Available online at www.sciencedirect.com
SCIENCE @ DIRECT®
ELSEVIER
C. R. Acad. Sci. Paris, Ser. I 337 (2003) 663–666
COMPTES RENDUS MATHEMATIQUE

Probability Theory

The Ehrhard inequality

Christer Borell

School of Mathematical Sciences, Chalmers University of Technology and Göteborg University, 412 96 Göteborg, Sweden

Received 17 August 2003; accepted 29 September 2003

Presented by Paul Deheuvels

Abstract

We prove Ehrhard's inequality for all Borel sets. ***To cite this article: C. Borell, C. R. Acad. Sci. Paris, Ser. I 337 (2003).***

Résumé

L'inégalité d'Ehrhard. Nous démontrons l'inégalité d'Ehrhard pour tous les ensembles boréliens. ***Pour citer cet article : C. Borell, C. R. Acad. Sci. Paris, Ser. I 337 (2003).***

1. Introduction

Throughout this paper let γ_n be the canonical Gaussian measure in $\mathbf{R}^n$, that is

$$\mathrm{d}\gamma_n(x) = \mathrm{e}^{-|x|^2/2}\frac{\mathrm{d}x}{\sqrt{2\pi}^{\,n}},$$

let $\Phi(a) = \gamma_1(]-\infty, a[)$ if $a \in \mathbf{R} \cup \{\pm\infty\}$, and let $\lambda \in]0,1[$. Furthermore, for any $A, B \subseteq \mathbf{R}^n$,

$$\lambda A + (1-\lambda)B = \{\lambda x + (1-\lambda)y;\ x \in A \text{ and } y \in B\}.$$

In [2] Antoine Ehrhard proves that

$$\Phi^{-1}\big(\gamma_n\big(\lambda A + (1-\lambda)B\big)\big) \geqslant \lambda\Phi^{-1}\big(\gamma_n(A)\big) + (1-\lambda)\Phi^{-1}\big(\gamma_n(B)\big)$$

for all convex bodies A and B in $\mathbf{R}^n$. Moreover Latała in [6] shows that Ehrhard's inequality is true if A is a convex body and B an arbitrary Borel set. This special case of Ehrhard's inequality, combined with some short but clever arguments, implies several well-known inequalities for Gaussian measures such as the isoperimetric inequality, the Bobkov inequality, and the Gross logarithmic Sobolev inequality. The Latała paper [7] gives an excellent account on these implications.

The purpose of this paper is to prove Ehrhard's inequality for all Borel sets. This solves Problem 1, p. 456, in the Ledoux and Talagrand book [8]. We here follow the convention that $\infty - \infty = -\infty + \infty = -\infty$.

Theorem 1.1. *The Ehrhard inequality is true for all Borel sets A and B in $\mathbf{R}^n$.*

E-mail address: borell@math.chalmers.se (C. Borell).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2003.09.031

Our proof of Ehrhard's inequality is inspired by a concavity maximum principle initiated by Korevaar in his study of elliptic and parabolic boundary value problems [5] further developed by Greco and Kawohl [3]. In contrast to [3] and [5] the space domain in this paper is unbounded.

It follows from the Ehrhard paper [2] that Theorem 1.1 is true in all dimensions if it is true in one dimension. Since a restriction to one dimension would not really simplify our proof below we will make no restriction on the dimension.

Let $\Delta = \nabla^2 = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$ be Laplace operator in $\mathbf{R}^n$. Given a positive solution u of the heat equation $\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u$ the first point in our proof of Ehrhard's inequality is to introduce the inverse Gaussian transformation $U = \Phi^{-1}(u)$. As $u = \Phi(U)$,

$$\frac{\partial u}{\partial t} = \varphi(U)\frac{\partial U}{\partial t}, \qquad \nabla u = \varphi(U)\nabla U$$

and $\Delta u = \varphi(U)(\Delta U - U|\nabla U|^2)$, where $\varphi(a) = \Phi'(a)$ if $a \in \mathbf{R}$. Thus

$$\frac{\partial U}{\partial t} = \frac{1}{2}\Delta U - \frac{1}{2}U|\nabla U|^2. \tag{1}$$

Let us note that $-U$ is a solution of (1) if U is. Moreover if $U(0,x) = ax + b$, where a and b are real constants, the function $U(t,x) = a(a^2t+1)^{-1/2}x + b(a^2t+1)^{-1/2}$ solves (1).

Our proof of Theorem 1.1 is based on an application of the methods in [3] and [5] to the parabolic differential equation in (1). In this context the Feynman–Kac formula fits very well as will be seen below. We are very grateful to Professor Stanislaw Kwapien for pointing out an alternative to the use of the Feynman–Kac formula in the proof of Theorem 1.1 and sketch his line of reasoning at the very end of Section 2.

2. Proof of Theorem 1.1

To prove Theorem 1.1 we assume without loss of generality that A and B are non-empty compact subsets of $\mathbf{R}^n$. Let $\varepsilon \in]0,1[$ be fixed and choose an infinitely many times differentiable function $F \in \mathcal{C}^\infty(\mathbf{R}^n)$ such that $0 \leqslant F \leqslant 1$, $F = 1$ on A and $F = 0$ off $A_\varepsilon = A + \overline{B}(0,\varepsilon)$, where $\overline{B}(0,\varepsilon)$ is the closed Euclidean in ball $\mathbf{R}^n$ with centre 0 and radius ε. Let $\delta \in]0,\varepsilon[$ and define $f = \delta + (1-\varepsilon)F$. Set $\alpha = \delta + 1 - \varepsilon$ and observe that $\alpha < 1$. In particular, $f \in \mathcal{C}^\infty(\mathbf{R}^n)$, $\delta \leqslant f \leqslant \alpha$, $f = \alpha$ on A, and $f = \delta$ off A_ε. In a similar way, choose a function $g \in \mathcal{C}^\infty(\mathbf{R}^n)$ such that $\delta \leqslant g \leqslant \alpha$, $g = \alpha$ on B, and $g = \delta$ off B_ε. Set

$$\kappa = \max\bigl(\Phi\bigl(\lambda\Phi^{-1}(\alpha) + (1-\lambda)\Phi^{-1}(\delta)\bigr), \Phi\bigl(\lambda\Phi^{-1}(\delta) + (1-\lambda)\Phi^{-1}(\alpha)\bigr)\bigr).$$

The construction shows that $\kappa \to 0$ as $\delta \to 0$. Next we choose a function $h \in \mathcal{C}^\infty(\mathbf{R}^n)$ such that $\kappa \leqslant h \leqslant \alpha$, $h = \alpha$ on $\lambda A_\varepsilon + (1-\lambda)B_\varepsilon$, and $h = \kappa$ off $(\lambda A_\varepsilon + (1-\lambda)B_\varepsilon)_\varepsilon$. The definitions give

$$\Phi^{-1}\bigl(h\bigl(\lambda x + (1-\lambda)y\bigr)\bigr) \geqslant \lambda\Phi^{-1}\bigl(f(x)\bigr) + (1-\lambda)\Phi^{-1}\bigl(g(y)\bigr) \quad \text{if } x, y \in \mathbf{R}^n. \tag{2}$$

Now consider the inequality

$$\Phi^{-1}\Bigl(\int_{\mathbf{R}^n} h\,\mathrm{d}\gamma_n\Bigr) \geqslant \lambda\Phi^{-1}\Bigl(\int_{\mathbf{R}^n} f\,\mathrm{d}\gamma_n\Bigr) + (1-\lambda)\Phi^{-1}\Bigl(\int_{\mathbf{R}^n} g\,\mathrm{d}\gamma_n\Bigr). \tag{3}$$

By first letting $\delta \to 0$ and then $\varepsilon \to 0$ in (3) we obtain the Ehrhard inequality for A and B. The inequality (3) will follow from a slightly more general inequality. Let for every $t \geqslant 0$ and $x \in \mathbf{R}^n$,

$$u_q(t,x) = \int_{\mathbf{R}^n} q(x + \sqrt{t}z)\,\mathrm{d}\gamma_n(z), \quad q = f, g, h.$$

Clearly, (3) follows if

$$\Phi^{-1}\bigl(u_h\bigl(t, \lambda x + (1-\lambda)y\bigr)\bigr) \geqslant \lambda\Phi^{-1}\bigl(u_f(t,x)\bigr) + (1-\lambda)\Phi^{-1}\bigl(u_g(t,y)\bigr) \tag{4}$$

for all $t \geqslant 0$ and $x, y \in \mathbf{R}^n$. The special case $t = 0$ reduces to (2) and the special case $t = 1$ and $x = y = 0$ is the same as (3). To prove (4) let q be any of f, g, or h and define the inverse Gaussian transformation of u_q by $U_q = \Phi^{-1}(u_q)$. Note that $\sup_{t\geqslant 0,\ x\in\mathbf{R}^n} |U_q| < \infty$. Moreover, if $i_1, \ldots, i_n \in \mathbf{N}$ it is readily seen that

$$\sup_{t\geqslant 0,\ x\in\mathbf{R}^n} \left| \frac{\partial^{i_1+\cdots+i_n}}{\partial x^{i_1}\cdots\partial x^{i_n}} U_q \right| < \infty. \tag{5}$$

We now introduce the function $C(t, x, y) = U_h(t, \lambda x + (1-\lambda)y) - \lambda U_f(t, x) - (1-\lambda)U_g(t, y)$ for all $t \geqslant 0$ and $x, y \in \mathbf{R}^n$. The inequality $C(t, x, y) \geqslant 0$ is equivalent to (4). To simplify notation, from now on let $\xi = (t, x)$, $\eta = (t, y)$, and $\varsigma = (t, \lambda x + (1-\lambda)y)$ so that

$$\nabla_x C = \lambda\big\{(\nabla U_h)(\varsigma) - (\nabla U_f)(\xi)\big\}, \tag{6}$$

$$\nabla_y C = (1-\lambda)\big\{(\nabla U_h)(\varsigma) - (\nabla U_g)(\eta)\big\}, \tag{7}$$

$$\Delta_x C = \lambda^2(\Delta U_h)(\varsigma) - \lambda(\Delta U_f)(\xi), \qquad \Delta_y C = (1-\lambda)^2(\Delta U_h)(\varsigma) - (1-\lambda)(\Delta U_g)(\eta)$$

and

$$\sum_{1\leqslant i\leqslant n} \frac{\partial^2 C}{\partial x_i \partial y_i} = \lambda(1-\lambda)(\Delta U_h)(\varsigma).$$

Thus introducing the differential operator

$$\mathcal{E} = \frac{1}{2}\left\{\Delta_x + 2\sum_{1\leqslant i\leqslant n} \frac{\partial^2}{\partial x_i \partial y_i} + \Delta_y\right\}, \qquad \mathcal{E}C = \frac{1}{2}\big\{(\Delta U_h)(\varsigma) - \lambda(\Delta U_f)(\xi) - (1-\lambda)(\Delta U_g)(\eta)\big\}.$$

Now using (1)

$$\mathcal{E}C = \frac{\partial U_h}{\partial t}(\varsigma) + \frac{1}{2}U_h(\varsigma)\big|(\nabla U_h)(\varsigma)\big|^2 - \lambda\frac{\partial U_f}{\partial t}(\xi) - \frac{\lambda}{2}U_f(\xi)\big|(\nabla U_f)(\xi)\big|^2 - (1-\lambda)\frac{\partial U_g}{\partial t}(\eta) - \frac{1-\lambda}{2}U_g(\eta)\big|(\nabla U_g)(\eta)\big|^2$$

or

$$\mathcal{E}C = \frac{\partial C}{\partial t} + \Psi(t, x, y)$$

with

$$\Psi(t, x, y) = \frac{1}{2}U_h(\varsigma)\big|(\nabla U_h)(\varsigma)\big|^2 - \frac{\lambda}{2}U_f(\xi)\big|(\nabla U_f)(\xi)\big|^2 - \frac{1-\lambda}{2}U_g(\eta)\big|(\nabla U_g)(\eta)\big|^2.$$

Here

$$\big|(\nabla U_f)(\xi)\big|^2 = \big|(\nabla U_h)(\varsigma)\big|^2 + \sum_{1\leqslant i\leqslant n}\left\{\frac{\partial U_f}{\partial x_i}(\xi) + \frac{\partial U_h}{\partial x_i}(\varsigma)\right\}\left\{\frac{\partial U_f}{\partial x_i}(\xi) - \frac{\partial U_h}{\partial x_i}(\varsigma)\right\}$$

and

$$\big|(\nabla U_g)(\eta)\big|^2 = \big|(\nabla U_h)(\varsigma)\big|^2 + \sum_{1\leqslant i\leqslant n}\left\{\frac{\partial U_g}{\partial x_i}(\eta) + \frac{\partial U_h}{\partial x_i}(\varsigma)\right\}\left\{\frac{\partial U_g}{\partial x_i}(\eta) - \frac{\partial U_h}{\partial x_i}(\varsigma)\right\}.$$

From these equations and (6) and (7) it follows that $\Psi(t, x, y) = \frac{1}{2}|(\nabla U_h)(\varsigma)|^2 C - b(t, x, y)\cdot\nabla_{(x,y)}C$ for an appropriate continuous function $b(t, x, y)$, which, depending on (5), for fixed t is Lipschitz continuous in the space variables with a Lipschitz constant uniformly bounded in t. Moreover,

$$\mathcal{E}C + b(t, x, y)\cdot\nabla_{(x,y)}C = \frac{\partial C}{\partial t} + \frac{1}{2}\big|(\nabla U_h)(\varsigma)\big|^2 C. \tag{8}$$

In what follows we interpret (∇_x, ∇_y) as an $2n$ by 1 matrice with the transpose matrice $(\nabla_x, \nabla_y)^*$ and have $\mathcal{E} = \frac{1}{2}(\nabla_x, \nabla_y)^* \sigma\sigma^* (\nabla_x, \nabla_y)$ for an appropriate $2n$ by $2n$ matrice σ. Let $T \in]0, \infty[$ be fixed and denote by (X, Y) the solution of the stochastic differential equation

$$\mathrm{d}\big(X(t), Y(t)\big) = b\big(T - t, X(t), Y(t)\big)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), \quad 0 \leqslant t \leqslant T,$$

with the initial value $(X(0), Y(0)) = (x, y)$, where W is a normalized Wiener process in $\mathbf{R}^{2n}$. The Feynman–Kac theorem ([4], p. 366) yields

$$C(T, x, y) = E\big[C\big(0, X(T), Y(T)\big)\,\mathrm{e}^{-\frac{1}{2}\int_0^T |(\nabla U_h)(T-\theta, \lambda X(\theta)+(1-\lambda)Y(\theta))|^2\,\mathrm{d}\theta}\big]$$

and, since $C(0, X(T), Y(T)) \geqslant 0$, we get $C(T, x, y) \geqslant 0$. This completes the proof of Theorem 1.1.

The Feynman–Kac formula can be avoided in the proof of Theorem 1.1. To explain this, again let $T \in]0, \infty[$ be fixed. The definitions of the functions f, g, and h imply that the lower limit of the function $\inf_{0 \leqslant t \leqslant T} C(t, x, y)$ as $|x| + |y| \to \infty$ is non-negative. Therefore, if $C(t, x, y) < 0$ at some point $(t, x, y) \in [0, T] \times \mathbf{R}^n \times \mathbf{R}^n$ there exists a strictly positive number ε such that the function $\varepsilon t + C(t, x, y)$ possesses a strictly negative minimum in $[0, T] \times \mathbf{R}^n \times \mathbf{R}^n$ at a certain point $P = (t_0, x_0, y_0)$ with $t_0 > 0$. Now

$$C(P) < 0, \quad \frac{\partial C}{\partial t}(P) \leqslant -\varepsilon, \quad \nabla_{(x,y)} C(P) = 0, \quad \text{and} \quad \mathcal{E}C(P) \geqslant 0$$

which contradict (8). Thus $C(t, x, y) \geqslant 0$.

3. The Ehrhard inequality in infinite dimension

Let E be a real, locally convex Hausdorff vector space and denote by $\mathcal{B}(E)$ the Borel σ-algebra in E. A Borel probability measure γ on E is a Gaussian Radon measure if each bounded linear functional on E has a Gaussian distribution relative to γ and if $\gamma_* = \gamma$ on $\mathcal{B}(E)$, where for any $A \subseteq E$, $\gamma_*(A) = \sup\{\gamma(K);\ K$ compact subset of $A\}$.

Theorem 3.1. *If γ is a Gaussian Radon measure on E,*

$$\Phi^{-1}\big(\gamma_*\big(\lambda A + (1-\lambda)B\big)\big) \geqslant \lambda\Phi^{-1}\big(\gamma(A)\big) + (1-\lambda)\Phi^{-1}\big(\gamma(B)\big)$$

for all $A, B \in \mathcal{B}(E)$.

Theorem 3.1 follows from Theorem 1.1 using the same line of reasoning as in the author's paper [1].

References

[1] Ch. Borell, Convex measures on locally convex spaces, Ark. Mat. 12 (1974) 239–252.
[2] A. Ehrhard, Symétrisation dans l'espace de Gauss, Math. Scand. 53 (1983) 281–301.
[3] A. Greco, B. Kawohl, Log-concavity in some parabolic problems, Electronic J. Differential Equations 19 (1999) 1–12.
[4] I. Karatzas, E. Shreve, Brownian Motion and Stochastic Calculus, 2nd edition, Springer, 1991.
[5] N.J. Korevaar, Convex solutions to non-linear elliptic and parabolic boundary value problems, Indiana Univ. Math. J. 32 (1983) 603–614.
[6] R. Latała, A note on the Ehrhard inequality, Studia Math. 118 (1996) 169–174.
[7] R. Latała, On some inequalities for Gaussian measures, Proc. ICM 2 (2002) 813–822.
[8] M. Ledoux, M. Talagrand, Probability in Banach Spaces, Springer, 1991.