

AN INEQUALITY OF HADAMARD TYPE FOR PERMANENTS*

ERIC CARLEN[†], ELLIOTT H. LIEB[‡], AND MICHAEL LOSS[†]

Abstract. Let F be an $N \times N$ complex matrix whose j th column is the vector \vec{f}_j in \mathbb{C}^N . Let $|\vec{f}_j|^2$ denote the sum of the absolute squares of the entries of \vec{f}_j . Hadamard's inequality for determinants states that $|\det(F)| \leq \prod_{j=1}^N |\vec{f}_j|$. Here we prove a sharp upper bound on the permanent of F , which is $|\text{perm}(F)| \leq \frac{N!}{N^{N/2}} \prod_{j=1}^N |\vec{f}_j|$, and we determine all of the cases of equality.

We also discuss the case in which $|\vec{f}_j|$ is replaced by the ℓ_p norm of the vector \vec{f} considered as a function on $\{1, 2, \dots, N\}$. We note a simple sharp inequality for $p = 1$, and obtain bounds for intermediate p by interpolation.

Key words. Permanent, Hadamard inequality, heat kernel

AMS subject classifications. 15A45, 49M20

1. Introduction. Let F be an $N \times N$ complex matrix whose j th column is the vector \vec{f}_j in \mathbb{C}^N . Let $|\vec{f}_j|^2$ denote the sum of the absolute squares of the entries of \vec{f}_j . Hadamard's inequality for determinants [4] states that $|\det(F)| \leq \prod_{j=1}^N |\vec{f}_j|$. Here we prove a sharp upper bound on the permanent of F :

THEOREM 1.1. *For any vectors $\vec{f}_1, \dots, \vec{f}_N$ in \mathbb{C}^N we have the inequality*

$$(1.1) \quad |\text{perm}(F)| \leq \frac{N!}{N^{N/2}} \prod_{j=1}^N |\vec{f}_j|.$$

For $N > 2$, there is equality in (1.1) if and only if at least one of the vectors \vec{f}_j is zero, or else F is a rank one matrix and, moreover, each of the vectors \vec{f}_j is a constant modulus vector; i.e., its entries all have the same absolute value.

The conditions for equality can be reformulated as follows: There is equality in (1.1) if and only if one or more of the vectors \vec{f}_j is zero, or else there are numbers r_j , ξ_j , ζ_j , $j = 1, \dots, N$, with each $r_j > 0$ and each ξ_j and ζ_j lying on the unit circle in the complex plane, so that

$$F_{j,k} = \xi_j \zeta_k r_k$$

for each j, k .

We shall give two proofs of this inequality. The first turns on recognizing (1.1) as a close relative of the Brascamp–Lieb type inequality that we recently proved [2] for integrals of products of functions on the sphere \mathbb{S}^{N-1} . To explain this way of viewing (1.1), we first introduce some notation and terminology.

*Received August 2, 2005; accepted for publication November 25, 2005.

[†]School of Mathematics, Georgia Tech, Atlanta, GA 30332, USA (carlen@math.gatech.edu; loss@math.gatech.edu). Work partially supported by U.S. National Science Foundation grant DMS 03-00349.

[‡]Departments of Mathematics and Physics, Jadwin Hall, Princeton University, P. O. Box 708, Princeton, NJ 08544, USA (lieb@math.princeton.edu). Work partially supported by U.S. National Science Foundation grant PHY 01-39984.

© 2006 by the authors. This paper may be reproduced, in its entirety, for non-commercial purposes.

Let \mathcal{S}^N denote the symmetric group on N letters; i.e., the group of all permutations σ of $\{1, \dots, N\}$. Let the (composition) product in \mathcal{S}^N be denoted by juxtaposition, and for each $1 \leq i, j \leq N$ with $i \neq j$, let $\sigma_{i,j}$ be the pair permutation with $\sigma_{i,j}(i) = j$, $\sigma_{i,j}(j) = i$, $\sigma_{i,j}(k) = k$ for $k \neq i, j$. Let μ denote the uniform probability measure on \mathcal{S}^N so that if g is any function on \mathcal{S}^N ,

$$(1.2) \quad \int_{\mathcal{S}^N} g(\sigma) d\mu = \frac{1}{N!} \sum_{\sigma \in \mathcal{S}^N} g(\sigma) .$$

We may identify vectors in \mathbb{C}^N with complex valued functions on $\{1, \dots, N\}$ as follows: If $f : \{1, \dots, N\} \rightarrow \mathbb{C}$, let \vec{f} be the vector in \mathbb{C}^N whose j th entry is $f(j)$. Conversely, given a vector \vec{f} in \mathbb{C}^N , define the function f by setting $f(j)$ equal to the j th component of \vec{f} .

For $1 \leq k \leq N$, define the function $\pi_k : \mathcal{S}^N \rightarrow \{1, \dots, N\}$ by

$$\pi_k(\sigma) = \sigma(k) .$$

If $f : \{1, \dots, N\} \rightarrow \mathbb{C}$, then

$$f \circ \pi_k : \mathcal{S}^N \rightarrow \mathbb{C} .$$

Let $\{f_1, \dots, f_N\}$ be any N complex valued functions on $\{1, \dots, N\}$. For $1 \leq j \leq N$, let \vec{f}_j denote the corresponding vector in \mathbb{C}^N , and let F denote the $N \times N$ matrix whose j th column is \vec{f}_j . Then

$$(1.3) \quad \int_{\mathcal{S}^N} \prod_{j=1}^N (f_j \circ \pi_j) d\mu = \frac{1}{N!} \text{perm}(F) .$$

Let $\|\cdot\|_p$ denote the L^p norm on (\mathcal{S}^N, μ) , and note that

$$|\vec{f}_j| = \sqrt{N} \|f_j \circ \pi_j\|_2$$

so that (1.1) is equivalent to

$$(1.4) \quad \int_{\mathcal{S}^N} \prod_{j=1}^N (f_j \circ \pi_j) d\mu \leq \prod_{j=1}^N \|f_j \circ \pi_j\|_2 .$$

In the from (1.4), the inequality (1.1) bears a striking resemblance to the Brascamp–Lieb type inequality on \mathcal{S}^{N-1} that we proved in [2]. For purposes of comparison, we recall this result.

Let ν denote the uniform probability measure on \mathbb{S}^{N-1} , the unit sphere in \mathbb{R}^N . For each $j = 1, 2, \dots, N$, let \vec{e}_j denote the j th standard basis vector in \mathbb{R}^N , so that for any \vec{x} in \mathbb{S}^{N-1} , $\pi_j(\vec{x}) = \vec{x} \cdot \vec{e}_j$ is the j th component of \vec{x} . Then:

THEOREM 1.2. *For all $N \geq 2$, given non-negative measurable functions f_1, \dots, f_N , on $[-1, 1]$,*

$$(1.5) \quad \int_{\mathbb{S}^{N-1}} \left(\prod_{j=1}^N f_j \circ \pi_j \right) d\nu \leq \prod_{j=1}^N \|f_j \circ \pi_j\|_{L^p(\mathbb{S}^{N-1})} .$$

for all $p \geq 2$. Moreover, the L^2 norm is optimal in that for each $p < 2$, there exist functions f_j so that $\|f_j \circ \pi_j\|_{L^p(S^{N-1})} < \infty$ for each j , while the integral on the left side of (1.5) diverges. Finally, for every $p \geq 2$ and $N \geq 3$, there is equality in (1.5) if and only if some function f_j vanishes identically, or else each f_j is constant.

Note that Theorem 1.2, about functions on the sphere, provides sharp information on the ratio

$$(1.6) \quad \frac{\int_{S^{N-1}} \left(\prod_{j=1}^N f_j \circ \pi_j \right) d\nu}{\prod_{j=1}^N \|f_j \circ \pi_j\|_{L^p(S^{N-1})}}$$

for all values of p . However, Theorem 1.1, about functions on the permutation group, only provides sharp information on the ratio

$$(1.7) \quad \frac{\int_{\mathcal{S}^N} \prod_{j=1}^N (f_j \circ \pi_j) d\mu}{\prod_{j=1}^N \|f_j \circ \pi_j\|_p}$$

for $p \geq 2$. Moreover, while the ratio in (1.6) can have an infinite numerator, and finite denominator for $p < 2$, this is clearly impossible for (1.7). In fact, it is easy to obtain a sharp upper bound on (1.7) for $p = 1$.

The fact that 2 is the critical L^p index in the inequality (1.5) for all values of N has interesting consequences for the subadditivity of entropy on the sphere \mathbb{S}^{N-1} , as explained in [2], and analogous entropy inequalities for \mathcal{S}^N would follow from (1.4) in the same way. A recent paper [1] of Barthe, Cordero-Erausquin and Maurey provides an illuminating geometric perspective on the criticality of $p = 2$ in (1.5), among other things.

Thus, while (1.4) bears a clear family resemblance to (1.5), there are differences. Nonetheless, (1.4) can be proved by the same means that were employed in [2] to prove (1.5), as we explain in Section 2. Next, in Section 3, we give an alternative proof of Theorem 1.1. This second proof gives a result that is a little more general than Theorem 1.1. In Section 4, we prove the bounds on (1.7) for $1 \leq p \leq 2$, and we discuss certain natural open problems. Finally, in an appendix we prove an interpolation theorem that is used in Section 4.

After submitting the paper for publication we were informed by Leonid Gurvits that a different proof of Theorem 1.1 can be obtained from Theorem 9.1.1 in Appendix 1 of the book by Nesterov and Nemirovskii [5]. Their statement, however, does not cover the cases of equality. We are grateful to Leonid Gurvits and Alex Samorodnitsky for informative email exchanges on this and related topics.

2. First Proof of Theorem 1.1. Theorem 1.2 was proved using a monotone heat kernel interpolation argument. Theorem 1.1 can be proved in the same manner. We will use the heat kernel to generate a norm preserving interpolation between arbitrary non-negative trial functions and constant functions. We will show that under this interpolation, the ratio in (1.7) is increasing. It turns out that this interpolation is very simple, and can be written down explicitly without reference to the heat kernel, as we explain after the proof; see (2.20). Nonetheless, the heat kernel formulation facilitates checking that the ratio in (1.7) increases along the interpolation.

We now explain the heat kernel interpolation. For $1 \leq i, j \leq N$ with $i \neq j$ define the operator $D_{i,j}$ on $L^2(\mathcal{S}^N, \mu)$ by

$$D_{i,j}g(\sigma) = g(\sigma\sigma_{i,j}) - g(\sigma) .$$

One easily sees that each $D_{i,j}$ is self adjoint, and indeed, that

$$(2.1) \quad D_{i,j}^2 = -2D_{i,j} ,$$

so that $D_{i,j}$ is actually non positive. The Laplacean Δ on \mathcal{S}^N is the operator

$$-\Delta = \sum_{i < j} D_{i,j}^2 .$$

Note that a function g on \mathcal{S}^N is of the form $f \circ \pi_j$ for some $f : \{1, \dots, N\} \rightarrow \mathbb{R}$ if and only if

$$D_{k,\ell}g = 0$$

whenever $k \neq j$ and $\ell \neq j$.

Next observe that the left hand side of (1.1) is only increased if we replace each entry in each \vec{f}_j by its absolute value, and the right hand side is unchanged by this operation. Hence, in proving the inequality, we may assume without loss of generality that each of the functions f_j is non negative, and of course, that none are identically zero.

For any non negative function f on $\{1, \dots, N\}$ and any $1 \leq j \leq N$, and any $1 \leq p < \infty$, consider the function defined by $(e^{t\Delta}(f \circ \pi_j)^p)^{1/p}$. Since Δ commutes with every $D_{k,\ell}$, whenever $k \neq j$ and $\ell \neq j$ we have that

$$D_{k,\ell}(e^{t\Delta}(f \circ \pi_j)^p) = e^{t\Delta}(D_{k,\ell}(f \circ \pi_j)^p) = 0 .$$

Hence $(e^{t\Delta}(f \circ \pi_j)^p)^{1/p}$ depends on σ only through $\pi_j(\sigma)$, and for $t \geq 0$, we can define the function $f(\cdot, t)$ on $\{1, \dots, N\}$ by

$$(2.2) \quad f(t, \pi_j) = (e^{t\Delta}(f \circ \pi_j)^p)^{1/p} .$$

Notice that since $e^{t\Delta}$ preserves the integrals of functions,

$$(2.3) \quad \|f(t, \pi_j)\|_p = \|f(0, \pi_j)\|_p = \|f \circ \pi_j\|_p$$

for all $t > 0$. Moreover, since the null space of Δ consists of the constant functions on \mathcal{S}^N and nothing else, and since all non zero eigenvalues of Δ are strictly negative, $\lim_{t \rightarrow \infty} e^{t\Delta}(f \circ \pi_j)^p = \int_{\mathcal{S}^N} (f \circ \pi_j)^p d\mu$, and hence

$$(2.4) \quad \lim_{t \rightarrow \infty} f(t, \pi_j) = \|f \circ \pi_j\|_p .$$

Now, given N non negative, non zero functions f_j on $\{1, \dots, N\}$, and $t > 0$, define

$$(2.5) \quad \eta_p(t) = \int_{\mathcal{S}^N} \left(\prod_{j=1}^N f_j(t, \pi_j) \right) d\mu .$$

(Note that by (2.2), the right hand side does depend on p , as indicated by the subscript on the left hand side.) By (2.4), $\lim_{t \rightarrow \infty} \eta_p(t) = \prod_{j=1}^N \|f_j \circ \pi_j\|_p$, while clearly $\eta_p(0) = \int_{\mathcal{S}^N} \prod_{j=1}^N f_j \circ \pi_j d\mu$. Hence the inequality (1.4) would be proved if we could show that $t \mapsto \eta_2(t)$ is non decreasing. Moreover, it is clear that for all $s, t > 0$ and all j ,

$$f(t+s, \pi_j) = (e^{s\Delta}(f(t, \pi_j))^p)^{1/p} .$$

Because of this semigroup property, it would suffice to show that

$$(2.6) \quad \left. \frac{d}{dt} \eta_2(t) \right|_{t=0} \geq 0 .$$

This is indeed what we shall do. (Notice that the differentiability of η_2 is not an issue in this finite dimensional setting.) The following lemmas prepare the way for the computation of the left side of (2.6).

LEMMA 2.1. *For any function g on \mathcal{S}^N ,*

$$(2.7) \quad \Delta g^2 = 2(\Delta g)g + 2 \sum_{i < j} |D_{i,j}g|^2 .$$

Proof. We compute

$$|D_{i,j}g(\sigma)|^2 = g^2(\sigma\sigma_{i,j}) + g^2(\sigma) - 2g(\sigma\sigma_{i,j})g(\sigma)$$

and

$$(2D_{i,j}g(\sigma))g(\sigma) = 2g(\sigma\sigma_{i,j})g(\sigma) - 2g^2(\sigma) .$$

Summing these equations, one has

$$(2.8) \quad \begin{aligned} |D_{i,j}g(\sigma)|^2 + (2D_{i,j}g(\sigma))g(\sigma) &= g^2(\sigma\sigma_{i,j}) - g^2(\sigma) \\ &= D_{i,j}g^2(\sigma) . \end{aligned}$$

Multiplying through by 2, summing over $i < j$, and making use of (2.1), one obtains (2.7). \square

Lemma 2.1 justifies the following notational convention: We define $|\nabla g|^2$ by

$$(2.9) \quad |\nabla g|^2 = \sum_{i < j} |D_{i,j}g|^2 .$$

LEMMA 2.2. *With $f(t, \sigma)$ defined by $f(t, \sigma) = (e^{t\Delta}(f \circ \pi_j)^2(\sigma))^{1/2}$,*

$$(2.10) \quad \left. \frac{\partial}{\partial t} f(t, \sigma) \right|_{t=0} = \Delta(f \circ \pi_j) + \frac{|\nabla(f \circ \pi_j)|^2}{(f \circ \pi_j)} .$$

Proof. This is a simple computation using (2.7). \square

The following lemma gives the modification to the Leibniz rule for the finite difference operation $D_{i,j}$. What is crucial for us is that the modification drops out if one of the functions does not depend on σ through either $\pi_i(\sigma)$ or $\pi_j(\sigma)$.

LEMMA 2.3. *For any two function g and h on \mathcal{S}^N , and any $1 \leq i < j \leq N$,*

$$D_{i,j}(gh)(\sigma) = (D_{i,j}g(\sigma))h(\sigma) + g(\sigma\sigma_{i,j})(D_{i,j}h(\sigma)) .$$

In particular, if $D_{i,j}h = 0$,

$$(2.11) \quad D_{i,j}(gh) = (D_{i,j}g)h .$$

Proof. This is an even simpler computation, in which one makes the obvious addition and subtraction. \square

LEMMA 2.4. *The inequality (1.4) is satisfied for any N non negative functions f_j on $\{1, \dots, N\}$*

Proof. We may freely assume that none of the functions is identically zero, since then (1.4) is trivially satisfied, with zero on both sides.

Define $f_j(t, \cdot)$ by

$$f_j(t, \pi_j) = (e^{t\Delta}(f_j \circ \pi_j)^2)^{1/2} .$$

and $\eta_2(t)$ by $\eta_2(t) = \int_{S^N} \left(\prod_{j=1}^N f_j(t, \pi_j) \right) d\mu$. Then by Lemma 2.2,

$$\frac{d}{dt} \eta_2(t) \Big|_{t=0} = \sum_{j=1}^N \int_{S^N} \left(\Delta(f_j \circ \pi_j) + \frac{|\nabla(f_j \circ \pi_j)|^2}{(f_j \circ \pi_j)} \right) \left(\prod_{k=1, k \neq j}^N (f_k \circ \pi_k) \right) d\mu .$$

Consider the contribution coming from

$$\int_{S^N} \Delta(f_j \circ \pi_j) \left(\prod_{k=1, k \neq j}^N (f_k \circ \pi_k) \right) d\mu .$$

Notice that

$$-\Delta(f_j \circ \pi_j) = \sum_{i>j} D_{j,i}^2(f_j \circ \pi_j) + \sum_{i<j} D_{i,j}^2(f_j \circ \pi_j) .$$

Since $D_{i,j}$ is self adjoint,

$$\begin{aligned} - \int_{S^N} \Delta(f_j \circ \pi_j) \left(\prod_{k=1, k \neq j}^N (f_k \circ \pi_k) \right) d\mu &= \sum_{i>j} \int_{S^N} D_{j,i}(f_j \circ \pi_j) D_{j,i} \left(\prod_{k=1, k \neq j}^N (f_k \circ \pi_k) \right) d\mu \\ &\quad + \sum_{i<j} \int_{S^N} D_{i,j}(f_j \circ \pi_j) D_{i,j} \left(\prod_{k=1, k \neq j}^N (f_k \circ \pi_k) \right) d\mu . \end{aligned}$$

(2.12)

Now by (2.11),

$$\begin{aligned} &\int_{S^N} D_{j,i}(f_j \circ \pi_j) D_{j,i} \left(\prod_{k=1, k \neq j}^N (f_k \circ \pi_k) \right) d\mu \\ &= \int_{S^N} D_{j,i}(f_j \circ \pi_j) D_{j,i}(f_i \circ \pi_i) \left(\prod_{k=1, k \neq j, i}^N (f_k \circ \pi_k) \right) d\mu \\ &= \int_{S^N} \frac{D_{j,i}(f_j \circ \pi_j)}{f_j \circ \pi_j} \frac{D_{j,i}(f_i \circ \pi_i)}{f_i \circ \pi_i} \left(\prod_{k=1}^N (f_k \circ \pi_k) \right) d\mu . \end{aligned}$$

(2.13)

Defining the non negative function ρ by $\rho = \prod_{k=1}^N (f_k \circ \pi_k)$ and making a similar computation for the sum on $i < j$, we obtain

$$(2.14) \quad \int_{\mathcal{S}^N} \Delta(f_j \circ \pi_j) \left(\prod_{k=1, k \neq j}^N (f_k \circ \pi_k) \right) d\mu = \sum_{i \neq j} \int_{\mathcal{S}^N} \left(\frac{D_{i,j}(f_j \circ \pi_j)}{f_j \circ \pi_j} \frac{D_{i,j}(f_i \circ \pi_i)}{f_i \circ \pi_i} \right) \rho d\mu .$$

From here we see that

$$(2.15) \quad \left. \frac{d}{dt} \eta(t) \right|_{t=0} = \frac{1}{2} \sum_{i,j} \int_{\mathcal{S}^N} \left(\frac{D_{i,j}(f_j \circ \pi_j)}{f_j \circ \pi_j} - \frac{D_{i,j}(f_i \circ \pi_i)}{f_i \circ \pi_i} \right)^2 \rho d\mu . \quad \square$$

LEMMA 2.5. *Let $N > 2$. If there is equality in (1.4), and if none of the functions f_j is identically zero, then each of them has a constant modulus. That is, for each j and k , $|f_j(k)| = \|f_j\|_2$.*

Proof. Suppose that for some N functions f_j on $\{1, \dots, N\}$, there is equality in (1.4), and that none of the functions vanishes identically. Then clearly there is still equality in (1.4) if we replace each f_j by $|f_j|$. Hence, we may freely assume that the functions are all non negative, and that none vanishes identically.

It now follows from the proof of Lemma 2.4 that $\frac{d}{dt} \eta_2(t) = 0$ for all t . However, for all $t > 0$, each $f_j(t, \cdot)$ will be strictly positive, and so from (2.15) we see that for each $t > 0$, we must have

$$(2.16) \quad \frac{D_{i,j}(f_j(t, \pi_j))}{f_j(t, \pi_j)} = \frac{D_{i,j}(f_i(t, \pi_i))}{f_i(t, \pi_i)}$$

for all i and j , at every σ .

Fix any $i \neq j$ in $\{1, \dots, N\}$. For $N > 2$, we can choose k from $\{1, \dots, N\}$ so that $k \neq i, j$. Note that

$$\pi_i(\sigma_{i,k} \sigma_{i,j}) = \pi_i(\sigma_{i,j}) = j .$$

Thus, the right side of (2.16) vanishes at $\sigma = \sigma_{i,k}$. Hence the left side vanishes there as well, and since

$$\pi_j(\sigma_{i,k} \sigma_{i,j}) = k \quad \text{while} \quad \pi_j(\sigma_{i,j}) = i .$$

We therefore conclude that for each j , $f_j(t, k) = f_j(t, i)$ for all $i, k \neq j$.

This is almost what we seek. To conclude the proof, consider the permutation σ_c with

$$\sigma_c(i) = j \quad \sigma_c(j) = k \quad \text{and} \quad \sigma_c(k) = i .$$

Since,

$$\pi_j(\sigma_c \sigma_{i,j}) = k \quad \text{and} \quad \pi_j(\sigma_c) = j ,$$

the numerator in the right side of (2.16) is $f_i(t, k) - f_i(t, j)$ at σ_c . By what we have seen above, this is zero. Therefore, the numerator of the left hand side of (2.16) vanishes at σ_c . This is $f_j(t, j) - f_j(t, i)$. Hence for each $t > 0$, and each j , $f_j(t, \cdot)$ is constant. By continuity, it follows that each $f_j(\cdot)$ is constant. \square

To complete the proof of Theorem 1.1, consider N functions f_j on $\{1, \dots, N\}$, none of which is identically zero, and for which equality holds (1.4). Then we know that each $|f_j(k)|$ is non zero, and so we can define a complex number $Z_{k,j}$ by

$$Z_{k,j} = f_j(k)/|f_j(k)| .$$

Clearly, each of the $Z_{j,k}$ lies on the unit circle.

Now let A be the $N \times N$ matrix with $A_{j,k} = |f_{j,k}|$, and let Z be the $N \times N$ matrix with entries $Z_{j,k}$. Then, if F is the $N \times N$ matrix with $F_{j,k} = f_j(k)$, $F = Z \cdot A$, where the right hand side is the Hadamard product of Z and A . Clearly,

$$(2.17) \quad |\text{perm}(Z \cdot A)| \leq |\text{perm}(A)|$$

and there is equality if and only if the quantity

$$(2.18) \quad \prod_{j=1}^N Z_{j,\sigma(j)}$$

does not depend on σ .

LEMMA 2.6. *Let Z be an $N \times N$ matrix such that for each j and k , the j, k th entry is a complex number $Z_{k,j}$ lying on the unit circle. Then, the product in (2.18) is independent of σ if and only if there are vectors $\vec{\xi}$ and $\vec{\zeta}$ in \mathbb{C}^N with each entry lying in the unit circle such that for each j and k , $Z_{j,k} = \xi_j \zeta_k$.*

Proof. Suppose that $Z_{j,k} = \xi_j \zeta_k$. Then

$$\prod_{j=1}^N Z_{j,\sigma(j)} = \prod_{j=1}^N \xi_j \zeta_{\sigma(j)} = \left(\prod_{j=1}^N \xi_j \right) \left(\prod_{k=1}^N \zeta_k \right) ,$$

which is independent of σ . This proves sufficiency.

To prove necessity, we use induction. The lemma is clearly true for $N \leq 2$, so we start with 3. Let us expand in the first row of Z .

$$\text{perm}(Z) = \sum_{i=1}^N Z_{1i} \text{perm}(Z^i)$$

where Z^i is the the $(N-1) \times (N-1)$ matrix cofactor of Z_{1i} in Z . Since $|\text{perm}(Z)| = N!$ and $|\text{perm}(Z^i)| \leq (N-1)!$ we must have

$$|\text{perm}(Z^i)| = (N-1)! .$$

By induction, Z^i must be of the form

$$Z_{j,k} = \xi_j \zeta_k \quad \text{for} \quad j \neq 1, k \neq i .$$

Likewise any $(N-1) \times (N-1)$ submatrix complementary to $Z_{\alpha,\beta}$ must have the form

$$(2.19) \quad Z_{j,k} = \xi_j \zeta_k \quad \text{for } j \neq \alpha, \quad k \neq \beta .$$

The $(N-1)$ dimensional vectors ξ and ζ depend on α and β , in principle, but this is not so, as we now show. Let $1 \leq j, k, l, m \leq N$ be four integers. There is an $\alpha \leq N$ that is different from j and l . Similarly, there is a β different from k and m . Then, equation (2.19) is valid, with the *same* ξ and ζ , for both $Z_{j,k}$ and for $Z_{l,m}$, that is, $Z_{j,k} = \xi_j \zeta_k$ and $Z_{l,m} = \xi_l \zeta_m$. From this we see that

$$Z_{j,k} Z_{l,m} = Z_{j,m} Z_{l,k}$$

for *any* quartet of indices $1 \leq j, k, l, m \leq N$. With $l = m = 1$ we then deduce that $Z_{j,k} = Z_{j,1} Z_{1,k} / Z_{1,1}$, and we are done. \square

Proof of Theorem 1.1. We have already explained that (1.4) is equivalent to (1.1), and (1.4) has been proved in Lemma 2.4. The statement concerning the cases of equality then follows from Lemmas 2.5 and 2.6. \square

REMARK. It is not hard to derive an explicit formula formula for the interpolation; i.e., the right hand side of (2.2). We will write it down using the following notation:

For any vector \vec{f} in \mathbb{R}^n , let $f(j)$ denote the j th component of \vec{f} . If all of the components are non negative, then for every $p > 0$ we can define the new vector \vec{f}^p by

$$\vec{f}^p(j) = (f(j))^p \quad \text{for all } j = 1, \dots, N .$$

Also, define the corresponding ‘‘constant vector’’ $\langle \vec{f} \rangle$ by

$$\langle \vec{f} \rangle(j) = \frac{1}{N} \sum_{k=1}^N f(k) \quad \text{for all } j = 1, \dots, N .$$

Notice that \vec{f} and $\langle \vec{f}^2 \rangle^{1/2}$ have the same ℓ_2 norm.

The following is the simplest norm preserving interpolation between \vec{f} and $\langle \vec{f}^2 \rangle^{1/2}$: For each $t > 0$, consider the vector

$$(2.20) \quad \left((\vec{f}^2 - \langle \vec{f}^2 \rangle) e^{-t} + \langle \vec{f}^2 \rangle (1 - e^{-t}) \right)^{1/2} .$$

It is easy to see that up to a rescaling of t , this is in fact an ℓ_2 norm preserving interpolation between \vec{f} and the constant vector $\langle \vec{f}^2 \rangle^{1/2}$.

It is also not hard to check that this is the same interpolation as defined in (2.2). To see this, let f be any non negative function on $\{1, \dots, N\}$, and let \vec{f} be the corresponding vector. It is easy to compute $\Delta f \circ \pi_k$ using the fact that $\Delta = 2 \sum_{i \neq j} D_{i,j}$. One finds that

$$(2.21) \quad \Delta f \circ \pi_k = 4N (\langle f \rangle - f \circ \pi_k) .$$

Once one has (2.21), one easily finds

$$e^{t\Delta} f \circ \pi_k = (f \circ \pi_k - \langle f \rangle) e^{-4Nt} + \langle f \rangle (1 - e^{-4Nt}) .$$

From here, one gets an explicit formula for our non linear heat kernel interpolation:

$$(2.22) \quad (e^{t\Delta} f^2 \circ \pi_k)^{1/2} = ((f^2 \circ \pi_k - \langle f^2 \rangle) e^{-4Nt} + \langle f^2 \rangle (1 - e^{-4Nt}))^{1/2} .$$

But for a rescaling of t , this is the interpolation (2.20). While the right side is elementary, the use of heat kernel identities seems to facilitate checking the monotonicity that is central to the proof. Nonetheless, the formula (2.20) does shed some light on issues raised in the final section.

3. Second Proof of Theorem 1.1. The second proof is based on induction and the arithmetic–geometric mean inequality. In that sense, it uses only elementary tools. However, as will be seen, they must be applied in a particularly judicious way. In any case, the structure of this proof leads naturally to a generalization of Theorem 1.1 to non square matrices F .

Given $K \leq N$ vectors in \mathbb{C}^N , we form the $K \times N$ matrix whose i th row is \vec{f}_i . Let $f_{i,j}$ denote the i, j th entry of this matrix; i.e., the j th entry of \vec{f}_i . Define the functional $\mathcal{P}(\vec{f}_1, \dots, \vec{f}_K)$ by

$$(3.1) \quad \mathcal{P}(\vec{f}_1, \dots, \vec{f}_K) = \left[\sum_{1 \leq j_1 < j_2 < \dots < j_K \leq N} \left(\text{perm} \begin{bmatrix} f_{1,j_1} & f_{1,j_2} & \dots & f_{1,j_K} \\ f_{2,j_1} & f_{2,j_2} & \dots & f_{2,j_K} \\ \dots & \dots & \dots & \dots \\ f_{K,j_1} & f_{K,j_2} & \dots & f_{K,j_K} \end{bmatrix} \right)^2 \right]^{1/2} .$$

Notice that each of the permanents in (3.1) is the permanent of a $K \times K$ matrix. Note also that for $K = N$, there is only one term in the sum, and

$$(3.2) \quad \mathcal{P}(\vec{f}_1, \dots, \vec{f}_N) = \text{perm}[\vec{f}_1, \dots, \vec{f}_N] .$$

We shall now prove:

THEOREM 3.1. *With the notation defined above, we have the inequality*

$$(3.3) \quad \mathcal{P}(\vec{f}_1, \dots, \vec{f}_K) \leq \sqrt{\binom{N}{K}} \frac{K!}{N^{K/2}} \prod_{j=1}^K |\vec{f}_j| .$$

If $K \geq 2$ and none of the vectors \vec{f}_i is the zero vector, then (3.3) holds with equality if and only if $[\vec{f}_1, \dots, \vec{f}_K]$ is a rank one matrix, and each the vectors \vec{f}_i is a constant modulus vector.

Because of (3.2) Theorem 3.1 reduces to Theorem 1.1 in the case $K = N$.

Proof. As before, when proving the inequality, we may assume that all entries of each vector are non negative. The proof proceeds by induction in K . The inequality is trivial in case $K = 1$, although any vector yields equality in this case. The first non trivial case is $K = 2$. We now treat this case explicitly, since the same sort of reasoning will be employed in the general inductive step.

$$(3.4) \quad \mathcal{P}(\vec{f}_1, \vec{f}_2)^2 = \sum_{i < j} (f_{1,i} f_{2,j} + f_{1,j} f_{2,i})^2 = \sum_{i < j} f_{1,i}^2 f_{2,j}^2 + f_{1,j}^2 f_{2,i}^2 + \sum_{i < j} 2 f_{1,i} f_{2,j} f_{1,j} f_{2,i} .$$

We will use the arithmetic–geometric mean inequality on the terms of the last sum, but there are two natural ways to do this, *and we need to use some of each*. Therefore, rewrite the last sum as

$$\sum_{i < j} 2f_{1,i}f_{2,j}f_{1,j}f_{2,i} = \alpha \sum_{i < j} 2(f_{1,i}f_{2,j})(f_{1,j}f_{2,i}) + (1 - \alpha) \sum_{i < j} 2(f_{1,i}f_{2,i})(f_{1,j}f_{2,j})$$

where $\alpha < 1$ will be determined later. By the arithmetic–geometric mean inequality,

$$(3.5) \quad \sum_{i < j} 2f_{1,i}f_{2,j}f_{1,j}f_{2,i} \leq \alpha \sum_{i < j} [(f_{1,i}f_{2,j})^2 + (f_{1,j}f_{2,i})^2] + (1 - \alpha) \sum_{i < j} [(f_{1,i}f_{2,i})^2 + (f_{1,j}f_{2,j})^2] .$$

Combing (3.4) and (3.5),

$$\mathcal{P}(\vec{f}_1, \vec{f}_2)^2 \leq (1 + \alpha) \sum_{i \neq j} f_{1,i}^2 f_{2,j}^2 + (1 - \alpha)(N - 1) \sum_i (f_{1,i}f_{2,i})^2 .$$

Choose $\alpha = (N - 2)/N$ which yields $1 + \alpha = (1 - \alpha)(N - 1) = 2(N - 1)/N$, and find

$$\mathcal{P}(\vec{f}_1, \vec{f}_2)^2 \leq 2 \frac{N - 1}{N} \sum_{i,j} f_{1,i}^2 f_{2,j}^2 = 2 \frac{(N - 1)}{N} |\vec{f}_1|^2 |\vec{f}_2|^2 ,$$

which is the desired result for $K = 2$.

For the general case we can write

$$\mathcal{P}(\vec{f}_1, \dots, \vec{f}_K)^2 = \frac{1}{K!} \sum'_{j_1, \dots, j_K} \left(\sum_{k=1}^K f_{1,j_k} a_{j_1, \dots, \widehat{j_k}, \dots, j_K} \right)^2$$

where the prime over the summation sign indicates that the sum is over distinct indices. Here

$$a_{j_1, \dots, \widehat{j_k}, \dots, j_K} = \text{perm} \begin{bmatrix} f_{2,j_1} & \dots & \widehat{f_{2,j_k}} & \dots & f_{2,j_K} \\ \dots & & & & \\ f_{K,j_1} & \dots & \widehat{f_{K,j_k}} & \dots & f_{K,j_K} \end{bmatrix}$$

is the permanent of the matrix where the f_1 row and the j_k column has been removed so that we take the permanent of an $(K - 1) \times (K - 1)$ matrix. The removal of the column is indicated by the circumflex symbol, as usual. Note that this expression is symmetric in the indices. Developing the square leads to

$$\begin{aligned} \frac{1}{K!} \sum'_{j_1, \dots, j_K} \sum_{k,l}^K f_{1,j_k} f_{1,j_l} a_{j_1, \dots, \widehat{j_k}, \dots, j_K} a_{j_1, \dots, \widehat{j_l}, \dots, j_K} &= \frac{1}{K!} \sum'_{j_1, \dots, j_K} \sum_{k=1}^K (f_{1,j_k})^2 (a_{j_1, \dots, \widehat{j_k}, \dots, j_K})^2 \\ &+ \frac{1}{K!} \sum'_{j_1, \dots, j_K} \sum_{k \neq l}^K f_{1,j_k} f_{1,j_l} a_{j_1, \dots, \widehat{j_k}, \dots, j_K} a_{j_1, \dots, \widehat{j_l}, \dots, j_K} \end{aligned}$$

As before, we estimate the last term using the arithmetic–geometric mean inequality *in two different ways* and obtain

$$(3.6) \quad \alpha \frac{1}{(K - 2)!} \sum'_{j_1, \dots, j_K} (f_{1,j_1})^2 (a_{j_2, \dots, j_K})^2 + (1 - \alpha) \frac{1}{(K - 2)!} \sum'_{j_1, \dots, j_K} (f_{1,j_1})^2 (a_{j_1, \widehat{j_2}, \dots, j_K})^2$$

where $0 < \alpha < 1$ will be chosen below.

Focusing on the very last term, we write it as

$$\begin{aligned}
& (1 - \alpha) \frac{1}{(K-2)!} \sum_{j_1, \dots, j_K} (f_{1,j_1})^2 (a_{j_1, \widehat{j_2}, \dots, j_K})^2 \\
&= (1 - \alpha)(N - K + 1) \frac{1}{(K-2)!} \sum_{j_1} (f_{1,j_1})^2 \sum_{j_2, \dots, j_{(K-1)} \neq j_1} (a_{j_1, \dots, j_{(K-1)}})^2 \\
&= (1 - \alpha) \frac{(N - K + 1)}{K-1} \frac{1}{(K-2)!} \sum_{j_1} (f_{1,j_1})^2 \sum_{l=1}^{K-1} \sum_{j_2, \dots, j_{(K-1)} \neq j_1} (a_{j_2, \dots, j_1, \dots, j_{(K-1)}})^2
\end{aligned} \tag{3.7}$$

where the index j_1 is in the l -th position. Note that

$$\sum_{l=1}^{K-1} \sum_{j_2, \dots, j_{(K-1)} \neq j_1} (a_{j_2, \dots, j_1, \dots, j_{(K-1)}})^2 = \sum_{j_2, \dots, j_K \neq j_1} (a_{j_2, \dots, j_K})^2.$$

Hence, collecting the terms we get

$$\begin{aligned}
\mathcal{P}(\vec{f}_1, \dots, \vec{f}_K)^2 &\leq (1 + (K-1)\alpha) \frac{1}{(K-1)!} \sum_{j_1} (f_{1,j_1})^2 \sum_{j_2, \dots, j_K \neq j_1} (a_{j_2, \dots, j_K})^2 \\
&\quad + (1 - \alpha)(N - K + 1) \frac{1}{(K-1)!} \sum_{j_1} (f_{1,j_1})^2 \sum_{j_2, \dots, j_K \neq j_1} (a_{j_2, \dots, j_K})^2
\end{aligned} \tag{3.8}$$

Now choose $\alpha = (N - K)/N$ so that

$$(1 + (K-1)\alpha) = (1 - \alpha)(N - K + 1) = \frac{K(N - K + 1)}{N},$$

and

$$\mathcal{P}(\vec{f}_1, \dots, \vec{f}_K)^2 \leq \frac{K(N - K + 1)}{N} \frac{1}{(K-1)!} \sum_{j_1} (f_{1,j_1})^2 \sum_{j_2, \dots, j_K} (a_{j_2, \dots, j_K})^2$$

By the inductive hypotheses,

$$\sum_{j_2, \dots, j_K} (a_{j_2, \dots, j_K})^2 \leq (K-1)! \binom{N}{K-1} \frac{((K-1)!)^2}{N^{K-1}} \prod_{j=2}^K |\vec{f}_j|^2,$$

which yields

$$\mathcal{P}(\vec{f}_1, \dots, \vec{f}_K)^2 \leq \binom{N}{K} \frac{(K!)^2}{N^K} \prod_{j=1}^N |\vec{f}_j|^2.$$

This proves the inequality.

To establish the cases of equality, we return to the step where the arithmetic-geometric mean inequality has been employed, of course still under the assumption that the entries are all non negative. Equality entails that

$$f_{1,j}f_{2,j} = f_{1,i}f_{2,i} \text{ and } f_{1,i}f_{2,j} = f_{1,j}f_{2,i} ,$$

for all $i \neq j$. Since the vectors \vec{f}_1 and \vec{f}_2 are not allowed to be the zero vectors there exists an index i so that $f_{1,i}$ is not zero. If $f_{2,i} = 0$ then it follows from $f_{1,i}f_{2,j} = f_{1,j}f_{2,i}$ that $f_{2,j} = 0$ for all $j \neq i$. This would mean that f_2 is identically zero, contrary to the conditions in the theorem. Thus, both, $f_{1,i}$ and $f_{2,i}$ are non zero. It now follows from $f_{1,j}f_{2,j} = f_{1,i}f_{2,i}$ that neither $f_{1,j}$ nor $f_{2,j}$ can be zero. Therefore

$$\frac{f_{2,j}}{f_{2,i}} = \frac{f_{1,j}}{f_{1,i}} = \frac{f_{2,i}}{f_{2,j}}$$

for all $i < j$ and hence the vectors \vec{f}_1 and \vec{f}_2 have to be constant vectors. Since the same argument applies to any two vectors the result follows. We may now treat the case that the entries are complex using the final lemma of the previous section. \square

COROLLARY 3.2. *For all p with $1 \leq p \leq 2$, the quantity*

$$(3.9) \quad \mathcal{P}_p(\vec{f}_1, \dots, \vec{f}_K) = \left[\sum_{1 \leq j_1 < j_2 < \dots < j_K \leq n} \left(\text{perm} \begin{bmatrix} f_{1,j_1} & f_{1,j_2} & \dots & f_{1,j_K} \\ f_{2,j_1} & f_{2,j_2} & \dots & f_{2,j_K} \\ \dots & \dots & \dots & \dots \\ f_{K,j_1} & f_{K,j_2} & \dots & f_{K,j_K} \end{bmatrix} \right)^p \right]^{1/p}$$

satisfies the inequality

$$\mathcal{P}_p(\vec{f}_1, \dots, \vec{f}_K) \leq \binom{N}{K}^{1/p} \frac{K!}{N^{K/2}} \prod_{j=1}^K |\vec{f}_j| .$$

Proof. By Hölder's inequality,

$$\mathcal{P}_p(\vec{f}_1, \dots, \vec{f}_K) \leq \binom{N}{K}^{1/p-1/2} \mathcal{P}(\vec{f}_1, \dots, \vec{f}_K) .$$

Now apply Theorem 3.1 to estimate $\mathcal{P}(\vec{f}_1, \dots, \vec{f}_K)$; this yields the result. \square

4. Bound for other values of p . For $1 \leq p < \infty$, and any vector \vec{f} in \mathbb{C}^N , define

$$(4.1) \quad |\vec{f}|_p = \left(\sum_{k=1}^N |(\vec{f})_k|^p \right)^{1/p} .$$

Note that if f is the corresponding function of $\{1, \dots, N\}$, for each j we have

$$(4.2) \quad |\vec{f}|_p = N^{1/p} \|f \circ \pi_j\|_p .$$

By (1.3) and (4.2),

$$(4.3) \quad \frac{\int_{\mathcal{S}^N} \prod_{j=1}^N (f_j \circ \pi_j) d\mu}{\prod_{j=1}^N \|f_j \circ \pi_j\|_p} = \frac{N!}{N^{N/p}} \frac{|\text{perm}[\vec{f}_1, \dots, \vec{f}_N]|}{\prod_{j=1}^N |\vec{f}_j|_p}.$$

Thus, we may study the ratio in (1.7) by studying the ratio on the right in (4.3).

Define the function

$$(4.4) \quad C(p) = \sup_{\vec{f}_1, \dots, \vec{f}_N \neq 0} \left\{ \frac{|\text{perm}[\vec{f}_1, \dots, \vec{f}_N]|}{\prod_{j=1}^N |\vec{f}_j|_p} \right\}.$$

We know from Theorem 1 that

$$(4.5) \quad C(2) = \frac{N!}{N^{N/2}}$$

with equality precisely when $[\vec{f}_1, \dots, \vec{f}_N]$ is the constant matrix. Moreover, it is easy to see that $C(1) = 1$: Observe that

$$(4.6) \quad |\text{perm}[\vec{f}_1, \dots, \vec{f}_N]| \leq \sum_{\sigma \in \mathcal{S}^N} \prod_{j=1}^N |f_{j, \sigma(j)}|$$

and that

$$(4.7) \quad \prod_{j=1}^N |\vec{f}_j| = \sum_{k_1, \dots, k_N} \prod_{j=1}^N |f_{j, k_j}|$$

and note that every term in (4.6) shows up in (4.7), and hence $C(1) \leq 1$. Choosing $\vec{f}_j = \vec{e}_j$ for each j , so that F equals the identity matrix, shows that

$$(4.8) \quad C(1) = 1.$$

In fact, there is equality only if in each row and each column of F there is a single non-zero entry.

Notice that the optimizers for $C(p)$ are different for $p = 1$ and $p = 2$: For $p = 1$ we get the optimal ratio by taking $\vec{f}_j = \vec{e}_j$ for each j , while for $p = 2$ we get the optimal ratio by taking \vec{f}_j to be constant for each j .

If $\vec{f}_j = \vec{e}_j$ for each j , then

$$\frac{|\text{perm}[\vec{f}_1, \dots, \vec{f}_N]|}{\prod_{j=1}^N |\vec{f}_j|_p} = 1.$$

However, if \vec{f}_j is a non zero constant vector for each j , then

$$\frac{|\text{perm}[\vec{f}_1, \dots, \vec{f}_N]|}{\prod_{j=1}^N |\vec{f}_j|_p} = \frac{N!}{N^{N/p}}.$$

Evidently,

$$(4.9) \quad C(p) \geq \max \left\{ 1, \frac{N!}{N^{N/p}} \right\}.$$

Note that there is equality at $p = 1$ and $p = 2$. Pietro Caputo, to whom we sent an early draft of this paper, has suggested to us that in fact there should be equality in (4.9) for $1 < p < 2$, as well. It is easy to see that this is true for $N = 2$. In that case, we may assume without loss of generality that $\vec{f}_1 = \begin{bmatrix} 1 \\ x \end{bmatrix}$ and $\vec{f}_2 = \begin{bmatrix} y \\ 1 \end{bmatrix}$ for some non negative numbers x and y . Then $\text{perm}[\vec{f}_1, \vec{f}_2] = 1 + xy$, and by Hölder's inequality,

$$1 + xy \leq (1 + x^p)^{1/p} (1 + y^q)^{1/q}$$

where $1/q = 1 - 1/p$. For $1 \leq p \leq 2$, $q \geq p$, and so

$$(1 + y^q)^{1/q} \leq (1 + y^p)^{1/p},$$

with equality for $p < 2 < q$ if and only if $y = 0$. We conclude that

$$1 + xy \leq (1 + x^p)^{1/p} (1 + y^p)^{1/p},$$

which is the desired inequality. Moreover, for $p = 2$, by the condition for equality in the Schwarz inequality, there is equality if and only if $x = y = 1$, while when $1 \leq p < 2$, there is equality if and only if $x = y = 0$.

Beyond this trivial case, we have not succeeded in proving (4.9), but we do have the following upper bound on $C(p)$:

THEOREM 4.1. *The function $\ln(C(p))$ is a convex function of $1/p$. In particular, for all $1 \leq p \leq 2$,*

$$C(p) \leq \left(\frac{N!}{N^{N/2}} \right)^{2-2/p}.$$

Proof. The first statement follows directly from the interpolation theorem in the appendix. This is a version of the Riesz–Thorin [7] interpolation theorem for multilinear forms. The usual proof of the Riesz–Thorin interpolation theorem for operators is, in fact, an interpolation theorem for bilinear forms; see for example, [3]. It easily extends to multilinear forms, and though this seems likely to be known, we have not found any reference. We therefore include the short proof in the appendix.

Given the logarithmic convexity, for t solving the equation $t + (1 - t)/2 = 1/p$,

$$C(p) \leq C(1)^t C(2)^{1-t}.$$

The rest now follows from our computation of $C(1)$ and $C(2)$. \square

One might try to compute $C(p)$ for $1 < p < 2$ by adapting either of the two computations we made for $p = 2$. Unfortunately, we have not been able to adapt either one. For the second computation, the trouble arises at the point where we “develop the square”. For the first, there is an even more fundamental problem: The interpolation used there simply *is not monotone* for $p < 2$.

To see this consider $N = 3$ and the vectors $\vec{f}_1 = \begin{bmatrix} 1 \\ x \\ y \end{bmatrix}$, $\vec{f}_2 = \begin{bmatrix} y \\ 1 \\ x \end{bmatrix}$ and $\vec{f}_3 = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ for some numbers x and y with $0 \leq x, y \leq 1$. Then

$$[\vec{f}_1, \vec{f}_2, \vec{f}_3] = \begin{bmatrix} 1 & y & x \\ x & 1 & y \\ y & x & 1 \end{bmatrix},$$

which is a circulant matrix.

Define the function $\phi(x, y)$ by

$$\phi(x, y) = \frac{\text{perm}[\vec{f}_1, \vec{f}_2, \vec{f}_3]}{|\vec{f}_1|_p |\vec{f}_2|_p |\vec{f}_3|_p} = \frac{1 + x^3 + y^3 + 3xy}{(1 + x^p + y^p)^{3/p}} .$$

It is easy to see, using (2.22), that the class of circulant matrices is preserved under the heat semigroup flow used in the proof of Theorem 1.1 provided in Section 2, so that this flow corresponds to a path $(x(t), y(t))$ on the graph of ϕ over the unit square $0 \leq x, y \leq 1$. This path starts from the given values of x and y and satisfies

$$\lim_{t \rightarrow \infty} (x(t), y(t)) = (1, 1) .$$

However, for $p < 2$, $\phi(x, y)$ has a strict local maximum at both $(x, y) = (1, 1)$ and $(x, y) = (0, 0)$. Thus, for $p < 2$, the heat semigroup flow is initially strictly monotone decreasing when started from $\vec{f}_j = \vec{e}_j$, $j = 1, 2, 3$. Nonetheless, further analysis of the function ϕ supports the conjecture that there is equality in (4.9); the example simply shows that no flow preserving the class of circulant matrices can be used to prove it.

5. Appendix. Here we prove the following multilinear generalization of the Riesz–Thorin interpolation theorem

Let J denote a multilinear functional of M -tuples of vectors \vec{f}_j in \mathbb{C}^N . Then with $f_{j,k}$ denoting the k th component of the vector \vec{f}_j , there are numbers J_{k_1, \dots, k_M} such that

$$(5.1) \quad J(\vec{f}_1, \dots, \vec{f}_M) = \sum_{k_1, \dots, k_M} J_{k_1, \dots, k_M} \prod_{j=1}^M f_{j, k_j} .$$

For each j , let p_j satisfy $0 \leq 1/p_j \leq 1$, and define the vector \vec{p} by

$$\vec{p} = (1/p_1, \dots, 1/p_M) .$$

Define the constant $C(\vec{p})$ by

$$C(\vec{p}) = \sup_{\vec{f}_1, \dots, \vec{f}_M \neq 0} \left\{ \frac{J(\vec{f}_1, \dots, \vec{f}_M)}{\prod_{j=1}^M |\vec{f}_j|_{p_j}} \right\} .$$

THEOREM 5.1. *The function $\ln(C(\vec{p}))$ is convex on $[0, 1]^M$.*

We remark that one can extend the theorem to a version for multilinear functionals in L^p spaces by the standard approximation argument with simple functions.

Proof. Suppose that \vec{p} , \vec{q} and \vec{r} are vectors in $[0, 1]^M$ such that for some t with $0 < t < 1$,

$$\vec{p} = t\vec{q} + (1-t)\vec{r} .$$

Define the numbers q_j and r_j by $\vec{q} = (1/q_1, \dots, 1/q_M)$ and $\vec{r} = (1/r_1, \dots, 1/r_M)$ so that q_j is related to \vec{q} and r_j is related to \vec{r} the same way p_j is related to \vec{p} . We must show that

$$C(\vec{p}) \leq C(\vec{q})^t C(\vec{r})^{1-t} .$$

Towards this end, let $\vec{f}_1, \dots, \vec{f}_M$ be any M non zero vectors. We can assume that $|\vec{f}_j|_{p_j} = 1$ for each j . Since the choice of the vectors is arbitrary apart from the normalization, it suffices to show that

$$(5.2) \quad J(\vec{f}_1, \dots, \vec{f}_M) \leq C(\vec{q})^t C(\vec{r})^{1-t} .$$

Define

$$\rho_{j,k} = |f_{j,k}|^{1/p}$$

so that for each j

$$\sum_{k=1}^N \rho_{j,k} = 1 .$$

Also, define $\alpha_{j,k}$ by $\alpha_{j,k} = f_{j,k}/|f_{j,k}|$ when $f_{j,k} \neq 0$, and $\alpha_{j,k} = 0$ otherwise.

Then, for each complex number z define the vector $\vec{\rho}_j(z)$ by

$$(5.3) \quad (\vec{\rho}_j(z))_k = \alpha_{j,k} \rho_{j,k}^{z/q_j + (1-z)/r_j} .$$

Note that, for each j ,

$$(5.4) \quad \vec{\rho}_j(t) = \vec{f}_j .$$

Moreover:

(i) The right side of (5.3) is an entire function of z .

(ii) Whenever $\Re(z) = 0$, then $|\vec{\rho}_j(z)|_{r_j} = 1$

(iii) Whenever $\Re(z) = 1$ then $|\vec{\rho}_j(z)|_{q_j} = 1$.

Next we define

$$G(z) = \left| \frac{J(\vec{\rho}_1(z), \dots, \vec{\rho}_M(z))}{C(\vec{q})^z C(\vec{r})^{1-z}} \right|$$

which is a subharmonic function of z . (By (i) and (5.1), it is the absolute value of a sum of products of entire functions.) By (ii), we know that on the line $\Re(z) = 0$, $G(z) \leq 1$, and by (iii), we know that on the line $\Re(z) = 1$, $G(z) \leq 1$. Hence, by the maximum principle for subharmonic functions, $G(z) \leq 1$. But by (5.4), this yields (5.2). \square

REFERENCES

- [1] F. BARTHE, D. CORDERO-ERAUSQUIN AND B. MAUREY, *Entropy of spherical marginals and related inequalities*, preprint, (2005)
- [2] E.A. CARLEN, E.H. LIEB AND M. LOSS, *A Sharp analog of Young's Inequality on \mathbf{S}^N and Related Entropy Inequalities*, Jour. Geom. Anal., 14 (2004), pp. 487–520. arXiv math.FA/0408030.
- [3] L. GÅRDING, *Some points of analysis and their history*, University Lecture Series, 11, A.M.S., Providence, (1997).
- [4] J. HADAMARD, *Résolution d'une question relatif aux déterminants*, Bull. Sci. Math., 28 (1893), pp. 240–246.
- [5] YURII NESTEROV AND ARKADII NEMIROVSKII, *Interior-point polynomial algorithms in convex programming*, volume 13 of SIAM Studies in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.
- [6] M. RIESZ, *Sur les maxima des fonctions bilinéaires et sur les fonctionelles linéaires*, Acta. Math., 49 (1927), pp. 465–497.
- [7] O. THORIN, *An extension of a convexity theorem due to M. Riesz*, Fys. Sällsk Förh., 8:14, (1938).

