Holomorphic Spaces
MSRI Publications
Volume **33**, 1998

Higher-Order Hankel Forms and Commutators

RICHARD ROCHBERG

Abstract. We discuss the algebraic structure of the spaces of higher-order Hankel forms and of the spaces of higher-order commutators. In both cases we find a close relationship between the space of order $n+1$ and the derivations of the underlying algebra of functions into the space of order n.

1. Introduction and Summary

Let $\mathcal{H}^2$ be the Hardy space of the unit circle, Γ; that is, $\mathcal{H}^2$ is the space of functions, $f = \sum a_n z^n$, holomorphic on the unit disk for which

$$\|f\| = \left(\sum |a_n|^2\right)^{1/2} < \infty.$$

Such an f will be identified with its boundary values on Γ. A *Hankel form* on $\mathcal{H}^2$ is a bilinear map $B = B_b : \mathcal{H}^2 \times \mathcal{H}^2 \to \mathbb{C}$ that has the characteristic form

$$B(f,g) = \int_\Gamma fg\bar{b}. \tag{1–1}$$

Here b is the (boundary value) of a holomorphic function, the *symbol function* of B. In terms of Taylor coefficients,

$$B(f,g) = \sum_{n,k\geq 0} \hat{f}(n)\hat{g}(k)\overline{\hat{b}(n+k)}.$$

When B is viewed as acting on functions, its characteristic property is that its value only depends on the product fg. When viewed as acting on the coefficients, the characteristic property of B is that its matrix,

$$\{\beta_{n,k}\} = \{\overline{\hat{b}(n+k)}\},$$

is a Hankel matrix; that is, a matrix on $\mathbb{Z}^+ \times \mathbb{Z}^+$ whose entries only depend on the sum of the indices. The associated Hankel operator is the linear map, $\mathbf{B}$, from $\mathcal{H}^2$ to its linear dual space which takes f to the linear functional $\mathbf{B}(f)(\,\cdot\,) = B(f,\cdot)$. The analytic properties of these forms and the associated operators have been studied extensively, with much attention given to the relationship between the properties of $\mathbf{B}$ and B, and those of b.

The idea of bilinear forms given by a representation such as (1–1), and thus only depending on the product of the arguments, can certainly be extended to other function spaces. One investigation of those more general forms is in [Janson et al. 1987]. In the more general contexts operators based on expressions such as (1–1) are sometimes called *small Hankel operators*. There is another generalization, the *large Hankel operators*; the two types agree for the Hardy space.

Recently there has also been consideration of more general classes, the Hankel forms of higher type or order. For each nonnegative integer n there is a class, H^n, of Hankel forms of type n. The elements of H^1 are the traditional Hankel forms, and $H^n \subset H^{n+1}$ for each n.

An example of a Hankel form of type 2 on the Hardy space is

$$E(f,g) = \int_\Gamma f' g \bar{b}. \tag{1–2}$$

The characteristic property of such a form is that for any polynomial, p, the new bilinear form $C_p(f,g) = E(pf,g) - E(f,pg)$ is a Hankel form. On the coefficient side the matrix is of the form

$$\{e_{n,k}\} = \{n\overline{\hat{b}(n+k)}\}.$$

Thus such forms are obtained by perturbing Hankel forms in a controlled way. Higher-order forms were introduced in [Janson and Peetre 1987]. The point of view there was that certain Lie groups have irreducible representations on the Hardy and Bergman spaces. The representations on the function spaces induce representations on the associated spaces of bilinear forms. However those representations are not irreducible. An analysis shows that the Hankel forms are the simplest irreducible component of the induced representation. The higher-order Hankel forms are, by definition, the other irreducible components. Thus the space of Hilbert–Schmidt bilinear forms on, say, $\mathcal{H}^2$ can be decomposed as $\bigoplus_n (H^n \ominus H^{n-1})$. In [Janson and Peetre 1987] the basic analytic properties of these forms (boundedness, Schatten ideal membership, and so on) are worked out using a mixture of harmonic analysis and representation theory. That point of view has been taken to other contexts; see [Rosengren 1996] and the references there.

It is also possible to develop a theory of higher-order Hankel forms in the absence of a group action. That is done for general spaces of holomorphic functions in [Peetre and Rochberg 1995], but those analytical results are in a much more primitive state than the results for forms on the Hardy and Bergman spaces. Here we continue to study higher-order forms, concentrating now on the algebraic structure of the spaces H^n. Even in the simplest case of the Hardy space the analytical theory of the higher-order Hankel forms is a bit recalcitrant. However there are algebraic relations between the higher-order forms and the classical forms. It may be that the algebraic structure can be used to carry the analytical results from the classical forms to the higher-order forms. The only instance of this so far is a Kronecker theorem (Theorem 2.12 below), which perhaps doesn't really qualify as an analytical result. A more intriguing possibility of doing this is discussed in Remark 4.4 (page 176). Also, there appears to be a rich, but not well understood, relation between algebraic aspects of the theory of higher-order Hankel forms and algebraic aspects of the theory of higher-order commutators. This is particularly intriguing because the commutators considered need not be linear. Here we present analogous algebraic results for the two topics side by side (in Sections 2 and 3) as a step in developing this relation.

The results for Hankel forms are in Section 2. We will look at bilinear forms on a space, K, of holomorphic functions. We show that, roughly, for each n there is a natural identification of the quotient space H^{n+1}/H^n with the space of derivations mapping K into H^n/H^{n-1}. By iterating this result we find that, roughly, all elements H^n are built from elements of H^1 and differential operators of order at most $n-1$. This gives a new proof of the Kronecker theorem characterizing higher-order Hankel forms of finite rank which was first proved in [Rochberg 1995]. We will also be able to extend that theorem to new contexts.

In Section 3 we discuss commutators. A basic example involves the nonlinear operator Λ defined densely on $L^2(\mathbb{R})$ by $\Lambda f = f \ln |f|$. Let P be the Cauchy–Szegő projection, the orthogonal projection of $L^2(\mathbb{R})$ to $\mathcal{H}^2(\mathbb{R})$, and for any $b \in L^\infty(\mathbb{R})$ denote by M_b the operator of pointwise multiplication by b. It is easy to see that the nonlinear operator $[M_b, \Lambda] = M_b\Lambda - \Lambda M_b$ is bounded on $L^2(\mathbb{R})$. Less obvious, but also true, is that the operator $[P, \Lambda]$ is bounded on $L^2(\mathbb{R})$. This commutator is, in a sense discussed a bit more fully in Section 3, a nonlinear analog of a Hankel operator acting on $\mathcal{H}^2(\mathbb{R})$. In this analogy the boundedness of $[P, \Lambda]$ is analogous to the basic boundedness result for Hankel operators on $\mathcal{H}^2(\mathbb{R})$. Recently there has also been consideration of higher-order commutators such as $[M_b, [M_b, \Lambda]]$, $[M_b, [M_b, [M_b, \Lambda]]]$, etc., which we also consider here. These operators arise in the study of the internal structure of interpolation theory but they also have applications to classical analysis; see for instance [Cwikel et al. 1989; Iwaniec 1995; Pérez 1996]. In Section 3 we show that many of the algebraic results of Section 2 for higher-order Hankel forms have analogs for higher-order commutators. In my opinion the main conclusion of that section is not those relatively straightforward algebraic observations, rather it is the evidence of a possible systematic relation between higher-order Hankel forms and higher-order commutators. Another reason for interest in the algebraic structure of commutators is in the hope of extracting analytical information. The theory in [Milman and Rochberg 1995] proves, for instance, the boundedness on $L^2(\mathbb{R})$

of an operator whose main term is

$$R = [P, \Lambda^2] - \Lambda[P, \Lambda]. \tag{1–3}$$

This is a *higher-order* extension of the boundedness of $[P, \Lambda]$. Estimates on operators such as R have been less useful in analysis than the more elementary boundedness result for $[P, \Lambda]$. The interaction of nonlinearities makes it quite difficult to extract analytical information from estimates on R. It may be that the algebraic viewpoint can help.

The last section contains some brief further comments.

For a broader view of the topics of higher-order Hankel forms and higher-order commutators as well as for further references we refer to [Peetre and Rochberg 1995; Milman and Rochberg 1995]. For more recent work see [Rosengren 1996; Cwikel et al. $\geq$ 1997; Carro et al. 1995a; 1995b] and the references listed there.

2. Higher-Order Hankel Forms

Background and Notation. Let K be a Hilbert space of holomorphic functions defined on some domain D in $\mathbb{C}^N$. We assume that K contains a dense subalgebra, A, of bounded functions, and that for all a in A and k in K we have the norm estimate

$$\|ak\|_K \leq \|a\|_\infty \|k\|_K. \tag{2–1}$$

The choice of the Bergman space of the open unit disk as K and of $\mathcal{H}^\infty$, the algebra of bounded holomorphic functions on the disk, as A is a basic example. That is, K equals

$$A^2(\mathbb{D}, dx\,dy) = \left\{ f : f \in \operatorname{Hol}(\mathbb{D}) \text{ and } \|f\|^2 = \int_{\mathbb{D}} |f|^2 \,dx\,dy < \infty \right\}.$$

We emphasize that we do not assume that A, or even K, contains the polynomials or even the constant function. In particular, we want to include the Bergman space of the upper half-plane as an example, namely,

$$K = A^2(\mathbb{R}^2_+, dx\,dy) = \left\{ f : f \in \operatorname{Hol}(\mathbb{R}^2_+) \text{ and } \|f\|^2 = \int_{\mathbb{R}^2_+} |f|^2 dx\,dy < \infty \right\}.$$

In this case a convenient choice for A would be the polynomials in $(z+i)^{-1}$. The Bergman spaces of the disk and the half-plane and their standard weighted variants are the type of examples we have in mind. But many of the results have analogs on, for instance, the Fock space, where there is no natural choice for A that would have the norm estimates (2–1).

Let $\operatorname{Bilin}(K)$ be the space of continuous bilinear maps from $K \times K$ to $\mathbb{C}$.

Definition 2.1. For $a \in A$ and $B \in \operatorname{Bilin}(A)$, define the elements aB and Ba of $\operatorname{Bilin}(A)$ by setting, for all $f, g \in A$,

$$(aB)(f, g) = B(af, g), \qquad (Ba)(f, g) = B(f, ag).$$

Definition 2.2. For each $a \in A$, define δ_a to be the map of $\mathrm{Bilin}(A)$ to itself given by $\delta_a B = aB - Ba$. Thus

$$\delta_a B(f,g) = B(af,g) - B(f,ag).$$

We now collect some computational properties of these maps, which follow directly from the definition.

Proposition 2.3. *For all $a, b \in A$ and all $B \in \mathrm{Bilin}(A)$ we have*:

(1) $\delta_a \delta_b B = \delta_b \delta_a B$.
(2) $\delta_{ab} B = a(\delta_b B) + (\delta_a B)b = b(\delta_a B) + (\delta_b B)a$.
(3) $\delta_{ab} B = a(\delta_b B) + b(\delta_a B) - \delta_b \delta_a B$.

We now define Hankel forms and higher-order Hankel forms. In [Peetre and Rochberg 1995] several different definitions were offered; although it was clear that they agree on the standard examples, the general story is not clear. Here we use the definitions in [Peetre and Rochberg 1995] that are based on pairs (A, K).

Definition 2.4. (1) $H^0 = \{0\}$.
(2) We say $B \in \mathrm{Bilin}(K)$ is a *Hankel form* if $\delta_a B = 0$ for all $a \in A$. We denote the collection of all such forms by H^1.
(3) For $n = 2, 3, \ldots$, we define H^n, the set of Hankel forms of *type* (or *order*) n, to be the set of all $B \in \mathrm{Bilin}(K)$ such that $\delta_a B \in H^{n-1}$ for all $a \in A$.
(4) For $n = 1, 2, \ldots$, set $J^n = H^n / H^{n-1}$.

Using part (1) of Proposition 2.3 and induction it is easy to check that

$$B \in H^n \text{ if and only if } (\delta_a)^n B = 0 \text{ for all } a \in A.$$

Here are some examples. Let K be the Bergman space of the upper half plane, U, and let A be the bounded analytic functions on U. Let μ be a finite measure supported on V, a compact subset of U. Define the bilinear forms B and C on K by $B(f,g) = \int fg\, d\mu$ and $C(f,g) = B(f',g) = \int f'g\, d\mu$. The fact that V is a compact subset of U insures that both B and C are continuous. It is then immediate that $B \in H^1$, and it follows from the product rule for differentiation that $C \in H^2$. It is also clear how to continue and construct elements of all the H^n, and that similar constructions will work as long as the functions in K and their derivatives have good bounds on compact subsets of D. Part of the content of Theorem 2.5 is that, in some sense, this method of constructing elements of H^2 from elements of H^1 gives all of H^2 and similarly for the higher-order forms.

Modules. Fix K and A. The elements of A multiply the elements of K on the left and on the right; thus K is a *bimodule* over A. For $a \in A$ we will use L_a or juxtaposition to denote left multiplication by a acting on K. Thus, for $B \in \mathrm{Bilin}(K)$ and $f, g \in K$, we have $(L_a B)(f,g) = (aB)(f,g) = B(af,g)$. Similarly for right multiplication: $(R_a B)(f,g) = (Ba)(f,g) = B(f,ag)$.

For any $a, b, c \in A$, the three operations δ_a, L_b, and R_c on $\mathrm{Bilin}(K)$ all commute. Using this it is easy to check that, for each n, both left and right multiplication map each element of H^n to another element of H^n, and thus that H^n is an A-bimodule. It then follows that there is an induced bimodule structure on the quotients, J^n. Here however even more is true. We observe that, given $a \in A$ and $B \in H^n$, then $L_a B - R_a B = \delta_a B \in H^{n-1}$. Thus, using the same notation for the induced maps on J^n, we see that $L_a = R_a$ as operators on J^n; that is for, $j \in J^n$ and $a \in A$, we have $aj = ja$ as elements of J^n.

For normed spaces V and W we denote by $\mathrm{Map}(V, W)$ the space of continuous linear maps from V to W.

If X is an A-bimodule, an element $D \in \mathrm{Map}(A, X)$ is called a *derivation* if $D(ab) = aD(b) + D(a)b$ for all $a, b \in A$. We denote the space of such derivations by $\mathrm{Deriv}(A, X)$. For instance, $M = \mathrm{Map}(\mathrm{Bilin}(K), \mathrm{Bilin}(K))$ is an A-bimodule if we define left and right multiplication by

$$(am)(B) = a(mB) \quad \text{and} \quad (mb)(B) = (mB)b$$

for all $a, b \in A$, $m \in M$, and $B \in \mathrm{Bilin}(K)$. If we now define a map, D, of A into M by setting $D(a) = \delta_a$, part (2) of Proposition 2.3 implies that $D \in \mathrm{Deriv}(A, M)$.

Given $B \in \mathrm{Bilin}(K)$, we define $\Delta(B) \in \mathrm{Map}(A, \mathrm{Bilin}(K))$ by

$$\Delta(B)(a) = \delta_a B.$$

Now fix $\alpha \in A$. We define a mapping ∇_α that takes $\mathrm{Map}(A, \mathrm{Bilin}(K))$ to densely defined bilinear forms by the following rule: For $\tilde{\Delta} \in \mathrm{Map}(A, \mathrm{Bilin}(K))$,

$$\nabla_\alpha(\tilde{\Delta})(f, g) = \tilde{\Delta}(f)(\alpha, g).$$

Thus $\nabla_\alpha(\tilde{\Delta})$ is defined for $f \in A$ and $g \in K$; recall that A is dense in K. If $1 \in A$, the choice $\alpha = 1$ is a natural one to consider.

The next theorem concerns the properties of these two maps.

THEOREM 2.5. *For $n = 1, 2, \ldots$ we have*:

(1) $\Delta : J^{n+1} \to \mathrm{Deriv}(A, J^n)$.
(2) $\nabla_\alpha : \mathrm{Deriv}(A, J^n) \to J^{n+1}$.
(3) $\nabla_\alpha(\Delta(B)) = \alpha B$ *for any* $B \in J^{n+1}$, *and* $\Delta(\nabla_\alpha(D)) = \alpha D$ *for any* $D \in \mathrm{Deriv}(A, J^n)$.

REMARK 2.6. We are abusing the notation slightly when we use Δ and ∇_α for induced maps. This should cause no problem.

REMARK 2.7. Informally, and most clearly for the case $\alpha = 1$, the theorem says that $J^{n+1} = \mathrm{Deriv}(A, J^n)$.

Proof of Theorem 2.5. (1) We know that $\Delta(B)$ takes A to bilinear forms. Pick $B \in H^{n+1}$. By (3) in Proposition 2.3 we know that, for all $a, b \in A$,

$$\begin{aligned}\Delta(B)(ab) &= \delta_{ab}B = a\delta_b B + b\delta_a B - \delta_a\delta_b B\\ &= a(\Delta(B)b) + b(\Delta(B)a) - \delta_a\delta_b B.\end{aligned}$$

The second line shows that $\Delta(B)(ab)$ is in H^n. Now note that $\delta_a\delta_b B \in H^{n-1}$; and hence, as a map into the quotient J^n, $\Delta(B)$ satisfies

$$\Delta(B)(ab) = a\Delta(B)(b) + b\Delta(B)(a).$$

Recall that left and right multiplication by A agree on J^n. Hence the previous equation establishes that $\Delta(B)$ is a derivation. Finally note that, if we change B by an element of H^n, the range of $\Delta(B)$ changes by elements of H^{n-1}. Hence, as an element of J^n, the image is unchanged. Thus our map is well-defined on J^{n+1}.

(2) We start with $D \in \operatorname{Deriv}(A, J^n)$. Define B by

$$B(x,y) = \nabla_\alpha(D)(x,y) = D(x)(\alpha, y).$$

Certainly B is bilinear. Pick $a \in A$. We have

$$\begin{aligned}(\delta_a B)(x,y) &= B(ax,y) - B(x,ay)\\ &= D(ax)(\alpha,y) - D(x)(\alpha,ay)\\ &= (aD(x) + xD(a))(\alpha,y) - D(x)(\alpha,ay).\end{aligned}$$

Here we used the fact that D is a derivation and the fact that its range is J^n, a module on which left and right multiplication by A agree. We continue with

$$\begin{aligned}(\delta_a B)(x,y) &= D(x)(a\alpha,y) + D(a)(\alpha x,y) - D(x)(\alpha,ay)\\ &= \delta_a(D(x))(\alpha,y) + D(a)(\alpha x,y).\end{aligned}$$

We want to show $(\delta_a B) \in H^n$. $D(a)$ is in H^n, hence so is the mapping from (x,y) to $D(a)(\alpha x,y)$. To finish we need to shown that C, defined by $C(x,y) = \delta_a(D(x))(\alpha,y)$, is in H^{n-1}. We do this by induction on n. First note that if $n=1$ then $D(x)$ is in H^1, hence $\delta_a(D(x)) \equiv 0$ and we are fine. Suppose now that we are fine up to index $n-1$. It is direct to check that $D \in \operatorname{Deriv}(A, J^n)$ implies that the map $\delta_a \circ D$ taking x to $\delta_a(D(x))$ is in $\operatorname{Deriv}(A, J^{n-1})$. Hence, by the computations in the proof of the case $n-1$ of the theorem, we know that $\nabla_\alpha(\delta_a \circ D) \in H^{n-1}$. Unwinding the definition of ∇_α we find that this is what we needed. Finally note that, if the choice of the representative of $D(x)$ in J^n is changed, then $\delta_a(D(x))$ changes by an element of H^{n-1}, so we get the same image of $\nabla_\alpha D$ in J^{n+1}; that is, the choice of representative doesn't change the outcome of the computation.

(3) Suppose $B \in J^{n+1}$. We have

$$\nabla_\alpha(\Delta(B))(x,y) = \Delta(B)(x)(\alpha,y) = (\delta_x B)(\alpha,y) = B(\alpha x,y) - B(\alpha,xy).$$

The bilinear form that takes (x, y) to $B(\alpha, xy)$ is in H^1. Hence the previous computation shows that $\nabla_\alpha(\Delta(B)) = \alpha B$ as maps into J^{n+1}. As before, note that the outcome doesn't depend on choices made for the representative of $\Delta(B)(x)$.

Now select $D \in \operatorname{Deriv}(A, J^n)$, $a \in A$, and $x, y \in K$. We have

$$\begin{aligned}\Delta(\nabla_\alpha D)(a)(x, y) &= \delta_a(\nabla_\alpha D)(x, y) = (\nabla_\alpha D)(ax, y) - (\nabla_\alpha D)(ax, y)\\ &= (aD(x) + xD(a))(\alpha, y) - D(x)(\alpha, ay)\\ &= D(x)(\alpha a, y) + D(a)(\alpha x, y) - D(x)(\alpha, ay)\\ &= (\delta_a D(x))(\alpha, y) + \alpha D(a)(x, y).\end{aligned}$$

The second term on the right is exactly what we wanted. We are working with derivations into J^n, so we are done if we show that the bilinear map of (x, y) to $(\delta_a D(x))(\alpha, y)$ is a map into H^{n-1}. As before, this follows by induction. If $n = 1$ then $\delta_a D(x)$ is the zero functional. Then we proceed as in the end of the proof of part (2), to see that $\nabla_\alpha(\delta_a \circ D) \in H^{n-1}$. □

We now refine these calculations to develop structure theorems. When doing this we make further assumptions: that K is a space of functions of one variable and that the polynomials, P, are contained in A. The first assumption is for notational convenience. We return to the second later in the section.

THEOREM 2.8. *Suppose $P \subset A$ and $n \geq 1$. If $B \in H^n$ then*

$$B(p, g) = B(1, pg) + \sum_{j=1}^{n-1} \frac{(-1)^{j+1}}{j!} (\delta_z^n B)(p^{(j)}, g). \tag{2–2}$$

for all $p \in P$ and $g \in K$.

PROOF. First we develop a combinatoric formula for δ_f, where $f \in P$. We apply Proposition 2.3 (3) to the function z^2 and obtain $\delta_{z^2} = 2z\delta_z + \delta_z^2$. This is our starting point for an inductive proof that

$$\delta_f = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!} f^{(n)} \delta_z^n. \tag{2–3}$$

Repeated application of Proposition 2.3 (3) clearly gives

$$\delta_{f(z)} = \sum_{n=1}^{\infty} \Lambda_n(f) \delta_z^n$$

for some linear operators Λ_n. By Proposition 2.3 (3) we have

$$\delta_{zf} = z\delta_f + f\delta_z - \delta_z \delta_f; \tag{2–4}$$

this insures that

$$\Lambda_1(zf) = z\Lambda_1(f) + f. \tag{2–5}$$

We already observed that $\Lambda_1(z^2) = 2z$. The required formula for Λ_1 for the remaining monomials follows from (2–5) by induction on degree and then for all of P by linearity. Equation (2–4) also implies that

$$\Lambda_n(zf) = z\Lambda_n(f) - \Lambda_{n-1}(f)$$

for $n > 1$. We now proceed by induction on n. Suppose the formula for Λ_{n-1} is established. Thus

$$\Lambda_n(zf) = z\Lambda_n(f) + \frac{(-1)^n}{(n-1)!} f^{(n-1)}. \tag{2–6}$$

It follows easily from (2–4) that $\Lambda_n(z^k) = 0$ for $k < n$ and that $\Lambda_n(z^n) = (-1)^{n+1}$. The formula for general monomials now follows by induction using (2–6) and, again, for general polynomials by linearity. This gives us (2–3).

The theorem now follows by writing $B(p, g) = B(1, pg) - (\delta_p B)(1, g)$, applying (2–3) to $\delta_p B$, and noting that, because $B \in H^n$, the series ends after the term involving δ_z^{n-1}. □

REMARK 2.9. Equation (2–3) can be viewed as a formal Taylor series and summed, yielding $\delta_{f(z)} = f(z) - f(z - \delta_z)$. This formula can also be derived formally by writing $\delta_{f(z)} = L_{f(z)} - R_{f(z)} = f(L) - f(R) = f(z) - f(z - \delta_z)$.

REMARK 2.10. The restriction to functions of a single variable was for notational convenience. The analog of (2–3), for instance, for polynomials in two variables is, symbolically,

$$\delta_{f(z,w)} = f(z, w) - f(z - \delta_z, w - \delta_w).$$

REMARK 2.11. From an algebraic point of view, Theorem 2.8 is a complete structure theorem for Hankel forms. If we also consider topology, the situation is not clear. Suppose $n = 2$. If we restrict to polynomials, p, q, then for $B \in H^2$

$$B(p, q) = B(1, pq) + \frac{1}{2}(\delta_z B)(p', q).$$

Formally the map $C(p, q) = B(1, pq)$ is a Hankel form of type 1; that is, $\delta_f C = 0$ for any polynomial f. Likewise, $\delta_z B$ is a form of type 1; thus the form of type 2 has been represented as a linear combination of a form of type 1 and a form of type 1 composed with differentiation (which produces a form of type 2). For general n the conclusion is that every form of type n is built from forms of lower type by composing with differentiation in certain explicit ways. Algebraically this is the whole story. However, it is not clear how to obtain the continuity results for the new forms: are there estimates of the form $|B(1, pq)| \le c\,\|p\|_K\,\|q\|_K$ and $|(\delta_z B)(p', q)| \le c\,\|p\|_K\,\|q\|_K$? (Because each form is a bounded perturbation of the other, the two estimates are equivalent.) It is not clear if this issue can be settled algebraically or needs analytical work. We return to this issue in Remark 4.4, where we settle a minor variation of this question for $n = 2$.

We now turn to forms of finite rank. We will say a bilinear form, B, is of *rank* k if there are $2k$ continuous linear functionals, $h_1, \dots, h_k, j_1, \dots, j_k$, such that

$$B(f,g) = \sum_{i=1}^{k} h_i(f) k_i(g)$$

for any $f, g \in K$. We denote the set of all such forms by $\mathcal{F}_k$ and write $\mathcal{F} = \bigcup_n \mathcal{F}_n$. If ζ is a point in D for which all the functionals $h_i(f) = f^{(i)}(\zeta)$ are continuous then $B(f,g) = f^{(i)}(\zeta) g^{(n-i)}(\zeta)$ is an example of a form in $H^{n+1} \cap \mathcal{F}_1$. The content of the next theorem is, roughly, that these are the only examples.

We will use the following convexity and continuity hypothesis on K.

Convexity hypothesis. For each point $\zeta \notin D$ and each $M > 0$ there is a polynomial $p \in K$ so that $\|p\|_K = 1$ and $|p(\zeta)| > M$.

Continuity hypothesis. For each $\zeta \in D$ and each integer n the functional which takes $f \in K$ to $f^{(n)}(\zeta)$ is continuous.

Theorem 2.12. *Suppose $P \subset K$ and that K satisfies the two preceding hypotheses. Given $B \in H^n \cap \mathcal{F}_k$ for some $n, k \geq 1$, it is possible to find constants $M = M(n,k)$ and $C = C(n,k)$, points $\zeta_1, \dots, \zeta_M \in D$, and scalars $c_{i,j,m}$ such that*

$$B(f,g) = \sum_{\substack{i \leq n-1 \\ j \leq C,\, m \leq M}} c_{i,j,m} (f^{(i)} g)^{(j)} (\zeta_m). \tag{2–7}$$

Remark 2.13. The form that takes (f,g) to $(f^{(i)}g)^{(j)}(\zeta_m)$ is of type $i+1$, independently of j. This can be verified by an elementary induction on i. Also note that it suffices to use only these asymmetric representations because, for example, $fg' = (fg)' - f'g$.

Proof of Theorem 2.12. We work by induction on n. The case $n = 1$ is in [Janson et al. 1987, Section 14]. The basic point there is that by restricting from $K \times K$ to $P \times P$ we end up with a problem about ideals in polynomial rings, which can then be analyzed using tools from commutative algebra. We continue to restrict our attention to the forms acting on $P \times P$ and will next show that forms, B, on $P \times P$ that satisfy $\delta_a^n B = 0$ for every $a \in P$ have representations of the form (2–7). Suppose $B \in H^2 \cap \mathcal{F}$. By Theorem 2.8,

$$B(f,g) = B(1, fg) + (\delta_z B)(f', g).$$

The maps $(f,g) \mapsto B(f,g)$ and $(f,g) \mapsto (\delta_z B)(f', g)$ are both in $\mathcal{F}$, hence so is the form $C(f,g) = B(1, fg)$. Thus, in terms of their action on polynomials, both C and $\delta_z B$ are in $H^1 \cap \mathcal{F}$. (That is, they are annihilated by δ_z; we are not considering continuity.) Hence the result for $n = 1$, for forms acting on $P \times P$, can be applied to both of them. This gives the required form. The rank of C is at most the sum of the ranks of B and $\delta_z B$ and hence at most $3k$; thus we keep control of the ranks. Now we suppose that, as a form on $P \times P$, we have

$B \in H^3 \cap \mathcal{F}$; the argument here will make the general induction step clear. By Theorem 2.8,

$$B(f,g) = B(1,fg) + (\delta_z B)(f',g) + c(\delta_z^2 B)(f'',g)$$

(with $c = -\frac{1}{2}$). We can write $C(f,g) = B(1,fg)$ as a sum of three terms, all in $\mathcal{F}$; hence $C(f,g)$ is in $\mathcal{F}$ and we can estimate its rank in terms of k. We now apply the case $n = 1$ of the theorem to C and to $\delta_z^2 B$, and the case $n = 2$ to δB, to obtain the required form. Clearly this approach will also deal with the general induction step.

We now need to check that the ζ_i's are in D. Suppose $\zeta_1 \notin D$. Suppose N is the highest-order derivative of f that is evaluated at ζ_1 in the representation. Pick $r \in P$ so that r vanishes at $\zeta_2, \dots, \zeta_M$ to order higher than the order of any derivative being evaluated at ζ_j, for $j \geq 2$, and vanishes to order $N-1$ at ζ_1. Pick $g \in P$ so that neither g nor any of its first N derivatives vanish at ζ_1. Let p be the polynomial whose existence is insured by the Convexity Hypothesis, which has unit norm and is large at ζ_1. We have $B(rp,g) = c(r,g,B)p(\zeta_1)$. This isn't compatible with the boundedness estimate

$$\bigl|B(rp,g)\bigr| \leq c_B \,\|rp\|\,\|g\| \leq c'_B \,\|p\|\,\|g\| = c'_B \,\|g\|.$$

Now that we know that all the points are in D, the representation extends to all of K by continuity, using the Continutity Hypothesis. □

Remark 2.14. Again the restriction to functions of a single variable is for notational convenience. The formulation of the more general result can be seen in [Rochberg 1995]. Those proofs are more computational than the ones here.

The requirement in Theorem 2.12 that the polynomials be dense precludes such basic examples as $K = A^2(\mathbb{R}^2_+)$, the Bergman space of the upper half-plane. We now show how to use the previous result (or rather its proof) to obtain results for $K = A^2(\mathbb{R}^2_+)$. Similar arguments could be used for other simple examples, for instance the spaces $A^2(\mathbb{R}^2_+; y^\alpha dx\,dy)$ with $\alpha > -1$, but the story for complicated choices of K is not clear.

Theorem 2.15. *Let* $K = A^2(\mathbb{R}^2_+)$ *and suppose* $B \in \mathrm{Bilin}(K)$. *If* $B \in H^n \cap \mathcal{F}_k$ *for some* $n, k \geq 1$ *then* B *has a representation of the type* (2–7).

Proof. B is of finite rank; hence we can find linear functionals h_i, k_i such that $B(f,g) = \sum_1^k h_i(f)k_i(g)$, and we may assume that the k_i are linearly independent. Set

$$V = \bigl\{f : B(f,g) = 0 \text{ for all } g \in K\bigr\}.$$

Because the k_i are linearly independent, $V = \bigcap \ker(h_i)$. Let α be a conformal map of the upper half-plane to the unit disk, and set

$$W = \bigl\{f : f, \alpha f, \dots, \alpha^{n-1} f \in V\bigr\} = \bigcap_{i,j} \bigl\{f : h_i(\alpha^j f) = 0\bigr\}.$$

From this it is clear that W is a closed subspace of K of finite codimension, $\dim(K/W) \leq kn$. We now claim that W is invariant under multiplication by α. Pick $f \in W$; we need to show that $B(\alpha^n f, g) = 0$ for all $g \in K$. To see this note that $B \in H^n$, so $\delta_a^n(B) = 0$. Hence

$$B(\alpha^n f, g) = \sum_{k=0}^{n-1} c_k B\big(\alpha^{n-1-k} f, \alpha^{k+1} g\big).$$

Because $f \in W$ the right hand side is 0. This shows that $\alpha W \subset W$.

Let X be the closed subalgebra of $\mathcal{H}^\infty(\mathbb{R}^2_+)$ generated by 1 and α. Let

$$Y = \{\beta \in X : \beta f \in W \text{ for all } f \in K\}.$$

It is not clear at first glance that Y has any nonzero elements, but we now show that it is rather large. Certainly Y is a closed ideal in X. We will show that Y contains a polynomial in α. Pick $f \in K \setminus W$. Because W has codimension at most nk, there is a polynomial, p_1, of degree $n_1 \leq nk$, such that $p_1(\alpha) f \in W$. Let $W_1 = \operatorname{span}\{f, \alpha f, \ \ldots, , \alpha^{n_1 - 1} f\}$. If $W \oplus W_1 \neq K$ we continue, picking $g \in (W \oplus W_1) \setminus W$. As before there is a polynomial, now p_2 of degree n_2, such that $p_2(\alpha) g \in W \oplus W_1$. Set $W_2 = \operatorname{span}\{g, \alpha g, \ \ldots, , \alpha^{n_2 - 1} g\}$. This process must eventually fill K and we will have

$$K = W \oplus \bigoplus_i W_i. \tag{2–8}$$

Set $Q_1(\alpha) = \prod p_i(\alpha)$. If $f \in K$ then $Q_1(\alpha) f \in W$. To see this, split f using the decomposition (2–8) as $f = g + g_1 + g_2 + \cdots + g_j$. It is enough to look at each summand, and g_2 is typical: multiplication by $p_2(\alpha)$ takes g_2 into $W \oplus W_1$, and hence further multiplication by $p_1(\alpha)$ takes the product into W. Because W is α-invariant, multiplication by the remaining factors of Q_1 does no harm. Y is an ideal and we have now seen that $Q_1 \in Y$. Hence $Q_1 X \subset Y$. The characteristic property of Q_1 is that $B(Q_1 f, g) = 0$ for all $f, g \in K$. In exactly the same way we can find Q_2 such that $B(f, Q_2 g) = 0$ for all $f, g \in K$. Set $Q = Q_1 Q_2$.

We now split K as an orthogonal direct sum

$$K = R \oplus Q(\alpha) K.$$

The fundamental property of this splitting is that, for any $r_1, r_2 \in R$ and f_1, f_2 in K,

$$B\big(r_1 + Q(\alpha) f_1, \ r_2 + Q(\alpha) f_2\big) = B(r_1, \ r_2). \tag{2–9}$$

Suppose that $Q(\alpha)$ vanishes to order n_i at ζ_i in $\mathbb{R}^2_+$, for $i = 1, \ldots, N$, and that this is a complete listing of the zeros of $Q(\alpha)$ in $\mathbb{R}^2_+$. If $k \in K$ and K vanishes at each ζ_i to order at least n_i then $k = Q(\alpha) h$ for some $h \in K$. Hence for each polynomial S there is a unique $k(S)$ in R that agrees with S to order n_i at ζ_i, for $i = 1, \ldots, N$. We now define a bilinear form on polynomials by setting

$$\tilde{B}(S, T) = B\big(k(S), k(T)\big).$$

Let γ be an element of $\mathcal{H}^\infty(\mathbb{R}^2_+)$ that agrees with the monomial Z to order n_i at ζ_i, for $i = 1, \dots, N$. For any polynomial U, we have

$$k(ZU) \equiv \gamma k(U) \pmod{Q(\alpha)K}.$$

Hence, taking into account (2–9),

$$\tilde{B}(ZS, T) - \tilde{B}(S, ZT) = B\big(\gamma k(S), k(T)\big) - B\big(k(S), \gamma k(T)\big).$$

Hence, for any polynomial U, and for all S, T,

$$(\delta^n_U \tilde{B})(S, T) = (\delta^n_{U(\gamma)} B)\big(k(S), k(T)\big).$$

Since $B \in H^n$, the expression on the right is always zero. This shows that $\tilde{B}$ is a Hankel form of type n on polynomials. The proof of Theorem 2.12 then shows that $\tilde{B}$ has the form (2.8), and certainly the ζ's that show up in (2.8) must be among our ζ_i's. Now note that if $k_i \in K$, for $i = 1, 2$, then $k_i = k(S_i) + Q(\alpha)h_i$ for some polynomials S_i and some $h_i \in K$. Also k_i, $k(S_i)$, and S_i all agree to order n_j at each ζ_j. Moreover,

$$B(k_1, k_2) = B\big(k(S_1), k(S_2)\big) = \tilde{B}(S_1, S_2).$$

We have seen that the expression on the right is a linear combination of values and derivatives of the S_i of a sort given by (2.8). Hence the expression on the left must be the same combination of values and derivatives of the k_i. That B has such a representation is what we wanted to show.

To finish we note that the degree of Q is controlled by n and k, and hence so is the number of terms in (2.8). □

3. Commutators

In this section we develop analogs of the results in the previous section for certain nonlinear operators. That we worked with bilinear forms before and now work with operators is not a major change in point of view. Those results for bilinear forms could have been formulated as results about operators. Given a bilinear form B on K there is an induced linear map T_B from K to its dual space K', given by $T_B(f)(g) = B(f, g)$. (We don't want to identify the Hilbert space K with its dual using the inner product because that map is conjugate linear rather than linear.) If we have an algebra A of functions that act on K by multiplication, we can consider the corresponding multiplication operators $M_a(f) = af$. Everything in the previous section that was formulated in terms of operators that send B to aB, Ba, and $aB - Ba$ can be recast in terms of the operators that send T_B to $T_B M_a$, $M_a^* T_B$, and $T_B M_a - M_a^* T_B$.

Before going further we mention why one might look for such analogies between the results in the previous sections and commutators.

We'll work on $L^2(\mathbb{R})$. For $f \in L^2$ set $\Lambda_1(f) = f \ln |f|$. For any function a defined on $\mathbb{R}$ let M_a be the operator of multiplication by a. Pick an unbounded function b in BMO($\mathbb{R}$). (We need nothing here about BMO except that it contains unbounded functions—for instance, $b(x) = \ln|x|$). Let $\Lambda_2 = M_b$. Let P be the Cauchy–Szegő projection acting on $L^2(\mathbb{R})$. For any operators acting on functions, linear or not, we define the commutator $[A, B]$ of A and B by

$$[A, B](f) = (AB - BA)(f) = A(B(f)) - B(A(f)).$$

It is elementary that both Λ_1 and Λ_2 are unbounded, that if a is a bounded function then the operator $[M_a, \Lambda_1]$ is bounded on $L^2(\mathbb{R})$, and that $\bigl[M_a, [M_a, \Lambda_1]\bigr] = \bigl[M_b, \Lambda_2\bigr] = 0$. Also true, but not elementary, is that $\bigl[\Lambda_1, P\bigr]$, $\bigl[\Lambda_2, P\bigr]$, and $\bigl[\Lambda_2, [\Lambda_2, P]\bigr]$ are all bounded on $L^2(\mathbb{R})$. The boundedness of all three can be given a unified proof using interpolation theory. (I don't think that $\bigl[\Lambda_1, [\Lambda_1, P]\bigr]$ is bounded. The story there is complicated by the nonlinearities. Set $\Omega f = \frac{1}{2} f\bigl(\ln|f|\bigr)^2$. The operator that is bounded is $[P, \Omega] - \Lambda_1[P, \Lambda_1]$.) To see how this is related to the previous section we define Hankel forms on $\mathcal{H}^2(\mathbb{R}) = P\bigl(L^2(\mathbb{R})\bigr)$ by the natural analog of (1–1), namely

$$B(f, g) = \int_{\mathbb{R}} f g \bar{b}. \tag{3–1}$$

Now $f, g \in \mathcal{H}^2(\mathbb{R})$ and b is the boundary value of a function holomorphic in the upper half-plane. It is direct to check that

$$B(f, g) = \bigl\langle g, \overline{(I - P)(\bar{b} f)} \bigr\rangle.$$

Hence the properties of B are all to be found in the theory of the linear operator that takes f to $(I - P)(\bar{b} f)$. Now note that if we choose for b in the definition of Λ_2 the function $\bar{b}$ in (3–1), we have

$$[\Lambda_2, P]f = \bar{b} P f - P(\bar{b} f) = \bar{b} f - P(\bar{b} f) = (I - P)(\bar{b} f).$$

Thus, from good information about $[\Lambda_2, P]$ we can derive equally good information about the Hankel operator with symbol b. Conversely, for any choice of b, the commutator $[\Lambda_2, P]$ can be written as an orthogonal direct sum of two operators, one a Hankel operator and the other unitarily equivalent to a Hankel operator. In sum, the theories of Hankel operators on $\mathcal{H}^2(\mathbb{R})$ and the theory of the operators $[\Lambda_2, P]$ are essentially equivalent. This suggests that the results for higher-order Hankel forms may have analogs for commutators (including possibly the nonlinear ones involving Λ_1, Ω, and related operators). The relations between the two topics, both those mentioned and those we develop in this section, suggest that there may be deeper connections.

(In another direction, the implication for Hankel operators and their generalizations of the boundedness of $\bigl[\Lambda_2, [\Lambda_2, P]\bigr]$, as well as higher-order commutators such as $\bigl[\Lambda_2, [\Lambda_2, [\Lambda_2, P]]\bigr]$ that are also bounded, is not clear.)

Another reason for suspecting that there may be a connection between the two topics is a similarity between some of the detailed computations that led to the results in [Janson and Peetre 1987] and those that led to the results in [Rochberg 1996].

Finally, the theory of Hankel forms is related to the study of bilinear forms on vector spaces that have additional structure; namely, there is a notion of pointwise products. The Hankel forms are those that only depend on the product of its arguments. The higher-order Hankel forms represent, in some sense, infinitesimal perturbations away from that situation PR. The commutators we consider arise in interpolation theory in ways related to infinitesimal changes in Banach space structure. The commutators that arise when working with the L^p scale involve the infinitesimal versions of changes which respect the multiplicative structure but not the linear structure. We return to this rather vague comment in Remark 4.5.

The setup. Our main interest here is in developing algebraic properties. We will be quite informal about the type of continuity possessed by the various maps considered.

Let X be a space of functions (on some space that we generally won't bother to specify) and let A be an algebra of functions with the property that, for $a \in A$ and $x \in X$, we have $ax \in X$ and $\|ax\|_X \le C\,\|a\|_A \|x\|_X$. For instance, A could be the bounded holomorphic functions on the disk and X the Bergman space of the disk, or A could be $L^\infty(\mathbb{R})$ and $X = L^2(\mathbb{R})$. We also assume $A \cap X$ is dense in X. (We could get by with less, but for now we just want to capture the basic examples.)

Suppose Λ is a map from X to functions and that $a \in A$. We *do not* assume that Λ is linear. We define left and right multiplication as before: for all $a \in A$ and $f \in X$, we set

$$
\begin{aligned}
(L_a\Lambda)(f) &= (a\Lambda)(f) = a(\Lambda f),\\
(R_a\Lambda)(f) &= (\Lambda a)(f) = \Lambda(af),\\
\delta_a\Lambda &= L_a\Lambda - R_a\Lambda.
\end{aligned}
$$

Continuing the analogy with the previous section we say that $\Lambda \in L^1$ if $\delta_a\Lambda = 0$ for all $a \in A$, and for $n = 1, 2, \dots$ we say that $\Lambda \in L^{n+1}$ if $\delta_a\Lambda \in L^n$ for all $a \in A$. If we are looking at linear operators this would be the setup for an operator version of the higher-order Hankel forms. The point now is that many of the algebraic results of the previous section still go through without the assumption of linearity.

Elementary properties. The following facts, valid for $a, b, c \in A$ and $\Lambda \in L^n$, follow immediately from the definitions:

(1) $L^n \subset L^{n+1}$.
(2) $L_a\Lambda$, $R_a\Lambda \in L^n$, and $\delta_a\Lambda \in L^{n-1}$.
(3) L_a, R_b, and δ_c commute.
(4) $\delta_a\delta_b = \delta_b\delta_a$.

(5) $\delta_{ab} = L_a\delta_b + R_b\delta_a = L_b\delta_a + R_a\delta_b$.
(6) $\delta_{ab} = a\delta_b + b\delta_a - \delta_b\delta_a$.
(7) $\Omega \in L^n$ if and only if $\delta_a^n\Omega = 0$ for all $a \in A$.

L^{n+1} and L^n are vector spaces and A-bimodules. Hence the quotient spaces, $I^{n+1} = L^{n+1}/L^n$, are A-bimodules and the induced action of A on the I^n is commutative; that is, for $a \in A$ and $i \in I^n$, we have $ai = ia$ as elements of I^n.

Having collected all of these facts, which are analogous to Proposition 2.3 and related facts in the proceeding section, we should note that the situation is not always that simple. Recall that a linear map D of A to itself is called a derivation if $D(ab) = aD(b) + bD(a)$ for all $a, b \in A$. Given such a D and given $\Lambda \in L^1$, is $\Lambda D \in L^2$? (This construction with the choice $D(f) = f'$ was used frequently in the previous section.) We compute

$$\begin{aligned}(\delta_a(\Lambda D))(f) &= a(\Lambda D)(f) - (\Lambda D)(af) = a\Lambda(Df) - \Lambda(D(af)) \\ &= \Lambda(aDf) - \Lambda(fDa + aDf).\end{aligned}$$

If Λ is linear we can continue with

$$(\delta_a(\Lambda D))(f) = \Lambda(aDf) - \Lambda(fDa) - \Lambda(aDf) = -Da\Lambda(f).$$

Thus $\delta_a(\Lambda D)$ is in L^1. Thus we see that the linearity of D is irrelevant but that this process will construct an element of L^2 from $\Lambda \in L^1$ only if Λ is linear. To go even further and show ΛD^2 is in L^3 we would also need D to be linear.

Examples. Having just seen that composition with derivations doesn't necessarily generate elements of L^n for $n > 1$, we now show how to generate elements of L^1 and, more generally, all the L^n.

PROPOSITION 3.1. *Suppose A is dense in X and suppose D is a derivation on A. For $n = 1, 2, \ldots$ the map Λ_n given by*

$$\Lambda_n(f) = f\Big(\frac{Df}{f}\Big)^n$$

is an element of L^{n+1}.

REMARK 3.2. Of course Λ_n is only defined on the dense subspace A and for each f there is a problem on the set where $f = 0$. We put aside these issues and concentrate on formal structure.

REMARK 3.3. We *do not* assume that D is linear.

PROOF. The case $n = 0$ is trivial. For $n = 1$ we compute

$$(\delta_a\Lambda_1)(f) = (\delta_a D)(f) = D(af) - aD(f) = aD(f) + D(a)f - aD(f) = D(a)f.$$

As required, $\delta_a\Lambda_1$ is an element of L^1. It is straightforward to complete the proof by induction. However, with an eye to later discussion, we take a slightly less direct route.

For $a \in A$ and $t \in \mathbb{R}$ we define

$$T_t(a) = a\exp(t(\frac{Da}{a})).$$

LEMMA 3.4. *For all $a, b \in A$ we have $T_t(ab) = T_t(a)T_t(b)$.*

PROOF.

$$\begin{aligned} T_t(ab) &= ab\exp\Big(t\Big(\frac{D(ba)}{ab}\Big)\Big) = ab\exp\Big(t\Big(\frac{aDb + bDa}{ab}\Big)\Big) \\ &= ab\exp\Big(t\Big(\frac{Db}{b} + \frac{Da}{a}\Big)\Big) = ab\Big(\exp t\Big(\frac{Da}{a}\Big)\Big)\Big(\exp t\Big(\frac{Db}{b}\Big)\Big) \\ &= T_t(a)T_t(b), \end{aligned}$$

as required. □

We write

$$T_t(a) = \sum \Gamma_n(a)t^n.$$

We now equate powers of t in the proposition. That gives

$$\Gamma_n(ab) = \sum c_k\Gamma_{n-k}(a)\Gamma_k(b) = a\Gamma_n(b) + \sum_{k=0}^{n-1} c_k\Gamma_{n-k}(a)\Gamma_k(b).$$

We now complete an inductive proof of the proposition. We need to show $\Gamma_n \in L^{n+1}$, and the case $n = 1$ is done. Suppose the cases up to $n - 1$ are done.

$$\begin{aligned} (\delta_a\Gamma_n)(b) &= a\Gamma_n(b) - \Gamma_n(ab) \\ &= a\Gamma_n(b) - \left(a\Gamma_n(b) + \sum_{k=0}^{n-1} c_k\Gamma_{n-k}(a)\Gamma_k(b)\right) \\ &= \sum_{k=0}^{n-1} c_k\Gamma_{n-k}(a)\Gamma_k(b). \end{aligned}$$

By the induction hypothesis this is of the required form, and the proposition is proved. □

To use the proposition to construct examples we need examples of operators D that are not assumed linear but satisfy

$$D(fg) = fD(g) - gD(f). \tag{3–2}$$

Suppose that D is of the form $D(f) = f\varphi(\ln(f))$ for some operator φ. To verify (3–2) we compute

$$D(fg) = fg\varphi(\ln(fg)) = fg\varphi(\ln(f) + \ln(g)).$$

Hence we will have the required form if φ is any linear operator. Similarly if $D(f) = f\varphi_1(\ln|f|)$ then (3–2) will be satisfied if φ_1 is linear.

Here are some examples of elements of L^n.

EXAMPLE 3.5. If D is a derivation, the proposition tells us that the map of f to $D(f)$ is in L^2.

EXAMPLE 3.6. Suppose A consists of holomorphic functions and set $\varphi(\ln f) = (\ln(f))'$. We have $\Lambda_1(f) = f'$. That example was central in the previous sections, as were the powers $\Lambda_1^n(f) = f^{(n)}$. However, the fact that the powers have properties which we can work with rests on the linearity of Λ_1.

For $n > 1$ the previous computations give us expressions that are in the L^{n+1}; we have $\Lambda_n(f) = f^{-n+1}(f')^n$. We don't know of other places where these operators arise.

EXAMPLE 3.7. Suppose $D(f) = f\varphi_1(\ln|f|) = f\ln|f|$. We have $\Lambda_n(f) = f(\ln|f|)^n$. This series of examples, starting with $\Lambda_1(f) = f\ln|f|$, is used to form some of the basic examples the nonlinear commutators that arise in interpolation theory. For instance, with $X = L^2(\mathbb{R})$ and P the Cauchy–Szegő projection, the operator

$$2[P,\Lambda_2] - \Lambda_1[P,\Lambda_1] \tag{3–3}$$

is bounded.

EXAMPLE 3.8. The previous two examples are related to the themes of this section and the preceding one. As soon as we move to other examples we encounter rather unfamiliar nonlinear operators. For instance, suppose $D(f) = f\varphi(\ln(f))$, with $\varphi(\ln f) = (\ln(f))''$; then

$$\Lambda_0(f) = f, \quad \Lambda_1(f) = \frac{ff'' - f'^2}{f}, \quad \ldots.$$

Derivations. The results in the previous section relating J^{n+1} to derivations into J^n used the fact that the J's are linear spaces but not the fact that the elements of the spaces were linear forms. Hence the results go through for I's. We need to change the details slightly because we are now dealing with operators rather than forms and we need to be attentive to the fact that the elements of the I's may be nonlinear. But the similarities are very strong, so we will be quick.

Let M denote the set of maps from X to functions. Given $\Lambda \in M$ we define a map from A to M by

$$\Delta(\Lambda)(a) = \delta_a\Lambda.$$

Pick and fix $\alpha \in A$. We define an operator, ∇_α, that takes a mapping, $\tilde{\Delta}$, of A into M to an element of M by

$$\nabla_\alpha(\tilde{\Delta})(f) = \tilde{\Delta}(f)(\alpha).$$

Again, if $1 \in A$ the choice $\alpha = 1$ is a natural one to consider.

We continue to denote by $\operatorname{Deriv}(A, M)$ the derivations of A into an A-module M.

THEOREM 3.9. *The following results hold for* $n = 0, 1, 2, \ldots$:

(1) Δ *maps* I^{n+1} *into* $\operatorname{Deriv}(A, I^n)$.
(2) ∇_α *maps* $\operatorname{Deriv}(A, I^n)$ *into* I^{n+1}.
(3) $\nabla_\alpha(\Delta(\Lambda)) = -\alpha\Lambda$ *for all* $\Lambda \in I^{n+1}$, *and* $\Delta(\nabla_\alpha(D)) = -\alpha D$ *for all* $D \in \operatorname{Deriv}(A, I^n)$.

REMARK 3.10. Again we abuse notation slightly, in using Δ and ∇_α for induced maps.

PROOF OF THEOREM 3.9. (1) Pick $\Lambda \in L^{n+1}$. For all $a, b \in A$, we have

$$\begin{aligned}\Delta(\Lambda)(ab) &= \delta_{ab}\Lambda = a\delta_b\Lambda + b\delta_a\Lambda - \delta_a\delta_b\Lambda \\ &= a(\Delta(\Lambda)b) + b(\Delta(\Lambda)a) - \delta_a\delta_b\Lambda.\end{aligned}$$

The second equality shows that $\Delta(\Lambda)(ab)$ is in L^n. Now $\delta_a\delta_b\Lambda \in L^{n-1}$, so

$$\Delta(\Lambda)(ab) = a\Delta(\Lambda)(b) + b\Delta(\Lambda)(a)$$

as maps into I^n. Thus $\Delta(\Lambda)$ is a derivation. Finally, if we change Λ by an element of L^n, the range of $\Delta(\Lambda)$ changes by elements of L^{n-1}; hence, as an element of I^n, the image is unchanged. Thus our map is well-defined on I^{n+1}.

(2) We start with $D \in \operatorname{Deriv}(A, I^n)$. Define Λ by

$$\Lambda(x) = \nabla_\alpha(D)(x) = D(x)(\alpha).$$

Pick $a, x \in A$. Then

$$\begin{aligned}(\delta_a\Lambda)(x) &= a\Lambda(x) - \Lambda(ax) = aD(x)(\alpha) - D(ax)(\alpha) \\ &= aD(x)(\alpha) - aD(x)(\alpha) - xD(a)(\alpha) = -xD(a)(\alpha).\end{aligned}$$

We now recall that the product xD is the product in the A-module $\operatorname{Deriv}(A, I^n)$. Thus $xD(a)(\alpha) = D(a)(\alpha x)$, which is an element of I^n applied to x. This shows that the map goes into L^{n+1}. The verification that the coset in I^{n+1} is unchanged if the choice of $D(x)$ is changed by an element of I^{n-1} is routine.

(3) Suppose $\Lambda \in I^{n+1}$. We have

$$\begin{aligned}\nabla_\alpha(\Delta(\Lambda))(x) &= \Delta(\Lambda)(x)(\alpha) = (\delta_x\Lambda)(\alpha) \\ &= x\Lambda(\alpha) - \Lambda(\alpha x) = x\Lambda(\alpha) - R_\alpha\Lambda(x).\end{aligned}$$

The operator that takes x to $x\Lambda(\alpha)$ is in L^1. Hence the previous computation shows that $\nabla_\alpha(\Delta(\Lambda)) = -R_\alpha\Lambda$ as maps into I^{n+1}. As before, the outcome doesn't depend on choices made for the representative of $\Delta(\Lambda)(x)$. Finally recall that I^n is a commutative A-module and hence $R_\alpha\Lambda = L_\alpha\Lambda = \alpha\Lambda$, as required.

Now select $D \in \operatorname{Deriv}(A, I^n)$ and $a, x \in A$. Recalling again that aD is the module product, we have

$$\begin{aligned}\Delta(\nabla_\alpha D)(a)(x) &= \delta_a(\nabla_\alpha D)(x) = a(\nabla_\alpha D)(x) - (\nabla_\alpha D)(ax) \\ &= a(D(x)(\alpha)) - D(ax)(\alpha) = D(x)(a\alpha) - (aD(x) + xD(a))(\alpha) \\ &= (\delta_a D(x))(\alpha) - \alpha D(a)(x).\end{aligned}$$

As in the previous section, an easy induction shows that the first term is of lower order and hence drops out when we pass to quotients. □

4. Further Remarks

Remark 4.1. Some of the results in the previous sections involved derivations of an algebra of functions A into an A-bimodule. Such results can be reformulated in terms of module cohomology. At this point we don't see a direct use for that viewpoint, so we only offer the observation. Background can be found in [Ferguson 1996] and the references there.

Remark 4.2. A *Foguel-type operator* is an operator of the form

$$R(X) = \begin{bmatrix} S^* & X \\ 0 & S \end{bmatrix},$$

where S is the unilateral shift on $l^2(\mathbb{Z}^+)$ and X is an operator on $l^2(\mathbb{Z}^+)$. Such operators have been considered in the investigation of polynomially bounded operators (see [Davidson and Paulsen 1997] for details and further references). One reason is that if $X = \Gamma_f$, the Hankel operator on $l^2(\mathbb{Z}^+)$ with symbol function f, then polynomials in $R(X)$ are particularly easy to compute. If p is a polynomial then

$$p(R(\Gamma_f)) = \begin{bmatrix} S^* & \Gamma_f \; p'(S) \\ 0 & S \end{bmatrix} = \begin{bmatrix} p(S^*) & \Gamma_{fp'} \\ 0 & p(S) \end{bmatrix}. \tag{4–1}$$

This follows from an elementary induction and the fact that Hankel operators satisfy $S^*\Gamma - \Gamma S = 0$. Let L_{S^*} denote multiplication on the left by S^* and R_S denote multiplication on the right by S. Set $\delta_z = L_{S^*} - R_S$. For any polynomial p we use the obvious extension of the notation and set $\delta_{p(z)} = L_{p(S^*)} - R_{p(S)}$. We are now in the notational set-up of Section 2, adapted to operators. The Hankel operators are exactly those X for which $\delta_{p(z)}(X) = 0$ for all polynomials p. (Equivalently, $\delta_z(X) = 0$.) To describe the general pattern, for any polynomial p in one variable define a new polynomial by

$$p^*(x, y) = \frac{p(x) - p(y)}{x - y}.$$

A quick induction on the degree of monomials shows that, for any p,

$$p(R(X)) = \begin{bmatrix} p(S^*) & p^*(L_{S^*}, R_S)(X) \\ 0 & p(S) \end{bmatrix}.$$

Writing $L_{S^*} = \delta_z + R_S$ we get

$$\begin{aligned} p(R(X)) &= \begin{bmatrix} p(S^*) & \sum_{n=1}^{\infty} p^{(n)}(R_S)\delta_z^{n-1}(X) \\ 0 & p(S) \end{bmatrix} \\ &= \begin{bmatrix} p(S^*) & \sum_{n=1}^{\infty} \delta_z^{n-1}(X) p^{(n)}(S) \\ 0 & p(S) \end{bmatrix}. \end{aligned}$$

The linear operators induced by Hankel forms of order k are exactly those X for which $\delta_z^k(X)=0$. For those X the infinite series ends and we have a slightly more complicated but still quite explicit formula for composition with polynomials. For instance, if X is a Hankel operator of order 2 we find

$$p(R(X)) = \begin{bmatrix} p(S^*) & Xp'(S)+\frac{1}{2}\delta_z(X)p''(S) \\ 0 & p(S) \end{bmatrix}.$$

REMARK 4.3. In Section 3 we defined classes of operators L^n by $\Lambda \in L^n$ if $\delta_u^n(\Lambda)=0$ for all u in some class of functions. However that discussion was motivated by the study of commutators that arise in interpolation theory and the results in interpolation theory also suggest another, slightly more sophisticated point of view. We mention it here because it raises a number of interesting questions. We don't pursue it further here because it seems less amenable to formal algebraic analysis than the ideas in Section 3.

The commutator results which arise in interpolation theory involve operators Λ acting on a space X. The operators are generically unbounded and often nonlinear. The basic results are that for certain linear operators T which are bounded on X it is also true that the commutator $[T,\Lambda]$ is bounded on X. Often the class of T for which this holds includes all operators of the form M_u for bounded functions u. However, the construction of Λ involves a number of choices and in [Cwikel et al. $\geq$ 1997], for instance, the operators Λ are really viewed modulo bounded operators. Furthermore, even when it is true that $\delta_u(\Lambda)=[M_u,\Lambda]$ is bounded it is generally not true that $\delta_u^2(\Lambda)=0$. This suggests considering bounded operators as being of type 1 and defining an operator to be of type n if $\delta_u^{n-1}(\Lambda)$ is a bounded operator for each bounded function u. This would come closer to the viewpoint of [Cwikel et al. $\geq$ 1997] and the results from interpolation theory generate a variety of type 2 operators that are not type 1. For example, if X is an L^p space and φ is any Lipschitz function, $\Lambda(f)=f\varphi(\ln|f|)$ will be of type 2 (in this sense) and not of type 1. However it is not generally true that $\delta_u^2(\Lambda)=0$, although it is true in the special case $\varphi(x)=x$.

The difficulty with this approach is that it is not clear how to generate objects of type 3 other than with the ideas used in Section 3. That is, the examples $\Lambda_n(f)=f(\ln|f|)^n$ give examples of objects of arbitrary type in Section 3. It is not clear what other examples we would have for the variation just described.

In this context we should mention the very interesting results of Kalton [1988], which say, roughly, that in some circumstances if Λ has the property that $[M_u,\Lambda]$ is bounded for each bounded function u then $[T,\Lambda]$ is also bounded for a much larger of linear operators T.

A final word on interpolation. As we mentioned in the introduction it may be that the computations in Section 3 can be used to extract information from the boundedness of (3–3) and the more complicated combinations in [Milman and Rochberg 1995].

Remark 4.4. We mentioned after Theorem 2.8 that a flaw in this representation is that we can't insure that the individual summands are bounded. For $n = 2$ we can do that if we accept a more complicated, but also more symmetrical, formulation. Suppose $B \in H^2$. By Theorem 2.8 we know that, for $p, q \in P$,

$$B(p,q) = B(1,pq) + \tfrac{1}{2}(\delta_z B)(p',q).$$

A similar argument in the second variable gives

$$B(p,q) = B(pq,1) - \tfrac{1}{2}(\delta_z B)(p,q').$$

Adding gives

$$2B(p,q) = B(pq,1) + B(1,pq) + (\delta_z B)(p',q) - (\delta_z B)(p,q'),$$

which we rewrite as

$$2B(p,q) = B(pq,1) + B(1,pq) + (\delta_D(\delta_z B))(p,q).$$

Computing $(\delta_p \delta_q B)(1,1)$ and recalling that $B \in H^2$ we get

$$B(pq,1) + B(1,pq) = B(p,q) + B(q,p). \tag{4–2}$$

Writing $B_{\mathrm{sym}}(p,q) = B(p,q) + B(q,p)$ we have

$$2B(p,q) = B_{\mathrm{sym}}(p,q) + (\delta_D(\delta_z B))(p,q). \tag{4–3}$$

which is the representation we wanted. Clearly

$$|B_{\mathrm{sym}}(p,q)| \le C\|p\|\,\|q\|$$

and hence

$$|(\delta_D(\delta_z B))(p,q)| \le C\|p\|\,\|q\|.$$

Using (4–2) we see that $B_{\mathrm{sym}} \in H^1$. Thus (4–3) gives a representation of B as a type 1 form plus a form built from a type 1 form by composing with differentiation. All the forms are continuous.

Unfortunately it is not clear how to continue this analysis to higher n.

Remark 4.5. We have mentioned that higher-order Hankel forms are related to deformation of multiplicative structure. One way to formulate this is to introduce the family of operators M_ε defined by $M_\varepsilon(f,g) = f(z+\varepsilon)g(z-\varepsilon)$. If we expand this in powers of ε we get

$$\sum B_n(f,g)\varepsilon^n = fg + (f'g - gf')\varepsilon + \cdots.$$

Hankel forms are those that only depend on $B_0(f,g) = fg$; Hankel forms of type 2 are linear functions of $B_0(f,g)$ and $B_1(f,g)$; and so on.

In Section 3 we introduced the operator $T_t(a) = a\exp(t(Da/a))$ and noted that $T_t(ab) = T_t(a)T_t(b)$ (Lemma 3.4). In the basic example from interpolation theory, when $D(f) = f \ln |f|$, this becomes $T_t(f) = f|f|^t$. The reason this operator plays a crucial role in interpolation theory is that

$$\|f\|_{L^p} = \|T_t(f)\|_{L^{p/(1+t)}}.$$

The crucial fact is that T_t moves functions through the scale of spaces without changing the norm. For general scales of spaces the operators Λ that arise are the derivative of a family of operators with similar properties. Thus, from the interpolation theoretic point of view, it is just an accident that the formula for $D(f) = f \ln |f|$ can be used to generate a multiplicative operator. However, this fact is basic to our algebraic computations.

Question 4.6. Can the algebraic approach in Section 2 be extended to trilinear forms in a natural way? Cobos, Kühn, and Peetre have developed a theory of trilinear forms acting on a Hilbert space in [Cobos et al. 1992].

Acknowledgements

My thanks to Nigel Kalton and to Mario Milman for interesting comments and useful suggestions while I was developing these ideas.

References

[Carro et al. 1995a] M. J. Carro, J. Cerdà, and J. Soria, "Commutators and interpolation methods", *Ark. Mat.* **33**:2 (1995), 199–216.

[Carro et al. 1995b] M. J. Carro, J. Cerdà, and J. Soria, "Higher order commutators in interpolation theory", *Math. Scand.* **77**:2 (1995), 301–319.

[Cobos et al. 1992] F. Cobos, T. Kühn, and J. Peetre, "Schatten–von Neumann classes of multilinear forms", *Duke Math. J.* **65**:1 (1992), 121–156.

[Cwikel et al. 1989] M. Cwikel, B. Jawerth, M. Milman, and R. Rochberg, "Differential estimates and commutators in interpolation theory", pp. 170–220 in *Analysis at Urbana* (Urbana, IL, 1986–1987), vol. II, edited by E. Berkson et al., London Math. Soc. Lecture Note Ser. **138**, Cambridge Univ. Press, Cambridge, 1989.

[Cwikel et al. $\geq$ 1997] M. Cwikel, M. Milman, and R. Rochberg, "A unified approach to derivation mappings Ω for a class of interpolation methods". In preparation.

[Davidson and Paulsen 1997] K. R. Davidson and V. I. Paulsen, "Polynomially bounded operators", *J. Reine Angew. Math.* **487** (1997), 153–170.

[Ferguson 1996] S. H. Ferguson, "Polynomially bounded operators and Ext groups", *Proc. Amer. Math. Soc.* **124**:9 (1996), 2779–2785.

[Iwaniec 1995] T. Iwaniec, "Integrability theory of the Jacobians", lecture notes, Bonn University, 1995.

[Janson and Peetre 1987] S. Janson and J. Peetre, "A new generalization of Hankel operators (the case of higher weights)", *Math. Nachr.* **132** (1987), 313–328.

[Janson et al. 1987] S. Janson, J. Peetre, and R. Rochberg, "Hankel forms and the Fock space", *Rev. Mat. Iberoamericana* **3**:1 (1987), 61–138.

[Kalton 1988] N. J. Kalton, *Nonlinear commutators in interpolation theory*, Memoirs of the American Mathematical Society **385**, Amer. Math. Soc., Providence, 1988.

[Milman and Rochberg 1995] M. Milman and R. Rochberg, "The role of cancellation in interpolation theory", pp. 403–419 in *Harmonic analysis and operator theory* (Caracas, 1994), edited by S. A. M. Marcantognini et al., Contemp. Math. **189**, Amer. Math. Soc., Providence, RI, 1995.

[Peetre and Rochberg 1995] J. Peetre and R. Rochberg, "Higher order Hankel forms", pp. 283–306 in *Multivariable operator theory* (Seattle, WA, 1993), edited by R. E. Curto et al., Contemp. Math. **185**, Amer. Math. Soc., Providence, 1995.

[Pérez 1996] C. Pérez, "Sharp estimates for commutators of singular integrals via iterations of the Hardy–Littlewood maximal function", preprint, Universidad Autónoma de Madrid, 1996.

[Rochberg 1995] R. Rochberg, "A Kronecker theorem for higher order Hankel forms", *Proc. Amer. Math. Soc.* **123**:10 (1995), 3113–3118.

[Rochberg 1996] R. Rochberg, "Higher order estimates in complex interpolation theory", *Pacific J. Math.* **174**:1 (1996), 247–267.

[Rosengren 1996] H. Rosengren, "Multilinear Hankel forms of higher order and orthogonal polynomials", preprint, Lund University, 1996. Available at http://www.maths.lth.se/matematiklu/personal/hjalmar/engHR.html.

Richard Rochberg
Department of Mathematics
Campus Box 1146
Washington University
St. Louis MO, 63130
United States
rr@math.wustl.edu