Characterization of equality in the correlation inequality for convex functions, the U-conjecture

Gilles Hargé

Equipe d’analyse et de probabilités, université d’Evry, rue du père Jarlan, bât. Maupertuis, 91025 Evry Cedex, France

Received 16 June 2003; received in revised form 28 May 2004; accepted 14 June 2004

Available online 22 September 2004

Abstract

We give a characterization of the equality in Hu’s inequality (which is a correlation inequality between two convex functions in \( \mathbb{R}^n \) with respect to the standard Gaussian measure). For this, we prove a new inequality which is slightly better than Hu’s inequality. Then, we obtain a result concerning the U-conjecture.

© 2004 Elsevier SAS. All rights reserved.

Résumé

On donne une caractérisation du cas d’égalité dans l’inégalité de Hu, qui est une inégalité de décorrélation entre deux fonctions convexes contre la mesure gaussienne dans \( \mathbb{R}^n \). Pour cela, on démontre une inégalité légèrement plus forte que l’inégalité de Hu. On obtient aussi un résultat concernant la conjecture U.

© 2004 Elsevier SAS. All rights reserved.

MSC: primary: 52A40, 60E15

Keywords: Gaussian measure; Ornstein–Uhlenbeck semigroup; Convexity; Correlation

1. Introduction

We denote by \( \mu_n \) the standard Gaussian measure on \( \mathbb{R}^n \). In 1973, Kagan, Linnik and Rao [11] considered the following question: if \( P \) and \( Q \) are two polynomials on \( \mathbb{R}^n \) independent with respect to \( \mu_n \) (such that if \( X \) is a random vector on \( \mathbb{R}^n \) of law \( \mu_n \), then \( P(X) \) and \( Q(X) \) are independent random variables), is it possible to find an orthogonal transformation \( U \) on \( \mathbb{R}^n \) and an integer \( k \) such that \( P \circ U \) is a function of \( (x_1, \ldots, x_k) \) and \( Q \circ U \) is a function of \( (x_{k+1}, \ldots, x_n) \) ? If the answer is positive, we say that \( P \) and \( Q \) are unlinked. This question is known as the U-conjecture. Kagan, Linnik and Rao gave a partial answer in [11]. Recently, Bhandari and Basu [5] 

E-mail address: gharge@maths.univ-evry.fr (G. Hargé).
have shown that if $P$ and $Q$ are two convex, positive polynomials independent with respect to $\mu_n$ and if $P(0) = 0$ then $P$ and $Q$ are unlinked. Related to this problem, Bhandari and DasGupta [6] proved in 1994 that two convex and even functions $f$ and $g$ are unlinked if they are uncorrelated (that is if $\int f g d\mu_n = \int f d\mu_n \int g d\mu_n$) under an additional hypothesis. This hypothesis is related to the Gaussian correlation conjecture which is still a conjecture in dimension greater than two (see [7] for further details and references on this conjecture), so, the result of Bhandari and DasGupta is not proved in the general case, that is, without this additional hypothesis. Recall now an inequality due to Hu [10] and which concerns two convex functions $f$ and $g$ in $L^2(\mu_n)$ (see [9] or [8] for the second term in the inequality):

$$\int fg d\mu_n \geq \int f d\mu_n \int g d\mu_n + \left( \int xf d\mu_n, \int xg d\mu_n \right). \quad (1.1)$$

If $(\int xf d\mu_n, \int xg d\mu_n) \geq 0$, this is a correlation inequality between $f$ and $g$ ($\langle \rangle$ is the usual scalar product on $\mathbb{R}^n$).

We will prove the following theorem, which shows in particular the result of Bhandari and DasGupta in the general case, and which gives a partial answer to the U-conjecture:

**Theorem 1.1.** Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be convex functions in $L^2(\mu_n)$ for which equality holds in (1.1). Then, there exist an orthogonal transformation $U$ on $\mathbb{R}^n$, two vectors $\alpha_1$ and $\alpha_2$ in $\mathbb{R}^n$, an integer $k \in \{0, \ldots, n\}$ and two convex functions $f : \mathbb{R}^k \to \mathbb{R}$ in $L^2(\mu_k)$ and $g : \mathbb{R}^{n-k} \to \mathbb{R}$ in $L^2(\mu_{n-k})$ such that, for all $x$ in $\mathbb{R}^n$:

$$f(Ux) = \langle \alpha_1, x \rangle + \tilde{f}(x_1, \ldots, x_k) \quad \text{and} \quad g(Ux) = \langle \alpha_2, x \rangle + \tilde{g}(x_{k+1}, \ldots, x_n) \quad (x = (x_1, \ldots, x_n)).$$

Furthermore $(\alpha_1)_i = 0$ if $i \leq k$ and $(\alpha_2)_i = 0$ if $i > k$ (if $k = 0$ or $k = n$, we make obvious conventions).

Of course, the condition is sufficient. Actually, it is easy to see that if $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a function in $L^2(\mu_n)$ and if $\alpha$ belongs to $\mathbb{R}^n$ then $\varphi$ and the function: $x \mapsto \langle \alpha, x \rangle$ satisfy (1.1) with equality.

Secondly, we will show a generalization of the result of Bhandari and Basu:

**Theorem 1.2.** Denote by $X$ a random vector of law $\mu_n$. Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be convex functions in $L^2(\mu_n)$. We assume that $f$ is an analytic function which verifies, for all $x$ in $\mathbb{R}^n$, $f(x) \geq f(0)$. If $f(X)$ and $g(X)$ are independent random variables then $f$ and $g$ are unlinked.

The proof of Theorem 1.1 uses a new proof of Hu’s inequality and the Ornstein–Uhlenbeck semigroup. In fact, we will prove a new inequality which is slightly better than Hu’s inequality.

## 2. Reinforcement of Hu’s inequality

Two functions $f$ and $g$ for which equality holds in (1.1) are not necessarily regular functions. Of course, it is possible to approximate such functions with regular functions but those last functions do not satisfy automatically (1.1) with equality. Nevertheless, that will be the case if we approximate $f$ and $g$ with the help of the Ornstein–Uhlenbeck semigroup. For this, we will show an inequality which is a reinforcement of Hu’s inequality (Theorem 2.1 below) and which concerns the Ornstein–Uhlenbeck semigroup.

For the last thirty years, semigroups have been used to prove various inequalities as correlation, concentration, Poincaré and log-Sobolev inequalities. Concerning correlation inequalities, it is possible to refer to the works of Pitt [17] (perhaps the first use of the Ornstein–Uhlenbeck semigroup to prove a correlation inequality), Bakry and Michel [3], Hu [10], or my former works [7,8]. There is also related work of Houdré, Pérez-Abreu and Surgailis [9]. Concerning concentration, Poincaré and log-Sobolev inequalities, see for example the surveys of Bakry and...
Ledoux [1,2,12–14]. Furthermore, we will compare the following result with a work of Beckner [4] concerning
Poincaré inequality.

We define the Ornstein–Uhlenbeck semigroup with the Mehler formula:

\[ \forall f \in L^2(\mu_n), \quad P_t f(x) = \int f(e^{-t}x + \sqrt{1-e^{-2t}}y) d\mu_n(y). \]

The Ornstein–Uhlenbeck operator \( L \) is defined by:

\[ D(L) = \left\{ f \in L^2(\mu), \frac{P_tf - f}{t} \text{ possesses a limit in } L^2(\mu) \text{ when } t \text{ goes to 0} \right\}, \]

\[ Lf = \lim_{t \to 0} \frac{P_tf - f}{t} \text{ for } f \in D(L). \]

It is well known that \( D(L) = \{ f \in L^2(\mu), \Delta f - \langle x, \nabla f \rangle \in L^2(\mu) \} \) and \( Lf = \Delta f - \langle x, \nabla f \rangle \) (where \( \Delta f - \langle x, \nabla f \rangle \) is taken in distribution sense). Furthermore, we have:

\[ \forall f \in D(L), \forall t > 0, \quad (P_tf) \in D(L) \quad \text{and} \quad \frac{d}{dt} P_tf = Lf = P_tL f \quad \text{(in } L^2(\mu)\text{).} \]

The properties of \( P_t \) and \( L \) we will use in the next theorems and remarks could be found in [15] or [16].

**Theorem 2.1.** Let \( f, g : \mathbb{R}^n \to \mathbb{R} \) be convex functions such that \( f \in L^2(\mu_n) \) and \( g \in L^2(\mu_n) \), then, for all \( t \geq 0 \):

\[ \int fg \, d\mu_n \geq \int f P_t g \, d\mu_n + (1-e^{-t}) \left( \int xf \, d\mu_n, \int xg \, d\mu_n \right). \]

**Remark 1.** The inequality obtained in this theorem can be compared to a generalized Poincaré inequality due to Beckner [4]. He showed, for \( f \in L^2(\mu_n) \):

\[ \int f^2 \, d\mu_n - \int (P_tf)^2 \, d\mu_n \leq (1-e^{-2t}) \int \|\nabla f\|^2 \, d\mu_n. \]

If we choose \( f = g \) in Theorem 2.1 and if we replace \( t \) by \( 2t \), we obtain, for a convex function \( f \) (using the fact that \( P_t \) is a symmetric semigroup with respect to \( \mu_n \)):

\[ (1-e^{-2t}) \left\| \int \nabla f \, d\mu_n \right\|^2 \leq \int f^2 \, d\mu_n - \int (P_{2t}f)^2 \, d\mu_n. \]

**Remark 2.** We deduce from Theorem 2.1, that if \( f \) and \( g \) are in \( L^2(\mu_n) \) and convex, then:

\[ \int \frac{\frac{f}{t} - P_t f}{t} \, d\mu_n \geq \frac{1-e^{-t}}{t} \left( \int xf \, d\mu_n, \int xg \, d\mu_n \right). \]

So, we obtain, when \( t \) goes to 0 and if \( g \in D(L) \):

\[ -\int f Lg \, d\mu_n \geq \left( \int xf \, d\mu_n, \int xg \, d\mu_n \right). \]

We will prove this inequality before Theorem 2.1.

**Remark 3.** Recall that, for \( g \in L^2(\mu_n) \), \( \lim_{t \to +\infty} P_t g = f \) (in \( L^2(\mu_n) \)). Consequently, if we let \( t \) goes to infinity in Theorem 2.1, we recover inequality (1.1).
Remark 4. In fact, the inequality of Theorem 2.1 is more accurate than inequality (1.1). Actually, Theorem 2.1 gives:

\[ \int f g \, d\mu_n \geq \int f P_t g \, d\mu_n + (1 - e^{-t}) \left( \int x f \, d\mu_n, \int x g \, d\mu_n \right). \]

Furthermore, if we apply inequality (1.1) to \( f \) and \( P_t g \) (which is convex with the help of Mehler formula), we obtain:

\[ \int f P_t g \, d\mu_n \geq \int f d\mu_n \int g \, d\mu_n + \left( \int x f \, d\mu_n, \int x g \, d\mu_n \right). \]

because \( \int x P_t g \, d\mu_n = \int g P_t x \, d\mu_n = e^{-t} \int x g \, d\mu_n \). Consequently:

\[ \int f g \, d\mu_n \geq \int f P_t g \, d\mu_n + (1 - e^{-t}) \left( \int x f \, d\mu_n, \int x g \, d\mu_n \right) \geq \int f d\mu_n \int g \, d\mu_n + \left( \int x f \, d\mu_n, \int x g \, d\mu_n \right). \]

Remark 5. Theorem 2.1 allows us to show that if \( f \) and \( g \) are convex functions and satisfy (1.1) with equality then \( P_t f \) and \( P_t g \) verify the same equality. Actually, we have:

\[ \int f P_t g \, d\mu_n \geq \int f P_t g \, d\mu_n + (1 - e^{-2t}) \left( \int x f \, d\mu_n, \int x g \, d\mu_n \right) \geq \int f d\mu_n \int g \, d\mu_n + \left( \int x f \, d\mu_n, \int x g \, d\mu_n \right). \]

If \( f \) and \( g \) satisfy (1.1) with equality then:

\[ \int f g \, d\mu_n = \int f P_t g \, d\mu_n \geq \int f P_t g \, d\mu_n + (1 - e^{-2t}) \left( \int x f \, d\mu_n, \int x g \, d\mu_n \right) \Rightarrow \int P_t f P_t g \, d\mu_n = \int P_t f d\mu_n \int P_t g d\mu_n + \left( \int x P_t f \, d\mu_n, \int x P_t g \, d\mu_n \right). \]

Before given the proof of Theorem 2.1, we will show the following result:

Theorem 2.2. If \( f, g : \mathbb{R}^n \to \mathbb{R} \) are convex functions in \( L^2(\mu_n) \) and if \( g \in D(L) \), then:

\[ -\int L g \, d\mu_n \geq \left( \int x f \, d\mu_n, \int x g \, d\mu_n \right). \]

Proof. We will prove this theorem for \( P_u f \) and \( P_u g \) \((u > 0)\) instead of \( f \) and \( g \). Then, we will deduce easily the result for \( f \) and \( g \) with the following convergences in \( L^2(\mu_n) \):

\[ \lim_{u \to 0} P_u f = f, \quad \lim_{u \to 0} P_u g = g, \quad \lim_{u \to 0} L P_u g = \lim_{u \to 0} P_u L g = L g. \]

We will use the following properties of \( P_t \) and \( L \):

- \( \forall h \in L^2(\mu_n), \, \forall t > 0, \, x \mapsto P_t h(x) \) is a \( C^\infty \) function on \( \mathbb{R}^n \), \( \frac{\partial}{\partial s} (P_t h) (x) \in L^2(\mu_n) \) and \( \forall x \in \mathbb{R}^n, \, \forall s \geq 0, \, \frac{\partial}{\partial s} (P_{s+t} h)(x) = e^{-t} P_t \left( \frac{\partial}{\partial s} (P_s h) \right)(x). \)
\[ \forall (h_1, h_2) \in D(L)^2, \int h_1 L h_2 \, d\mu_n = - \int \langle \nabla h_1, \nabla h_2 \rangle \, d\mu_n. \]

We notice that \( P_u f \) and \( P_u g \) are convex, \( C^\infty \) and in \( D(L) \).

Define \( \theta(t) = \int (\nabla P_u f, P_t \nabla P_u g) \, d\mu_n \). Because \( u \) is strictly positive, \( \nabla P_u f \) and \( \nabla P_u g \) are in \( L^2(\mu_n) \). Consequently:

\[ \lim_{t \to \infty} \theta(t) = \left( \int \nabla P_u f \, d\mu_n, \int \nabla P_u g \, d\mu_n \right). \]

Moreover:

\[ \theta'(t) = \sum_{i=1}^n \int \left\langle \nabla \left[ \frac{\partial P_u f}{\partial x_i} \right], \nabla \left[ \frac{\partial P_u g}{\partial x_i} \right] \right\rangle \, d\mu_n = - \sum_{i=1}^n \int \left\langle \nabla \left[ \frac{\partial P_u f}{\partial x_i} \right], \nabla \left[ \frac{\partial P_u g}{\partial x_i} \right] \right\rangle \, d\mu_n. \]

It is possible to justify this equality by saying that \( P_t \left( \frac{\partial P_u f}{\partial x_1} \right) \) belongs to \( D(L) \) and that \( \frac{\partial P_u f}{\partial x_1} \) belongs to \( D(L) \) (because \( \frac{\partial P_u f}{\partial x_1} = e^{-u/2} P_u f \)). We write:

\[ P_t \left( \frac{\partial P_u g}{\partial x_1} \right) = e^t \left( \frac{\partial P_u g}{\partial x_1} \right). \]

then, we obtain:

\[ \theta'(t) = -e^t \sum_{i=1}^n \int \left\langle \nabla \left[ \frac{\partial P_u f}{\partial x_i} \right], \nabla \left[ \frac{\partial P_{t+u} g}{\partial x_i} \right] \right\rangle \, d\mu_n = -e^t \int \text{Tr}(\text{Hess} \, P_u f \, \text{Hess} \, P_{t+u} g) \, d\mu_n. \]

Because \( P_u f \) and \( P_{t+u} g \) are convex, we deduce \( \theta'(t) \leq 0 \). Consequently:

\[ \int (\nabla P_u f, \nabla P_u g) \, d\mu_n \geq \left( \int \nabla P_u f \, d\mu_n, \int \nabla P_u g \, d\mu_n \right), \]

which gives:

\[ - \int P_u f L P_u g \, d\mu_n \geq \left( \int x P_u f \, d\mu_n, \int x P_u g \, d\mu_n \right). \]

\[ \textbf{Proof of Theorem 2.1.} \] For two convex functions \( f \) and \( g \) in \( L^2(\mu_n) \), we define:

\[ \xi(t) = \int f P_t g \, d\mu_n - \left( \int x f \, d\mu_n, \int x P_t g \, d\mu_n \right) \]

\[ = \int f P_t g \, d\mu_n - e^{-t} \left( \int x f \, d\mu_n, \int x g \, d\mu_n \right). \]

It is sufficient to show that \( \xi \) is a decreasing function.

\[ \xi'(t) = \int f L P_t g \, d\mu_n - \left( \int x f \, d\mu_n, \int x L P_t g \, d\mu_n \right) \]

\[ = \int f L P_t g \, d\mu_n - \left( \int x f \, d\mu_n, \int L x P_t g \, d\mu_n \right) \]

\[ = \int f L P_t g \, d\mu_n + \left( \int x f \, d\mu_n, \int x P_t g \, d\mu_n \right). \]

We apply the previous theorem to \( f \) and \( P_t g \) to obtain \( \xi'(t) \leq 0 \). \[ \square \]
3. Characterization of equality in Hu’s inequality

We begin to prove some elementary facts concerning convex functions. The following reasoning process is inspired by the one of Bhandari and DasGupta [6].

Lemma 3.1. If \( \varphi : \mathbb{R} \to \mathbb{R} \) is a convex and non-constant function then \( \lim_{t \to +\infty} \varphi(t) = +\infty \) or \( \lim_{t \to -\infty} \varphi(t) = +\infty \) [6].

We deduce, by using the convexity of \( \varphi(x) - \langle \alpha, x \rangle \):

If \( \varphi : \mathbb{R}^n \to \mathbb{R} \) is convex and verify

\[
\exists \alpha \in \mathbb{R}^n, \exists b \in \mathbb{R}, \forall x \in \mathbb{R}^n, \varphi(x) \leq \langle \alpha, x \rangle + b,
\]

then: \( \forall x \in \mathbb{R}^n, \varphi(x) = \langle \alpha, x \rangle + \varphi(0) \).

Lemma 3.2. Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a convex function. We define:

\[
\mathcal{E}(f) = \{ h \in \mathbb{R}^n, \exists (a, b) \in \mathbb{R}^2, \forall t \in \mathbb{R}, f(th) \leq at + b \}.
\]

\( \mathcal{E}(f) \) is a linear space contained in \( \mathbb{R}^n \) and

\[
\exists! \alpha \in \mathcal{E}(f), \forall y_1 \in \mathcal{E}(f), \forall y_2 \in \mathbb{R}^n, f(y_1 + y_2) = \langle \alpha, y_1 \rangle + f(y_2).
\]

Proof. It is obvious to see that \( \mathcal{E}(f) \) is a linear space. Let \( (e_1, \ldots, e_r) \) be an orthonormal basis of \( \mathbb{R}^n \) and define:\n
\[
\tilde{\varphi}(x) = \int \varphi \left( x + \sum_{i=1}^{k} y_i e_i \right) d\mu_k(y) \text{ where } y = (y_1, \ldots, y_k).
\]

Unicity of \( \alpha \) is obvious. \( \square \)

Remark 6. If we choose \( y_2 = 0 \) and \( h \in \mathcal{E}(f) \) we see that \( \forall t \in \mathbb{R}, f(th) = t \langle \alpha, h \rangle + f(0) \).

Lemma 3.3. Let \( \varphi : \mathbb{R}^n \to \mathbb{R} \) be a convex function in \( L^2(\mu_n) \). Let \( h \in \mathbb{R}^n \), we assume that the map: \( t \mapsto \varphi(th) \) is a non-constant function. Let \( (e_1, \ldots, e_k) \) be an orthonormal family in \( \mathbb{R}^n \) (k \( \leq \) n) and define:

\[
\tilde{\varphi}(x) = \int \varphi \left( x + \sum_{i=1}^{k} y_i e_i \right) d\mu_k(y) \text{ where } y = (y_1, \ldots, y_k).
\]

Then:

\[
\exists \varepsilon \in [-1, 1], \lim_{t \to +\infty} \tilde{\varphi}(th) = +\infty,
\]

\[
\exists \varepsilon \in [-1, 1], \forall s \geq 0, \lim_{t \to +\infty} (P_s \varphi)(th) = +\infty.
\]
Proof. We know from Lemma 3.1 that
\[ \exists \varepsilon \in \{-1, 1\}, \lim_{t \to +\infty} \varphi(th) = +\infty. \]

Let \( y \in \mathbb{R}^n : \varphi(\frac{t}{2}h) \leq \frac{1}{2}(\varphi(th + y) + \varphi(-y)) \) so \( \lim_{t \to +\infty} \varphi(th + y) = +\infty. \) Furthermore:
\[ \exists A > 0, \varepsilon t > A \Rightarrow \varphi\left(\frac{t}{2}h\right) \geq 0. \]

With Fatou's lemma, we write:
\[ \int \left( \liminf_{t \to +\infty} \left[ \varphi\left( th + \sum_{i=1}^{k} y_i e_i \right) + \varphi\left( -\sum_{i=1}^{k} y_i e_i \right) \right] d\mu_k(y) \right) \leq \liminf_{t \to +\infty} \int \left[ \varphi\left( th + \sum_{i=1}^{k} y_i e_i \right) + \varphi\left( -\sum_{i=1}^{k} y_i e_i \right) \right] d\mu_k(y) \]
\[ \Rightarrow \lim_{t \to +\infty} \tilde{\varphi}(th) = +\infty. \]

We can prove the result for \( Ps\varphi \) in the same way.

Lemma 3.4. Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a convex function in \( L^2(\mu) \). Let \( h \in \mathbb{R}^n \) and \( s \geq 0 \). Like in the previous lemma, we associate \( \tilde{f} \) to \( f \), then:
\[ h \notin E(f) \Rightarrow (h \notin E(\tilde{f}) \text{ and } h \notin E(Ps f)). \]

Proof. \( h \notin E(f) \) so, for every \( a \in \mathbb{R} \), the map \( t \mapsto f(th) - at \) is non-constant. Define \( f_a(x) = f(x) - a(x, \frac{h}{\|h\|}) \). \( f_a \) is convex and the map \( t \mapsto f_a(th) \) is non-constant, so, with the previous lemma:
\[ \exists \varepsilon \in [-1, 1], \lim_{t \to +\infty} \tilde{f}_a(th) = +\infty. \]

Moreover:
\[ \int f_a(th + \sum_{i=1}^{k} y_i e_i) d\mu_k(y) = \int f(th + \sum_{i=1}^{k} y_i e_i) d\mu_k(y) - at = \tilde{f}(th) - at. \]

We deduce that the map \( t \mapsto f(th) - at \) is non-constant for all \( a \). So \( h \notin E(\tilde{f}) \). We can prove the result for \( Ps\varphi \) in the same way.

Remark 7. In fact, the following equalities are true:
\[ E(f) = E(\tilde{f}) = E(Ps f). \]
Actually, we have for all \( \beta \in [0, \sqrt{2}] \)
\[ \int f(\beta x) d\mu_n(x) < +\infty. \]
Let \( h \in E(f) \) then \( \exists(a, b) \in \mathbb{R}^2, \forall t \in \mathbb{R}, f(th) \leq at + b. \) For \( y \in \mathbb{R}^n \) and \( \beta \in [1, \sqrt{2}], \) we write:
\[ th + y = (1 - \frac{1}{\beta}) \frac{h}{\|h\|} th + \frac{1}{\beta} f(\beta y). \]
We deduce:
\[ h \in E(\tilde{f}) \text{ and } h \in E(Ps f). \]

Now, we can prove Theorem 1.1.

Proof of Theorem 1.1. We choose \((e_1, \ldots, e_n)\) an orthonormal basis of \( \mathbb{R}^n \) such that:

- \((e_1, \ldots, e_{r+k})\) is an orthonormal basis of \( E(f)^\perp, \)
- \((e_{r+1}, \ldots, e_{r+k})\) is an orthonormal basis of \( E(f)^\perp \cap E(g). \)
We use Lemma 3.2 to construct $\alpha_1$ associated to $f$ and $\alpha_2$ associated to $g$. We obtain:

\[
\begin{align*}
&f\left(\sum_{i=1}^{n} x_i e_i\right) = f\left(\sum_{i=1}^{r+k} x_i e_i\right) + \left(\alpha_1, \sum_{i=r+k+1}^{n} x_i e_i\right), \\
g\left(\sum_{i=1}^{n} x_i e_i\right) = g\left(\sum_{i=1}^{r} x_i e_i + \sum_{i=r+k+1}^{n} x_i e_i\right) + \left(\alpha_2, \sum_{i=r+1}^{r+k} x_i e_i\right).
\end{align*}
\]

We want to prove $r = 0$ so we assume $r \neq 0$. We will use the fact that $\sum_{i=1}^{r} x_i e_i \notin \mathcal{E}(f) \cup \mathcal{E}(g)$ if $(x_1, \ldots, x_r) \neq 0$.

Define (abuse of notations):

\[
\tilde{f}(x_1, \ldots, x_r) = \int f\left(\sum_{i=1}^{r} x_i e_i\right) d\mu_k(x_{r+1}, \ldots, x_{r+k}),
\]

\[
\tilde{g}(x_1, \ldots, x_r) = \int g\left(\sum_{i=1}^{r} x_i e_i + \sum_{i=r+k+1}^{n} x_i e_i\right) d\mu_{n-r-k}(x_{r+k+1}, \ldots, x_n).
\]

Recall that if $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a function in $L^2(\mu_n)$ and if $\alpha \in \mathbb{R}^n$ then $\varphi$ and the map $x \mapsto (\alpha, x)$ satisfy (1.1) with equality. So, because $f$ and $g$ satisfy (1.1) with equality, it is the case for $\tilde{f}$ and $\tilde{g}$ against $\mu_r$ in $\mathbb{R}^r$. Furthermore (Lemma 3.4) $\sum_{i=1}^{r} x_i e_i \notin \mathcal{E}(f) \cup \mathcal{E}(g)$ if $(x_1, \ldots, x_r) \neq 0$. We deduce $\mathcal{E}(\tilde{f}) = \mathcal{E}(\tilde{g}) = \{0\}$.

Using Remark 4 and because $\tilde{f}$ and $\tilde{g}$ satisfy (1.1) with equality, we have:

\[
\forall u \geq 0, \quad \int \tilde{f} \tilde{g} d\mu_r = \int f \tilde{P}_u \tilde{g} d\mu_r + (1 - e^{-u})\left(\int x \tilde{f} d\mu_r, \int x \tilde{g} d\mu_r\right),
\]

(the Ornstein–Uhlenbeck semigroup we use here is the one of $\mathbb{R}^r$).

For the continuation of the proof, we only need the existence of one real $u > 0$ such that this equality is verified. The function $\xi$ used in the proof of Theorem 2.1 is decreasing, so we deduce:

\[
\forall s \in [0, u], \quad \xi'(s) = 0.
\]

Consequently, for all $s \in [0, u]$:

\[
- \int \tilde{f} P_{s/2} \tilde{g} d\mu_r = \left(\int x \tilde{f} d\mu_r, \int x P_{s/2} \tilde{g} d\mu_r\right)
\]

\[
\Rightarrow - \int P_{s/2} \tilde{f} L P_{s/2} \tilde{g} d\mu_r = \left(\int x \tilde{f} d\mu_r, \int P_{s/2} x P_{s/2} \tilde{g} d\mu_r\right)
\]

\[
= e^{s/2} \left(\int x \tilde{f} d\mu_r, \int x P_{s/2} \tilde{g} d\mu_r\right) = \left(\int x P_{s/2} \tilde{f} d\mu_r, \int x P_{s/2} \tilde{g} d\mu_r\right)
\]

\[
\Rightarrow \int (\nabla P_{s/2} \tilde{f}, \nabla P_{s/2} \tilde{g}) d\mu_r = \left(\int \nabla P_{s/2} \tilde{f} d\mu_r, \int \nabla P_{s/2} \tilde{g} d\mu_r\right).
\]

We define, like in the proof of Theorem 2.2, the function:

\[
\theta(t) = \int (\nabla P_{t/2} \tilde{f}, P_t \nabla P_{t/2} \tilde{g}) d\mu_r,
\]
So we have, for every $t \geq 0$, $\theta'(t) = 0$, consequently:

$$\forall t \geq 0, \quad \int \text{Tr}(\text{Hess } P_{s/2} \bar{f} \text{ Hess } P_{t+s/2} \bar{g}) \, d\mu_r = 0.$$ 

Using the formula $\text{Hess } P_{t+s/2} \bar{g} = e^{-2t} P_t(\text{Hess } P_{s/2} \bar{g})$, we obtain:

$$\forall t \geq 0, \quad \int \text{Tr}(\text{Hess } P_{s/2} \bar{f} P_t(\text{Hess } P_{s/2} \bar{g})) \, d\mu_r = 0.$$ 

Let $t$ goes to infinity, then:

$$\text{Tr}\left( \int \text{Hess } P_{s/2} \bar{f} \, d\mu_r \int \text{Hess } P_{s/2} \bar{g} \, d\mu_r \right) = 0. \quad (3.1)$$

Matrices $\int \text{Hess } P_{s/2} \bar{f} \, d\mu_r$ and $\int \text{Hess } P_{s/2} \bar{g} \, d\mu_r$ are symmetric and positive. It is easy to see that if both are invertible then equality (3.1) is impossible. So, we can assume that $\int \text{Hess } P_{s/2} \bar{f} \, d\mu_r$ is not invertible. We deduce there exists an element $h$ of $\mathbb{R}'$, $h \neq 0$, such that:

$$\left[ \left( \int \text{Hess } P_{s/2} \bar{f} \, d\mu_r \right) h, h \right] = 0$$

$$\Rightarrow \int (\text{Hess } P_{s/2} \bar{f} h) \, d\mu_r = 0$$

$$\Rightarrow \forall x \in \mathbb{R}', \quad (\text{Hess } P_{s/2} \bar{f}(x) h, h) = 0.$$ 

Define $\zeta(\lambda) = P_{s/2} \bar{f}(\lambda h)$, we obtain, for all $\lambda$, $\zeta''(\lambda) = 0$, consequently:

$$\exists (a, b) \in \mathbb{R}^2, \quad \forall x \in \mathbb{R}, \quad P_{s/2} \bar{f}(\lambda h) = a\lambda + b.$$ 

So $h \in \mathcal{E}(P_{s/2} \bar{f})$. Then, we deduce from Lemma 3.4 that $h \in \mathcal{E}(\bar{f})$ but this is impossible because $\mathcal{E}(\bar{f}) = \{0\}$. $\Box$

**Remark 8.** Actually, we have proved that if $f$ and $g$ are two convex functions for which equality holds in Theorem 2.1 for a fixed $t > 0$, then the conclusion of Theorem 1.1 remains valid.

**Remark 9.** We obtain in the proof of the theorem: $\alpha_1 \in \mathcal{E}(f)$ and $\alpha_2 \in \mathcal{E}(g)$.

We deduce immediately from Theorem 1.1:

**Corollary 3.5.** Let $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ be two convex functions in $L^2(\mu_n)$ such that $\int f \, g \, d\mu_n = \int f \, d\mu_n \int g \, d\mu_n$ and $(\int f \, d\mu_n, \int g \, d\mu_n) \geq 0$. Then $\langle \int f \, d\mu_n, \int g \, d\mu_n \rangle = 0$ and there exist an orthogonal transformation $U$ on $\mathbb{R}^n$, two vectors $\alpha_1$ and $\alpha_2$ in $\mathbb{R}^n$, an integer $k \in \{0, \ldots, n\}$ and two convex functions $\bar{f} : \mathbb{R}^k \rightarrow \mathbb{R}$ in $L^2(\mu_k)$ and $\bar{g} : \mathbb{R}^{n-k} \rightarrow \mathbb{R}$ in $L^2(\mu_{n-k})$ such that, for every $x \in \mathbb{R}^n$:

$$f(Ux) = \langle \alpha_1, x \rangle + \tilde{f}(x_1, \ldots, x_k) \quad \text{and} \quad g(Ux) = \langle \alpha_2, x \rangle + \tilde{g}(x_{k+1}, \ldots, x_n) \quad (x = (x_1, \ldots, x_n)).$$

We have $\langle \alpha_1, \alpha_1 \rangle = 0$ if $i \leq k$ and $\langle \alpha_2, \alpha_2 \rangle = 0$ if $i > k$.

Moreover, if $\int f \, d\mu_n = \int g \, d\mu_n = 0$ then $\langle \int f \, d\mu_n, \int g \, d\mu_n \rangle = 0$ is verified and $\alpha_1 = \alpha_2 = 0$ (consequently, $f$ and $g$ are unlinked).

**Proof.** We start with Hu’s inequality:

$$\int fg \, d\mu_n \geq \int f \, d\mu_n \int g \, d\mu_n + \left( \int xf \, d\mu_n, \int xg \, d\mu_n \right) \geq \int f \, d\mu_n \int g \, d\mu_n.$$
We obtain: $\int xf \, d\mu_n, \int xg \, d\mu_n = 0$. Then, we use Theorem 1.1 to construct $U$.

Now, assume that $\int xf \, d\mu_n = 0$. Denote $x = (y_1, y_2)$ where $y_1 = (x_1, \ldots, x_k)$ and $y_2 = (x_{k+1}, \ldots, x_n)$, $\alpha_1 = (0, \tilde{\alpha})$ with $\tilde{\alpha} \in \mathbb{R}^{n-k}$. We obtain: $\int f(Ux) \, d\mu_n(x) = (\int f(y_1) \, d\mu_k(y_1), \tilde{\alpha})$ and $\tilde{\alpha} = 0$. □

**Remark 10.** The second part of this corollary generalizes and proves the result of Bhandari and DasGupta [6] in any dimension (if $f$ and $g$ are even functions then $\int xf \, d\mu_n = \int xg \, d\mu_n = 0$).

### 4. The $U$-conjecture

In the following, we denote by $X = (X_1, \ldots, X_n)$ a random vector of law $\mu_n$. We deduce from the previous corollary:

**Theorem 4.1.** Let $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ be two convex functions in $L^2(\mu_n)$. Assume that $\int xf \, d\mu_n = 0$ and that $f(X)$ and $g(X)$ are independent random variables then $f$ and $g$ are unlinked.

**Proof.** The equality $\int xg \, d\mu_n = 0$ is verified. With notations and results of the first part of the previous corollary, we obtain $\alpha_1 = 0$. So, it is possible to find $U$ and $\tilde{f}$ such that:

$$f(Ux) = \tilde{f}(x_1, \ldots, x_k) \text{ and } g(Ux) = ax_1 + \tilde{g}(x_{k+1}, \ldots, x_n).$$

It is easy to see that $f(UX)$ and $g(UX)$ are independent random variables. Denote $Y = (X_2, \ldots, X_k)$ and $Z = (X_{k+1}, \ldots, X_n)$. $\tilde{f}(X_1, Y)$ and $aX_1 + \tilde{g}(Z)$ are independent random variables, so it is the case for $\tilde{f}(X_1, Y)$ and $aX_1$. Recall that in the proof of Theorem 1.1, we have obtained:

$$\tilde{f}(x_{r+1}, \ldots, x_{r+k}) = \tilde{f}\left(\sum_{i=r+1}^{r+k} x_i e_i\right),$$

where, for $r+1 \leq i \leq r+k$, $e_i \in \mathcal{E}(f)^{\perp}$. Consequently, $\sum_{i=r+1}^{r+k} x_i e_i \notin \mathcal{E}(f)$ if $(x_{r+1}, \ldots, x_{r+k}) \neq 0$. So, we have $\mathcal{E}(\tilde{f}) = \{0\}$. Let $\psi \in L^2(\mu_1)$ and let us assume $a \neq 0$ and $\tilde{f}$ depends on $x_1$.

$$E\left(\tilde{f}(X_1, Y)\psi(aX_1)\right) = E\left(\tilde{f}(X_1, Y)E\left(\psi(aX_1)\right)\right) = E\left(\tilde{f}(X_1, y) \, d\mu_{k-1}(y) \, \psi(aX_1)\right).$$

Define: $\tilde{f}(x_1) = \int \tilde{f}(x_1, y) \, d\mu_{k-1}(y)$. We have $\mathcal{E}(\tilde{f}) = \{0\}$ (Lemma 3.4). Furthermore:

$$E\left(\tilde{f}(X_1)\psi(aX_1)\right) = E\left(\tilde{f}(X_1)\right)E\left(\psi(aX_1)\right).$$

We choose $\psi(x_1) = \tilde{f}(\frac{1}{a} x_1)$. We obtain: $\tilde{f}(X_1) = E(\tilde{f}(X_1))$ almost surely. So, $\tilde{f}$ is a constant function, which is a contradiction with $\mathcal{E}(f) = \{0\}$. We deduce $a = 0$ or $\tilde{f}$ does not depend on $x_1$. □

The following result is to be compared to the one of Bhandari and Basu [5] who show that if $P$ and $Q$ are two convex, positive polynomials independent with respect to $\mu$ and if $P(0) = 0$ then $P$ and $Q$ are unlinked.

**Corollary 4.2.** Let $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ be two convex functions in $L^2(\mu_n)$. We assume that $f$ and $g$ are bounded below, that $\int xf \, d\mu_n, \int xg \, d\mu_n = 0$ and that $\int fg \, d\mu_n = \int f \, d\mu_n \int g \, d\mu_n$ then $f$ and $g$ are unlinked.

**Proof.** We use notations and results of the first part of Corollary 3.5. Because $f$ and $g$ are bounded below, it is easy to see that $\alpha_1 = \alpha_2 = 0$. □

Now, we will prove Theorem 1.2. We begin with a lemma:
Lemma 4.3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function such that:

$$\forall x \in \mathbb{R}^n, \quad x \neq 0 \Rightarrow f(x) > f(0).$$

For $\varepsilon > 0$, define $\delta(\varepsilon) = \sup \{ \| x \|, \ x \in \mathbb{R}^n, \ f(x) - f(0) \leq \varepsilon \}$ then $\lim_{\varepsilon \to 0} \delta(\varepsilon) = 0$.

Proof. Assume that $\delta(\varepsilon)$ does not go to 0. We construct $\eta > 0$ and a sequence $(\varepsilon_p)_{p \geq 1}$ such that $\lim_{p \to +\infty} \varepsilon_p = 0$ and $\delta(\varepsilon_p) > \eta$ for all $p$. We associate to $\varepsilon_p$ an element $x_p$ of $\mathbb{R}^n$ such that $\| x_p \| > \eta$ and $f(x_p) - f(0) \leq \varepsilon_p$. We can assume that $\frac{x_p}{\| x_p \|}$ goes to $e$ (an element of $\mathbb{R}^n$ of norm equal to 1). Then:

$$f\left(\frac{\eta}{\| x_p \|} x_p\right) \leq \left(1 - \frac{\eta}{\| x_p \|}\right) f(0) + \frac{\eta}{\| x_p \|} f(x_p) \leq f(0) + \varepsilon_p.$$  

We obtain: $f(\eta e) \leq f(0)$, so $f(\eta e) = f(0)$, but it is not possible because $\eta e \neq 0$. \hfill $\square$

Proof of Theorem 1.2. We use here the same orthonormal basis as in the proof of Theorem 1.1. So, we construct $\alpha_1$ and $\alpha_2$ associated to $f$ and $g$. Because $f$ is bounded below, we have $\alpha_1 = 0$. Consequently:

$$f\left(\sum_{i=1}^{n} x_i e_i\right) = f\left(\sum_{i=r+1}^{r+k} x_i e_i\right),$$

$$g\left(\sum_{i=1}^{n} x_i e_i\right) = g\left(\sum_{i=1}^{n} x_i e_i + \sum_{i=r+k+1}^{n} x_i e_i\right) + \left(\alpha_2, \sum_{i=r+1}^{r+k} x_i e_i\right).$$

We assume $r \geq 1$. Define:

$$\tilde{g}(x_1, \ldots, x_r) = \int \left(\sum_{i=1}^{n} x_i e_i + \sum_{i=r+k+1}^{n} x_i e_i\right) d\mu_{n-r-k}(x_{r+1}, \ldots, x_n),$$

$$f(x_1, \ldots, x_{r+k}) = f\left(\sum_{i=1}^{r+k} x_i e_i\right).$$

We have $E(\tilde{g}) = \{0\}$ (Lemma 3.4) and $E(f_1) = \{0\}$ (because $(e_1, \ldots, e_{r+k})$ is an orthonormal basis of $E(f)^\perp$). Let $\psi : \mathbb{R} \to \mathbb{R}$ be measurable and bounded, we obtain:

$$\int \left(\tilde{g}(x_1, \ldots, x_r) + \left(\alpha_2, \sum_{i=r+1}^{r+k} x_i e_i\right)\right) \psi(f_1)(x_1, \ldots, x_{r+k}) d\mu_{r+k}(x_1, \ldots, x_{r+k})$$

$$= \int g(x) \psi \circ f(x) \ d\mu_n(x)$$

$$= \int g(x) \ d\mu_n(x) \int \psi \circ f(x) \ d\mu_n(x)$$

$$= \int \left(\tilde{g}(x_1, \ldots, x_r) + \left(\alpha_2, \sum_{i=r+1}^{r+k} x_i e_i\right)\right) d\mu_{r+k}(x_1, \ldots, x_{r+k})$$

$$\times \int \psi(f_1)(x_1, \ldots, x_{r+k}) d\mu_{r+k}(x_1, \ldots, x_{r+k}).$$

Let consider $Y$ a random vector in $\mathbb{R}^{r+k}$ of law $\mu_{r+k}$ and choose, for $\varepsilon > 0$, $\psi(t) = 1_{[-\varepsilon, \varepsilon]}(t - f_1(0))$. We write $Y = (Y_1, Y_2)$ where $Y_1 \in \mathbb{R}^r$, $Y_2 \in \mathbb{R}^k$. So we have:

$$E\left(\tilde{g}(Y_1) + \langle \alpha_2, Y_2 \rangle\right) \frac{1_{f_1(Y) - f_1(0) \leq \varepsilon}}{P(|f_1(Y) - f_1(0)| \leq \varepsilon)} = E\left(\tilde{g}(Y_1) + \langle \alpha_2, Y_2 \rangle\right) = E\left(\tilde{g}(Y_1)\right).$$
If there exists \( x \neq 0 \) such that \( f_1(x) = f_1(0) \), and as \( f_1 \) is convex, we obtain:
\[
\forall t \in [0,1], \quad f_1(tx) = f_1(0).
\]
But \( f_1 \) is an analytic function, so:
\[
\forall t \in \mathbb{R}, \quad f_1(tx) = f_1(0).
\]
Consequently \( x \in \mathcal{E}(f_1) \), which implies \( x = 0 \). We have obtained: \( \forall x \neq 0, \ f_1(x) > f_1(0) \). Now, we use Lemma 4.3 for \( f_1 \). We denote: \( \delta(\varepsilon) = \sup \{ \| x \|, \ x \in \mathbb{R}^{r+k}, f_1(x) - f_1(0) \leq \varepsilon \} \).

\[
E \left( \left| \tilde{g}(Y_1) + \langle \alpha_2, Y_2 \rangle - \bar{g}(0) \right| \right) \leq E \left( \left| \tilde{g}(Y_1) + \langle \alpha_2, Y_2 \rangle - \bar{g}(0) \right| 1_{1_{f_1(Y) - f_1(0)} \leq \varepsilon} \right).
\]

We deduce:
\[
\lim_{\varepsilon \to 0} E \left( \left| \tilde{g}(Y_1) + \langle \alpha_2, Y_2 \rangle \right| 1_{1_{f_1(Y) - f_1(0)} \leq \varepsilon} \right) = \tilde{g}(0).
\]

Consequently:
\[
E \left( \tilde{g}(Y_1) \right) = \bar{g}(0).
\]
The map \( \tilde{g} \) is convex, that means that there is equality in Jensen’s inequality. We deduce \( \bar{g} \) is an affine map, but this is impossible because \( \mathcal{E}(\tilde{g}) = \{ 0 \} \). Consequently \( r = 0 \). Then, it is possible to find an orthogonal transformation \( U \) such that:
\[
f(Ux) = \bar{f}(x_{r+1}, \ldots, x_{r+k}), \quad g(Ux) = \bar{g}(x_{r+k+1}, \ldots, x_n) + ax_{r+1}.
\]
But \( f(Ux) \) and \( g(Ux) \) are independent random variables, so, \( a = 0 \) or \( \bar{f} \) does not depend on \( x_{r+1} \) (see the proof of Theorem 4.1).

**Remark 11.** The idea of using a function like \( 1_{[-\varepsilon,\varepsilon]}(t - f_1(0)) \) and the case of equality in Jensen’s inequality is due to Bhandari and Basu [5]. If we compare the proof given here to their proof, the novelty is the use of \( \mathcal{E}(\bar{g}) \), the choice of the orthonormal basis, the extension to an entire function \( f \) and the end of the proof (to prove \( a = 0 \) or \( \bar{f} \) does not depend on \( x_{r+1} \)).

**Remark 12.** In the proof, we could use \( \{ x \in \mathbb{R}^n, \ \forall t \in \mathbb{R}, \ f(tx) = f(0) \} \) instead of \( \mathcal{E}(f) \). However, it is essential to work with \( \mathcal{E}(g) \). We have to notice that \( f(x) \geq f(0) \) for all \( x \) implies \( \{ x \in \mathbb{R}^n, \ \forall t \in \mathbb{R}, \ f(tx) = f(0) \} = \mathcal{E}(f) \) (by using \( a_1 = 0 \) and Remark 6).

**References**


