Invent math. 115, 463–482 (1994)

Inventiones mathematicae

Sharp uniform convexity and smoothness inequalities for trace norms

Keith Ball[1,*], Eric A. Carlen[2,], and Elliott H. Lieb[3,***]**

[1] Department of Mathematics, Texas A&M University, College Station, TX, 77843, USA
[2] School of Mathematics, Georgia Institute of Technology, Atlanta, GA, 30332, USA
[3] Departments of Mathematics and Physics, Princeton University, P.O. Box 708, Princeton, NJ, 08544, USA

Oblatum 7-VII-1993

Summary. We prove several sharp inequalities specifying the uniform convexity and uniform smoothness properties of the Schatten trace ideals C_p, which are the analogs of the Lebesgue spaces L_p in non-commutative integration. The inequalities are all precise analogs of results which had been known in L_p, but were only known in C_p for special values of p. In the course of our treatment of uniform convexity and smoothness inequalities for C_p we obtain new and simple proofs of the known inequalities for L_p.

I Introduction

The concepts of uniform convexity and its dual property, uniform smoothness, play an important role in analysis. After reviewing these concepts in the L_p function spaces, we shall consider their extension to the Schatten trace ideals, C_p, i.e., the setting in which functions are replaced by operators, and integrals are replaced by traces. The emphasis throughout will be on the optimal constants appearing in the various inequalities. These optimal constants are "natural", as will be explained later in the introduction: they are the constants one would obtain from an informed guess using elementary calculus. However, as is often the case in such matters, no ready-made arguments suffice to validate the informed guesses.

A normed space X is said to be **uniformly convex** if, for each $\varepsilon > 0$, there is a $\delta > 0$ such that if x and y are *unit* vectors in X with $\|x - y\| \geqq 2\varepsilon$, then the average $(x + y)/2$ has norm at most $1 - \delta$. A normed space X is said to be **uniformly smooth** if, for all $\varepsilon > 0$, there is a $\tau > 0$ such that if x and y are *unit* vectors in X with $\|x - y\| \leqq 2\tau$, then the average $(x + y)/2$ has norm at least $1 - \varepsilon\tau$.

* Work partially supported by US National Science Foundation grant DMS 88-07243
** Work partially supported by US National Science Foundation grant DMS 92-07703
*** Work partially supported by US National Science Foundation grant PHY90-19433 A02

Figuratively speaking, the unit ball of a uniformly convex space is uniformly free of "flat spots", and the unit ball of a uniformly smooth space is uniformly free of "corners". Since the unit ball of X^*, the dual of X, is the polar conjugate of the unit ball of X, it is not difficult to show that X is uniformly convex (and hence reflexive) if and only if X^* is uniformly smooth [D].

Many applications of uniform convexity and smoothness require quantitative versions of these notions. The function δ_X given by

$$\delta_X(\varepsilon) := \inf\{1 - \tfrac{1}{2}\|x + y\| : \|x\| = \|y\| = 1,\ \|x - y\| \geqq 2\varepsilon\} \tag{1.1}$$

is called the **modulus of convexity** of X. (N.B. The function δ_X is frequently defined with ε in place of 2ε. The definition used here simplifies several of the formulae involving δ_X and fits more naturally with the definition of the modulus of smoothness given below.) Clearly, X is uniformly convex if and only if δ_X is strictly positive for every $\varepsilon > 0$.

It might seem natural to define the modulus of smoothness by setting it equal, at τ, to

$$\sup\{1 - \tfrac{1}{2}\|x + y\| : \|x\| = \|y\| = 1,\ \|x - y\| \leqq 2\tau\} \tag{$*$}$$

Clearly, X is uniformly smooth if and only if this supremum is $o(\tau)$ at $\tau = 0$. The definition $(*)$, however, would not be well adapted to the duality between uniform convexity and uniform smoothness. Instead, the function ρ_X given by

$$\rho_X(\tau) := \sup\left\{\frac{\|u + v\| + \|u - v\|}{2} - 1 : \|u\| = 1,\ \|v\| = \tau\right\} \tag{1.2}$$

is called the **modulus of smoothness** of X. This definition arises from $(*)$ if we rewrite the quantity to be maximized there in terms of $u = (x + y)/2$ and $v = (x - y)/2$, and change the constraint from $\|u + v\| = \|u - v\| = 1$ to simply $\|u\| = 1$. For small τ, there is no substantial difference, and it is easy to show (see [Kö]) that X is uniformly smooth if and only if $\lim_{\tau \to 0} \rho_X(\tau)/\tau = 0$.

Lindenstrauss [L] has shown that with these definitions, the modulus of convexity of a normed space X and the modulus of smoothness of its dual X^* are related by

$$\rho_{X^*}(\tau) = \sup\{\tau\varepsilon - \delta_X(\varepsilon) : 0 \leqq \varepsilon \leqq 1\}. \tag{1.3}$$

This is a quantitative versions of Day's duality theorem [D].

Uniform convexity was introduced by Clarkson [C] who proved that every L_p space with $1 < p < \infty$ is both uniformly convex and uniformly smooth. Clarkson proved inequalities which give bounds of the form

$$\delta_{L_p}(\varepsilon) \geqq (\varepsilon/K_{p,r})^r \tag{1.4}$$

where $r = p$ for $2 \leqq p < \infty$, and $r = p/(p-1)$ for $1 < p \leqq 2$.

Lower bounds of the form (1.4) for the modulus of convexity δ of a normed space are especially useful and, in many applications, particular importance attaches to the best possible value of r. Evidently, such an inequality cannot hold for any $r < 2$. A normed space X is said to be *r-uniformly convex* in case $\delta_X(\varepsilon) \geqq (\varepsilon/C)^r$ for some constant C. (After Eq. (2.6) below, we make an apparently more restrictive definition of r-uniform convexity. The two definitions will be shown to be consistent in Proposition 7, and the present definition is the simplest to use in the introduction.) Clarkson's bounds (1.4) only show that L_p is r-uniformly convex with $r > 2$ for all $p \neq 2$ while, actually, L_p is 2-uniformly convex for $1 < p \leqq 2$.

The 2-uniform convexity of L_p for $1 < p \leqq 2$ follows from a result of Hanner [H], who proved an inequality from which δ_{L_p} can be easily *computed.* Hanner's result is recalled in part (a) of Theorem 2 below. The best constant $K_{p,2}$ in (1.4) seems to have been first determined by Ball and Pisier [BP], who gave a simple direct proof, independent of Hanner's calculation, that L_p is 2-uniformly convex for such p. Their optimal 2-uniform convexity inequality is:

$$\delta_{L_p}(\varepsilon) \geqq \frac{p-1}{2}\varepsilon^2 \quad \text{for} \quad 1 < p \leqq 2. \tag{1.5}$$

Because of the dual nature of the notions of uniform convexity and smoothness, the modulus of smoothness of L_p for $2 \leqq p < \infty$ satisfies an inequality of the form $\rho_{L_p}(\tau) \leqq (\hat{K}_{p,2}\tau)^2$. Again, this is a better estimate than that which follows from Clarkson's inequalities. A more detailed history of these and related inequalities will be presented in Sect. II of our paper.

Less is known about the corresponding inequalities for the trace classes C_p. Clarkson's inequalities were extended to C_p partly by Dixmier [Di], and fully by Klaus [Si], with precisely the same constants and exponents as in the L_p case.

Tomczak-Jaegermann later showed that, as with L_p, C_p is actually 2-uniformly convex for $1 < p \leqq 2$:

$$\delta_{C_p}(\varepsilon) \geqq (\varepsilon/\tilde{K}_{p,2})^2 \quad \text{for} \quad 1 < p \leqq 2. \tag{1.6}$$

Her proof proceeds by establishing the C_p analog of Hanner's inequality when p is an even integer, then deducing the 2-uniform smoothness of C_p for all $p \geqq 2$ from this by interpolation, and then using Lindenstrauss's duality result to obtain the 2-uniform convexity of C_p for $1 < p \leqq 2$. Implicit in her proof is the fact that when $p = 2k/(2k-1)$ for some positive integer k, the sharp constants $\tilde{K}_{p,2}$ for C_p coincide with those of L_p; i.e. $\tilde{K}_{p,2} = (p-1)/2$ for such values of p.

The principal results in our paper are the determination of the best possible constants for all p in Tomczak-Jaegermann's theorem, and the proof that the C_p analog of Hanner's inequality holds for $1 < p \leqq 4/3$, and in the dual range $4 \leqq p < \infty$. Our two main theorems are the following (in which $\|\cdot\|_p$ denotes the L_p or the C_p norm):

Theorem 1 (Optimal 2-uniform convexity) *For* $1 \leqq p \leqq 2$, *the inequality*

$$\left(\frac{\|X+Y\|_p^p + \|X-Y\|_p^p}{2}\right)^{2/p} \geqq \|X\|_p^2 + (p-1)\|Y\|_p^2. \tag{1.7}$$

holds in the following cases:

(a) *X and Y are functions in L_p.*
(b) *X and Y are matrices in C_p.*

If $2 \leqq p \leqq \infty$, *the inequality is reversed.*

The title of this theorem will be explained more fully in Sect. II; the point, of course, is that validity of the inequality (1.7) implies 2-uniform convexity. Part (a) is an unpublished result of Ball and Pisier, and the cases of part (b) for $p = 2k/(2k-1)$ are, as we have said before, implicit in the paper [TJ] of Tomczak-Jaegermann. The rest is new. The constant $p-1$ in (1.7) is clearly seen to be optimal as well as natural from the point of view of elementary calculus: if X and Y are real numbers with $|Y|$ much smaller than $|X|$, then the two sides of (1.7) agree to second order in Y. The fact that (1.7) is true for *all* pairs of *numbers* is Gross's two-point inequality [G].

Theorem 2 (Extension of Hanner's inequality to C_p) *For* $1 \leqq p \leqq 2$, *the inequality*

$$\|X + Y\|_p^p + \|X - Y\|_p^p \geqq (\|X\|_p + \|Y\|_p)^p + |\|X\|_p - \|Y\|_p|^p \tag{1.8}$$

holds in the following cases:

(a) *X and Y are functions in* L_p.
(b) $p \leqq \frac{4}{3}$ *and X and Y are matrices in* C_p.
(c) *X and Y are matrices in* C_p *such that both* $X + Y$ *and* $X - Y$ *are positive semidefinite.*

For $2 \leqq p < \infty$, *the inequality is reversed and the restriction in* (b) *becomes* $p \geqq 4$, *and the restriction in* (c) *changes to the restriction that X and Y are positive semidefinite.*

Part (a) is Hanner's inequality, and the cases of part (b) in which $p = 2k$ are due to Tomczak-Jaegermann [TJ]. The rest is new. As we explain in the first proof of Proposition 3 below, the inequality (1.8), whenever it holds, implies the inequality (1.7). Thus, if the conditions under which we establish (1.8) were not more restrictive than those under which we establish (1.7), Theorem 1 would be a corollary of Theorem 2.

The paper is organized as follows: In Sect. II, we review the large number of inequalities bearing on the uniform convexity of L_p spaces. Thus, Sect. II consists largely of known results which are presented because of the light they shed on the problems solved in this paper regarding the uniform convexity of C_p, and those that remain open. To our knowledge, such a systematic compendium of these inequalities has not appeared before, and we hope it will be found useful. There are however some new results and some new, simpler proofs. Finally in Sect. III we prove Theorem 1, and in Sect. IV we prove Theorem 2.

Theorem 1 has been applied by Carlen and Lieb [CL] to prove a conjecture of Gross, which arose in his work on quantum field theory. Other applications of the kinds of the inequalities that we discuss here are given, for example, in Pisier's book [P].

Although all of our theorems are stated and proved in the language of matrices, the proofs go through without any change in the context of linear operators on a Hilbert space. By the results of Ruskai [Ru], they can even be extended to a natural Von Neumann algebra context.

II Uniform convexity and smoothness in L_p

While this section is largely focused on inequalities relating to L^p spaces, we state certain definitions and prove certain results in the general normed space setting so that they are available to us in the next section.

For the rest of the paper, *q denotes the dual index of p*, i.e., $1/p + 1/q = 1$.

The notion of uniform convexity was introduced by Clarkson who proved four inequalities. The two that imply the uniform convexity of L_p spaces are the following, in which x and y are functions in L_p:

$$\left(\left\|\frac{x+y}{2}\right\|_p^p + \left\|\frac{x-y}{2}\right\|_p^p\right)^{1/p} \leqq \left(\frac{\|x\|_p^p + \|y\|_p^p}{2}\right)^{1/p} \quad \text{for} \quad 2 \leqq p \leqq \infty \tag{2.1}$$

and

$$\left(\left\|\frac{x+y}{2}\right\|_p^q + \left\|\frac{x-y}{2}\right\|_p^q\right)^{1/q} \leqq \left(\frac{\|x\|_p^p + \|y\|_p^p}{2}\right)^{1/p} \quad \text{for} \quad 1 \leqq p \leqq 2. \tag{2.2}$$

In the cases $1 \leqq p \leqq 2$ and $2 \leqq p \leqq \infty$, the inequalities in (2.1) and (2.2), respectively, hold in the reversed sense. These are the other two Clarkson inequalities – the ones which imply that the L_p spaces are uniformly smooth. They follow from (2.1) and (2.2) by an elementary duality argument.

The inequality (2.1), which involves *only* p powers and not q powers as well, is simpler to prove, and is known as the "easy" Clarkson inequality. In fact, (2.1) is not only easier to prove, it is actually a consequence of (2.2). This is so because both inequalities can be viewed as statements about the norms of certain linear operators. Viewed as such, (2.1) is weaker than the dual inequality to (2.2). More concretely, for $1 \leqq s,\ t \leqq \infty$, equip $L_s \times L_s$ with the norm $\|\cdot\|_{s,t}$ given by $\|(x, y)\|_{s,t} = ((\|x\|_s^t + \|y\|_s^t)/2)^{1/t}$. Also, define the operator $B: L_{s,t} \to L_{s,t}$ by $B(x, y) = ((x + y)/2, (x - y)/2)$. Then (2.2) is equivalent to the statement that B is a bounded operator from $L_{p,p}$ to $L_{p,q}$ for $1 \leqq p \leqq 2$ with norm $2^{1/p}$. But since B is self-adjoint, it has the same norm as an operator between the dual spaces; i.e. B is bounded from $L_{q,p}$ to $L_{q,q}$ with norm $2^{1/p}$ for $1 \leqq p \leqq 2$. Finally, since $p \leqq q$, $\|(x, y)\|_{q,p'} \leqq \|(x, y)\|_{q,q}$, and hence B has norm $2^{1/p}$ from $L_{q,q}$ to $L_{q,q}$ for $2 \leqq q \leqq \infty$, which is clearly equivalent to (2.1).

Since these bounds on the norm of B (which are equivalent to (2.1) and (2.2)) are log-linear in $1/p$, they can be proved by interpolation between the elementary cases $p = 1$ (Minkowski's inequality for L_1), $p = 2$ (the parallelogram law) and $p = \infty$ (Minkowski's inequality for L_∞), as observed by Boas [Bo]. This same approach was later used by Klaus [Si] to establish the C_p analogs of Clarkson's inequalities.

It is convenient also to have the following inequality, obtained from (2.2) by rearranging some powers of 2. If x and y belong to L_s where either $s = p$ or $s = q$, then

$$\left(\frac{\|x+y\|_s^q + \|x-y\|_s^q}{2}\right)^{1/q} \leqq (\|x\|_s^p + \|y\|_s^p)^{1/p} \quad \text{for} \quad 1 \leqq p \leqq 2. \tag{2.3}$$

Replacing x and y respectively with $x + y$ and $x - y$, and rearranging some powers of 2, we see that the inequality reverses when p and q are interchanged. That is,

$$\left(\frac{\|x+y\|_s^p + \|x-y\|_s^p}{2}\right)^{1/p} \geqq (\|x\|_s^q + \|y\|_s^q)^{1/q} \quad \text{for} \quad 1 \leqq p \leqq 2. \tag{2.4}$$

It follows directly from (2.3) that if x and y are unit vectors in one of the spaces L_p or L_q, and $\|x - y\| = 2\varepsilon$ then

$$\left\|\frac{x+y}{2}\right\| \leqq (1 - \varepsilon^q)^{1/q} \leqq 1 - \frac{\varepsilon^q}{q},$$

so that $\delta_{L_s}(\varepsilon) \geqq \varepsilon^q/q$. Similarly, from (2.4) one sees that $\rho_{L_s}(\tau) \leqq \tau^p/p$.

We next turn to the sort of inequalities that figure in our Theorem 1. We begin with some definitions and general considerations intended to clarify the relations among all the inequalities that we consider.

Let X be a uniformly convex normed space, and suppose that $\delta_X(\varepsilon) \geqq (\varepsilon/C)^r$ for some C and $r > 1$. Then with $\|x\| = \|y\| = 1$ and $\|x - y\| = 2\varepsilon$, we have that

$$\left\|\frac{x+y}{2}\right\|^r \leqq (1 - (\varepsilon/C)^r)^r \leqq 1 - \left(\frac{r-1}{r}\right)^{r-1} (\varepsilon/C)^r,$$

and thus, with $K = C(1/t)^{-1/t}$, $1/t + 1/r = 1$,

$$\left\|\frac{x+y}{2}\right\|^r + \left\|\frac{x-y}{2K}\right\|^r \leqq \frac{\|x\|^r + \|y\|^r}{2} \tag{2.5}$$

for all x and y such that $\|x\| = \|y\|$. By replacing x with $x + y$ and y with $x - y$, we find that

$$\frac{\|x+y\|^r + \|x-y\|^r}{2} \geqq \|x\|^r + \|K^{-1}y\|^r \tag{2.6}$$

for all x and y such that $\|x + y\| = \|x - y\|$.

As promised in the introduction (after (1.4)), we now impose a definition of r-uniform convexity that may seem more restrictive than the one we gave before. Proposition 7 below shows that the two definitions are equivalent, *up to constants*. It is the constant in (2.6), figuring in the second definition, that is the main object of our attention.

A normed space X is said to be ***r*-uniformly convex** for some $r \in [2, \infty)$ if there is a constant K such that (2.6) holds for *all* $x, y \in X$. The best constant K is called the ***r*-uniform convexity constant** of X. When X is r-uniformly convex, so that (2.6) and hence (2.5) hold, it is immediate from the latter that $\delta_X(\varepsilon) \geqq (\varepsilon/K)^r$. Thus r-uniform convexity implies the validity of a lower bound of the form $\delta_X(\varepsilon) \geqq (\varepsilon/C)^r$ for the modulus of convexity; i.e. the condition under which we called X r-uniformly convex in the introduction.

Similarly, X is said to be ***t*-uniformly smooth** for some $t \in (1, 2]$ if

$$\frac{\|x+y\|^t + \|x-y\|^t}{2} \leqq \|x\|^t + \|Ky\|^t, \tag{2.7}$$

for some K and *all* $x, y \in X$. The best constant K is called the ***t*-uniform smoothness constant** of X. We shall show at the end of this section that the t-uniform smoothness constant of a normed space X equals the r-uniform convexity constant of its dual X^* where, as usual, $1/r + 1/t = 1$.

When (2.7) holds, we have that for all x and y with $\|x\| = 1$ and $\|y\| = \tau$

$$\frac{\|x+y\| + \|x-y\|}{2} - 1 \leqq \left(\frac{\|x+y\|^t + \|x-y\|^t}{2}\right)^{1/t} - 1 \leqq (K(1/t)^{1/t}\tau)^t.$$

Hence, by (1.1), t-uniform smoothness implies an estimate of the form $\rho_X(\tau) \leqq (C\tau)^t$. Proposition 7 shows that the reverse implication holds as well.

The parallelogram identity shows that Hilbert space is 2-uniformly convex and 2-uniformly smooth, and it is readily seen that the exponent 2 is the best that can occur for each property. Clarkson's inequality shows that when $1 \leqq p \leqq 2$ then each of L_p and L_q is q-uniformly convex and p-uniformly smooth. As we have remarked, these *exponents* are not in general the best possible, despite the fact that the *constants* in Clarkson's inequalities are always sharp.

The actual situation is the following:

For $1 < p \leqq 2$, L_p *is 2-uniformly convex though no better than p-uniformly smooth while for* $2 \leqq q < \infty$, L_q *is 2-uniformly smooth as well as q-uniformly convex.*

These facts follow from Hanner's inequality (Theorem 2(a) of the introduction) which determines exactly the moduli of convexity and smoothness of all L_p spaces. The optimal 2-uniform convexity inequality is the following:

Proposition 3 (Optimal 2-uniform convexity for L_p) *If* $1 \leqq p \leqq 2$ *and* x *and* $y \in L_p$, *then*

$$\frac{\|x+y\|_p^2 + \|x-y\|_p^2}{2} \geqq \|x\|_p^2 + (p-1)\|y\|_p^2. \tag{2.8}$$

For $2 \leqq p < \infty$, *the inequality is reversed.*

Remark. The inequality (2.8) holds for any normed space for which (1.8) holds, as we will soon show. Inequality (2.8) does not seem to appear in the literature in quite this form but it is probably folk-lore. Ball and Pisier noticed that it follows from Gross's two-point inequality using arguments which (in the context of general Banach lattices) go back to Figiel [F].

The reader will note that (2.8) is not identical to (1.7), but we shall soon see that the validity of (2.8) for *all* L_p spaces implies the validity of the apparently stronger (1.7).

First proof. To deduce (2.8) from Hanner's inequality recall that a special case of Gross's inequality [G] states that if $1 \leqq p \leqq 2$ and a and b are real,

$$\left(\frac{|a+b|^p + |a-b|^p}{2}\right)^{1/p} \geqq (a^2 + (p-1)b^2)^{1/2}.$$

Now if $x, y \in L_p$,

$$\begin{aligned}\left(\frac{\|x+y\|_p^2 + \|x-y\|_p^2}{2}\right)^{1/2} &\geqq \left(\frac{\|x+y\|_p^p + \|x-y\|_q^p}{2}\right)^{1/p} \\ &\geqq \left[\frac{(\|x\|_p + \|y\|_p)^p + |\|x\|_p - \|y\|_p|^p}{2}\right]^{1/p} \\ &\geqq (\|x\|_p^2 + (p-1)\|y\|_p^2)^{1/2},\end{aligned}$$

where we have used, in succession, Hölder's inequality, Hanner's inequality (1.8), and Gross's inequality with $a = \|x\|_p$ and $b = \|y\|_p$. □

Second proof. An alternative proof of Proposition 3 consists simply of showing that for $\|y\|$ small,

$$\frac{\|x+y\|_p^2+\|x-y\|_p^2}{2} \geqq \|x\|_p^2+(p-1)\|y\|_p^2+o(\|y\|_p^2), \tag{2.9}$$

and then observing that this infinitesimal form of (2.8) is equivalent to the full statement. That is, (2.9) is the same as

$$\frac{d^2}{ds^2}\left(\frac{\|x+sy\|_p^p+\|x-sy\|_p^p}{2}\right)\Bigg|_{s=0} \geqq p(p-1)\|x\|_p^{p-2}\|y\|_p^2, \tag{2.10}$$

which is easy to establish for L_p functions by elementary calculus. Proposition 3 follows from this by integration with respect to $s\in[0,1]$. □

As will be shown later, inequality (2.10) is also true for matrices in C_p, and this will form the basis of the proof of Theorem 1 given in Sect. III. But there is an important difference between the commutative and non-commutative cases. For *functions* x and y in L_p,

$$\frac{d^2}{ds^2}\left(\frac{\|x+sy\|_p^p+\|x-sy\|_p^p}{2}\right)\Bigg|_{s=0} = p(p-1)\int|x|^{p-2}|y|^2, \tag{2.11}$$

and the latter dominates $p(p-1)\|x\|_p^{p-2}\|y\|_p^2$ by Hölder's inequality (since $p<2$). For *matrices*, the analogue of (2.11) is false: one always has

$$\frac{d^2}{ds^2}\left(\frac{\|X+sY\|_p^p+\|X-sY\|_p^p}{2}\right)\Bigg|_{s=0} \leqq p(p-1)\,Tr|X|^{p-2}|Y|^2$$

and equality need not hold. The problem is to find a replacement for (2.11) in the non-commutative setting.

Inequality (2.8), as we said, is apparently weaker than (1.7) since

$$\left(\frac{\|x+y\|^p+\|x-y\|^p}{2}\right)^{1/p} \leqq \left(\frac{\|x+y\|^2+\|x-y\|^2}{2}\right)^{1/2}$$

for $1\leqq p\leqq 2$. However, a simple doubling argument shows that (1.7) is actually a consequence of the fact that (2.8) holds for all L_p spaces. This argument is also valid for C_p, and we give it in detail in Sect. III, Eq. (3.5)–(3.6). The reader can easily translate the C_p version into the L_p version.

Note also that in the first proof of Proposition 3, which was based on Hanner's inequality, we actually arrived at (1.7) in an intermediate step. In the C_p setting, we do not possess a full analog of Hanner's inequality, and so we shall prove Theorem 1 by adapting the second proof of Proposition 3 to the C_p setting.

The following diagram shows the relationships between the several expressions mentioned above. Connecting lines indicate inequality between the expressions. All of the indicated inequalities hold in both C_p and L_p except for that indicated by the line labeled Hanner, which is only known to hold in C_p in the special cases specified in Theorem 2. In each expression x and y are elements of L_p or of C_p, and q is the index conjugate to p. For $1\leqq p\leqq 2$, the quantities *increase* as one goes up the page; for $2\leqq p\leqq\infty$, they *decrease*.

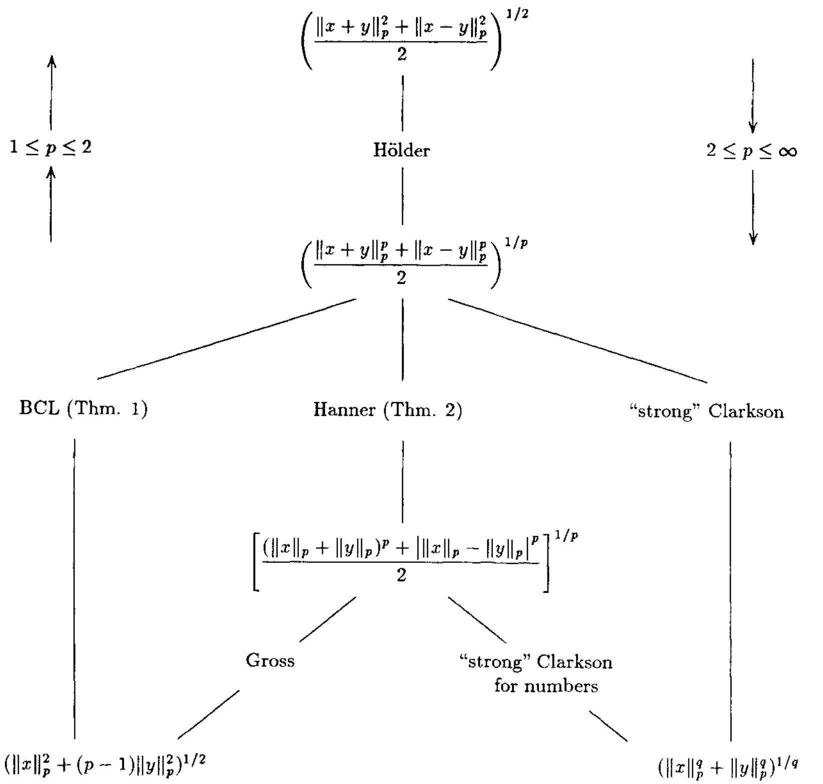

Fig. 1. Relationships among the inequalities

We turn now to a proof of Hanner's inequality for L_p, which yields Clarkson's inequalities along the way, and to a simple duality result which shows that optimal constants obtained for q-uniform convexity of a normed space immediately yield optimal constants for p-uniform smoothness of its dual (and conversely).

Lemma 4 (Variational characterization of sums of p^{th} powers) *For* $1 < p < \infty$ *define* $\alpha = \alpha_p$: $[0, \infty) \to [0, \infty)$ *by*

$$\alpha(r) = (1+r)^{p-1} + |1-r|^{p-1} \operatorname{sign}(1-r).$$

Then for all $x, y \in \mathbb{R}$

$$|x+y|^p + |x-y|^p = \left\{\begin{pmatrix}\sup\\ \inf\end{pmatrix}\right\} \{\alpha(r)|x|^p + \alpha(1/r)|y|^p \colon 0 < r < \infty\}$$

the sup *or* inf *being taken according as* $p \leqq 2$ *or* $p \geqq 2$.

Proof. Assume $1 < p \leqq 2$; the proof for $p \geqq 2$ is similar. Plainly, it may be assumed that $0 < y \leqq 1 = x$. For $r = y$; one readily checks that

$$\alpha(r) + \alpha\left(\frac{1}{r}\right)y^p = (1+y)^p + (1-y)^p.$$

To see that $(1+y)^p + (1-y)^p \geqq \alpha(r) + \alpha(1/r)y^p$ for all r, it suffices to check that the latter quantity attains its maximum when $r = y$. But

$$\begin{aligned}\frac{d}{dr}(\alpha(r) + \alpha(1/r)y^p) &= \alpha'(r) - \frac{1}{r^2}\alpha'(1/r)y^p \\ &= (p-1)\left[(1+r)^{p-2} - |1-r|^{p-2} - \frac{y^p}{r^2}\left(\left(\frac{1}{r}+1\right)^{p-2} - \left|\frac{1}{r}-1\right|^{p-2}\right)\right] \\ &= (p-1)\left(1-\left(\frac{y}{r}\right)^p\right)[(1+r)^{p-2} - |1-r|^{p-2}].\end{aligned}$$

Since $p - 2 \leqq 0$ and $1 + r \geqq |1 - r|$, the last factor is non-positive. Thus, the whole is non-negative for $0 < r < y$ and non-positive for $r > y$. □

Proof of Theorem 2(a) (*Hanner's inequality*) Again assume $1 < p \leqq 2$ and let $x, y \in L_p$. Then

$$\begin{aligned}\|x+y\|^p + \|x-y\|^p &= \int(|x+y|^p + |x-y|^p) = \int \sup_r \{\alpha(r)|x|^p + \alpha(1/r)|y|^p\} \\ &\geqslant \sup_r \int(\alpha(r)|x|^p + \alpha(1/r)|y|^p) = \sup_r \{\alpha(r)\|x\|^p + \alpha(1/r)\|y\|^p\} \\ &= (\|x\| + \|y\|)^p + |\|x\| - \|y\||^p.\end{aligned}$$ □

Proof of (2.2) Let us also show how Lemma 4 can be used to deduce the "hard" Clarkson inequality. This time we shall prove it in the uniform smoothness range, i.e. for L_q with $2 \leqq p < \infty$. Since

$$\|x+y\|_p^p + \|x-y\|_p^p \leqq \alpha(r)\|x\|_p^p + \alpha(1/r)\|y\|_p^p$$

for all r, it is enough to find an r for which the right side equals $2(\|x\|^q + \|y\|^q)^{p/q}$.

Set $u = \|x\|_p^q$, $v = \|y\|_p^q$, $r = v/u$, and assume that $v \leqq u$. Then

$$\begin{aligned}\alpha(r)\|x\|_p^p + \alpha(1/r)\|y\|_p^p &= \alpha(v/u)u^{p-1} + \alpha(u/v)v^{p-1} = 2(u+v)^{p-1} \\ &= 2(\|x\|_p^p + \|y\|_p^p)^{p/q}.\end{aligned}$$ □

We now prove the duality results mentioned earlier. These results hold in general; no reference to L_p or C_p is made.

Lemma 5 (Duality for q-uniform convexity and p-uniform smoothness) *Let X be a normed space with dual X^*. The p-uniform smoothness constant of X (the constant K in* (2.7)) *is equal to the q-uniform convexity constant of X^* (the constant K in* (2.6).

Proof. Suppose that the q-uniform convexity of X^* is K and let $x, y \in X$. We denote norms in X and X^* indiscriminately by $\|\cdot\|$ and trust that the meaning will be clear. There are unit vectors λ and μ in X^* such that $\lambda(x+y) = \|x+y\|$

and $\mu(x-y)=\|x-y\|$. Define $\phi,\psi\in X^*$ by $\phi=Z^{-1/q}\|x+y\|^{p-1}\lambda$ and $\psi=Z^{-1/q}\|x-y\|^{p-1}\mu$ with $Z=(\|x+y\|^p+\|x-y\|^p)/2$. Then, $\|\phi\|^q+\|\psi\|^q=2$ and we have

$$\left(\frac{\|x+y\|^p+\|x-y\|^p}{2}\right)^{1/p}=\frac{\phi(x+y)+\psi(x-y)}{2}=\left(\frac{\phi+\psi}{2}\right)(x)+\left(\frac{\phi-\psi}{2}\right)(y)$$

$$\leqq\left(\left\|\frac{\phi+\psi}{2}\right\|^q+\left\|\frac{\phi-\psi}{2K}\right\|^q\right)^{1/q}(\|x\|^p+\|Ky\|^p)^{1/p}$$

$$\leqq\left(\frac{\|\phi\|^q+\|\psi\|^q}{2}\right)^{1/q}(\|x\|^p+\|Ky\|^p)^{1/p}$$

$$=(\|x\|^p+\|Ky\|^p)^{1/p}.$$

The first inequality is Hölder's inequality for numbers, and the second is (2.5) with $r=q$. The other implication is similar. □

Lemma 6 (Duality for Hanner's inequality) *Let X be a normed space with dual X^*. Let $1<p\leqq 2$ and $1/p+1/q=1$. Then the validity of*

$$\|\phi+\psi\|^q+\|\phi-\psi\|^q\leqq(\|\phi\|+\|\psi\|)^q+|\,\|\phi\|-\|\psi\|\,|^q \tag{2.12}$$

for all $\phi,\psi\in X^*$ implies the validity of

$$\|y+z\|^p+\|y-z\|^p\geqq(\|y\|+\|z\|)^p+|\,\|y\|-\|z\|\,|^p \tag{2.13}$$

for all $y,z\in X$. Similarly, the validity of (2.13) in X implies the validity of (2.12) in X^*.

Proof. Suppose first that (2.12) holds in X^*. To establish (2.13), we first rewrite it in terms of $u=y+z$ and $v=y-z$, so that what we must show is:

$$2^p(\|u\|^p+\|v\|^p)\geqq(\|u+v\|+\|u-v\|)^p+|\,\|u+v\|-\|u-v\|\,|^p. \tag{2.14}$$

We may assume without loss of generality that $\|u+v\|=1$, and that $r:=\|u-v\|\leqq 1$. Then the right side of (2.14), which we call R^p, can be rewritten (as in Lemma 4) as

$$R^p=\alpha\|u+v\|^p+\beta\|u-v\|^p=(1+r)^p+(1-r)^p,$$

where $\alpha=(1+r)^{p-1}+(1-r)^{p-1}$ and $\beta=r^{1-p}[(1+r)^{p-1}-(1-r)^{p-1}]$. As in the proof of Lemma 5, we choose unit vectors λ and μ in X^* such that $\lambda(u+v)=\|u+v\|$ and $\mu(u-v)=\|u-v\|$. Then we define

$$\phi=\alpha R^{-p/q}\|u+v\|^{p-1}\lambda\quad\text{and}\quad\psi=\beta R^{-p/q}\|u-v\|^{p-1}\mu.$$

Thus,

$$R=\phi(u+v)+\psi(u-v)=(\phi+\psi)(u)+(\phi-\psi)(v)$$

$$\leqq\|\phi+\psi\|\,\|u\|+\|\phi+\psi\|\,\|v\|\leqq(\|\phi+\psi\|^q+\|\phi-\psi\|^q)^{1/q}(\|u\|^p+\|v\|^p)^{1/p}.$$

To complete the demonstration of (2.14), we have to show that $T:=\|\phi+\psi\|^q+\|\phi-\psi\|^q\leqq 2^q$. By (2.12),

$$T\leqq R^{-p}[\alpha\|u+v\|^{p-1}+\beta\|u-v\|^{p-1}]^q+R^{-p}[\alpha\|u+v\|^{p-1}-\beta\|u-v\|^{p-1}]^q$$

$$=R^{-p}[\alpha+\beta r^{p-1}]^q+R^{-p}[\alpha-\beta r^{p-1}]^q=2^q.$$

A similar proof works in the other direction to go from (2.13) to (2.12) □

Remark. In our application of Lemma 6 to the proof of Theorem 2, we need the following refinement whose truth is evident from the proof given above. We take $X = C_p$ and $X^* = C_q$. Then the validity of (2.13) with the *extra constraint* that $y + z$ and $y - z$ are *positive semidefinite* matrices implies the validity of (2.12) when ϕ and ψ are *positive semidefinite* matrices.

We close this section with a proposition showing the consistency of the two definitions that we have given for r-uniform convexity; i.e., the ones following (1.4) and (2.6).

Proposition 7 (Equivalence of definitions of r-uniform convexity) *Let X be a normed space. Then* (2.5) *holds for some constant K and all $x, y \in X$ if and only if $\delta_X(\varepsilon) \geqq (\varepsilon/C)^r$ for some constant C. Similarly,* (2.7) *holds for some constant K and all $x, y \in X$ if and only if $\rho_X(\tau) \leqq (C\tau)^r$ for some constant C.*

Proof. We have already seen that (2.5) and (2.7) imply the indicated bounds on δ_X and ρ_X respectively.

Suppose first that $\rho_X(\tau) \leqq (C\tau)^r$ for some constant C. Of course, $1 < r \leqq 2$. Then for all $\|x\| = 1$ and $\|y\| \leqq 1$,

$$\frac{\|x + y\| + \|x - y\|}{2} - 1 \leqq (C\|y\|)^r.$$

Define numbers b and β by

$$b := \frac{\|x + y\| + \|x - y\|}{2} \quad \text{and} \quad \beta := \frac{\|x + y\| - \|x - y\|}{\|x + y\| + \|x - y\|}.$$

Then

$$\left(\frac{\|x + y\|^r + \|x - y\|^r}{2}\right)^{1/r} - \frac{\|x + y\| + \|x - y\|}{2} = b\left[\left(\frac{(1 + \beta)^r + (1 - \beta)^r}{2}\right)^{1/r} - 1\right]. \tag{2.15}$$

The function of β on the right side in (2.15) vanishes quadratically at the origin. Thus, a simple estimation using Taylor's theorem shows that it is no greater than $D_r\beta^2$ for some constant D_r depending only on r. Then, since $|\beta| \leqq \|y\|/b \leqq \|y\|$ and $1 \leqq r \leqq 2$, we have from (2.15) and the assumption on ρ_X that

$$\frac{\|x + y\|^r + \|x - y\|^r}{2} \leqq \left(1 + C^r\|y\|^r + D_r\|y\|^2\right)^r \leqq 1 + (K_r\|y\|)^r, \tag{2.16}$$

for all x and y with $\|y\| \leqq \|x\| = 1$, where K_r depends only on C and r. Therefore,

$$\frac{\|x + y\|^r + \|x - y\|^r}{2} \leqq \|x\|^r + K_r^r\|y\|^r \tag{2.17}$$

for all x and y with $\|y\| \leqq \|x\|$. Finally, since we may assume that $K_r \geqq 1$, (2.17) holds for all x and y.

Next, suppose that $\delta_X(\varepsilon) \geqq (\varepsilon/C)^r$ for some constant C. Then by (1.3),

$$\rho_{X^*}(\tau) = \sup\{\tau\varepsilon - \delta_X(\varepsilon) \colon 0 \leqq \varepsilon \leqq 1\} \leqq \sup\{\tau\varepsilon - (\varepsilon/C)^r \colon 0 \leqq \varepsilon \leqq \infty\} = (C\tau)^{r'}$$

where $1/r + 1/r' = 1$.Then, by what we have shown above, there is a constant K so that (2.7) is valid in X^*.But then by Lemma 5, (2.5) is valid in X with the same constant K. □

III Optimal 2-uniform convexity inequalities for trace norms

Norms on the space of $n \times n$ matrices which are non commutative analogs of the L_p norms can be defined in terms of the trace by

$$\| X \|_p = (\mathrm{Tr}((X^* X)^{p/2})^{1/p} = (\mathrm{Tr}((XX^*)^{p/2})^{1/p}, \tag{3.1}$$

for $1 \leqq p < \infty$. For $p = \infty$, $\| X \|_p$ denotes the operator norm of X, as usual. The analogy can be made quite close, and it has been developed in a von Neumann algebra context by Segal [Se] and Dixmier [Di] as part of their theories of non-commutative integration.

Many familiar inequalities for L_p norms also hold for the C_p norms. This is true, in particular, of the Hölder inequality

$$\| XY \|_r \leqq \| X \|_p \| Y \|_q, \; 1/r = 1/p + 1/q.$$

There are, however, other inequalities for L_p norms which *do not* hold for the C_p norms. Many examples are connected with the poor behavior of the map

$$X \mapsto |X| = (X^* X)^{1/2}. \tag{3.2}$$

For example, if f and g are complex valued functions in some L_p space, then $\| |f| - |g| \|_p \leqq \| f - g \|_p$. This is not true for C_p, and, when $p = 2$, the factor of $\sqrt{2}$ in the Araki-Yamagami inequality [ArY] $\| |X| - |Y| \|_2 \leqq \sqrt{2} \| X - Y \|_2$ is optimal whenever $n \geqq 2$.

As we have asserted in Theorems 1 and 2, however, almost all of the optimal inequalities expressing uniform convexity and smoothness properties of L_p spaces have exact analogs which hold for the C_p norms. Most of this section is devoted to the proof of Theorem 1. Before giving the proof we briefly discuss the history of uniform convexity inequalities for C_p as we know it.

The first such inequality was established by Dixmier [Di] who proved the C_p analog of the "easy" Clarkson inequality (2.1) by means of interpolation. As with L_p, this implies that for $2 \leqq p < \infty$, $\delta_{C_p}(\varepsilon) \geqq (1/p)\varepsilon^p$.

Interpolation was later used to establish the analog of the "hard" Clarkson inequality (2.2) which implies the uniform convexity of L_p for $1 < p \leqq 2$. Such a proof has been given by Martin Klaus, and is sketched in [Si]; to some extent it is modeled on Boas' proof [Bo] of (2.2) for L_p. This result implies that for $1 < p \leqq 2$, C_p is at least q-uniformly convex.

Later, Tomczak-Jaegermann showed that for $1 < p \leqq 2$, C_p is actually 2-uniformly convex and C_q is 2-uniformly smooth. Moreover, she showed for $q = 2k$, that the sharp 2-uniform smoothness constants of C_q are the same as those for L_q (so that the corresponding equalities of 2-uniform convexity constants hold by Lemma 5).

Before we prove Theorem 1, note that $t \mapsto t^{p/2}$ is concave for $1 < p \leqq 2$ and therefore (1.7) immediately implies that

$$\frac{\|X+Y\|_p^2 + \|X-Y\|_p^2}{2} \geqq \|X\|_p^2 + (p-1)\|Y\|_p^2, \tag{3.3}$$

which expresses the 2-uniform convexity of C_p in the usual way. It follows that

$$\delta_{C_p}(\varepsilon) \geqq \frac{(p-1)}{2}\varepsilon^2 \quad \text{for} \quad 1 < p \leqq 2, \tag{3.4}$$

and thus the analog of (1.5) holds for C_p.

We now observe that (1.7) is only formally stronger than (3.3). To see that (3.3) implies (1.7), consider the $2n \times 2n$ matrices given in block form by

$$Z = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix}, \quad W = \begin{bmatrix} Y & 0 \\ 0 & -Y \end{bmatrix}. \tag{3.5}$$

Then

$$\operatorname{Tr}|Z+W|^p = \operatorname{Tr}|Z-W|^p = (\operatorname{Tr}|X+Y|^p + \operatorname{Tr}|X-Y|^p)$$

and thus,

$$\|Z+W\|_p^2 = \|Z-W\|_p^2 = (\|X+Y\|_p^p + \|X-Y\|_p^p)^{2/p}.$$

Since also $\|Z\|_p^2 = 2^{2/p}\|X\|_p^2$ and $\|W\|_p^2 = 2^{2/p}\|Y\|_p^2$, (3.3) implies

$$\begin{aligned}\|X\|_p^2 + (p-1)\|Y\|_p^2 &= 2^{-2/p}(\|Z\|_p^2 + (p-1)\|W\|_p^2 \\ &\leqq 2^{-2/p}\left(\frac{\|Z+W\|_p^2 + \|Z-W\|_p^2}{2}\right) \\ &= \left(\frac{\|X+Y\|_p^p + \|X-Y\|_p^p}{2}\right)^{2/p},\end{aligned} \tag{3.6}$$

which is (1.7).

Proof of Theorem 1 First, we reduce to the case in which X and Y are self-adjoint. Consider the $2n \times 2n$ matrices given in block form by

$$C = \begin{bmatrix} 0 & X \\ X^* & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & Y \\ Y^* & 0 \end{bmatrix}.$$

Clearly, if (1.7) holds for the $2n \times 2n$ matrices C and D, it holds for X and Y. Since C and D are self-adjoint, it suffices to prove inequality (3.3) for such matrices. We therefore assume without loss of generality that X and Y are self-adjoint.

Let Z and W be defined in terms of X and Y as in (3.5). Then we can rewrite (1.7) as

$$\operatorname{Tr}(|Z+rW|^p)^{2/p} \geqq (\operatorname{Tr}|Z|^p)^{2/p} + r^2(p-1)(\operatorname{Tr}|W|^p)^{2/p}. \tag{3.7}$$

First, note that without loss of generality we may assume by continuity that the union of the ranges of Z and W span $\mathbb{C}^{2n}$. Then $\det(Z+rW)$ is a polynomial of order exactly $2n$ in r, and it has at most $2n$ zeros for $0 \leqq r \leqq 1$. We will avoid these values of r below in our computations. We define $\psi(r)$ by

$$\psi(r) = \operatorname{Tr}|Z+rW|^p = \operatorname{Tr}((Z^2 + r(ZW+WZ) + r^2W^2)^{p/2}).$$

Then $\psi(r)$ is continuously differentiable and

$$\frac{d}{dr}\psi(r) = (p/2)\mathrm{Tr}((Z^2 + r(ZW + WZ) + r^2 W^2)^{p/2-1})((ZW + WZ) + 2rW^2).$$

With the aid of the integral representation

$$(Z^2 + r(ZW + WZ) + r^2 W^2)^{(p/2-1)} = \beta_p \int_0^\infty t^{(p/2-1)} \frac{1}{t + (Z^2 + r(ZW + WZ) + r^2 W^2)} dt, \tag{3.8}$$

we see that $d\psi(r)/dr$ is again continuously differentiable.

Now, both sides of (3.7) agree at $r = 0$, and the first derivatives in r of both sides vanish there as well. Moreover, the second derivative in r of the left hand side of (3.7) satisfies

$$\frac{d^2}{dr^2}(\psi(r))^{2/p} \geqq \frac{2}{p}\psi(r)^{(2-p)/p}\frac{d^2}{dr^2}\psi(r),$$

while the second derivative on the right side of (3.7) is just $2(p-1)(\mathrm{Tr}|W|^p)^{2/p}$. It therefore suffices to show that

$$\frac{1}{p}\psi(r)^{(2-p)/p}\frac{d^2}{dr^2}\psi(r) \geqq (p-1)(\mathrm{Tr}|W|^p)^{2/p} \tag{3.9}$$

for all $0 < r < 1$. By redefining Z to be $Z + rW$, it suffices to establish (3.9) at $r = 0$. Since $Z + rW$ is non-singular, after the redefinition, $|Z|$ will be strictly positive.

We now claim that

$$\left.\frac{d^2}{dr^2}\mathrm{Tr}|Z + rW|^p\right|_{r=0} \geqq \left.\frac{d^2}{dr^2}\mathrm{Tr}||Z| + rW|^p\right|_{r=0}. \tag{3.10}$$

To see this, note that by the integral formula (3.8),

$$\left.\frac{d^2}{dr^2}\mathrm{Tr}|Z + rW|^p\right|_{r=0} = \mathrm{Tr}|Z|^{p-2}W^2 - (p/2)\beta_p \int_0^\infty t^{p/2-1}\mathrm{Tr} \times\left(\frac{1}{Z^2+t}(ZW + WZ)\frac{1}{Z^2+t}(ZW + WZ)\right)dt. \tag{3.11}$$

The trace under the integral sign consists of four terms which, using the cyclicity of the trace, can be rewritten as

$$\mathrm{Tr}\left(WZ\frac{1}{Z^2+t}WZ\frac{1}{Z^2+t}\right) + 3\mathrm{Tr}\left(W\frac{1}{Z^2+t}WZ^2\frac{1}{Z^2+t}\right).$$

Since only Z^2 enters the second of these two terms, this term is unchanged when Z is replaced by $|Z|$. Upon writing out the first term in a basis that diagonalizes Z, that term becomes

$$\sum_{i,j=1}^{n}\left(\frac{1}{z_i^2+t}\right)\left(\frac{1}{z_j^2+t}\right)|w_{ij}|^2 z_i z_j.$$

Clearly this term, and hence the integral in (3.11), increases when Z is replaced by $|Z|$. The first term in (3.11), being a function of Z^2, is invariant under the substitution, and the assertion (3.10) is established.

Therefore, without loss of generality we may assume that $Z > 0$. Then, of course, $Z + rW > 0$ for all r sufficiently small, and we no longer need to square $Z + rW$ to obtain a positive operator whose powers can be expressed as an integral over its resolvent. Working directly with $Z + rW$, we can use the simpler integral representation

$$(Z + rW)^{(p-1)} = \gamma_p \int_0^\infty t^{(p-1)} \left[\frac{1}{t} - \frac{1}{t + (Z + rW)} \right] dt \tag{3.12}$$

to conclude that

$$\psi''(0) = p\gamma_p \int_0^\infty t^{(p-1)} \mathrm{Tr} \left[\frac{1}{t + Z} W \frac{1}{t + Z} W \right] dt. \tag{3.13}$$

Consider the right side of (3.13) as a function of Z for fixed W. We claim that it is convex in Z. To prove this, it suffices to prove the following inequality for every self-adjoint matrix A:

$$\Delta(A) := \frac{d^2}{ds^2} \mathrm{Tr} \left(\frac{1}{t + (Z + sA)} W \frac{1}{t + (Z + sA)} W \right) \Bigg|_{s=0} \geqq 0.$$

There are six terms. If we define $C = (t + Z)^{-1/2} A (t + Z)^{-1/2}$ and $D = (t + Z)^{-1/2} W (t + Z)^{-1/2}$, then the result of the computation is

$$\Delta(A) = 4 \mathrm{Tr}\, C^2 D^2 + 2 \mathrm{Tr}\, CDCD.$$

But by the Schwarz inequality,

$$|\mathrm{Tr}(CDCD)| \leqq \{\mathrm{Tr}(CD^2C)\}^{1/2} \{\mathrm{Tr}(DC^2D)\}^{1/2} = \mathrm{Tr}\, C^2 D^2.$$

Thus, $\Delta(A) \geqq 0$ and the integrand in (3.13) is a convex function of Z.

Now fix W and t, and define

$$F(Z) = \mathrm{Tr} \left[\frac{1}{t + Z} W \frac{1}{t + Z} W \right].$$

Clearly, when U is any unitary matrix that commutes with W, $F(UZU^*) = F(Z)$. Let $\{e_1, \ldots, e_{2n}\}$ be an orthonormal basis of eigenvectors of W. Let $\{U_j\}_{1 \leqq j \leqq 2^{2n}}$ be some enumeration of the 2^{2n} unitary matrices with the property that $U_j e_k = \pm e_k$ for each k. Clearly each of these unitaries commutes with W. Thus, by the convexity of F which we have established in the last paragraph,

$$F(Z) = 2^{-2n} \sum_{j=1}^{2^{2n}} F(U_j Z U_j^*) \geqq F\left(2^{-2n} \sum_{j=1}^{2^{2n}} U_j Z U_j^* \right) = F(Z_{\mathrm{diag}}),$$

where Z_{diag} is the matrix whose diagonal entries, in the basis specified above, are those of Z, and whose off-diagonal entries are all zero. Replacing Z by Z_{diag} in (3.13), the integration can be carried out, and we obtain

$$\psi''(0) \geqq p(p-1) \left(\sum_{j=1}^{2n} z_j^{(p-2)} w_j^2 \right)$$

where z_j and w_j, respectively, denote the jth diagonal entries of Z and W in the W-basis specified above.

Now consider $\psi(0 = \mathrm{Tr}(Z^p)$ as a function of Z. It is clearly convex. Thus,by the averaging method just employed, we obtain

$$\psi(0) \geqq \left(\sum_{j=1}^{2n} z_j^p\right).$$

To establish (3.9), it only remains to check that

$$\left(\sum_{j=1}^{2n} z_j^p\right)^{(2-p)/p} \left(\sum_{j=1}^{2n} z_j^{(p-2)} w_j^2\right) \geqq \left(\sum_{j=1}^{2n} |w_j|^p\right)^2.$$

but this follows immediately from Hölder's inequality. □

IV Hanner's inequality for matrices

This section is devoted to the proof of parts (b) and (c) of Theorem 2. We begin with the proof of Theorem 2(c), and then show how that implies Theorem 2(b).

Proof of Theorem 2(c) First, let Y be a fixed self-adjoint $n \times n$ matrix, and consider the set M_Y of $n \times n$ self-adjoint matrices given by

$$M_Y := \{X\colon X + Y > 0 \quad \text{and} \quad X - Y > 0\}.$$

Clearly M_Y is convex, and if $X \in M_Y$, then $X > 0$.

We claim that

$$G(X) := \|X + Y\|_p^p + \|X - Y\|_p^p - 2\|X\|_p^p \tag{4.1}$$

is a convex function on M_Y.

By the averaging method employed in the proof of Theorem 1, this convexity would imply that

$$\|X + Y\|_p^p + \|X - Y\|_p^p - \alpha\|X\|_p^p \geqq \|X_{\mathrm{diag}} + Y\|_p^p + \|X_{\mathrm{diag}} - Y\|_p^p - \alpha\|X_{\mathrm{diag}}\|_p^p \tag{4.2}$$

for any $0 \leqq \alpha \leqq 2$, where X_{diag} denotes the diagonal part of X in a basis diagonalizing Y. (Note that if $X \in M_Y$, then $X_{\mathrm{diag}} \in M_Y$.) By Lemma 4 and Hanner's inequality in l_p.

$$\|X_{\mathrm{diag}} + Y\|_p^p + \|X_{\mathrm{diag}} - Y\|_p^p - \alpha(r)\|X_{\mathrm{diag}}\|_p^p \geqq \alpha(1/r)\|Y\|_p^p,$$

for all r, where $\alpha(r)$ is the function defined in Lemma 4. (Here we are making use of the easily checked fact that for $1 \leqq p \leqq 2$, $\alpha(r)$ and $\alpha(1/r)$ never exceed 2.) Combining this with (4.2), we would obtain

$$\|X + Y\|_p^p + \|X - Y\|_p^p \geqq \alpha(r)\|X\|_p^p + \alpha(1/r)\|Y\|_p^p.$$

Then, by another application of Lemma 4, the inequality (1.8) would be established for $1 \leqq p \leqq 2$ for all matrices X and Y such that $X + Y$ and $X - Y$ are positive semidefinite. By Lemma 6, and the remark that follows it, (1.8) would be established for $2 \leqq p < \infty$ and all positive semidefinite matrices X and Y.

It remains to establish the convexity of $G(X)$. We choose a self-adjoint matrix A and define

$$\phi(s) = \|(X + sA) + Y\|_p^p + \|(X + sA) - Y\|_p^p - 2\|(X + sA)\|_p^p. \tag{4.3}$$

Then

$$\frac{d}{ds}\phi(s) = p\mathrm{Tr}[((X+sA)+Y)^{p-1} + ((X+sA)-Y)^{p-1} - 2(X+sA)^{p-1}]A.$$

Using the integral representation formula to compute the next derivative, we have

$$\phi''(0) = p\gamma_p \int_0^\infty t^{(p-1)}\mathrm{Tr}\left(\left[\frac{1}{t+X+Y}A\frac{1}{t+X+Y} + \frac{1}{t+X-Y}A\frac{1}{t+X-Y} - 2\frac{1}{t+X}A\frac{1}{t+X}\right]A\right)dt.$$

This is positive by the convexity of

$$X \mapsto \mathrm{Tr}\frac{1}{t+X}A\frac{1}{t+X}A,$$

which we established in the last section. □

Proof of Theorem 2(b) We use a power doubling argument inspired by the 2-convexification method developed by Figiel and Johnson [FJ]. First, consider the case $p \geqq 4$. As in the proof of Theorem 1, we can assume that X and Y are self-adjoint $n \times n$ matrices. The spectrum of the $2n \times 2n$ matrix

$$\begin{pmatrix} X & Y \\ Y & X \end{pmatrix} \tag{4.4}$$

consists of the union of the spectra of $X+Y$ and of $X-Y$. Thus, the pth power of its C_p norm equals the left side of (1.8). By the same spectral considerations, one sees that the pth power of the C_p norm of the 2×2 matrix

$$\begin{pmatrix} \|X\| & \|Y\| \\ \|Y\| & \|X\| \end{pmatrix} \tag{4.5}$$

equals the right side of (1.8). Thus, our problem is to show that the C_p norm of the 2×2 matrix in (4.5) exceeds the C_p norm of the $2n\times 2n$ matrix in (4.4).

Now

$$\left\|\begin{pmatrix} X & Y \\ Y & X \end{pmatrix}\right\|_p^2 = \left\|\begin{pmatrix} X^2+Y^2 & XY+YX \\ XY+YX & X^2+Y^2 \end{pmatrix}\right\|_{p/2}. \tag{4.6}$$

The second matrix is positive semidefinite, and it has the special block form $\begin{pmatrix} A & B \\ B & A \end{pmatrix}$. Block matrices of this form are characterized by the fact that they commute with $\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$, where I is the $n\times n$ identity matrix. Evidently, all powers of a positive semidefinite block matrix of this special form have the same special form.

Thus, if r is the index conjugate to $p/2$, there is a positive matrix $\begin{pmatrix} C & D \\ D & C \end{pmatrix}$ whose C_r-norm is 1 with the property that the norm in (4.6) is realized as

$$\mathrm{Tr}\begin{pmatrix} X^2+Y^2 & XY+YX \\ XY+YX & X^2+Y^2 \end{pmatrix}\begin{pmatrix} C & D \\ D & C \end{pmatrix} = 2\mathrm{Tr}(X^2+Y^2)C + 2\mathrm{Tr}(XY+YX)D$$

$$\leqq 2\|C\|_r(\|X^2\|_{p/2} + \|Y^2\|_{p/2}) + 4\|D\|_r\|X\|_p\|Y\|_p$$

by the Hölder inequality for traces of matrices.

Let us define $\| C \| := \| C \|_r$, $\| D \| := \| D \|_r$, $\| X \| := \| X \|_p$ and $\| Y \| := \| Y \|_p$. The last expression is

$$2\| C \|(\| X \|^2 + \| Y \|^2) + 4\| D \| \, \| X \| \, \| Y \|$$
$$= \mathrm{Tr}\begin{pmatrix} \| C \| & \| D \| \\ \| D \| & \| C \| \end{pmatrix}\begin{pmatrix} \| X \|^2 + \| Y \|^2 & 2\| X \| \, \| Y \| \\ 2\| X \| \, \| Y \| & \| X \|^2 + \| Y \|^2 \end{pmatrix}$$
$$\leqq \left\| \begin{pmatrix} \| C \| & \| D \| \\ \| D \| & \| C \| \end{pmatrix} \right\|_r \left\| \begin{pmatrix} \| X \| & \| Y \| \\ \| Y \| & \| X \| \end{pmatrix} \right\|_p^2 .$$

The positivity of the matrix $\left(\begin{smallmatrix} C & D \\ D & C \end{smallmatrix}\right)$ guarantees that both $C + D$ and $C - D$ are positive. Since $1 \leqq r \leqq 2$, Theorem 2(c) implies that

$$\left\| \begin{pmatrix} \| C \| & \| D \| \\ \| D \| & \| C \| \end{pmatrix} \right\|_r \leqq \left\| \begin{pmatrix} C & D \\ D & C \end{pmatrix} \right\|_r = 1.$$

Consequently,

$$\left\| \begin{pmatrix} X & Y \\ Y & X \end{pmatrix} \right\|_p^2 \leqq \left\| \begin{pmatrix} \| X \| & \| Y \| \\ \| Y \| & \| X \| \end{pmatrix} \right\|_p^2 ,$$

as required. Finally, by Lemma 6, we obtain the validity of (1.8) for $1 \leqq p \leqq \frac{4}{3}$. □

Remark. For all p, (1.8) holds (with the appropriate direction of inequality) in C_p when $\| Y \| = \| X \|$ since (1.8) is then a special case of the "easy" Clarkson inequality (2.1) which was extended to C_p by Dixmier [Di]. It also holds to leading order for small Y, as one can verify using Theorem 1. We make the natural conjecture that (1.8) holds in C_p for $1 \leqq p \leqq 2$, without the restrictions imposed in part (c) of Theorem 2.

References

[ArYa] Araki, H., Yamagami, S.: An inequality for the Hilbert-Schmidt norm. Commun. Math. Phys. **81**, 89–96 (1981)

[BP] Ball, K., Pisier, G.: Unpublished result; private communication.

[Bo] Boas, R.P.: Some uniformly convex spaces. Bull. Am. Math. Soc. **46**, 304-311 (1940)

[C] Clarkson, J.A.: Uniformly convex spaces. Trans. Am. Math. Soc. **40**, 396-414 (1936)

[CL] Carlen, E., Lieb, E.: Optimal hypercontractivity for fermi fields and related non-commutative integration inequalities. Commun. Math. Phys. "**155**, 27-46 (1993); for a slightly different presentation, see: Optimal two-uniform convexity and fermion hypercontractivity. In: Araki, H., Ito, K.R., Kishimoto, A., Ojima, I. (eds,) Quantum and non-commutative analysis. London New York. Kluwer (in press)

[D] Day, M.: Uniform convexity in factor and conjugate spaces. Ann. Math. **45**, 375–385 (1944)

[Di] Dixmier, J.: Formes linéaires sur un anneau d'opérateurs. Bull. Soc. Math. Fr. **81**, 222–245 (1953)

[F] Figiel, T.: On the moduli of convexity and smoothness. Studia Math. **56**, 121-155 (1976)

[FJ] Figiel, T., Johnson, S.B.: A uniformly convex Banach space which contains no C_p. Compos. Math. **29**, 179-190 (1974)

[Gr] Gross, L.: Logarithmic Sobolev inequalities. Am. J. Math. **97**, 1061–1083 (1975)

[H] Hanner, O.: On the uniform convexity of L^p and l^p. Ark. Math. **3**, 239–244 (1956)

[Kö] Köthe, G.: Topologische lineare Räume, Die Grundlehren der mathematischen Wissen schaften in Einzeldarstellungen, Bd. 107, Springer Berlin Heidelberg New York: 1960

[L] Lindenstrauss, J.: On the modulus of smoothness and divergent series in Banach spaces. Mich. Math. J. **10**, 241–252 (1963)

[P] Pisier, G.: The volume of convex bodies and Banach space geometry. Cambridge: Cambridge University Press, 1989

[Ru] Ruskai, M.B.: Inequalities for traces on Von Neumann algebras. Commun. Math. Phys. **26**, 280–289 (1972)

[Se] Segal, I.E.: A non-commutative extension of abstract integration. Ann. Math. **57**, 401–457 (1953)

[Si] Simon, B.: Trace ideals and their applications. (See p. 22) Cambridge: Cambridge University Press, 1979

[TJ] Tomczak-Jaegermann, N.: The moduli of smoothness and convexity and Rademacher averages of trace classes $S_p(1 \leqq p < \infty)$. Studia Math. **50**, 163–182 (1974)