On Martingale Inequalities in Non-commutative Stochastic Analysis

Eric A. Carlen*

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332

and

Paul Krée†

Département de Mathématiques, Université Paris VI, 4, Place Jussieu,
75252 Paris Cedex 05, France

Received May 15, 1997; accepted May 19, 1998

We develop a non-commutative $L^p$ stochastic calculus for the Clifford stochastic integral, an $L^2$ theory of which has been developed by Barnett, Streater, and Wilde.

The main results are certain non-commutative $L^p$ inequalities relating Clifford integrals and their integrands. These results are applied to extend the domain of the Clifford integral from $L^2$ to $L^1$ integrands, and we give applications to optional stopping of Clifford martingales, proving an analog of a Theorem of Burkholder: The stopped Clifford process $F_T$ has zero expectation provided $E \sqrt{T} < \infty$. In proving these results, we establish a number of results relating the Clifford integral to the differential calculus in the Clifford algebra. In particular, we show that the Clifford integral is given by the divergence operator, and we prove an explicit martingale representation theorem. Both of these results correspond closely to basic results for stochastic analysis on Wiener space, thus furthering the analogy between the Clifford process and Brownian motion.

1. INTRODUCTION

The Clifford stochastic integral [BSW] is a non-commutative analog of the Ito integral. Since it will be useful to clarify this analogy before stating our results, we begin by recalling some familiar features of Brownian motion from an algebraic perspective.

Let $\mathcal{S}$ be the subspace of $L^2(\mathbb{R}_+)$ generated by the indicator functions $1_{[s,t]}$ of all bounded intervals $(s, t)$. The one dimensional Brownian motion $t \mapsto B(t)$ has the following basic property: For any interval $(s, t)$, $B(t) - B(s)$ is Gaussian and independent of the past.

* Work supported by U.S. National Science Foundation grant no. DMS 95-00840.
† UA 213, CNRS.

© 1998 Academic Press
has a centered normal distribution with variance \( t - s \), and for any finite, non-overlapping collection of intervals \( (s_j, t_j) \), the joint law of the \( B(t_j) - B(s_j) \) is simply the product of these marginals.

Now consider the symmetric algebra \( S(\mathcal{F}) \), which is the algebra of all polynomial functions \( A \) on \( \mathcal{F} \)

\[
A(z) = \sum \limits_{\alpha} c_{\alpha} (z \cdot h_1)^{\alpha_1} (z \cdot h_2)^{\alpha_2} \cdots (z \cdot h_n)^{\alpha_n}
\]

where \( \alpha = (\alpha_1, \alpha_2, ..., \alpha_n) \) is a usual multi-index, and only finitely many of the \( c_{\alpha} \) are non-zero, \( (z \cdot h) \) denotes the inner product in \( L^2(\mathbb{R}_+) \), and finally, the \( h_j \) are non-overlapping indicator functions.

The algebra \( S(\mathcal{F}) \) is of course the union of all of its finite dimensional subalgebras \( S(\mathcal{F}_F) \), generated by finite collections \( F \) of non-overlapping intervals. On each of these subalgebras \( S(\mathcal{F}_F) \), which is naturally isomorphic to the algebra of polynomials over a Euclidean space, there is an expectation functional \( E_F \) induced by the standard Gaussian integral on the Euclidean space. Moreover, when \( F_1 \subset F_2 \), \( S(\mathcal{F}_{F_1}) \subset S(\mathcal{F}_{F_2}) \), and for all \( A \in S(\mathcal{F}_{F_1}) \), \( E_{F_2} A = E_{F_1} A \). Because of this consistency, there exists a positive linear functional \( E \) on \( \mathcal{F} \) inducing all the \( E_F \) on the corresponding subalgebras, and the set-theoretic Wiener measure can be recovered from this functional in a well-known way.

Moreover, still at the level of the algebra \( \mathcal{F} \), one has another basic feature of the Brownian motion: its filtration. That is, let \( \mathcal{F}_t \) denote the subspace of \( \mathcal{F} \) generated by indicator functions of intervals that vanish on \( [t, \infty) \). Then the algebra \( \mathcal{F}(\mathcal{F}_t) \) corresponds to an algebra of functions on Wiener space that are measurable at time \( t \). Hence the basic probabilistic concept of adapted processes has an algebraic interpretation. This provides the basic framework necessary for the development of stochastic integration. To produce a robust approach to stochastic integration, one needs estimates controlling the size of the integrals in terms of the integrands. It is natural to look for bounds of this type involving the \( L^p \) norms associated to \( E \), and such bounds form an essential part of the theory of the Ito integral, and commutative stochastic integrals in general [Bu]. Our aim here is to extend these developments to the Clifford setting.

In the theory of Clifford stochastic integrals as it has been developed by Barnett, Streater and Wilde, the commutative, symmetric algebra \( S(\mathcal{F}) \) is replaced by the non-commutative Clifford algebra \( \mathfrak{C}(\mathcal{F}) \). The next section consists of a brief introduction to Clifford algebras, and all results and terminology that we use here are fully explained or referenced there. In particular, the book of Plymel and Robinson [PR] is an excellent source for the background. Here in this introduction, we shall rely somewhat on analogy with the familiar case of Brownian motion to make ourselves understood, in order to avoid repeating what is explained in the references.
In this analogy, the expectation functional $E$ is replaced by the Segal state $\mathcal{E}(\mathcal{F})$, still denoted by $E$. Noncommutative $L^p$ norms can then be defined in terms of $E$, and again there is the natural filtration of subalgebras $\mathcal{E}(\mathcal{F}_t)$, and hence a natural notion of adapted processes. Moreover, there is a distinguished process $t \mapsto Q(t) \in \mathcal{E}(\mathcal{F}_t)$ called the Clifford Process, and at the $L^2$ level, a close analog of the Ito theory of stochastic integrals with respect to it.

One goal of this paper is to establish $L^p$ estimates for these Clifford stochastic integrals. To do this it is natural to work in a somewhat different setting than that of [BSW]. We work here in a larger Clifford algebra $\mathcal{E}(Q, P)$ in which the underlying space $\mathcal{F}$ is doubled to the sum of two copies $\mathcal{F}_Q \oplus \mathcal{F}_P$. The advantage of working in this larger algebra—at least when one seeks $L^p$ estimates—was demonstrated in [CL], where sharp hypercontractivity for fermions was proved. In fact, a significant part of our methodology here is based on the methodology of that paper.

Another significant part is built on the differential calculus in Clifford algebras. In fact, as is well known now, the differential calculus on Wiener space naturally enters such basic issues as martingale representation, at least if one requires an explicit form for the representation. We show that the same is true in the Clifford setting.

Before stating our results, we establish some notation and terminology. First, there is of course a natural filtration of subalgebras $\mathcal{E}(Q, P)_t$ just as in $\mathcal{E}(\mathcal{F})$. Then there are two Clifford processes $t \mapsto Q(t)$ and $t \mapsto P(t)$, which may be thought of as two independent, but non-commuting, copies of the same process.

A function $G: \mathbb{R}_+ \to \mathcal{E}(\mathcal{F})$ of the form

$$G(t) = \sum_{j=1}^{N} G_j(h_j(t)),$$  \hspace{1cm} (1.2)

where each $G_j$ is in $\mathcal{E}(Q, P)_{h_j}$, is called a simple adapted Clifford process. Its Clifford stochastic integral with respect to $dQ$ is defined by

$$\int G(t) \, dQ(t) = \sum_{j=1}^{N} G_j(Q(t_j) - Q(s_j)).$$  \hspace{1cm} (1.3)

The only difference with the definition in [BSW] is that we are working with a larger filtration. Naturally, one makes the analogous definition for the $dP$ integral. Hence we shall be working with a pair $(G_Q, G_P)$ of integrands and considering the quantity

$$\int G_Q \, dQ + \int G_P \, dP.$$
Work done so far in this setting has been in an $L^2$ (with respect to $E$) setting. Indeed, the basis of the Barnett, Streater and Wilde theory is the isometry

$$E \left[ \left| G(t) \, dQ(t) \right|^2 \right] = E \left( \sum_{j=1}^{N} G^*_j G_j(t_j - s_j) \right).$$

(1.4)

Of course, the same isometry holds for the $dP$ integrals, and moreover given two integrands $G_Q$ and $G_P$, the corresponding integrals $\int G_Q \, dQ$ and $\int G_P \, dP$ are orthogonal in $L^2$:

$$E \left( \left( \int G_Q \, dQ \right)^* \left( \int G_P \, dP \right) \right) = 0.$$

(1.5)

(The *, which is an operator adjoint, is explained in the next section.)

Here we extend many results to $L^p$ for $1 \leq p < \infty$. To do this, it is natural to introduce the square function $\langle G, G \rangle$ of the integrand $G$:

$$\langle G, G \rangle = \sum_{j} G^*_j G_j(t_j - s_j).$$

(1.6)

Since we shall often deal with two-component integrals $\int G_Q \, dQ + \int G_P \, dP$, we define

$$\langle \mathbf{G}, \mathbf{G} \rangle = \langle G_Q, G_Q \rangle + \langle G_P, G_P \rangle \quad \text{for} \quad \mathbf{G} = (G_Q, G_P).$$

(1.7)

Experience with Brownian motion suggests that we should try to control the $L^p$ norm of $\int G(t) \, dQ(t)$ in terms of the $L^p$ norm of $\langle G, G \rangle^{1/2}$.

Here an important difference between the commutative and non-commutative theories emerges: The quantities $\| \langle G, G \rangle^{1/2} \|_p$ and $\| \langle G^*, G^* \rangle^{1/2} \|_p$, where $G^*(t)$ denotes the adjoint of $G(t)$, are not comparable. For any given $\varepsilon > 0$, it is possible to find $G$ such that either one of the two is less than $\varepsilon$, and the other is bigger than $1/\varepsilon$. Thus it is rather pleasing that the minimum of these two quantities controls Clifford integration for $1 \leq p < 2$.

**Theorem 1.** For all $0 < p \leq 2$, there is a finite constant $K_p$ such that for all simple Clifford integrals $\int G_Q \, dQ + \int G_P \, dP$,

$$\| \int G_Q \, dQ \|_p + \| \int G_P \, dP \|_p \leq K_p \text{min} \left\{ \| \langle \mathbf{G}, \mathbf{G} \rangle^{1/2} \|_p, \| \langle \mathbf{G}^*, \mathbf{G}^* \rangle^{1/2} \|_p \right\}$$

(1.8)

with $K_p = (2/(2 - p))^{1/p} ((2 - p)/p)^{1/2}$. 

478 CARLEN AND KRÉE
Here of course $G^* = (G_Q^*, G_P^*)$.

This theorem has been extended by Pisier and Xu [PX], who have proved the corresponding converse bound. We shall explain this shortly.

First we give an application of this to optionally stopped stochastic integrals, and obtain an analog of a theorem of Burkholder for optionally stopped Ito integrals. First, we need a definition [BL].

This requires some care. First, as we shall explain in the next section, $\mathcal{C}(Q, P)$ may be faithfully embedded in a von Neumann algebra of operators on some Hilbert space $\mathcal{H}$. This Hilbert space can be built out of our $\mathcal{C}(Q, P)$ itself using the GNS construction, but the particular construction does not matter at this point. Consider any positive operator $T$ on $\mathcal{H}$, possibly unbounded. By the spectral theorem, there are orthogonal projections $\Pi_t$; i.e., the orthogonal projections onto the spectral subspaces of $T$ on which $T > t$, such that

$$T = \int_0^\infty \Pi_t dt$$

(1.9)

weakly on the domain of $T$.

A non-negative operator $T$ is said to be a stopping time [BL, Hu] if each of its spectral projections $\Pi_t$ as in (1.9) satisfy

$$\Pi_t \in L^2(\mathcal{C}(Q, P)_t) \quad \text{for all } t.$$  

(1.10)

That is, $\Pi_t$ is in the $L^2$ closure of the past at time $t$, $\mathcal{C}(Q, P)_t$.

Now, consider a Clifford integral $\int G_Q dQ + G_P dP$ where $(G_Q, G_P)$ satisfies

$$\min\left\{ \| \langle G, G \rangle \|, \| \langle G^*, G^* \rangle \| \right\} < \infty.$$  

(1.11)

Let $T$ be a stopping time. We define $\int_0^T G_Q dQ + \int_0^T G_P dP$, the Clifford integral stopped at $T$, by (with $\gamma$ the principle automorphism)

$$\int_0^T G_Q dQ + \int_0^T G_P dP = \int \Pi_t G_Q(t) \gamma(\Pi_t) \, dQ(t) + \int \Pi_t G_P(t) \gamma(\Pi_t) \, dP(t).$$

(1.12)

The motivation for this is that if we were stopping an Ito integral, the analog of $\Pi_t$ is $1_{\{T > t\}}$ and we would have

$$\int_0^T G \, dB = \int 1_{\{T > t\}} G \, dB = \int 1_{\{T > t\}} G 1_{\{T > t\}} \, dB.$$

The analogs of the second and third term are not equal in our setting. Barnett and Lyons [BL] take the first choice as the basis of their definition.
of the stopped process, and here we have taken the second, modified slightly so that if each \( M(t) \) is self adjoint, \( M_T \) will be as well.

An interesting case is that in which \( G_Q \) or \( G_P \) is identically 1. Then (1.11) is not satisfied, but one can still try to use (1.12) to define a stopped Clifford process:

\[
Q(T) = \int \Pi_t \gamma(\Pi_t) \, dQ(t) \quad \text{and} \quad P(T) = \int \Pi_t \gamma(\Pi_t) \, dP(t). \tag{1.13}
\]

The question then is: For which \( T \) does this make sense? Moreover, it is an easy consequence of the Ito isometry property that the Brownian motion stopped at an integrable stopping time has zero expectation. This property is used very often, and so it is important to know the optimal conditions under which it is true that a stopped Brownian motion will have zero expectation. Burkholder [Bu] proved that this is the case when the square root of the stopping time is integrable, and that this condition is sharp; i.e., integrability of any lower power of the stopping time does not suffice. We obtain the corresponding result here. (The square root of \( T \) in our case is defined using the spectral theorem.)

**Theorem 2.** Let \( T \) be a stopping time, and consider \((G_Q, G_P)\) with (1.11) satisfied. Then

\[
\left\| \int_0^T G_Q \, dQ + \int_0^T G_P \, dP \right\|_1 \leq \min \{ \| \langle G, G \rangle^{1/2} \|_1, \| \langle G^*, G^* \rangle^{1/2} \|_1 \}. \tag{1.14}
\]

Moreover, in case \( E(\sqrt{T}) < \infty \), the stochastic integrals in (1.13) are well defined and

\[
E_Q(T) = 0 \quad \text{and} \quad E_P(T) = 0. \tag{1.15}
\]

**Proof.** Note that \((\Pi_t G(t) \gamma \Pi_t)^* (\Pi_t G(t) \gamma \Pi_t) \leq G^*(t) G(t)\) and \((\Pi_t G(t) \gamma \Pi_t \Pi_t \gamma \Pi_t)^* \leq G(t) G^*(t)\). Hence Theorem 1 gives (1.14). Next, \((\Pi_t \gamma \Pi_t)^* (\Pi_t \gamma \Pi_t) \leq \gamma \Pi_t\), so that

\[
\int_0^\infty (\Pi_t \gamma \Pi_t)^* (\Pi_t \gamma \Pi_t) \, dt \leq \gamma \int_0^\infty \Pi_t \, dt = \gamma(T).
\]

Hence Theorem 1 is exactly what we need to ensure that the integrals in (1.12) and (1.13) are well defined under the stated hypothesis on \( T \). Finally, Theorem 1 and an obvious approximation argument now yields the validity of (1.15).
Before proceeding further we explain Pisier and Xu’s extension of what we proved in Theorem 1. First, observe that if $G = A + B$ is any decomposition of $G$, then by the Minkowski inequality Theorem 1 implies that

$$\left\| \int G_0 \, dQ + \int G_1 \, dP \right\|_p \leq K_p \left\{ \left\| \langle A, A \rangle^{1/2} \right\|_p + \left\| \langle B^*, B^* \rangle^{1/2} \right\|_p \right\}$$

and hence

$$\left\| \int G_0 \, dQ + \int G_1 \, dP \right\|_p \leq K_p \text{min} \left\{ \left\| \langle A, A \rangle^{1/2} \right\|_p + \left\| \langle B^*, B^* \rangle^{1/2} \right\|_p \right\}$$

where the minimum is over all such decompositions. The right hand side defines a norm, the $H^p$ norm, on processes $G$:

$$\|G\|_{H^p} = \text{min} \left\{ \left\| \langle A, A \rangle^{1/2} \right\|_p + \left\| \langle B^*, B^* \rangle^{1/2} \right\|_p \right\}$$

with the infimum as above.

Pisier and Xu [PX] showed, motivated by an earlier preprint version of this paper, that in this form, the bound of Theorem 1 has a converse: For all $p$ with $1 < p < 2$, there is a constant $K_p$ such that

$$\left\| \int G_0 \, dQ + \int G_1 \, dP \right\|_p \geq K_p \|G\|_{H^p}.$$ 

The work of Pisier and Xu also gives another proof of Theorem 1. We have kept our original proof in this paper because it is entirely different, and its more “stochastic process” oriented methodology is of independent interest. Our original proof of Theorem 1 is provided in Section 4 starting from an explicit martingale representation formula, which require the differential calculus on the Clifford algebra $\mathcal{C}(\mathcal{S})$, not only for its proof, but also for its formulation. Therefore Section 3 contains several results relating the Clifford integral and the differential calculus on $\mathcal{C}(\mathcal{S})$, the basic features of which are introduced in the next section.

In particular, there is a natural notion of a gradient $V$ in $\mathcal{C}(Q, P)$, and then by duality, of a divergence $\delta$ on $\mathcal{C}(Q, P) \otimes (\mathcal{S}_Q \otimes \mathcal{S}_P)$. Of course there are natural imbeddings of $\mathcal{C}(Q, P) \otimes \mathcal{S}_Q$ and $\mathcal{C}(Q, P) \otimes \mathcal{S}_P$ into $\mathcal{C}(Q, P) \otimes (\mathcal{S}_Q \otimes \mathcal{S}_P)$, and hence there are restrictions $\delta_Q$ and $\delta_P$ of $\delta$ to these spaces.

Now, note that one can view a simple adapted process as in (1.2) as belonging to one of these spaces if we write it as $\sum_{i=1}^n G_i \otimes 1_{(a_i, b_i)}$. It turns out that such processes are in the domain of the divergence, and the divergence is essentially the Clifford integral. We say “essentially” because the exact statement involves the principle automorphism $\gamma$ of the Clifford
algebra, which is explained in the next section. We also define $A_2$ to be the natural Hilbert space completion of the simple past adapted processes. Then the precise result is:

**Theorem 3.** For all $p$ with $1 \leq p \leq 2$, all past adapted processes $G(\cdot)$ with

$$\min\{\|G, G\|^p, \|G^* G^*\|^p\} < \infty$$

are in the domains of $\delta_Q$ and $\delta_P$, with

$$\int G \, dQ = \delta_Q(\gamma G(\cdot)) \quad \text{and} \quad \int G \, dP = \delta_P(\gamma G(\cdot)).$$

In particular, $\delta_Q: A_2 \to L^2$ and $\delta_P: A_2 \to L^2$ are isometries.

This theorem is proved in Section 3 immediately following Proposition 3.1 which is an $L^2$ version.

Next we use this to produce an explicit martingale representation theorem. The martingale terminology will be explained later; the result shows that all elements of the completion of $\mathcal{E}(Q, P)$ in the $L^1$ norms can be written as Clifford stochastic integrals, and explicitly identifies the integrands. This involves the partial gradients $V_Q$ and $V_P$ corresponding to the subspaces $S_Q$ and $S_P$ of $S_Q \oplus S_P$ out of which $\mathcal{E}(Q, P)$ is built. Again, all of this is explained in the next section.

**Theorem 4.** For any $p$ with $1 < p \leq 2$, any element $A$ of $L^p(\mathcal{E}(Q, P))$ can be uniquely represented as a Clifford integral

$$A = E(A) + \int G_Q \, dQ + \int G_P \, dP$$

where $G_Q(\cdot)$ and $G_P(\cdot)$ belong to $A_2$. Moreover, these unique processes satisfy

$$\|G\|_{H^p} \leq k_p \|A\|_p$$

and are given by

$$G_Q(\cdot) = \gamma(\text{Ad}(V_Q A)) = \text{Ad}(\gamma(V_Q A))$$

$$G_P(\cdot) = \gamma(\text{Ad}(V_P A)) = \text{Ad}(\gamma(V_P A))$$

where $\text{Ad}: L^2(\mathbb{R}_+, L^2(\mathcal{E}(Q, P))) \to L^2(\mathcal{E}(Q, P))$ denotes the orthogonal projection onto the adapted square integrable processes.
This theorem is proved in Section 3 immediately following Proposition 3.2 which is an \( L^2 \) version.

Next we apply this in the case that \( A \) is the absolute square of a Clifford integral. In this case it is possible to evaluate all of the derivatives and projections explicitly. The result shows how to express one function of a stochastic integral—the absolute square—as a stochastic integral, and hence is a special case of a non-commutative Ito formula. We shall see how to extend this to other more interesting functions in Section 4, where this is done in the course of proving Theorem 1.

**Theorem 5.** Let \( G \) and \( H \) belong to \( \mathcal{A}_2 \). For all \( t \geq 0 \), let

\[
M(t) = M_Q(t) + M_P(t)
\]

where

\[
M_Q(t) = \int_{(0,t)} G_Q dQ \quad \text{and} \quad M_P(t) = \int_{(0,t)} G_P dP.
\]

Then,

\[
|M|^2 = \left( |G_Q|^2 + |G_P|^2 \right) dt + \left( (M_Q(t))^* G_Q(t) + (G_Q(t))^* M_Q(t) + M_P^* G_Q + \gamma(G_Q^* M_P) \right) dQ(t)
\]

\[
+ \left( (M_P(t))^* G_P(t) + (G_P(t))^* M_P(t) + M_Q^* G_P + \gamma(G_P^* M_Q) \right) dP(t)
\]

and each of the Clifford integrals converges in \( L^1 \).

Theorem 1 allows the integration of adapted integrands under minimal conditions. However, one would expect strong \( L^p \) regularity for the integrals given strong \( L^p \) regularity for the integrands. Our next result is an analog of the Zakai inequalities for Ito integrals.

**Theorem 6.** For all simple Clifford integrals \( G \),

\[
\left\| G dQ \right\|_p \leq (p - 1)^{1/2} \left( \int |G(s)|_p^2 ds \right)^{1/2} \quad \text{for} \quad p \geq 2 \quad (1.21)
\]

and

\[
\left\| G dQ \right\|_p \geq (p - 1)^{1/2} \left( \int |G(s)|_p^2 ds \right)^{1/2} \quad \text{for} \quad p \leq 2 \quad (1.22)
\]
The same results hold for Clifford integrals $\int G \, dP$, and the constant $(p-1)^{1/2}$ is best possible.

2. A BRIEF INTRODUCTION TO CLIFFORD CALCULUS

2. Basic definitions

For the reader’s convenience, this section, the lengthiest in the paper, has been divided into subsections.

A quadratic space is any pair $(X, \phi)$ where $\phi$ is a quadratic form on the real vector space $X$. Let $\phi(x, y)$ denote the associated bilinear form obtained by polarization. The Clifford algebra of $(X, \phi)$ is an algebra with unit into which $X$ is imbedded so that the product in the algebra is determined by $\phi$ in a particular way.

More precisely, a map $h$ from $(X, \phi)$ to the real unital algebra $A$ is called a Clifford map in case $h$ is linear and $h(x)^2 = \phi(x) 1$ for all $x \in X$. The Clifford algebra of $(X, \phi)$ is a pair $(J, \mathcal{C}(X))$ where $J : X \rightarrow \mathcal{C}(X)$ is Clifford and the following universality property holds: For any algebra $A$ and any Clifford map $h : X \rightarrow A$, there is a unique homomorphism $\tilde{h} : \mathcal{C}(X) \rightarrow A$ such that $h = \tilde{h} \circ J$.

The existence of such a pair is a standard construction in the tensor algebra over $X$. And clearly, because of the universal property, the Clifford algebra is unique up to isomorphism. Moreover, if $(Y, \psi)$ is another quadratic space, and $L : X \rightarrow Y$ is a linear map such that $\psi(Lx) = \phi(x)$ for all $x$ in $X$, then $J \circ L : X \rightarrow \mathcal{C}(Y)$ is a Clifford map, and thus, by the universality property, $L$ extends to a homomorphism $\mathcal{C}(L) : \mathcal{C}(X) \rightarrow \mathcal{C}(Y)$.

In the case $(Y, \psi) = (X, \phi)$, this construction yields automorphisms of $\mathcal{C}(X)$. A case of particular importance is the principle automorphism $\gamma$ which is induced by $x \mapsto -x$. Evidently, $\gamma^2 = 1$, and $\gamma$ induces a grading

$$\mathcal{C}(X) = \mathcal{C}_{\text{even}}(X) \oplus \mathcal{C}_{\text{odd}}(X)$$

where $\gamma(A) = A$ for all $A$ in $\mathcal{C}_{\text{even}}(X)$, and $\gamma(A) = -A$ for all $A$ in $\mathcal{C}_{\text{odd}}(X)$. Moreover,

$$\mathcal{J}(x).A = \gamma(A) \mathcal{J}(x).$$

(2.1)

One can also use $\mathcal{J}$ to obtain a basis for $\mathcal{C}(X)$ from a basis for $X$. Let $\{e_j \mid j \in \mathbb{N}\}$ be any ordered basis for $X$, which we then denote $\epsilon_j = \mathcal{J}(e_j)$ are a set of generators for $\mathcal{C}(X)$. To work with these conveniently, introduce fermionic multi-indices $\alpha$ where $\alpha$ is the ordered subset of indices

$$\alpha = \{x_1', x_2', \ldots, x_n'\} \subset \mathbb{N}$$

(2.2)
for some integer \( n \) with \( x'_1 < x'_2 < \cdots < x'_n \). We then define \( x_j = 1 \) in case \( j \in \pi \), and \( x_j = 0 \) otherwise, and we let \( |\pi| \) denote the cardinality of \( \pi \). Finally, we shall write 0 for \( \pi \) when \( \pi \) is empty.

Now for any fermionic multi-index \( \pi \), put

\[
\tilde{\xi}^\pi = \xi_{\pi_{i_1}} \xi_{\pi_{i_2}} \cdots \xi_{\pi_{i_m}}
\]

and in the special case that \( \pi = 0 \), put \( \tilde{\xi}^0 = 1 \). It is clear that the \( \tilde{\xi}^\pi \) form a spanning set; we shall soon see that they are a basis.

So far, we have not been concerned with whether or not the space \( X \) is real or complex. However, if \( X \) is real, one can form the complexification \( X^c = X + iX \), in which \( X \) is naturally imbedded as \( X \oplus i0 \). Then let \( \phi^c \) denote the complexification \( \phi \) on \( X^c = X + iX \):

\[
\phi(x + iy) = \phi(x) - \phi(y) + 2i\phi(x, y).
\]

Then the Clifford algebra of \( (X^c, \phi^c) \) is given by the complexification of \( \mathcal{C} : X^c \to \mathcal{C}(X)^c = \mathcal{C}(X) + i\mathcal{C}(X) \). Hence, as for symmetric algebras, \( \mathcal{C}(X)^c \) is isomorphic to \( \mathcal{C}(X^c) \), and we shall henceforth write \( \mathcal{C}(X) \) to denote this complex Clifford algebra.

The passage to the complex case is important since it permits the addition of structure making \( \mathcal{C}(X) \) a \( \ast \)-algebra: As shown in [PR], the natural conjugation \( x + iy \mapsto x - iy \) on \( X^c \) extends to a unique sesquilinear involution \( \ast \) on \( \mathcal{C}(X) \) satisfying:

(i) \( (\mathcal{C}(x))^\ast = (\mathcal{C}(x)) \) for all \( x \) in \( X \).

(ii) \( (AB)^\ast = (A^\ast B^\ast) \) for all \( A \) and \( B \) in \( \mathcal{C}(X) \).

Using these properties, it is easy to see that \( \ast \) commutes with \( \gamma \).

2.2. The Normalized Trace

To begin with non commutative probability theory, all we need now is an expectation functional on the algebra \( \mathcal{C}(X) \), which we think of as a non-commutative analog of the ring of measurable functions on a probability space. There is linear functional \( E \) on any Clifford algebra \( \mathcal{C}(X) \) that is uniquely specified by the following properties:

(i) \( E(1) = 1 \).

(ii) \( E(\mathcal{C}(x)) = 0 \) for all \( x \) in \( X \).

(iii) For all \( A \in \mathcal{C}(X_1) \) and \( B \in \mathcal{C}(X_2) \), where \( X_1 \) and \( X_2 \) are any pair of orthogonal subspaces of \( X \)

\[
E(AB) = E(A) E(B).
\]
This last property is called the independence property, and in the context of ordinary probability measures on Euclidean spaces, the corresponding property characterizes the Gaussian measures, so that as Segal has emphasized [Se56], \( E \) is a non-commutative analog of expectation with respect to a Gaussian law.

The uniqueness of such a functional, provided it exists, is a clear consequence of the properties listed above. Furthermore, note that

\[
\xi^a = (-1)^{|a|} (|a| - 1)/2 (\xi^a)^*,
\]

(2.5)

then since \((\xi^a)^* \xi^a = 1\) for all \( a \), while \((\xi^a)^* \xi^b = (\pm 1) \xi^c\) for some \( \xi \) with \(|\xi| > 0\) for \( a \) different from \( b \), and since

\[
E\xi^a = 0 \quad \text{for all } \xi \text{ with } |\xi| > 0
\]

(2.6)

by (ii) and (iii), the \( \xi^a \) are orthonormal for the inner product

\[
\langle A, B \rangle = E(A^* B)
\]

(2.7)

and in particular, are a basis, as claimed above.

Next, suppose \( A = \sum_a A_a \xi^a \) and \( B = \sum_a B_a \xi^a \). Then by (2.5)

\[
E(AB) = \sum_a (\xi a|a| - 1)/2 A_a B_a = E(BA)
\]

(2.8)

which means that \( E \) is tracial.

Now a Clifford algebra over a vector space \( X \) of finite dimension \( 2N \) is naturally isomorphic to the algebra of all \( 2N \times 2N \) matrices. If one wishes to make this identification explicit, there are a number of ways to concretely realize \( \mathcal{C}(X) \) as an algebra of operators on some Hilbert space. In the two most common, this Hilbert space is either exterior algebra over \( X \); i.e., a Fock space representation, or the \( N \)-fold tensor product of \( \mathcal{C} \), with itself; i.e., a spin-chain representation. In the latter case, \( N \) is one-half the dimension of \( Y \) if this is even. See [Se56], [Gr72] for the Fock-space construction and [BW], [CL1] for the spin chain construction.

The unique linear functional \( \tau \) on a matrix algebra that is tracial and normalized (i.e., has \( \tau(I) = 1 \)) is the normalized trace. Hence, \( E \) must be the normalized trace in finite dimensions. (The conclusion is easily seen to follow also when the dimension of \( X \) is odd.) On the other hand, the normalized trace is easily seen to satisfy the defining properties (i), (ii) and (iii). This proves the existence of \( E \) in the finite dimensional case.

The infinite dimensional case is more subtle. One considers first the sub-algebras generated by the finite dimensional subspaces of \( X \). Let \( \mathcal{F} \) denote the collection of all such subspaces. On each \( \mathcal{C}(Y) \), \( Y \) in \( \mathcal{F} \), we have the
normalized trace functional $E_Y$ through the identification with a matrix algebra.

Moreover, one easily sees that whenever $Y_1, Y_2$ are subspaces of $X$ with $Y_1 \subset Y_2$,

$$E_{Y_1}A = E_{Y_2}A \quad \text{for all } A \in \mathcal{C}(Y_1)$$  \hspace{1cm} (2.9)

which is a direct analog of the Kolmogorov consistency condition on finite dimensional marginals. Now since any $A$ in $\mathcal{C}(X)$ belongs to $\mathcal{C}(Y)$ for some $Y \in \mathcal{F}$, the consistency provides the existence of the linear functional $E$ on $\mathcal{C}(X)$.

2.3. Non-commutative Integration in Clifford Algebras

To have analogs of such results the monotone convergence theorem that we need in any probabilistic setting, we must extend $E$ to a von Neumann algebra of operators on some Hilbert space $\mathcal{H}$.

To do this in an intrinsic way, one can use the GNS construction. That is, for each $A$ in $\mathcal{C}(X)$, let $L_A$ denote the operator of left multiplication on $\mathcal{C}(X)$:

$$L_A B = AB$$

Let $\| \cdot \|_2$ denote the Hilbertian norm corresponding to the inner product (2.4). Then (see [PR]) for $A$ in $\mathcal{C}(X)$, there is a constant $c_A$ so that

$$\| L_A B \|_2 \leq c_A \| B \|_2$$ \hspace{1cm} (2.10)

Hence, if we let $L^2(\mathcal{C}(X))$ denote the completion of $\mathcal{C}(X)$ in the norm $\| \cdot \|_2$, $L_A$ uniquely extends to a bounded operator on $L^2(\mathcal{C}(X))$, and since $\mathcal{C}(X)$ contains a unit and $E(A^*A) = 0$ implies that $A = 0$, it is clear that $A \mapsto L_A$ provides a realization of $\mathcal{C}(X)$ as an algebra of operators on $L^2(\mathcal{C}(X))$.

If one then takes the closure of this operator algebra in the strong operator topology, which is the same as closing it in the weak operator topology, one obtains a von Neumann algebra $\mathcal{M}(X)$ contained in the algebra of all bounded linear operators on $L^2(\mathcal{C}(X))$. In [PR], it is then shown that $E$ extends uniquely to $\mathcal{M}(X)$, and this extension, still denoted by $E$, is 

faithful, meaning that $E(A^*A) = 0$ implies that $A = 0$, and normal, meaning that if $\{ A_j \}_{j \in J}$ is any indexed family of elements of $\mathcal{M}(X)$ such that $A_j \leq A_k$ whenever $j \leq k$, and $A_j$ increases to $A \in \mathcal{M}(X)$ in the weak operator topology (pointwise weakly), then $E A_j$ increases to $E A$.

It is this extended normal linear functional $E$ that is the analog of the expectation with respect to a countably additive probability measure provided by the Kolmogorov theorem. The fact that it is normal permits us to do all of the estimates we need in the later sections at the finite dimensional level.

Now let the $L^p$ norms, $1 \leq p < \infty$, be defined by

$$\| A \|_p = (E(A^*A)^{p/2})^{1/p}$$ \hspace{1cm} (2.11)
In a von Neuman algebra of operators, the spectral theorem can be used, and the $p/2$ power is well defined.

We have already denoted the completion of $\mathcal{C}(X)$ in the Hilbertian metric associated to the inner product (2.4) by $L^2(\mathcal{C}(X))$. For all $p$ with $1 \leq p < \infty$, we define the space $L^p(\mathcal{C}(X))$ to be the closure of $\mathcal{C}(X)$ in the $L^p$ norm. As shown in [Se53], [Ne74], $L^p(\mathcal{C}(X))$ is a topological subspace of the space $\mathcal{M}(X)$ of measurable operators. See these references for further information. The elements of these spaces are considered as non-commutative random variables with an integrable $p$th absolute power.

Here however we also need $L^p$ spaces of stochastic process: both as input and output of the Clifford integrals studied here.

2.4. Clifford Processes

Therefore, let $\mathcal{S}$ be the span of all indicator functions of intervals of $\mathbb{R}_+$ as in the introduction. Then any element $G$ of $\mathcal{C}(X) \otimes \mathcal{S}$ can be written in the form

$$G = \sum_{j=1}^{N} G_j \otimes 1_{(s_j, t_j)}$$

(2.12)

where $s_j < t_j < s_{j+1}$ for all $j$. Evidently, this may be identified with the map $t \mapsto G(t)$ form $\mathbb{R}_+$ to $\mathcal{C}(X)$ given by

$$G(t) = \sum_{j=1}^{N} G_j 1_{(s_j, t_j)}(t).$$

(2.13)

Therefore, we refer to the elements of $\mathcal{C}(X) \otimes \mathcal{S}$ as simple Clifford valued process.

There is of course a natural $L^2$ norm on $\mathcal{C}(X) \otimes \mathcal{S}$. To express it neatly, we fix the convention that in any expression of the type (2.12),

$$A_j t = t_j - s_j.$$

(2.14)

Then we can define a norm $\| \cdot \|_{H^2}$, called the $H^2$ norm, by

$$\|G\|_{H^2} = \mathbb{E} \left( \sum_{j=1}^{N} G_j^* G_j A_j t \right).$$

(2.15)

To extend this to other values of $p$, it is natural to start from an analog of the square function for $G$. An important difference with the commutative theory is that there are two of these:

$$\left( \sum_{j=1}^{N} G_j^* G_j \right)^{1/2} \quad \text{and} \quad \left( \sum_{j=1}^{N} G_j (G_j^*) \right)^{1/2}.$$
These two self adjoint operators can have very different eigenvalues. For example, suppose $p$ is a rank-one projection, and $U_j$, $j = 1$ to $N$, are unitary operators that map the range of $P$ into mutually orthogonal sub-spaces. Then put $G_j = U_j P$. In this case, the only non-zero eigenvalue of $(\sum_{j=1}^N G_j(G_j^*)^{1/2}$ is 1, which has multiplicity $N$. On the other hand, the only non-zero eigenvalue of $(\sum_{j=1}^N G_j^* G_j)^{1/2}$ is $N$, which is simple. Thus for this example,

$$\left\| \sum_{j=1}^N G_j(G_j^*) \right\|_p^{1/p} = N^{1/p} \quad \text{and} \quad \left\| \left( \sum_{j=1}^N G_j^* G_j \right)^{1/2} \right\|_p = N^{1/2}. \quad (2.17)$$

These agree for $p = 2$, but otherwise, letting $N$ become large, one sees that the two norms considered above are incomparable.

Here we make the definition of the $H^p$ norm of a Clifford process:

$$\|G\|_{H^p} = \left\| \left( \sum_{j=1}^N G_j^* G_j \right)^{1/2} \right\|_p \quad (2.18)$$

and note that the other choice amounts to $\|G^*\|_{H^p}$. It is not immediately obvious that this norm satisfies the triangle inequality, but they have been considered before, and a simple proof can be found in [LPP].

This summarizes what we need of integration in Clifford algebras, and we now turn to the differential calculus in this setting.

2.5. Differential Calculus in Clifford Algebras: Directional and Global Derivatives

Let $X^*$ denote the space of linear functionals on $X$. As shown on [Bo], there exists for any given $\nu$ in $X^*$ a unique endomorphism $\nabla_\nu$ of $\mathcal{E}(X)$ that has the following properties:

(i) $\nabla_\nu 1 = 0$.

(ii) $\nabla_\nu \mathcal{J}(x) = \nu(x)$ for all $x$ in $X$.

(iii) For all $A$ and $B$ in $\mathcal{E}(X)$,

$$\nabla_\nu(AB) = (\nabla_\nu A) B = \gamma(A) \nabla_\nu B. \quad (2.19)$$

This last property is called the graded Leibniz rule. Note that (i) and (ii) say that $\nabla_\nu$ behaves like the usual directional derivative on linear functions. Moreover, for any $\nu, \nu' \in X^*$, $\nabla_\nu$ and $\nabla_{\nu'}$ anticommute, in contrast with usual directional derivatives. To give an explicit formula in a basis, let $\{e_i | i \in I\}$ be an ordered basis for $X$, and let $\{u_i | i \in I\}$ be the dual basis of $X^*$. That is, $u_i(e_j) = \delta_{i,j}$ for all $i$ and $j$. Put $\nabla_i = \nabla_{u_i}$ and $\xi_j = \mathcal{J}(e_j)$ as before. Then for any Fermion multi-index

$$\nabla_i \xi^a = (-1)^{k-1} \xi^{a-i} \quad (2.20)$$
where in the case that \( i \in \alpha, \), \( k \) is the place of \( i \) in \( \alpha \), and in case \( i \notin \alpha \), \( \xi^{a\rightarrow(i)} \) is defined to be 0.

The definition given above is valid in all Clifford algebras without restriction on the quadratic form \( \phi \). This is useful since one case of interest is \( \phi = 0 \), in which case one gets the Grassman algebra over \( X \). However, if \( \phi \) is non-degenerate, one can identify \( X \) and \( X^* \), and define the operators \( \nabla_y \) for \( y \in X \). These satisfy properties (i) and (iii) above, and

\[
(ii') \quad \nabla_y \phi(x) = \phi(y, x) \quad \text{for all } x \in X.
\]

It is easy to see that for any given \( A, u \mapsto \nabla_u A \) is linear. This enables us to define the global derivative map

\[
\nabla: \mathcal{C}(X) \to \mathcal{C}(X) \otimes X
\]

associated to these directional derivatives. First, we equip \( \mathcal{C}(X) \otimes X \) with the inner product

\[
\langle B \otimes u, A \otimes w \rangle = \langle B, A \rangle \langle u, w \rangle.
\]

Then \( \nabla A \) is defined by

\[
\langle \nabla A, B \otimes u \rangle = \langle \nabla_u A, B \rangle
\]

for all \( u \) in \( X \) and \( B \) in \( \mathcal{C}(X) \). The following global version clearly holds

\[
\nabla(AB) = (\nabla A)B + \gamma(A)VB.
\]

Moreover, if \( X \) is the sum of two orthogonal spaces \( X_1 \) and \( X_2 \), then we have by restriction the partial gradients

\[
\nabla^{(j)}: \mathcal{C}(X) \to \mathcal{C}(X) \otimes X_j \quad j = 1, 2
\]

and \( \nabla = \nabla^{(1)} + \nabla^{(2)} \). Also, if \( u \in X^* \) vanishes on \( X_1 \), the \( \nabla_u \) vanishes on \( \mathcal{C}(X_1) \).

2.6. Differential Calculus in Clifford Algebras: Divergence

Then the divergence \( \delta \) is the adjoint of \( \nabla \) so that

\[
\delta: \mathcal{C}(X) \otimes X \to \mathcal{C}(X).
\]

Hence, for any \( G \) in \( \mathcal{C}(X) \otimes X \), \( \delta G \) is the unique element of \( \mathcal{C}(X) \) such that

\[
\langle \delta G, b \rangle = \langle G, \nabla B \rangle
\]

for all \( B \) in \( \mathcal{C}(X) \).
From these relations, it follows that \( V \) and \( \delta \) are both closable operators on Hilbert spaces obtained by completing the quadratic spaces on which they are defined.

In the case of the divergence, if \( X \) is the sum of two orthogonal spaces \( X_1 \) and \( X_2 \), we define \( \delta^{(j)}: \mathcal{H}(X) \otimes X_j \rightarrow \mathcal{H}(X) \) as the adjoint to \( \nabla^{(j)} \). The resulting operator is just the restriction of \( \delta \) to the smaller domain

\[ \mathcal{H}(X) \otimes X_j \subset \mathcal{H}(X) \otimes X. \]

2.7. The Clifford Algebra \( \mathbb{C}(Q, P) \)

So far we have presented the theory in general terms. We now turn to the concrete example that is of particular interest here: Let \( S \) denote, as before, the span of all of the indicator functions of intervals \( 1_{(s, t)} \) on \( \mathbb{R}_+ \) with the quadratic form

\[ \phi(h) = \int_{\mathbb{R}} h(t)^2 \, dt. \]  

Note that this algebra \( \mathcal{H}(\mathcal{S}) \) has a natural filtration of subalgebras \( \mathcal{H}(\mathcal{S}_t) \subset \mathcal{H}(\mathcal{S}_u) \) for \( t \leq u \); this is simply induced by the filtration of subspaces \( \mathcal{S}_t = \mathcal{S} \cap L^2([0, t]) \).

Next define a stochastic process \( Q(t) \) by putting

\[ Q(t) = \mathcal{J}(1_{(s, t)}) \in \mathcal{H}(\mathcal{S}_t). \]  

Then notice that

\[ (Q(u) - Q(t))^2 = u - t \]  

for all \( u > t \), even without taking an expectation, and that the increments corresponding to non-overlapping intervals are independent. This is the Fermionic Browninan motion introduced by Barnett, Streater and Wilde [BSW]. They developed a theory of non-commutative stochastic integration for this process in an \( L^2 \) setting as we have explained in the introduction.

In further developing the subject, it is advantageous to work in a somewhat broader setting: Consider the complexification \( \mathcal{S}_c \) of \( \mathcal{S} \) as a real Hilbert space; with \( \mathcal{S}_Q \) and \( \mathcal{S}_P \) denoting its real and imaginary parts, which are linearly independent over \( \mathbb{R} \):

\[ \mathcal{S}_c = \mathcal{S}_Q \oplus \mathcal{S}_P. \]

Thus, \( \mathcal{S}_Q \) and \( \mathcal{S}_P \) are real quadratic spaces, and \( \mathcal{S}_P \) is identified with \( \mathcal{S}_Q \) by \( \mathcal{S}_P = i\mathcal{S}_Q \). The quadratic form \( \phi \) on \( \mathcal{S} \) extends naturally to a quadratic form on \( \mathcal{S}_Q \oplus \mathcal{S}_P \) that we also denote by \( \phi \).
Now let \((\mathcal{J}, \mathcal{E}(Q, P))\) denote the Clifford algebra of \((\mathcal{J}^Q \oplus \mathcal{J}^P, \phi)\). We write \(\mathcal{E}(Q)\) and \(\mathcal{E}(P)\) to denote the subalgebras corresponding to \(\mathcal{J}^Q\) and \(\mathcal{J}^P\) respectively.

The \(Q\) and \(P\) notation comes from mathematical physics where such a construction is associated with the introduction of phase-space variables for fermionic systems. Here we will see that there are purely probabilistic reasons for working in this larger setting. The key fact is that the larger algebra \(\mathcal{E}(Q, P)\) is rich in ordinary Bernoulli variables (in a commutative subalgebra of \(\mathcal{E}(Q, P)\)) that are indexed by the intervals \((s, t)\) in a way that is consistent with the natural filtration. This fact too comes from mathematical physics, specifically the Jordan–Wigner transform \([JoWi, SML]\).

Here we will develop this methodology as a tool in the theory of Clifford integrals.

In this setting, we have two fermionic Brownian motions \(Q(t)\) and \(P(t)\) whose corresponding increments anti-commute:

\[
(Q(t) - Q(s))(P(t) - P(s)) = -(P(t) - P(s))(Q(t) - Q(s)).
\]

That is, for any \(h\) in \(\mathcal{J}\), let \(h^Q = h \oplus 0\) denote the natural imbedding of \(h\) into \(\mathcal{J}^Q \oplus \mathcal{J}^P\) through the first summand, and let \(h^F = 0 \oplus h\) denote the natural imbedding of \(h\) into \(\mathcal{J}^Q \oplus \mathcal{J}^F\) through the second summand. Then

\[
Q(t) = \mathcal{J}(1_{[0,t]}^Q) \quad P(t) = \mathcal{J}(1_{[0,t]}^P).
\]

Since they are generated by orthonormal subspaces, they are independent.

Next, we clearly have two gradient operators \(\nabla_Q\) and \(\nabla_P\) such that

\[
\nabla = \nabla_Q + \nabla_P
\]

and the corresponding divergence operators \(\delta_Q\) and \(\delta_P\).

We now prove a lemma we shall use in the next section; its proof is given here to tie together some of the material introduced in this section.

**Lemma 2.1.** For any bounded time interval \((s, t)\), and any \(A\) in \(\mathcal{E}(Q, P)\),

\[
\mathbb{E}(\nabla_Q^{t,s} A) = \mathbb{E}((Q(t) - Q(s)) A) \quad (2.34)
\]

\[
\mathbb{E}(\nabla_P^{t,s} A) = \mathbb{E}((P(t) - P(s)) A). \quad (2.35)
\]

**Hence**

\[
\delta_Q(1_{(s,t)}) = (Q(t) - Q(s)) \quad \text{and} \quad \delta_P(1_{(s,t)}) = (P(t) - P(s)). \quad (2.36)
\]

**Proof.** The element \(A\) is a polynomial in variables of the form \(Q(t_j) - Q(s_j)\) and \(P(t_j) - P(s_j)\) for some finite family of non-overlapping intervals \((s_j, t_j)\). Now both sides of the equalities we are trying to prove are additive in the
interval \((s, t)\). Hence, refining the partition if necessary, we may assume that either \((s, t)\) is disjoint from each \((s_j, t_j)\), or that \((s, t) = (s_j, t_j)\) for some \(j\).

Now since \(\mathcal{Q}_1\left(Q(t_k) - Q(s_k)\right) = \delta_{j,k}(t_j - s_j)\), and \(\mathcal{P}_1\left(P(t_k) - P(s_k)\right) = 0\) for all \(j\) and \(k\), \(\mathbb{E}\mathcal{Q}_1\mathcal{Q}_1 = 0\) unless \((s, t) = (s_j, t_j)\) for some \(j\), and \(Q(t_j) - Q(s_j)\) is a factor in \(A\). In this case, we compute \(\mathcal{Q}_1 A\) using the graded Leibnitz rule, and compute \((Q(t) - Q(s))A\) by anticommuting \((Q(t) - Q(s))\) through to the \(j\)th place, and then using \((Q(t) - Q(s))^2 = t - s\). In this case we have

\[
\mathcal{Q}_1 A = (Q(t) - Q(s))A
\]

even without taking the expectation. But if \((Q(t) - Q(s))\) is not a factor of \(A\), then \(\mathbb{E}((Q(t) - Q(s))A) = 0\) by the independence property. Thus the first equality is proved, and the companion statement concerning \(P\) is proved in the same way. The final statement follows by duality.

3. CLIFFORD DERIVATIVES AND CLIFFORD STOCHASTIC INTEGRALS

The \(L^2\) theory [BSW] has been described in the introduction. As far as the \(L^2\) theory of the integration of simple adapted processes is concerned, the only differences between what they have done and what we need here stem from the fact that we are working in the larger Clifford algebra \(\mathbb{H}(Q, P)\), and hence have the two Clifford processes defined in (2.33). The Clifford integrals of simple integrands with respect to these processes are defined as in (1.3).

It is easy to see that the isometry property holds for both of these integrals:

\[
\mathbb{E}\int G(t) dQ(t) = \mathbb{E}\int G(t) dP(t) = \mathbb{E}\left(\sum_{j=1}^{N} G_j^* G_j A_j \right).
\]

Thus, the two maps

\[
G \mapsto \int G dQ \quad \text{and} \quad G \mapsto \int G dP
\]

extend to isometries form \(\mathcal{A}_2\) to \(L^2(\mathbb{H}(Q, P))\).

We begin by proving that \(\mathcal{A}_2\) is contained in the domain of \(\mathcal{D}_\gamma\), and that on \(\mathcal{A}_2\), \(\mathcal{D}\) and the Clifford integral coincide. At present, we are working in \(L^2\), and the following theorem refers to the \(L^2\) Clifford integral. Once we have proved Theorem 1, we will be able to extend this result to \(L^1\), and obtain Theorem 3.
Proposition 3.1 (The Clifford Integral as a Divergence). All square integrable past adapted process \( G(\cdot) \) are in the domains of \( \delta_Q \) and \( \delta_P \), with

\[
\int G \, dQ = \delta(\gamma_d G(\cdot)) \quad \text{and} \quad \int G \, dP = \delta(\gamma_p G(\cdot)).
\]  

(3.1)

In particular, \( \delta_Q : \mathcal{A}_2 \to L^2 \) and \( \delta_P : \mathcal{A}_2 \to L^2 \) are isometries.

**Proof.** We treat the \( Q \) integration; the proof is identical for the \( P \) integration. Because the divergence is a closed operator, it suffices to establish (3.1) for simple adapted processes. By linearity, it suffices to consider processes of the form \( G(\cdot) = A \otimes 1_{(t_1, t_2)} \) where \( A \in \mathcal{C}(Q, P)_{(t_1)} \). Finally, since set of elements of \( \mathcal{C}(Q, P)_{(t_1)} \) is dense in \( L^2(\mathcal{C}(Q, P)) \), it suffices to check that

\[
\langle A(Q(t_2) - Q(t_1)), BC \rangle = \langle \delta(\gamma(A \otimes 1_{(t_1, t_2)})), BC \rangle
\]  

(3.2)

for such \( B \) and \( C \).

Using the cyclicity of the trace and the independence property (2.4), we have that

\[
\langle \delta(\gamma(A \otimes 1_{(t_1, t_2)})), BC \rangle = \langle \gamma(A \otimes 1_{(t_1, t_2)}), \nabla(BC) \rangle
\]

\[
= \langle \gamma(A \otimes 1_{(t_1, t_2)}), \nabla B + \gamma(B) \nabla C \rangle
\]

\[
= \langle \gamma(A), \nabla_{1_{(t_1, t_2)}} B \rangle C + \gamma(A), \nabla_{1_{(t_1, t_2)}} B \rangle C \rangle.
\]  

(3.3)

However, since \( B \in \mathcal{C}(Q, P)_{(t_1)}, \nabla_{1_{(t_1, t_2)}} B = 0 \). The remaining term in (3.3) is:

\[
\langle \gamma(A \otimes 1_{(t_1, t_2)}), \nabla B \rangle \nabla C \rangle
\]

\[
= \langle \gamma(A), \nabla_{1_{(t_1, t_2)}} B \rangle \nabla C \rangle
\]

\[
= \langle \gamma(A), \nabla_{1_{(t_1, t_2)}} B \rangle \nabla C \rangle = \langle \gamma(A), \nabla B \rangle \nabla C \rangle
\]

where we have once again used the independence property (2.4).
Finally, by a Lemma 2.1,
\[ E(V_{\mathcal{Y}_\theta}(C)) = E((F(t_2) - F(t_1))C). \]
Thus, \( \langle \delta \gamma (A \otimes 1_{(t_1, t_2)}), BC \rangle = \langle A, B \rangle \langle (F(t_2) - F(t_1)), C \rangle \) which establishes the result.

**Proof of Theorem 3.** One simply applies Theorem 1 and the obvious approximation argument to extend the validity of Proposition 3.1. No circularity occurs as the proof of Theorem 1 itself will only require Proposition 3.1.

Proposition 3.1 may now be applied to obtain a martingale representation theorem that extends a result of Barnett, Streater and Wilde by giving an explicit expression for the integrand that is a close analog of the Clark–Hausman–Ustenel formula for Wiener martingales. The proof is similar in form to one given in the Gaussian case in [Kr].

**Proposition 3.2.** Any element \( A \) of \( L^2(\mathcal{G}Q, \mathcal{P}) \) can be uniquely represented as a Clifford integral
\[
A = E(A) + \int G_Q dQ + \int G_P dP \tag{3.4}
\]
where \( G_Q(\cdot) \) and \( G_P(\cdot) \) belong to \( \mathcal{A}_2 \). Moreover, these unique processes are given by
\[
G_Q(\cdot) = \gamma(\text{Ad}_Q A) = \text{Ad}(\gamma \text{V}_Q A)
\]
\[
G_P(\cdot) = \gamma(\text{Ad}_P A) = \text{Ad}(\gamma \text{V}_P A) \tag{3.5}
\]
where \( \text{Ad}: L^2(\mathbb{R}_+, L^2(\mathcal{G}Q, \mathcal{P})) \rightarrow L^2(\mathcal{G}Q, \mathcal{P}) \) denotes the orthogonal projection onto the adapted square integrable processes.

**Corollary 3.3.** For all \( A \) in \( L^2(\mathcal{G}Q, \mathcal{P}) \),
\[
\delta(\text{Ad}_Q A) = A - E(A) \tag{3.6}
\]

**Proof.** This follows immediately from Propositions 3.1 and 3.2.

**Proof of Proposition 3.2.** We first prove uniqueness, and identify the integrands \( G_Q(\cdot) \) and \( G_P(\cdot) \) at the same time. First, we may assume that \( E(A) = 0 \). Then, suppose that \( A = \delta G_Q(\cdot) + \delta G_P(\cdot) \) for some past-adapted square integrable processes \( G_Q \) and \( G_P \). By Proposition 3.1, this is the same as assuming that \( A \) has some representation of the form (3.4).

Then for any \( B \in \mathcal{G}Q, \mathcal{P} \), we have
\[
\langle G_Q(\cdot), \text{V}_Q B \rangle = \langle G_Q, \text{Ad}(\text{V}_Q B) \rangle = \langle \delta G_Q(\cdot), \text{Ad}(\text{V}_Q B) \rangle
\]
since $Q$ is an isometry by Proposition 3.1. The same result clearly holds with $Q$ substituted by $P$.

But then using the hypothesis that $\delta(\gamma G_{Q}(\cdot)) + \delta(\gamma G_{P}(\cdot)) = A$, together with the fact that $\nabla_{Q}B + \nabla_{P}B = \nabla B$,

$$\langle G_{Q}(\cdot), \nabla_{Q}B \rangle + \langle G_{P}(\cdot), \nabla_{P}B \rangle = \langle A, \delta(\gamma \text{Ad}(\nabla B)) \rangle = \langle \gamma(\text{Ad}(A)), \nabla B \rangle.$$

Then integrating by parts, we have

$$\int G_{Q} \, dQ + \int G_{P} \, dP - \delta(\gamma \text{Ad}A) = 0$$

or

$$\int (G_{Q} - \text{Ad} \nabla A) \, dQ + \int (G_{P} - \text{Ad} \nabla A) \, dP = 0.$$

This proves the uniqueness since both of these integrals are orthogonal, and hence must vanish separately. The isometry property then implies that both integrands vanish, and thus establishes the forms of $G_{Q}(\cdot)$ and $G_{P}(\cdot)$.

We now prove existence. Consider the span of the identity together with the elements $A$ of $\mathcal{F}(Q, P)$ of the form

$$A = \prod_{i=1}^{N} (Q(t_{i}) - Q(s_{i}))^{x_{i}} (P(t_{i}) - P(s_{i}))^{y_{i}} s_{i} < t_{i} < t_{i+1}$$

and with each $x_{i}$ and $y_{i}$ equal to either 0 or 1. We may assume that for each $i$, at least one of $x_{i}$ and $y_{i}$ is non-zero. It is clear that this is dense in $L^{2}(\mathcal{F}(Q, P))$.

For such $A$, there are three cases to consider. If $x_{N} = 0$, define the process $G_{P}(\cdot)$ defined by

$$G_{P}(t) = \prod_{i=1}^{N-1} (Q(t_{i}) - Q(s_{i}))^{x_{i}} (P(t_{i}) - P(s_{i}))^{y_{i}} \otimes 1_{(t_{i}, t_{i+1})}(t).$$

Evidently, $G_{P}$ is past-adapted, and $\int G_{P} \, dP = A$.

A corresponding prescription works when $y_{N} = 0$. The remaining case is when $x_{N} = y_{N} = 1$.

To treat this case, consider the problem of representing $Q(1)P(1)$ through Clifford integrals; clearly, our problem easily reduces to this. For each integer $m$, let $t_{j} = j/m$ for $j = 0, 1, ..., m$, and put $\delta_{j}Q = Q(t_{j+1}) - Q(t_{j})$, and similarly for $\delta_{j}P$. Then
\[ Q(1) P(1) = \left( \sum_{j=1}^{m} \delta_j Q \right) \left( \sum_{k=1}^{m} \delta_k P \right) = \sum_{j=1}^{m} \delta_j Q P(t_j) + \sum_{k=1}^{m} Q(t_k) \delta_k P + \sum_{j=1}^{m} \delta_j Q P. \]

Then one easily computes
\[ E\left| \sum_{j=1}^{m} \delta_j Q P \right|^2 = 1/m. \]

Therefore
\[ Q(1) P(1) = \lim_{m \to \infty} \left( -\sum_{j=1}^{m} P(t_j) \delta_j Q + \sum_{k=1}^{m} Q(t_k) \delta_k P \right) = \int G_Q dQ + \int G_P dP \]

where
\[ G_Q(t) = -P(t) \quad \text{and} \quad G_P(t) = Q(t). \]

Now it is clear that for our problem we should define
\[ G_Q(t) = -\prod_{i=1}^{N-1} (Q(t_i) - Q(s_i))^\gamma (P(t_i) - P(s_i))^\delta P(t) \]

and
\[ G_P(t) = \prod_{i=1}^{N-1} (Q(t_i) - Q(s_i))^\gamma (P(t_i) - P(s_i))^\delta Q(t) \]

for \( t \) in \( (t_{N-1}, t_N) \), and both to be 0 otherwise.

Thus, for a dense subspace in \( L^2(\mathcal{C}(Q, P)) \), there does exist a representation of the form (3.4).

**Proof of Theorem 4.** We now approximate any \( A \in L^p(\mathcal{C}(Q, P)) \), \( 1 \leq p \leq 2 \) by a sequence \( A^* \in \mathcal{C}(Q, P) \). Assume that all of the expectations are zero. Proposition 3.2 applies, and so we let \( G_Q^* \) and \( G_P^* \) be the integrands such that
\[ A^* = \int G_Q^* dQ + \int G_P^* dP. \]

Then by the Pisier-Xu lower bound,
\[ \|G_Q^* - G_Q^m\|_{\mathcal{H}} \leq K_p \|A^* - A^m\|_p \]
and hence the \( G_Q^\ast \) and the \( G_P^\ast \) are a Cauchy sequence in \( H^p \). We may now take the limit and obtain

\[
A = \int G_Q \, dQ + \int G_P \, dP.
\]

The statements concerning \( L^2 \) in Theorem 4 are already proved.

We now obtain an explicit Clifford stochastic integral representation for the absolute square of a simple Clifford stochastic integral. This may be regarded as an “Ito’s formula” for quadratic functionals of Clifford integrals. While the consequences of this quadratic Ito’s formula are much more limited than in the classical case, we shall be able to apply it in the next section on non-quadratic functionals.

The following formulation is a provisional version of Theorem 5. We formulate it now as a proposition for simple integrands. We need Theorem 1 to extend this even to \( L^2 \) integrands, since for these the integrand in (3.7) below will only be \( L^1 \).

**Proposition 3.4.** Let \( G_Q(\cdot) \) and \( G_P(\cdot) \) be simple past adapted processes. For all \( t \geq 0 \), let \( M(t) = M_Q(t) + M_P(t) \) where

\[
M_Q(t) = \int 1_{(0, t]} G_Q \, dQ \quad \text{and} \quad M_P(t) = \int 1_{(0, t]} G_P \, dP.
\]

Then,

\[
|M|^2 = \int (|\gamma G_Q|^2 + |\gamma G_P|^2) \, dt
\]

\[
+ \int \left( (M_Q(t))^\ast G_Q(t) + \gamma (G_Q(t))^\ast M_Q(t) + M_P^\ast G_Q + \gamma (G_Q^\ast M_P) \right) dQ(t)
\]

\[
+ \int \left( (M_P(t))^\ast G_P(t) + \gamma (G_P(t))^\ast M_P(t) + M_Q^\ast G_P + \gamma (G_P^\ast M_Q) \right) dP(t).
\]

**Remark.** Barnett, Streater and Wilde proved a result which would be formulated here as the statement that \( |M_Q(t)|^2 - \int 1_{(0, t]} |\gamma G_Q(s)|^2 \, ds \) is a Clifford martingale, and hence a Clifford integral. (see [BSW] and Section 4). However, they do not give the explicit form (3.8) of the Clifford integral. In fact one cannot do this in any generality in a purely \( L^2 \) setting: the integrands in (3.8) will in general only belong to \( H^1 \) when \( G_Q \) and \( G_P \) belong to \( H^2 \).
Proof. We may write

\[ G_Q(s) = \sum_{i=1}^{n} G_Q^i \otimes L_{(t_i, u_i)}(s) \quad \text{and} \quad G_P(s) = \sum_{i=1}^{n} G_P^i \otimes L_{(t_i, u_i)}(s) \quad (3.9) \]

with \( t_i < u_i < t_{i+1} \) and \( G_Q^i, G_P^i \in \mathcal{C}(Q, P)_{t_i} \) for all \( i \). We may freely suppose that \( u_0 \leq t \).

We shall first show that

\[ |M_Q|^2 = \sum_{i=1}^{n} \left| \int (M_Q(t))^* G_Q(t) + \gamma(G_Q(t))^* M_Q(t) \right| dQ(t). \]

Using the notation \( A_Q := Q(u_i) - Q(t_i) \), etc., we have from Proposition 3.2 and the isometry property that

\[ E \left[ \int G_Q(s) dQ(s) \right]^2 = \sum_{i=1}^{n} A_Q \int G_Q(s) dQ(s) \quad (3.10) \]

To proceed, note that

\[ \sum_{i=1}^{n} A_Q \int G_Q(s) dQ(s) = \sum_{i=1}^{n} (\gamma(G) A_Q) \]

and split the sum into three pieces, according to whether \( i < j, i = j, \) or \( i > j \).

Starting with the \( i = j \) terms, note that by

\[ (\gamma(G) A_Q = \gamma(G) A_Q \quad (3.11) \]

for all \( G \) in the past of time \( t_j \), and by the fact that and \( (A_Q)^2 = A_Q t \),

\[ \sum_{i=1}^{n} A_Q \int G_Q^i dQ(s) = \sum_{i=1}^{n} |G_Q^i|^2 A_Q t, \]

so that by (3.6),

\[ \left( \delta_Q \text{Ad} \sum_{i=1}^{n} |G_Q^i|^2 A_Q t \right) = \sum_{i=1}^{n} |G_Q^i|^2 A_Q t - \sum_{i=1}^{n} E |G_Q^i|^2 A_Q t. \]

Thus, the \( i = j \) contribution to the second term on the right in (3.10) cancels out the first term, and replaces it with the first term on the right in (3.8).

We now turn to the \( i < j \) contribution. First, by the commutation property (3.11),

\[ \sum_{j=1}^{n} \sum_{i=1}^{j-1} A_Q G_Q^i \otimes G_Q^i A_Q = \sum_{j=1}^{n} \left( \sum_{i=1}^{j-1} \gamma(G_Q^i \otimes G_Q^i A_Q) \right) A_Q \quad (3.12) \]
The right hand side is clearly the Clifford integral of a simple adapted process. By Corollary 3.3, $\delta \cdot \text{Ad} \cdot V$ is the identity on such integrals. Thus
\[
\delta \text{AdV} \left( \sum_{j=1}^{n} \sum_{i=1}^{j-1} A_j Q G_Q^i \ast G_Q^i A_i Q \right) = \sum_{j=1}^{n} \sum_{i=1}^{j-1} A_j Q G_Q^i \ast g_Q^i A_i Q \\
= \sum_{j=1}^{n} \gamma G_Q^i \ast \gamma \left( \sum_{i=1}^{j-1} g_Q^i A_i Q \right) A_j Q.
\]
In the same way, but without need of commutation, we compute
\[
\delta \text{AdV} \sum_{i=1}^{n} \sum_{j=1}^{i-1} A_j Q G_Q^i \ast G_Q^i A_i Q = \sum_{j=1}^{n} \sum_{i=1}^{j-1} G_Q^i A_i Q \ast G_Q^i A_j Q.
\]
Altogether, we have:
\[
\left| \int G_Q(s) \, dQ(s) \right|^2 = \sum_{i=j}^{n} |\gamma G_Q^i|^2 A_i t + \sum_{j=1}^{n} \left( \sum_{i=1}^{j-1} G_Q^i A_i Q \right) \ast G_Q^i A_j Q \\
- \sum_{j=1}^{n} \gamma G_Q^i \ast \gamma \left( \sum_{i=1}^{j-1} G_Q^i A_i Q \right) A_j Q.
\]
This is (3.8) written as a sum. Simply exchanging $Q$ for $P$ in the above argument gives
\[
|M_P|^2 = \left| \int |\gamma G_P|^2 \, dt + \int \left( (M_P(t))^* G_P(t) + (\gamma G_P(t))^* (M_P(t)) \right) \, dP(t) \right|.
\]
It remains to consider $M_Q^* M_P + M_P^* M_Q$. Since $M_Q$ and $M_P$ are orthogonal, each of these has zero expectation. Consider the first of these, which is
\[
\sum_{i,j=1}^{n} A_j Q G_Q^i \ast g_P^j A_i P.
\]
We split the double sum into parts as before. The main difference is with the $i = j$ terms:
\[
\sum_{i=1}^{n} A_i Q G_Q^i \ast G_P^i A_i P = \sum_{i=1}^{n} \gamma (G_Q^i \ast G_P^i) A_i Q A_i P.
\]
As in the proof of Proposition 3.3, we consider
\[
\left| \sum_{i=1}^{n} \gamma (G_Q^i \ast G_P^i) A_i Q A_i P \right|_2 \leq \max_{i=1, \ldots, n} \left( |G_Q^i \ast G_P^i| \right)_2 \left( \sum_{i=1}^{n} (A_i t)^2 \right).
Now one can represent the same integrand in terms of finer and finer partitions: One just uses the original $G_Q$ and $G_P$ on all of the subintervals obtain from the original $(t_i, u_i)$. Doing this, $\max_{i=1,...,n} \| G_Q^* G_P^r \|^2$ is unchanged, while $\sum_{i=1}^n (A_i)^2$ is decreased. Thus, these terms make no contribution in the limit as we refine the partition.

The sum of the terms corresponding to $i < j$ and $i > j$ yield, arguing exactly as before,

$$\int M^*_Q G_P \, dP + \sum \gamma(G^*_M G)^P \, dQ.$$ Notice that, expressed this way, the result is independent of which refinement of the original partition we might be using.

Doing the same computations for $M^*_M M_Q$, and combining results, we have

$$M^*_Q M_P + M^*_M M_Q = \int (M^*_Q G_Q + \gamma(G^*_M M_P) \, dQ + \int (M^*_Q G_P + \gamma(G^*_M M_Q) \, dP.$$

4. MOMENT INEQUALITIES FOR CLIFFORD INTEGRALS

We now prove Theorem 1, our main inequality relating Clifford integrals and their square functions. Our strategy follows Novikov in his proof of the Burkholder–Gundy inequalities for classical Wiener–Ito integrals. His proof used the Ito formula and certain simple convexity and monotonicity arguments. The adaptation to our context is not automatic, since some of the convex functions he used are not convex as operator functions.

Moreover, we have only a very restricted analog of the Ito formula; namely the one for quadratic functions given by Proposition 3.4. We shall be able to get by with this by combining it with an integral representation for fractional powers of positive operators. The methods used here are similar to the methods used in [BCL] to prove certain sharp convexity properties of trace norms.

Proof of Theorem 1. Let $M_i$ be a simple Clifford integral

$$M_i := \int \chi_{(0, i)} G_Q \, dQ + \int \chi_{(0, i)} G_P \, dP,$$

and let $\langle M, M \rangle_i$ denote the square function process:

$$\langle M, M \rangle_i = \int \chi_{(0, i)} (|G_G(x)|^2 + |G_P(x)|^2) \, ds.$$
Let $a$ and $b$ be positive constants to be determined later, and define a process $S_t$ by

$$S_t := a + |M_t|^2 + b(M, M)_t,$$  \hfill (4.1)

It is clear that $S_t > 0$ for $t \geq 0$.

Next we again suppose that $G_Q$ and $G_P$ have the form (3.9). We shall again use the fact that this representation is not unique: If $0 = s_0 < s_1 < \cdots < s_n = t$ is any partition such that each $(s_j, s_{j+1})$ is a subinterval of $(t_i, u_i) \cap (0, t)$ for some $(t_i, u_i)$ form the partition in (3.9), then we also have

$$G_Q(s) = \sum_{j=1}^N \hat{G}_Q \otimes 1_{(s_j, s_{j+1})}(s) \quad \text{and} \quad G_P(s) = \sum_{j=1}^N \hat{G}_P \otimes 1_{(s_j, s_{j+1})}(s)$$  \hfill (4.2)

where each of the $\hat{G}_Q$ and $\hat{G}_P$ is one of the $G_Q^i$ and $G_P^i$ in (3.9), namely, the one with $(s_j, s_{j+1}) = (t_i, u_i)$.

Our object is to use the stochastic calculus now at our disposal to estimate $\mathbb{E}|S_t|^p$. To do this, recall the integral representation for powers $A^p$ of a positive operator $A$:

$$A^p = c_p \int_0^{\infty} x^{p-1} \left( \frac{1}{x} \right)^{\frac{1}{p}} \mathrm{d}x < 2$$  \hfill (4.3)

where $c_p$ is a finite constant. (This follows directly from the spectral theorem and the corresponding result for real numbers.)

Clearly, an upper bound on $\mathbb{E}|S_t|^p$ for $A = S_t$ gives us an upper bound on $|M_t|_p$, and to get the former, because of (4.3), we seek a lower bound on $\mathbb{E}(x + S_t)^{-1}$.

To do this, choose a partition $0 = s_0 < s_1 < \cdots < s_n = t$ as above. For $1 \leq j \leq n$, put

$$A_j S = S_j - S_{j-1} \quad \text{and} \quad A_j s = s_j - s_{j-1}.$$  

Then, writing out a telescoping sum,

$$\frac{1}{x + S_t} = \frac{1}{x} - \sum_{j=1}^n \left( \frac{1}{x + S_{j-1} + A_j S} - \frac{1}{x + S_{j-1}} \right).$$  \hfill (4.4)

The terms in the telescoping sum can be expanded in what amounts to a Taylor series by repeated use of the because of the so-called “resolvent identity”

$$\frac{1}{C + D} - \frac{1}{C} = \frac{1}{C + D} \frac{1}{C}$$  \hfill (4.5)
valid for all positive operators $C$ and all selfadjoint operators $D$ such that $C + D$ is positive. Iterating this identity, we obtain

$$\frac{1}{C+D} \frac{1}{C} = \frac{1}{C} D \frac{1}{C} + \frac{1}{C} \frac{1}{C} \frac{1}{C} - \frac{1}{C+D} \frac{1}{C} D \frac{1}{C} \frac{1}{C}.$$  \hspace{1cm} (4.6)

We apply this by putting, for fixed $j$ and $x$,

$$C = x + S_{t_{j-1}} \quad \text{and} \quad D = A_j S.$$

Clearly $C$ and $C + D$ are positive. Also, note that

$$\frac{1}{C} D \frac{1}{C} \frac{1}{C} \frac{1}{C} \geq 0.$$

Next we claim that

$$\mathbb{E} \left| \frac{1}{C+D} \frac{1}{C} \frac{1}{C} \frac{1}{C} \right| \leq K(A_j t) \frac{1}{2}$$

where $K$ is a finite constant independent of the partition, and depending only on the process.

To see this, we need an upper bound on the $L^p$ norm of $D = S_j$ for has the form

$$S_j = A_j A_j Q + B_j A_j P + C_j \beta_s,$$

where $A_j, B_j$ and $C_j$ can be easily explicitly computed. Thus, using the cyclicity of the trace:

$$\mathbb{E} \left( \frac{1}{x+S_{t_j}} - \frac{1}{x+S_{t_{j+1}}} \right) \leq \mathbb{E} \left( \left( \frac{1}{x+S_{t_{j+1}}} - S_{t_j} \right)^2 \right) + K(t_{j+1} - t_j) \frac{1}{2}.$$  \hspace{1cm} (4.7)

Then by Proposition 2.3,

$$S_{t_{j+1}} - S_t = \int_{t_j}^{t_{j+1}} H(s) \, dF(s) + (1 + b) \int_{t_j}^{t_{j+1}} d\langle M, M \rangle_s.$$

Inserting this in (4.7), the part involving the Clifford integral drops out when we take the expectation, so we get:

$$\mathbb{E} \left( \left( \frac{1}{x+S_{t_j}} \right)^2 \right) \left( S_{t_{j+1}} - S_{t_j} \right) = (1 + b) \mathbb{E} \left( \left( \frac{1}{x+S_{t_{j+1}}} \right)^2 \right) \int_{t_j}^{t_{j+1}} |\gamma G(s)|^2 \, ds.$$  

Summing, and letting the mesh of the partition tend to zero, this yields

$$\mathbb{E} \left( \frac{1}{x+S_{t_j}} \right) \leq (1 + b) \mathbb{E} \left( \int_{t_j}^{t_{j+1}} \left( \frac{1}{x+S_{s}} \right)^2 |\gamma G(s)|^2 \, ds \right).$$
Thus,
\[ E \left| S_t \right| \leq a^{p/2} + (1 + b) \cdot c_p \int_0^{\infty} x^{p/2 - 1} E \left( \left( \frac{1}{x + S_x} \right)^2 \| yG(s) \| ^2 \right) \, dx. \]

Changing the order of integration, the integration in \( x \) can be explicitly carried out; the constants work out as they would have to in the case where the operators are scalars. Hence,
\[ E(S_t) \leq a^{p/2}(b + 1) \int_0^t E \left( S_{r}^{p/2 - 1} \langle M \rangle_r \right) \, dr. \]

Now clearly \( E \left| M_r \right| \leq E(S_r)^{p/2} \). Also, \( A \mapsto A' \) is operator monotone decreasing for \(-1 < r < 0\), and thus, since \(-1 < p/2 - 1 < 0\),
\[ S_{r}^{p/2 - 1} \leq b^{p/2 - 1} \langle M \rangle_r^{p/2 - 1}. \]

We finally obtain
\[ E \left| M_r \right| \leq a^{p/2} + (1 + b) b^{p/2 - 1} \int_0^t \left( \langle M \rangle_r \right)^{p/2 - 1} \frac{d}{dr} \langle M \rangle_r \, dr = (b + 1) b^{p/2 - 1} E \left( \langle M \rangle_t^{p/2} \right). \]

The optimal value of \( a \) is zero, and the optimal value of \( b \) is found to be \( b = (2 - p)/p \) at which value \((b + 1) b^{p/2 - 1} = (2/p)(2 - p/p)^{p/2 - 1} \).

This proves that \( \| G \, dF \|_p \leq K_p \| G \| _{H_p} \). Applying the same argument to \( G^* \, dE \) we obtain \( \| G^* \, dF \|_p \leq K_p \| G^* \| _{H^*} \). Then since \( \| G^* \, dF \|_p = \| G \, dF \|_p \) we obtain the stated result.

5. EMBEDDED BERNOULLI ALGEBRAS AND L^P ESTIMATES

We begin this section by explaining how the algebra \( \mathcal{U}(Q, P) \) contains ordinary Bernoulli variables that are indexed by the intervals \((s, t)\) in a way that is consistent with the natural filtration. This fact will be used to prove the Zakai type inequalities for Clifford integrals.

Consider a finite dimensional subspace \( X \) of \( \mathcal{F} \) spanned by the indicator functions of a finite collection of non-overlapping intervals. Let \( X = X_1 \oplus X_2 \) be the orthogonal decomposition of \( X \) into two subspaces. Let \( X_1^Q \oplus X_1^F \), \( X_2^Q \oplus X_2^F \) and \( X_1^Q \oplus X_2^F \) be the corresponding subspaces of \( \mathcal{F}_Q \oplus \mathcal{F}_p^r \). Now any \( x \) in \( X_1^Q \oplus X_2^F \) can be uniquely expressed as \( x = u + v \) with \( u \) in \( X_1^Q \oplus X_1^F \) and \( v \) in \( X_2^Q \oplus X_2^F \). The map \( u + v \mapsto -u + v \) is clearly an automorphism of
the quadratic space $X^Q \oplus X^P$ whose square is the identity. Hence it induces an automorphism $\gamma_X$ of $\mathcal{E}(X^Q \oplus X^P)$.

Now, the dimension of $X^Q \oplus X^P$ is $2N$ for some integer $N$, and so $\mathcal{E}(X^Q \oplus X^P)$ is isomorphic to the algebra of all $2^N \times 2^N$ matrices. Hence, all of its automorphisms are inner, and there is an element $U_{X_1}$ of $\mathcal{E}(X^Q \oplus X^P)$, unique up to a constant multiple, so that

$$\gamma_X(A) = U_{X_1}^{-1}AU_{X_1}$$

(5.1)

for all $A$ in $\mathcal{E}(X^Q \oplus X^P)$. Since $\gamma_X^2$ is the identity, $U_{X_1}^2$ is a constant multiple of the identity. We fix the arbitrary constant multiple in the definition of $U_{X_1}$, apart from a sign, by $U_{X_1}^2 = 1$. This makes $U_{X_1}$ both unitary and self-adjoint.

Now for each $j$ indexing a finite partition of $[0, t]$, define

$$AQ_j = U_{j-1}AQ_j$$

(5.2)

where $U_{j-1}$ is the unitary corresponding to reflection in the span of $\{AQ_1, ..., AQ_{j-1}\}$ as described above. It is easy to see that all of the $AQ_j$ commute with one another. Thus the restriction of $E$ to the algebra they generate provides us with an ordinary probability space. Thus the non-commutative probability theory associated with the Clifford algebra contains everything in ordinary probability that can be built out of Bernoulli variables. A subsequent paper will develop this observation in an infinite dimensional setting in a way that permits limits to be taken.

**Proof of Theorem 6.** Let $T > 0$ and an integer $N$ be given, and let $AQ = Q(Tj/N) - Q(T(j-1)/N)$ and $AP = P(Tj/N) - P(T(j-1)/N)$ for $j = 1, ..., N$. Consider a simple process of the special form

$$G(\cdot) = \sum_{j=1}^{N} G_{AQ}^{j} 1_{(j-1)/N, j/N)}$$

(5.3)

where the $G_{AQ}^{j}$ are polynomials in $AQ_1, AP_1, ..., AQ_{j-1}, AP_{j-1}$. As these polynomials and $T$ and $N$ vary, the resulting class of processes is clearly dense for any of the norms under consideration.

Consider a fixed process $G(\cdot)$ of the form (5.2). The assertion to be proved about $G(\cdot)$ is clearly true for $N=1$. We shall prove it general by induction. Hence we define $A$ and $B$ by

$$A = \sum_{j=1}^{N-1} G_{AQ}^{j} AQ \quad \text{and} \quad B = G_{AQ}^{N} AQ_n$$

(5.3)
and make the inductive hypotheses that
\[ \|A\|_p^2 \leq (p-1) \sum_{j=1}^{N-1} (\|G_j\|_p^2 (T/N)) \quad \text{for } p \geq 2 \quad (5.4) \]

and
\[ \|A\|_p^2 \geq (p-1) \sum_{j=1}^{N-1} (\|G_j\|_p^2 (T/N)) \quad \text{for } 1 \leq p \leq 2 \quad (5.5) \]

Now let \( \hat{Q}_j \) and \( \hat{P}_j \) denote the Bernoulli variables corresponding to \( Q_j \) and \( P_j \). Then since
\[ \hat{Q}_j = (\hat{Q}_1 \hat{P}_1) \cdots (\hat{Q}_{N-1} \hat{P}_{N-1}) \hat{Q}_N \]
if we put
\[ C := G_N(\hat{P}_{N-1} \hat{Q}_{N-1}) \cdots (\hat{P}_2 \hat{Q}_2) (\hat{P}_1 \hat{Q}_1). \]

Then \( B = G_N \hat{Q}_N = C \hat{Q}_N \), and if we concretely represent these operators in the Brauer–Weyl spin-chain representation, so that they are operators on the \( N \)-fold tensor product \( \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 \), \( A \) and \( C \) operate only on the first \( N-1 \) factors, while \( \hat{Q}_N \) operates only on the last factor. Moreover, the spectrum of \( \hat{Q}_N \) is \[ \{ |T/N|^{1/2}, -(T/N)^{1/2} \} \]. Thus, the set of singular values of \( A + B \) is exactly the union of the sets of singular values of \( A + (T/N)^{1/2} C \) and \( A - (T/N)^{1/2} C \), and
\[ E(\|A + B\|_p^p) = (1/2)(\|A + (T/N)^{1/2} C\|_p^p + \|A - (T/N)^{1/2} C\|_p^p). \quad (5.6) \]

Then for \( p > 2 \), the optimal 2-uniform smoothness inequality of [BCL] yields
\[ \left( \frac{\|A + (T/N)^{1/2} C\|_p^p + \|A - (T/N)^{1/2} C\|_p^p}{2} \right)^{2/p} \leq \|A\|_p^2 + (p-1)\|C\|_p^2 (T/N). \quad (5.7) \]

Since \((P_{N-1}Q_{N-1}) \cdots (P_2Q_2)(P_1Q_1)\) is unitary, \( \|C\|_p = \|G_N\|_p \), and it then follows from (5.3) and (5.6) that
\[ \left\| \int G \ dF \right\|_p = (E(\|A + B\|_p^p))^{2/p} \leq (p-1) \sum_{j=1}^{N} (\|G_j\|_p^2 (T/N)) \]
for \( p > 2 \). Since the inequality (5.7) reverses for \( 1 < p < 2 \), we obtain the other inequality in the same manner.
REFERENCES


[TJ74] N. Tomczak-Jaegermann, The moduli of smoothness and convexity and Rademacher averages of trace classes \( S_{p}^{*} 1 \leq p < \infty \), *Studia Math.* 50 (1974), 163–182.

