<table>
<thead>
<tr>
<th>Title</th>
<th>Feynman's Operational Calculi for Noncommuting Operators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>KIM, Byoung Soo</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 (2012), 1797: 144-166</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2012-06</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/172950">http://hdl.handle.net/2433/172950</a></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>

Kyoto University
Feynman’s Operational Calculi for Noncommuting Operators

By

Byoung Soo KIM*

Abstract

We survey the author’s recent development of Jefferies and Johnson’s theory of Feynman’s operational calculi. Extraction of a linear factor in Feynman’s operational calculi can simplify disentangling under various conditions. We also introduce a first order infinitesimal calculus for a function of n noncommuting operators. Further, we consider a measure permutation formula and its applications in Feynman’s operational calculi.

§ 1. Introduction

It is important in several areas of mathematics and its applications to be able to form functions of operators. If one has a single self-adjoint operator or several commuting self-adjoint operators, the spectral theorem provides an extremely rich functional calculus. However, as soon as we have two or more noncommuting operators, the functional calculus becomes much more complicated even if the operators are self-adjoint.

Feynman’s 1951 paper [3] on the operational calculus for noncommuting operators arose out of his ingenious work on quantum electrodynamics and was inspired in part by his earlier work on the Feynman path integral. Indeed, Feynman thought of his operational calculus as a kind of generalized path integral. Much surprisingly varied work on the subject has been done since by mathematicians and physicists. References can be found in the recent books of Johnson and Lapidus [11] and Nazaikinskii, Shatalov and Sternin [13].

Recently, Jefferies and Johnson [4], [5], [6] introduced a mathematically rigorous approach to Feynman’s operational calculi. The central objects of this theory are the disentangling algebra, a commutative Banach algebra, and the disentangling map which carries the commutative structure into the noncommutative algebra of operators.

2010 Mathematics Subject Classification(s): 47A60, 46J32, 28M60
Key Words: Feynman’s operational calculi, disentangling, extraction of a linear factor, derivation formula, Newton’s formula with remainder, Taylor’s formula with remainder, measure permutation formula, commutation formula, Pauli matrices

*School of Liberal Arts, Seoul National University of Science and Technology, Seoul 139-743, Korea
This paper is a survey article on the author's recent development of the Jefferies and Johnson's theory of Feynman's operational calculi. We summarize the results in [1], [2], [7], [9], [10]. Sections 2 and 3 are concerned with the formulas which simplify disentangling in Feynman's operational calculi under various conditions. In Section 4, we explore the differential calculus associated with the disentangled operators arising from Feynman's operational calculi. Sections 5 and 6 are concerned with a measure permutation formula, which correspond to the index permutation formula in Maslov's discretized version of Feynman's operational calculus. We will not give the detailed proof and just state the properties.

We turn now to reviewing the basic definitions of the Jefferies and Johnson's theory of Feynman's operational calculi.

Given a positive integer $n$ and $n$ positive numbers $r_1, \ldots, r_n$, let $A(r_1, \ldots, r_n)$ be the space of complex-valued functions of $n$ complex variables $f(z_1, \ldots, z_n)$, which are analytic at $(0, \ldots, 0)$, and are such that their power series expansion

\begin{equation}
    f(z_1, \ldots, z_n) = \sum_{m_1, \ldots, m_n=0}^{\infty} c_{m_1, \ldots, m_n} z_1^{m_1} \cdots z_n^{m_n}
\end{equation}

converges absolutely, at least on the closed polydisk $|z_1| \leq r_1, \ldots, |z_n| \leq r_n$.

For $f \in A(r_1, \ldots, r_n)$ given by (1.1), let

\begin{equation}
    \|f\| = \|f\|_{A(r_1, \ldots, r_n)} = \sum_{m_1, \ldots, m_n=0}^{\infty} |c_{m_1, \ldots, m_n}| r_1^{m_1} \cdots r_n^{m_n}.
\end{equation}

The functions on $A(r_1, \ldots, r_n)$ defined by (1.1) and (1.2) make $A(r_1, \ldots, r_n)$ into a Banach algebra with identity under pointwise multiplication of the functions involved [4].

Let $X$ be a Banach space and $A_1, \ldots, A_n$ nonzero operators from $\mathcal{L}(X)$, the space of bounded linear operators acting on $X$. Except for the numbers $\|A_1\|, \ldots, \|A_n\|$, which will serve as weights, we ignore for the present the nature of $A_1, \ldots, A_n$ as operators and introduce a commutative Banach algebra consisting of "analytic functions" $f(\tilde{A}_1, \ldots, \tilde{A}_n)$, where $\tilde{A}_1, \ldots, \tilde{A}_n$ are treated as purely formal commuting objects.

Now consider the collection $D(A_1, \ldots, A_n)$ of all expressions of the form

\begin{equation}
    f(\tilde{A}_1, \ldots, \tilde{A}_n) = \sum_{m_1, \ldots, m_n=0}^{\infty} c_{m_1, \ldots, m_n} \tilde{A}_1^{m_1} \cdots \tilde{A}_n^{m_n}
\end{equation}

where $c_{m_1, \ldots, m_n} \in \mathbb{C}$ for all $m_1, \ldots, m_n = 0, 1, \ldots$, and

\begin{equation}
    \|f(\tilde{A}_1, \ldots, \tilde{A}_n)\| = \|f(\tilde{A}_1, \ldots, \tilde{A}_n)\|_{D(A_1, \ldots, A_n)} = \sum_{m_1, \ldots, m_n=0}^{\infty} |c_{m_1, \ldots, m_n}| \|A_1\|^{m_1} \cdots \|A_n\|^{m_n} < \infty.
\end{equation}
The space $\mathcal{D}(A_1, \ldots, A_n)$ equipped with pointwise operations and the norm (1.4) is a commutative Banach algebra with identity. In fact, if we take $\|A_j\| = r_j$ for $j = 1, \ldots, n$, then the two Banach algebras are the same except for a renaming of the indeterminants. We refer to $\mathcal{D}(A_1, \ldots, A_n)$ as the disentangling algebra associated with the $n$-tuple $(A_1, \ldots, A_n)$ of bounded linear operators acting on $X$.

Let $A_1, \ldots, A_n$ be nonzero operators from $\mathcal{L}(X)$ and $\mu_1, \ldots, \mu_n$ continuous probability measures defined at least on $\mathcal{B}[0, T]$, the Borel class of $[0, T]$ (such measures are continuous provided that each single point set has measure 0). We wish to define the disentangling mapping

$$T_{\mu_1, \ldots, \mu_n}: \mathcal{D}(A_1, \ldots, A_n) \rightarrow \mathcal{L}(X)$$

according to the rule determined by the measures $\mu_1, \ldots, \mu_n$. Putting it another way, given any analytic function $f \in \mathcal{A}(\|A_1\|, \ldots, \|A_n\|)$, we wish to form the function $f_{\mu_1, \ldots, \mu_n}(A_1, \ldots, A_n)$ of not necessarily commuting operators $A_1, \ldots, A_n$ as directed by $\mu_1, \ldots, \mu_n$.

Given nonnegative integers $m_1, \ldots, m_n$, let

$$p^{m_1, \ldots, m_n}(z_1, \ldots, z_n) = z_1^{m_1} \cdots z_n^{m_n}$$

so that

$$p^{m_1, \ldots, m_n}(\tilde{A}_1, \ldots, \tilde{A}_n) = \tilde{A}_1^{m_1} \cdots \tilde{A}_n^{m_n}.$$  

For each $m = 0, 1, \ldots$, let $S_m$ denote the set of all permutations of the integers $\{1, \ldots, m\}$, and given $\pi \in S_m$, we set

$$\Delta_m(\pi) = \{(s_1, \ldots, s_m) \in [0, T]^m; 0 < s_{\pi(1)} < \cdots < s_{\pi(m)} < T\}.$$  

We now define the mapping $T_{\mu_1, \ldots, \mu_n}$. For $j = 1, \ldots, n$ and all $s \in [0, T]$, let

$$A_j(s) = A_j$$

and for $i = 1, \ldots, m$, we define

$$C_i(s) := \begin{cases} 
A_1(s) & \text{if } i \in \{1, \ldots, m_1\}, \\
A_2(s) & \text{if } i \in \{m_1 + 1, \ldots, m_1 + m_2\}, \\
& \vdots \\
A_n(s) & \text{if } i \in \{m_1 + \cdots + m_{n-1} + 1, \ldots, m\}, 
\end{cases}$$

for all $0 \leq s \leq T$.

Note that $A_j$ in (1.9) is time independent. This implies that the integrands can be pulled outside of the integral in all of the cases involved. The measures effect the weights on the various terms as usual but, even more centrally for this theory, the measures effect the ordering of the operators (the time dependent setting has been developed elsewhere, for example, in [8], [14], [15]).
**Definition 1.1.** For any nonnegative integers \(m_1, \ldots, m_n\) and \(m = m_1 + \cdots + m_n\),
\[ T_{\mu_1, \ldots, \mu_n} P^{m_1, \ldots, m_n}(\tilde{A}_1, \ldots, \tilde{A}_n) \]
(1.11)
\[ = \sum_{\pi \in S_m} \int_{\Delta_m(\pi)} C_{\pi(m)}(s_{\pi(m)}) \cdot \cdots \cdot C_{\pi(1)}(s_{\pi(1)}) (\mu_1^{m_1} \times \cdots \times \mu_n^{m_n})(ds_1, \ldots, ds_m). \]

Then, for \(f(\tilde{A}_1, \ldots, \tilde{A}_n) \in D(A_1, \ldots, A_n)\) given by
\[ f(\tilde{A}_1, \ldots, \tilde{A}_n) = \sum_{m_1, \ldots, m_n=0}^{\infty} c_{m_1, \ldots, m_n} \tilde{A}_1^{m_1} \cdots \tilde{A}_n^{m_n}, \]
we set
\[ T_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n) = \sum_{m_1, \ldots, m_n=0}^{\infty} c_{m_1, \ldots, m_n} T_{\mu_1, \ldots, \mu_n} P^{m_1, \ldots, m_n}(\tilde{A}_1, \ldots, \tilde{A}_n). \]

We will often use the alternative notation:
\[ P_{\mu_1, \ldots, \mu_n}^{m_1, \ldots, m_n}(A_1, \ldots, A_n) = T_{\mu_1, \ldots, \mu_n} P_{\mu_1, \ldots, \mu_n}^{m_1, \ldots, m_n}(\tilde{A}_1, \ldots, \tilde{A}_n) \]
(1.14)
and
\[ f_{\mu_1, \ldots, \mu_n}(A_1, \ldots, A_n) = T_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n). \]

For elementary properties of the disentangling algebra \(D(A_1, \ldots, A_n)\) and the disentangling map \(T_{\mu_1, \ldots, \mu_n}\), see the introductory papers [4], [5], [6], [7].

\[ \S \ 2. \ \text{Extraction of a Linear Factor} \]

The operation of ‘disentangling’ is the key to Feynman’s operational calculi for noncommuting operators. Hence, formulas which simplify this procedure under various conditions are central to the subject. The main results of this section make it possible to carry out the disentangling in an iterative manner. In this section we summarize the results in [9].

Let \(A_1, \ldots, A_n\) belong to \(L(X)\), where \(X\) is a Banach space and \(\mu_1, \ldots, \mu_n\) be continuous probability measures on \(\mathcal{B}[0, T]\).

**Theorem 2.1.** Suppose that the probability measures \(\mu_1, \ldots, \mu_k\) are supported by \([a, b] \subset [0, T]\) and that the probability measures \(\mu_{k+1}, \ldots, \mu_n\) are supported by \([0, a] \cup [b, T]\). Let \(m_1, \ldots, m_n\) be nonnegative integers. Let
\[ K_{m_1, \ldots, m_k} = P_{\mu_1, \ldots, \mu_k}^{m_1, \ldots, m_k}(A_1, \ldots, A_k) \]
and \(\mu_0\) any continuous probability measure supported by \([a, b]\). Then
\[ P_{\mu_1, \ldots, \mu_n}^{m_1, \ldots, m_n}(A_1, \ldots, A_n) = T_{\mu_0, \mu_{k+1}, \ldots, \mu_n}^{1, m_k+1, \ldots, m_n}(K_{m_1, \ldots, m_k}, A_{k+1}, \ldots, A_n). \]
Theorem 2.2 (Extraction of a Linear Factor). Let $\mu_1, \ldots, \mu_n$ be given as in Theorem 2.1. Assume that $g(\tilde{A}_1, \ldots, \tilde{A}_k) \in D(A_1, \ldots, A_k)$ and $h(\tilde{A}_{k+1}, \ldots, \tilde{A}_n) \in D(A_{k+1}, \ldots, A_n)$. Let

$$(2.3) \quad f(z_1, \ldots, z_n) = g(z_1, \ldots, z_k) h(z_{k+1}, \ldots, z_n).$$

Let $K = T_{\mu_1, \ldots, \mu_k} g(\tilde{A}_1, \ldots, \tilde{A}_k)$ and $\mu_0$ any continuous probability measure supported by $[a, b]$. Then $f(\tilde{A}_1, \ldots, \tilde{A}_n) \in D(A_1, \ldots, A_n)$ and

$$(2.4) \quad T_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n) = T_{\mu_0, \mu_{k+1}, \ldots, \mu_n} F(\tilde{K}, \tilde{A}_{k+1}, \ldots, \tilde{A}_n),$$

where $F(z_0, z_{k+1}, \ldots, z_n) = z_0 h(z_{k+1}, \ldots, z_n)$.

Theorem 2.2 extends to Jefferies and Johnson's theory of Feynman's operational calculi a computational technique which is due to Maslov and which is used many times both in his book [12] and in the book by Nazaikinskii, Shatalov and Sternin [13]. The technique is referred to in [13] as "the extraction of a linear factor" and is used in conjunction with the "autonomous bracket" notation of Maslov.

The technique of extracting a linear factor is used in [12], [13] to establish equalities while doing calculations with noncommuting operators. Of course, such calculations can be organized so that the goal is to show that a related operator expression is equal to the zero operator. Our first corollary below will make it easy to apply Theorem 2.2 in the manner which we have just discussed.

Corollary 2.3. Let the hypotheses of Theorem 2.2 be satisfied and suppose that

$$(2.5) \quad K = T_{\mu_1, \ldots, \mu_k} g(\tilde{A}_1, \ldots, \tilde{A}_k) = O.$$ 

Then $T_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n)$ also equals the zero operator.

Next we give an example illustrating the use of Theorem 2.2 and Corollary 2.3. This example depends on a special relationship between $A_1, \ldots, A_k$.

Example 2.4. Let bounded operators $A_1, \ldots, A_4$ be given and suppose that $g(z_1, z_2) = z_1^3 - z_1^2 z_2 + z_1 z_2^2 - z_2^3$ and $h(z_3, z_4) \in A(\|A_3\|, \|A_4\|)$. Further, assume that $\mu_1, \ldots, \mu_4$ are continuous probability measures on $[0, T]$ with the supports of $\mu_1$ and $\mu_2$ being subsets of $[a, b]$ and the supports of $\mu_3$ and $\mu_4$ being subsets of $[0, a] \cup [b, T]$ where $0 < a < b < T$. Finally, we take $f(z_1, z_2, z_3, z_4) = g(z_1, z_2) h(z_3, z_4)$ and ask for the computation of $T_{\mu_1, \ldots, \mu_4} f(\tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \tilde{A}_4)$ in the case where $A_2 = A_1$. Since $A_2 = A_1$, $A_1$ and $A_2$ certainly commute. Hence, by Proposition 3.1 of [4], all of the functional calculi $T_{v_1, v_2}$ acting on $(\tilde{A}_1, \tilde{A}_2)$ agree with the usual commutative functional calculi. Thus,

$$T_{\mu_1, \mu_2} g(\tilde{A}_1, \tilde{A}_2) = g(A_1, A_2) = A_1^3 - A_1^2 A_2 + A_1 A_2^2 - A_2^3 = O$$

where the last equality comes from the fact that $A_2 = A_1$. It follows immediately from Corollary 2.3 that $T_{\mu_1, \ldots, \mu_4} f(\tilde{A}_1, \ldots, \tilde{A}_4)$ equals the zero operator.
Corollary 2.5. Suppose that the probability measures $\mu_1, \ldots, \mu_k$ are supported by $[0, a]$ and that the probability measures $\mu_{k+1}, \ldots, \mu_n$ are supported by $[a, T]$. Let $f, g$ and $h$ be given as in Theorem 2.2. Then

\begin{equation}
\mathcal{T}_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n) = \mathcal{T}_{\mu_{k+1}, \ldots, \mu_n} h(\tilde{A}_{k+1}, \ldots, \tilde{A}_n) \mathcal{T}_{\mu_1, \ldots, \mu_k} g(\tilde{A}_1, \ldots, \tilde{A}_k).
\end{equation}

Results in this section can be combined with various results in [4], [5], [6] to yield further corollaries.

§ 3. Methods for Iterative Disentangling

The results of this section permit us, under appropriate assumptions, to disentangle in an iterative manner. The main result 'extraction of a multilinear factor', Theorem 3.3, is a multilinear version of Theorem 2.2. Theorem 3.7 also produces a linear term at each stage of the iteration but, unlike Theorem 3.3, the new linear term depends on the previous one. In this section we summarize the results in [7].

Let $d$ be a positive integer. For each $j = 1, \ldots, d$, let $I_j$ be the nonempty subset of $I = \{1, \ldots, n\}$ such that

$I_j = \{i_{j-1} + 1, \ldots, i_j\}$

where $i_0 = 0$ and let

$I_0 = I - (I_1 \cup \cdots \cup I_d)$.

Now we introduce the abbreviated notation.

**Notation.** We write

\begin{equation}
P_{\mu_i, i \in I}^{m_i, i \in I}(A_i, i \in I) = P_{\mu_1, \ldots, \mu_n}^{m_1, \ldots, m_n}(A_1, \ldots, A_n),
\end{equation}

as well as

$f(\tilde{A}_i, i \in I) = f(\tilde{A}_1, \ldots, \tilde{A}_n), \quad \mathbb{D}(A_i, i \in I) = \mathbb{D}(A_1, \ldots, A_n),$

$f(z_i, i \in I) = f(z_1, \ldots, z_n), \quad \mathcal{T}_{\mu_i, i \in I} = \mathcal{T}_{\mu_1, \ldots, \mu_n}.$

Theorems 3.3 and 3.7 below can be established via induction and Theorem 2.2. The two monomial cases Theorems 3.1 and 3.5 could be obtained as special cases of Theorems 3.3 and 3.7 respectively. However the monomial cases are easier to understand, and so we will state them as separate theorems.

**Theorem 3.1.** Let $a_j, b_j, j = 1, \ldots, d$ be real numbers such that

\begin{equation}
0 \leq a_1 < b_1 \leq a_2 < b_2 \leq \cdots \leq a_d < b_d \leq T.
\end{equation}
Suppose that $\mu_i, i \in I_j$ have supports contained within $[a_j, b_j]$ for $j = 1, \ldots, d$. Let $\nu_j, j = 1, \ldots, d$, be any continuous probability measures having supports contained within $[a_j, b_j]$. Given nonnegative integers $m_1, \ldots, m_n$, let

$$K_j = \mathcal{T}_{\mu_i, i \in I_j}^{m_i, i \in I_j}(A_i, i \in I_j)$$

for $j = 1, \ldots, d$. Then

$$P_{\mu_1, \ldots, \mu_n}^{m_1, \ldots, m_n}(A_1, \ldots, A_n) = P_{\nu_1, \ldots, \nu_d, \mu_i, i \in I_0}^{1, \ldots, 1, m_i, i \in I_0}(K_1, \ldots, K_d; A_i, i \in I_0).$$

**Corollary 3.2.** Let $\mu_1, \ldots, \mu_n$ be given as in Theorem 3.1. Suppose $I_0 = \emptyset$, that is, $I = I_1 \cup \cdots \cup I_d$. For any nonnegative integers $m_1, \ldots, m_n$,

$$P_{\mu_1, \ldots, \mu_n}^{m_1, \ldots, m_n}(A_1, \ldots, A_n) = K_d \cdots K_1 = \mathcal{T}_{\mu_i, i \in I_d}^{m_i, i \in I_d}(A_i, i \in I_d) \cdots \mathcal{T}_{\mu_i, i \in I_1}^{m_i, i \in I_1}(A_i, i \in I_1).$$

Now we come to the theorem which allows us to iteratively disentangle a multilinear factor from $\mathcal{T}_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n)$ where $f(z_1, \ldots, z_n)$ is an appropriately factorable analytic function of $z_1, \ldots, z_n$. Note that the linear factors $K'_1, \ldots, K'_d$ defined below can be computed independently of one another; this will not be true of Theorem 3.7.

**Theorem 3.3 (Extraction of a Multilinear Factor).** Let $\mu_1, \ldots, \mu_n$ and $\nu_1, \ldots, \nu_d$ be given as in Theorem 3.1. Assume that $g_j(\tilde{A}_i, i \in I_j) \in \mathbb{D}(A_i, i \in I_j)$ for $j = 1, \ldots, d$ and $h(\tilde{A}_i, i \in I_0) \in \mathbb{D}(A_i, i \in I_0)$. Let

$$f(z_1, \ldots, z_n) = \left[ \prod_{j=1}^{d} g_j(z_i, i \in I_j) \right] h(z_i, i \in I_0)$$

and

$$K'_j := \mathcal{T}_{\mu_i, i \in I_j} g_j(\tilde{A}_i, i \in I_j)$$

for $j = 1, \ldots, d$. Then $f(\tilde{A}_1, \ldots, \tilde{A}_n) \in \mathbb{D}(A_1, \ldots, A_n)$ and

$$\mathcal{T}_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n) = \mathcal{T}_{\nu_1, \ldots, \nu_d, \mu_i, i \in I_0} F(\tilde{K}'_1, \ldots, \tilde{K}'_d; \tilde{A}_i, i \in I_0),$$

where $F(w_1, \ldots, w_d; z_i, i \in I_0) = w_1 \cdots w_d h(z_i, i \in I_0)$.

**Corollary 3.4.** Let $\mu_1, \ldots, \mu_n$ and $f(z_1, \ldots, z_n)$ be given as in Theorem 3.3. Suppose $I = I_1 \cup \cdots \cup I_d$. Then

$$\mathcal{T}_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n) = \tilde{K}'_d \cdots \tilde{K}'_1 = \mathcal{T}_{\mu_i, i \in I_d} g_d(\tilde{A}_i, i \in I_d) \cdots \mathcal{T}_{\mu_i, i \in I_1} g_1(\tilde{A}_i, i \in I_1).$$

The results in the rest of this section are consequences of the main results of Section 2. However, the supports of the clusters of measures beyond the first are in nonabutting pairs of
intervals arranged so that Theorem 2.2 is applicable and there is linearity at each stage of the disentangling but not multilinearity of the processes as a whole. The difference in the results is perhaps most clearly seen by comparing the multilinear expression \( K_d \cdots K_1 \) in formula (3.4) of Corollary 3.2 with (3.13) of Corollary 3.6 below. Perhaps the most cogent point is that the operators \( K_1, \ldots, K_d \) in Corollary 3.2 can be calculated independently of one another and in any order whereas this is not true of \( L_1, \ldots, L_d \) in Corollary 3.6.

**Theorem 3.5.** Let \( a_j, b_j, j = 1, \ldots, d \) be real numbers such that

\[
0 \leq a_d \leq \cdots \leq a_2 \leq a_1 < b_1 \leq b_2 \leq \cdots \leq b_d \leq T.
\]

Suppose that \( \mu_i, i \in I_j \) have supports contained within \([a_{j-1}, a_j] \cup [b_{j-1}, b_j]\) for \( j = 1, \ldots, d \) where \( a_0 = b_1 \) and \( b_0 = a_1 \). Let \( \eta_j, j = 1, \ldots, d \), be any continuous probability measures having supports contained within \([a_j, b_j]\). Given nonnegative integers \( m_1, \ldots, m_n \), let

\[
L_j = P_{\eta_{j-1};\mu_i, i \in I_j}^{1;m_i, i \in I_j}(L_{j-1}; A_i, i \in I_j)
\]

for \( j = 1, \ldots, d \) where \( L_0 \) is the identity operator and \( \eta_0 \) is any continuous probability measure having support contained within \([a_1, b_1]\). Then

\[
P_{\mu_1, \ldots, \mu_n}^{m_1, \ldots, m_n}(A_1, \ldots, A_n) = P_{\eta_0}^{1}(L_d; A_i, i \in I_0)
\]

**Corollary 3.6.** Let \( \mu_1, \ldots, \mu_n \) and \( \eta_1, \ldots, \eta_d \) be given as in Theorem 3.5. Suppose \( I = I_1 \cup \cdots \cup I_d \). For any nonnegative integers \( m_1, \ldots, m_n \),

\[
P_{\mu_1, \ldots, \mu_n}^{m_1, \ldots, m_n}(A_1, \ldots, A_n) = L_d = P_{\eta_d}^{1;m_i, i \in I_d}(L_{d-1}; A_i, i \in I_d)
\]

where \( L_{d-1} \) is given inductively by the formula (3.10). Equation (3.12) can be expressed more explicitly by the formula

\[
P_{\mu_1, \ldots, \mu_n}^{m_1, \ldots, m_n}(A_1, \ldots, A_n) = P_{\eta_d}^{1;m_i, i \in I_d}(P_{\eta_{d-1}}^{1;m_i, i \in I_{d-1}}(\cdots
\]

Now we come to one of the main results of this section.

**Theorem 3.7.** Let \( \mu_1, \ldots, \mu_n \) and \( \eta_1, \ldots, \eta_d \) be given as in Theorem 3.5 and \( f(z_1, \ldots, z_n) \) be given as in Theorem 3.3. For each \( j = 1, \ldots, d \), let

\[
F_{j-1}(w_{j-1}; z_i, i \in I_j) = w_{j-1}g_j(z_i, i \in I_j)
\]

and

\[
L'_j = T_{\eta_{j-1};\mu_i, i \in I_j}F_{j-1}(L'_{j-1}; \tilde{A}_i, i \in I_j)
\]
where $w_0 = 1$, $L'_0$ is the identity operator and $\eta_0$ is any continuous probability measure having support contained within $[a_1, b_1]$. Then

$$T_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n) = T_{\eta_0; \mu_i, i \in I_0} F(I'_d; A_i, i \in I_0)$$

(3.16)

where $F(w_d; z_i, i \in I_0) = w_d h(z_i, i \in I_0)$.

**Corollary 3.8.** Let $\mu_1, \ldots, \mu_n$, $\eta_1, \ldots, \eta_d$ and $f(z_1, \ldots, z_n)$ be given as in Theorem 3.7. Suppose $I = I_1 \cup \cdots \cup I_d$. Then we have

$$T_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n) = T_{\eta_d; \mu_i, i \in I_d} F_{d-1}(\tilde{L}_{d-1}; A_i, i \in I_d)$$

(3.17)

where $F_{d-1}$ and $L'_{d-1}$ are given inductively by the formulas (3.14) and (3.15), respectively.

We use the linearity of the disentangling mapping again to extend Theorem 3.3 to finite sums of functions of the form (3.5). In the language of tensor products, the extension is from elementary tensors to algebraic tensors.

**Corollary 3.9.** Let the $a_j$'s and $b_j$'s, $A_1, \ldots, A_n, \mu_1, \ldots, \mu_n$ and $v_1, \ldots, v_d$ be exactly as in Theorem 3.3. Further, let

$$f(z_1, \ldots, z_n) = \sum_{l=1}^{N} f_l(z_1, \ldots, z_n)$$

(3.18)

where each $f_l$ has the form

$$f_l(z_1, \ldots, z_n) = \prod_{j=1}^{d} g_{jl}(z_i, i \in I_j) h_l(z_i, i \in I_0)$$

(3.19)

and satisfies the conditions of Theorem 3.3. For any $l \in \{1, \ldots, N\}$, let

$$K'_{jl} := T_{\mu_i, i \in I_j} g_{jl}(\tilde{A}_i, i \in I_j)$$

(3.20)

for $j = 1, \ldots, d$. Then $f(\tilde{A}_1, \ldots, \tilde{A}_n) \in D(A_1, \ldots, A_n)$ and

$$T_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n) = \sum_{l=1}^{N} T_{\nu_1, \ldots, \nu_d; \mu_i, i \in I_0} F_l(\tilde{K}'_{1l}, \ldots, \tilde{K}'_{dl}; A_i, i \in I_0)$$

(3.21)

$$= \sum_{l=1}^{N} T_{\nu_1, \ldots, \nu_d; \mu_i, i \in I_0} [K'_{1l} \cdots K'_{dl} h_l(\tilde{A}_i, i \in I_0)],$$

where $F_l(w_1, \ldots, w_d; z_i, i \in I_0) = w_1 \cdots w_d h_l(z_i, i \in I_0)$, $l = 1, \ldots, N$. 


A variety of further corollaries could be written down using the results of this section either by themselves or in conjunction with earlier results.

§ 4. The derivation formula and higher-order expansions

In this section we summarize the results in [10]. We explore the differential or derivational calculus associated with the disentangled operators arising from Feynman’s operational calculi for noncommuting operators. We deal with a first order infinitesimal calculus for a function of \( n \) noncommuting operators. Here the first derivatives are replaced by the first order derivational derivatives. The derivational derivatives of the first and higher order have been useful in, for example, operator algebras, noncommutative geometry and Maslov’s discrete form of Feynman’s operational calculus.

Let \( A \) be a nonzero operator in \( \mathcal{L}(X) \) and let \( f(x) = \sum_{m=0}^{\infty} c_m x^m \) be an entire function. Then for any continuous probability measure \( \mu \) on \( B[0, T] \),

\[
\mathcal{T}_\mu f(A) = \sum_{m=0}^{\infty} c_m A^m
\]

and we let

\[
f(A) = \sum_{m=0}^{\infty} c_m A^m.
\]

In this section we are interested in the term of order \( \epsilon \) in the expression

\[
f(A + \epsilon B) = f(A) + \epsilon C_1 + \epsilon^2 C_2 + \cdots + O(\epsilon^n)
\]

where \( B \) is an operator in \( \mathcal{L}(X) \). The coefficient of \( \epsilon \) can be obtained as

\[
C_1 = \frac{d}{d\epsilon} f(A + \epsilon B)|_{\epsilon=0}.
\]

We will find an expression for a more general “derivative”, of which this is a particular case. First order derivations will play the central role here.

A derivation of \( \mathcal{L}(X) \) is an arbitrary linear mapping

\[
D : \mathcal{L}(X) \rightarrow \mathcal{L}(X)
\]

satisfying the Leibniz rule

\[
D(AB) = D(A)B + AD(B), \quad A, B \in \mathcal{L}(X).
\]

Derivations of the form

\[
D(A) = D_B(A) = BA - AB,
\]
where $B$ is an arbitrary element of $\mathcal{L}(X)$, are called inner derivations of $\mathcal{L}(X)$.

Now we claim that

\begin{equation}
D[f(A)] = T_{\mu, \mu} f_1(\tilde{A}, \overline{D(A)}),
\end{equation}

where $\mu$ is any continuous probability measure on $B[0, T]$ and $f_1(x,y) = f'(x)y$. In fact we will show a more general result in Theorems 4.1 and 4.3 below.

**Theorem 4.1.** For each $f(\tilde{A}_1, \ldots, \tilde{A}_n) \in D(A_1, \ldots, A_n)$ and for an inner derivation $D$ of $\mathcal{L}(X)$, we have

\begin{equation}
D[T_{\mu_1, \ldots, \mu_n} f(\tilde{A}_1, \ldots, \tilde{A}_n)] = \sum_{j=1}^{n} T_{\mu_1, \ldots, \mu_j, \mu_j, \mu_j+1, \ldots, \mu_n} F_j(\tilde{A}_1, \ldots, \tilde{A}_j, \overline{D(A_j)}, \tilde{A}_{j+1}, \ldots, \tilde{A}_n),
\end{equation}

where

\begin{equation}
F_j(x_1, \ldots, x_j, y, x_{j+1}, \ldots, x_n) = \frac{\partial}{\partial x_j} f(x_1, \ldots, x_n)y.
\end{equation}

**Lemma 4.2.** Let $D$ be a derivation of $\mathcal{L}(X)$. Define $\mathcal{D} : \mathcal{L}(\mathcal{L}(X)) \rightarrow \mathcal{L}(\mathcal{L}(X))$ by

\begin{equation}
\mathcal{D}(T) = [D, T] = DT - TD.
\end{equation}

Then $\mathcal{D}$ is an inner derivation of $\mathcal{L}(\mathcal{L}(X))$. Furthermore, for any $A \in \mathcal{L}(X)$,

\begin{equation}
\mathcal{D}(L_A) = L_{D(A)}.
\end{equation}

**Theorem 4.3.** Formula (4.6) is valid for an arbitrary derivation of $\mathcal{L}(X)$.

**Remark.** We have considered disentangling maps and derivations associated with the algebra $\mathcal{L}(X)$, where $X$ is a Banach space. In fact there is no change in the definitions of the disentangling map and the derivation if we replace $\mathcal{L}(X)$ by an arbitrary algebra $\mathcal{A}$. Moreover Theorem 4.1 through Theorem 4.3 remain true in this new setting.

Consider the algebra $\mathcal{L}(X)_{\{t\}}$ whose elements are (infinitely differentiable) families of elements of $\mathcal{L}(X)$ depending on a numerical parameter $t$. Clearly, the mapping

\begin{equation}
\frac{d}{dt} : \mathcal{L}(X)_{\{t\}} \rightarrow \mathcal{L}(X)_{\{t\}}
\end{equation}

taking each family $A_t \in \mathcal{L}(X)_{\{t\}}$ into its $t$-derivative is a derivation of the algebra $\mathcal{L}(X)_{\{t\}}$. Let $f(x) = \sum_{m=0}^{\infty} c_m x^m$ be an entire function. Then by Theorem 4.3 and its remark, we obtain

\begin{equation}
\frac{d}{dt} f(A_t) = T_{\mu, \mu} f_1(\tilde{A}_t, \overline{\frac{dA_t}{dt}}),
\end{equation}

where $\mu$ is any continuous probability measure on $\mathcal{B}[0,T]$ and $f_1(x,y) = f'(x)y$.

From (4.11) we easily derive a formula for the coefficient (4.2). Namely, take $A_\epsilon = A + \epsilon B$, then $\frac{dA}{d\epsilon} = B$ and

\begin{equation}
C_1 = T_{\mu,\mu}f_1(\tilde{A},\tilde{B}),
\end{equation}

i.e.,

\begin{equation}
f(A + \epsilon B) = f(A) + \epsilon T_{\mu,\mu}f_1(\tilde{A},\tilde{B}) + O(\epsilon^2).
\end{equation}

We will develop here some special cases of higher order expansions.

**Theorem 4.4.** Let $f(x) = \sum_{m=0}^{\infty} c_m x^m$ be an entire function. For each $n = 1, 2, \ldots$,

\begin{equation}
\frac{d^n}{dt^n} f(A + tB) = T_{\mu,\mu}f_n(\tilde{A} + tB, \tilde{B}),
\end{equation}

where $\mu$ is any continuous probability measure on $\mathcal{B}[0,T]$ and $f_n(x,y) = f^{(n)}(x)y^n$.

Thus we can write down the usual Maclaurin expansion

\begin{equation}
f(A + \epsilon B) = f(A) + \sum_{n=1}^{N} \frac{\epsilon^n}{n!} T_{\mu,\mu}f_n(\tilde{A},\tilde{B}) + O(\epsilon^{N+1}).
\end{equation}

Our next theorem will give an expansion of arbitrarily high order accompanied by an explicit remainder term.

**Theorem 4.5.** (Newton's formula with remainder) Let $A, B$ be nonzero operators in $\mathcal{L}(X)$ and let $\epsilon > 0$. Let $\mu$ be a continuous probability measure on $\mathcal{B}[0,T]$. For any positive integer $N$, we have

\begin{equation}
f(A + \epsilon B) - f(A) = \sum_{n=1}^{N} \frac{\epsilon^n}{n!} T_{\mu,\mu}f_n(\tilde{A},\tilde{B}) + Q_N,
\end{equation}

where $f_n(x,y) = f^{(n)}(x)y^n$ for $n = 1, 2, \ldots, N$ and $Q_N$ is given by the formula

\begin{equation}
Q_N = \frac{\epsilon^{N+1}}{(N+1)!} T_{\mu,\mu}g_N(\tilde{A},\tilde{B}),
\end{equation}

where $g_N(x,y) = f^{(N+1)}(x+\epsilon_1 y)y^{N+1}$ for some $0 < \epsilon_1 < \epsilon$.

**Theorem 4.6.** (Taylor's formula with remainder) Let $A, C$ be nonzero operators in $\mathcal{L}(X)$ and let $\mu_1, \mu_2$ be continuous probability measures on $\mathcal{B}[0,T]$. For any positive integer $N$, we have

\begin{equation}
f(C) - f(A) = \sum_{n=1}^{N} \frac{1}{n!} T_{\mu_1,\mu_2} h_n(\tilde{A}, \tilde{C}) + R_N,
\end{equation}
where \( h_n(x,y) = f^{(n)}(x)(y-x)^n \) for \( n = 1, 2, \ldots, N \) and \( R_N \) is given by the formula

\[
R_N = \frac{1}{(N+1)!} T_{\mu_1, \mu_2} k_N(\tilde{A}, \tilde{C}),
\]

where \( k_N(x,y) = f^{(N+1)}(x+t(y-x))(y-x)^{N+1} \) for some \( 0 < t < 1 \).

**Remark.** Unlike Newton's formula with remainder in (4.16), Taylor's formula with remainder in (4.18) does not generally provide an expression in power of \( \epsilon \) if \( C = A + \epsilon B \).

5. Measure permutation formulas

In this section we summarize the results in [2]. Let \( A \) and \( B \) belong to \( \mathcal{L}(X) \), where \( X \) is a Banach space. Let \( \mu_1 \) and \( \mu_2 \) be continuous probability measures and let \( f(x,y) = x^m y^n \) for nonnegative integer \( m \). Then, as one can expect, we have

\[
T_{\mu_1, \mu_2} f(\tilde{A}, \tilde{B}) = T_{\mu_2, \mu_1} f(\tilde{B}, \tilde{A}),
\]

that is, the measures and operators can be exchanged simultaneously for \( f(x,y) = x^m y^n \). Proposition 2.11 of [4] is a generalization of the above identity. But we can not expect that

\[
T_{\mu_1, \mu_2} f(\tilde{A}, \tilde{B}) = T_{\mu_2, \mu_1} f(\tilde{A}, \tilde{B})
\]

in general. To understand this, let us consider a simple example. Let \( f(x,y) = xy \). Then

\[
T_{\mu_1, \mu_2} f(\tilde{A}, \tilde{B}) = AB(\mu_1 \times \mu_2)[s_2 < s_1] + BA(\mu_1 \times \mu_2)[s_1 < s_2]
\]

and

\[
T_{\mu_2, \mu_1} f(\tilde{A}, \tilde{B}) = AB(\mu_2 \times \mu_1)[s_2 < s_1] + BA(\mu_2 \times \mu_1)[s_1 < s_2],
\]

where, for example, \((\mu_1 \times \mu_2)[s_2 < s_1]\) denotes \((\mu_1 \times \mu_2)((s_1, s_2) : s_2 < s_1)\). Hence the equality in (5.2) can not be true unless

\[
(\mu_1 \times \mu_2)[s_2 < s_1] = (\mu_2 \times \mu_1)[s_2 < s_1].
\]

Moreover, the conditions to ensure (5.2) are much more complicate if, for example, \( f(x,y) = x^2 y^3 \) or \( f(x,y) = xy^5 + x^2 y^2 \), etc.

In this section we are interested in the difference between the two operators \( T_{\mu_1, \mu_2} f(\tilde{A}, \tilde{B}) \) and \( T_{\mu_2, \mu_1} f(\tilde{A}, \tilde{B}) \). Our first lemma below is easy to obtain but it is useful to simplify the proof of the Measure permutation formula I in Theorem 5.2.

**Lemma 5.1.** Let \( m_1 \) and \( m_2 \) be positive integers and let \( \mu \) be a continuous probability measure on \( B[0,T] \). Let

\[
K_{m_1} = m_1 P_{\mu, \mu}^{m_1-1} (A, [A,B]).
\]
Then the identity

\[(5.4) \quad A^{m_1}B^{m_2} - B^{m_2}A^{m_1} = B^{m_2-1}K_{m_1} + B^{m_2-2}K_{m_1}B + \cdots + K_{m_1}B^{m_2-1}\]

holds.

Now we are ready to prove measure permutation formulas, the main results in this section.

Theorem 5.2. (Measure permutation formula I) Let \(\mu_1\) and \(\mu_2\) be continuous probability measures on \(\mathcal{B}[0, T]\). Assume that \(\mu_1\) is supported by \([0, a]\) and \(\mu_2\) is supported by \([a, T]\) for some \(a \in [0, T]\). Then for any positive integers \(m_1\) and \(m_2\),

\[(5.5) \quad P^{m_1, m_2}_{\mu_2, \mu_1}(A, B) = P^{m_1, m_2}_{\mu_1, \mu_2}(A, B) + m_2P^{m_2-1, 1}_{\mu, \mu}(B, K_{m_1}),\]

where \(K_{m_1}\) is given by (5.3) and \(\mu\) is any continuous probability measure on \(\mathcal{B}[0, T]\). Further assume that \(f(\tilde{A}, \tilde{B}) \in \mathcal{D}(A, B)\) and let

\[(5.6) \quad f(x, y) = \sum_{m_1, m_2=0}^{\infty} c_{m_1, m_2}x^{m_1}y^{m_2},\]

then we have

\[(5.7) \quad T^{\mu_2, \mu_1}f(\tilde{A}, \tilde{B}) = T^{\mu_1, \mu_2}f(\tilde{A}, \tilde{B}) + \sum_{m_1, m_2=1}^{\infty} c_{m_1, m_2}m_2P^{m_2-1, 1}_{\mu, \mu}(B, K_{m_1}).\]

Let \(\delta f(x, y, y_1)\) denote the difference derivative of \(f(x, y)\) with respect to \(y\), that is,

\[(5.8) \quad \delta f(x, y, y_1) = \begin{cases} \frac{f(x, y) - f(x, y_1)}{y - y_1}, & \text{if } y \neq y_1 \\ \frac{\partial^2 f}{\partial y^2}(x, y), & \text{if } y = y_1. \end{cases}\]

The infinite series of operators on the right hand side of (5.7) can be expressed as a disentangling of a function which is related with \(f\). We present it as Measure permutation formula II in the following theorem. The technique of 'extraction of a linear factor' [7, 9] plays an important role in the next theorem.

Theorem 5.3. (Measure permutation formula II) Under the assumptions of Theorem 5.2 one has

\[(5.9) \quad T^{\mu_2, \mu_1}f(\tilde{A}, \tilde{B}) = T^{\mu_1, \mu_2}f(\tilde{A}, \tilde{B}) + T_{\nu_1, \nu_2, \nu_1, \nu_3}(A, B) \frac{\delta^2 f}{\partial x \partial y}(\tilde{A}; \tilde{A}; \tilde{B}, \tilde{B}),\]

where \(\nu_1, \nu_2\) and \(\nu_3\) are continuous probability measures supported on \([0, b_1]\), \([b_1, b_2]\) and \([b_2, T]\) for any \(b_1, b_2\) with \(0 < b_1 < b_2 < T\), respectively. Moreover (5.9) can be simplified as

\[(5.10) \quad T^{\mu_2, \mu_1}f(\tilde{A}, \tilde{B}) = T^{\mu_1, \mu_2}f(\tilde{A}, \tilde{B}) + T_{\nu_2, \nu_1, \nu_1, \nu_3}(A, B) \delta f(\tilde{A}; \tilde{B}, \tilde{B}).\]
Now we give some corollaries and an example of Measure permutation formulas I and II. For $f(\tilde{A}) \in \mathbb{D}(A)$, let us denote

$$f(A) = f_{\mu}(A) = \mathcal{T}_{\mu}f(\tilde{A})$$

for any continuous probability measure $\mu$ on $\mathcal{B}[0,T]$. Our first corollary, called commutation formula, gives a relationship between $f(A)B$ and $Bf(A)$. It is an easy consequence of Theorems 5.2 or 5.3.

**Corollary 5.4. (Commutation formula I)** Let $f(\tilde{A}) \in \mathbb{D}(A)$ and let

$$(5.11) \quad f(x) = \sum_{m=0}^{\infty} c_{m}x^{m}.$$ \hspace{1cm} \text{Then we have} \hspace{1cm}

$$(5.12) \quad f(A)B = Bf(A) + \mathcal{T}_{\mu,\mu}f_{1}(\tilde{A},[A,B]),$$

where $\mu$ is any continuous probability measure on $\mathcal{B}[0,T]$ and $f_{1}(x,z) = f'(x)z$.

The following is a generalized commutation formula. If we take $g(y) = y$ in Corollary 5.5, then Corollary 5.4 is immediate from Corollary 5.5.

**Corollary 5.5. (Commutation formula II)** Let $f(\tilde{A}) \in \mathbb{D}(A)$ and $g(\tilde{B}) \in \mathbb{D}(B)$. Then we have

$$(5.13) \quad f(A)g(B) = g(B)f(A) + \mathcal{T}_{\nu_{2},\nu_{3}}(f \otimes g)_{1}(\tilde{A},[A,B];\tilde{B},\tilde{B}),$$

where $\nu_{1},\nu_{2},\nu_{3}$ are the measures as in Theorem 5.3 and

$$(5.14) \quad (f \otimes g)_{1}(x,z;y,y_{1}) = f'(x)z\frac{\delta g}{\delta y}(y,y_{1}).$$

Derivation formula (4.6) in Section 4 plays a key role in a differential calculus for Feynman's operational calculi. (For details, see Sections 3 and 4 of [10].) In the following corollary we obtain the derivation formula (4.6) as a corollary of the commutation formula (5.13) in Corollary 5.5.

**Corollary 5.6.** Let $D$ be a derivation of $\mathcal{L}(X)$ and let $f(\tilde{A}) \in \mathbb{D}(A)$. Then we have

$$(5.15) \quad D[f(A)] = \mathcal{T}_{\mu,\mu}f_{1}(\tilde{A},\overline{D(A)}),$$

where $\mu$ is any continuous probability measure on $\mathcal{B}[0,T]$ and $f_{1}(x,z) = f'(x)z$.

In our next example we illustrate the measure permutation formula for a function of the Pauli matrices.
Example 5.7. Let's consider the simplest non-commuting example of the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$ 

Let $f(\tilde{\sigma}_1, \tilde{\sigma}_2) \in D(\sigma_1, \sigma_2)$ and let $\mu_1, \mu_2$ be the measures given as in Theorem 5.2. Now let us rewrite the operator $T_{\mu_2, \mu_1} f(\tilde{\sigma}_1, \tilde{\sigma}_2)$ in the form $T_{\mu_1, \mu_2} f(\tilde{\sigma}_1, \tilde{\sigma}_2)$ so that the matrix $\sigma_1$ acts first in every terms of the expression. Since $[\sigma_1, \sigma_2] = 2i\sigma_3,$

$$T_{\mu_2, \mu_1} f(\tilde{\sigma}_1, \tilde{\sigma}_2) = T_{\mu_1, \mu_2} f(\tilde{\sigma}_1, \tilde{\sigma}_2) + 2i T_{\nu_1, \nu_2, \nu_1, \nu_3} \sigma_3 \frac{\delta^2 f}{\delta x \delta y} (\tilde{\sigma}_1, \tilde{\sigma}_1; \tilde{\sigma}_2, \tilde{\sigma}_2),$$

where $\nu_1, \nu_2, \nu_3$ are the measures given as in Theorem 5.3. Now let us investigate the second term on the right hand side of the above equation. Assume that

$$f(x, y) = \sum_{m_1, m_2 = 0}^{\infty} c_{m_1, m_2} x^{m_1} y^{m_2}.$$ 

Then we have

$$T_{\nu_1, \nu_2, \nu_1, \nu_3} \sigma_3 \frac{\delta^2 f}{\delta x \delta y} (\tilde{\sigma}_1, \tilde{\sigma}_1; \tilde{\sigma}_2, \tilde{\sigma}_2) = \sum_{m_1, m_2 = 1}^{\infty} c_{m_1, m_2} m_1 \left( \sum_{l=0}^{m_2-1} p_{v_2, v_2, v_1, v_3}^{1, m_1-1, l, m_2-l-1}(\tilde{\sigma}_3, \tilde{\sigma}_1; \tilde{\sigma}_2, \tilde{\sigma}_2) \right).$$

We can extract a linear factor to obtain

$$p_{v_2, v_2, v_1, v_3}^{1, m_1-1, l, m_2-l-1}(\tilde{\sigma}_3, \tilde{\sigma}_1; \tilde{\sigma}_2, \tilde{\sigma}_2) = p_{v_2, v_1, v_3}^{1, l, m_2-l-1}(H_{m_1}; \tilde{\sigma}_2, \tilde{\sigma}_2),$$

where

$$H_{m_1} = p_{v_2, v_2}^{1, m_1-1}(\tilde{\sigma}_3, \tilde{\sigma}_1).$$

Note that the Pauli matrices $\sigma_1$ and $\sigma_3$ satisfy

$$\sigma_1^2 = id, \quad \sigma_1 \sigma_3 + \sigma_3 \sigma_1 = O, \quad \sigma_3 + \sigma_1 \sigma_3 \sigma_1 = O.$$ 

Hence if $m_1$ is odd, then

$$H_{m_1} = \frac{1}{m_1} (\sigma_3^{m_1-1} + \sigma_3^{m_1-2} \sigma_1 + \cdots + \sigma_1 \sigma_3^{m_1-2} + \sigma_3^{m_1-1}) = \frac{1}{m_1} \sigma_3.$$ 

Similarly, if $m_1$ is even, then we know that $H_{m_1} = O.$ Hence

$$m_1 p_{v_2, v_2, v_1, v_3}^{1, m_1-1, l, m_2-l-1}(\tilde{\sigma}_3, \tilde{\sigma}_1; \tilde{\sigma}_2, \tilde{\sigma}_2) = \begin{cases} p_{v_2, v_1, v_3}^{1, l, m_2-l-1}(\tilde{\sigma}_3; \tilde{\sigma}_2, \tilde{\sigma}_2), & \text{if } m_1 \text{ is odd} \\ O, & \text{if } m_1 \text{ is even} \end{cases}.$$
and so
\[
\mathcal{T}_{v_2,v_2,v_1,v_3} \frac{\delta^2 f}{\delta x \delta y} (\sigma_1, \sigma_1; \sigma_2, \sigma_2)
\]
\[
= \sum_{m_1, m_2 = 1; m_1 = \text{odd}}^{\infty} c_{m_1, m_2} \left( \sum_{l=0}^{m_2-1} P_{v_2,v_1,v_3}^{1,l,m_2-l-1} (\sigma_3; \sigma_2, e^{i 2 \pi}) \right)
\]
\[
= \mathcal{T}_{v_2,v_1,v_3} \frac{\delta^2 f}{\delta x \delta y} (1, -1; \sigma_2, \sigma_2).
\]

Finally, we have
\[
\mathcal{T}_{\mu_2,\mu_1} f(\sigma_1, \sigma_2) = \mathcal{T}_{\mu_1,\mu_2} f(\tilde{A}, \tilde{B}) = \mathcal{T}_{\mu_2,\mu_1} f(\tilde{A}, \tilde{B}).
\]

§ 6. Generalized measure permutation formulas

In this section we summarize the results in [1]. Let $A$ and $B$ be bounded linear operators acting on $X$, where $X$ is a Banach space. Let $\mu_1$ and $\mu_2$ be continuous probability measures on $B[0, T]$. In Section 5, we introduced relationships between the two operators $\mathcal{T}_{\mu_1,\mu_2} f(\tilde{A}, \tilde{B})$ and $\mathcal{T}_{\mu_2,\mu_1} f(\tilde{A}, \tilde{B})$, and called the relationships as measure permutation formulas. But in Section 5 we considered the case where the numbers of operators and measures involved are exactly two. In this section we generalize the results for more than three operators and measures.

The technique of 'extraction of a linear factor' [7, 9] plays an important role not only in Theorem 6.2 below but also in the rest of this paper. Let's start with a simple lemma.

**Lemma 6.1.** Let $\mu$ and $\nu$ be continuous probability measures on $B[0, T]$. For any nonnegative integer $n$, we have

\[
P_{\mu,\nu}^{1,n}(A+B,C) = P_{\mu,\nu}^{1,n}(A,C) + P_{\mu,\nu}^{1,n}(B,C)
\]

Of course this lemma is an easy consequence as anyone can expect. But (6.1) may not be true if the power associated with the operator $A + B$ is greater than 1. Here is a simple example, where we take the operator $C$ to be the identity operator $I$ and $n$ to be 1. Now it is obvious to see that

\[
P_{\mu,\nu}^{1,1}(A+B,I) = (A+B)^n
\]

and

\[
P_{\mu,\nu}^{1,1}(A,I) + P_{\mu,\nu}^{1,1}(B,I) = A^n + B^n,
\]

and these two operators may not be equal unless $m = 1$. 

Theorem 6.2. (Generalized measure permutation formula I) Let $a_1, a_2, a_3$ be real numbers such that $0 \leq a_1 < a_2 < a_3 \leq T$. Suppose that $\mu_1$ and $\mu_2$ be continuous probability measures on $\mathcal{B}[0,T]$ having supports contained in $[a_1,a_2]$ and $[a_2,a_3]$, respectively. Let $\lambda$ be any continuous probability measure having support contained in $[0,a_1] \cup [a_3,T]$. Then for any positive integers $m_1, m_2$ and nonnegative integer $n$, we have

$$P_{\mu_2,\mu_1,\lambda}^{m_1,m_2,n}(A, B, C) = P_{\mu_1,\mu_2,\lambda}^{m_1,m_2,n}(A, B, C) + m_2 P_{\mu,\mu,\lambda}^{m_2-1,1,n}(B, K_{m_1}, C),$$

where $\mu$ is any continuous probability measure having supports contained in $[a_1,a_3]$ and $K_{m_1}$ is given as in (5.3). Further assume that $f(\tilde{A}, \tilde{B}, \tilde{C}) \in D(A,B,C)$ and let

$$f(x,y,z) = \sum_{m_1,m_2,n=0}^{\infty} c_{m_1,m_2,n} x^{m_1} y^{m_2} z^n.$$

Then we have

$$\mathcal{T}_{\mu_2,\mu_1,\lambda} f(\tilde{A}, \tilde{B}, \tilde{C}) = \mathcal{T}_{\mu_1,\mu_2,\lambda} f(\tilde{A}, \tilde{B}, \tilde{C}) + \sum_{m_1,m_2=1,n=0}^{\infty} c_{m_1,m_2,n} m_2 P_{\mu,\mu,\lambda}^{m_2-1,1,n}(B, K_{m_1}, C).$$

By introducing additional operators $C_1, \ldots, C_k$ and measures $\lambda_1, \ldots, \lambda_k$ we obtain the following corollary. It is obtained simply by extracting a linear factor.

Corollary 6.3. Let $a_1, a_2, a_3, \mu_1$ and $\mu_2$ be given as in Theorem 6.2. Let $\lambda_1, \ldots, \lambda_k$ be any continuous probability measures having supports contained in $[0,a_1] \cup [a_3,T]$. For any positive integers $m_1, m_2$ and nonnegative integers $n_1, \ldots, n_k$, we have

$$P_{\mu_2,\mu_1,\lambda_1,\ldots,\lambda_k}^{m_1,m_2,n_1,\ldots,n_k}(A, B, C_1, \ldots, C_k) = P_{\mu_1,\mu_2,\lambda_1,\ldots,\lambda_k}^{m_1,m_2,n_1,\ldots,n_k}(A, B, C_1, \ldots, C_k) + m_2 P_{\mu,\mu,\lambda_1,\ldots,\lambda_k}^{m_2-1,1,n_1,\ldots,n_k}(B, K_{m_1}, C_1, \ldots, C_k),$$

where $\mu$ and $K_{m_1}$ are given as in (5.3). Further assume that $f(\tilde{A}, \tilde{B}, \tilde{C}_1, \ldots, \tilde{C}_k) \in D(A,B,C_1,\ldots,C_k)$ and let

$$f(x,y,z_1,\ldots,z_k) = \sum_{m_1,m_2,n_1,\ldots,n_k=0}^{\infty} c_{m_1,m_2,n_1,\ldots,n_k} x^{m_1} y^{m_2} z_1^{n_1} \cdots z_k^{n_k}.$$

Then we have

$$\mathcal{T}_{\mu_2,\mu_1,\lambda_1,\ldots,\lambda_k} f(\tilde{A}, \tilde{B}, \tilde{C}_1, \ldots, \tilde{C}_k) = \mathcal{T}_{\mu_1,\mu_2,\lambda_1,\ldots,\lambda_k} f(\tilde{A}, \tilde{B}, \tilde{C}_1, \ldots, \tilde{C}_k) + \sum_{m_1,m_2=1,n_1,\ldots,n_k=0}^{\infty} c_{m_1,m_2,n_1,\ldots,n_k} m_2 P_{\mu,\mu,\lambda_1,\ldots,\lambda_k}^{m_2-1,1,n_1,\ldots,n_k}(B, K_{m_1}, C_1, \ldots, C_k).$$
Measure permutation formula II in Theorem 5.3 can also be generalized as follows. That is, the infinite series of operators on the right hand side of (6.4) can be expressed as a second order difference derivative of $f$.

**Theorem 6.4. (Generalized measure permutation formula II)** Under the assumptions of Theorem 6.2 one has

\[
\mathcal{T}_{\mu_{2},\mu_{1},\lambda}f(\tilde{A},\tilde{B},\tilde{C}) = \mathcal{T}_{\mu_{1},\mu_{2},\lambda}f(\tilde{A},\tilde{B},\tilde{C}) + \mathcal{T}_{\nu_{2},\nu_{2},\nu_{1},\nu_{3},\lambda}[A,B] \frac{\delta^{2}f}{\delta x \delta y} (\tilde{A},\tilde{B},\tilde{B};\tilde{C};\tilde{C}),
\]

where $\nu_{1}, \nu_{2}$ and $\nu_{3}$ are continuous probability measures supported on $[a_{1},b_{1}], [b_{1},b_{2}]$ and $[b_{2},a_{3}]$ for any $b_{1}, b_{2}$ with $a_{1} < b_{1} < b_{2} < a_{3}$, respectively. Moreover (6.8) can be simplified as

\[
\mathcal{T}_{\mu_{2},\mu_{1},\lambda}f(\tilde{A},\tilde{B},\tilde{C}) = \mathcal{T}_{\mu_{1},\mu_{2},\lambda}f(\tilde{A},\tilde{B},\tilde{C}),
\]

**Corollary 6.5.** Under the assumptions of Corollary 6.3 one has

\[
\mathcal{T}_{\mu_{2},\mu_{1},\lambda_{1},\ldots,\lambda_{k}}f(\tilde{A},\tilde{B},\tilde{C}) = \mathcal{T}_{\mu_{1},\mu_{2},\lambda_{1},\ldots,\lambda_{k}}f(\tilde{A},\tilde{B},\tilde{C}) + \mathcal{T}_{\nu_{2},\lambda_{1},\nu_{2},\nu_{1},\nu_{3}}[A,B] \frac{\delta f_{x}}{\delta y} (\tilde{A};\tilde{B};\tilde{C}),
\]

where $\nu_{1}, \nu_{2}$ and $\nu_{3}$ are the measures given as in Theorem 6.4.

In Theorem 6.2 through Corollary 6.5 we permute $\mu_{2}$ and $\mu_{1}$ where these are the first and the second measures in the definition of the disentangling map. Theorem 6.6 below says that we can permute the second and the third, or the first and the third measures to get similar formulas as in Theorem 6.4.

**Theorem 6.6. (Generalized measure permutation formula II)** Under the assumptions of Theorem 6.4 one has

\[
\mathcal{T}_{\lambda_{2},\mu_{1},\lambda_{1}}f(\tilde{A},\tilde{B},\tilde{C}) = \mathcal{T}_{\lambda_{1},\mu_{1},\lambda_{2}}f(\tilde{A},\tilde{B},\tilde{C}) + \mathcal{T}_{\nu_{2},\nu_{2},\nu_{1},\nu_{3}}[A,B] \frac{\delta f_{x}}{\delta y} (\tilde{A};\tilde{B};\tilde{C}),
\]

and

\[
\mathcal{T}_{\mu_{2},\lambda_{1},\mu_{1}}f(\tilde{A},\tilde{B},\tilde{C}) = \mathcal{T}_{\mu_{1},\lambda_{1},\mu_{2}}f(\tilde{A},\tilde{B},\tilde{C}) + \mathcal{T}_{\nu_{2},\lambda_{1},\nu_{2},\nu_{3}}[A,B] \frac{\delta f_{x}}{\delta y} (\tilde{A};\tilde{B};\tilde{C}),
\]

where $\nu_{1}, \nu_{2}$ and $\nu_{3}$ are given as in Theorem 6.4.
By extracting a linear factor, the following corollary for additional operators and measures are easy consequences of Theorem 6.6.

**Corollary 6.7.** Under the assumptions of Corollary 6.3 one has

\begin{align*}
T_{\lambda_1, \ldots, \lambda_k, \mu_2, \mu_1} f(\tilde{C}_1, \ldots, \tilde{C}_k, \tilde{A}, \tilde{B}) &= T_{\lambda_1, \ldots, \lambda_k, \mu_2, \mu_1} f(\tilde{C}_1, \ldots, \tilde{C}_k, \tilde{A}, \tilde{B}) \\
&+ T_{\lambda_1, \ldots, \lambda_k, \mu_2, \mu_1} [A, B] \frac{\delta^2 f}{\delta x \delta y} (\tilde{C}_1, \ldots, \tilde{C}_k; \tilde{A}, \tilde{A}; \tilde{B}, \tilde{B}) \\
(6.13)
&= T_{\lambda_1, \ldots, \lambda_k, \mu_2, \mu_1} f(\tilde{C}_1, \ldots, \tilde{C}_k, \tilde{A}, \tilde{B}) \\
&+ T_{\lambda_1, \ldots, \lambda_k, \mu_2, \mu_1} [A, B] \frac{\delta f_x}{\delta y} (\tilde{C}_1, \ldots, \tilde{C}_k; \tilde{A}; \tilde{B}, \tilde{B})
\end{align*}

and

\begin{align*}
T_{\mu_2, \lambda_1, \ldots, \lambda_k, \mu_1} f(\tilde{A}, \tilde{C}_1, \ldots, \tilde{C}_k, \tilde{B}) &= T_{\mu_2, \lambda_1, \ldots, \lambda_k, \mu_1} f(\tilde{A}, \tilde{C}_1, \ldots, \tilde{C}_k, \tilde{B}) \\
&+ T_{\mu_2, \lambda_1, \ldots, \lambda_k, \mu_1} [A, B] \frac{\delta^2 f}{\delta x \delta y} (\tilde{A}, \tilde{A}; \tilde{C}_1, \ldots, \tilde{C}_k; \tilde{B}, \tilde{B}) \\
(6.14)
&= T_{\mu_2, \lambda_1, \ldots, \lambda_k, \mu_1} f(\tilde{A}, \tilde{C}_1, \ldots, \tilde{C}_k, \tilde{B}) \\
&+ T_{\mu_2, \lambda_1, \ldots, \lambda_k, \mu_1} [A, B] \frac{\delta f_x}{\delta y} (\tilde{A}; \tilde{C}_1, \ldots, \tilde{C}_k; \tilde{B}, \tilde{B})
\end{align*}

where \(v_1, v_2\) and \(v_3\) are given as in Theorem 6.4.

Until now in this section, we permute two measures in the disentangling map when the two measures have supports in \([a_1, a_2]\) and \([a_2, a_3]\), respectively, and none of the other measures have supports in \([a_1, a_3]\). But in our next theorem we permute two measures when another measure have support in between the supports of the two measures.

**Theorem 6.8. (Iterative measure permutation formula)** Suppose that \(\mu_1, \mu_2\) and \(\mu_3\) are continuous probability measures having supports contained in \([0, a_1], [a_1, a_2]\) and \([a_2, T]\) with \(0 < a_1 < a_2 < T\), respectively. Let \(f(\tilde{A}, \tilde{B}, \tilde{C}) \in D(A, B, C)\). Then we have

\begin{align*}
T_{\mu_3, \mu_2, \mu_1} f(\tilde{A}, \tilde{B}, \tilde{C}) &= T_{\mu_1, \mu_2, \mu_3} f(\tilde{A}, \tilde{B}, \tilde{C}) \\
&+ T_{\mu_2, \mu_1, \mu_3} [A, B] \frac{\delta^2 f}{\delta x \delta y} (\tilde{A}; \tilde{A}; \tilde{B}; \tilde{B}; \tilde{C}) \\
(6.15)
&+ T_{\mu_2, \mu_1, \mu_3} [A, C] \frac{\delta^2 f}{\delta x \delta z} (\tilde{A}; \tilde{A}; \tilde{B}; \tilde{C}) \\
&+ T_{\mu_2, \mu_1, \mu_3} [B, C] \frac{\delta^2 f}{\delta y \delta z} (\tilde{A}; \tilde{B}; \tilde{A}; \tilde{B}; \tilde{C})
\end{align*}

where \(v_1, v_2\) and \(v_3\) are continuous probability measures having supports on \([a_1, b_1]\), \([b_1, b_2]\) and \([b_2, a_2]\) for \(b_1, b_2 \) with \(a_1 < b_1 < b_2 < a_2\), respectively.
Here are some comments on the above theorem.

**Remark.** 1. We can not permute $\mu_1$ and $\mu_3$ directly because the support of $\mu_2$ lies in between the supports of $\mu_1$ and $\mu_3$.  
2. In Theorem 6.8, we permute $\mu_3$ and $\mu_2$ in the first step, and then $\mu_2$ and $\mu_1$, and finally $\mu_3$ and $\mu_2$ to obtain $T_{\mu_1,\mu_2,\mu_3}f(\tilde{A},\tilde{B},\tilde{C})$ from $T_{\mu_3,\mu_2,\mu_1}f(\tilde{A},\tilde{B},\tilde{C})$. Of course this procedure is not mandatory. That is, we can permute $\mu_2$ and $\mu_1$ in the first step, and then $\mu_3$ and $\mu_2$, and finally $\mu_2$ and $\mu_1$. But this new steps induce the same identity as (6.15).

In Example 5.7 we applied measure permutation formula to derive an identity for a function of the Pauli matrices. Our final example is a further investigation of the example.

**Example 6.9.** Consider the simplest noncommuting example of the Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$  

Let $f(\bar{\sigma}_1,\bar{\sigma}_2) \in D(\sigma_1,\sigma_2)$ be given by $f(x,y) = \sum_{m_1,m_2=0}^{\infty} c_{m_1,m_2} x^{m_1} y^{m_2}$ and let $\mu_1,\mu_2$ be the measures supported on $[0,a]$ and $[a,T]$, respectively. In Example 5.7, we permute the measures $\mu_1$ and $\mu_2$ so that the matrix $\sigma_1$ acts first in every term of the expression, and obtained

$$(6.16) \quad T_{\mu_1,\mu_2} f(\bar{\sigma}_1,\bar{\sigma}_2) = T_{\mu_1,\mu_2} f(\bar{\sigma}_1,\bar{\sigma}_2) + 2i T_{\nu_1,\nu_2,\nu_3,\sigma_3} \frac{\delta^2 f}{\delta x \delta y} (1,-1;\bar{\sigma}_2,\bar{\sigma}_2),$$  

where $\nu_1,\nu_2$ and $\nu_3$ are continuous probability measures supported on $[0,b_1],[b_1,b_2]$ and $[b_2,T]$ for $0 < b_1 < b_2 < T$, respectively. Let us permute the measures $\nu_2$ and $\nu_3$ in the second term on the right hand side of (6.16), which are associated with the matrices $\sigma_3$ and $\sigma_2$, respectively, so that the matrix $\sigma_2$ acts first in every term of the expression. By Theorem 6.6, we have

$$(6.17) \quad T_{\nu_1,\nu_2,\sigma_3} \frac{\delta^2 f}{\delta x \delta y} (1,-1;\bar{\sigma}_2,\bar{\sigma}_2)$$

where $\eta_1,\eta_2$ and $\eta_3$ are continuous probability measures supported on $[b_1,c_1],[c_1,c_2]$ and $[c_2,T]$ for any $b_1 < c_1 < c_2 < T$, respectively, and $g$ is given by

$$g(z,y,y_1) = \frac{\delta^2 f}{\delta x \delta y} (1,-1;y,y_1).$$

Now let us investigate each terms on the right hand side of (6.17). By extracting a linear factor we know that the first term on the right hand side of (6.17) is reduced to $T_{\nu_1,\nu_2} \sigma_3 \tilde{K} = \sigma_3 K$, where

$$K = T_{\nu_1,\nu_2} \frac{\delta^2 f}{\delta x \delta y} (1,-1;\bar{\sigma}_2,\bar{\sigma}_2).$$
But since

$$\frac{\delta^2 f}{\delta x \delta y}(1, -1; y, y_1) = \sum_{m_1, m_2 = 1, m_1 = \text{odd}}^\infty c_{m_1, m_2} \left( \sum_{l=0}^{m_2-1} y_1^l y_{1}^{m_2-l-1} \right),$$

we have

$$K = \sum_{m_1, m_2 = 1, m_1 = \text{odd}}^\infty c_{m_1, m_2} m_2 \sigma_2^{m_2-1} = T_{v_2} \frac{\delta f_y}{\delta x}(1, -1; \tilde{\sigma}_2)$$

and so

$$T_{v_3, v_1, v_2} \frac{\delta^2 f}{\delta x \delta y}(1, -1; \tilde{\sigma}_2, \tilde{\sigma}_2) = T_{v_3, v_2} \frac{\delta f_y}{\delta x}(1, -1; \tilde{\sigma}_2) = T_{\mu_2, \mu_1} \frac{\delta f_y}{\delta x}(1, -1; \tilde{\sigma}_2).$$

For the second term on the right hand side of (6.17), note that \([\sigma_3, \sigma_2] = -2i \sigma_1\) and

$$\frac{\delta^2 g}{\delta z \delta y_1}(z, z_1; y; y_1, y_2) = \frac{\delta^3 f}{\delta x \delta y \delta y_1}(1, -1; y, y_1, y_2) = \sum_{m_1, m_2 = 1, m_1 = \text{odd}}^\infty c_{m_1, m_2} \left( \sum_{l=0}^{m_2-2m_2-l-2} \sum_{r=0}^{l-r-2} y_1^l y_{1}^{r} y_{2}^{m_2-l-r-2} \right).$$

Hence the term is reduced to

$$M = 2i \sum_{m_1, m_2 = 1, m_1 = \text{odd}}^\infty c_{m_1, m_2} \left( \sum_{l=0}^{m_2-2m_2-l-2} \sum_{r=0}^{l-r-2} P_{\eta_2, \eta_1, \eta_3}^{1, l, r, m_2-l-r-2}(\sigma_1, \sigma_2, \sigma_2, \sigma_2) \right).$$

We extract a linear factor to obtain

$$P_{\eta_2, \eta_1, \eta_3}^{1, l, r, m_2-l-r-2}(\sigma_1, \sigma_2, \sigma_2, \sigma_2) = P_{\eta_2, \eta_1, \eta_3}^{1, l+r, m_2-l-r-2}(\sigma_1, \sigma_2, \sigma_2, \sigma_2)$$

and so

$$M = 2i T_{\eta_2, \eta_1, \eta_3} k(\sigma_1, \sigma_2, \sigma_2) = 2i T_{v_2, v_1, v_3} k(\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_2),$$

where

$$k(z, y, y_2) = \sum_{m_1, m_2 = 1, m_1 = \text{odd}}^\infty c_{m_1, m_2} \left( \sum_{l=0}^{m_2-2m_2-l-2} \sum_{r=0}^{l-r-2} z y_1^l y_{1}^{r} y_{2}^{m_2-l-r-2} \right)$$

$$=z \sum_{m_1, m_2 = 1, m_1 = \text{odd}}^\infty c_{m_1, m_2} \left( \sum_{s=1}^{m_2} s y_{2}^{s-1} y_{2}^{m_2-s-1} \right)$$

$$=z \left( \frac{\delta^2 f}{\delta x \delta y} \right)_y(1, -1; y, y_2).$$

Finally from (6.16) we have

$$T_{\mu_2, \mu_1} f(\sigma_1, \sigma_2) = T_{\mu_1, \mu_2} f(\sigma_1, \sigma_2) + 2i T_{\mu_2, \mu_1} \sigma_3 \frac{\delta f_y}{\delta x}(1, -1; \tilde{\sigma}_2)$$

$$+(2i)^2 T_{v_2, v_1, v_3} \sigma_1 \left( \frac{\delta^2 f}{\delta x \delta y} \right)_y(1, -1; \tilde{\sigma}_2, \tilde{\sigma}_2).$$
One can proceed with these manipulations. That is, using Theorem 6.4 we can permute $v_2$ and $v_1$ in the third term of the right hand side of (6.18).

References