A note on the Ehrhard inequality*

Rafał Latała (Warsaw)

Abstract

We prove that for $\lambda \in [0,1]$ and A,B two Borel sets in R^n with A convex the following inequality holds true

$$\Phi^{-1}(\gamma_n(\lambda A + (1-\lambda)B)) \geq \lambda\Phi^{-1}(\gamma_n(A)) + (1-\lambda)\Phi^{-1}(\gamma_n(B)),$$

where γ_n is a canonical gaussian measure in R^n and Φ^{-1} is the inverse of gaussian distribuant.

Introduction. Let γ_n be a canonical gaussian measure in R^n i.e. the measure with density

$$\gamma_n(dx) = (2\pi)^{-n/2}\exp(-\frac{|x|^2}{2})dx$$

and let

$$\Phi(x) = \gamma_1((-\infty, x)) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}dy \qquad \text{for } x \in R.$$

A.Ehrhard proved in [1] the following Brunn-Minkowski like inequality for convex Borel sets A,B in R^n:

$$\Phi^{-1}(\gamma_n(\lambda A + (1-\lambda)B)) \geq \lambda\Phi^{-1}(\gamma_n(A)) + (1-\lambda)\Phi^{-1}(\gamma_n(B)).$$

This is an important result which has found numerous applications in the theory of gaussian processes and elsewhere.

It is still an open problem if this result remains true if we assume only that A and B are Borel sets. In the book of M.Ledoux and M.Talagrand [2] it is listed as Problem 1.

In this paper we generalize the result of Ehrhard to the case, when one of the sets A and B is convex.

Let us start with the following lemma:

* *1991 Mathematics Subject Classification* 60E15

Lemma 1 *Let $A = (a, b)$ be a finite open interval and numbers $\Gamma_B, \lambda \in (0, 1)$ be given. Then there exists an interval (c,d) with $\gamma_1((c, d)) = \Gamma_B$ such that for each finite sum of intervals B with $\gamma_1(B) = \Gamma_B$ we have*

$$\gamma_1(\lambda A + (1-\lambda)(c, d)) \le \gamma_1(\lambda A + (1-\lambda)B).$$

Proof: Let a positive integer n be fixed. We will look for the minimum of $\gamma_1(\lambda A + (1-\lambda)B)$ over all sets B - sums of at most n open intervals (some of them may be infinite) with the fixed gaussian measure $\gamma_1(B) = \Gamma_B$. It is easy to notice that the minimum is taken for some set

$$B_0 = (c_1, d_1) \cup \ldots \cup (c_k, d_k), \quad c_1 < d_1 < c_2 < \ldots < d_k, k \le n.$$

We are to show that k=1. Let us assume that $k > 1$, we have

$$\lambda A + (1-\lambda)B_0 = \bigcup_{i=1}^{k} (\lambda a + (1-\lambda)c_i, \lambda b + (1-\lambda)d_i).$$

If for some $i < k$ $[\lambda a + (1-\lambda)c_i, \lambda b + (1-\lambda)d_i] \cap [\lambda a + (1-\lambda)c_{i+1}, \lambda b + (1-\lambda)d_{i+1}] \ne \emptyset$ then we can find $\bar{c_1} > c_1$ and $\bar{d_i} > d_i$ such that for $\bar{B} = (\bar{c_1}, d_1) \cup \ldots \cup (c_i, \bar{d_i}) \cup \ldots \cup (c_k, d_k)$ we have $\gamma_1(\bar{B}) = \gamma_1(B_0)$ and $\gamma_1(\lambda A + (1-\lambda)\bar{B}) < \gamma_1(\lambda A + (1-\lambda)B_0)$, but this contradicts the minimality of B_0. Hence the intervals $[\lambda a + (1-\lambda)c_i, \lambda b + (1-\lambda)d_i]$ $i = 1, \ldots, k$ are disjoint.

If $c_1 = -\infty$ we define the function $\varphi(\varepsilon)$ for $\varepsilon > 0$ such that

$$\gamma_1((-\infty, d_1)) = \gamma_1((\varphi(\varepsilon), d_1 + \varepsilon)) \tag{1}$$

and let $B_\varepsilon = (\varphi(\varepsilon), d_1 + \varepsilon) \cup \bigcup_{i=2}^{k} (c_i, d_i)$ then for small enough ε $\gamma_1(B_\varepsilon) = \gamma_1(B_0)$ so

$$\gamma_1(\lambda A + (1-\lambda)B_\varepsilon) \ge \gamma_1(\lambda A + (1-\lambda)B_0). \tag{2}$$

From (1) we obtain

$$\Phi(d_1 + \varepsilon) - \Phi(d_1) = \Phi(\varphi(\varepsilon))$$

and from (2)

$$\Phi(\lambda b + (1-\lambda)(d_1 + \varepsilon)) - \Phi(\lambda b + (1-\lambda)d_1) \ge \Phi(\lambda a + (1-\lambda)\varphi(\varepsilon)),$$

but that is a contradiction since $\lim_{\varepsilon\to 0}\frac{\Phi(\lambda a+(1-\lambda)\varphi(\varepsilon))}{\Phi(\varphi(\varepsilon))}=\infty$ and $lim_{\varepsilon\to 0}\frac{\Phi(\lambda b+(1-\lambda)(d_1+\varepsilon))-\Phi(\lambda b+(1-\lambda)d_1)}{\Phi(\varphi(\varepsilon))}=\frac{\Phi(\lambda b+(1-\lambda)(d_1+\varepsilon))-\Phi(\lambda b+(1-\lambda)d_1)}{\Phi(d_1+\varepsilon)-\Phi(d_1)}<\infty.$ So $c_1>-\infty$ and analogously $d_k<\infty$.

Now let us define for small enough $|\varepsilon|$ the function $\varphi(\varepsilon)$ by the condition

$$\gamma_1((c_1+\varepsilon,d_1)\cup(c_2+\varphi(\varepsilon),d_2))=\gamma_1((c_1,d_1)\cup(c_2,d_2)).$$

It means that

$$\Phi(c_1+\varepsilon)+\Phi(c_2+\varphi(\varepsilon))=\Phi(c_1)+\Phi(c_2)$$

so $\varphi(0)=0$ and

$$\varphi^{'}(0)=-\exp(\frac{c_2^2}{2}-\frac{c_1^2}{2}). \tag{3}$$

Let $B_\varepsilon=(c_1+\varepsilon,d_1)\cup(c_2+\varphi(\varepsilon),d_2)\cup\bigcup_{i=3}^{k}(c_i,d_i)$ and $\psi(\varepsilon)=\gamma_1(\lambda A+(1-\lambda)B_\varepsilon)$. By definition of $\varphi(\varepsilon)$ we have $\gamma_1(B_\varepsilon)=\gamma_1(B_0)$ hence $\psi(\varepsilon)\geq\psi(0)$ and $\psi^{'}(0)=0$. Since

$$\begin{aligned}\psi(\varepsilon) &= \psi(0)+\Phi(\lambda a+(1-\lambda)c_1)+\Phi(\lambda a+(1-\lambda)c_2)+\\ &\quad -\Phi(\lambda a+(1-\lambda)(c_1+\varepsilon))-\Phi(\lambda a+(1-\lambda)(c_2+\varphi(\varepsilon)))\end{aligned}$$

we obtain from (3)

$$\psi^{'}(0)=\frac{1-\lambda}{\sqrt{2\pi}}[-\exp(-\frac{(\lambda a+(1-\lambda)c_1)^2}{2})+\exp(-\frac{(\lambda a+(1-\lambda)c_2)^2}{2}+\frac{c_2^2}{2}-\frac{c_1^2}{2})]$$

so since $\psi^{'}(0)=0$

$$-\frac{(\lambda a+(1-\lambda)c_1)^2}{2}=-\frac{(\lambda a+(1-\lambda)c_2)^2}{2}+\frac{c_2^2}{2}-\frac{c_1^2}{2}.$$

Therefore

$$(c_1+c_2)(2-\lambda)=2(1-\lambda)a$$

and since $c_2>c_1$

$$(2-\lambda)c_2>(1-\lambda)a. \tag{4}$$

In the same way we prove that

$$(2-\lambda)d_1<(1-\lambda)b. \tag{5}$$

Finally we find for small enough $|\varepsilon|$ the function $\varphi(\varepsilon)$ such that

$$\gamma_1((c_1, d_1+\varepsilon) \cup (c_2+\varphi(\varepsilon), d_2)) = \gamma_1((c_1, d_1) \cup (c_2, d_2))$$

that is

$$\Phi(d_1+\varepsilon) - \Phi(c_2+\varphi(\varepsilon)) = \Phi(d_1) - \Phi(c_2)$$

so $\varphi(0) = 0$ and

$$\varphi'(\varepsilon) = \exp(\frac{(c_2+\varphi(\varepsilon))^2}{2} - \frac{(d_1+\varepsilon)^2}{2}). \tag{6}$$

For $B_\varepsilon = (c_1, d_1+\varepsilon) \cup (c_2+\varphi(\varepsilon), d_2) \cup \bigcup_{i=3}^k (c_i, d_i)$ and $\psi(\varepsilon) = \gamma_1(\lambda A + (1-\lambda)B_\varepsilon)$ we have $\gamma_1(B_\varepsilon) = \gamma_1(B_0)$ hence $\psi(\varepsilon) \geq \psi(0)$ so $\psi'(0) = 0$ and $\psi''(0) \geq 0$. Since

$$\begin{aligned}\psi(\varepsilon) &= \psi(0) - \Phi(\lambda b + (1-\lambda)d_1) + \Phi(\lambda a + (1-\lambda)c_2) + \\ &\quad + \Phi(\lambda b + (1-\lambda)(d_1+\varepsilon)) - \Phi(\lambda a + (1-\lambda)(c_2+\varphi(\varepsilon)))\end{aligned}$$

we deduce from (6)

$$\begin{aligned}\psi'(\varepsilon) = \tfrac{1-\lambda}{\sqrt{2\pi}} \quad &[\exp(-\tfrac{(\lambda b+(1-\lambda)(d_1+\varepsilon))^2}{2}) + \\ &-\exp(-\tfrac{(\lambda a+(1-\lambda)(c_2+\varphi(\varepsilon)))^2}{2} + \tfrac{(c_2+\varphi(\varepsilon))^2}{2} - \tfrac{(d_1+\varepsilon)^2}{2})]\end{aligned} \tag{7}$$

and

$$\begin{aligned}\psi''(0) = \tfrac{1-\lambda}{\sqrt{2\pi}} \quad &\Big[-(1-\lambda)(\lambda b + (1-\lambda)d_1)\exp(-\tfrac{(\lambda b+(1-\lambda)d_1)^2}{2}) + \\ &-\Big(-(1-\lambda)(\lambda a + (1-\lambda)c_2)\exp(\tfrac{c_2^2}{2} - \tfrac{d_1^2}{2}) + c_2\exp(\tfrac{c_2^2}{2} - \tfrac{d_1^2}{2}) - d_1\Big) \cdot \\ &\cdot \exp(-\tfrac{(\lambda a+(1-\lambda)c_2)^2}{2} + \tfrac{c_2^2}{2} - \tfrac{d_1^2}{2})\Big].\end{aligned} \tag{8}$$

Since $\psi'(0) = 0$ by (7) we have

$$\exp(-\frac{(\lambda b + (1-\lambda)d_1)^2}{2}) = \exp(-\frac{(\lambda a + (1-\lambda)c_2)^2}{2} + \frac{c_2^2}{2} - \frac{d_1^2}{2}),$$

so from $\psi''(0) \geq 0$ and (8) we obtain that

$$\exp(\frac{d_1^2}{2})((1-(1-\lambda)^2)d_1 - \lambda(1-\lambda)b) \geq \exp(\frac{c_2^2}{2})((1-(1-\lambda)^2)c_2 - \lambda(1-\lambda)a) \tag{9}$$

But by (4) the right-hand-side of (9) is positive and by (5) the left-hand-side is negative. This contradiction shows that $k = 1$ and the proof of Lemma is completed.

Corollary 1 *If A=(a,b) and B is a Borel set in R then for* $\lambda \in (0,1)$ *the following inequality holds*

$$\Phi^{-1}(\gamma_1(\lambda A + (1-\lambda)B)) \geq \lambda\Phi^{-1}(\gamma_1(A)) + (1-\lambda)\Phi^{-1}(\gamma_1(B)) \qquad (10)$$

Proof: By simple approximation arguments it is enough to show (10) when B is a finite sum of open intervals.Then Lemma 1 reduces this case to the situation when B is an interval. Therefore (10) holds by the result of Ehrhard.

Theorem 1 *If A and B are Borel sets in* R^n *and A is convex then for* $\lambda \in (0,1)$ *the following inequality holds*

$$\Phi^{-1}(\gamma_n(\lambda A + (1-\lambda)B)) \geq \lambda\Phi^{-1}(\gamma_n(A)) + (1-\lambda)\Phi^{-1}(\gamma_n(B)) \qquad (11)$$

Proof: We use the same method as Ehrhard in the proof of Théorème 3.1. in [1]. We refer to this paper for the definitions of gaussian k-symmetrizations.

For $n = 1$ the theorem follows from Collolary 1.

Now let n=2 and f be an arbitrary 1-symmetrization in R^2. Then one can easily deduce from the already established case n=1 that

$$\lambda f[A] + (1-\lambda)f[B] \subset f[\lambda A + (1-\lambda)B]. \qquad (12)$$

Let us assume that (11) is false that is

$$\lambda\Phi^{-1}(\gamma_2(A)) + (1-\lambda)\Phi^{-1}(\gamma_2(B)) - \Phi^{-1}(\gamma_2(\lambda A + (1-\lambda)B)) = \varepsilon > 0 \quad (13)$$

then since symmetrizations doesn't change gaussian measure by (12) we have

$$\lambda\Phi^{-1}(\gamma_2(f[A])) + (1-\lambda)\Phi^{-1}(\gamma_2(f[B])) - \Phi^{-1}(\gamma_2(\lambda f[A] + (1-\lambda)f[B])) \geq \varepsilon. \quad (14)$$

By Théorème 3.1. of [1] f[A] is again convex set, therefore we can inductively prove (14) for each finite composition f of 1-symmetrizations in R^2. But by Théorème 1.6. of [1] we can choose a sequence f_j of compositions of 1-symmetrizations such that

$$\lim_{j\to\infty}\left(\lambda\Phi^{-1}(\gamma_2(f_j[A])) + (1-\lambda)\Phi^{-1}(\gamma_2(f_j[B])) - \Phi^{-1}(\gamma_2(\lambda f_j[A] + (1-\lambda)f_j[B]))\right) = 0.$$

this contradiction shows that Theorem holds for n=2.

Finally let $n \geq 3$. Then as above we prove (12) for arbitrary 2-symmetrization in R^n. So if we assume (13) we derive (14) (with γ_n instead of γ_2) for f a composition of 2-symmetrization. But each n-symmetrization in R^n is a composition of some 2-symmetrizations (Corollaire 2.3. in [1]) and for n-symmetrization f we obviously have

$$\lambda\Phi^{-1}(\gamma_n(f[A])) + (1-\lambda)\Phi^{-1}(\gamma_n(f[B])) - \Phi^{-1}(\gamma_n(\lambda f[A] + (1-\lambda)f[B])) = 0.$$

This contradicts (14) and completes the proof.

References

[1] A.Ehrhard, *Symétrisation dans l'espace de Gauss,* Math.Scand. 53 (1983), 281-301

[2] M.Ledoux, M.Talagrand *Probability in Banach spaces,* Springer Verlag 1991

Institute of Mathematics,Warsaw Univercity,
ul.Banacha 2,02-097 Warszawa,
Poland
email: rlatala@mimuw.edu.pl