Noname manuscript No.
(will be inserted by the editor)

Heat equation and the sharp Young's inequality

Giuseppe Toscani

the date of receipt and acceptance should be inserted later

Abstract We show that the sharp Young's inequality for convolutions first obtained by Bechner [2] and Brascamp-Lieb [7] can be derived from the monotone in time evolution of a Lyapunov functional of the convolution of two solutions to the heat equation, with different diffusion coefficients. Our proof is based on a suitable adaptation of an old idea of Stam [19] and Blachman [6], used to obtain Shannon's entropy power inequality.

Keywords Heat equation · Young's inequality

1 Introduction

The goal of this note is to present a new proof of the Young inequality in the sharp form obtained by Bechner [2],

$$\|f * g\|_r \le (A_p A_q A_{r'})^n \|f\|_p \|g\|_q. \tag{1}$$

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

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

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

G.Toscani
Department of Mathematics
University of Pavia
Via Ferrata 1,
27100 Pavia, Italy
Tel.: +39-0382-985640
Fax: +39-0382-985602
E-mail: giuseppe.toscani@unipv.it

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 (1) is saturated by Gaussian densities. This principle has been largely utilized by Lieb in a more recent paper [15]. 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 inequality also holds in the so-called reverse case

$$\|f * g\|_r \geq (A_p A_q A_{r'})^n \|f\|_p \|g\|_q, \tag{3}$$

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

$$A_m = \left(\frac{m^{1/m}}{|m'|^{1/|m'|}} \right)^{1/2}. \tag{4}$$

The proof of this sharp reverse Young 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 inequality (1) and its reverse form (3). As a matter of fact, the proof of the main result in [1] relies on a parametrization of functions which was used in [13] 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]. In their paper, Young's inequality is derived by looking at the closure properties, with respect to the so-called heat inequalities, of certain functionals of the solution to the heat equation. Even if not explicitly mentioned in the paper, this idea connects 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 inequality with other inequalities has been enlightened by Lieb in [14]. He proved in fact that, by letting $p, q, r \to 1$ in (1), the sharp form of Young's inequality reduces to another well-known inequality in information theory, known as Shannon's entropy power inequality [18].

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

$$\exp\left(2H(X+Y)\right) \geq \exp\left(2H(X)\right) + \exp\left(2H(Y)\right), \tag{5}$$

with equality if X and Y are Gaussian random variables. Shannon's entropy functional of the probability density function $f(x)$ of X is defined by

$$H(X) = H(f) = -\int_{I\!R} f(x) \log f(x)\, dx. \tag{6}$$

Note that Shannon's entropy functional coincides to Boltzmann's H-functional [10] up to a change of sign. The entropy-power

$$N(X) = N(f) = \exp\left(2H(X)\right)$$

(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 (5) 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 (5) was given by Stam [19] (see also Blachman [6] for the generalization to n-dimensional random vectors), and was based on an identity which couples Fisher's information with Shannon's entropy functional [12].

The original proofs of Blachman and Stam make a substantial use of the solution to the heat equation

$$\frac{\partial f(x,t)}{\partial t} = \Delta f(x,t), \tag{7}$$

that is, for $t \geq 0$, of the function $f(x,t) = f * M_{2t}(x)$, where $M_t(x)$ denotes the Gaussian density in $\mathbb{R}^n$ of variance t

$$M_t(x) = \frac{1}{(2\pi t)^{n/2}} \exp\left(\frac{|v|^2}{2t}\right). \tag{8}$$

Other 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 [22].

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

$$\frac{d^2}{dt^2}\left(N(f * M_{2t})\right) \leq 0 \tag{9}$$

Inequality (9) is due to Costa [11]. More recently, a short and simple proof of (9) has been obtained by Villani [21], using an old idea by McKean [17].

Summarizing, the proof of Stam is based on the following argument. Let $f(x,t) = f * M_{\gamma(t)}$ and $g(x,t) = g * M_{\eta(t)}$ be two solutions of the heat equation (7) corresponding to the initial data $f(x)$ (respectively $g(x)$), with $\gamma(t)$ and $\eta(t)$ increasing functions of time. If the entropies of the initial data are finite, one considers the evolution in time of the functional $\Theta_{f,g}(t)$ defined by

$$\Theta_{f,g}(t) = \frac{\exp\{2H(f(t))\} + \exp\{2H(g(t))\}}{\exp\{2H(f(t) * g(t))\}}. \tag{10}$$

Evaluating the time derivative of $\Theta_{f,g}(t)$, and using a key inequality for Fisher information on convolutions, shows that, for a particular choice of the functions $\gamma(t)$ and $\eta(t)$, $\Theta_{f,g}(t)$ is increasing in time, and converges towards the constant value $\Theta_{f,g}(+\infty) = 1$, thus proving inequality (5). Note that this method of proof also determines the cases of equality in (5).

It is fundamental to remark that the evaluation of the limit of $\Theta_{f,g}(t)$, as $t \to \infty$, is made easy in reason of a scaling property. Indeed, the (Lyapunov) functional $\Theta(f,g)$ is invariant with respect to the scaling (dilation)

$$f(x) \to f_a(x) = \frac{1}{a^n} f\left(\frac{x}{a^n}\right), \quad a > 0, \tag{11}$$

which preserves the total mass of the function f. The importance of this property will be clarified later on.

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 (6).

In the rest of the paper, inspired by the Stam's approach to the proof of Shannon's entropy power inequality, we will present a *physical* proof of both direct and reverse Young's inequalities, which is based on the two ingredients specified above: a suitable use of two solutions to the heat equation, corresponding to different coefficients of diffusion, coupled with the scaling invariance property (11). Our proof is alternative to the proof of [4], and relies on a result which generalizes Stam proof of subadditivity of Fisher information. Moreover, as one learns form Stam's proof of Shannon's entropy power, where the exact form of the functions $\gamma(t)$ and $\eta(t)$ is found in a constructive way, our proof of Young inequality is also constructive, in that it characterizes in a clear and direct way the unique possible set of values the two coefficients of diffusion can assume.

In the following Section, we will describe how the method works by proving Hölder inequality. Even if this is a well-known result [8, **?**, **?**], it will give indication on the underlying methodology. Next, Section 3 will contain our main result.

To make computations as simple as possible, we will present all proofs in one dimension. Without loss of generality, in fact, one can easily argue that identical proofs hold in dimension n, with $n > 1$. Also, if not strictly necessary, we will restrict ourselves to consider as functions probability density functions. Since our computations involve solutions to the heat equation, all details involving regularity and the various integrations by parts can be assumed to hold true.

The basic idea used here is that many inequalities can be viewed as the tendency of various Lyapunov functionals of the solution to the heat equation to reach their extremal values as time tends to infinity. The discovery of a Lyapunov functional which allows to prove Young inequality is only one of the possible application of this idea [3, **?**, **?**]. In a forthcoming paper [20], we are going to develop this strategy by revisiting various well-known inequalities in terms of the monotonicity of suitable Lyapunov functionals.

These results do justice to the ideas of various researchers who worked in information theory in the last fifty years, with sharp results which are not so well-known to the audience of mathematicians.

2 Heat equation and Hölder inequality

We begin by showing that Hölder's inequality can be viewed as a consequence of the time monotonicity of a suitable Lyapunov functional of the solution to the heat equation [8,4].

Hölder's inequality for integrals states that, if $p, q > 1$ are such that $1/p + 1/q = 1$

$$\int_{I\!R} |f(x)g(x)|\, dx \le \left(\int_{I\!R} |f(x)|^p\, dx\right)^{1/p} \left(\int_{I\!R} |g(x)|^q\, dx\right)^{1/q}. \tag{12}$$

Moreover, there is equality in (12) if and only f and g are such that there exist positive real numbers a and b such that $af^p(x) = bg^q(x)$ almost everywhere. Hölder's inequality can be proven in many ways, for example resorting to Young's inequality for constants, which states that, if $1/p + 1/q = 1$

$$cd \le \frac{c^p}{p} + \frac{d^q}{q}, \tag{13}$$

for all nonnegative c and d, where equality is achieved if and only if $c^p = d^q$.

Without loss of generality, one can assume that the functions f, g in (12) are nonnegative. A different way to achieve inequality (12) is contained into the following

Theorem 1 *Let $\Phi_{u,v}(t)$ be the functional*

$$\Phi_{u,v}(t) = \int_{I\!R} u(x,t)^{1/p} v(x,t)^{1/q}\, dx, \tag{14}$$

where $1/p + 1/q = 1$, and $u(x,t)$ and $v(x,t)$, $t > 0$, are solutions to the heat equation corresponding to the initial values $u(x) \in L^1(I\!R)$ (respectively $v(x) \in L^1(I\!R)$). Then $\phi_{u,v}(t)$ is increasing in time from

$$\Phi_{u,v}(t = 0) = \int_{I\!R} u(x)^{1/p} v(x)^{1/q}\, dx,$$

to

$$\lim_{t\to\infty} \Phi_{u,v}(t) = \left(\int_{I\!R} u(x)\, dx\right)^{1/p} \left(\int_{I\!R} v(x)\, dx\right)^{1/q}.$$

Proof We outline that the functional $\Phi_{u,v}(t)$ is invariant with respect to the mass preserving scaling (11). Moreover, the condition $u(x), v(x) \in L^1(I\!R)$ is enough to ensure that $\Phi_{u,v}(t) \in L^1(I\!R)$ at any time $t \ge 0$. Indeed, inequality (13) implies

$$u(x,t)^{1/p} v(x,t)^{1/q} \le \frac{1}{p} u(x,t) + \frac{1}{q} v(x,t),$$

where, since $u(x,t)$ and $v(x,t)$ are solution to the heat equation,

$$\int_{I\!R} u(x,t)\,dx = \int_{I\!R} u(x)\,dx, \quad \int_{I\!R} v(x,t)\,dx = \int_{I\!R} v(x)\,dx.$$

Let us first proceed in a formal way. However, by resorting to the smoothness properties of the solution to the heat equation, all mathematical details can be rigorously justified.

Let us evaluate the time derivative of $\Phi(t)$. It holds

$$\Phi'_{u,v}(t) = \int_{I\!R} \left[(u(x,t)^{1/p})_t v(x,t)^{1/q} + u(x,t)^{1/p}(v(x,t)^{1/q})_t \right] dx =$$

$$\int_{I\!R} \left[\frac{1}{p} u^{1/p-1} v^{1/q} u_{xx} + \frac{1}{q} u^{1/p} v^{1/q-1} v_{xx} \right] dx =$$

$$\int_{I\!R} \left[\frac{1}{p} u^{-1/q} v^{1/q} u_{xx} + \frac{1}{q} u^{1/p} v^{-1/p} v_{xx} \right] dx.$$

Integrating by parts we end up with

$$\Phi'_{u,v}(t) = \frac{1}{pq} \int_{I\!R} u^{1/p} v^{1/q} \left[\left(\frac{u_x}{u}\right)^2 - 2\frac{u_x}{u}\frac{v_x}{v} + \left(\frac{v_x}{v}\right)^2 \right] dx =$$

$$\frac{1}{pq} \int_{I\!R} u^{1/p}(x,t) v^{1/q}(x,t) \left(\frac{u_x(x,t)}{u(x,t)} - \frac{v_x(x,t)}{v(x,t)} \right)^2 dx \ge 0. \tag{15}$$

Hence the functional $\Phi_{u,v}(t)$ is increasing in time. Note that the time derivative of the functional is equal to zero if and only if, for every $t > 0$

$$\frac{u_x(x,t)}{u(x,t)} - \frac{v_x(x,t)}{v(x,t)} = 0$$

for all points $x \in I\!R$. This condition can be rewritten as

$$\frac{d}{dx} \log \frac{u(x,t)}{v(v,t)} = 0.$$

Consequently $\Phi'(t) = 0$ if and only if

$$u(x,t) = c\, v(x,t) \tag{16}$$

for some positive constant c. Thus, unless condition (16) is verified almost everywhere at time $t = 0$, the functional $\Phi(t)$ 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 Φ. In fact, at each time $t > 0$, the value of $\Phi_{u,v}(t)$ does not change if we scale $u(x,t)$ and $v(x,t)$ according to

$$\begin{aligned} u(x,t) &\to U(x,t) = \sqrt{1+2t}\, u(x\sqrt{1+2t}, t) \\ v(x,t) &\to V(x,t) = \sqrt{1+2t}\, v(x\sqrt{1+2t}, t). \end{aligned} \tag{17}$$

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

$$\lim_{t\to\infty} U(x,t) = M_1(x) \int_{I\!R} u(x)\,dx \qquad \lim_{t\to\infty} V(x,t) = M_1(x) \int_{I\!R} v(x)\,dx, \qquad (18)$$

where, according to (8) $M_1(x)$ is the Gaussian density in $I\!R$ of variance equal to 1. Therefore, passing to the limit one obtains

$$\lim_{t\to\infty} \Phi_{u,v}(t) = \left(\int_{I\!R} u(x)\,dx\right)^{1/p} \left(\int_{I\!R} v(x)\,dx\right)^{1/q} \int_{I\!R} M_1(x)^{1/p} M_1(x)^{1/q}\,dx =$$
$$\left(\int_{I\!R} u(x)\,dx\right)^{1/p} \left(\int_{I\!R} v(x)\,dx\right)^{1/q} \int_{I\!R} M_1(x)\,dx =$$
$$\left(\int_{I\!R} u(x)\,dx\right)^{1/p} \left(\int_{I\!R} v(x)\,dx\right)^{1/q}.$$

Since

$$\lim_{t\to 0^+} \Phi_{u,v}(t) = \int_{I\!R} u(x)^{1/p} v(x)^{1/q}\,dx,$$

the monotonicity of the functional $\Phi(t)$ implies the inequality

$$\int_{I\!R} u(x)^{1/p} v(x)^{1/q}\,dx \le \left(\int_{I\!R} u(x)\,dx\right)^{1/p} \left(\int_{I\!R} v(x)\,dx\right)^{1/q}, \qquad (19)$$

with equality if and only if (16) is verified at time $t = 0$, that is

$$u(x) = cv(x), \qquad (20)$$

for some positive constant c. Setting $f = u^{1/p}$ and $g = v^{1/q}$ proves both Hölder inequality (12) and the equality cases.

Despite its apparent complexity, this way of proof is based on a solid physical argument, namely the monotonicity in time of a Lyapunov functional of the solution to the heat equation. This gives a clear indication that many inequalities reflect the physical principle of the tendency of a system to move towards the state of maximum entropy. In the next Section we will see how this idea applies to prove Young's inequality.

3 Young's inequality and Lyapunov functionals

The proof of the sharp Young's inequality follows along the same lines of the proof of Hölder's inequality we presented in Section 2. In this case the key functional to study is the one considered by Bennett and Bez [4]

$$\Psi_{u,v}(t) = \left(\int_{I\!R} \left(u(x,t)^{1/p} * v(x,t)^{1/q}\right)^r\,dx\right)^{1/r}, \qquad (21)$$

where, as in Young's inequality, $1/p+1/q=1+1/r$. With respect to the notations of the previous Section, there is a substantial difference in the meaning of the functions $u(x,t)$ and $v(x,t)$. Here $u(x,t)$ and $v(x,t)$ are still solutions of the heat equation corresponding to the initial data $u(x)$ (respectively $v(x)$). However, these solutions correspond to two different heat equations, with different coefficients of diffusions, say α and β. In other words, $u(x,t)$ solves the diffusion equation

$$u_t = \alpha u_{xx}, \tag{22}$$

while $v(x,t)$ solves

$$v_t = \beta v_{xx}. \tag{23}$$

Hence $u(x,t)$ and $v(x,t)$ diffuse at different velocities. It is a simple exercise to verify that, in view of the relationship between p,q and r, the functional $\Psi_{u,v}(t)$ is invariant with respect to the mass preserving scaling (11).

Theorem 2 *Let $\Psi_{u,v}(t)$ be the functional (21), where $1/p+1/q=1+1/r$, and $u(x,t)$ and $v(x,t)$, $t>0$, are solutions to the heat equation corresponding to the initial values $u(x)\in L^1(I\!R)$ (respectively $v(x)\in L^1(I\!R)$). Then, if $p,q,r>1$, and the diffusion coefficients in (22) and (23) are given by $\alpha=q'/p$ (respectively $\beta=p'/q$), or by a multiple of them, $\Psi_{u,v}(t)$ is increasing in time from*

$$\Psi_{u,v}(t=0) = \left(\int_{I\!R}\left(u(x)^{1/p} * v(x)^{1/q}\right)^r dx\right)^{1/r},$$

to the limit value

$$\lim_{t\to\infty}\Psi_{u,v}(t) = (A_pA_qA_{r'})^{1/2}\left(\int_{I\!R} u(x)\,dx\right)^{1/p}\left(\int_{I\!R} v(x)\,dx\right)^{1/q}. \tag{24}$$

If on the contrary $0<p,q,r<1$, and the diffusion coefficients in (22) and (23) are given by $\alpha=|q'|/p$ (respectively $\beta=|p'|/q$), or by a multiple of them, $\Psi_{u,v}(t)$ is decreasing in time from

$$\Psi_{u,v}(t=0) = \left(\int_{I\!R}\left(u(x)^{1/p} * v(x)^{1/q}\right)^r dx\right)^{1/r},$$

to the limit value (24), where now A_m is given by (4). In both cases $\Psi'_{u,v}(t)=0$ if and only if $u(x,t)$ and $v(x,t)$ are Gaussian functions.

Proof Let us consider first the case in which $p,q,r>1$. Without loss of generality, we will assume that both the initial data $u(x)$ and $v(x)$ are probability density functions. This is sufficient to show that, for any time $t>0$, the functional $\Psi_{u,v}(t)$ is bounded. Indeed we can write

$$\int_{I\!R} u(x-y)^{1/p}v(y)^{1/q}\,dy =$$

$$\int_{\{u(x-y)\le v(y)\}} u(x-y)^{1/p}v(y)^{1/q}\,dy + \int_{\{u(y)>v(x-y)\}} u(y)^{1/p}v(x-y)^{1/q}\,dy.$$

Now, since $r > 1$, and $v(x,t)$ has mass equal to 1, Jensen's inequality implies

$$\left(\int_{\{u(x-y)\le v(y)\}} u(x-y)^{1/p} v(y)^{1/q}\,dy\right)^r =$$

$$\left(\int_{\{u(x-y)\le v(y)\}} u(x-y)^{1/p} v(y)^{1/q-1} v(y)\,dy\right)^r \le$$

$$\int_{\{u(x-y)\le v(y)\}} u(x-y)^{r/p} v(y)^{r/q-r} v(y)\,dy =$$

$$\int_{\{u(x-y)\le v(y)\}} u(x-y)^{r/p} v(y)^{r/q-r+1}\,dy.$$

Note that

$$\frac{r}{p} + \frac{r}{q} - r + 1 = 2, \qquad \frac{r}{p} > 1.$$

Therefore, on the set $\{u(x-y) \le v(y)\}$, since the exponent of $v(y)$ is smaller than 1,

$$u(x-y)^{r/p} v(y)^{r/q-r+1} \le u(x-y)v(y). \tag{25}$$

Inequality (25) follows simply dividing by u^2. Therefore

$$\left(\int_{\{u(x-y)\le v(y)\}} u(x-y)^{1/p} v(y)^{1/q}\,dy\right)^r \le \int_{\{u(x-y)\le v(y)\}} u(x-y)v(y)\,dy \le 1.$$

Identical computations show that

$$\left(\int_{\{u(y)>v(x-y)\}} u(y)^{1/p} v(x-y)^{1/q}\,dy\right)^r \le \int_{\{u(x-y)\le v(y)\}} u(y)v(x-y)\,dy \le 1.$$

Therefore

$$\int_{I\!R} \left(u(x)^{1/p} * v(x)^{1/q}\right)^r \le 2c_r,$$

where c_r is the positive constant in the inequality

$$(a+b)^r \le c_r(a^r + b^r).$$

We proceed now to compute the time derivative of the functional $\Psi_{u,v}(t)$. To shorten, let us denote

$$h(x,t) = u(x,t)^{1/p} * v(x,t)^{1/q}. \tag{26}$$

Since $u(x,t)$ and $v(x,t)$ are solutions to the heat equation

$$\frac{\partial}{\partial t} h(x,t) = \frac{\partial}{\partial t} \int_{I\!R} u(x-y,t)^{1/p} v(y,t)^{1/q}\,dy =$$

$$\frac{1}{p}\int_{I\!R} u(x-y,t)^{1/p-1} u_t(x-y,t) v(y,t)^{1/q}\,dy +$$

$$\frac{1}{q}\int_{I\!R} u(x-y,t)^{1/p}v(y,t)^{1/q-1}v_t(y,t)\,dy =$$
$$\frac{\alpha}{p}\int_{I\!R} u(x-y,t)^{1/p-1}u_{xx}(x-y,t)v(y,t)^{1/q}\,dy +$$
$$\frac{\beta}{q}\int_{I\!R} u(x-y,t)^{1/p}v(y,t)^{1/q-1}v_{yy}(y,t)\,dy.$$

On the other hand, we have

$$\frac{\partial^2}{\partial x^2}h(x,t) = \int_{I\!R}\frac{\partial}{\partial x}\left(\frac{1}{p}u(x-y,t)^{1/p-1}u_x(x-y,t)\right)v(y,t)^{1/q}\,dy =$$
$$\frac{1}{p}\int_{I\!R} u(x-y,t)^{1/p-1}u_{xx}(x-y,t)v(y,t)^{1/q}\,dy +$$
$$\frac{1}{p}\left(\frac{1}{p}-1\right)\int_{I\!R} u(x-y,t)^{1/p-2}u_x^2(x-y,t)v(y,t)^{1/q}\,dy. \tag{27}$$

Hence

$$\frac{1}{p}\int_{I\!R} u(x-y,t)^{1/p-1}u_{xx}(x-y,t)v(y,t)^{1/q}\,dy =$$
$$\frac{\partial^2}{\partial x^2}h(x,t) + \frac{1}{pp'}\int_{I\!R} u(x-y,t)^{1/p}v(y,t)^{1/q}\left(\frac{u_x}{u}\right)^2(x-y,t)\,dy. \tag{28}$$

Analogous formula for the last integral in (27). Therefore we have

$$\frac{\partial}{\partial t}h(x,t) = (\alpha+\beta)\frac{\partial^2}{\partial x^2}h(x,t) +$$
$$\frac{\alpha}{pp'}\int_{I\!R} u(x-y,t)^{1/p}v(y,t)^{1/q}\left(\frac{u_x}{u}\right)^2(x-y,t)\,dy +$$
$$\frac{\beta}{qq'}\int_{I\!R} u(x-y,t)^{1/p}v(y,t)^{1/q}\left(\frac{v_x}{v}\right)^2(y,t)\,dy. \tag{29}$$

Making use of formula (29), we obtain

$$\frac{d}{dt}\int_{I\!R} h^r(x,t)\,dx = r\int_{I\!R} h(x,t)^{r-1}h_t(x,t)\,dx =$$
$$r(\alpha+\beta)\int_{I\!R} h^{r-1}(x,t)h_{xx}(x,t)\,dx +$$
$$\int_{I\!R} h^{r-1}(x,t)\left[\frac{\alpha}{pp'}\int_{I\!R} u(x-y,t)^{1/p}v(y,t)^{1/q}\left(\frac{u_x}{u}\right)^2(x-y,t)\,dy +\right.$$
$$\left.\frac{\beta}{qq'}\int_{I\!R} u(x-y,t)^{1/p}v(y,t)^{1/q}\left(\frac{v_x}{v}\right)^2(y,t)\,dy\right]dx.$$

Since it holds

$$\frac{\left(u^{1/p}\right)_x}{u^{1/p}} = \frac{1}{p}\frac{u_x}{u}, \qquad \frac{\left(v^{1/q}\right)_x}{v^{1/p}} = \frac{1}{q}\frac{v_x}{v} \tag{30}$$

we obtain

$$\frac{\alpha}{pp'}\int_{\mathbb{R}} u(x-y,t)^{1/p} v(y,t)^{1/q}\left(\frac{u_x}{u}\right)^2(x-y,t)\,dy =$$

$$\alpha\frac{p}{p'}\int_{\mathbb{R}}\frac{\left(u_x^{1/p}(x-y,t)\right)^2}{u^{1/p}(x-y)}v(y,t)^{1/q}\,dy,$$

and

$$\frac{\beta}{qq'}\int_{\mathbb{R}} u(x-y,t)^{1/p} v(y,t)^{1/q}\left(\frac{v_x}{v}\right)^2(y,t)\,dy =$$

$$\beta\frac{q}{q'}\int_{\mathbb{R}} u^{1/p}(x-y)\frac{\left(v_y^{1/q}(y,t)\right)^2}{v^{1/q}(y)}\,dy.$$

Finally

$$\frac{1}{r}\frac{d}{dt}\int_{\mathbb{R}} h^r(x,t)\,dx = -(\alpha+\beta)(r-1)\int_{\mathbb{R}} h^{r-2}(x,t)\left(h_x(x,t)\right)^2\,dx +$$

$$\alpha\frac{p}{p'}\int_{\mathbb{R}} h^{r-1}(x,t)A(u^{1/p},v^{1/q})(x,t)\,dx+\beta\frac{q}{q'}\int_{\mathbb{R}} h^{r-1}(x,t)B(u^{1/p},v^{1/q})(x,t)\,dx. \tag{31}$$

In (31) we defined

$$A(f,g)(x,t) = \int_{\mathbb{R}}\frac{\left(f_x(x-y,t)\right)^2}{f(x-y)}g(y,t)\,dy, \tag{32}$$

and

$$B(f,g)(x,t) = \int_{\mathbb{R}} f(x-y)\frac{\left(g_y(y,t)\right)^2}{g(y)}\,dy. \tag{33}$$

Since

$$\frac{d\Psi_{u,v}(t)}{dt} = \Psi_{u,v}(t)^{1-r}\,\frac{d}{dt}\int_{\mathbb{R}} h^r(x,t)\,dx,$$

the sign of the time derivative of the functional $\Psi_{u,v}(t)$ depends of the sign of the expression on the right-hand side of (31). In order to determine this sign, the following Lemma will be of paramount importance.

Lemma 3 *Let $f(x)$ and $g(x)$ be probability density functions such that both $A(f,g)$ and $B(f,g)$, given by (32) and (33), are well defined. Then, for all positive constants a, b and $r > 0$*

$$\left(a^2+b^2+2abr\right)\int_{\mathbb{R}}(f*g)^{r-2}\left((f*g)_x\right)^2\,dx \le$$

$$a^2\int_{\mathbb{R}}(f*g)^{r-1}A(f,g)\,dx + b^2\int_{\mathbb{R}}(f*g)^{r-1}B(f,g)\,dx. \tag{34}$$

Moreover, there is equality in (34) if and only if, for any positive constant c and constants m_1, m_2, f and g are Gaussian densities, $f(x) = M_{ca}(x-m_1)$ and $g(x) = M_{cb}(x-m_2)$.

Proof The proof will follow along the same lines of the analogous one for Fisher information, given by Blachman [6]. First of all, to easily justify computations, let us prove the lemma by considering smooth functions f and g. Then the proof for general f and g will follow owing to the convexity properties of A and B [16]. This can be easily done by considering $f * M_t$ and $g * M_t$ solutions to the heat equation for some $t > 0$. Let

$$k(x) = f * g(x).$$

Then, for any pair of positive constants a, b

$$(a+b)k'(x) = a \int_{I\!R} f'(x-y)g(y)\,dy + b \int_{I\!R} f(x-y)g'(y)\,dy.$$

Therefore

$$(a+b)\frac{k'(x)}{k(x)} = a \int_{I\!R} \frac{f'(x-y)}{f(x-y)}\frac{f(x-y)g(y)}{k(x)}\,dy + b \int_{I\!R} \frac{g'(y)}{g(y)}\frac{f(x-y)g(y)}{k(x)}\,dy =$$

$$\int_{I\!R} \left(a\frac{f'(x-y)}{f(x-y)} + b\frac{g'(y)}{g(y)} \right) d\mu_x(y),$$

where we denoted

$$d\mu_x(y) = \frac{f(x-y)g(y)}{k(x)}\,dy.$$

Note that, for every $x \in I\!R$, $d\mu_x$ is a unit measure on $I\!R$. Consequently, by Jensen's inequality

$$(a+b)^2 \left[\frac{k'(x)}{k(x)} \right]^2 = \left[\int_{I\!R} \left(a\frac{f'(x-y)}{f(x-y)} + b\frac{g'(y)}{g(y)} \right) d\mu_x(y) \right]^2 \le$$

$$\int_{I\!R} \left(a\frac{f'(x-y)}{f(x-y)} + b\frac{g'(y)}{g(y)} \right)^2 d\mu_x(y). \tag{35}$$

Hence, for every constant $r > 0$

$$(a+b)^2 \int_{I\!R} k^r(x) \left[\frac{k'(x)}{k(x)} \right]^2 dx \le$$

$$\int_{I\!R} k^r(x) \int_{I\!R} \left(a\frac{f'(x-y)}{f(x-y)} + b\frac{g'(y)}{g(y)} \right)^2 \frac{f(x-y)g(y)}{k(x)}\,dy\,dx =$$

$$\int_{I\!R} k^{r-1}(x) \left[a^2 \int_{I\!R} \frac{(f'(x-y))^2}{f(x-y)} g(y)\,dy + b^2 \int_{I\!R} \frac{(g'(y))^2}{g(y)} f(x-y)\,dy \right] dx +$$

$$2ab \int_{I\!R} k^{r-1}(x) \int_{I\!R} f'(x-y)g'(y)\,dy\,dx.$$

On the other hand,

$$\int_{I\!R} f'(x-y)g'(y)\,dy = k''(x),$$

so that

$$\int_{I\!R} k^{r-1}(x) \int_{I\!R} f'(x-y)g'(y)\,dy\,dx =$$

$$\int_{I\!R} k^{r-1}(x)k''(x)\,dx = -(r-1)\int_{I\!R} k^{r-2}(x)(k'(x))^2\,dx.$$

This concludes the proof of the lemma. The cases of equality are easily found resorting to the following argument. Equality follows if, after application of Jensen's inequality, there is equality in (35). On the other hand, for any convex function φ and unit measure $d\mu$ on the set Ω, equality in Jensen's inequality

$$\varphi(\int_\Omega f\,d\mu) \le \int_\Omega \varphi(f)\,d\mu$$

holds true if and only if f is constant, so that

$$f = \int_\Omega f\,d\mu.$$

In our case, this means that there is equality if and only if the function

$$a\frac{f'(x-y)}{f(x-y)} + b\frac{g'(y)}{g(y)}$$

does not depend on y. If this is the case, taking the derivative with respect to y, and using the identity

$$\frac{d}{dy}\left(\frac{f'(x-y)}{f(x-y)}\right) = -\frac{d}{dx}\left(\frac{f'(x-y)}{f(x-y)}\right),$$

we conclude that f and g have to satisfy

$$a\frac{d^2}{dx^2}\log f(x-y) = b\frac{d^2}{dy^2}\log g(y). \tag{36}$$

Note that (36) can be verified if and only if the functions on both sides are constant. Thus, there is equality if and only if

$$\log f(x) = b_1x^2 + c_1x + d_1, \quad \log g(x) = b_2x^2 + c_2x + d_2. \tag{37}$$

By coupling (37) with (36), we obtain that there is equality in (34) if and only if f and g are gaussian densities, of variances ca and cb, respectively, for any given positive constant c.

The case $r = 1$ has been treated in Blachman [6], as part of his proof of the entropy power inequality (5). In this case

$$I(f) = \int_{I\!R} \frac{(f'(x))^2}{f(x)}\,dx \tag{38}$$

denotes the Fisher information of the probability density f, and inequality (34) becomes

$$(a+b)^2 I(f*g) \le a^2 I(f) + b^2 I(g).$$

We remark that the validity of (34) is not restricted to probability density functions. Indeed, it continues to hold for nonnegative functions of any given mass.

Let us apply the result of Lemma 3 to control the sign of the right-hand side in formula (31). If we choose $a^2 = \alpha p/p'$, $b^2 = \beta q/q'$ in (34), then the coefficient of the term on the left-hand side of inequality (34) assumes the value

$$a^2 + b^2 + 2abr = \alpha \frac{p}{p'} + \beta \frac{q}{q'} + 2\sqrt{\alpha\beta}\sqrt{\frac{pq}{p'q'}} r.$$

Let us introduce, for any given $r > 1$ the function

$$\Gamma(\alpha, \beta) = (\alpha + \beta)(r - 1) - \left(\alpha \frac{p}{p'} + \beta \frac{q}{q'} + 2\sqrt{\alpha\beta}\sqrt{\frac{pq}{p'q'}} r \right). \tag{39}$$

It is clear that, as soon as for some values of α, β the function $\Gamma(\alpha, \beta) \le 0$, the expression on the right-hand side of (31) is nonnegative, and the functional $\Psi_{u,v}(t)$ is increasing. In order to check its sign, consider that the function Γ is jointly convex, and it is such that, for any positive constant c

$$\Gamma(c\alpha, c\beta) = c\Gamma(\alpha, \beta).$$

Therefore, if a point $(\alpha = \alpha_0, \beta = \beta_0)$ is an extremal point, also the point $(c\alpha_0, c\beta_0)$ is an extremal point, and Γ admits the half-line $\beta_0\alpha = \alpha_0\beta$ of extremals. Since

$$\frac{\partial \Gamma}{\partial \alpha} = r - 1 - \frac{p}{p'} - \sqrt{\frac{\alpha}{\beta}}\sqrt{\frac{pq}{p'q'}} r,$$

by adding and subtracting the quantity pr/q' we obtain

$$\frac{\partial \Gamma}{\partial \alpha} = r + \frac{p}{q'} r - 1 - \frac{p}{p'} - \sqrt{\frac{\alpha}{\beta}}\sqrt{\frac{pq}{p'q'}} r + \frac{p}{q'} r =$$

$$pr\left(\frac{1}{p} - \frac{1}{q'}\right) - p\left(\frac{1}{p} + \frac{1}{p'}\right) - \sqrt{\frac{\alpha}{\beta}}\sqrt{\frac{pq}{p'q'}} r + \frac{p}{q'} r.$$

Since

$$\frac{1}{p} - \frac{1}{q'} = \frac{1}{q} - \frac{1}{p'} = \frac{1}{r}, \tag{40}$$

one obtains

$$\frac{\partial \Gamma}{\partial \alpha} = -\sqrt{\frac{\alpha}{\beta}}\sqrt{\frac{pq}{p'q'}} r + \frac{p}{q'} r = 0$$

if the point (α, β) belong to the half-line

$$\beta = \frac{p}{q'} \cdot \frac{p'}{q} \alpha. \tag{41}$$

Same result is obtained if we impose the vanishing of the partial derivative of Γ with respect to β. On the other hand, thanks to identity (40)

$$\Gamma\left(\frac{q'}{p}, \frac{p'}{q}\right) = \left(\frac{q'}{p} + \frac{p'}{q}\right)(r - 1) - \frac{q'}{p'} - \frac{p'}{q'} - 2r =$$

$$q'r\left(\frac{1}{p}-\frac{1}{q'}\right)+p'r\left(\frac{1}{q}-\frac{1}{p'}\right)-q'-p'=0.$$

Hence, along the line (41), in view of lemma 3 the functional $\Phi_{u,v}(t)$ is increasing with respect to t. Proceeding as in Section 2, namely by scaling $u(x,t)$ and $v(x,t)$ as in (17), we conclude that the functional will keep its maximum value as time goes to infinity, and

$$\lim_{t\to\infty}\Phi_{u,v}(t)=\left(\int_{I\!R}u(x)\,dx\right)^{1/p}\left(\int_{I\!R}v(x)\,dx\right)^{1/q}C(p,q,r), \qquad (42)$$

where

$$C(p,q,r)=\left(\int_{I\!R}\left(M_{q'/p}(x)^{1/p}*M_{p'/q}(x)^{1/q}\right)^{r}\,dx.\right)^{1/r}. \qquad (43)$$

Using that the convolution of Gaussian functions is a Gaussian function, $M_\alpha * M_\beta = M_{\alpha+\beta}$, we compute

$$\left(\int_{I\!R}\left(M_\alpha^{1/p}*M_\beta^{1/q}\right)^{r}\,dx.\right)^{1/r}=\left[\frac{p\alpha}{\alpha^{1/p}}\frac{q\beta}{\beta^{1/q}}\frac{(p\alpha+q\beta)^{1/r}}{r^{1/r}(p\alpha+q\beta)}\right]^{1/2}.$$

The choice $\alpha=q'/p$, $\beta=p'/q$ gives

$$C(p,q,r)=\left(A_pA_qA_{r'}\right)^{1/2},$$

where A_m is defined by (2). This concludes the proof of the first part of Theorem 2.

The case in which $1/p+1/q=1+1/r$, but $0<p,q,r<1$ can be treated likewise. In this case the dual exponents p',q',r' are negative, and formula (31) takes the form

$$\frac{1}{r}\frac{d}{dt}\int_{I\!R}h^r(t)\,dx=-\left[-(\alpha+\beta)(1-r)\int_{I\!R}h^{r-2}(t)\left(h_x(t)\right)^2\,dx+\right.$$

$$\left.\frac{\alpha p}{|p'|}\int_{I\!R}h^{r-1}(t)A(u^{1/p},v^{1/q})(t)\,dx+\frac{\beta q}{|q'|}\int_{I\!R}h^{r-1}(t)B(u^{1/p},v^{1/q})(t)\,dx\right] \qquad (44)$$

To control the sign of the quantity into square brackets, we introduce now the function

$$\tilde{\Gamma}(\alpha,\beta)=(\alpha+\beta)(1-r)-\left(\alpha\frac{p}{|p'|}+\beta\frac{q}{|q'|}+2\sqrt{\alpha\beta}\sqrt{\frac{pq}{|p'||q'|}}r\right). \qquad (45)$$

In this case

$$\frac{\partial\Gamma}{\partial\alpha}=1-r+\frac{p}{p'}+\sqrt{\frac{\alpha}{\beta}}\sqrt{\frac{pq}{|p'||q'|}}r.$$

By adding and subtracting the quantity pr/q' we obtain as before

$$\frac{\partial\Gamma}{\partial\alpha}=\sqrt{\frac{\alpha}{\beta}}\sqrt{\frac{pq}{|p'||q'|}}r+\frac{p}{q'}r=0$$

if the point (α, β) belong to the half-line

$$\beta = \frac{p}{|q'|} \cdot \frac{|p'|}{q} \alpha. \tag{46}$$

This choice however implies that the right-hand side in (44) is non positive, and the functional $\Psi_{u,v}(t)$ decreases. This leads to the reverse Young's inequality (3).

4 Conclusions

In this paper we presented a new proof of the sharp form of Young's inequality for convolutions, as well as and its reverse form. For the sake of simplicity, this proof has been done in dimension $n = 1$. Looking at the details of the computations, it appears evident that the proof still holds in dimension $n > 1$, since the computations in higher dimension do not affect the constants in formulas (31) and (34), which are at the basis of the whole procedure. The main difference relays in the fact that the Gaussian functions are n-dimensional Gaussians, which lead to the additional presence of the exponent n in the sharp constant. Hence Theorem 2 leads to the sharp inequality (1) in any dimension, without any additional (if not computational) difficulty. Also, both Young's inequality and its reverse form are here derived by a unique well understandable physical principle, in the form of time monotonicity of a Lyapunov functional.

Acknowledgment: The author acknowledge support by MIUR project "Optimal mass transportation, geometrical and functional inequalities with applications".

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 $I\!R^n$. *Proc. Nat. Acad. Sci. USA* **72**, (2), 638–641, (1975).
3. J. Bennett, A. Carbery, M. Christ and T. Tao, The BrascampLieb 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 Youngs 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, 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. Wiley, 2006.
13. 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).
14. E.H. Lieb, Proof of an entropy conjecture of Wehrl. *Commun. Math. Phys.*, **62**, (1), 35–41, (1978).
15. E.H. Lieb, Gaussian kernels have only gaussian maximizers. *Invent. Math.*, **102** 179–208 (1990).
16. P.L. Lions, G. Toscani, A strenghened central limit theorem for smooth densities. *J. Funct. Anal.* **129** 148–167 (1995).
17. H.P. McKean Jr., Speed of approach to equilibrium for Kacs caricature of a Maxwellian gas. *Arch. Rat. Mech. Anal.* **21**, 343–367, (1966).
18. C.E. Shannon, A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 Jul.(1948), 623–656, Oct. (1948).
19. A.J. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Contr.* **2**, 101–112, (1959).
20. G. Toscani, Heat equation, Lyapunov functionals and inequalities. (preprint) (2012)
21. C. Villani A short proof of the concavity of entropy power. *IEEE Trans. Info. Theory* **46**, (4) 1695-1696, (2000).
22. R. Zamir and M. Feder, A generalization of the entropy power inequality with applications. *IEEE Trans. Inf. Theory*, **39**, (5) 1723–1728, (1993).