About the limit distribution of the Diaphony created by Mehlers kernel

Peter Jez

Technical Report 2013-01 January 2013

Department of Computer Sciences

Jakob-Haringer-Straße 2
5020 Salzburg
Austria
www.cosy.sbg.ac.at

Technical Report Series

About the limit distribution of the Diaphony created by Mehler's kernel

Peter Jez; Department of Computer Sciences,University of Salzburg
email: peter.jez@aon.at

January 25, 2013

Abstract

In the following we investigate the limit distribution of the Diaphony created by the Mehler kernel. The classical Diaphony was introduced by Zinterhof [5]. In [6] a Diaphony has been defined for reproducing kernel Hilbert spaces over an abstract set E. The limit distribution of the classical Diaphony has been investigated by H. Leeb [3].

1 Introduction

For $\lambda = (\lambda_1, \ldots, \lambda_s) \in \mathbb{R}^s$, $\nu = (n_1, \ldots, n_s) \in \mathbb{N}_0^s$ and $k \in \mathbb{N}$ we use the notation

$$\lambda^{\nu} := \lambda_1^{n_1} \ldots \lambda_s^{n_s}$$

$$\lambda^{k\nu} := \lambda_1^{kn_1} \ldots \lambda_s^{kn_s}$$

and by $\varphi_\nu(x)$ we denote the normalized Hermite polynomials given by

$$\varphi_\nu(x) := \prod_{j=1}^{s} \frac{H_{n_j}(x_j)}{\sqrt{2^{n_j} n_j! \sqrt{\pi}}}$$

where $H_j(x_j)$ denotes the Hermite polynomial with degree j. For a given $\lambda \in \mathbb{R}^s$ with $|\lambda_j| < 1$ the Mehler kernel is given by (see [4])

$$K : \mathbb{R}^s \times \mathbb{R}^s \to \mathbb{R}$$

$$K(x, y) = \sum_{\nu \in \mathbb{N}_0^s} \varphi_\nu(x) \varphi_\nu(y) \lambda^{2\nu} =$$

$$= \frac{1}{\sqrt{\pi}^s \sqrt{\prod_{j=1}^{s} (1 - \lambda_j^4)}} \times$$

$$\times \exp\left(\frac{1}{2} \sum_{j=1}^{s} \frac{\lambda_j^2}{1 + \lambda_j^2} (x_j + y_j)^2 - \frac{1}{2} \sum_{j=1}^{s} \frac{\lambda_j^2}{1 - \lambda_j^2} (x_j - y_j)^2 \right).$$

The space

$$H_\lambda := \left\{ f(x) : \mathbb{R}^s \to \mathbb{R} \,\middle|\, f(x) = \sum_{\nu \in \mathbb{N}_0^s} a_\nu \varphi_\nu(x) \lambda^\nu, \sum_{\nu \in \mathbb{N}_0^s} |a_\nu|^2 < \infty \right\}$$

equipped with the scalar product defined by

$$\langle \lambda^\mu \varphi_\mu, \lambda^\nu \varphi_\nu \rangle = \delta_{\mu\nu} \tag{1}$$

forms a reproducing kernel Hilbert space with kernel $K(x,y)$. The Diaphony is given by

$$D_N(x_n) := \left(\frac{1}{N^2} \sum_{n,k=1}^{N} K(x_n, x_k) - \frac{1}{\sqrt{\pi}^s} \right)^{\frac{1}{2}} =$$

$$= \left(\frac{1}{N^2} \sum_{n,k=1}^{N} \left(K(x_n, x_k) - \frac{1}{\sqrt{\pi}^s} \right) \right)^{\frac{1}{2}} =$$

$$= \left(\frac{1}{N^2} \sum_{n,k=1}^{N} \sum_{\nu \in \mathbb{N}_0^s \setminus (0,\ldots,0)} \varphi_\nu(x_n) \varphi_\nu(x_k) \lambda^{2\nu} \right)^{\frac{1}{2}} =$$

$$= \left(\sum_{\nu \in \mathbb{N}_0^s \setminus (0,\ldots,0)} \lambda^{2\nu} \left(\frac{1}{N} \sum_{n=1}^{N} \varphi_\nu(x_n) \right)^2 \right)^{\frac{1}{2}}$$

which appears in the estimation

$$\left| \frac{1}{N} \sum_{k=1}^{N} f(x_k) - \frac{1}{\sqrt{\pi}^s} \int_{\mathbb{R}^s} f(x) e^{-|x|^2} dx \right| \le D_N(x_n) \,.\, \|f\|$$

for $f(x) \in H_\lambda$ and $\|.\|$ denotes the norm induced by the scalar product (1).

2 The limit distribution of $ND_N^2(x_n)$

In the following let $E_s := diag(1, \ldots, 1) \in \mathbb{R}^{s \times s}$. We investigate the limit distribution of $ND_N^2(x_n)$ for $x_1, \ldots, x_N$ independent identically distributed $N\left(0, \frac{1}{\sqrt{2}} E_s\right)$ random variables. Let

$$I_M := \{(\nu_1, \ldots, \nu_s) \in \mathbb{N}_0^s, 0 \le \nu_j \le M, j = 1, \ldots, s\}.$$

We set

$$Y_N^{(M)} := N \sum_{\nu \in I_M \setminus (0,\ldots,0)} \lambda^{2\nu} \left(\frac{1}{N} \sum_{n=1}^{N} \varphi_\nu(x_n) \right)^2 =$$

$$= \sum_{\nu \in I_M \setminus (0,\dots,0)} \lambda^{2\nu} \left(\frac{1}{\sqrt{N}} \sum_{n=1}^{N} \varphi_\nu (x_n) \right)^2 .$$

Lemma 1. *1. We have*

$$E\left(\left| N D_N^2 (x_n) - Y_N^{(M)} \right| \right) \to 0$$

for $M \to \infty$ uniformly with respect to N.

2. Let $h(x) : \mathbb{R} \to \mathbb{R}$ be a bounded and Lipschitz continuous function. Then we have for all given $\varepsilon > 0$ an index $M_0(\varepsilon, h)$ (independent from N) with the property that

$$\left| E\left(h\left(N D_N^2 \right) \right) - E\left(h\left(Y_N^{(M)} \right) \right) \right| < \varepsilon$$

for all $M > M_0(\varepsilon, h)$.

Proof. ad 1) We have

$$\left| N D_N^2 (x_n) - Y_N^{(M)} \right| = \left| \sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu} \left(\frac{1}{\sqrt{N}} \sum_{n=1}^{N} \varphi_\nu (x_n) \right)^2 \right| =$$

$$= \sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu} \left(\frac{1}{\sqrt{N}} \sum_{n=1}^{N} \varphi_\nu (x_n) \right)^2 = N D_N^2 (x_n) - Y_N^{(M)}$$

and therefor

$$E\left(\left| N D_N^2 (x_n) - Y_N^{(M)} \right| \right) =$$

$$= E\left(N D_N^2 (x_n) - Y_N^{(M)} \right) = E\left(\sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu} \left(\frac{1}{\sqrt{N}} \sum_{n=1}^{N} \varphi_\nu (x_n) \right)^2 \right) =$$

$$= E\left(\sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu} \left(\frac{1}{N} \sum_{n=1}^{N} \varphi_\nu^2 (x_n) + \frac{1}{N} \sum_{n,k=1; n \neq k} \varphi_\nu (x_n) \varphi_\nu (x_k) \right) \right) =$$

$$= \sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu} N \frac{1}{N} E\left(\varphi_\nu^2 (x) \right) = \frac{1}{\sqrt{\pi}^s} \sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu}$$

by orthogonality of the φ_ν. For all $\varepsilon > 0$ there is an index $M(\varepsilon)$ (of course independent from N) with the property that

$$\frac{1}{\sqrt{\pi}^s} \sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu} < \varepsilon$$

for all $M > M(\varepsilon)$ by geometric series and therefor

$$E\left(\left| N D_N^2 (x_n) - Y_N^{(M)} \right| \right) \to 0$$

for $M \to \infty$ uniformly with respect to N.
ad 2) Let $h(x) : \mathbb{R} \to \mathbb{R}$ be a bounded and Lipschitz continuous function. Then we have

$$\left|E\left(h\left(ND_N^2\right)\right) - E\left(h\left(Y_N^{(M)}\right)\right)\right| = \left|E\left(h\left(ND_N^2\right) - h\left(Y_N^{(M)}\right)\right)\right| \leq$$

$$\leq E\left(\left|h\left(ND_N^2\right) - h\left(Y_N^{(M)}\right)\right|\right).$$

$h(x)$ is Lipschitz continuous and therefor a constant $L(h)$ exists with the property

$$|h(x) - h(y)| \leq L(h) \, . \, |x - y|$$

for all $x, y \in \mathbb{R}$. So we get

$$\left|E\left(h\left(ND_N^2\right)\right) - E\left(h\left(Y_N^{(M)}\right)\right)\right| \leq L(h) \, .E\left(\left|ND_N^2 - Y_N^{(M)}\right|\right).$$

From part 1 of the Lemma there is an index $M_0(\varepsilon)$ independent from N with the property that

$$E\left(\left|ND_N^2 - Y_N^{(M)}\right|\right) < \varepsilon$$

for $M > M_0(\varepsilon)$. □

Lemma 2. *Let*

$$Y^{(M)} = \sum_{\nu \in I_M \setminus (0,\ldots,0)} \lambda^{2\nu} X_\nu^2$$

where the random variables X_ν are independent and identically distributed $N\left(0, \pi^{-\frac{1}{4}}\right)$ random variables. Then we have

$$Y_N^{(M)} \to_d Y^{(M)}$$

for $N \to \infty$.

Proof. We consider the vector $v_N \in \mathbb{R}^{(M+1)^s - 1}$ defined by

$$v_N := \left(\frac{1}{\sqrt{N}} \sum_{n=1}^{N} \varphi_\nu(x_n)\right)_{\nu \in I_M \setminus (0,\ldots,0)}.$$

By multidimensional central limit theorem [1] this vector converges in distribution to

$$v := (X_\nu)_{\nu \in I_{M \setminus (0,\ldots,0)}} \sim N\left(0, \pi^{-\frac{1}{4}} E_{(M+1)^s - 1}\right)$$

because of the orthogonality of the Hermite polynomials. If a vector is normal distributed with a regular covariance matrix it has a density function and if the covariance matrix is a diagonal matrix this density function is separable. Because of this separability the components of the vector are independent random variables. The quadratic form $A(z) : \mathbb{R}^{(M+1)^s - 1} \to \mathbb{R}$ defined by

$$A(z) = z^T A z$$

with $A = diag\left(\lambda^{2\nu}\right)_{\nu \in I_M \setminus (0,\dots,0)} \in \mathbb{R}^{((M+1)^s-1)\times((M+1)^s-1)}$ is of course continuous and by continuous mapping theorem [1] we have

$$Y_N^{(M)} = A\left(v_N\right) \to_d A\left(v\right) = Y^{(M)}.$$

□

Lemma 3. *Let*

$$Y := \sum_{\nu \in \mathbb{N}_0^s \setminus (0,\dots,0)} \lambda^{2\nu} X_\nu^2$$

with i.i.d $X_\nu \sim N\left(0, \pi^{-\frac{1}{4}}\right)$.

1. *Then we have*

   $$E\left(\left|Y - Y^{(M)}\right|\right) \to 0$$

   for $M \to \infty$.

2. *Let* $h(x) : \mathbb{R} \to \mathbb{R}$ *be a bounded and Lipschitz continuous function. Then we have for all given* $\varepsilon > 0$ *an index* $M_0(\varepsilon, h)$ *with the property*

   $$\left|E\left(h\left(Y^{(M)}\right)\right) - E\left(h\left(Y\right)\right)\right| < \varepsilon$$

   for all $M > M_0(\varepsilon, h)$.

Proof. ad 1) We have

$$\left|Y - Y^{(M)}\right| = \left|\sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu} X_\nu^2\right| = \sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu} X_\nu^2$$

and therefor

$$E\left(\left|Y - Y^{(M)}\right|\right) = \frac{1}{\sqrt{\pi}^s} \sum_{\nu \in \mathbb{N}_0^s \setminus I_M} \lambda^{2\nu} \to 0$$

for $M \to \infty$. Part 2 is proved in the same way as in Lemma (1). □

Theorem 1. *Let* $x_k \sim N\left(0, \frac{1}{\sqrt{2}} E_s\right)$ *for* $k = 1, \dots, N$ *i.i.d RV. Then we have*

$$N D_N^2\left(x_n\right) \to_d Y$$

for $N \to \infty$.

Proof. If we have

$$E\left(h\left(X_n\right)\right) \to E\left(h\left(X\right)\right)$$

for $n \to \infty$ for all real valued, bounded and Lipschitz continuous functions h then we have $X_n \to_d X$ by Portmanteau Theorem (see [2]). Let $\varepsilon > 0$ be arbitrary and $h(x)$ be an arbitrary bounded and Lipschitz continuous function. According to Lemma (1), part 2 there is an index $M_0(\varepsilon, h)$ with the property

$$\left|E\left(h\left(ND_N^2(x_n)\right)\right) - E\left(h\left(Y_N^{(M)}\right)\right)\right| < \varepsilon$$

for all $M > M_0(\varepsilon, h)$. According to Lemma (3) there is an index $M_1(\varepsilon, h)$ with the property

$$\left|E\left(h\left(Y^{(M)}\right)\right) - E\left(h\left(Y\right)\right)\right| < \varepsilon$$

for all $M > M_1(\varepsilon, h)$. Now let $M > \max\left(M_0(\varepsilon, h), M_1(\varepsilon, h)\right)$. Then we have

$$\left|E\left(h\left(ND_N^2(x_n)\right)\right) - E\left(h\left(Y\right)\right)\right| \leq$$

$$\leq \left|E\left(h\left(ND_N^2(x_n)\right)\right) - E\left(h\left(Y_N^{(M)}\right)\right)\right| + \left|E\left(h\left(Y_N^{(M)}\right)\right) - E\left(h\left(Y^{(M)}\right)\right)\right| +$$

$$+ \left|E\left(h\left(Y^{(M)}\right)\right) - E\left(h\left(Y\right)\right)\right| < 2\varepsilon + \left|E\left(h\left(Y_N^{(M)}\right)\right) - E\left(h\left(Y^{(M)}\right)\right)\right|.$$

By Lemma (2) we have weak convergence of $Y_N^{(M)}$ to $Y^{(M)}$ so we can find an index $N(\varepsilon, M, h)$ with the property that

$$\left|E\left(h\left(Y_N^{(M)}\right)\right) - E\left(h\left(Y^{(M)}\right)\right)\right| < \varepsilon$$

for all $N > N(\varepsilon, M, h)$ and summing up we have

$$\left|E\left(h\left(ND_N^2(x_n)\right)\right) - E\left(h\left(Y\right)\right)\right| < 3\varepsilon.$$

□

References

[1] A. DasGupta. *Asymptotic Theory of Statistics and Probability*. Springer Texts in Statistics. Springer-Verlag, 2008.

[2] A. Klenke. *Probability Theory*. Springer-Verlag, 2006.

[3] H. Leeb. The Asymptotic Distribution of Diaphony in One Dimension, 1996.

[4] S. Thangavelu. *Hermite and Laguerre expansions*. Mathematical Notes 42. Princeton University Press, Princeton, 1993.

[5] P. Zinterhof. Über einige Abschätzungen bei der Approximation von Funktionen mit Gleichverteilungsmethoden. *Sitzungsber. Österr. Akad. Wiss. Math.-Natur. Kl. II*, 185:121–132, 1976.

[6] P. Zinterhof and C. Amstler. Uniform distribution, Discrepancy and reproducing kernel Hilbert spaces. *Journal of Complexity*, 17:497–515, 2001.