RTG PAPER - YOUNG'S INEQUALITY

DANIEL REICH

Preface

The analyses of sharp constants are important to study for two reasons. One is that they are occasionally useful, and even important. The second is that they provide us with good examples of "hard analysis" problems that can be carried to completion, which is usually not the case.
-Elliot H. Lieb & Michael Loss

This paper follows the proof of Young's inequality in *Analysis* by Lieb & Loss. Virtually all steps are shown in great detail. Calculations may at times seem repetative, but this was done to make the proof readable to students not yet overly experienced in the realm of analysis (much like myself at the present time).

Contents

1. YOUNG'S INEQUALITY

Theorem 1.1 (Young's Inequality)**.** *Let* $p, q, r \geq 1$ *and* $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$.

(i) *Let* $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$, *and* $h \in L^r(\mathbb{R}^n)$.

$$\begin{aligned} \left| \int_{\mathbb{R}^n} f(x)(g * h)(x) dx \right| &= \left| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x) g(x-y) h(y) dx dy \right| \\ &\leq C_{p,q,r;n} \|f\|_p \|g\|_q \|h\|_r \end{aligned} \tag{1}$$

where the sharp constant,

$$C_{p,q,r;n} = (C_p C_q C_r)^n, \text{ where } {C_p}^2 = \frac{p^{\frac{1}{p}}}{p'^{\frac{1}{p'}}} \text{ with } \frac{1}{p} + \frac{1}{p'} = 1 \tag{2}$$

(a) *note that* $C_p = \frac{1}{C_{p'}}$
(b) *also note that*

$$\frac{1}{p'} + \frac{1}{q'} + \frac{1}{r'} = 1 - \frac{1}{p} + 1 - \frac{1}{q} + 1 - \frac{1}{r} = 3 - \left(\frac{1}{p} + \frac{1}{q} + \frac{1}{r} \right) = 1 \tag{3}$$

(ii) *If p,q,r > 1, then equality can occur in* (1) *if and only if f,g and h are Gaussian functions*

$$\begin{aligned} f(x) &= A \exp[-p'(x-a, J(x-a)) + ik \cdot x], \\ g(x) &= B \exp[-q'(x-b, J(x-b)) - ik \cdot x], \\ h(x) &= C \exp[-r'(x-c, J(x-c)) + ik \cdot x], \end{aligned} \tag{4}$$

where $A, B, C \in \mathbb{C}$; $a, b, c, k \in \mathbb{R}^n$ *with* $a = b + c$; *and J is any real, symmetric, positive-definite matrix.*

2. BACKGROUND THEOREMS AND DEFINITIONS

Theorem 2.1 (Hölder's Inequality for $\mathbb{R}^n$)**.** *Let* $1 \leq p,q \leq \infty$ *s.t.* $\frac{1}{p} + \frac{1}{q} = 1$. *Let* $f \in L^p(\mathbb{R}^n)$ *and* $g \in L^q(\mathbb{R}^n)$. *Then the pointwise product, given by (fg)(x) = f(x)g(x), is in* $L^1(\mathbb{R}^n)$ *and*

$$\left| \int_{\mathbb{R}^n} f(x) g(x) dx \right| \leq \int_{\mathbb{R}^n} |f(x)||g(x)| dx \leq \|f\|_p \|g\|_q. \tag{5}$$

(i) The first inequality in (5) is an equality if and only if $f(x)g(x) = e^{i\theta}|f(x)||g(x)|$ for some real constant θ and for almost every x.

(ii) If $f \not\equiv 0$ the second inequality in (5) is an equality if and only if there is a constant $\lambda \in \mathbb{R}$ such that
  (a) If $1 < p < \infty$, $|g(x)| = \lambda |f(x)|^{p-1}$ for almost every x.
  (b) If $p = 1$, $|g(x)| \leq \lambda$ for almost every x and $|g(x)| = \lambda$ when $f(x) \not\equiv 0$.
  (c) If $p = \infty$, $|f(x)| \leq \lambda$ for almost every x and $|f(x)| = \lambda$ when $g(x) \not\equiv 0$.

The generalized Hölder's inequality states that if $f_1, f_2, ..., f_m$ are functions on $\mathbb{R}^n$ with $f_i \in L^{p_i}(\mathbb{R}^n)$ and $\sum_{j=1}^{m} \frac{1}{p_i} = 1$, then

$$\left| \int_{\mathbb{R}^n} \prod_{j=1}^{m} f_i(x)dx \right| \leq \int_{\mathbb{R}^n} \prod_{j=1}^{m} |f_i(x)|dx \leq \prod_{j=1}^{m} \|f_i\|_{p_i} \tag{6}$$

which can be proved by induction on m.

Theorem 2.2 (Fubini's Theorem for $\mathbb{R}^n$)**.** *Let f be a function on $\mathbb{R}^n \times \mathbb{R}^m$. If*

$$\int_{\mathbb{R}^n \times \mathbb{R}^m} |f(x, y)|dxdy < \infty \tag{7}$$

then

$$\begin{aligned} &\int_{\mathbb{R}^n \times \mathbb{R}^m} f(x, y)dxdy \\ =&\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f(x, y)dy \right) dx \\ =&\int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x, y)dx \right) dy \end{aligned} \tag{8}$$

Let us note that for most integrals considered in this paper the requisite condition for changing the order of integration (7) *is satisfied. This is explicitly shown in the simple version of Young's inequality in the proof of equations* (24) *and* (25)*.*

Definition (Convolution)**.** *When f and g are two (complex-valued) functions on $\mathbb{R}^n$ we define their* **convolution** *to be the function $f * g$ given by*

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y)g(y)dy \tag{9}$$

Note that by changing variables

$$f * g = g * f \tag{10}$$

Definition (Strong Convergence in $\mathbb{R}^n$)**.** *Let $f, f^i \in L^p(\mathbb{R}^n)$ for $i = 1, 2, 3, ...$ We say that f^i* **converges strongly** *to f in $L^p(\mathbb{R}^n)$ if*

$$\|f^i - f\|_p \to 0 \text{ as } i \to \infty \tag{11}$$

We denote this by

$$f^i \to f \text{ as } i \to \infty$$

Theorem 2.3 (Approximation by C^∞-functions)**.** *Let j be in $L^1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} j = 1$. For $\epsilon > 0$, define $j_\epsilon(x) := \epsilon^{-n} j\left(\frac{x}{\epsilon}\right)$, so that $\int_{\mathbb{R}^n} j_\epsilon = 1$ and $\|j_\epsilon\|_1 = \|j\|_1$. Let $f \in L^p(\mathbb{R}^n)$ for some $1 \leq p < \infty$. Then*

$$j_\epsilon * f \in L^p(\mathbb{R}^n) \text{ and } \|j_\epsilon * f\|_p \leq \|j\|_1 \|f\|_p \tag{12}$$

$$j_\epsilon * f \to f \text{ strongly in } L^p(\mathbb{R}^n) \text{ as } \epsilon \to 0 \tag{13}$$

If j is a Gaussian function in $\mathbb{R}^n$, then $j_\epsilon * f \in C^\infty(\mathbb{R}^n)$

Definition (Weak Convergence in $\mathbb{R}^n$)**.** *If $f, f^1, f^2, f^3,...$ is a sequence of functions in $L^p(\mathbb{R}^n)$, we say that f^i* **converges weakly to** *f (and write $f^i \rightharpoonup f$) if*

$$\lim_{i\to\infty} L(f^i) = L(f) \tag{14}$$

for every continuous linear functional $L \in L^p(\mathbb{R}^n)^*$.

A function $g \in L^{p'}(\mathbb{R}^n)$ acts on arbitrary functions $f \in L^p(\mathbb{R}^n)$ by

$$L_g(f) = \int_{\mathbb{R}^n} g(x)f(x)dx \tag{15}$$

The Riesz-Representation Theorem tells that for any $L \in L^p(\mathbb{R}^n)^*$ there exists a unique $g \in L^{p'}(\mathbb{R}^n)$ such that for every $f \in L^p(\mathbb{R}^n)$ we have

$$L(f) = \int_{\mathbb{R}^n} g(x)f(x)dx \tag{16}$$

This proves that $L^{p'}(\mathbb{R}^n) = L^p(\mathbb{R}^n)^*$ when $\frac{1}{p} + \frac{1}{p'} = 1$.

Theorem 2.4 (Bounded sequences have weak limits)**.** *Let $\Omega \subset \mathbb{R}^n$ be a measurable set and consider $L^p(\Omega)$ with $1 < p < \infty$. Let $f^1, f^2,...$ be a sequence of functions, bounded in $L^p(\Omega)$. Then there exist a subsequence $f^{n_1}, f^{n_2},...$ (with $n_1 < n_2 <...$) and an $f \in L^p(\Omega)$ such that $f^{n_i} \rightharpoonup f$ weakly in $L^p(\Omega)$ as $i \to \infty$.*

Theorem 2.5 (Lower semicontinuity of norms in $\mathbb{R}^n$)**.** *For $1 \leq p \leq \infty$ the L^p-norm is* **weakly lower semicontinuous***, i.e, if $f^j \rightharpoonup f$ weakly in $L^p(\mathbb{R}^n)$, then*

$$\liminf_{j\to\infty} \|f^j\|_p \geq \|f\|_p \tag{17}$$

If $p = \infty$ we use the extra fact that our measure space is sigma-finite. Moreover, if $1 < p < \infty$ and if $\lim_{j\to\infty} \|f^j\|_p = \|f\|_p$, then $f^j \to f$ strongly as $j \to \infty$.

Definition (Pointwise Convergence in $\mathbb{R}^n$)**.** *Let $f, f^1, f^2,...$ be complex-valued measurable functions on $\mathbb{R}^n$. $f^j(x)$* **converges pointwise** *to $f(x)$ if for every $x \in \mathbb{R}^n$, $f^j(x) \to f(x)$ in the $\epsilon - \delta$ sense.*

Theorem 2.6 (Dominated Convergence in $\mathbb{R}^n$)**.** *Let $f^1, f^2,...$ be complex-valued functions on $\mathbb{R}^n$ with finite integrals and assume that these functions converge to a function f pointwise for almost every x. If there exists a non-negative function $G(x)$ with finite integral on $\mathbb{R}^n$ such that $|f^j(x)| < G(x)$ for all $j = 1, 2, ...,$ then $|f(x)| < G(x)$ and*

$$\lim_{j\to\infty} \int_{\mathbb{R}^n} f^j(x)dx = \int_{\mathbb{R}^n} f(x)dx \tag{18}$$

Theorem 2.7 (Stone-Weierstrass for $\mathbb{R}^n$)**.** *Let $f(x)$ be a continuous function on $\mathbb{R}^n$ with compact support, i.e, $f(x) = 0$ for $|x| > r$, where r is some finite radius. Then f is the uniform limit of polynomials on $|x| \leq r$.*

Theorem 2.8 (A Weaker form of Minkowski's Inequality for $\mathbb{R}^n$)**.** *Let f be a non-negative function on $\mathbb{R}^n \times \mathbb{R}^m$ which is $dx \times dy$ measurable. Let $1 \leq p < \infty$. Then*

$$\left(\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^m} f(x,y)dy\right)^p dx\right)^{\frac{1}{p}} \leq \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^n} f(x,y)^p dx\right)^{\frac{1}{p}} dy \tag{19}$$

in the sense that the finiteness of the right hand side implies the finiteness of the left hand side.

Equality and finiteness in (19) for $1 < p < \infty$ imply the existence of a dx-measurable function $\alpha : \mathbb{R}^n \to \mathbb{R}^+$ and a dy-measurable function $\beta : \mathbb{R}^m \to \mathbb{R}^+$ such that

$$f(x,y) = \alpha(x)\beta(y) \text{ for almost every } (x,y) \tag{20}$$

Theorem 2.9 (Monotone Convergence in $\mathbb{R}^n$)**.** *Let $f^1, f^2, \ldots$ be an increasing sequence of functions on $\mathbb{R}^n$ with finite integrals then*

$$\lim_{j\to\infty}\int_{\mathbb{R}^n} f^j(x)dx = \int_{\mathbb{R}^n}\lim_{j\to\infty} f^j(x)dx \tag{21}$$

3. Dimensional Analysis

Why must $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$ be a necessary condition for Young's inequality (1) ?

To answer this, let us investigate a change of scaling in the functions $f(x), g(x-y)$ and $h(y)$ on both sides of Young's inequality. Under a change in scaling, the right hand side becomes

$$\begin{aligned}&\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(\lambda x)g(\lambda(x-y))h(\lambda y)dxdy\right| \\ =&\frac{1}{\lambda^2}\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(x-y)h(y)dxdy\right|\end{aligned} \tag{22}$$

And the left hand side becomes

$$\begin{aligned}&\left(\int_{\mathbb{R}^n}|f(\lambda x)|^p dx\right)^{\frac{1}{p}}\left(\int_{\mathbb{R}^n}|g(\lambda z)|^q dz\right)^{\frac{1}{q}}\left(\int_{\mathbb{R}^n}|h(\lambda y)|^r dy\right)^{\frac{1}{r}} \\ =&\left(\frac{1}{\lambda}\int_{\mathbb{R}^n}|f(x)|^p dx\right)^{\frac{1}{p}}\left(\frac{1}{\lambda}\int_{\mathbb{R}^n}|g(z)|^q dz\right)^{\frac{1}{q}}\left(\frac{1}{\lambda}\int_{\mathbb{R}^n}|h(y)|^r dy\right)^{\frac{1}{r}} \\ =&\frac{1}{\lambda}^{\frac{1}{p}+\frac{1}{q}+\frac{1}{r}}\|f\|_p\|g\|_q\|h\|_r\end{aligned} \tag{23}$$

Case 1: If $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}<2$, take the limit as $\lambda\to 0$. The left hand side goes to ∞ more quickly than the right hand side and therefore our inequility must be false.

Case 2: If $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}>2$, take the limit as $\lambda\to\infty$. The right hand side goes to zero more quickly than the left hand side and therefore our inequility must be false.

Thus, for Young's inequality to be true $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}$ must equal 2.

4. SIMPLE VERSION OF YOUNG'S INEQUALITY WITHOUT THE SHARP CONSTANT

Lemma 4.1. *The sharp constant in Young's inequality* (2) *is bounded by 1.*

Proof. We will use the generalized form of Holder's inequality in our proof. Let $p',q',r'\geq 1$ and $\frac{1}{p'}+\frac{1}{q'}+\frac{1}{r'}=1$, and let $\alpha\in L^{r'}(\mathbb{R}^n)$, $\beta\in L^{p'}(\mathbb{R}^n)$, and $\gamma\in L^{q'}(\mathbb{R}^n)$, so that the necessary conditions of equation (6) are satisfied.

To satisfy the requisite condition of Young's inequality that $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=2$, we know from equation (3) that our choice of p,q,r can be $p=\frac{p'}{p'-1}, q=\frac{q'}{q'-1}$, and $r=\frac{r'}{r'-1}$ (in which case clearly $p,q,r>1$). Also let $f\in L^p(\mathbb{R}^n)$, $g\in L^q(\mathbb{R}^n)$, and $h\in L^r(\mathbb{R}^n)$, so that both requisite conditions of equation (1) are satisfied.

We will show that

$$\begin{aligned}&\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(x-y)h(y)dxdy\right|\\ &\leq\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|f(x)g(x-y)h(y)|dxdy\\ &=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|\alpha(x,y)\beta(x,y)\gamma(x,y)|dxdy\end{aligned}\tag{24}$$

which by Hölder's inequality (6) is less than or equal to

$$\|\alpha\|_{r'}\|\beta\|_{p'}\|\gamma\|_{q'}=\|f\|_p\|g\|_q\|h\|_r\tag{25}$$

Start by letting

$$\begin{aligned}\alpha(x,y)&=f(x)^{\frac{p}{r'}}g(x-y)^{\frac{q}{r'}},\\ \beta(x,y)&=g(x-y)^{\frac{q}{p'}}h(y)^{\frac{r}{p'}},\\ \gamma(x,y)&=f(x)^{\frac{p}{q'}}h(y)^{\frac{r}{q'}}\end{aligned}$$

Then,

$$\begin{aligned}
\alpha(x,y)\beta(x,y)\gamma(x,y) &= f(x)^{\frac{p}{r'}+\frac{p}{q'}} g(x-y)^{\frac{q}{p'}+\frac{q}{r'}} h(y)^{\frac{r}{p'}+\frac{r}{q'}} \\
&= f(x)^{p\left(\frac{1}{r'}+\frac{1}{q'}\right)} g(x-y)^{q\left(\frac{1}{p'}+\frac{1}{r'}\right)} h(y)^{r\left(\frac{1}{p'}+\frac{1}{q'}\right)} \\
&= f(x)^{p\left(1-\frac{1}{p'}\right)} g(x-y)^{q\left(1-\frac{1}{q'}\right)} h(y)^{r\left(1-\frac{1}{r'}\right)} \\
&= f(x)^{p\left(\frac{p'-1}{p'}\right)} g(x-y)^{q\left(\frac{q'-1}{q'}\right)} h(y)^{r\left(\frac{r'-1}{r'}\right)} \\
&= f(x)^{p\left(\frac{1}{p}\right)} g(x-y)^{q\left(\frac{1}{q}\right)} h(y)^{r\left(\frac{1}{r}\right)} \\
&= f(x)g(x-y)h(y)
\end{aligned} \tag{26}$$

Thus we have shown the equality necessary for equation (24).
To complete the proof, we need to prove equation (25).

$$\begin{aligned}
\|\alpha\|_{r'} &= \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\alpha(x,y)|^{r'} dxdy\right)^{\frac{1}{r'}} \\
&= \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x)^{\frac{p}{r'}} g(x-y)^{\frac{q}{r'}}|^{r'} dxdy\right)^{\frac{1}{r'}} \\
&= \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x)|^p |g(x-y)|^q dxdy\right)^{\frac{1}{r'}} \\
&= \left(\int_{\mathbb{R}^n} |f(x)|^p \int_{\mathbb{R}^n} |g(x-y)|^q dydx\right)^{\frac{1}{r'}} \\
&= \left(\int_{\mathbb{R}^n} |f(x)|^p dx \int_{\mathbb{R}^n} |g(-z)|^q dz\right)^{\frac{1}{r'}} \\
&= \|f\|_p^{\frac{p}{r'}} \|g\|_q^{\frac{q}{r'}}
\end{aligned} \tag{27}$$

where the 3rd to last equality in (27) is achieved via Fubini's theorem (8) and the 2nd to last equality in (27) is obtained by substituting $z = y - x$ and $dz = dy$.

Similarly,

$$\begin{aligned}
\|\beta\|_{p'} &= \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\beta(x,y)|^{p'} dxdy\right)^{\frac{1}{p'}} \\
&= \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |g(x-y)^{\frac{q}{p'}} h(y)^{\frac{r}{p'}}|^{p'} dxdy\right)^{\frac{1}{p'}} \\
&= \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |g(x-y)|^{q} |h(y)|^{r} dxdy\right)^{\frac{1}{p'}} \\
&= \left(\int_{\mathbb{R}^n} |h(y)|^{r} \int_{\mathbb{R}^n} |g(x-y)|^{q} dxdy\right)^{\frac{1}{p'}} \\
&= \left(\int_{\mathbb{R}^n} |g(z)|^{q} dz \int_{\mathbb{R}^n} |h(y)|^{r} dy\right)^{\frac{1}{p'}} \\
&= \|g\|_q^{\frac{q}{p'}} \|h\|_r^{\frac{r}{p'}}
\end{aligned} \tag{28}$$

where the 2nd to last equality in (28) is obtained by substituting $z = x - y$ and $dz = dx$,

And,

$$\begin{aligned}
\|\gamma\|_{q'} &= \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\gamma(x,y)|^{q'} dxdy\right)^{\frac{1}{q'}} \\
&= \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x)^{\frac{p}{q'}} h(y)^{\frac{r}{q'}}|^{q'} dxdy\right)^{\frac{1}{q'}} \\
&= \left(\int_{\mathbb{R}^n} |f(x)|^{p} dx \int_{\mathbb{R}^n} |h(y)|^{r} dy\right)^{\frac{1}{q'}} \\
&= \|f\|_p^{\frac{p}{q'}} \|h\|_r^{\frac{r}{q'}}
\end{aligned} \tag{29}$$

Combining the results of equations (27) (28) and (29)

$$\begin{aligned}
\|\alpha\|_{r'}\|\beta\|_{p'}\|\gamma\|_{q'} &= \|f\|_p^{\frac{p}{r'}} \|g\|_q^{\frac{q}{r'}} \|g\|_q^{\frac{q}{p'}} \|h\|_r^{\frac{r}{p'}} \|f\|_p^{\frac{p}{q'}} \|h\|_r^{\frac{r}{q'}} \\
&= \|f\|_p^{\frac{p}{q'}+\frac{p}{r'}} \|g\|_q^{\frac{q}{r'}+\frac{q}{p'}} \|h\|_r^{\frac{r}{p'}+\frac{r}{q'}} \\
&= \|f\|_p^{p\left(\frac{1}{q'}+\frac{1}{r'}\right)} \|g\|_q^{q\left(\frac{1}{r'}+\frac{1}{p'}\right)} \|h\|_r^{r\left(\frac{1}{p'}+\frac{1}{q'}\right)} \\
&= \|f\|_p^{p\left(1-\frac{1}{p'}\right)} \|g\|_q^{q\left(1-\frac{1}{q'}\right)} \|h\|_r^{r\left(1-\frac{1}{r'}\right)} \\
&= \|f\|_p^{p\left(\frac{p'-1}{p'}\right)} \|g\|_q^{q\left(\frac{q'-1}{q'}\right)} \|h\|_r^{r\left(\frac{r'-1}{r'}\right)} \\
&= \|f\|_p^{p\left(\frac{1}{p}\right)} \|g\|_q^{q\left(\frac{1}{q}\right)} \|h\|_r^{r\left(\frac{1}{r}\right)} \\
&= \|f\|_p \|g\|_q \|h\|_r
\end{aligned} \tag{30}$$

The proof of equations (24) and (25) is complete, so

$$\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(x-y)h(y)dxdy\right| \leq \|f\|_p\|g\|_q\|h\|_r \tag{31}$$

and the sharp constant in Young's inequality (1) is therefore bounded by 1. □

5. Equivalent form of Young's Inequality

Taking an equivalent form of Young's inequality involving only the norm of a convolution of functions simplifies the problem by not having to concentrate on the additional function f.

Lemma 5.1. *Young's inequality* (1) *can be restated as an inequality involving the norm of a convolution:*

$$\|g*h\|_p \leq (C_qC_r/C_p)^n\|g\|_q\|h\|_r = C_{p',q,r;n}\|g\|_q\|h\|_r \tag{32}$$

with $\frac{1}{q}+\frac{1}{r}=1+\frac{1}{p}$

Proof. In our original form of the Young's inequality (1), the left hand side convolution $(g*h)(x)$ can be written in the equivalent form $e^{i\theta(x)}|(g*h)(x)|$, where $\theta(x)$ is the function specifying the direction of the convolution in the complex plane for all x. By choosing $f(x)=e^{-i\theta(x)}|(g*h)(x)|^{\frac{p'}{p}}$ so that the phase of f cancels with the phase $e^{i\theta(x)}$ of the convolution $(g*h)$, we can prove that (1) ⇒ (32).

We will then use Hölder's inequality to prove that (32) ⇒ (1).

(1) ⇒ (32) is as follows:

$$\left|\int_{\mathbb{R}^n} f(x)(g*h)(x)dx\right| = \int_{\mathbb{R}^n}|(g*h)(x)|^{\frac{p'}{p}+1}dx = \|g*h\|_{p'}^{p'} \tag{33}$$

where $\frac{p'}{p}+1=p'$. By Young's inequality

$$\|g*h\|_{p'}^{p'} = \int_{\mathbb{R}^n}|(g*h)(x)|^{\frac{p'}{p}}|(g*h)(x)|dx \leq C_{p,q,r;n}\|(g*h)^{\frac{p'}{p}}\|_p\|g\|_q\|h\|_r \tag{34}$$

where

$$\begin{aligned} \|(g*h)^{\frac{p'}{p}}\|_p &= \left(\int_{\mathbb{R}^n}(|(g*h)(x)|^{\frac{p'}{p}})^p dx\right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^n}|(g*h)(x)|^{p'}dx\right)^{\frac{1}{p}} \\ &= \|(g*h)\|_{p'}^{\frac{p'}{p}} \end{aligned} \tag{35}$$

Substituting the result of (35) back into the result of (34), we have

$$\|g*h\|_{p'}^{p'} \leq C_{p,q,r;n}\|g*h\|_{p'}^{\frac{p'}{p}}\|g\|_q\|h\|_r \tag{36}$$

Multiplying both sides by $\|g*h\|_{p'}^{\frac{-p'}{p}}$ yields

$$\|g*h\|_{p'}^{p'-\frac{p}{p'}} = \|g*h\|_{p'} \leq C_{p,q,r;n}\|g\|_q\|h\|_r \tag{37}$$

where $p' - \frac{p}{p'} = 1$. Since we have from equation (2) that $C_p = \frac{1}{C_{p'}}$ with $\frac{1}{p} + \frac{1}{p'} = 1$, substituting p for p' into equation (37), we obtain equation (32).

We have shown that (1) ⇒ (32). We need to prove that (32) ⇒ (1). Use Hölder's inequality (5) on the left hand side of Young's inequality (1) as follows:

$$\left|\int_{\mathbb{R}^n} f(x)(g*h)(x)dx\right| \leq \|f\|_p\|g*h\|_{p'} \tag{38}$$

substituting the result of (37) into the right hand side of (38), we obtain equation (1). Thus we have shown that equations (1) and (32) are equivalent forms of Young's inequality. □

6. The symmetry of Young's inequality

Upon inspection of Young's inequality, one might notice a certain symmetry present. As it turns out, the sharp constant is dependent on the values of p, q and r, where the contributions from those three factors can be separated, specifically, $C_{p,q,r;n} = (C_pC_qC_r)^n$.

Let us denote $\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(x-y)h(y)dxdy$ by $I[f(x), g(x-y), h(y)]$.

Lemma 6.1.

$$\begin{aligned} I[f(x), g(x-y), h(y)] &= I[g(x), f(x-y), h(-y)] \\ &= I[f(x), h(x-y), g(y)] \\ &= I[h(x), g(y-x), f(y)] \end{aligned} \tag{39}$$

Proof. We will show the case for $\mathbb{R}$, which can be easily adapted for $\mathbb{R}^n$.

$$\begin{aligned}
I[h(x), g(y-x), f(y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)g(x-y)h(y)dxdy \\
&= \int_{\infty}^{-\infty}\int_{\infty}^{-\infty} f(-x)g(-x+y)h(-y)d(-x)d(-y) \\
&= \int_{\infty}^{-\infty}\int_{\infty}^{-\infty} f(-x)g(-x+y)h(-y)dxdy \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(-x)g(-x+y)dx\right] h(-y)dy \\
&= \int_{-\infty}^{\infty}\left[\int_{\infty}^{-\infty} f(z-y)g(z)(-dz)\right] h(-y)dy \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(z-y)g(z)dz\right] h(-y)dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(z)f(z-y)h(-y)dzdy \\
&= I[g(x), f(x-y), h(-y)]
\end{aligned} \tag{40}$$

Here the third line of (40) is obtained since $d(-x) = -dx, d(-y) = -dy$. Switching one set of bounds in the fourth line would introduce a sign change, but by switching both the sign changes cancel out one another. In the fifth line, we let $z = -x + y$ and $dz = -dx$, and switch the inner integral's bounds of integration. In the sixth line, we switch the bounds on the inner integral and introduce a factor of -1 which cancels out the -1 multiplying dz. In the last line, we substite x back in for z, which is just a change of notation.

The next form follows by Fubini's theorem (8),

$$\begin{aligned}
I[g(x), f(x-y), h(-y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)f(x-y)h(-y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)f(x-y)h(-y)dydx \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x-y)h(-y)dy\right] g(x)dx \\
&= \int_{-\infty}^{\infty}\left[\int_{\infty}^{-\infty} f(z)h(z-x)(-dz)\right] g(x)dx \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(z)h(z-x)dz\right] g(x)dx \\
&= I[f(x), h(x-y), g(y)]
\end{aligned} \tag{41}$$

Here the the second line of (41) is obtained by Fubini's theorem (8). In the fourth line, we let $z = x - y$ and $dz = -dy$, and switch the inner integral's bounds of integration. In the fifth line, we switch the inner integral's bounds of integration and introduce a factor of -1 which cancels out the -1 multiplying dz. In the last line we swap y for x then x for z, which are just changes in notation.

Finally, one can see that by using Fubini's theorem (8) once again and simply swapping x and y in the integrals,

$$I[f(x), g(x-y), h(y)] = I[h(x), g(y-x), f(y)]$$

$\square$

7. Gaussian Optimizers

We will consider the real one-dimensional case of equation (4)

Theorem 7.1. *If* $p, q, r > 1$*, and if* f, g *and* h *are Gaussian functions*

$$\begin{aligned} f(x) &= Ae^{-p'(x-a)^2}, \\ g(x) &= Be^{-q'(x-b)^2}, \\ h(x) &= Ce^{-r'(x-c)^2}, \end{aligned} \tag{42}$$

then equality in Young's inequality (1) *is achieved,*
where $A, B, C, a, b, c \in \mathbb{R}$ *with* $a = b + c$.

Proof. First notice,

$$\begin{aligned} &\int_{\mathbb{R}}\int_{\mathbb{R}} f(x+a)g(x-y+b)h(y+c)dxdy \\ =&\int_{\mathbb{R}}\int_{\mathbb{R}} f(x)g(x-a-y+b)h(y+c)dxdy \\ =&\int_{\mathbb{R}}\int_{\mathbb{R}} f(x)g(x-y-c)h(y+c)dxdy \\ =&\int_{\mathbb{R}}\int_{\mathbb{R}} f(x)g(x-y)h(y)dxdy \end{aligned} \tag{43}$$

In equation (43), we shift x by $-a$ in line two. In line three, we use that $b - a = c$ and in line four we shify y by $-c$. This shows that for our proof it is enough to show the case where f, g and h are chosen in (42) with $a = b = c = 0$. We can also notice that multiplying the three functions by constants A, B and C scales both sides of Young's inequality (1) by the same factor, so it is enough to prove the case where $A = B = C = 1$.

We will calculate the left hand side of Young's inequality (1), then calculate the right hand side and show that they are equal.

The Left Hand Side:

$$\begin{aligned}
&\left|\int_{\mathbb{R}}\int_{\mathbb{R}} f(x)g(x-y)h(y)dxdy\right| \\
=&\left|\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-p'x^2}e^{-q'(x-y)^2}e^{-r'y^2}dxdy\right| \\
=&\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-(p'+q')x^2+q'2xy-(r'+q')y^2}dxdy \\
=&\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-\left(\sqrt{p'+q'}x-\frac{q'}{\sqrt{p'+q'}}y\right)^2+\frac{q'^2}{p'+q'}y^2-(r'+q')y^2}dxdy \\
=&\int_{\mathbb{R}} e^{\left(\frac{q'^2}{p'+q'}-r'-q'\right)y^2}\int_{\mathbb{R}} e^{-\left(\sqrt{p'+q'}x-\frac{q'}{\sqrt{p'+q'}}y\right)^2}dxdy \\
=&\int_{\mathbb{R}} e^{\left(\frac{q'^2}{p'+q'}-r'-q'\right)y^2}dy\int_{\mathbb{R}} e^{-(p'+q')z^2}dz
\end{aligned}\tag{44}$$

Line four of (44) is found by completing the square. In line six, we set $z = x + \frac{q'}{p'+q'}y$ and $dz = dx$, which is simply a shift in x. Line six uses the calculation $\frac{q'^2}{p'+q'} - r' - q' = \frac{q'^2}{p'+q'} - r'\frac{p'+q'}{p'+q'} - q'\frac{p'+q'}{p'+q'}$. Calculating the inner integral in line six, we have

$$\begin{aligned}
&\left|\int_{\mathbb{R}}\int_{\mathbb{R}} f(x)g(x-y)h(y)dxdy\right| \\
=&\int_{\mathbb{R}} e^{\left(\frac{q'^2}{p'+q'}-r'-q'\right)y^2}\sqrt{\frac{\pi}{p'+q'}}dy \\
=&\sqrt{\frac{\pi}{p'+q'}}\int_{\mathbb{R}^n} e^{-\left(\frac{r'p'+r'q'+p'q'}{p'+q'}\right)y^2}dy \\
=&\sqrt{\frac{\pi}{p'+q'}}\sqrt{\frac{\pi(p'+q')}{r'p'+r'q'+p'q'}} \\
=&\frac{\pi}{\sqrt{r'p'+r'q'+p'q'}} \\
=&\frac{\pi}{\sqrt{p'q'r'}}
\end{aligned}\tag{45}$$

The last line of (45) uses a simplification obtained by multiplying both sides of equation (3) by $p'q'r'$.

The Right Hand Side:

$$C_{p,q,r;1} = C_pC_qC_r = \left(\frac{p^{\frac{1}{2p}}}{p'^{\frac{1}{2p'}}}\right)\left(\frac{q^{\frac{1}{2q}}}{q'^{\frac{1}{2q'}}}\right)\left(\frac{r^{\frac{1}{2r}}}{r'^{\frac{1}{2r'}}}\right)\tag{46}$$

$$\begin{aligned}\|f\|_p &= \left(\int_{\mathbb{R}} |e^{-p'x^2}|^p dx\right)^{\frac{1}{p}}\\ &= \left(\int_{\mathbb{R}} e^{-pp'x^2} dx\right)^{\frac{1}{p}}\\ &= \left(\sqrt{\frac{\pi}{pp'}}\right)^{\frac{1}{p}}\\ &= \left(\frac{\pi}{pp'}\right)^{\frac{1}{2p}}\end{aligned}\tag{47}$$

Similarly,

$$\|g\|_q = \left(\frac{\pi}{qq'}\right)^{\frac{1}{2q}}\tag{48}$$

And,

$$\|h\|_r = \left(\frac{\pi}{rr'}\right)^{\frac{1}{2r}}\tag{49}$$

Combining equations (47),(48) and (49)

$$\begin{aligned}\|f\|_p\|g\|_q\|h\|_r &= \left(\frac{\pi}{pp'}\right)^{\frac{1}{2p}}\left(\frac{\pi}{qq'}\right)^{\frac{1}{2q}}\left(\frac{\pi}{rr'}\right)^{\frac{1}{2r}}\\ &= (\pi)^{\frac{1}{2p}+\frac{1}{2q}+\frac{1}{2r}}(pp')^{\frac{-1}{2p}}(qq')^{\frac{-1}{2q}}(rr')^{\frac{-1}{2r}}\\ &= \pi(pp')^{-\frac{1}{2p}}(qq')^{-\frac{1}{2q}}(rr')^{-\frac{1}{2r}}\end{aligned}\tag{50}$$

We must multiply equation (50) by the sharp constant (46) to complete the calculation of the Right Hand Side.

$$\begin{aligned}C_{p,q,r;1}\|f\|_p\|g\|_q\|h\|_r &= \left(\frac{p^{\frac{1}{2p}}}{p'^{\frac{1}{2p'}}}\right)\left(\frac{q^{\frac{1}{2q}}}{q'^{\frac{1}{2q'}}}\right)\left(\frac{r^{\frac{1}{2r}}}{r'^{\frac{1}{2r'}}}\right)\pi(pp')^{-\frac{1}{2p}}(qq')^{-\frac{1}{2q}}(rr')^{-\frac{1}{2r}}\\ &= \pi(p')^{-\frac{1}{2p}-\frac{1}{2p'}}(q')^{-\frac{1}{2q}-\frac{1}{2q'}}(r')^{-\frac{1}{2r}-\frac{1}{2r'}}\\ &= (p')^{-\frac{1}{2}}(q')^{-\frac{1}{2}}(r')^{-\frac{1}{2}}\\ &= \frac{\pi}{\sqrt{p'q'r'}}\end{aligned}\tag{51}$$

Thus the Right Hand Side equals the Left Hand Side.

□

8. THE FULL PROOF WITH THE SHARP CONSTANT

Outline of the Proof.

- Sharp Constant is at most 1 (Proved in Section 4).
- Define an auxiliary problem that is true for all ϵ, δ.
- The sharp constant in the auxiliary problem is attained.

- The sharp constant in the auxiliary problem is separable into a product of one dimensional sharp constants.
- Optimizing functions of the auxiliary problem must factorize.
- Optimizing functions of the auxiliary problem must be Gaussians.
- Optimizing functions of Young's inequality must be Gaussians.
- The sharp constant of the auxiliary problem equals the sharp constant of Young's inequality.
- The sharp constant in Young's inequality is $C_{p,q,r;n}$ (Shown in Section 7).

8.1. **Auxiliary Problem.** We will use an auxiliary problem to show that maximizers f, g and h exist in equation (32) and that we can compute them. Let us introduce the Gaussian function

$$j(x) = \pi^{-\frac{n}{2}} e^{-|x|^2} \tag{52}$$

and note that $j \in L^1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} j = \|j\|_1 = 1$.

We define our auxiliary function: for $0 \leq \epsilon, \delta \ll 1$,

$$K^{\epsilon,\delta}_{g,h}(x) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} J^{\epsilon,\delta}_n g(y)h(z)dydz \tag{53}$$

where

$$J^{\epsilon,\delta}_n = (\pi\epsilon)^{-\frac{n}{2}} \exp\left(-\frac{|x-y-z|^2}{\epsilon} - \delta|y|^2 - \delta|x|^2 - \delta|z|^2\right) \tag{54}$$

Notice that since $J^{\epsilon,\delta}_n \geq 0$ and $0 \leq \exp[-\delta(|y|^2 + |z|^2 + |x|^2)] \leq 1$

$$J^{\epsilon,\delta}_n \leq J^{\epsilon,\delta-\eta}_n \leq J^{\epsilon,0}_n \tag{55}$$

where $0 \leq \eta \leq \delta$. Therefore for $g(y)$ and $h(z)$ positive (we will show later that maximizers can be thought of as positive functions)

$$K^{\epsilon,\delta}_{g,h} \leq K^{\epsilon,\delta-\eta}_{g,h} \leq K^{\epsilon,0}_{g,h} \tag{56}$$

The motivation for introducing this auxiliary function is contained in the following informal calculation:
In the limit $\epsilon, \delta \to 0$, $J^{\epsilon,\delta}_n \to \delta(y-(x-z))$. So in the limit $\epsilon, \delta \to 0$, our auxiliary function becomes

$$\begin{aligned} K^{0,0}_{g,h}(x) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \delta(y-(x-z))g(y)h(z)dydz \\ &= \int_{\mathbb{R}^n} g(z-x)h(z)dz \\ &= (g*h)(x) \end{aligned} \tag{57}$$

In order to introduce our auxiliary problem, we need the following calculation:

$$\begin{aligned}
\|K^{\epsilon,0}_{g,h}\|_p &= \left(\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(\pi\epsilon)^{-\frac{n}{2}}e^{-\frac{|x-y-z|^2}{\epsilon}}g(y)h(z)dydz\right|^p dx\right)^{\frac{1}{p}}\\
&= \left(\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(\pi\epsilon)^{-\frac{n}{2}}e^{-\frac{|x-u|^2}{\epsilon}}g(y)h(u-y)dudy\right|^p dx\right)^{\frac{1}{p}}\\
&= \left(\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n}(\pi\epsilon)^{-\frac{n}{2}}e^{-\frac{|x-u|^2}{\epsilon}}(g*h)(u)du\right|^p dx\right)^{\frac{1}{p}}\\
&= \left(\int_{\mathbb{R}^n}\left|[j_{\sqrt{\epsilon}}*(g*h)](x)\right|^p dx\right)^{\frac{1}{p}}\\
&= \|j_{\sqrt{\epsilon}}*(g*h)\|_p\\
&\le \|g*h\|_p\\
&\le C^{\epsilon,\delta}_n\|g\|_q\|h\|_r
\end{aligned} \tag{58}$$

In line 2 of (58) we use Fubini's Theorem (8) to interchange dy and dz. Since we are integrating dz, we can then fix y and substitute $u = y+z, du = dz$. To obtain line 3, we use Fubini's Theorem once again, interchange du and dy, and then intergrate dy. Line 4 follows from the definition of a convolution (9), where $j_\epsilon = \epsilon^{-n}j(\frac{x}{\epsilon})$, and $j(x)$ is the Gaussian specified in equation (52). Line 6, is the direct application of equation (12) with $\|j\|_1 = 1$. In line 7 we are introducing the sharp constant for our auxiliay function where

$$C^{\epsilon,\delta}_n \le C^{\epsilon,0}_n \le C_{p',q,r;n} \le 1 \tag{59}$$

The first inequality in (59) is a direct result of equation (55). The second inequality is clear from (58). And the bound of 1 was proved in (31).

Considering (56) and (58), we can now state our auxiliary problem as

$$\|K^{\epsilon,\delta}_{g,h}\|_p \le C^{\epsilon,\delta}_n\|g\|_q\|h\|_r \tag{60}$$

which is in the form of Young's inequality equation (32) where $\frac{1}{q}+\frac{1}{r} = 1+\frac{1}{p}$.

8.2. **The sharp constant in the auxiliary problem is attained.**

Lemma 8.1. *There exist functions g and h with $\|g\|_q = \|h\|_r = 1$ such that $\|K^{\epsilon,\delta}_{g,h}\|_p = C^{\epsilon,\delta}_n$.*

Proof. Let g_i, h_i be a maximizing sequence of pairs of functions with $\|g_i\|_q = \|h_i\|_r = 1$ so that $\|K^{\epsilon,\delta}_{g_i,h_i}\|_p \to C^{\epsilon,\delta}_n$. By Theorem 2.4 (Bounded sequences have weak limits) on page 4 there exist $g \in L^q(\mathbb{R}^n), h \in L^r(\mathbb{R}^n)$ such that $g_i \rightharpoonup g$ and $h_i \rightharpoonup h$ weakly in $L^q(\mathbb{R}^n)$ and $L^r(\mathbb{R}^n)$ respectively. By Theorem 2.5 (Lower semicontinuity of norms), we have that $\|g\|_q \le \liminf_{i\to\infty}\|g_i\|_q = 1$ and similarly $\|h\|_r \le 1$. We need to show that these inequalities are strict equalities. In order to show strict equality we must show that that $K^{\epsilon,\delta}_{g_i,h_i}(x)$ converges strongly (page 3) in $L^p(\mathbb{R}^n)$ as $i \to \infty$ to the function $K^{\epsilon,\delta}_{g,h}(x)$.

To show strong convergence of $K^{\epsilon,\delta}_{g_i,h_i}(x)$ to $K^{\epsilon,\delta}_{g,h}(x)$, we will show

(a) Pointwise convergence of $K^{\epsilon,\delta}_{g_i,h_i}(x)$ to $K^{\epsilon,\delta}_{g,h}(x)$
(b) Uniform boundedness of $K^{\epsilon,\delta}_{g_i,h_i}(x)$ in x and in i
(c) Conditions (a) and (b) hold even if we multiply $K^{\epsilon,\delta}_{g_i,h_i}(x)$ by $e^{\gamma|x|^2}$ for some sufficiently small $\gamma > 0$, so we have that $K^{\epsilon,\delta}_{g_i,h_i}(x)$ decays sufficiently as $x \to \infty$
(d) Applying the Dominated Convergence Theorem (18), we have strong convergence.

(a) Pointwise Convergence: We will decompose our equation for pointwise convergence into three equations, which can all be made arbitrarily small. In order to accomplish this, we will introduce a new continuous function with compact support $W^\epsilon_n(x,y,z)$. By Theorem 2.7 (Stone-Weierstrass) we can write this new function as a limit of polynomials, i.e., we can write $W^\epsilon_n(x,y,z) = \sum_{m=1}^N \alpha_m(y,x)\beta_m(z,x)$ where $\alpha_m(y,x)$ and $\beta_m(z,x)$ are polynomials of degree m and N is sufficiently large (possibly infinite). The construction of this function is justified as follows:

We know that $e^{\frac{|x-y-z|^2}{\epsilon}} \to 0$ for $|x|,|y|,|z| \to \infty$. Therefore we can define a new continuous function with compact support by $W^\epsilon_n(x,y,z) = e^{\frac{|x-y-z|^2}{\epsilon}}$ for $|x|,|y|,|z| < w$ and $W^\epsilon_n(x,y,z)$ goes to zero linearly at w and remains zero for all $|x|,|y|,|z| \to \infty$. Therefore, for w large enough, we have $\|W^\epsilon_n(x,\cdot,\cdot) - e^{\frac{|x-\cdot-\cdot|^2}{\epsilon}}\|_\infty < \epsilon'$, where ϵ' can be made arbitrarily small. We are now ready to show pointwise convergence:

(61)

$$\begin{aligned}
&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left[K^{\epsilon,\delta}_{g_i,h_i}(x) - K^{\epsilon,\delta}_{g,h}(x)\right]dydz\\
=&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left[J^{\epsilon,\delta}_n(x,y,z) - (\pi\epsilon)^{-\frac{n}{2}}W^\epsilon_n(x,y,z)e^{-\delta(|x|^2+|y|^2+|z|^2)}\right]g_i(y)h_i(z)dydz\\
+&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(\pi\epsilon)^{-\frac{n}{2}}W^\epsilon_n(x,y,z)e^{-\delta(|x|^2+|y|^2+|z|^2)}\Big[g_i(y)h_i(z) - g(y)h(z)\Big]dydz\\
+&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left[(\pi\epsilon)^{-\frac{n}{2}}W^\epsilon_n(x,y,z)e^{-\delta(|x|^2+|y|^2+|z|^2)} - J^{\epsilon,\delta}_n(x,y,z)\right]g(y)h(z)dydz
\end{aligned}$$

Since $\|W^\epsilon_n(x,\cdot,\cdot) - e^{-\frac{|x-\cdot-\cdot|^2}{\epsilon}}\|_\infty < \epsilon'$, the second line of equation (61) becomes arbitrarily small as follows:

(62)

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left[J^{\epsilon,\delta}_n(x,y,z) - (\pi\epsilon)^{-\frac{n}{2}}W^\epsilon_n(x,y,z)e^{-\delta(|x|^2+|y|^2+|z|^2)}\right]g_i(y)h_i(z)dydz\right|\\
=&(\pi\epsilon)^{-\frac{n}{2}}\left|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left[W^\epsilon_n(x,y,z) - e^{-\frac{|x-y-z|^2}{\epsilon}}\right]e^{-\delta(|x|^2+|y|^2+|z|^2)}g_i(y)h_i(z)dydz\right|\\
\leq&(\pi\epsilon)^{-\frac{n}{2}}\|W^\epsilon_n(x,\cdot,\cdot) - e^{-\frac{|x-\cdot-\cdot|^2}{\epsilon}}\|_\infty\int_{\mathbb{R}^n}e^{-\delta|y|^2}|g_i(y)|dy\int_{\mathbb{R}^n}e^{-\delta|z|^2}|h_i(z)|dz\\
\leq&\epsilon'(\pi\epsilon)^{-\frac{n}{2}}\|e^{-\delta|\cdot|^2}\|_{q'}\|e^{-\delta|\cdot|^2}\|_{r'} \leq \frac{\epsilon''}{2}
\end{aligned}$$

In the third line of (62) we seperated our integrals and used Hölder's inequality (5) on both integrals. Since $e^{-\delta|\cdot|^2}$ has finite norm in every space, we can make (62) arbitrarily small.

The third line of equation (61) goes to zero by (14) as follows:

$$\begin{aligned}
&\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(\pi\epsilon)^{-\frac{n}{2}}W_n^\epsilon(x,y,z)e^{-\delta(|x|^2+|y|^2+|z|^2)}\left[g_i(y)h_i(z)-g(y)h(z)\right]dydz\\
=&\Bigg[(\pi\epsilon)^{-\frac{n}{2}}e^{-\delta|x|^2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\sum_{m=1}^{N}\alpha_m(y,x)\beta_m(z,x)e^{-\delta(|y|^2+|z|^2)}g_i(y)h_i(z)dydz\\
&-(\pi\epsilon)^{-\frac{n}{2}}e^{-\delta|x|^2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\sum_{m=1}^{N}\alpha_m(y,x)\beta_m(z,x)e^{-\delta(|y|^2+|z|^2)}g(y)h(z)dydz\Bigg]\\
=&\Bigg[(\pi\epsilon)^{-\frac{n}{2}}e^{-\delta|x|^2}\sum_{m=1}^{N}\int_{\mathbb{R}^n}\alpha_m(y,x)e^{-\delta|y|^2}g_i(y)dy\int_{\mathbb{R}^n}\beta_m(z,x)e^{-\delta|z|^2}h_i(z)dz\\
&-(\pi\epsilon)^{-\frac{n}{2}}e^{-\delta|x|^2}\sum_{m=1}^{N}\int_{\mathbb{R}^n}\alpha_m(y,x)e^{-\delta|y|^2}g(y)dy\int_{\mathbb{R}^n}\beta_m(z,x)e^{-\delta|z|^2}h(z)dz\Bigg]\\
=&0
\end{aligned}\tag{63}$$

In the fourth line of (63) as $i \to \infty$, we applied (14) and then we are subtracting identical functions. The last bracketed equation in (61) can be made arbitrarily small in exactly the same way as (62). Thus $K^{\epsilon,\delta}_{g_i,h_i}(x) \to K^{\epsilon,\delta}_{g,h}(x)$ pointwise.

(b) Uniform boundedness: From $e^{\frac{|x-y-z|^2}{\epsilon}}e^{-\delta|x|^2}$ being bounded by 1 in equation (53), we have

$$|K^{\epsilon,\delta}_{g_i,h_i}(x)| \le \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(\pi\epsilon)^{-\frac{n}{2}}e^{-\delta|z|^2}e^{-\delta|y|^2}|g_i(y)h_i(z)|dydz \tag{64}$$

Separating the integral in (64) and applying Hölder's inequality (5) to both integrals, we have $|K^{\epsilon,\delta}_{g_i,h_i}(x)| < (\pi\epsilon)^{-\frac{n}{2}}\|e^{-\delta|\cdot|^2}\|_{q'}\|e^{-\delta|\cdot|^2}\|_{r'}$. Since $e^{-\delta|\cdot|^2}$ has finite norm in every space, $K^{\epsilon,\delta}_{g_i,h_i}(x)$ is uniformly bounded (in x and in i).

(c) Multiply $K^{\epsilon,\delta}_{g_i,h_i}(x)$ by $e^{\gamma|x|^2}$ and show (a) and (b) again: We know that $K^{\epsilon,\delta}_{g_i,h_i}(x)$ is uniformly bounded, so if it is also converges to 0 for large x, then we will know that it has finite integral on $\mathbb{R}^n$. Multiplying $K^{\epsilon,\delta}_{g_i,h_i}(x)$ by $e^{\gamma|x|^2}$ for $\gamma < \delta$, we can prove pointwise convergence to $e^{\gamma|x|^2}K^{\epsilon,\delta}_{g,h}(x)$ using the exact same steps as in (i) above. Notice that $e^{\gamma|x|^2}$ is just a constant with respect to the integrals over dy and dz. Additionally, we can see that $e^{\gamma|x|^2}e^{\frac{|x-y-z|^2}{\epsilon}}e^{-\delta|x|^2}$ is bounded by 1 and uniform boundedness (ii) follows directly. Since $e^{\gamma|x|^2}K^{\epsilon,\delta}_{g_i,h_i}(x)$ both converges pointwise and is uniformly bounded in (x and i), $K^{\epsilon,\delta}_{g_i,h_i}(x)$ must go to zero for large x and therefore must be in $L^p(\mathbb{R}^n)$.

(d) Dominated Convergence We bound $|K^{\epsilon,\delta}_{g_i,h_i}(x)-K^{\epsilon,\delta}_{g,h}(x)|^p$ by $G(x)^p$, where $G(x) = \sup_i |K^{\epsilon,\delta}_{g_i,h_i}(x)|$. Since $G(x)$ is bounded (all $K^{\epsilon,\delta}_{g_i,h_i}(x)$ are bounded), it is an $L^p(\mathbb{R}^n)$ function (all $K^{\epsilon,\delta}_{g_i,h_i}(x)$ are in $L^p(\mathbb{R}^n)$). We can therefore apply the Dominated Convergence Theorem (18) and

$$\lim_{i\to\infty}\int_{\mathbb{R}^n}|K^{\epsilon,\delta}_{g_i,h_i}(x)-K^{\epsilon,\delta}_{g,h}(x)|^p dx = \int_{\mathbb{R}^n}|K^{\epsilon,\delta}_{g,h}(x)-K^{\epsilon,\delta}_{g,h}(x)|^p dx = 0 \tag{65}$$

Thus $K^{\epsilon,\delta}_{g_i,h_i}(x)$ converges strongly in $L^p(\mathbb{R}^n)$ as $i\to\infty$ to the function $K^{\epsilon,\delta}_{g,h}(x)$.

Now that we have strong convergence, the proof that we have maximizing functions $g(y)$ and $h(z)$ with $\|g\|_q = \|h\|_r = 1$ can be completed by showing that the inequalities we had, i.e., $\|g\|_q, \|h\|_r \le 1$ are indeed strict equalities. We know that $\|K^{\epsilon,\delta}_{g_i,h_i}\|_p \to C^{\epsilon,\delta}_n$, so strong convergence of $K^{\epsilon,\delta}_{g_i,h_i}$ to $K^{\epsilon,\delta}_{g,h}$ implies that $\|K^{\epsilon,\delta}_{g,h}\|_p = C^{\epsilon,\delta}_n$. We also know that if the norms of f and g were strictly less than 1, then the ratio $\frac{\|K^{\epsilon,\delta}_{g,h}\|_p}{\|g\|_q\|h\|_r}$ would be strictly greater than $C^{\epsilon,\delta}_n$, which is a contradiction to equation (60). Thus g and h are a maximizing pair of functions with $\|g\|_q = \|h\|_r = 1$ such that $\|K^{\epsilon,\delta}_{g,h}\|_p = C^{\epsilon,\delta}_n$. □

8.3. **The sharp constant in the auxiliary problem is separable.**

Lemma 8.2.

$$C^{\epsilon,\delta}_{n+m} = C^{\epsilon,\delta}_n C^{\epsilon,\delta}_m \tag{66}$$

Proof. We will show that $C^{\epsilon,\delta}_{n+m} \le C^{\epsilon,\delta}_n C^{\epsilon,\delta}_m$ and then show that we can saturate the inequality, i.e., $C^{\epsilon,\delta}_{n+m} = C^{\epsilon,\delta}_n C^{\epsilon,\delta}_m$. In the proof to follow, we will use Fubini's Theorem (8) several times to interchange the order of integration.

By Lemma 7.1,

$$C^{\epsilon,\delta}_{n+m} = \left(\int_{\mathbb{R}^{n+m}}\left|\int_{\mathbb{R}^{n+m}}\int_{\mathbb{R}^{n+m}} J^{\epsilon,\delta}_{n+m}(x,y,z)g(y)h(z)dydz\right|^p dx\right)^{\frac{1}{p}} \tag{67}$$

Let us write a point $x \in \mathbb{R}^{n+m}$ as $x = (x_1, x_2)$, where $x_1 \in \mathbb{R}^n$ and $x_2 \in \mathbb{R}^m$. Since $J^{\epsilon,\delta}_{n+m} = J^{\epsilon,\delta}_n J^{\epsilon,\delta}_m$ which can be seen in equation (54), we have

$$\begin{aligned} C^{\epsilon,\delta}_{n+m} = \Bigg(\int_{\mathbb{R}^m}\int_{\mathbb{R}^n}\Bigg|\int_{\mathbb{R}^{2m}}\int_{\mathbb{R}^{2n}} & J^{\epsilon,\delta}_m(x_2,y_2,z_2)J^{\epsilon,\delta}_n(x_1,y_1,z_1) \\ & \times g(y_1,y_2)h(z_1,z_2)dy_1dz_1dy_2dz_2\Bigg|^p dx_1dx_2\Bigg)^{\frac{1}{p}} \end{aligned} \tag{68}$$

Shifting the absolute value inside the third integral,

$$\begin{aligned} C^{\epsilon,\delta}_{n+m} \leq \Bigg(\int_{\mathbb{R}^m}\int_{\mathbb{R}^n}\Bigg(\int_{\mathbb{R}^{2m}}\Bigg|\int_{\mathbb{R}^{2n}} J^{\epsilon,\delta}_m(x_2,y_2,z_2)J^{\epsilon,\delta}_n(x_1,y_1,z_1) \\ \times g(y_1,y_2)h(z_1,z_2)dy_1dz_1\Bigg|dy_2dz_2\Bigg)^p dx_1dx_2\Bigg)^{\frac{1}{p}} \end{aligned} \tag{69}$$

Putting the equation in the form of Minkowski's inequality (19) with $f(x,y) = \Big|J^{\epsilon,\delta}_m(x_2,y_2,z_2)\int_{\mathbb{R}^{2n}} J^{\epsilon,\delta}_n(x_1,y_1,z_1)g(y_1,y_2)h(z_1,z_2)dy_1dz_1\Big|$,

$$\begin{aligned} C^{\epsilon,\delta}_{n+m} \leq \Bigg(\int_{\mathbb{R}^m}\Bigg[\Bigg(\int_{\mathbb{R}^n}\Bigg(\int_{\mathbb{R}^{2m}}\Bigg|J^{\epsilon,\delta}_m(x_2,y_2,z_2)\int_{\mathbb{R}^{2n}} J^{\epsilon,\delta}_n(x_1,y_1,z_1) \\ \times g(y_1,y_2)h(z_1,z_2)dy_1dz_1\Bigg|dy_2dz_2\Bigg)^p dx_1\Bigg)^{\frac{1}{p}}\Bigg]^p dx_2\Bigg)^{\frac{1}{p}} \end{aligned} \tag{70}$$

Applying Minkowski's inequality (19) to the integral in square brackets,

$$\begin{aligned} C^{\epsilon,\delta}_{n+m} \leq \Bigg(\int_{\mathbb{R}^m}\Bigg[\int_{\mathbb{R}^{2m}}\Bigg(\int_{\mathbb{R}^n}\Bigg|J^{\epsilon,\delta}_m(x_2,y_2,z_2)\int_{\mathbb{R}^{2n}} J^{\epsilon,\delta}_n(x_1,y_1,z_1) \\ \times g(y_1,y_2)h(z_1,z_2)dy_1dz_1\Bigg|^p dx_1\Bigg)^{\frac{1}{p}} dy_2dz_2\Bigg]^p dx_2\Bigg)^{\frac{1}{p}} \end{aligned} \tag{71}$$

Moving $J^{\epsilon,\delta}_m$ outside the integrals over $\mathbb{R}^n$,

$$\begin{aligned} C^{\epsilon,\delta}_{n+m} \leq \Bigg(\int_{\mathbb{R}^m}\Bigg(\int_{\mathbb{R}^{2m}} J^{\epsilon,\delta}_m(x_2,y_2,z_2)\Bigg[\int_{\mathbb{R}^n}\Bigg|\int_{\mathbb{R}^{2n}} J^{\epsilon,\delta}_n(x_1,y_1,z_1) \\ \times g(y_1,y_2)h(z_1,z_2)dy_1dz_1\Bigg|^p dx_1\Bigg]^{\frac{1}{p}} dy_2dz_2\Bigg)^p dx_2\Bigg)^{\frac{1}{p}} \end{aligned}$$

Applying equation (60) to the integrals in square brackets

$$\begin{aligned} C^{\epsilon,\delta}_{n+m} \leq \Bigg(\int_{\mathbb{R}^m}\Bigg(\int_{\mathbb{R}^{2m}} J^{\epsilon,\delta}_m(x_2,y_2,z_2) \\ \times C^{\epsilon,\delta}_n\|g(\cdot,y_2)\|_q\|h(\cdot,z_2)\|_r dy_2dz_2\Bigg)^p dx_2\Bigg)^{\frac{1}{p}} \end{aligned}$$

Moving $C^{\epsilon,\delta}_n$ outside and applying equation (60) once again

$$C^{\epsilon,\delta}_{n+m} \leq C^{\epsilon,\delta}_n C^{\epsilon,\delta}_m \|\,\|g(\cdot,y_2)\|_q\|_q\ \|\,\|h(\cdot,z_2)\|_r\|_r$$

where

$$\begin{aligned} \|\,\|g(\cdot,y_2)\|_q\|_q &= \Bigg(\int_{\mathbb{R}^m}\Bigg|\int_{\mathbb{R}^n}|g(y_1,y_2)|^p dy_1\Bigg|^{p\frac{1}{p}} dy_2\Bigg)^{\frac{1}{p}} \\ &= \Bigg(\int_{\mathbb{R}^{n+m}}|g(y)|^p dy\Bigg)^{\frac{1}{p}} \\ &= \|g\|_q \end{aligned}$$

Similarly $\| \|h(\cdot, z_2)\|_r\|_r = \|h\|_r$, so we have

$$C_{n+m}^{\epsilon,\delta} \leq C_n^{\epsilon,\delta} C_m^{\epsilon,\delta} \|g\|_q \|h\|_r$$

By Lemma 7.1 $\|g\|_q = \|h\|_r = 1$ and therefore

$$C_{n+m}^{\epsilon,\delta} \leq C_n^{\epsilon,\delta} C_m^{\epsilon,\delta} \tag{72}$$

To complete the proof we need to show that we can saturate inequality (72) for optimizers g_1, h_1 and g_2, h_2,

$$\begin{aligned} C_n^{\epsilon,\delta} C_m^{\epsilon,\delta} = &\left(\int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} J_n^{\epsilon,\delta}(x_1, y_1, z_1) g_1(y_1) h_1(z_1) dy_1 dz_1 \right|^p dx_1 \right)^{\frac{1}{p}} \\ &\times \left(\int_{\mathbb{R}^m} \left| \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} J_m^{\epsilon,\delta}(x_2, y_2, z_2) g_2(y_2) h_2(z_2) dy_2 dz_2 \right|^p dx_2 \right)^{\frac{1}{p}} \end{aligned}$$

Placing the equation corresponding to $C_m^{\epsilon,\delta}$ inside the equation corresponding to $C_n^{\epsilon,\delta}$, which can be done since it is just a constant.

$$\begin{aligned} C_n^{\epsilon,\delta} C_m^{\epsilon,\delta} = &\left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} \left| \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} J_m^{\epsilon,\delta}(x_2, y_2, z_2) g_2(y_2) h_2(z_2) dy_2 dz_2 \right|^p dx_2 \right) \right. \\ &\left. \times \left| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} J_n^{\epsilon,\delta}(x_1, y_1, z_1) g_1(y_1) h_1(z_1) dy_1 dz_1 \right|^p dx_1 \right)^{\frac{1}{p}} \end{aligned}$$

Rearranging terms,

$$\begin{aligned} C_n^{\epsilon,\delta} C_m^{\epsilon,\delta} = &\left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} \left| \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} J_n^{\epsilon,\delta}(x_1, y_1, z_1) g_1(y_1) h_1(z_1) dy_1 dz_1 \right. \right. \right. \\ &\left. \left. \left. \times J_m^{\epsilon,\delta}(x_2, y_2, z_2) g_2(y_2) h_2(z_2) dy_2 dz_2 \right|^p dx_1 dx_2 \right) \right)^{\frac{1}{p}} \end{aligned}$$

Finally, letting $g(y_1, y_2) := g_1(y_1) g_2(y_2)$ and $h(z_1, z_2) := h_1(z_1) h_2(z_2)$

$$\begin{aligned} C_n^{\epsilon,\delta} C_m^{\epsilon,\delta} = &\left(\int_{\mathbb{R}^{n+m}} \left| \int_{\mathbb{R}^{n+m}} \int_{\mathbb{R}^{n+m}} \right. \right. \\ &\left. \left. \times J_{n+m}^{\epsilon,\delta}(x, y, z) g(y) h(z) dy dz \right|^p dx \right)^{\frac{1}{p}} = C_{n+m}^{\epsilon,\delta} \end{aligned} \tag{73}$$

□

8.4. **Optimizers Must Factorize.**

Lemma 8.3. *If $g(y_1, y_2)$ and $h(z_1, z_2)$ are any pair of optimizers of the $m + n$-dimensional problem, then $g(y_1, y_2) = g_1(y_1) g_2(y_2)$ and $h(z_1, z_2) = h_1(z_1) h_2(z_2)$, where $g_1(y_1), h_1(y_1)$ and $g_2(y_2) h_2(y_2)$ are optimizers of the corresponding n- and m-dimensional problems respectively.*
Note: an immediate consequence is that all optimizers must be products of optimizers of the one-dimensional problem.

Proof. We know that when two optimizers $g(y_1, y_2)$ and $h(z_1, z_2)$ are multiplied, their product is positive for all x, e.g. both functions are negative or both functions are positive for all x. For these optimizers, We must have equality in all equations following equation (67). The first inequality (69) due to shifting the absolute value inwards is an equality as a result of the positivity of the product f times g. We must also have equality in the application of Minkowski's inequality (71), i.e.

$$\begin{aligned}(74)\quad &\Bigg(\int_{\mathbb{R}^n}\Bigg(\int_{\mathbb{R}^{2m}} J_m^{\epsilon,\delta}(x_2, y_2, z_2)\int_{\mathbb{R}^{2n}} J_n^{\epsilon,\delta}(x_1, y_1, z_1)\\ &\qquad\qquad \times g(y_1, y_2)h(z_1, z_2)dy_1dz_1dy_2dz_2\Bigg)^p dx_1\Bigg)^{\frac{1}{p}}\\ =&\int_{\mathbb{R}^{2m}}\Bigg(\int_{\mathbb{R}^n}\Bigg(J_m^{\epsilon,\delta}(x_2, y_2, z_2)\int_{\mathbb{R}^{2n}} J_n^{\epsilon,\delta}(x_1, y_1, z_1)\\ &\qquad\qquad \times g(y_1, y_2)h(z_1, z_2)dy_1dz_1\Bigg)^p dx_1\Bigg)^{\frac{1}{p}} dy_2dz_2\end{aligned}$$

In order to achieve equality in (74), by Theorem 2.8 (A Weaker form of Minkowski's inequality) there must exist two functions $A_{x_2}(x_1)$ and $B_{x_2}(y_2, z_2)$ (let us assume both functions are positive everywhere) such that

$$\begin{aligned}(75)\quad &J_m^{\epsilon,\delta}(x_2, y_2, z_2)\int_{\mathbb{R}^{2n}} J_n^{\epsilon,\delta}(x_1, y_1, z_1)g(y_1, y_2)h(z_1, z_2)dy_1dz_1\\ =&A_{x_2}(x_1)B_{x_2}(y_2, z_2)\end{aligned}$$

in which case we have the needed equality in (71), i.e.,

$$\begin{aligned}&\Bigg(\int_{\mathbb{R}^n}\Bigg(\int_{\mathbb{R}^{2m}} A_{x_2}(x_1)B_{x_2}(y_2, z_2)dy_2dz_2\Bigg)^p dx_1\Bigg)^{\frac{1}{p}}\\ (76)\quad =&\Bigg(\int_{\mathbb{R}^n} A_{x_2}(x_1)^p dx_1\Bigg)^{\frac{1}{p}}\int_{\mathbb{R}^{2m}} B_{x_2}(y_2, z_2)dy_2dz_2\\ =&\int_{\mathbb{R}^{2m}}\Bigg(\int_{\mathbb{R}^n}\Bigg(A_{x_2}(x_1)B_{x_2}(y_2, z_2)\Bigg)^p dx_1\Bigg)^{\frac{1}{p}} dy_2dz_2\end{aligned}$$

Multiplying both sides of (75) by $J_m^{\epsilon,\delta}(x_2, y_2, z_2)^{-1}$, so that both sides are independent of x_2, we can form functions $C(x_1)$ and $D(y_2, z_2)$ such that

$$(77)\quad C(x_1)D(y_2, z_2) = J_m^{\epsilon,\delta}(x_2, y_2, z_2)^{-1}A_{x_2}(x_1)B_{x_2}(y_2, z_2)$$

in which case the integrand in the inner bracket of the first half of equation (74) becomes

$$\begin{aligned}(78)\quad &\\ &\int_{\mathbb{R}^{2m}} J_m^{\epsilon,\delta}(x_2, y_2, z_2)\int_{\mathbb{R}^{2n}} J_n^{\epsilon,\delta}(x_1, y_1, z_1)g(y_1, y_2)h(z_1, z_2)dy_1dz_1dy_2dz_2\\ =&\int_{\mathbb{R}^{2m}} A_{x_2}(x_1)B_{x_2}(y_2, z_2)dy_2dz_2\\ =&\int_{\mathbb{R}^{2m}} J_m^{\epsilon,\delta}(x_2, y_2, z_2)C(x_1)D(y_2, z_2)dy_2dz_2\\ =&C(x_1)E(x_2)\end{aligned}$$

for some function E.

We can show that optimizers factorize. If we look at our auxiliary problem (60) as a function of f, g and h, i.e.,

$$\int_{\mathbb{R}^{m+n}} \int_{\mathbb{R}^{m+n}} \int_{\mathbb{R}^{m+n}} J^{\epsilon,\delta}_{m+n} f(x)g(y)h(z)dxdydz \leq C^{\epsilon,\delta}_{m+n} \|f\|_{p'} \|g\|_q \|h\|_r \tag{79}$$

(recall that in this case $f \in L^{p'}(\mathbb{R}^n)$) then for optimizers f, g and h, by substituing our result in (78), equation (70) becomes

$$\begin{aligned} C^{\epsilon,\delta}_{n+m} &= \left(\int_{\mathbb{R}^m} \int_{\mathbb{R}^n} f(x_1, x_2) \Big(C(x_1)E(x_2) \Big)^p dx_1 dx_2 \right)^{\frac{1}{p}} \\ &\leq \|C(\cdot)E(\cdot)\|_p \|f\|_{p'} \end{aligned} \tag{80}$$

where by Hölder's inequality (5) (ii) we have equality in (80) only when $f(x_1, x_2) = \lambda[C(x_1)E(x_2)]^{p-1}$, where λ is a real constant.

Since f, g and h play a symmetric role (39), we can conclude that all optimizers must factorize. By Lemma 7.2, it is clear that each of these factors must be an optimizer of the corresponding n- or m-dimensional problem. ☐

8.5. **Optimizers are Gaussian functions.**

Lemma 8.4. *The optimizers of the auxiliary problem* (60) *are Gaussian functions.*

Proof. We will show that strictly positive C^∞-optimizers must be Gaussians and then show by approximation of C^∞-functions that the limiting case, which is not neccesarily a C^∞-optimizer, must indeed be a Guassian function. The assumption that our functions are stricly positive is not a problem since optimizers must multiply together to form a positive function, therefore we can assume them to be positive individually.

Let g and h be optimizers of the one dimensional problem. Consider

$$G(x_1, x_2) = g\left(\frac{x_1 + x_2}{\sqrt{2}}\right) g\left(\frac{x_1 - x_2}{\sqrt{2}}\right)$$

and

$$H(x_1, x_2) = h\left(\frac{x_1 + x_2}{\sqrt{2}}\right) h\left(\frac{x_1 - x_2}{\sqrt{2}}\right)$$

It is crucial that

$$\begin{aligned} &J^{\epsilon,\delta}_1 \left(\frac{x_1 + x_2}{\sqrt{2}}, \frac{y_1 + y_2}{\sqrt{2}}, \frac{z_1 + z_2}{\sqrt{2}} \right) J^{\epsilon,\delta}_1 \left(\frac{x_1 - x_2}{\sqrt{2}}, \frac{y_1 - y_2}{\sqrt{2}}, \frac{z_1 - z_2}{\sqrt{2}} \right) \\ =&J^{\epsilon,\delta}_1(x_1, y_1, z_1) J^{\epsilon,\delta}_1(x_2, y_2, z_2) \end{aligned} \tag{81}$$

or equaivalently (where we leave out the factor $(\pi\epsilon)^{-\frac{n}{2}}$ for simplicity):

(82)

$$\begin{aligned}
&\exp\Bigg(-\frac{\left|\frac{x_1+x_2}{\sqrt{2}}-\frac{y_1+y_2}{\sqrt{2}}-\frac{z_1+z_2}{\sqrt{2}}\right|^2}{\epsilon}-\delta\left|\frac{y_1+y_2}{\sqrt{2}}\right|^2-\delta\left|\frac{x_1+x_2}{\sqrt{2}}\right|^2-\delta\left|\frac{z_1+z_2}{\sqrt{2}}\right|^2\\
&\qquad-\frac{\left|\frac{x_1-x_2}{\sqrt{2}}-\frac{y_1-y_2}{\sqrt{2}}-\frac{z_1-z_2}{\sqrt{2}}\right|^2}{\epsilon}-\delta\left|\frac{y_1-y_2}{\sqrt{2}}\right|^2-\delta\left|\frac{x_1-x_2}{\sqrt{2}}\right|^2-\delta\left|\frac{z_1-z_2}{\sqrt{2}}\right|^2\Bigg)\\
&=\exp\Bigg(-\frac{|x_1-y_1-z_1|^2}{\epsilon}-\delta|y_1|^2-\delta|x_1|^2-\delta|z_1|^2\\
&\qquad\qquad-\frac{|x_2-y_2-z_2|^2}{\epsilon}-\delta|y_2|^2-\delta|x_2|^2-\delta|z_2|^2\Bigg)
\end{aligned}$$

We can show equality in (82) in the δ terms as follows:

$$\begin{aligned}
\left|\frac{y_1+y_2}{\sqrt{2}}\right|^2+\left|\frac{y_1-y_2}{\sqrt{2}}\right|^2&=\frac{(y_1+y_2)^2}{2}+\frac{(y_1-y_2)^2}{2}\\
&=\frac{y_1^2+2y_1y_2+y_2^2}{2}+\frac{y_1^2-2y_1y_2+y_2^2}{2}=|y_1|^2+|y_2|^2
\end{aligned}$$

and the same can be shown for x and z. Equality in the ϵ terms is as follows:

$$\begin{aligned}
&\left|\frac{x_1+x_2}{\sqrt{2}}-\frac{y_1+y_2}{\sqrt{2}}-\frac{z_1+z_2}{\sqrt{2}}\right|^2+\left|\frac{x_1+x_2}{\sqrt{2}}-\frac{y_1+y_2}{\sqrt{2}}-\frac{z_1+z_2}{\sqrt{2}}\right|^2\\
=&\frac{1}{2}(x_1+x_2-y_1-y_2-z_1-z_2)^2+\frac{1}{2}(x_1-x_2-y_1+y_2-z_1+z_2)^2\\
=&\frac{1}{2}[(x_1-y_1-z_1)+(x_2-y_2-z_2)]^2+\frac{1}{2}[(x_1-y_1-z_1)-(x_2-y_2-z_2)]^2\\
=&|x_1-y_1-z_1|^2+|x_2-y_2-z_2|^2
\end{aligned}$$

By Lemma 7.4, since G is an optimizer, we have that

$$g\left(\frac{x_1+x_2}{\sqrt{2}}\right)g\left(\frac{x_1-x_2}{\sqrt{2}}\right)=u(x_1)v(x_2) \tag{83}$$

for some functions u and v, where $u(x_1)v(x_2)\in L^q(\mathbb{R}^2)$. We will prove that this relation implies that g must be a Gaussian function.

Assume that g is in C^∞ and strictly positive. Then the functions $\eta(x):=\log g(x), \mu(x):=\log u(x)$ and $\nu(x):=\log v(x)$ are also in C^∞ and satisfy the relation:

$$\eta\left(\frac{x_1+x_2}{\sqrt{2}}\right)+\eta\left(\frac{x_1-x_2}{\sqrt{2}}\right)=\mu(x_1)+\nu(x_2) \tag{84}$$

Diffentiating equation (84) with respect to x_1, (where $\frac{\partial\eta(x(x_1,x_2))}{\partial x_1}=\eta'(x(x_1,x_2))\frac{\partial x(x_1,x_2)}{\partial x_1}$)

$$\frac{1}{\sqrt{2}}\eta'\left(\frac{x_1+x_2}{\sqrt{2}}\right)+\frac{1}{\sqrt{2}}\eta'\left(\frac{x_1-x_2}{\sqrt{2}}\right)=\mu'(x_1) \tag{85}$$

and differentiating (85) with respect to x_2

$$\frac{1}{2}\eta''\left(\frac{x_1+x_2}{\sqrt{2}}\right)-\frac{1}{2}\eta''\left(\frac{x_1-x_2}{\sqrt{2}}\right)=0 \tag{86}$$

which can be simplified to

$$\eta''\left(\frac{x_1+x_2}{\sqrt{2}}\right)=\eta''\left(\frac{x_1-x_2}{\sqrt{2}}\right) \tag{87}$$

for all x_1 and x_2, which implies that $\eta''(x)$ must be a constant d. We know that $\eta''=\frac{g(x)''g(x)-g'(x)^2}{g(x)^2}$, where $g(x), g'(x)^2$ are always positive, and $g''(x)=dg(x)^2+g'(x)^2$. Since $g\in L^q(\mathbb{R})$ and g is always positive, $g''(x)$ must be negative for some x and therefore d must be a negative constant. Integrating $\eta''(x)$ twice, we have $\eta(x)=-ax^2+2bx+c$ for some constants b and c and $a=-d>0$. Therefore $g(x)=\exp[-ax^2+bx+c]$, i.e., g is a Gaussian function. But we have made the assumption that g is in C^∞. To apply this argument to the original function g which is only in $L^q(\mathbb{R})$, we consider

$$\begin{aligned}
g_\lambda(x)&=\left(\frac{\lambda}{\pi}\right)^{\frac{1}{2}}\int_{\mathbb{R}}e^{-\lambda(x-y)^2}g(y)dy\\
u_\lambda(x)&=\left(\frac{\lambda}{\pi}\right)^{\frac{1}{2}}\int_{\mathbb{R}}e^{-\lambda(x-y)^2}u(y)dy\\
v_\lambda(x)&=\left(\frac{\lambda}{\pi}\right)^{\frac{1}{2}}\int_{\mathbb{R}}e^{-\lambda(x-y)^2}v(y)dy
\end{aligned}$$

where since g is non-negative, g_λ is strictly positive and clearly in C^∞. We will show that (83) holds for $g_\lambda, u_\lambda, v_\lambda$ in place of g, u, v respectively, as follows:

$$\begin{aligned}
&g_\lambda\left(\frac{x_1+x_2}{\sqrt{2}}\right)g_\lambda\left(\frac{x_1-x_2}{\sqrt{2}}\right)\\
=&\int_{\mathbb{R}^2}e^{-\lambda\left(\frac{x_1+x_2}{\sqrt{2}}-y\right)^2}e^{-\lambda\left(\frac{x_1-x_2}{\sqrt{2}}-z\right)^2}g(y)g(z)dydz\\
=&\int_{\mathbb{R}^2}e^{-\lambda\left(\frac{x_1+x_2}{\sqrt{2}}-\frac{y_1+y_2}{\sqrt{2}}\right)^2}e^{-\lambda\left(\frac{x_1-x_2}{\sqrt{2}}-\frac{y_1-y_2}{\sqrt{2}}\right)^2}g\left(\frac{y_1+y_2}{\sqrt{2}}\right)g\left(\frac{y_1-y_2}{\sqrt{2}}\right)dy_1dy_2\\
=&\int_{\mathbb{R}^2}e^{-\lambda\left(\frac{x_1+x_2}{\sqrt{2}}-\frac{y_1+y_2}{\sqrt{2}}\right)^2}e^{-\lambda\left(\frac{x_1-x_2}{\sqrt{2}}-\frac{y_1-y_2}{\sqrt{2}}\right)^2}u(y_1)v(y_2)dy_1dy_2\\
=&\int_{\mathbb{R}^2}e^{-\lambda(x_1-y_1)^2}e^{-\lambda(x_2-y_2)^2}u(y_1)v(y_2)dy_1dy_2\\
=&u_\lambda(x_1)v_\lambda(x_2)
\end{aligned} \tag{88}$$

In line three of (88) we introduce the change of variable in 2-dimension $y=\frac{y_1+y_2}{\sqrt{2}}$ and $z=\frac{y_1-y_2}{\sqrt{2}}$ with Jacobian

$$\frac{\partial(z,y)}{\partial(y_1,y_2)}=\begin{vmatrix}\frac{\partial z}{\partial y_1} & \frac{\partial z}{\partial y_2}\\ \frac{\partial y}{\partial y_1} & \frac{\partial y}{\partial y_2}\end{vmatrix}=\begin{vmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{vmatrix}=1$$

In line four we substitute in the result of (83). Line five is achieved as a result of the following calculation:

$$\begin{aligned}
&[(x_1+x_2)-(y_1+y_2)]^2+[(x_1-x_2)-(y_1-y_2)]^2\\
&=[(x_1-y_1)+(x_2-y_2)]^2+[(x_1-y_1)-(x_2-y_2)]^2\\
&=[(x_1-y_1)^2+2(x_1-y_1)(x_2-y_2)\\
&+(x_2-y_2)^2]+[(x_1-y_1)^2-2(x_1-y_1)(x_2-y_2)+(x_2-y_2)^2]\\
&=2[(x_1-y_1)^2+(x_2-y_2)^2]
\end{aligned}$$

Now that we have shown equation (83) holds for $g_\lambda, u_\lambda, v_\lambda$, the same argument as before follows and therefore

$$g_\lambda(x)=\exp[-a_\lambda x^2+b_\lambda x+c_\lambda]$$

with $a_\lambda > 0$. By Theorem 2.3 (Approximation by C^∞-functions) there exists a sequence $\lambda_j \to \infty$ such that $g_{\lambda_j}(x) \to g(x)$ for almost every $x \in \mathbb{R}$. Therefore $a_\lambda, b_\lambda x$ and c_λ must converge and we call the limits a, b and c respectively.

The result for h can be shown in exactly the same way. Thus the optimizers of the auxiliary inequality (60) are given by Gaussian functions. □

8.6. **The Sharp Constant.**

Lemma 8.5. *The sharp constant in Young's inequality is the limit of the sharp constant in the auxiliary problem, i.e.,*

$$\lim_{\epsilon,\delta\to 0} C_1^{\epsilon,\delta} = C_q C_r C_{p'}$$

Proof. We will show

(a)
$$\lim_{\delta\to 0}\|K_{g,h}^{\epsilon,\delta}\|_p=\|K_{g,h}^{\epsilon,0}\|_p$$

(b)
$$\lim_{\delta\to 0} C_1^{\epsilon,\delta}=C_1^{\epsilon,0}$$

(c)
$$C_1^{\epsilon,0}=C_qC_rC_{p'}$$

(d)
$$\lim_{\epsilon\to 0} C_1^{\epsilon,0}=C_qC_rC_{p'}$$

so that the sharp constant in Young's inequality is obtained.

(a) $\lim_{\delta\to 0}\|K_{g,h}^{\epsilon,\delta}\|_p=\|K_{g,h}^{\epsilon,0}\|_p$**:**
Since $J_n^{\epsilon,\delta}$ (54) is a smooth Gaussian function, $K_{g,h}^{\epsilon,\delta}$ converges pointwise in x to $K_{g,h}^{\epsilon,0}$. Since we have seen in equations (56) and (58) that $K_{g,h}^{\epsilon,\delta} \le K_{g,h}^{\epsilon,0}$ and $\|K_{g,h}^{\epsilon,0}\|_p < \infty$, we can apply the Dominated Convergence Theorem (18). Thus, $\lim_{\delta\to 0}\|K_{g,h}^{\epsilon,\delta}\|_p=\|K_{g,h}^{\epsilon,0}\|_p$

(b) $\lim_{\delta\to 0} C_1^{\epsilon,\delta} = C_1^{\epsilon,0}$ **:**
We know that $C_1^{\epsilon,\delta}$ is non-increasing in δ as a result of (55). We have already seen (59) that $C_1^{\epsilon,\delta}$ is bounded by $C_1^{\epsilon,0}$. For any $\eta > 0$ there exists non-negative, normalized optimizing functions g, h such that $\|K_{g,h}^{\epsilon,0}\|_p \geq C_1^{\epsilon,0} - \eta$. Our auxiliary equation (60) clearly shows that $\|K_{g,h}^{\epsilon,\delta}\|_p \leq C_1^{\epsilon,\delta}$ and therefore using the Monotone Convergence Theorem (21)

$$\lim_{\delta\to 0} C_1^{\epsilon,\delta} \geq \lim_{\delta\to 0} \|K_{g,h}^{\epsilon,\delta}\|_p = \|K_{g,h}^{\epsilon,0}\|_p \geq C_1^{\epsilon,0} - \eta \tag{89}$$

where the equality in (89) was shown in part (a). Since η can be made arbitrarily small, $\lim_{\delta\to 0} C_1^{\epsilon,\delta} = C_1^{\epsilon,0}$.

(c) $C_1^{\epsilon,0} = C_q C_r C_{p'}$**:**
We can write

$$C_1^{\epsilon,0} = \sup_{\delta>0} C_1^{\epsilon,\delta} = \sup_{\delta>0}\sup_{g,h} \|K_{g,h}^{\epsilon,\delta}\|_p \tag{90}$$

where g, h are non-negative Gaussians normalized with respect to q, r. By interchanging the supremums in (90) we can see that $C_1^{\epsilon,0}$ can be computed by taking the supremum over Gaussian functions. However at this point we only know that we have optimizing functions with δ dependence. Consider only normalized functions. We will show that the supremums can be interchanged and that we do have optimizing functions independent of δ in the following four steps:

(i) $\|K_{g,h}^{\epsilon,\delta}\|_p$ is continuous with respect to $\delta > 0$ for any $g \in L^p, h \in L^r$.
(ii) $\sup_{g,h} \|K_{g,h}^{\epsilon,\delta}\|_p$ is continuous with respect to $\delta > 0$ for Gaussian optimizers with δ dependence.
(iii) $\sup_{\delta>0}\sup_{g,h} \|K_{g,h}^{\epsilon,\delta}\|_p \geq \sup_{g,h}\sup_{\delta>0} \|K_{g,h}^{\epsilon,\delta}\|_p$.
(iv) $\sup_{\delta>0}\sup_{g,h} \|K_{g,h}^{\epsilon,\delta}\|_p \leq \sup_{g,h}\sup_{\delta>0} \|K_{g,h}^{\epsilon,\delta}\|_p$.

and then we will can interchange the supremums.

(i) $\|K_{g,h}^{\epsilon,\delta}\|_p$ **is continuous with respect to** $\delta > 0$ **for any** $g \in L^p, h \in L^r$**:**
We can use the Dominated Convergence Theorem (18) as in part (a) to interchange the limit and integral, where the limit is over some $\eta \to 0$ of $K_{g,h}^{\epsilon,\delta-\eta}$ where g, h are fixed with respect to δ.

(ii) $\sup_{g,h} \|K_{g,h}^{\epsilon,\delta}\|_p$ **is continuous with respect to** $\delta > 0$ **for Gaussian optimizers with** δ **dependence:**
For each optimizing Gaussian function, fixed with respect to δ, we have that $K_{g,h}^{\epsilon,\delta}$ increases as $\delta \to 0$ by equation (56) and is continuous as a result of part (i). Consider new optimizers at each value of δ (obviously δ dependent). Let us denote the optimizers at δ' by g', h' and at δ'' by g'', h''. Then for $\delta'' - \delta' < \eta$, we have that $\|K_{g'',h''}^{\epsilon,\delta''}\|_p - \|K_{g',h'}^{\epsilon,\delta'}\|_p$ can be made arbirarily small by taking η sufficiently small. Therefore, $\sup_{g,h} \|K_{g,h}^{\epsilon,\delta}\|_p$ is continuos with respect to δ.

(iii) $\sup_{\delta>0}\sup_{g,h}\|K^{\epsilon,\delta}_{g,h}\|_p \geq \sup_{g,h}\sup_{\delta>0}\|K^{\epsilon,\delta}_{g,h}\|_p$**:**
We have that from part (ii) above that we can come arbitrarily close to $\delta = 0$ with optimizing functions (possibly δ dependent). Therefore for every $\delta > 0$, $\sup_{g,h}\|K^{\epsilon,0}_{g,h}\|_p \geq \|K^{\epsilon,0}_{g_o,h_o}\|_p$, where g, h are Gaussian optimizers (possible with δ dependence) and g_o, h_o are Gaussian functions independent of δ which we can now introduce to be the most optimizing function for $\delta = 0$ (at this point, not yet optimizers of the auxiliary inequality (60)), i.e., $\|K^{\epsilon,0}_{g_o,h_o}\|_p > \sup_{g*,h*}\|K^{\epsilon,0}_{g*,h*}\|_p$, where $g*, h*$ are not δ dependent. Now by taking the supremum over δ we have $\sup_{\delta>0}\sup_{g,h}\|K^{\epsilon,\delta}_{g,h}\|_p \geq \sup_{g*,h*}\|K^{\epsilon,\delta}_{g_o,h_o}\|_p$.

(iv) $\sup_{\delta>0}\sup_{g,h}\|K^{\epsilon,\delta}_{g,h}\|_p \leq \sup_{g,h}\sup_{\delta>0}\|K^{\epsilon,\delta}_{g,h}\|_p$
Let us denote optimizing functions for a fixed δ by g', h'. Then

$$\begin{aligned}
&\sup_{\delta>0}\sup_{g,h}\|K^{\epsilon,\delta}_{g,h}\|_p = \sup_{\delta>0}\|K^{\epsilon,\delta}_{g',h'}\|_p\\
\leq&\|K^{\epsilon,0}_{g',h'}\|_p \leq \sup_{g,h}\|K^{\epsilon,0}_{g,h}\|_p\\
=&\sup_{g,h}\sup_{\delta>0}\|K^{\epsilon,\delta}_{g,h}\|_p
\end{aligned}$$

where the final line above is a direct consequence of the Dominated Convergence Theorem justified in part (a).

Combining the results of part (iii) and part (iv), we can interchange the supremums in (90) and compute $C^{\epsilon,0}_1 = SC$, where SC is the computed sharp constant.

(d) $\lim_{\epsilon\to 0} C^{\epsilon,0}_1 = C_q C_r C_{p'}$**:**
We already know that $C^{\epsilon,0}_1 \leq C_{p',q,r;1}$ from equation (59). For every $\eta > 0$ there exist normalized g, h such that $\|g * h\|_p \geq C_{p',q,r;1} - \eta$. From eqaution (58) we have $C^{\epsilon,0}_1 \geq \|j_{\sqrt{\epsilon}} * (g * h)\|_p$. Since by Theorem 2.3 $j_{\sqrt{\epsilon}} * (g * h) \to g * h$ in $L^p(\mathbb{R})$ and since $\|j_{\sqrt{\epsilon}} * (g * h)\|_p$ is continuous by the simple version of Young's inequality, we have that $\liminf_{\epsilon\to 0} C^{\epsilon,0}_1 \geq C_{p',q,r;1} - \eta$. Thus $C_{p',q,r;1} = SC$.

We have already shown in Section 7 (Gaussian Optimizers) that $C_{p',q,r;1} = C_q C_r C_{p'}$. Thus concludes the proof of Young's inequality. □

9. An Application of the Sharp Constant

A. Wherl introduced a new definition of the "classical" entropy corresponding to a quantum system and proved that it had several interesting properties. Elliot H. Lieb used the sharp constant to prove A. Wherl's conjecture about the minimum value of this "classical" entropy. (Proof of an Entropy Conjecture by Wehrl, Computations in Mathematical Physics 1978)

ACKNOWLEDGEMENTS

Thank you to Daniel Ueltschi and Jan Wehr for much needed help in working through this proof.

REFERENCES

ELLIOTT H. LIEB AND MICHAEL LOSS, *Analysis*, AMS, 2001.

GERALD B. FOLLAND, *Real Analysis*, John Wiley and Sons, Inc., 1999.

FURTHER READING

H.J. BRASCAMP AND ELLIOTT H. LIEB, *Best constants in Young's inequality, its converse, and its generalization to more than three functions*, Adv. in Math. **20** (1976), 151-173.

ELLIOTT H. LIEB, *Gaussian kernels have only Gaussian maximizers*, Invent. Math. **102** (1990), 179-208.

INDEX