DOI: 10.1478/AAPP.91S1A18

AAPP | Atti della Accademia Peloritana dei Pericolanti
Classe di Scienze Fisiche, Matematiche e Naturali
ISSN 1825-1242

Vol. 91, Suppl. No. 1, A18 (2013)

LYAPUNOV FUNCTIONALS FOR THE HEAT EQUATION AND SHARP INEQUALITIES

GIUSEPPE TOSCANI [a*]

ABSTRACT. The heat equation represents a powerful instrument to obtain a number of mathematical inequalities in sharp form. This maybe not so well-known property goes back more or less to half a century ago, when independently from each others, researchers from information theory [22, 6] and kinetic theory [20] established a useful connection between Boltzmann's H-functional and Fisher information exactly by means of the solution to the heat equation. In this note, we briefly discuss these original ideas, together with some new application.

1. Introduction

In the years between the late fifties to mid sixties the solution to the heat equation started to be used as a powerful instrument to connect Lyapunov functionals. To our knowledge, the first application of this idea can be found in two papers by Stam [22] and Linnik [19], published in the same year and concerned with two apparently disconnected arguments. Stam [22] was motivated by the finding of a rigorous proof of Shannon's entropy power inequality [21], and made a substantial use of the so-called DeBruijn's identity, which, connects Shannon's measure of information of a probability density function $f(x)$ of a random variable X (Boltzmann H-functional up to a change of sign)

$$H(X) = H(f) = -\int_{\mathbb{R}^n} f(v) \log f(v)\, dv, \tag{1}$$

with the Fisher information of a random variable with density

$$I(X) = I(f) = \int_{\mathbb{R}^n} \frac{|\nabla f(v)|^2}{f(v)}\, dv. \tag{2}$$

Linnik [19] used the information measures of Shannon and Fisher in a proof of the central limit theorem of probability theory.

Work written within the activities of the National Group of Mathematical Physics of INDAM (National Institute of High Mathematics). The author was partially supported by MIUR project "Optimal mass transportation, geometrical and functional inequalities with applications".

DeBruijn's identity is obtained starting from the heat equation [13]. Indeed, if $f(v,t)$ denotes the solution to the heat equation

$$\frac{\partial f(v,t)}{\partial t} = \Delta f(v,t), \tag{3}$$

integration by parts immediately leads to the relationship

$$I(f(t)) = \frac{d}{dt} H(f(t)), \quad t > 0. \tag{4}$$

Note that for $t \geq 0$, the solution to the heat equation (3) can be written as $f(v,t) = f * M_{2t}(v)$, where as usual $*$ denotes convolution, and $M_t(v)$ is the Gaussian density in $\mathbb{R}^n$ of variance nt

$$M_t(v) = \frac{1}{(2\pi t)^{n/2}} \exp\left(\frac{|v|^2}{2t}\right). \tag{5}$$

Some years later, McKean [20] used DeBruijn's identity in the context of kinetic theory of rarefied gases, to study the convergence towards equilibrium of the one-dimensional Kac caricature of Maxwellian molecules. It is interesting to remark that McKean was aware of the work of Linnik, in reason of the fact that he refers to Fisher's information (2) as to Linnik's functional. The analysis of McKean pushes further the study of the subsequent derivatives of Shannon's measure of information. Aiming in proving the old conjecture that subsequent derivatives of Boltzmann's H-functional alternate in sign, he went to consider not only DeBruijn's identity (4), but also the derivative of Fisher's information along the solution to the heat equation

$$J(f(t)) = -\frac{d}{dt} I(f(t)). \tag{6}$$

For this functional McKean was able to prove that it satisfies a lower bound which involves Fisher information [20]. This result was used in more recent times to give alternative proofs both of logarithmic Sobolev inequality [23], and of concavity of entropy power [25].

In all of these papers, it appears clearly the reason of the importance of the connection, through the solution to the heat equation, between Shannon's and Fisher's measures of information. Due to its quadratic nature, most of the inequalities concerned with the logarithmic entropy are easier to prove by means of Fisher's functional.

Also, a careful reading of Stam's proof [22] allows to conclude that, even if not explicitly written, the proof is based on the time-monotonicity of a functional which is invariant with respect to the scaling

$$f(v) \to f_a(v) = a^n f(av), \quad a > 0, \tag{7}$$

which preserves the total mass of the probability density function f. This property allows to reckon the (bounded) limit value of the underlying functional, as time goes to infinity.

In what follows, we briefly resume how this idea of using heat equation, coupled with the scaling invariance, has been seminal to obtain both new inequalities, or new proofs of known inequalities, which have the advantage to be well motivated from a physical point of view.

2. Boltzmann H-functional and Gibbs's lemma

A simple example of the use of the heat equation to get inequalities for entropies is concerned with Shannon's entropy (1). Thanks to DeBruijn's identity (4), the derivative of Shannon's entropy is positive, and it converges to infinity as time goes to infinity. Indeed, Shannon's entropy is not scaling invariant, since

$$H(f_a) = H(f) - n \log a \tag{8}$$

and one of the two ingredients in Stam's proof of the entropy power inequality is missed. Clearly, there are various ways to obtain the scaling invariance of H by adding or multiplying it by suitable quantities. We resort here to the second moment of f. It is easily checked that the second moment of a probability density function scales according to

$$E(f_a) = \int_{\mathbb{R}^n} |v|^2 f_a(v)\, dv = \frac{1}{a^2} E(f). \tag{9}$$

Hence, if the probability density has bounded second moment, a scaling invariant functional is obtained by coupling Shannon's entropy of f with the logarithm of the second moment of f

$$\Lambda_f(t) = H(f(t)) - \frac{n}{2} \log E(f(t)), \tag{10}$$

where $f(v,t)$ solves the heat equation (3). For this functional, in fact, $\Lambda_f(t) = \Lambda_{f_a}(t)$. If we compute the time derivative of $\Lambda_f(t)$, we obtain

$$\frac{d}{dt}\Lambda_f(t) = I(f(t)) - \frac{n^2}{E(f(t))}, \tag{11}$$

which is a direct consequence of DeBruijn's identity (4), and of the time evolution of the second moment of the solution to the heat equation,

$$\frac{d}{dt}E(f(t)) = \frac{d}{dt}\int_{\mathbb{R}^n} |v|^2 f(v)\, dv = 2n \int_{\mathbb{R}^n} f(v)\, dv = 2n.$$

The right-hand side of (11) depends of Fisher's information only, and it is nonnegative. This can be easily shown by an argument which is often used in this type of proofs. One obtains

$$0 \le \int_{\mathbb{R}^n} \left(\frac{\nabla f(v)}{f(v)} + \frac{nv}{E(f(v))} \right)^2 f(v)\, dv =$$

$$I(f) + \frac{n^2}{E(f)^2} \int_{\mathbb{R}^n} |v|^2 f(v)\, dv + 2\frac{n}{E(f(v))} \int_{\mathbb{R}^n} v \cdot \nabla f(v)\, dv =$$

$$I(f) + \frac{n^2}{E(f)} - 2\frac{n^2}{E(f)} = I(f) - \frac{n^2}{E(f)}. \tag{12}$$

Note that, since $f(t)$ is the (smooth) solution to the heat equation, equality to zero in (12) holds if and only if

$$\frac{\nabla f(v)}{f(v)} + \frac{nv}{E(f(v))} = 0$$

for all $v \in \mathbb{R}^n$. This condition can be rewritten as

$$\nabla \left(\log f(v) - \frac{n}{E(f(v))} \frac{v^2}{2} \right) = 0 \tag{13}$$

which identifies the probability density $f(v)$ as a Gaussian density in $\mathbb{R}^n$. By (12), this also shows that, among all densities with the same second moment, Fisher's information takes its minimum value in correspondence to the Gaussian density

$$M_\sigma(v) = \frac{1}{(2\pi\sigma)^{n/2}} \exp\left(\frac{|v|^2}{2\sigma}\right), \tag{14}$$

where $\sigma = E(f)/n$.

Thus, unless the initial value is a Gaussian density, the functional $\Lambda(t)$ is monotone increasing, and it will reach its eventual maximum value as time $t \to \infty$. The computation of the limit value uses in a substantial way the scaling invariance of Λ. In fact, at each time $t > 0$, the value of $\Lambda_f(t)$ does not change if we scale $f(v,t)$ according to

$$f(v,t) \to F(v,t) = \left(\sqrt{1+2t}\right)^n \, f(v\sqrt{1+2t}, t). \tag{15}$$

On the other hand, it is well-known that [9]

$$\lim_{t\to\infty} F(v,t) = M_1(v) \int_{\mathbb{R}^n} f(v)\, dv \tag{16}$$

where, according to (5) $M_1(x)$ is the Gaussian density in $\mathbb{R}^n$ of variance equal to n. Therefore, passing to the limit one obtains

$$\Lambda_f(0) = H(f) - \frac{n}{2}\log E(f) \le \Lambda_{M_1} = \frac{n}{2}\log\frac{2\pi e}{n}. \tag{17}$$

This inequality holds for all probability density functions, and does not require that the second moment of f equals the second moment of the Gaussian density. If this is the case, we obtain the well-known bound (Gibbs's lemma)

$$H(f) \le \frac{n}{2}\log 2\pi e.$$

3. The entropy power inequality

Historically, the first proof of an inequality which has been obtained by means of the connection between Shannon's and Fisher's measures of information is the entropy power inequality [21]. In its original version, Shannon's entropy power inequality gives a lower bound on Shannon's entropy functional of the sum of independent random variables X, Y with densities

$$\exp\left(\frac{2}{n}H(X+Y)\right) \ge \exp\left(\frac{2}{n}H(X)\right) + \exp\left(\frac{2}{n}H(Y)\right), \quad n \ge 1, \tag{18}$$

with equality if X and Y are Gaussian random variables.

The entropy-power

$$N(X) = N(f) = \exp\left(\frac{2}{n}H(X)\right)$$

(variance of a Gaussian random variable with the same Shannon's entropy functional) is maximum and equal to the variance when the random variable is Gaussian, and thus, the essence of (18) is that the sum of independent random variables tends to be *more Gaussian* than one or both of the individual components.

The first rigorous proof of inequality (18) was given by Stam [22] (see also Blachman [6] for the generalization to n-dimensional random vectors), and was based on identity (4) which couples Fisher's information with Shannon's entropy functional [12].

In more details, the proof of Stam is based on the following argument. Let $f(v,t)$ and $g(v,t)$ be two solutions of the heat equation (3) corresponding to the initial data $f(v)$ (respectively $g(v)$) and to different diffusion coefficients, that can in general depend of time. If the entropies of the initial data are finite, one considers the time-evolution of the functional $\Theta_{f,g}(t)$ defined by

$$\Theta_{f,g}(t) = \frac{\exp\{\frac{2}{n}H(f(t))\} + \exp\{\frac{2}{n}H(g(t))\}}{\exp\{\frac{2}{n}H(f(t) * g(t))\}}. \tag{19}$$

Evaluating the time derivative of $\Theta_{f,g}(t)$, and using a key inequality for Fisher's information on convolutions, shows that $\Theta_{f,g}(t)$ is increasing in time, and converges towards the constant value $\Theta_{f,g}(+\infty) = 1$, thus proving inequality (18). Note that this method of proof also determines the cases of equality in (18).

It is interesting to remark that the evaluation of the limit of $\Theta_{f,g}(t)$, as $t \to \infty$, is made easy in reason of the scaling property. Indeed, the (Lyapunov) functional $\Theta(f,g)$ is invariant with respect to the scaling (7), which preserves the total mass of the function f.

The proof by Stam is a *physical* proof, in the spirit of Boltzmann H-theorem [10] in kinetic theory of rarefied gases, where convergence towards the Maxwellian equilibrium is shown in consequence of the monotonicity in time of the logarithmic entropy (1).

In reason of (4) it can be easily checked that the time derivative of the functional (19) depends of Fisher's information, and its sign can be controlled by owing to a property which is typical of Fisher's functional, namely the so-called Blackman-Stam inequality. For any given positive constants a and b, Fisher's information of the convolution of two probability densities f and g is controlled by the Fisher's information of f and g, and

$$(a+b)^2 I(f * g) \le a^2 I(f) + b^2 I(g). \tag{20}$$

Moreover there is equality in (20) if and only if both f and g are Gaussian densities, of variances proportional to a and b, respectively. The proof of (20) is simple, and can be found in [6]. Hence, like in (12), an inequality for Fisher's information leads to the entropy power inequality.

4. The concavity of entropy power

Variations of the entropy–power inequality are present in the literature. Costa's strengthened entropy–power inequality [11], in which one of the variables is Gaussian, and a generalized inequality for linear transforms of a random vector due to Zamir and Feder [26].

Also, other properties of Shannon's entropy-power $N(f)$ have been investigated so far. In particular, the *concavity of entropy power* theorem, which asserts that

$$\frac{d^2}{dt^2}\left(N(f * M_t)\right) \le 0 \tag{21}$$

provided that $f * M_t$ denotes the solution at time t to the heat equation. Inequality (21) is due to Costa [11]. Later, the proof has been simplified in [14, 15], by an argument based on the Blachman-Stam inequality [6]. More recently, a short and simple proof has been obtained by Villani [25], resorting to the functional $J(t)$ considered by McKean [20], we defined in (6).

The proof of concavity requires to evaluate two time derivatives of the entropy power, along the solution to the heat equation. The first derivative of the entropy power is easily evaluated resorting to DeBruijn's identity (4)

$$\frac{d}{dt}N(f(t)) = \frac{2}{n}\exp\left\{\frac{2}{n}H(f(t))\right\}\frac{d}{dt}H(f(t)) = \frac{2}{n}\exp\left\{\frac{2}{n}H(f(t))\right\} I(f(t)).$$

Let us set

$$\Upsilon_f(t) = \exp\left\{\frac{2}{n}H(f(t))\right\} I(f(t)). \tag{22}$$

Since the H-functional scales according to (8), while Fisher's information scales according to

$$I(f_a) = \int_{\mathbb{R}^n} \frac{|\nabla f_a(v)|^2}{f_a(v)}\, dv = a^2 \int_{\mathbb{R}^n} \frac{|\nabla f(v)|^2}{f(v)}\, dv = a^2 I(f), \tag{23}$$

the functional $\Upsilon_f(t)$ is invariant respect to the scaling (7). Consequently, the concavity of entropy power can be rephrased as the decreasing in time property of the scaling invariant functional $\Upsilon_f(t)$. Using definition (6) we obtain

$$\frac{d}{dt}\Upsilon_f(t) = \exp\left\{\frac{2}{n}H(f(t))\right\}\left(\frac{dI(f(t))}{dt} + \frac{2}{n}I(f(t))^2\right) =$$

$$\exp\left\{\frac{2}{n}H(f(t))\right\}\left(-J(f(t)) + \frac{2}{n}I(f(t))^2\right).$$

Hence, $\Upsilon_f(t)$ is non increasing if and only if

$$J(f(t)) \geq \frac{2}{n}I(f(t))^2. \tag{24}$$

In one dimension, inequality (24) is essentially due to McKean [20]. Let us repeat his computations. In the one dimensional case one obtains

$$I(f) = \int_{\mathbb{R}} \frac{f'(v)^2}{f(v)}\, dv,$$

while

$$J(f) = 2\left(\int_{\mathbb{R}} \frac{f''(v)^2}{f(v)}\, dv - \frac{1}{3}\int_{\mathbb{R}} \frac{f'(v)^4}{f(v)^3}\, dv\right). \tag{25}$$

McKean observed that $J(f)$ is positive. In fact, resorting to integration by parts, $J(f)$ can be rewritten as

$$J(f) = 2\int_{\mathbb{R}} \left(\frac{f''(v)}{f(v)} - \frac{f'(v)^2}{f(v)^2}\right)^2 f(v)\, dv \geq 0. \tag{26}$$

Having this formula in mind, consider that, for any constant $\lambda > 0$

$$0 \leq 2\int_{\mathbb{R}} \left(\frac{f''(v)}{f(v)} - \frac{f'(v)^2}{f(v)^2} + \lambda\right)^2 f(v)\, dv =$$

$$J(f) + 2\lambda^2 + 4\lambda \int_{\mathbb{R}} \left(f''(v) - \frac{f'(v)^2}{f(v)}\right) dv = J(f) + 2\lambda^2 - 4\lambda I(f).$$

Choosing $\lambda = I(f)$ shows (24) for $n = 1$. The same argument was used by Villani [25] to obtain (24) for $n > 1$. Once more, it is important to remark that equality in (24) holds if and only if f is a Gaussian density. In fact, the condition

$$\frac{f''(v)}{f(v)} - \frac{f'(v)^2}{f(v)^2} + \lambda = 0,$$

can be rewritten as

$$\frac{d^2}{dv^2} \log f(v) = -\lambda,$$

which corresponds to

$$\log f(v) = -\lambda v^2 + bv + c. \tag{27}$$

Joining condition (27) with the fact that $f(v)$ has to be a probability density, we conclude.

As before, unless the initial value is a Gaussian density, the functional $\Upsilon(t)$ is monotone decreasing, and it will reach its eventual minimum value as time $t \to \infty$. Once again, the computation of the limit value uses in a substantial way the scaling invariance property. Passing to the limit one obtains

$$\Upsilon_f(0) = \exp\left\{\frac{2}{n} H(f)\right\} I(f) \geq \Upsilon_{M_1} = 2n\pi e. \tag{28}$$

Inequality (28) is know under the name of *Isoperimetric Inequality for Entropies* (cfr. [15] for a different proof).

5. Hölder's inequality revisited

From the results of the previous Sections, one may get the impression that the solution to the heat equation is useful to obtain inequalities only in the particular cases in which Shannon's entropy and its variants are involved. As we shall see, this is not the case. Indeed, other classical inequalities, which are completely unrelated with entropy, can derived by methods similar to the ones used before. This is possible, for example for the classical Hölder's inequality, as well as for Young's inequality for convolutions, we will briefly treat in the next Section.

Without loss of generality, let $0 \leq f(v) \in L^1(\mathbb{R}^n)$ (respectively $0 \leq g(v) \in L^1(\mathbb{R}^n)$), and let $f(v,t)$ and $g(v,t)$ be the solutions to the heat equation (3) corresponding to the initial values f and g, respectively. If p and q are conjugate exponents, $1/p + 1/q = 1$, we consider the functional

$$\Phi_{u,v}(t) = \int_{\mathbb{R}} f(v,t)^{1/p} g(v,t)^{1/q}\, dv, \tag{29}$$

Note that this functional is invariant with respect to the scaling (7). It is only matter of simple computation to show that (29) is increasing in time from

$$\Phi_{f,g}(t=0) = \int_{\mathbb{R}^n} f(v)^{1/p} g(v)^{1/q}\, dv,$$

to

$$\lim_{t\to\infty} \Phi_{f,g}(t) = \left(\int_{\mathbb{R}^n} f(v)\, dv\right)^{1/p} \left(\int_{\mathbb{R}^n} g(v)\, dv\right)^{1/q}.$$

The detailed computation can be found in [24]. Clearly, Hölder's inequality follows by choosing $F(v) = f(v)^{1/p}$ and $G(v) = g(v)^{1/q}$. As for the previous inequalities, the proof via the heat equation allows to obtain automatically the cases of equality $(F(v) = cG(v))$. Also, the proof is in this case longer than the usual one, based on Young's inequality for constants. However, it shows that Hölder's inequality expresses the tendency of the solutions to the heat equation to converge towards the self-similar Gaussian solution.

6. Young's inequality for convolutions

Last, we will deal with Young's inequality for convolutions. In the sharp form obtained by Bechner [2] Young's inequality reads

$$\|f * g\|_r \le (A_p A_q A_{r'})^n \|f\|_p \|g\|_q. \tag{30}$$

In (30) $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$, $1 < p, q, r < \infty$ and $1/p + 1/q = 1 + 1/r$. Moreover, the constant A_m which defines the sharp constant is given by

$$A_m = \left(\frac{m^{1/m}}{m'^{1/m'}} \right)^{1/2} \tag{31}$$

where primes always denote dual exponents, $1/m + 1/m' = 1$.

The best constants in Young's inequality were found by Beckner [2], using tensorisation arguments and rearrangements of functions. In [7], Brascamp and Lieb derived them from a more general inequality, which is nowadays known as the Brascamp-Lieb inequality. The expression of the best constant, in the case in which both f and g are probability density functions, is obtained by noticing that inequality (30) is saturated by Gaussian densities. This principle has been largely utilized by Lieb in a more recent paper [18]. Among many other results, this paper contains a new proof of the Brascamp-Lieb inequality. In [7], Brascamp and Lieb noticed that the sharp form of Young's inequality also holds in the so-called reverse case

$$\|f * g\|_r \ge (A_p A_q A_{r'})^n \|f\|_p \|g\|_q, \tag{32}$$

where now $0 < p, q, r < 1$ while, as in Young's inequality (30), $1/p + 1/q = 1 + 1/r$. In this case, however, the dual exponents p', q', r' are negative, and

$$A_m = \left(\frac{m^{1/m}}{|m'|^{1/|m'|}} \right)^{1/2}. \tag{33}$$

The proof of this sharp reverse Young's inequality was subsequently simplified by Barthe [1]. While the original proof in [7] was rather complicated, and used tensorisation, Schwarz symmetrization, Brunn-Minkowski and some not so intuitive phenomenon for the measure in high dimension, the new proof in [1] was based on relatively more elementary arguments and gave a unified treatment of both cases, the Young's inequality (30) and its reverse form (32). As a matter of fact, the proof of the main result in [1] relies on a parametrization of functions which was used in [16] and was suggested by Brunn's proof of the Brunn-Minkowski inequality.

In a recent paper, Young's inequality has been seen in a different light by Bennett and Bez [4] (cfr. also [3, 5, 8]). There, Young's inequality is derived by looking at suitable properties of the solution to the heat equation. Even if not explicitly mentioned in the paper, this idea links Young's inequality in sharp form with other inequalities, for which the proof exactly moved along the same idea.

The connections of the sharp form of Young's inequality with other inequalities has been enlightened by Lieb in [17]. He proved in fact that, by letting $p, q, r \to 1$ in (30), the sharp form of Young's inequality reduces to Shannon's entropy power inequality discussed in Section **3**.

Motivated by this connection, a physical proof of Young's inequality, along the line drawn by Blachman and Stam [6, 22] has been given in [24]. Once more, the starting point is the study of the evolution in time of a Lyapunov functional, invariant under scaling. In this case the key functional to study is the one considered by Bennett and Bez [4]

$$\Psi_{f,g}(t) = \left(\int_{\mathbb{R}^n} \left(f(v,t)^{1/p} * g(v,t)^{1/q} \right)^r \, dv \right)^{1/r}, \tag{34}$$

where, as in Young's inequality, $1/p + 1/q = 1 + 1/r$. Similarly to what happens in the proof of the entropy power inequality $f(v,t)$ and $g(v,t)$ are solutions of the heat equation corresponding to the initial data $f(v)$ (respectively $g(v)$). However, these solutions correspond to two different heat equations, with different coefficients of diffusions, say α and β. The proof is constructive, and allows to identify the values of the two (unique) diffusion constants α and β which render the functional $\Psi_{f,g}(t)$ monotonically increasing in time. As in the case of Hölder inequality, Young's inequality follows by proving that, for suitable values of α and β the functional is monotonically increasing from

$$\Psi_{f,g}(t=0) = \left(\int_{\mathbb{R}^n} \left(f(v)^{1/p} * g(v)^{1/q} \right)^r \, dv \right)^{1/r},$$

to the limit value

$$\lim_{t\to\infty} \Psi_{f,g}(t) = (A_p A_q A_{r'})^{n/2} \left(\int_{\mathbb{R}^n} f(v)\, dv \right)^{1/p} \left(\int_{\mathbb{R}^n} g(v)\, dv \right)^{1/q}. \tag{35}$$

It is remarkable that the proof in [24] is based on the generalization of the inequality (20) for Fisher's information, which was the key point to prove the entropy power inequality. Hence, the pioneering studies of Stam [22], were seminal to obtain physical proofs of apparently only theoretical inequalities.

References

[1] F. Barthe, "Optimal Young's inequality and its converse: a simple proof", *Geom. Funct. Anal.* **8** (2), 234–242 (1998).

[2] W. Bechner, "Inequalities in Fourier Analysis on $\mathbb{R}^n$", *Proc. Nat. Acad. Sci. USA* **72**, (2), 638–641, (1975).

[3] J. Bennett, A. Carbery, M. Christ and T. Tao, "The Brascamp-Lieb Inequalities: Finiteness, Structure and Extremals", *Geometric And Functional Analysis* **17** (5), 1343–1415 (2007).

[4] J. Bennett and N. Bez, "Closure Properties of Solutions to Heat Inequalities", *J. Geom. Anal.* **19**, 584–600 (2009).

[5] J. Bennett, "Heat-flow monotonicity related to some inequalities in euclidean analysis", *Contemporary Mathematics* **505**, 85–96 (2010).

[6] N. M. Blachman, "The convolution inequality for entropy powers", *IEEE Trans. Inform. Theory* **2**, 267–271, (1965).

[7] H. J. Brascamp and E. H. Lieb, "Best constants in Young's inequality, its converse and its generalization to more than three functions", *Adx. Math.* **20**, 151–173 (1976).

[8] E. A. Carlen, E. H. Lieb, and M. Loss, "A sharp analog of Young's inequality on S^N and related entropy inequalities", *J. Geom. Anal.* **14** (3), 487–520 (2004).
[9] J. A. Carrillo and G. Toscani, "Exponential convergence toward equilibrium for homogeneous Fokker-Planck-type equations", *Math. Methods Appl. Sci.*, **21**, 1269-1286 (1998).
[10] C. Cercignani, *Theory and application of the Boltzmann equation* (Scottish Academic Press, Edinburgh, 1975).
[11] M. H. M. Costa, "A new entropy power inequality", *IEEE Trans. Inf. Theory*, **IT-31** (6), 751–760, (1985).
[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (J. Wiley & Sons, Hoboken, New Jersey, USA, 2006).
[13] J. Crank, *The Mathematics of Diffusion* (Clarendon Press, Oxford, UK 1956).
[14] A. Dembo, "A simple proof of the concavity of the entropy power with respect to the variance of additive normal noise", *IEEE Trans. Inform. Theory* **35**, 887–888 (1989).
[15] A. Dembo, T. Cover, and J. Thomas, "Information theoretic inequalities", *IEEE Trans. Inform. Theory* **37** (6), 1501–1518 (1991).
[16] R. Henstock and A.M. Macbeath, "On the measure of sum sets. (I) the theorems of Brunn, Minkowski and Lusternik", *Proc. London Math. Soc.* **3**, 182–194 (1953).
[17] E. H. Lieb, "Proof of an entropy conjecture of Wehrl", *Commun. Math. Phys.*, **62** (1), 35–41, (1978).
[18] E.H. Lieb, "Gaussian kernels have only gaussian maximizers", *Invent. Math.*, **102**, 179–208 (1990).
[19] Yu.V. Linnik, "An information-theoretic proof of the central limit theorem with the Lindeberg condition", *Theory Probab. Appl.* **4**, 288-299 (1959).
[20] H. P. McKean Jr., "Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas", *Arch. Rat. Mech. Anal.* **21**, 343–367, (1966).
[21] C. E. Shannon, "A mathematical theory of communication", *Bell Syst. Tech. J.* **27**, 379–423 Jul.(1948), 623–656, Oct. (1948).
[22] A.J. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon", *Inf. Contr.* **2**, 101–112, (1959).
[23] G. Toscani, "Sur l'inégalité logaritmique de Sobolev, *C.R. Acad. Sci. Paris*, **324** Série I, 689-694 (1997).
[24] G. Toscani, "Heat equation and the sharp Young's inequality", Preprint arXiv:1204.2086 (2012).
[25] C. Villani, "A short proof of the concavity of entropy power", *IEEE Trans. Info. Theory* **46** (4), 1695-1696, (2000).
[26] R. Zamir and M. Feder, "A generalization of the entropy power inequality with applications", *IEEE Trans. Inf. Theory*, **39** (5), 1723–1728, (1993).

$^{\alpha}$ Università degli Studi di Pavia
Dipartimento di Matematica
Via Ferrata 1
27100 Pavia, Italy

* Email: giuseppe.toscani@unipv.it

Article contributed to the Festschrift volume in honour of Giuseppe Grioli on the occasion of his 100th birthday.

Communicated 13 April 2012; published online 29 January 2013
