N., Ono, H. (eds.) Algorithms and Computation: 14th International Symposium, ISAAC Proceedings, Kyoto, Japan, December 15–17, 2003, pp. 676–685. Springer, Berlin (2003)
12. Orey, S.: Lecture Notes on Limit Theorems for Markov Chain Transition Probabilities. Van Nostrand Reinhold Co., London/New York/Toronto (1971)
13. Tanaka, H.: Propagation of chaos for certain purely discontinuous Markov processes with interaction. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **17**, 253–272 (1970)
14. Wild, E.: On Boltzmann's equation in the kinetic theory of gases. Proc. Camb. Phil. Soc. **47**, 602–609 (1951)