OPTIMAL HYPERCONTRACTIVITY FOR FERMI FIELDS AND RELATED NON-COMMUTATIVE INTEGRATION INEQUALITIES

Eric A. Carlen*
Department of Mathematics
Princeton University
Princeton, New Jersey 08544

Elliott H. Lieb**
Departments of Mathematics and Physics
Princeton University
Princeton, New Jersey 08544–0708

Dedicated to Prof. Huzihiro Araki on his 60^{th} birthday

Abstract: Optimal hypercontractivity bounds for the fermion oscillator semigroup are obtained. These are the fermion analogs of the optimal hypercontractivity bounds for the boson oscillator semigroup obtained by Nelson. In the process, several results of independent interest in the theory of non-commutative integration are established.

* On leave from School of Math., Georgia Institute of Technology, Atlanta, GA 30332
** Work supported by U.S. National Science Foundation grant no. PHY90–19433–A01.

I. INTRODUCTION

Observables pertaining to the configuration of a quantum system with n degrees of freedom are operators $Q_1, Q_2, \ldots, Q_n$ which, depending on the system, may or may not commute. Our main concern is with the case in which the configuration variables are amplitudes of certain field modes.

For boson fields, these configuration observables do commute, and the state space $\mathcal{H}$ can be taken as the space of all complex square integrable functions on their joint spectrum. This is the Schrödinger q-space representation, and the fact that in it the state space is a function space, and not just an abstract Hilbert space, is very helpful in the analysis of such systems. As one example, it sometimes turns out that physically interesting operators preserve the cone of positive functions, and this opens the way to the application of the Perron-Frobenius theorem in the study of ground states of such systems.

For fermion fields, the configuration observables do not commute, and this simple q-space representation is not available. However, the non-commutative integration theory of Irving Segal [Se53] permits the construction of a suitable substitute, and in fact it was created with such a purpose in view. This approach to the study of fermion systems has been extensively developed by Gross [Gr72] who, among other things, proved a version of the Perron-Frobenius theorem adapted to the setting and applied it to prove existence and uniqueness of ground states for certain fermion quantum field models.

The main estimate which enabled Gross to apply his Perron-Frobenius type theorem to fermion fields was a hypercontractivity estimate for the fermion oscillator semigroup. Corresponding hypercontractivity estimates for boson fields had been introduced earlier by Nelson [Ne66] who later obtained the optimal such bound for bosons [Ne73].

Our main result, Theorem 4 below, is the corresponding optimal hypercontractivity bound for fermions. Before stating this theorem, we describe its mathematical and physical contexts in some detail because it cannot even be formulated naturally in the conventional Fock space language. (Of course its perturbation theoretic *consequences* can be expressed quite naturally in the usual language.)

The state space for a system of n fermion degrees of freedom is conventionally realized as the Fock space

$$\mathcal{F} = \oplus_{j=o}^{n}\big((\mathbb{C}^n)^{\wedge j}\big) \tag{1.1}$$

The basic "free Hamiltonian" on this space is the fermion number operator

$$\widehat{H}_0 = \sum_{j=1}^{n} c_j^* c_j \tag{1.2}$$

where c_j and c_j^* are the usual fermion annihilation and creation operators acting on $\mathcal{F}$. The fermion oscillator semigroup is the semigroup of operators $\exp(-t\widehat{H}_0)$ that it generates.

For n boson degrees of freedom, the state space may be realized *either* as the boson Fock space, or as $L^2(\mathbb{R}^n, (2\pi)^{n/2}e^{-x^2/2}\mathrm{d}x)$. The natural isomorphism between these spaces was pointed out by Segal, and the latter may be regarded as the Schrödinger q-space realization. The boson oscillator semigroup is the semigroup generated by the boson number operator. Though we have just defined it in Fock space terms, it may also be considered as an operator semigroup on $L^2(\mathbb{R}^n, (2\pi)^{n/2}e^{-x^2/2}\mathrm{d}x)$. In this setting, it is **hypercontractive**; i.e., for any finite p greater than 2, there is t_p suffciently large, for which the semigroup is a contraction from $L^2(\mathbb{R}^n, (2\pi)^{n/2}e^{-x^2/2}\mathrm{d}x)$ to $L^p(\mathbb{R}^n, (2\pi)^{n/2}e^{-x^2/2}\mathrm{d}x)$ for all $t \geq t_p$. Since t_p is independent of n, this result proved useful in treating perturbation theoretic problems for boson fields. This hypercontractivity inequality cannot be formulated naturally in the Fock space setting because no natural notion of "L^p" can be introduced there. The $L^p(\mathbb{R}^n, (2\pi)^{n/2}e^{-x^2/2}\mathrm{d}x)$ setting is essential.

In order to to pass from the Fock space description for a system of n fermion degrees of freedom to a non-commutative analog of the Schrödinger q-space description (in which we have non-commutative analogs of L^p norms), we introduce configuration observables $Q_j = c_j + c_j^*$, and let $\mathcal{C}(\mathcal{Q})$ be the algebra with unit which they generate. This is an operator algebra on the 2^n–dimensional Hilbert space $\mathcal{F}$. As we explain in Section II, it is a Clifford algebra naturally associated to $\mathbb{C}^n$ with its usual inner product.

What follows here shall be explained in more detail in Sections II–IV, but the key fact for our present considerations is that the 2^n distinct monomials in the Q_j are a basis for $\mathcal{C}(\mathcal{Q})$ as a *vector space*. Thus, any A in $\mathcal{C}(\mathcal{Q})$ can be uniquely written as

$$A = \alpha I + \sum_j \alpha_j Q_j + \sum_{j<k} \alpha_{j,k} Q_j Q_k + \cdots + \alpha_{1,\ldots,n} Q_1 \cdots Q_n \quad . \tag{1.3}$$

It is not hard to see that any vector v in $\mathcal{F}$ can be written as $v = A\Omega$ for some unique $A \in \mathcal{C}(\mathcal{Q})$ acting on the vacuum Ω. The operator $\exp(-t\widehat{H}_0)$ acts on such vectors v. Thus, $\exp(-t\widehat{H}_0)v = B\Omega$ for some unique $B \in \mathcal{C}(\mathcal{Q})$ (which depends on t and A, of course). The map $P_t : A \mapsto B$ is clearly linear, and a simple calculation (elucidated in Section IV) shows that

$$P_t A = \alpha I + e^{-t}\sum_j \alpha_j Q_j + e^{-2t}\sum_{j<k}\alpha_{j,k}Q_jQ_k + \cdots + e^{-nt}\alpha_{1,\ldots,n}Q_1\cdots Q_n \quad . \tag{1.4}$$

Instead of considering Fock space, $\mathcal{F}$, and the operator $\exp(-t\widehat{H}_0)$, Segal's idea is to consider $\mathcal{C}(\mathcal{Q})$ and the linear operator P_t that acts on $\mathcal{C}(\mathcal{Q})$ as above. Following Segal, we can construct L^p norms on $\mathcal{C}(\mathcal{Q})$, the most important case being $p = 2$, in terms of which one can do quantum mechanics without mentioning Fock space.

To construct the norms, define the linear functional $\tau : \mathcal{C}(\mathcal{Q}) \to \mathbb{C}$ by

$$\tau(A) = \langle \Omega, A\Omega\rangle_{\mathcal{F}} \quad . \tag{1.5}$$

It turns out, as we explain in Section III that $\tau(A)$ is, for $A \in \mathcal{C}(\mathcal{Q})$, nothing other than the usual normalized trace on $\mathcal{F}$ applied to A.

The basic ingredients of a non-commutative integration theory are an algebra of operators such as $\mathcal{C}(\mathcal{Q})$, and a linear functional on it such as τ. Norms analogous to the L^p norms of commutative integration theory can now be introduced by

$$\|A\|_p = \left(\tau(A^*A)^{p/2})\right)^{1/p} \qquad \text{for} \qquad 1 \le p < \infty \quad ;$$

$\|A\|_\infty$ denotes the operator norm of A. Let $\mathcal{C}^p(\mathcal{Q})$ denote $\mathcal{C}(\mathcal{Q})$ equipped with the corresponding norm.

Having expanded our horizons beyond Hilbert space, we can ask for bounds between $\|P_tA\|_q$ and $\|A\|_p$ for different q, p and t Our main result, Theorem 4, is the optimal **fermion hypercontractivity** inequality; i.e. for all $1 < p \le q < \infty$ and all A in $\mathcal{C}^p(\mathcal{Q})$,

$$\|P_tA\|_q \le \|A\|_p \qquad \text{when} \qquad e^{-2t} \le \frac{p-1}{q-1} \quad , \tag{1.6}$$

and the t saturating the inequality on the right is the smallest for which the inequality on the left always holds. We prove this for n degrees of freedom with n an arbitrary finite integer. Since the estimate is independent of n, a theorem of Gross [Gr72] implies that it holds as well with infinitely many degrees of freedom.

The result (1.6) was conjectured by Gross [Gr75] who proved it [Gr72] in the special case $p = 2$, $q = 4$. The cases $p = 2$, $q = 2m$ where m is an integer were proved by Lindsay and Meyer [LiMe] following earlier work by Lindsay [Lin] on the case $p = 2$, $q = 2^m$. Since P_t is self adjoint, duality yields a corresponding family of results in which $q = 2$. Until now all other cases (except when $n = 1$ or 2) had remained open. The optimal relation between t, p and q found here for fermion hypercontractivity is the same as that found by Nelson [Ne73] for boson hypercontractivity.

By now there are many proofs of Nelson's inequality. Neveu's elegant proof [Nev], like Nelson's original proof, is based on probabilistic methods. Proofs based on geometric methods have been given by Carlen and Loss [CL90] and by Lieb [Li90] who considers generalizations in which the Mehler kernel (the kernel for the boson oscillator semigroup) is replaced by an arbitrary Gaussian kernel. More references can be found in the bibliography to Gross's article [Gr89]. However, none of the existing approaches to the boson problem has been found to solve the fermion problem.

As a corollary to the optimal fermion hypercontractivity inequality, we obtain the optimal **fermion logarithmic Sobolev inequality**:

$$\tau\big(|A|^2 \ln |A|^2\big) - \big(\|A\|_2^2 \ln \|A\|_2^2\big) \le 2\tau(A^* H_0 A) \tag{1.7}$$

where $|A| = (A^*A)^{1/2}$.

This inequality was conjectured by Gross, and proved by him in a weaker form with the constant on the right increased by a factor of $\ln 3$. In studying perturbations of H_0 by multiplication operators V, this inequality plays the same role as does the usual Sobolev inequality in studying perturbations of $-\frac{1}{2}\Delta$ by multiplication operators V on $L^2(\mathbb{R}^n, \mathrm{d}^n x)$ [Fe69]. (The multiplication operator associated to a self-adjoint element V of $\mathcal{C}(\mathcal{Q})$ is defined to be the average of left and right multiplication by V, and is denoted here by the same symbol V.) Again, in the standard Fock space setting, there is no natural way to formulate such an a-priori regularity inequality, or even to introduce the notion of a multiplication operator.

This paper is organized as follows: In Section II we study the structure of $\mathcal{C}(\mathcal{Q})$ for finite n. It is actually simpler, as well as technically advantageous, to consider it as a subalgebra of the algebra $\mathcal{C}(\mathcal{K})$ generated by the identity, the configuration observables $Q_1, \dots, Q_n$ and their conjugate momenta $P_1 \dots, P_n$. Since they both turn out to be certain Clifford algebras, their structure has been worked out long ago with the representation theory of the orthogonal groups. Thus, this section contains no new result but simply introduces notation and prepares the way for what follows. Of particular use are an explicit spin– chain representation of the operator algebra $\mathcal{C}(\mathcal{K})$, and the Jordan-Wigner transform identifying it with the algebra generated by n "hard core bosons".

Section III concerns properties of the spaces $\mathcal{C}^p(\mathcal{K})$ and their norms. The main result here is an optimal uniform convexity inequality for $\mathcal{C}^p(\mathcal{K})$, $1 < p \leq 2$, which is joint work with Keith Ball. We need only a special case of this inequality here, and a proof is provided in an appendix for the reader's convenience.

Section IV introduces a convenient expression for P_t in terms of the conditional expectation $\pi_{\mathcal{Q}}$ of $\mathcal{C}(\mathcal{K})$ with respect to $\mathcal{C}(\mathcal{Q})$. The main result in this section is an inequality for conditional expectations which enables us to prove that

$$\sup\{\|P_t A\|_q \ : \ \|A\|_p = 1\} \qquad = \qquad \sup\{\|P_t A\|_q \ : \ A \geq 0 \text{ and } \|A\|_p = 1\} \quad .$$

Thus, to establish (1.6) in general, we need only consider positive A.

In Section V we establish the optimal fermion hypercontractivity bounds and the corresponding optimal fermion logarithmic Sobolev inequality. This is done in several steps. First, using results collected in Sections II and III we establish that (1.7) holds for $1 < p \leq 2$ and $q = 2$. At $t = 0$ this is an equality; differentiating it there yields the logarithmic Sobolev inequality (1.7). Gross showed that (1.6) would follow from (1.7), if it were true (as we show here), for all *self adjoint* A. His result rests on a deep inequality he established for positive operators. Since it is *not* in general true that $\|P_t A\|_q \leq \|P_t |A|\|_q$, as is trivially true in the commutatative case, Gross's result does not allow us to conclude (1.6) for *general* (i.e. non-self adjoint) A from (1.7). The results of the Section IV do, however, allow us to draw this conclusion.

Finally, in Section VI we show that the same hypercontractivity relation (1.6) holds for a mixed system of bosons and fermions.

Non-commutative probability theory has grown into a substantial branch of analysis with a number of physical applications. The mathematical theory is reviewed and developed in [Me85] and [Me86], while other sorts of physical applications, besides those discussed here, are treated in [Da76] and [HuPa] for example.

It is a pleasure to thank Leonard Gross for discussing his results and conjectures with us, and for encouraging us to take up the latter. We are indebted to Keith Ball for his collaboration on the subject of convexity inequalities that led to Theorem 1 [BCL], which is one of the key ingredients in the present work. Thanks are also due to G.-F. Dell'Antonio, A. Jaffe and A. Wightman for useful discussions.

II. FERMIONS AND THE CLIFFORD ALGEBRA

We begin by recalling for later use some well known facts about fermions.

The fundamental observables for a system of n fermion degrees of freedom are **configuration operators** $Q_1, Q_2, \ldots, Q_n$ together with their **conjugate momenta operators** $P_1, P_2, \ldots, P_n$ all acting as operators on a complex Hilbert space $\mathcal{H}$ and satisfying the **canonical anticommutation relations**:

$$Q_jQ_k + Q_kQ_j = 2\delta_{jk} \quad , \quad P_jP_k + P_kP_j = 2\delta_{jk} \quad , \quad P_jQ_k + Q_kP_j = 0 \quad . \tag{2.1}$$

We denote the complex algebra generated by the identity and the configuration observables by $\mathcal{C}(\mathcal{Q})$, and the complex algebra generated by the identity, the configuration observables and the momentum observables all together by $\mathcal{C}(\mathcal{K})$. $\mathcal{C}(\mathcal{Q})$ is the object of primary interest; but many aspects of its structure are most readily seen within the larger algebra $\mathcal{C}(\mathcal{K})$.

This algebra can be concretely represented as the algebra of observables for a spin–chain as follows. We define the matrices

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad , \quad U = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad , \quad Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad P = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \quad .$$

Let $\mathcal{H}$ denote the n-fold tensor product of $\mathbb{C}^2$ with itself:

$$\mathcal{H} = \underbrace{\mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2}_{n \text{ times}} \quad ,$$

and on $\mathcal{H}$ define the operators

$$Q_j = U \otimes \cdots U \otimes Q \otimes I \otimes \cdots \otimes I \quad , \quad P_j = U \otimes \cdots U \otimes P \otimes I \otimes \cdots \otimes I \tag{2.2}$$

where the Q and the P occur in the jth places.

The operators $Q_1, \ldots, Q_n, P_1 \ldots, P_n$ just defined are easily seen to satisfy the canonical anticommutaion relations. Of great use in studying the algebra $\mathcal{C}(\mathcal{K})$ that they generate is the fact that it is also the algebra generated by n "hard core boson" degrees of freedom. More explicitly, put

$$\widehat{Q}_j = I \otimes \cdots I \otimes Q \otimes I \otimes \cdots \otimes I \quad , \quad \widehat{P}_j = I \otimes \cdots I \otimes P \otimes I \otimes \cdots \otimes I \quad , \tag{2.3}$$

and call the algebra generated by the operators $\widehat{Q}_1, \ldots, \widehat{Q}_n, \widehat{P}_1 \ldots, \widehat{P}_n$ the **hard core boson algebra**, and denote it by $\widehat{\mathcal{C}}(\mathcal{K})$. To see that the two algebras coincide, put

$$U_j = I \otimes \cdots I \otimes U \otimes I \otimes \cdots \otimes I \qquad \text{and} \qquad V_k = \prod_{j=1}^{k-1} U_j \quad . \tag{2.4}$$

Then since $P_j Q_j = \widehat{P}_j \widehat{Q}_j = iU_j$, each V_k belongs to both the hard core boson algebra and $\mathcal{C}(\mathcal{K})$. Moreover

$$\widehat{Q}_k = V_k Q_k \quad \text{and} \quad \widehat{P}_k = V_k P_k \tag{2.5}$$

with the inverse relation given as well by

$$Q_k = V_k \widehat{Q}_k \quad \text{and} \quad P_k = V_k \widehat{P}_k \quad . \tag{2.6}$$

Thus, Q_k,P_k are in $\widehat{\mathcal{C}}(\mathcal{K})$, and $\widehat{Q}_k$,$\widehat{P}_k$ are in $\mathcal{C}(\mathcal{K})$.

What we call the hard core boson algebra was initially introduced by Jordan and Klein [JoKl] as a first attempt to implement the Pauli exclusion principle mathematically. The transformation of observables (2.6) was discovered by Jordan and Wigner [JoWi] and used by them to write the algebraic relations characterizing the algebra in the familiar covariant form (2.1); today it is known as the Jordan–Wigner transform. It has been used many times since – for example, in a solution of the two dimensional Ising model by Schultz, Mattis and Lieb, whose paper [SML] can be consulted for references to other applications. It is also the key to Brauer and Weyl's treatment [BrWe] of the Clifford algebra on which the exposition in the rest of this section is largely based. Following them, we now show that $\mathcal{C}(\mathcal{K})$ is the full matrix algebra on $\mathcal{H}$.

First, we introduce a basis in $\mathcal{H}$. Let

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad e_{-1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

be the standard basis of $\mathbb{C}^2$. For each $j = 1, \ldots, n$, let σ_j be either 1 or -1. Then the unit vectors

$$e_{\sigma_1, \ldots, \sigma_n} = e_{\sigma_1} \otimes \cdots \otimes e_{\sigma_n}$$

provide a natural orthonormal basis for $\mathcal{H}$.

Next introduce the fermion creation and annihilation operators $c_j^* = \frac{1}{2}(Q_j - iP_j)$ and $c_j = \frac{1}{2}(Q_j + iP_j)$ and their hard core boson analogs $\widehat{c}_j^* = \frac{1}{2}(\widehat{Q}_j - i\widehat{P}_j)$ and $\widehat{c}_j = \frac{1}{2}(\widehat{Q}_j + i\widehat{P}_j)$. Now put $L(\sigma_1, \ldots, \sigma_n) = \prod_{j=1}^n B_j$ where $B_j = \widehat{c}_j$ if $\sigma_j = 1$ and $B_j = \widehat{c}_j \widehat{c}_j^*$ if $\sigma_j = -1$. Then $L(\sigma_1, \ldots, \sigma_n) e_{\sigma_1, \ldots, \sigma_n} = \Omega$ where

$$\Omega = \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{n \text{ times}} = e_{-1,-1,\ldots,-1} \tag{2.7}$$

is a distinguished unit vector in $\mathcal{H}$ called the **ground state**. Also, $L(\sigma_1, \ldots, \sigma_n)$ annihilates all the other basis vectors. Moreover, $L^*(\sigma_1, \ldots, \sigma_n)\Omega = e_{\sigma_1, \ldots, \sigma_n}$. Thus,

$$L^*(\tau_1, \ldots, \tau_n) L(\sigma_1, \ldots, \sigma_n) e_{\sigma_1, \ldots, \sigma_n} = e_{\tau_1, \ldots, \tau_n} \quad ,$$

and this operator annihilates all other basis vectors. Manifestly this operator belongs to $\widehat{\mathcal{C}}(\mathcal{K})$, and hence to $\mathcal{C}(\mathcal{K})$ as well, and the 2^n operators of this kind form a basis for the full matrix algebra on $\mathcal{H}$.

This concrete description of $\mathcal{C}(\mathcal{K})$ in terms of spin–chain observables is the most useful for many purposes. Still, it is also useful to have a characterization of $\mathcal{C}(\mathcal{K})$ which is less dependent on coördinates; i.e. on the choice of fermi configuration observables $Q_1, \ldots, Q_n$.

Toward this end, consider the standard n–dimensional Hilbert space $\mathbb{C}^n$ equipped with its standard inner product $(\cdot,\cdot)$ and complex conjugation. Let $\mathcal{K}$ denote $\mathbb{C}^n$ considered as a real $2n$–dimensional Hilbert space equipped with the inner product

$$\langle x, y\rangle_{\mathcal{K}} = \Re(x, y) \quad .$$

Then complex conjugation on $\mathbb{C}^n$ induces an involutory orthogonal transformation J on $\mathcal{K}$. Let $\mathcal{Q}$ and $\mathcal{P}$ respectively denote the eigenspaces of J corresponding to the eigenvalues $+1$ and -1. The bilinear form on $\mathcal{K}$ given by $\Im(x, y)$ is symplectic so that $\mathcal{K}$ is naturally endowed with the structure possessed by the classical phase space of a system of n linear degrees of freedom. $\mathcal{Q}$ is called the **configuration space**, $\mathcal{P}$ the **momentum space**, and the complex conjugation J is usually identified with time reversal.

The Clifford algebra $\mathcal{C}(\mathcal{K})$ is characterized up to automorphism as the algebra with unit I such that:

(i) There is a linear imbedding $\mathcal{J} : \mathcal{K} \to \mathcal{C}(\mathcal{K})$, and $\mathcal{J}(\mathcal{K})$ generates $\mathcal{C}(\mathcal{K})$.

(ii) For all $x, y \in \mathcal{K}$,

$$\mathcal{J}(x)\mathcal{J}(y) + \mathcal{J}(y)\mathcal{J}(x) = \langle x, y\rangle_{\mathcal{K}} \quad . \tag{2.8}$$

To make contact with our previous concrete description, let $\{q_1, \ldots, q_n\}$ be an orthonormal basis of $\mathbb{C}^n$ consisting of purely real vectors. For each j, let $p_j = iq_j$. Evidently $\mathcal{Q}$ is spanned by $\{q_1, \ldots, q_n\}$ and $\mathcal{P}$ is spanned by $\{p_1, \ldots, p_n\}$. Any $x \in \mathcal{K}$ can then be written as $x = \sum_{j=1}^n \xi_j q_j + \sum_{j=1}^n \eta_j p_j$. Using the notation introduced above, put $\mathcal{J} : \mathcal{K} \to \mathcal{C}(\mathcal{K})$ by

$$\mathcal{J}(x) = \sum_{j=1}^n \xi_j Q_j + \sum_{j=1}^n \eta_j P_j \quad .$$

Let $\{x_1, \ldots, x_{2n}\}$ be any orthonormal basis of $\mathcal{C}(\mathcal{K})$. Then the monomials

$$\mathcal{J}(x_{\alpha_1})\mathcal{J}(X_{\alpha_2}) \cdots \mathcal{J}(x_{\alpha_k})$$

together with I form a basis for the algebra. It is easy to see that the product of any two such monomials is a third. Though the multiplication rule can be simply expressed in terms of certain contraction rules, its precise form is not useful to us here. What is useful to observe is that since the right side of (2.8) is invariant under orthogonal transformations of $\mathcal{K}$, the multiplication law of these monomials does not depend on the choice of the orthonormal basis. For this reason, any orthogonal transformation R of $\mathcal{K}$ induces an automorphism of $\mathcal{C}(\mathcal{K})$ which we shall also denote by R. Indeed, with $R : \mathcal{C}(\mathcal{K}) \to \mathcal{C}(\mathcal{K})$ defined by

$$R\big(\mathcal{J}(x_{\alpha_1})\cdots\mathcal{J}(x_{\alpha_k})\big) = \mathcal{J}(R(x_{\alpha_1}))\cdots\mathcal{J}(R(x_{\alpha_k})) \quad ,$$

R is evidently invertible and by the remarks made just above one easily sees that for all $A, B \in \mathcal{C}(\mathcal{K})$, $R(AB) = R(A)R(B)$ which is to say that R is an automorphism.

Finally, we remark that $\mathcal{C}(\mathcal{K})$ is a $*$-algebra; there is a unique conjugate linear involutory antiautomorphism $A \mapsto A^*$ which is the identity on $\mathcal{J}(\mathcal{K})$. It is given by

$$R\big(\mathcal{J}(x_{\alpha_1})\cdots\mathcal{J}(x_{\alpha_k})\big)^* = \mathcal{J}(x_{\alpha_k})\cdots\mathcal{J}(x_{\alpha_1}) \quad .$$

Of course, on $\mathcal{C}(\mathcal{K})$ regarded as a matrix algebra, this is just the usual adjoint. Evidently the automorphism R of $\mathcal{C}(\mathcal{K})$ induced by an orthogonal transformation R of $\mathcal{K}$ is a $*$-automorphism; i.e. $R(A^*) = \big(R(A)\big)^*$.

The facts that orthogonal transformations of $\mathcal{K}$ induce automorphisms of $\mathcal{C}(\mathcal{K})$, that $\mathcal{C}(\mathcal{K})$ is a full matrix algebra, and that all automorphisms of full matrix algebras are inner; i.e. of the form $A \mapsto SAS^{-1}$ for some nonsingular matrix S, are the basis of Brauer and Weyl's treatment of the spin representations of the orthogonal groups [BrWe]. We will use the fact that all automorphisms of $\mathcal{C}(\mathcal{K})$ are inner several times in what follows.

III. ANALYSIS ON THE CLIFFORD ALGEBRA

Let $\mathcal{A}$ be a von Neumann algebra of operators on some finite dimensional Hilbert space. By a trace on $\mathcal{A}$ we shall mean a linear functional τ which is positive in the sense that $\tau(A^*A) > 0$ for all non zero A in $\mathcal{A}$, and cyclic in the sense that $\tau(AB) = \tau(BA)$ for all A and B in $\mathcal{A}$. Such a functional is evidently unitarily invariant in the sense that whenever A and U belong to $\mathcal{A}$, and U is unitary, then $\tau(U^*AU) = \tau(A)$. Since $\mathcal{C}(\mathcal{K})$ is a full matrix algebra, it contains all unitaries. Hence any trace on $\mathcal{C}(\mathcal{K})$ must assign the same value to all rank one projections, and thus must be a scalar multiple of the standard trace Tr on the matrix algebra. Henceforth, τ shall denote this trace normalized by the condition that $\tau(I) = 1$ and Tr shall denote the standard unnormalized trace.

In the non-commutative integration theories of Dixmier [Di53] and Segal [Se53], the trace functional τ is the non-commutative analog of the functional that assigns to an integrable function its integral. When the Hilbert space is infinite dimensional, some further regularity properties are required of τ in order to obtain a useful analog. Since all of our estimations will be carried out in the finite dimensional setting, we shall not go into this here, but shall simply refer the reader to these original papers as well as the accounts in [Gr72] and [Ne74].

Norms on $\mathcal{C}(\mathcal{K})$ which are the non commutative analogs of the L^p norms can now be introduced; namely for $1 \le p < \infty$ we put

$$\|A\|_p = \left(\tau\left((A^*A)^{p/2}\right)\right)^{1/p} \quad , \tag{3.1}$$

and denote the operator norm of A by $\|A\|_\infty$.

$\mathcal{C}^p(\mathcal{K})$ shall denote $\mathcal{C}(\mathcal{K})$ equipped with the norm $\|\cdot\|_p$; evidently $\mathcal{C}^2(\mathcal{K})$ is the Hilbert space of $2^n \times 2^n$ matrices equipped with the Hilbert-Schmidt norm.

Consider the monomials

$$E_{[\alpha_1\dots,\alpha_j;\beta_1,\dots,\beta_k]} = Q_{\alpha_1}\cdots Q_{\alpha_j}P_{\beta_1}\cdots P_{\beta_k} \tag{3.2}$$

where $\alpha_1 > \cdots > \alpha_j$ and $\beta_1 > \cdots > \beta_k$ and $j + k > 0$. Evidently

$$E^*_{[\alpha_1\dots,\alpha_j;\beta_1,\dots,\beta_k]}E_{[\alpha_1\dots,\alpha_j;\beta_1,\dots,\beta_k]} = I$$

and thus $\|E_{[\alpha_1\dots,\alpha_j;\beta_1,\dots,\beta_k]}\|_p = 1$ for all p. Moreover

$$\tau(E_{[\alpha_1\dots,\alpha_j;\beta_1,\dots,\beta_k]}) = 0 \quad . \tag{3.3}$$

To see this, first consider the case in which $j + k$ is odd. The inversion $x \mapsto -x$ on $\mathcal{K}$ is orthogonal. Hence it induces an automorphism of $\mathcal{C}(\mathcal{K})$, and hence there is an invertible S in $\mathcal{C}(\mathcal{K})$ so that

$$-E_{[\alpha_1\dots,\alpha_j;\beta_1,\dots,\beta_k]} = SE_{[\alpha_1\dots,\alpha_j;\beta_1,\dots,\beta_k]}S^{-1} \quad .$$

Then, by using cyclicity of the trace we get the desired result. Next consider the case in which $j+k$ is even and, say, $j > 0$. Then write $E_{[\alpha_1 \ldots, \alpha_j; \beta_1, \ldots, \beta_k]} = Q_{\alpha_1} X$, and note that by (2.1), $Q_{\alpha_1} X = -X Q_{\alpha_1}$. Again the desired conclusion follows from cyclicity of the trace.

It is easy to see from this that

$$\tau(E^*_{[\alpha_1 \ldots, \alpha_j; \beta_1, \ldots, \beta_k]} E_{[\gamma_1 \ldots, \gamma_m; \delta_1, \ldots, \delta_n]}) = 0 \tag{3.4}$$

unless the two monomials coincide. Thus, together with the identity, the assemblage of such monomials forms an orthonormal basis for $\mathcal{C}^2(\mathcal{K})$. Finally observe that since $Qe_- = e_+$

$$\langle \Omega, E_{[\alpha_1 \ldots, \alpha_j]} \Omega \rangle = 0 \tag{3.5}$$

whenever $j \geq 1$, and, as indicated, $k = 0$. It now follows that, restricted to $\mathcal{C}(\mathcal{Q})$,

$$\tau(A) = \langle \Omega, A\Omega \rangle \tag{3.6}$$

for all A in $\mathcal{C}(\mathcal{K})$.

Formula (3.6) is very important for us. It permits us to calculate the "physically" relevant quantity $\langle \Omega, A\Omega \rangle$ in terms of the apparently mathematically simpler quantity $\tau(A)$.

Many familiar inequalities for L^p norms hold for the $\mathcal{C}^p$ norms as well [Di53]. This is true in particular of the Hölder inequality

$$\|AB\|_r \leq \|A\|_p \|B\|_q \qquad \frac{1}{r} = \frac{1}{p} + \frac{1}{q} \quad .$$

Certain optimal inequalities expressing the uniform convexity properties of the L^p norms also hold for the $\mathcal{C}^p$ norms, and this fact constitutes one cornerstone of our analysis.

The modulus of convexity δ_p of $\mathcal{C}^p(\mathcal{K})$ is defined by

$$\delta_p(\epsilon) = \inf\left\{1 - \frac{1}{2}\|A+B\|_p : \|A\|_p = \|B\|_p = 1 \ , \ \|A-B\|_p = \epsilon\right\} \tag{3.7}$$

for $0 < \epsilon < 2$. For $1 < p < \infty$, δ_p is always positive which means these norms are uniformly convex. Useful geometric information is contained in the rate at which $\delta_p(\epsilon)$ tends to zero with ϵ. It is known [TJ74] that for $2 \leq p < \infty$, $\delta_p(\epsilon) \sim \epsilon^p$, but that for $1 < p \leq 2$, $\delta_p(\epsilon) \sim \epsilon^2$.

An optimal expression of this fact is given by the following theorem which was proved jointly with Keith Ball [BCL]:

Theorem 1: (Optimal 2-uniform convexity for matrices). *For all $m \times m$ matrices A and B and all p for $1 \le p \le 2$,*

$$\left(\frac{Tr|A+B|^p + Tr|A-B|^p}{2}\right)^{2/p} \ge \left(Tr|A|^p\right)^{2/p} + (p-1)\left(Tr|B|^p\right)^{2/p} \quad . \tag{3.8}$$

For $1 < p < 2$, there is equality only when $B = 0$.

This result, which we interpret here as a statement about $\mathcal{C}^p(\mathcal{K})$, is proved in the appendix in the special case that both $A+B$ and $A-B$ are positive; this is the only case in which we shall use it here, and the proof is considerably simpler in this case. The full result is proved in [BCL], in which other geometric inequalities for trace norms are proved as well.

The theorem implies that

$$\delta_p(\epsilon) \ge \frac{(p-1)}{2}\left(\frac{\epsilon}{2}\right)^2 \qquad \text{for} \qquad 1 < p \le 2 \tag{3.9}$$

as one sees by considering $A = (C+D)/2$, $B = (C-D)/2$, $\|C\|_p = \|D\|_p = 1$ and $\|C-D\|_p = \epsilon$. It is easily seen that the constant $(p-1)/8$ cannot be improved.

We make our main application of this result in Section V. There we will also need to know that the norms on $\mathcal{C}^p(\mathcal{K})$ are continuously differentiable away from the origin for $1 < p < \infty$. This is known [Gr75], but a simple proof can be based on inequalities of the form $\delta_p(\epsilon) \ge K_p\epsilon^{r(p)}$ such as we have found above for $1 < p < 2$. This proof, moreover, gives the modulus of continuity of the derivative, and is sketched in the appendix as well. Again, these estimates are independent of the dimension and therefore apply to the case of infinitely many degrees of freedom.

IV. CONDITIONAL EXPECTATIONS AND THE FERMION OSCILLATOR SEMIGROUP

We are particularly concerned with the subalgebra $\mathcal{C}(\mathcal{Q})$ of $\mathcal{C}(\mathcal{K})$, and the conditional expectation [Di53][Um54] with respect to it shall play a basic role in our investigation. For any A in $\mathcal{C}(\mathcal{K})$, the **conditional expectation** $\pi_{\mathcal{Q}}(A)$ of A with respect to $\mathcal{C}(\mathcal{Q})$ is defined to be the unique element of $\mathcal{C}(\mathcal{Q})$ such that $\tau(B^*\pi_{\mathcal{Q}}(A)) = \tau(B^*A)$ for all B in $\mathcal{C}(\mathcal{Q})$. Otherwise said, $\pi_{\mathcal{Q}}$ is the orthogonal projection from $\mathcal{C}^2(\mathcal{K})$ onto $\mathcal{C}^2(\mathcal{Q})$. It is well known that the conditional expectation is positivity preserving; a familiar argument shows that $\pi_{\mathcal{Q}}(A^*A) \geq \pi_{\mathcal{Q}}(A)^*\pi_{\mathcal{Q}}(A)$.

We can use the conditional expectation to give a useful expression for the oscillator semigroup for fermion fields.

Let R_θ be the orthogonal transformation of $\mathcal{K}$ given by

$$R_\theta(q_j) = (\cos\theta)q_j + (\sin\theta)p_j \tag{4.1}$$

for each j. Of course R_θ gives the evolution at time θ on phase space $\mathcal{K}$ generated by the classical oscillator Hamiltonian $H(\mathbf{p},\mathbf{q}) = \sum_{j=1}^n \frac{1}{2}(p_j^2 + q_j^2)$. Let R_θ denote the automophism of $\mathcal{C}(\mathcal{K})$ generated by the orthogonal transformation R_θ as in the first section. For each $t \geq 0$, define $\theta(t) = \arccos(e^{-t})$ and define the operator P_t on $\mathcal{C}^2(\mathcal{K})$ by

$$P_t A = \pi_{\mathcal{Q}} \circ R_{\theta(t)} \circ \pi_{\mathcal{Q}}^* A \quad . \tag{4.2}$$

Note that $\pi_{\mathcal{Q}}^*$ is the **natural imbedding** of $\mathcal{C}(\mathcal{Q})$ into $\mathcal{C}(\mathcal{K})$, and regarded as such, it is a $*$-automorphism. Formula (4.2) is the analog of the familiar expression for the boson oscillator semigroup on $L^2(\mathcal{Q}, (2\pi)^{-n/2}e^{-q^2/2}\mathrm{d}^n q)$, i.e. the Mehler semigroup

$$P_t^{(boson)}A(\mathbf{q}) = \int_{\mathcal{Q}} A\Big(e^{-t}\mathbf{q} + (1-e^{-2t})^{1/2}\mathbf{p}\Big)(2\pi)^{-n/2}e^{-p^2/2}\mathrm{d}^n p \quad .$$

Note that since all of the operators on the right in (4.2) are positivity preserving, so is P_t. Also, since the first two operations on the right preserve the $\mathcal{C}^p$-norms, and since the conditional expectation is readily seen to be a contraction from $\mathcal{C}^p(\mathcal{K})$ to $\mathcal{C}^p(\mathcal{Q})$ for each p, it is readily seen that P_t possesses this property as well.

To obtain a more familiar expression for P_t, note that

$$R_{\theta(t)} \circ \pi_{\mathcal{Q}}^*\big(E_{[\alpha_1,\ldots,\alpha_k]}\big) = e^{-kt}E_{[\alpha_1,\ldots,\alpha_k]} + \big(\text{terms annihilated by } \pi_{\mathcal{Q}}\big) \quad .$$

Hence

$$P_t\big(E_{[\alpha_1,\ldots,\alpha_k]}\big) = e^{-kt}E_{[\alpha_1,\ldots,\alpha_k]}. \tag{4.3}$$

Evidently $\{P_t : t \geq 0\}$ is generated by H_0 where $H_0\big(E_{[\alpha_1,\ldots,\alpha_k]}\big) = kE_{[\alpha_1,\ldots,\alpha_k]}$. It is easy to see that under the unitary equivalence between $\mathcal{C}^2(\mathcal{Q})$ and fermion Fock space $\mathcal{F}$ described by Segal [Se56], H_0 is equivalent to the usual number operator, or in other words, oscillator Hamiltonian on $\mathcal{F}$.

Our primary goal is to prove optimal hypercontractivity bounds for P_t. That is, given $1 < p < q < \infty$ we want to show that for some finite t, P_t is a contraction from $\mathcal{C}^p(\mathcal{Q})$ to $\mathcal{C}^q(\mathcal{Q})$, and to find the smallest such t. Let

$$\|P_t\|_{p\to q} = \sup\{\|P_tA\|_q \ : \ \|A\|_p = 1 \ \} \quad . \tag{4.4}$$

As a first reduction, we shall show that the supremum on the right in (4.4) can be restricted to the positive operators A with $\|A\|_p = 1$. In the boson case this follows immediately from the fact that, in ordinary probability theory, the absolute value of a conditional expectation is no greater than the conditional expectation of the absolute value.

In general, matters concerning the absolute value in the non-commutative setting are more troublesome than in the commutative setting. An example is provided by the Araki-Yamagami inequality [ArYa] which, specialized to our context, asserts that the map $A \mapsto |A|$ is Lipschitz continuous on $\mathcal{C}^2(\mathcal{K})$ with constant $\sqrt{2}$ instead of the constant 1 which we would have in the commutative setting.

Thus while the conditional expectation in an operator algebra has many properties analogous to those of the conditional expectation in ordinary probability theory [Um54], it is not in general true that $|\pi_{\mathcal{Q}}(A)|$ will be a smaller operator than $\pi_{\mathcal{Q}}(|A|)$. The following theorem expresses a useful property in this direction which does hold, and after proving it we shall show by example that stronger properties do not hold. The theorem and its proof are easily extended to a more general von Neumann algebra setting by the methods in [Ru72].

Theorem 2: (A Schwarz inequality for conditional expectations). *For all A in*

$\mathcal{C}(\mathcal{K})$ *and all* p *with* $1 \le p \le \infty$,

$$\|\pi_{\mathcal{Q}}(A)\|_p \le \|\pi_{\mathcal{Q}}(|A|)\|_p^{1/2}\|\pi_{\mathcal{Q}}(|A^*|)\|_p^{1/2} \quad . \tag{4.5}$$

Remark: If we let $F(A)$ denote $\|\pi_{\mathcal{Q}}A\|_p$, then the same argument which we shall use to prove Theorem 2 also establishes that

$$F(A^*B) \le F(A^*A)^{1/2}F(B^*B)^{1/2} \quad . \tag{4.6}$$

In this form, the term "Schwarz inequality", by which we referred to (4.5), is more evidently appropriate. Moreover, inequalities of the type (4.6) are well known in matrix analysis for many familiar functions; for example when F(A) is the determinant of A or the spectral radius of A. Further examples can be found in [MeDS].

In [Li76] it is shown that for a function F that satisfies (4.6), and which is monotone increasing; i.e. satisfies $F(B) \ge F(A)$ for all $B \ge A \ge 0$, the following inequalities hold:

$$F(\sum_{j=1}^m A_j^*B_j) \le F(\sum_{j=1}^m (A_j^*A_j)^{1/2}F(\sum_{j=1}^m (B_j^*B_j)^{1/2}$$

and

$$F(\sum_{j=1}^m A_j) \le F(\sum_{j=1}^m (|A_j|)^{1/2}F(\sum_{j=1}^m (|A_j^*|)^{1/2} \quad .$$

In particular, these inequalities hold for $F(A) = \|\pi_{\mathcal{Q}}A\|_p$.

Specializing the last inequality to the case $m = 1$ then yields (4.5). In our present case however, the proof of the (4.6) is essentially the same as the direct proof of (4.5). Nonetheless, it should not be considered novel that by taking $|A^*|$ into consideration as well as $|A|$, we can obtain a suitable bound on $\|\pi_{\mathcal{Q}}A\|_p$.

Proof: Let $A = U|A|$ be the polar decomposition of A. Then $\|\pi_{\mathcal{Q}}(A)\|_p = \tau(CU|A|)$ for some C in $\mathcal{C}(\mathcal{Q})$ with $\|C\|_{p'} = 1$. Let $C = V|C|$ be the polar decomposition of C. Both V and $|C|$ belong to $\mathcal{C}(\mathcal{Q})$ as well. Thus

$$\|\pi_{\mathcal{Q}}(A)\|_p = \tau\big(CU|A|^{1/2}|A|^{1/2}\big) = \tau\big(|C|^{1/2}U|A|^{1/2}|A|^{1/2}V|C|^{1/2}\Big)$$

$$\leq \tau\big(|C|^{1/2}U|A|U^*|C|^{1/2}\big)^{1/2}\tau\big(|C|^{1/2}V^*|A|V|C|^{1/2}\big)^{1/2}$$

$$= \tau\big(|C|(U|A|U^*)\big)^{1/2}\tau\big((V|C|V^*)|A|\big)^{1/2} \leq \|\pi_{\mathcal{Q}}(U|A|U^*)\|_p^{1/2}\|\pi_{\mathcal{Q}}(|A|)\|_p^{1/2} \quad .$$

Finally, we note that $U|A|U^* = |A^*|$. □

Example Let A be the matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$. Then

$$|A| = \sqrt{2}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad |A^*| = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad .$$

Note that $|A^*|$ is in $\mathcal{C}(\mathcal{Q})$, but A and $|A|$ are not. One easily finds

$$\pi_{\mathcal{Q}}(A) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad , \quad \pi_{\mathcal{Q}}(|A|) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad , \quad \pi_{\mathcal{Q}}(|A^*|) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} .$$

$$\|\pi_{\mathcal{Q}}(A)\|_\infty = 1 \qquad , \qquad \|\pi_{\mathcal{Q}}(|A|)\|_\infty = \frac{1}{\sqrt{2}} \qquad \text{and} \qquad \|\pi_{\mathcal{Q}}(|A^*|)\|_\infty = \sqrt{2} \quad .$$

Theorem 3: (P_t has positive maximizers). *The norm of P_t from $\mathcal{C}^p(\mathcal{Q})$ to $\mathcal{C}^q(\mathcal{Q})$ is achieved on the positive operators; i.e.*

$$\|P_t\|_{p\to q} = \sup\{\|P_tA\|_q \ : \ A \geq 0 \quad \text{and} \quad \|A\|_p = 1 \ \} \quad . \tag{4.7}$$

Proof: Since R_θ and $\pi_{\mathcal{Q}}^*$ are both $*$-automorphisms, $|R_{\theta(t)} \circ \pi_{\mathcal{Q}}^* A| = R_{\theta(t)} \circ \pi_{\mathcal{Q}}^*|A|$ and $|(R_{\theta(t)} \circ \pi_{\mathcal{Q}}^* A)^*| = R_{\theta(t)} \circ \pi_{\mathcal{Q}}^*|A^*|$. Thus $\|P_tA\|_q \leq \|P_t|A|\|_q^{1/2}\|P_t|A^*|\|_q^{1/2}$, and of course $\||A^*|\|_p = \||A|\|_p = \|A\|_p$. □

V. HYPERCONTRACTIVITY FOR FERMIONS

Our main result is the following theorem which is established in this section.

Theorem 4: (Optimal fermion hypercontractivity). *For all $1 < p \leq q < \infty$, $\|P_t\|_{p\to q} = 1$ exactly when*

$$e^{-2t} \leq \frac{p-1}{q-1} \quad .$$

The heart of the matter is the following lemma:

Lemma: *For all* $1 < p \le 2$, $\|P_t\|_{p\to 2} = 1$ *exactly when* $e^{-2t} \le (p-1)$.

Proof: Fix a positive element A of $\mathcal{C}(\mathcal{Q})$. Pick a basis $\{q_1, \ldots, q_n\}$ of $\mathcal{Q}$, and let $\mathcal{C}(\mathcal{Q}_{(n-1)})$ denote the Clifford algebra associated with the span of the first $n-1$ of these basis elements. It is evident from the form of our standard basis of $\mathcal{C}(\mathcal{Q})$ that A can be uniquely decomposed as $A = B + CQ_n$ where B and C belong to $\mathcal{C}(\mathcal{Q}_{(n-1)})$. Then using the Jordan-Wigner transform we can write $A = B + CV_n\widehat{Q}_n$. Now write

$$\mathcal{H} = \mathcal{H}_{(n-1)} \otimes \mathbb{C}^2 \tag{5.1}$$

so that B, C and V_nC can be considered as operators on the first factor $\mathcal{H}_{(n-1)}$. Let $u_\pm = (e_+ \pm e_-)/\sqrt{2}$, so that $Qu_\pm = \pm u_\pm$. Then if v is any vector in $\mathcal{H}_{(n-1)}$,

$$\langle (v \otimes u_\pm), A(v \otimes u_\pm)\rangle_{\mathcal{H}} = \langle v, \big(B \pm CV_n\big)v\rangle_{\mathcal{H}_{(n-1)}} \quad . \tag{5.2}$$

We see from this that since $A \ge 0$, so are both $B + CV_n \ge 0$ and $B - CV_n \ge 0$. Now let Tr_1 and Tr_2 denote the partial traces over the first and second factors in (5.1), so that with Tr still denoting the full trace, we have $Tr = Tr_1Tr_2$. Now applying Theorem 1 in the special case which is proved in the appendix

$$\|A\|_p^2 = \left(\frac{1}{2^n}Tr|B + CV_n\widehat{Q}_n|^p\right)^{2/p}$$

$$= \left(\frac{1}{2^{(n-1)}}\right)^{2/p}\left(\frac{Tr_1|B + CV_n|^p + Tr_1|B - CV_n|^p}{2}\right)^{2/p}$$

$$\ge \left(\frac{1}{2^{(n-1)}}\right)^{2/p}\left((Tr_1|B|^p)^{2/p} + (p-1)(Tr_1|C|^p)^{2/p}\right)$$

since V_n is unitary.

Thus $\|A\|_p^2 \ge \|B\|_p^2 + (p-1)\|C\|_p^2$ where the norms on the right are all norms on $\mathcal{C}(\mathcal{Q}_{(n-1)})$.

Now we make the inductive assumption that the lemma has been established for $\mathcal{C}(\mathcal{Q}_{(n-1)})$. This is clearly the case when $n=1$. We then have

$$\|A\|_p^2 \geq \|P_t B\|_2^2 + (p-1)\|P_t C\|_2^2$$

where $e^{-2t} = (p-1)$. But clearly from (4.3) $\|P_t C Q_n\|_2^2 = (p-1)\|P_t C\|_2^2$ where the norms are once again norms on $\mathcal{C}(\mathcal{Q})$. Moreover by (3.4) $P_t B$ and $P_t C Q_n$ are orthogonal. Thus $\|P_t B\|_2^2 + (p-1)\|P_t C\|_2^2 = \|P_t(B + CQ_n)\|_2^2 = \|P_t A\|_2^2$. $\square$

Theorem 5: (Optimal fermion logarithmic Sobolev inequality). *For all $A \in \mathcal{C}(\mathcal{Q})$,*

$$\tau\big(|A|^2 \ln |A|^2\big) - \big(\|A\|_2^2 \ln \|A\|_2^2\big) \leq 2\langle A, H_0 A\rangle \quad . \tag{5.3}$$

Proof: By the lemma, $\|P_t A\|_2^2 \leq \|A\|_{(1+e^{-2t})}^2$ and there is equality at $t=0$. Both sides are continuously differentiable, and comparing derivatives at $t=0$ we obtain the result. Indeed,

$$\frac{\mathrm{d}}{\mathrm{d}p}\|A\|_p = \frac{1}{p}\|A\|_p^{1-p}\Big(\tau(|A|^p \ln |A|) - \|A\|_p^p \ln \|A\|_p\Big) \tag{5.4}$$

and of course

$$\frac{\mathrm{d}}{\mathrm{d}t}\|P_t A\|_2^2 = \frac{\mathrm{d}}{\mathrm{d}t}\langle A, P_{2t} A\rangle = -2\langle A, H_0 A\rangle \quad . \square$$

Gross refers to the quadratic form on the right side of (5.3) as the Clifford Dirichlet form since it shares many properties of Dirichlet forms in the ordinary commutative setting. An approach to the development of a theory of Dirichlet forms in the non-commutative setting can be found in [AlHK].

Proof of Theorem 4: By a deep result of Gross, when $A \geq 0$ and $1 < p < \infty$,

$$\langle A^{p/2}, H_0 A^{p/2}\rangle \leq \frac{(p/2)^2}{p-1}\langle A, H_0 A^{p-1}\rangle \quad .$$

Replacing A in (5.3) by $A^{p/2}$ and using the inequality just quoted we obtain, following Gross's ideas [Gr75],

$$\tau(A^p \ln A) - \|A\|_p^p \ln \|A\|_p \leq \frac{p/2}{p-1}\langle A^{p-1}, H_0 A\rangle \quad .$$

By combining this with (5.4) a differential inequality is obtained which implies that $\|P_t A\|_{q(t)}$ is a decreasing function of t when $q(t) = 1 + e^{2t}(p-1)$. This establishes the result for $A \geq 0$, and by Theorem 3 it is established in general. By (4.3), $P_t I = I$, and therefore $\|P_t\|_{p\to q}$ is always at least 1 for all p and q. That the inequality is best possible follows from a direct computation with one degree of freedom. To be precise, $\|P_t(I + Q_1)\|_q = \|I + e^{-t}Q_1\|_q$ is easily computed and compared with $\|I + Q_1\|_p$ [Gr72]. The first quantity is greater than the second if $e^{-2t} > (p-1)/(q-1)$. □

VI. HYPERCONTRACTIVITY FOR BOSONS AND FERMIONS TOGETHER

As a result of the present work and of earlier work on bosons, we know that for t given by $e^{-2t} = (p-1)/(q-1)$, both the fermion and the boson oscillator semigroups are contractive from the appropriate p–spaces to the appropriate q–spaces, and that this value of t is optimal for each case *separately.*

It is natural to expect that the same condition governs hypercontractivity in a situation in which we have bosons and fermions *together.* This is indeed the case, as we now show using Minkowski's inequality in an argument based on Segal's method for showing that the optimal conditions for hypercontractivity with m boson degrees of freedom are the same as for one degree of freedom.

Let $\mu(\mathrm{d}x) = (2\pi)^{-m/2}e^{-x^2/2}\mathrm{d}x$ be the unit Gauss measure on $\mathbb{R}^m$. Then in our mixed setting, with m boson degrees of freedom and n fermion degrees of freedom, the relevant p–space is

$$\mathcal{B}^p = L^p(\mathbb{R}^m, \mu) \otimes \mathcal{C}^p(\mathcal{Q}_{(n)}),$$

which may be regarded as consisting of $\mathcal{C}^p(\mathcal{Q}_{(n)})$ valued measurable functions $x \mapsto A(x)$ such that

$$|\!|\!|A|\!|\!|_p^p = \int_{\mathbb{R}^m} \|A(x)\|_p^p \mu(\mathrm{d}x)$$

is finite. This equation defines the norm on $\mathcal{B}^p$. For $p=2$, $\mathcal{B}^p$ is naturally isomorphic to the tensor product of the symmetric tensor algebra over $\mathbb{C}^m$ and the antisymmetric tensor algebra over $\mathbb{C}^n$ as shown by Segal. On the latter space we have the mixed oscillator semigroup generated by the sum of the boson and fermion number operators

$$\exp\left\{-t\left[\sum_{j=1}^{m} a_j^* a_j + \sum_{j=1}^{n} c_j^* c_j\right]\right\}.$$

Considered as operators on $\mathcal{B}^p$, the operators $\mathcal{P}_t$ which constitute this semigroup are given by

$$\mathcal{P}_t A(x) = \int_{R^M} M_t(x,x')P_t\big[A(x')\big]\mu(\mathrm{d}x') \tag{6.1}$$

where $M_t(x,x')$ is the Mehler kernel; i.e., the positive integral kernel for the boson oscillator semigroup $P_t^{(boson)}$ discussed in Section IV. Of course P_t denotes the fermion oscillator semigroup studied throughout this paper.

Now successively applying Minkowski's inequality, our theorem on optimal fermion hypercontractivity, and Nelson's theorem on optimal boson hypercontractivity, we have for $e^{-2t} \le (p-1)/(q-1)$:

$$\begin{aligned}
\|\mathcal{P}_t A\|_q^q &= \int_{\mathbb{R}^m} \left\| \int_{\mathbb{R}^m} M_t(x,x')P_t\big[A(x')\big]\mu(\mathrm{d}x')\right\|_q^q \mu(\mathrm{d}x) \\
&\le \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^m} M_t(x,x')\|P_t A(x')\|_q \mu(\mathrm{d}x')\right)^q \mu(\mathrm{d}x) \\
&\le \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^m} M_t(x,x')\|A(x')\|_p \mu(\mathrm{d}x')\right)^q \mu(\mathrm{d}x) \\
&\le \Big(\int_{\mathbb{R}^m} \|A(x)\|_p^p \mu(\mathrm{d}x)\Big)^{q/p} = \|A\|_p^q \quad .
\end{aligned}$$

APPENDIX

Proof of Theorem 1 when $A \pm B \ge 0$: Let Z and W be the $2m \times 2m$ matrices given by

$$Z = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \quad , \quad W = \begin{bmatrix} B & 0 \\ 0 & -B \end{bmatrix} \quad .$$

Our goal is to establish that for all r with $0 \le r \le 1$,

$$\left(\frac{Tr(A+rB)^p + Tr(A-rB)^p}{2}\right)^{2/p} \ge (Tr(A)^p)^{2/p} + r^2(p-1)(Tr|B|^p)^{2/p} \quad ,$$

or what is the same,

$$Tr(Z+rW)^{2/p} \ge (Tr(Z)^p)^{2/p} + r^2(p-1)(Tr|W|^p)^{2/p} \quad . \tag{A.1}$$

First, note that the null space of $Z + rW$ is exactly the null space of Z for $0 \le r < 1$. Thus by carrying out all of the following computations on the orthogonal complement of this fixed null space, we may freely assume that $Z + rW > 0$ for all $0 \le r < 1$. Next, both sides of (A.1) agree at $r = 0$, and the first derivatives in r of both sides vanish there as well. We define $\psi(r)$ to be $Tr(Z+rW)^p$. Then the second derivative in r of the left side of (A.1) satisfies

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}\big(\psi(r)\big)^{2/p} \ge \frac{2}{p}\psi(r)^{(2-p)/p}\frac{\mathrm{d}^2}{\mathrm{d}r^2}\psi(r) \quad .$$

The second derivative on the right side is just

$$2(p-1)(Tr|W|^p)^{2/p} \quad ,$$

and we are left with showing that

$$\frac{1}{p}\psi(r)^{(2-p)/p}\frac{\mathrm{d}^2}{\mathrm{d}r^2}\psi(r) \ge (p-1)(Tr|W|^p)^{2/p} \tag{A.2}$$

for all $0 < r < 1$. By redefining Z to be $Z + rW$, it suffices to establish (A.2) at $r = 0$.

Now $\frac{\mathrm{d}}{\mathrm{d}r}\psi(r) = p\big(Tr(Z+rW)^{(p-1)}W\big)$, since $A \pm B \ge 0$, $Z + rW \ge 0$ for small r, and we can use the integral representation

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

to conclude that

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}\psi(0) = pc_p\int_0^\infty t^{(p-1)}Tr\left[\frac{1}{t+Z}W\frac{1}{t+Z}W\right]dt \quad . \tag{A.3}$$

Consider the right side as a function, $f(Z)$, of Z for fixed W. It is easy to see that f is convex in Z. (Simply replace Z by $Z + tX$, with X self–adjoint, and then differentiate twice with respect to t; the positivity follows from the Schwarz inequality for traces.) Also, $f(UZU^*) = f(Z)$ provided U is unitary and U commutes with W. In a basis in which W is diagonal, we form the set $\mathcal{U}$ consisting of the 2^{2m} distinct diagonal unitary matrices, each with $+1$ or -1 in each diagonal entry. Each of these clearly commutes with W. Then

$$f(Z) = 2^{-2m} \sum_{U \in \mathcal{U}} f(UZU^*) \geq f(2^{-2m} \sum_{U \in \mathcal{U}} UZU^*) = f(Z_{\text{diag}}),$$

where Z_{diag} is the matrix that is diagonal in the basis diagonalizing W, and whose diagonal entries are those of Z in this basis.

Replacing Z by Z_{diag} in (A.3), the integration can be carried out, and we obtain

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}\psi(0) \geq p(p-1)\left(\sum_{j=1}^{2m} z_j^{(p-2)} w_j^2\right)$$

where z_j and w_j, respectively, denote the jth diagonal entries of Z and W in a basis diagonalizing W.

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

$$\psi(0) \geq \left(\sum_{j=1}^{2m} z_j^p\right) \quad .$$

To establish (A.2), we are only left with showing that

$$\left(\sum_{j=1}^{2m} z_j^p\right)^{(2-p)/p} \left(\sum_{j=1}^{2m} z_j^{(p-2)} w_j^2\right) \geq \left(\sum_{j=1}^{2m} |w_j|^p\right)^{2/p} \quad , \tag{A.4}$$

but this follows immediately from Hölder's inequality.

To complete the proof, observe that equality in (A.1) for $r = 1$ and $1 < p < 2$ implies equality in (A.4) for almost every r in $[0, 1]$. Here, recall that z_j in (A.4) really denotes the jth diagonal element of $Z + rW$; these are the numbers $z_j + rw_j$, where z_j denotes the jth diagonal element of Z.

Let us assume that $w_j \neq 0$ for some j. Then equality in Hölder's inequality (A.4) requires that the vector with positive components $z_j + rw_j$ be proportional to the vector with components $|w_j|$. Thus, for almost every r in $[0, 1]$ we require

$$z_j + rw_j = c(r)|w_j|$$

for some number $c(r)$ that depends on r but not on j. The left side above is a linear function, and thus $c(r) = a + rb$ for some numbers a and b. But then clearly $b = w_j/|w_j|$, and all non-zero eigenvalues of W would necessarily have the same sign. This is impossible since $TrW = 0$. □

We now give an application of the uniform convexity implied by this theorem to the differentiability of the $\mathcal{C}^p(\mathcal{K})$ norms. First we recall that for $2 \le p < \infty$, the modulus of convexity is given by an analog of an inequality of Clarkson for integrals which Dixmier [Di53] established for traces. Specializing to $\mathcal{C}^p(\mathcal{K})$, this inequality reads

$$\left\|\frac{A+B}{2}\right\|_p^p + \left\|\frac{A-B}{2}\right\|_p^p \le \frac{1}{2}\left[\|A\|_p^p + \|B\|_p^p\right] \qquad 2 \le p < \infty \quad ,$$

which implies that in this range (*cf.* (3.7))

$$\delta_p(\epsilon) \ge \frac{1}{p}\left(\frac{\epsilon}{2}\right)^p \quad . \tag{A.5}$$

For any non-zero A in $\mathcal{C}^p(\mathcal{K})$, define $\mathcal{D}(A)$ by

$$\mathcal{D}(A) = \|A\|_p^{(1-p)}|A|^{(p-1)}U^*$$

where $A = U|A|$ is the polar decomposition of A. Let p' be defined by $1/p + 1/p' = 1$. Then for $1 < p < \infty$, $\|\mathcal{D}(A)\|_{p'} = 1$ and $\tau(\mathcal{D}(A)A) = \|A\|_p$. Moreover, $\mathcal{D}(A)$ is the unique element of $\mathcal{C}^{p'}(\mathcal{K})$ with this property. We call the map $A \mapsto \mathcal{D}(A)$ the **gradient map** on $\mathcal{C}^p(\mathcal{K})$. The next theorem sharpens a result of Gross [Gr75].

Theorem 6: (Hölder continuity of the gradient map on $\mathcal{C}^p(\mathcal{K})$). *For all $1 < p < \infty$, the gradient map is norm continuous. Moreover,*

$$\|\mathcal{D}(A) - \mathcal{D}(B)\|_{p'} \le 2\left(p'\frac{\|A-B\|_p}{\|A+B\|_p}\right)^{(p-1)} \quad for \quad 1 < p \le 2 \quad , \tag{A.6}$$

and

$$\|\mathcal{D}(A) - \mathcal{D}(B)\|_{p'} \le 4(p-1)\left(\frac{\|A-B\|_p}{\|A+B\|_p}\right) \quad for \quad 2 \le p < \infty \quad . \tag{A.7}$$

Proof: First observe that for all $1 \leq p \leq \infty$

$$\|\mathcal{D}(A)+\mathcal{D}(B)\|_{p'}\|A+B\|_p \geq \Re\tau\big((\mathcal{D}(A)+\mathcal{D}(B))(A+B)\big)$$

$$= 2\big(\|A\|_p+\|B\|_p\big) - \Re\tau\big((\mathcal{D}(A)-\mathcal{D}(B))(A-B)\big)$$

$$\geq 2\big(\|A+B\|_p\big) - \|\mathcal{D}(A)-\mathcal{D}(B)\|_{p'}\|A-B\|_p \quad .$$

Thus,

$$1-\left\|\frac{\mathcal{D}(A)+\mathcal{D}(B)}{2}\right\|_{p'} \leq \left\|\frac{\mathcal{D}(A)-\mathcal{D}(B)}{2}\right\|_{p'}\left(\frac{\|A-B\|_p}{\|A+B\|_p}\right) \quad . \tag{A.8}$$

But, by (A.5) and (3.7), we have

$$1-\left\|\frac{\mathcal{D}(A)+\mathcal{D}(B)}{2}\right\|_{p'} \geq \frac{1}{p'}\left\|\frac{\mathcal{D}(A)-\mathcal{D}(B)}{2}\right\|_{p'}^{p'} \tag{A.9}$$

for $1 < p \leq 2$. By combining (A.9) and (A.8) we obtain (A.6). Similarly, by combining (3.9) with (A.8) we obtain (A.7). □

The continuity of the gradient map for $\mathcal{C}^p(\mathcal{K})$ has been established by Gross [Gr72], but his proof is more involved and does not yield an estimate of the modulus of continuity. It is now easy to establish continuous differentiability of the $\mathcal{C}^p(\mathcal{K})$ norms away from the origin since, with $h(t) = \|A+tB\|_p$, with A different from 0 and with t and s sufficiently small, we have

$$\Re\tau\big(\mathcal{D}(A+tB)B\big) \leq \frac{h(t+s)-h(t)}{s} \leq \Re\tau\big(\mathcal{D}(A+(t+s)B)B\big) \tag{A.10}$$

when s is positive. To see this, observe that $h(t) = \Re\tau\big(\mathcal{D}(A+tB)(A+tB)\big)$, and that $h(t+s) \geq \Re\tau\big(\mathcal{D}(A+tB)(A+(t+s)B)\big)$ by Hölder's inequality. By subtracting the expression for $h(t)$ from the estimate for $h(t+s)$ and dividing by s, we obtain the inequality on the left in (A.10). The inequality on the right is obtained in an analogous manner. When s is negative, the inequalities are clearly reversed. Letting s tend to zero, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|A+tB\|_p\bigg|_{t=0} = \Re\tau\big(\mathcal{D}(A)B\big) \quad .$$

REFERENCES

[AlHK] Albeverio, S., Høegh-Krohn, Dirichlet forms and Markov semigroups on C^*-algebras, Commun. Math. Phys., **56** (1977) 173-187.

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

[BCL] Ball, K., Carlen, E.A., Lieb, E.H.: preprint 1992.

[BrWe] Brauer, R., Weyl, H.: Spinors in n dimensions, Am. Jour. Math., **57** (1935) 425-449.

[CL91] Carlen, E.A., Loss, M.: Extremals of functionals with competing symmetries, Jour. Func. Analysis **88** (1991) 437-456

[Da76] Davies, E.B.: *Quantum Theory of Open Systems*, Academic Press, New York, 1976.

[Di53] Dixmier, J.: Formes linéaires sur un anneau d'opérateurs, Bull. Soc. Math. France **81** (1953) 222-245.

[Fe69] Federbush, P.: A partially alternate derivation of a result of Nelson, Jour. Math. Phys., **10** (1969) 50-52.

[Gr72] Gross, L.: Existence and uniqueness of physical ground states, Jour. Funct. Analysis, **10** (1972) 52-109.

[Gr75] Gross, L.: Hypercontractivity and logarithmic Sobolev inequalities for the Clifford-Dirichlet form, Duke Math. J., **43** (1975) 383-396.

[Gr89] Gross, L.: Logarithmic Sobolev inequalities for the heat kernel on a Lie group and a bibliography on logarithmic Sobolev inequalities and hypercontractivity, pp. 108-130 in *White Noise Analysis, Mathematics and Applications*, eds. Hida et al., World Scientific, Singapore 1990.

[HuPa] Hudson, R, Parthasarathy, K.R.: Quantum Itô's formula and stochastic evolutions, Commun. Math. Phys., **93** (1984) 301-323.

[JoKl] Jordan, P., Klein, O.: Züm Mehrkörperproblem der Quantentheorie, Zeits. für Phys., **45** (1927) 751-765.

[JoWi] Jordan, P., Wigner, E.P.: Über das Paulische Äquivalenzverbot, Zeits. für Phys., **47** (1928) 631-651.

[Li76] Lieb, E.H.: Inequalities for some operator and matrix functions, Adv. Math. **20**

(1976) 174-178

[Li90] Lieb, E.H.: Gaussian kernels have only Gaussian maximizers, Invent. Math. **102** (1990) 179-208

[Lin] Lindsay, M.: Gaussian hypercontractivity revisited, Jour. Funct. Analysis **92** (1990) 313-324.

[LiMe] Lindsay, M., Meyer, P.A.: preprint, 1991.

[MeDS] Merris, R.,Dias da Silva,J.A.: Generalized Schur functions, Jour. Lin. Algebra **35** (1975) 442-448.

[Me85] Meyer, P.A.: Eléments de probabilités quantiques, exposés I-V, pp.186-312 in Sem. de Prob. XX, Lecture notes in Math. **1204**, Springer, New York, 1985.

[Me86] Meyer, P.A.: Eléments de probabilités quantiques, exposés VI-VIII, pp. 27-80 in Sem. de Prob. XXI, Lecture notes in Math. **1247**, Springer, New York, 1986.

[Ne66] Nelson, E.: A quartic interaction in two dimensions, in *Mathematical Theory of Elementary Particles*, R. Goodman ans I. Segal eds., MIT Press, Cambridge Mass., 1966.

[Ne73] Nelson, E.: The free Markov field, Jour. Funct. Analysis, **12** (1973) 211-227.

[Ne74] Nelson, E.: Notes on non-commutative integration, Jour. Funct. Analysis, **15** (1974) 103-116.

[Nev] Neveu, J.: Sur l'esperance conditionelle par rapport à un mouvement Brownien, Ann. Inst. H. Poincaré Sect. B. (N.S.) **12** (1976) 105-109

[Ru72] Ruskai, M.B.: Inequalities for traces on Von Neumann algebras, Commun. Math. Phys., **26** (1972) 280-289.

[SML] Schultz, T.D., Mattis, D.C., Lieb, E.H.: Two dimensional Ising model as a soluble problem of many fermions, Rev. Mod. Phys. **36** (1964) 856-871.

[Se53] Segal, I.E.: A non-commutative extension of abstract integration, Annals of Math., **57** (1953) 401-457.

[Se56] Segal, I.E.: Tensor algebras over Hilbert spaces II, Annals of Math., **63** (1956) 160-175.

[Se70] Segal, I.E.: Construction of non-linear local quantum processes: I, Annal of Math., **92** (1970) 462-481.

[TJ74] Tomczak-Jaegermann, N.: The moduli of smoothness and convexity and Rademacher averages of trace classes $S_p(1 \leq p < \infty)$, Studia Mathematica **50** (1974) 163-182.

[Um54] Umegaki, H.: Conditional expectation in operator algebras I, Tohoku Math. J., **6** (1954) 177-181.