HIGHER ORDER ESTIMATES IN COMPLEX INTERPOLATION THEORY

RICHARD ROCHBERG

Suppose $\{A_\theta\}_{0\leq\theta\leq1}$ is a scale of Banach spaces generated from the couple (A_0, A_1) by complex interploation. If T is a linear operator which is bounded on the couple then T is bounded on the entire scale. Also, associated to the scale is an operator Ω_1 which is generally nonlinear and unbounded on $A_{1/2}$ such that the commutator $[T, \Omega_1]$ is bounded on $A_{1/2}$. Here we extend the construction and produce a sequence $\Omega_2, \Omega_3, \ldots$ which are increasingly nonlinear and unbounded but such that certain combinations with T are bounded. The first example is $[T, \Omega_2] - \Omega_1[T, \Omega_1]$.

1. Background and Introduction.

For some time the author and others have explored second order estimates obtained in interpolation theory. (See, for instance, [RW], [JRW], [CJMR], [M1], [MS], [K1], and [K2].) Roughly, associated with an interpolation scale of Banach spaces is an operator Ω which is generally unbounded on spaces of the scale and which is often nonlinear. However, if T is a linear operator which is bounded on the scale then the commutator $[T, \Omega]$ will also be bounded on these spaces.

It was noted in the final comments in [RW] that in the context of complex interpolation it was possible to go further, to obtain third order and higher order estimates. However it was not clear at that time how to organize those computations and results effectively. It is now clear how to give relatively compact presentation of those results: the systematic presentation of those higher order estimates, both in the abstract and in the particular instances of classical function spaces, is the main content of this paper.

When these ideas are particularized to classical operators on classical function spaces we obtain results about boundedness of multiterm nonlinear expressions. The individual terms are unbounded and the final boundedness is due to a subtle cancellation. (See, for instance, Corollaries 6.1, 6.5, and 6.6.) This is similar to what was found in [RW] and what is found in recent work on compensated compactness [CLMS]. Here, however, the degree of unboundedness of the individual terms is greater as is the degree of nonlinearity.

The results on second order estimates are known to be related to recent work on compensated compactness. For more on that see [M2], [M3], and [IS]. We suspect that this relation persists for these higher order estimates but we have no specific evidence.

Here, roughly, are the contents. In §2 we recall what we need of complex interpolation theory and present the basic construction of our generalized interpolation spaces $X_0^{(n)}$. In §3 we study the norm on $X_0^{(n)}$ and some of its subspaces. We also describe the norm using the nonlinear operator Ω and its higher order generalizations. In §4 we prove a duality theorem for $X_0^{(n)}$. In §5 we prove an interpolation theorem for $X_0^{(n)}$ and from that we derive the commutator theorems. In §6 we particularize our general results to the special cases of weighted L^2 spaces, L^p spaces, and Banach lattices. A brief final section contains a discussion of possible further research.

My thanks to Mario Milman with whom I discussed these ideas as they were being developed.

2. The construction of $X_0^{(n)}$.

We will work in the context of the theory of complex interpolation of families of Banach spaces as introduced in [CCRSW1] and [CCRSW2]. We now recall that formalism as presented in [RW]. (For the reader who is not familiar with this extension of classical complex interpolation theory we note that if we restrict attention to boundary data which is constant on two halves of the unit circle then by conformally mapping the strip to the unit disk we recapture the classical complex interpolation theory of A. P. Calderón. The proofs here adapt to that context.)

We will present our results for families of finite dimensional spaces but we will not allow factors in our estimates unless they are independent of the families of spaces being considered, and in particular are independent of the dimension of the spaces being considered. All that we say extends to appropriate families of infinite dimensional spaces but the proofs then require a distracting layer of technicalities. However the details in [CCRSW2], which, in the language we will introduce here, show how to construct $X_0^{(0)}$ in the infinite dimensional context would extend to allow construction of $X_0^{(n)}$ in that context. (An alternative scheme for working with infinite dimensional function spaces is in [K1].)

We begin with a family $\mathcal{F}$ of N-dimensional Banach spaces parameterized by the points $e^{i\theta}, 0 \leq \theta \leq 2\pi$, of Γ, the unit circle. Thus, corresponding to $e^{i\theta}$ we have X_θ which is normed by $\| \cdot \|_\theta$. Of course we suppose this is

measurable in θ. We denote the Euclidean norm on $\mathbb{C}^N$ by $|\cdot|$. Set

$$A_\theta = \sup_{v \in \mathbb{C}^N} \frac{\|v\|_\theta}{|v|}$$

$$a_\theta = \inf_{v \in \mathbb{C}^N} \frac{\|v\|_\theta}{|v|}.$$

We make the technical assumption

$$(2.1) \qquad \int_\Gamma |\log(A_\theta/a_\theta)|\, d\theta < \infty.$$

Let $H_{\mathcal{F}}^\infty$ be the space of $\mathbb{C}^N$-valued holomorphic functions, F, on the unit disk which have radial boundary values $a.e.$ (also denoted by F) and for which

$$\|F\| = \operatorname{ess\,sup} \|F(e^{i\theta})\|_\theta < \infty.$$

Under the assumption (2.1) the correspondence between the functions and their boundary values is perfect, exactly as for traditional H^∞.

For $F \in H_{\mathcal{F}}^\infty$ write $F = \sum \widehat{F}(n) z^n$. Pick and fix a positive integer n. Let $X_0^{(n)}$ be the vector space $(\mathbb{C}^N)^n$ given the following norm. For $v_0, v_1, \ldots, v_{n-1} \in \mathbb{C}^N$ set $V = (v_0, \ldots, v_{n-1})$. Set

$$(2.2) \qquad \|V\|_{X_0^{(n)}} = \inf\left\{\|F\|; F \in H_{\mathcal{F}}^\infty, \widehat{F}(j) = v_j, j = 0, \ldots, n-1\right\}.$$

The subscript "0", which we will often omit, is to remind us that we are working with the coefficients of the expansion of F about 0. As in [**RW**] we could also consider other base points.

A normal families argument shows that the infimum in (2.2) is attained. It will also be convenient to insure that this extremal function is unique. One way to do this is to require that

(2.3)   for a set of θ of positive measure the Banach spaces X_θ

are uniformly convex.

This will be a standing assumption for the rest of the paper. In this case the extremal function is unique and we denote it by $A^{(n)}(v_0, \ldots, v_{n-1}) = A^{(n)}(V)(z)$. Hence

$$\|V\|_{X_0^{(n)}} = \left\|A^{(n)}(V)\right\|_{H_{\mathcal{F}}^\infty}.$$

We write

$$(2.4) \qquad A^{(n)}(V)(z) = \sum_{j=0}^\infty \Omega_{j,n}(V) z^j.$$

Thus
$$\Omega_{j,n}(V) = v_j, \qquad j = 0, 1, \ldots, n-1.$$

For convenience we set $\Omega_j = \Omega_{j,j}$. (In the absence of (2.3) $A^{(n)}$ is not well defined. However much of what follows remains true if we add "for any choice of extremal function.")

3. The description of the norm.

In this section we show how to estimate the norm on $X^{(n)}$ in terms of the norm in $X^{(0)}$ and the operators $\Omega_{j,k}$ introduced in (2.4). Throughout this section and the remainder of the paper we will write $\| \cdot \|$ for the norm on $X^{(0)}$:

$$\|a\| = \|a\|_{X^{(0)}}.$$

Note that $X^{(0)}$ is the interpolation space (at the origin) associated with the boundary data X_θ.

Suppose n and k are given positive integers. We wish to define three maps; the shift $S_{n,k}$, the extension $E_{n,k}$ and the restriction $R_{n,k}$. $S_{n,k}$ maps $X^{(n)}$ to $X^{(n+k)}$. For $V = (v_0, \ldots, v_{n-1}) \in X^{(n)}$

(3.1) $$S_{n,k}(V) = (\underbrace{0, \ldots, 0}_{k}, v_0, \ldots, v_{n-1}).$$

$E_{n,k}$ maps $X^{(n)}$ to $X^{(n+k)}$. For $V = (v_0, \ldots, v_{n-1}) \in X^{(n)}$

(3.2) $$E_{n,k}(V) = (v_0, \ldots, v_{n-1}, \Omega_{n,n}(V), \ldots, \Omega_{n+k-1,n}(V)).$$

$R_{n,k}$ maps $X^{(n+k)}$ to $X^{(n)}$. For $W = (w_0, \ldots, w_{n+k-1}) \in X^{(n+k)}$

(3.3) $$R_{n,k}(W) = (w_0, \ldots, w_{n-1}).$$

Proposition 3.1.
   (1) $S_{n,k}$ *is an isometry.*
   (2) $E_{n,k}$ *is an isometry.*
   (3) $R_{n,k}$ *is a contraction.*

Proof. To prove (1) it suffices to verify that for V in $X^{(n)}$

$$z^k A^{(n)}(V)(z) = A^{(n+k)}(S_{n,k}V)(z).$$

This follows at once from the fact that $|z^k| = 1$ on Γ, the definitions, and the fact that $A^{(n+k)}(S_{n,k}V)$ must be divisible by z^k.

To prove (2) first note that $A^{(n)}(V)(z)$ is a competing function in the extremal problem defining the norm of $E_{n,k}(V)$. Hence

$$\left\|E_{n,k}(V)\right\| \leq \left\|A^{(n)}(V)\right\|_{H_{\mathcal{F}}^{\infty}} = \|V\|_{X^{(n)}}.$$

To obtain an estimate in the other direction we note that,

$$\|E_{n,k}(V)\| = \|A^{(n+k)}(E_{n,k}(V))\|_{H_{\mathcal{F}}^{\infty}} \geq \|V\|_{X^{(n)}}.$$

The previous line also shows why (3) holds.   □

We now recall the notion of *twisted direct sum* of Banach spaces. Given two Banach spaces, A and B, and a map $\Omega : A \to B$ (which is not assumed to be linear) we define the twisted direct sum to be the vector space $\{(a, b) : a \in A, b \in B\}$. We measure size in this space by

$$(3.4) \qquad \|(a, b)\|_{\Omega} = \|a\|_A + \|b - \Omega a\|_B.$$

This vector space together with the functional in (3.4) is denoted $A_{\Omega} \oplus B$. Note that we are abusing notation in (3.4), $\|\cdot\|_{\Omega}$ is not actually a norm. We could give general conditions on Ω which force this functional to be comparable to a norm. However when we use the construction we will have a norm at hand for comparison.

For $v \in X^{(0)}$ set $\Omega^{(j)}(v) = (\Omega_1 v, \dots, \Omega_j v)$.

Proposition 3.2.

$$X^{(n)} \sim X^{(0)}{}_{\Omega^{(j)}} \oplus X^{(n-1)}.$$

That is, for $V = (v_0, \dots, v_{n-1}) \in X^{(n)}$

$$(3.5) \quad \|V\|_{X^{(n)}} \sim \|v_0\|_{X^{(0)}} + \|(v_1 - \Omega_1 v_0, \dots, v_{n-1} - \Omega_{n-1} v_0)\|_{X^{(n-1)}}.$$

The constants implicit in "$\sim$" in (3.5) can be selected to depend only on n.

Proof. Let $W = (v_1, \dots, v_{n-1})$. We need to establish

$$(3.6) \qquad \|V\|_{X^{(n)}} \sim \|v_0\|_{X^{(0)}} + \left\|W - \Omega^{(n-1)}(v_0)\right\|_{X^{(n-1)}}.$$

We have, by direct inspection,

$$(3.7) \qquad V = E_{1,n-1}(v_0) + S_{n-1,1}(W - \Omega^{(n-1)}(v_0)).$$

Hence

$$\|V\| \leq \|E_{1,n-1}(v_0)\|_{X^{(n)}} + \left\|S_{n-1,1}\left(W - \Omega^{(n-1)}(v_0)\right)\right\|_{X^{(n)}}.$$

An appeal to Proposition 3.1 (1) and (2) now establishes half of (3.6). To go in the other direction first note that $v_0 = R_{1,n-1}(V)$ and hence by Proposition 3.1 (3)

$$(3.8) \qquad \|v_0\| \leq \|V\|_{X^{(n)}}.$$

Finally, using Proposition 3.1 (1) and then (3.7)

$$\left\| W - \Omega^{(n-1)}(v_0) \right\|_{X^{(n-1)}} \leq \left\| S_{n-1,1}\left(W - \Omega^{(n-1)}(v_0) \right) \right\|_{X^{(n)}}$$
$$\leq \| V - E_{1,n-1}(v_0) \|_{X^{(n)}}$$
$$\leq \| V \|_{X^{(n)}} + \| E_{1,n-1}(v_0) \|_{X^{(n)}}.$$

Finally, using Proposition 3.1 (1) and then (3.7)

$$\leq \| V \|_{X^{(n)}} + \| v \|$$
$$\leq 2\| V \|_{X^{(n)}}.$$

Thus the right hand side of (3.6) is dominated by three times the left hand side and we are done. $\qquad \square$

In order to unravel this description we introduce more notation. Given $(v_0, \dots, v_{n-1}) \in X^{(n)}$ we set

$$w_0 = v_0$$
$$w_1 = v_1 - \Omega_1 w_0$$
$$\vdots$$
$$w_{n-1} = v_{n-1} - \Omega_1 w_{n-2} - \dots - \Omega_{n-1} w_0.$$

Theorem 3.3.
$$\|(v_0, \dots, v_{n-1})\|_{X^{(n)}} \sim \sum_{j=0}^{\infty} \|w_j\|.$$

The constants of comparability depend only on n.

Proof. Proposition 3.2 and induction. $\qquad \square$

For $n = 1$ this is the result from [**RW**]; namely

$$\|(a,b)\|_{X^{(1)}} \sim \|a\| + \|b - \Omega_1 a\|.$$

The next case is

$$(3.9) \quad \|(a,b,c)\|_{X^{(2)}} \sim \|a\| + \|b - \Omega_1 a\| + \|c - \Omega_1(b - \Omega_1 a) - \Omega_2 a\|.$$

There was another path we could have taken to study $X^{(2)}$. We could have written

$$(a, b, c) = (a, b, \Omega_{2,2}((a, b))) + (0, 0, c - \Omega_{2,2}((a, b))).$$

Starting from this and using Proposition 3.2 we would have found

$$\|(a, b, c)\|_{X^{(2)}} \sim \|(a, b)\|_{X^{(1)}} + \|c - \Omega_{2,2}((a, b))\|$$

$$(3.10) \qquad\qquad \sim \|a\| + \|b - \Omega_1 a\| + \|c - \Omega_{2,2}((a, b))\|.$$

This is less satisfactory than (3.9) because in the case of the classical function spaces we know how to compute $\Omega_2 = \Omega_{2,1}$ but don't know how to compute $\Omega_{2,2}$. In fact, comparing (3.9) and (3.10) gives us a first step toward a formula for $\Omega_{2,2}$:

$$(3.11) \qquad\qquad \Omega_{2,2}((a, b)) = -\Omega_1(b - \Omega_1 a) - \Omega_2(a) + \text{error}$$

with

$$\|\text{error}\| \le c \|(a, b)\|_{X^{(1)}}.$$

This type of error term is not bad as it looks at first glance. Neither side of (3.11) is $O(\|(a, b)\|_{X^{(1)}})$. Hence an error term which is $O(\|(a, b)\|_{X^{(1)}})$ means we have a partial answer to the problem of describing $\Omega_{2,2}$.

A similar, but more intricate, analysis is possible for $X^{(j)}$ for $j > 2$. We summarize the first few cases.

Corollary 3.4.

$$\|(a, 0)\|_{X^{(1)}} \sim \|a\| + \|\Omega_1 a\|$$

$$\|(a, 0, 0)\|_{X^{(2)}} \sim \|a\| + \|\Omega_1 a\| + \|(\Omega_1^2 - \Omega_2)a\|$$

$$\|(a, 0, 0, 0)\|_{X^{(3)}} \sim \|(a, 0, 0)\|_{X^{(2)}} + \|(-\Omega_1(\Omega_1^2 - \Omega_2) + \Omega_2\Omega_1 - \Omega_3)a\|.$$

4. Duality.

We denote the ordinary sesquilinear pairing on $\mathbb{C}^N$ by juxtaposition; if $a = (a_1, \ldots, a_N)$ and $b = (b_1, \ldots, b_N)$ then $ab = \sum a_i \bar{b}_i$.

Given the boundary data X_θ, $0 \le \theta \le 2\pi$, let X_θ^* be the vector space $\mathbb{C}^N$ normed to be the isometric dual of the space X_θ with respect to the pairing just described. Let $X^{*(n)}$ be the spaces constructed from the spaces X_θ^* by the construction in §2. Let $A^{*(n)}, \Omega_{j,k}^*$ and Ω_j^* be the analogs for these spaces of the A's and Ω's of (2.4). We denote the isometric dual of $X^{(n)}$ by $X^{(n)*}$.

For $V = (v_0, \ldots, v_{n-1}) \in X^{(n)}$ and $U = (u_0, \ldots, u_{n-1}) \in X^{*(n)}$ we introduce the bilinear pairing

$$B(V, U) = \sum_{k=0}^{n-1} v_k u_{n-k}.$$

This pairing lets us regard $A \in X^{*(n)}$ as an element of the dual space $X^{(n)*}$. Because all the spaces being considered are finite dimensional it is easy to check that this map of $X^{*(n)}$ into $X^{(n)*}$ is a bijection which allows identification of the two vector spaces. We now want to compare the norms of these two spaces.

Theorem 4.1.

$$X^{*(n)} \sim X^{(n)*}.$$

That is, there are constants c_n which depend only on n so that for any $A \in X^{*(n)}, A \neq 0$

(4.1) $$c_n^{-1} \leq \|A\|_{X^{*(n)}} / \|A\|_{X^{(n)*}} \leq c_n.$$

Note: Although the results of the previous sections extend rather mechanically to infinite dimensional spaces, this result doesn't. In particular, when $n = 0$ we have a (generalized version of) the classical duality theorem for complex interpolation. That theorem is well known to require additional hypotheses in the infinite dimensional case.

Proof. The case $n = 0$ is the duality theorem for complex interpolation and is known. See [**CCRSW1**] or [**CCRSW2**].
     Pick $V \in X^{(n)}, U \in X^{*(n)}$

$$B(V, U) = \int A^{(n)}(V)(e^{i\theta}) \, A^{*(n)}(U)(e^{i\theta}) \, e^{-i(k-1)\theta} \, d\theta.$$

Hence

$$|B(V, U)| \leq \sup_\theta \left\| A^{(n)}(V)(e^{i\theta}) \right\|_{X_\theta} \left\| A^{*(n)}(U)(e^{i\theta}) \right\|_{X_\theta^*}.$$

Hence

(4.2) $$|B(V, U)| \leq \|V\|_{X^{(n)}} \|U\|_{X^{*(n)}}$$

which yields

$$\|A\|_{X^{(n)*}} \leq \|A\|_{X^{*(n)}}$$

which is half of what we want.

We now use the notation of Theorem 3.3. Rewriting the equations relating the v's and the w's gives

$$v_0 = w_0$$
$$v_1 = \Omega_1 w_0 + w_1$$
$$v_2 = \Omega_2 w_0 + \Omega_1 w_1 + w_2$$
$$\text{etc.}$$

Hence

$$V = (w_0, \Omega_1 w_0, \dots, \Omega_{n-1} w_0) + (0, w_1, \Omega_1 w_1, \dots, \Omega_{n-2} w_1) + \dots.$$

Using the notation of Proposition 3.1 and setting

$$S_j = S_{n,j}, \ E = E_{1,n-1}, \ R_j = R_{n,j}$$

we can rewrite our equation for V as

(4.3) $\qquad V = E(w_0) + R_1 S_1 E(w_1) + \dots + R_{n-1} S_{n-1} E(w_{n-1}).$

By Theorem 3.3

$$\|V\| \sim \sum_{j=0}^{n-1} \|w_j\|.$$

Suppose for now that

(4.4) $\qquad\qquad\qquad \|V\| \sim \|w_{n-1}\|$

and

(4.5) $\qquad\qquad \|V\| >> \|w_j\| \quad j = 0, \dots, k-2.$

Pick $a \in X^{*(0)}, \|a\|_{X^{*(0)}} = 1$ so that

(4.6) $\qquad\qquad\qquad w_{n-1} a = \|w_{n-1}\|_{X^{(1)}}.$

That we can do this is insured by the $n = 0$ case of the theorem which we noted earlier is already known. (Including the fact that we can take $c_0 = 1$ in (4.1).) Let

$$A = E(a) = (a, \Omega_1^* a, \dots, \Omega_{n-1}^* a).$$

Using (4.3)

$$B(V, A) = E(w_0)A + \ldots + R_{n-1}S_{n-1}E(w_{n-1})A.$$

The last term is

$$(0, \ldots, 0, w_{n-1})(a, \Omega_1^*, \ldots, \Omega_{j-1}^* a) = w_{n-1}a.$$

By (4.6) and (4.4) this is comparable to $\|V\|$. For the other terms, using (4.2), Proposition 3.1, and (4.5)

$$\begin{aligned}
|R_j S_j E(w_j)A| &\leq \|R_j S_j E(w_j)\| \|A\| \\
&\leq \|w_j\| \|a\| \\
&<< \|V\|.
\end{aligned}$$

Combining these estimates we find

(4.7) $$B(V, A) \sim \|V\|$$

and, by Proposition 3.1,

(4.8) $$\|A\| = 1.$$

Next we note that we can obtain (4.7) and (4.8) without (4.4) and (4.5). Let J be the smallest index such that

(4.9) $$\|V\| \sim \|w_J\|$$

and

(4.10) $$\|V\| >> \|w_s\| \quad s = 0, \ldots, J - 1.$$

We had been looking at the extreme case $J = n - 1$ but a moment's thought shows that there must be some such J, possibly $J = 0$. Pick a so that $\|a\|_{X^{*(1)}} = 1$ and $w_J a = \|w_J\|$. Let

$$A = (0, 0, \ldots, a, \Omega_1 a, \ldots, \Omega_{n-J}).$$

When we go through the previous argument we find

$$|R_j S_j E(w_j)A| \begin{cases} << \|V\| & \text{if} \quad j < J \\ \sim \ \|V\| & \text{if} \quad j = J \\ = 0 & \text{if} \quad j > J \end{cases}$$

and we finish as before.

We have shown that the norm of V as a linear functional (on $X^{*(n)}$) dominates the norm of V as an element of the interpolation space $X^{(n)}$. This actually completes the demonstration that the norms of $X^{(n)}$ and $X^{*(n)*}$ are comparable. Applying this conclusion the family of spaces $\{X_\theta^*\}$ gives the conclusion we want, the comparability of the norms of $X^{*(n)}$ and $X^{(n)*}$. (The reason for this particular organization of the ideas was to let us use the notation of Theorem 3.3.) The proof is complete. □

Question: How do the constants c_n depend on n? This isn't even clear when $N = 1$ in which case it becomes a question in classical function theory.

5. Interpolation and Commutator Theorems.

For T a linear map of $\mathbb{C}^N$ to $\mathbb{C}^N$ we donote by $\|T\|_\theta$ the operator norm of T as a map of X_θ to itself. Let $T^{(n)}$ be the extension of T to $(\mathbb{C}^N)^n$ defined as follows; for $V = (v_0, \dots, v_{n-1})$ set $T^{(n)}V = (Tv_0, \dots, Tv_{n-1})$.

Theorem 5.1.   *Suppose T is a linear map of $\mathbb{C}^N$ to itself and that $\sup_\theta \|T\|_\theta \le 1$. Then the operator norm of $T^{(n)}$ as a map of $X^{(n)}$ to itself is at most 1.*

Proof. Pick $V \in X^{(n)}$. $T\left(A^{(n)}(V)(z)\right)$ is a $(\mathbb{C}^N)^n$ valued holomorphic function and by the hypotheses on T

$$\left\|TA^{(n)}(V)(e^{i\theta})\right\|_{X_\theta} \le \left\|A^{(n)}(V)(e^{i\theta})\right\|_{X_\theta}.$$

Hence, by the definition of $A^{(n)}$,

$$\left\|TA^{(n)}(V)(e^{i\theta})\right\|_{X_\theta} \le \|V\|_{X^{(n)}}.$$

Hence

$$\|(Tv_0, Tv_1, \dots, Tv_{n-1})\|_{X^{(n)}} \le \|V\|_{X^{(n)}}.$$

That is, $\left\|T^{(n)}V\right\| \le \|V\|$, which is the desired conclusion. □

Note: As in [**RW**] this result can be extended to take into account more detailed information about $\|T\|_\theta$ as a function of θ and can be extended to apply to families of linear maps $T(z)$ which depend holomorphically on z.

When $n = 0$ this is the basic theorem of complex interpolation. When $n = 1$ we obtain estimates of the form

$$\|(Ta, Tb)\|_{X^{(1)}} \le \|(a, b)\|_{X^{(1)}}.$$

When this is combined with Theorem 3.3 we obtain

$$\|Ta\| + \|Tb - \Omega_1 Ta\| \leq c(\|a\| + \|b - \Omega_1 a\|).$$

We now make the choice $b = \Omega_1 a$. Making the left hand side smaller we obtain

$$\|T\Omega_1 a - \Omega_1 Ta\| \leq c\|a\|$$

or, more compactly,

$$\|[T, \Omega_1]\| \leq c.$$

This estimate is the main theme of [**RW**].

We now move on to $n = 2$. We know

$$\|(Ta, Tb, Tc)\|_{X^{(2)}} \leq \|(a, b, c)\|_{X^{(2)}}.$$

We appeal to Theorem 3.3 to interpret the norms. This gives

$$(5.1) \quad \|Ta\| + \|Tb - \Omega_1 Ta\| + \|Tc - \Omega_1(Tb - \Omega_1 Ta) - \Omega_2 Ta\|$$
$$\leq c(\|a\| + \|b - \Omega_1 a\| + \|c - \Omega_1(b - \Omega_1 a) - \Omega_2 a\|).$$

We set $b = \Omega_1 a, c = \Omega_2 a$ and, making the left hand side smaller we find

$$(5.2) \qquad \qquad \|[T, \Omega_2] - \Omega_1[T, \Omega_1]\| \leq c.$$

This is the result mentioned in the comments in the last section of [**RW**]. (There is a slight change of notation — here we work with the Taylor coefficients rather than the derivatives.)

We will go on to derive a general pattern. However before doing this we should note that even at this stage by making other choices (i.e. other than $b = \Omega a$ and $c = \Omega_2 a$) we could obtain other complicated looking and not well understood interpolation results. For instance by restricting $T^{(2)}$ to the subspace of $X^{(2)}$ consisting of vectors of the form $(a, 0, 0)$ (i.e. by setting $b = c = 0$ in (5.1)), we find T is bounded on a space with norm comparable to

$$\|a\| + \|\Omega_1 a\| + \|(\Omega_2 - \Omega_1^2)a\|.$$

This is equivalent to noting that T maps the subspace of $X^{(3)}$ of elements of the form $(a, 0, 0)$ boundedly to itself and then invoking Corollary 3.4.

The generalization of (5.2) can be written efficiently if we introduce some notation. We set

$$K_0 = T$$
$$K_1 = [T, \Omega_1]$$

$$K_2 = [T, \Omega_2] - \Omega_1 K_1$$

$$\vdots$$

$$K_{n-1} = [T, \Omega_{n-1}] - \Omega_1 K_{n-2} - \cdots - \Omega_{n-2} K_1.$$

Corollary 5.2. *With the hypotheses of Theorem 5.1, K_{n-1} maps X boundedly to itself. That is*

$$\|K_{n-1}\| \leq c$$

with c depending only on n.

Proof. Let $V = (v, \Omega_1 v, \dots, \Omega_{n-1} v)$. Hence $T^{(n)} V = (Tv, T\Omega_1 v, \dots, T\Omega_{n-1} v)$. By the previous theorem

$$\left\| T^{(n)} V \right\|_{X^{(n)}} \leq c\|V\|_{X^{(n)}}.$$

By Proposition 3.1 (2) we obtain

$$\|V\|_{X^{(n)}} \leq c\|v\|.$$

We now want to use Theorem 3.3 to estimate the left hand side. Direct computation using the definitions of the K_j's and Theorem 3.3 gives

$$\|V\| \sim \sum_{j=0}^{n-1} \|K_j v\|.$$

Combining these estimates and dropping terms from the sum gives the required

$$\|K_{n-1} v\| \leq c\|v\|$$

as required. $\square$

In the next section we will see how the K_j simplify in some cases.

6. Examples.

We will compute the Ω_j's for some of the basic examples of complex interpolation spaces. We will obtain the simplest formulas if we work with complex interpolation families on the strip $S = \{z \in \mathbb{C} : |\operatorname{Re}(z)| \leq 1\}$ and then restrict to Γ to get our boundary data. The reiteration theorem for complex interpolation then insures that the space $X_0^{(1)}$ will be the space at 0 from the original family. We refer to [RW] as well as [CCRSW1], [CCRSW2], and [K2] for justification of various interpolation theoretic claims, such as explicit formulas for $A^{(n)}$, which we make in this section.

A. Weighted L^2 spaces.

Suppose we have a measure space (Y, μ) and a function W on Y which is positive $a.e. - \mu$. To each point $z \in S$ we associate the Hilbert space

$$L^2(Y, W^{-2\operatorname{Re}(z)}\mu) = \left\{ f : \int_Y |f|^2 W^{-2\operatorname{Re}(z)}\, d\mu < \infty \right\}.$$

Thus our boundary data on Γ is

$$X_\theta = L^2(Y, W^{-2\cos\theta}\mu)$$

and

$$X_0^{(0)} = L^2(Y, \mu).$$

(If we want to insure that all the spaces are finite dimensional we would insist that μ be supported on a finite set. However, as we noted, our results extend mechanically to the infinite dimensional context.) Given v in $L^2(Y, \mu)$ we have

$$A^{(0)}(v)(z) = e^{z \log W} v$$
$$= \sum_0^\infty v \frac{(\log W)^n}{n!} z^n.$$

Hence

$$\Omega_n(v) = \frac{(\log W)^n}{n!} v.$$

Let $Mv = (\log w)v$. We have $\Omega_n = M^n/n!$. Thus all of the Ω_n live in a commmuting family of linear operators. This lets us simplify the formula for K_n. It is slightly more convenient to do the computations for operators T which are given by integration against a kernel $L(x, y)$. That is

$$(Tv)(x) = \int_Y v(y) L(x, y)\, d\mu(y).$$

Set $b = \log w$. The integration kernel for $K_1 = TM - MT$ is

(6.1) $(b(y) - b(x)) L(x, y).$

We claim that the integration kernel for K_n is

(6.2) $\dfrac{(b(y) - b(y))^n}{n!} L(x, y).$

We will establish (6.2) by induction. (6.1) lets us start. Suppose we have (6.2). We know

$$K_{n+1} = [T, \Omega_{n+1}] - \Omega_1 K_n - \Omega_2 K_{n-1} - \ldots - \Omega_n K_1$$

$$= \frac{1}{(n+1)!} [T, M^{n+1}] - \sum_{j=1}^{n} \frac{1}{j!} M^j K_{n+1-j}.$$

Using the induction hypothesis and passing to the integration kernels we find the integration kernel for $(n+1)! \, K_{n+1}$ is

$$b^{n+1}(y) - b^{n+1}(x) - \sum_{j=1}^{n} \frac{(n+1)!}{j!} \, b^j(x) \frac{1}{(n+1-j)!} \, (b(y) - b(x))^{n+1-j}$$

$$= b^{n+1}(y) - b^{n+1}(x) - \sum_{j=1}^{n} \binom{n+1}{j} b^j(x)(b(y) - b(x))^{n+1-j}.$$

By the Binomial Theorem the sum equals

$$(b(x) + (b(y) - b(x)))^{n+1} - b^{n+1}(x) - (b(y) - b(x))^{n+1}$$
$$= b^{n+1}(x) - b^{n+1}(y) - (b(y) - b(x))^{n+1}.$$

Hence the entire expression is

$$(b(y) - b(x))^{n+1}$$

as required.

Hence

$$K_n = \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} (-1)^j M^j T M^{n-j}$$

$$= \frac{1}{n!} [\ldots [T, M], M], M], \ldots, M].$$

Actually we have only established this for T which is given by integration against a kernel. However the same combinatoric pattern works in general.

We noted in [**RW**] that the $n = 0$ case of this result can be applied to classical singular integral operators. The same comment applies here and we obtain the boundednes of the higher commutators which was noted in [**CRW**].

Corollary 6.1. *Suppose T is a Calderón-Zygmund singular integral operator which is bounded on $L^2(\mathbb{R}^n)$ and that b is in $BMO(\mathbb{R}^n)$. Let M_b be the operator of multiplication by b. For each n the operator*

$$\underbrace{[M_b, [M_b[\ldots [M_b, T]]\ldots]}_{n \text{ times}}$$

262 RICHARD ROCHBERG

is bounded on $L^2(\mathbb{R}^n)$.

Proof. The theory of Calderón-Zygmund operators insures that if T is bounded on L^2 it will also be bounded on $L^2(\mathbb{R}^n, e^{\varepsilon b}\, dx)$ for all ε with $|\varepsilon|$ sufficiently small. Hence K_n is bounded on L^2. $\qquad\qquad\square$

For $n = 0$ this corollary insures that for $f \in L^2, Tbf - bTf$ will be in L^2. Let $\langle \cdot, \cdot \rangle$ be the inner product on L^2. For g which is also in L^2 we have

$$c_b \|f\|\|g\| \geq |\langle Tbf - bTf, g \rangle|$$
$$= |\langle fT^*g - gTf, b \rangle|.$$

This estimate is the basis for showing that the map from $L^2 \times L^2$ which takes (f, g) to $fT^*g - gTf$ in L^1 actually maps into the Hardy space H^1 which is the predual of BMO (see [**CRW**] for details). This result is closely related to recent work on bilinear maps with cancellation ([**CLMS**], [**CG**], [**G**]). Hence it is of some interest to write down the analog for $n > 0$. We will do this for $n = 1$ but the general pattern will be clear. First we need an extension of the previous corollary.

Corollary 6.2. *Let T and b be the operator and function of the previous corollary and suppose that d is in $BMO(\mathbb{R}^n)$. Let M_d be the operator of multiplication by d. It follows that*

$$[M_d, [M_b, T]]$$

is bounded on $L^2(\mathbb{R}^n)$.

Proof. This doesn't quite follow from the previous result. One way to prove it is to note that Calderón-Zygmund theory insures that for small ε_1 and ε_2 T is bounded on the space $L^2(\mathbb{R}^n, \exp(\varepsilon_1 b)\exp(\varepsilon_2 d)\, dx)$. If we regard ε_2 as fixed then we have T bounded on the family of weighted L^2 spaces given by letting ε_1 vary. Then, as in the previous corollary, we have $[M_b, T]$ bounded on all the spaces $L^2(\mathbb{R}^n, \exp(\varepsilon_2 d)\, dx)$. We now use the Corollary again to obtain the boundedness of $[M_d, [M_b, T]]$. $\qquad\qquad\square$

Once we have this corollary we also have the estimate

$$|\langle d[M_b, T]f - [M_b, T]df, g \rangle| \leq c\|f\|\|g\|.$$

Hence

$$|\langle g[M_b, T]f - f[M_b, T]^*g, d \rangle| \leq c\|f\|\|g\|.$$

Because we have this for all d in BMO we can conclude that the left hand side in the inner product is in H^1. We have proved the following result.

Corollary 6.3.   *Suppose b is in BMO and T is a Calderón-Zygmund operator which is bounded on L^2. Consider the bilinear map B defined on $L^2 \times L^2$ by*

$$B(f,g) = gbTf - gTbf + fbT^*g - fT^*bg.$$

By Corollary 6.1 (with $n = 0$) B maps into L^1. In fact however B maps into H^1.

We should make a final comment on Corollary 6.1. There is nothing particular about L^2 in the analysis nor about the use of weight functions. If we are given a general Banach space R and a positive linear operator W defined on R then we can associate to $z \in S$ the space Domain(W^{-z}) which is the closure of $\{f \in R : W^{-z}f \in R\}$ normed by $\|f\| = \|W^{-z}f\|_R$. We will then have, for $v \in R$,

$$A^{(0)}(v)(z) = W^z v$$

and

$$\Omega_n v = \frac{1}{n!} (\log W)^n v.$$

The previous analysis then gives a formula for K_n. (Of course we are over-simplifying. We need enough context to insure that Domain(W^{-z}) is an interpolation family and enough of a functional calculus to make sense of $\log W$. Although these are real restrictions, they allow a much broader range of examples than multiplication by a positive function on L^2.)

B. L^p, variable p.

We now look at the scale obtained in the strip S by interpolating between L^1 and L^∞. We accept the formulas from Section 3D of [**RW**] and refer to that paper to see the minor modifications needed to consider the slightly more general situation of interpolation between L^{p_1} and L^{p_2}.

We start with a measure space (Y, μ) and to each z in S we associate the Banach space $L^p(Y, \mu)$ with $p = \frac{2}{1+x}$. If we regard these spaces as the boundary spaces on Γ then they are also the interpolation spaces inside the disk and, in particular

$$X_0^{(0)} = L^2(Y, \mu).$$

From Section 3D of [**RW**] we have, for $v \in L^2$

$$A^{(0)}(v)(z) = v \left(\frac{|v|}{\|v\|_2} \right)^z$$

$$= \sum \frac{1}{n!} v \left(\log \frac{|v|}{\|v\|_2} \right)^n z^n.$$

Thus

$$\Omega_n v = \frac{1}{n!} v \left(\log \frac{|v|}{\|v\|_2} \right)^n.$$

In contrast to the previous case, Ω_n is not linear and the operators Ω_1, $\Omega_2, \ldots$ don't commute. When $n = 1$ we conclude that the operator K_1 is bounded.

Corollary 6.4. *If T is a linear operator bounded on L^1 and L^∞ then for $v \in L^2$*

$$\left\| T \left(v \log \frac{|v|}{\|v\|_2} \right) - Tv \left(\log \frac{|v|}{\|v\|_2} \right) \right\| \leq c \|v\|_2.$$

Note: It was pointed out in [**RW**] that this result implies the slightly cleaner conclusion

$$\| T(v \log |v|) - Tv \log |Tv| \|_2 \leq c \|v\|_2.$$

We now move on to $n = 2$. We have, for any $v \in L^2$, and writing $\| \cdot \|$ for $\| \cdot \|_2$

$$\Omega_1 v = v \log \frac{|v|}{\|v\|}$$

$$\Omega_2 v = \frac{1}{2} v \left(\log \frac{|v|}{\|v\|} \right)^2.$$

Corollary 5.2 insures.

Corollary 6.5. *If T is bounded on L^1 and L^∞ then*

(6.3) $$[T, \Omega_2] - \Omega_1 [T, \Omega_1]$$

is bounded on L^2.

When we write this out in full it is a bit hard to interpret. However we can gain some additional insight into what is going on by restricting to special cases. Suppose that T is unitary on L^2. For instance we could work with the Hilbert Transform of $L^2(\mathbb{R})$. (This would require that we work on a subscale of (L^1, L^∞) but that isn't a problem.) If we aply the operator in (6.3) to a function f with $\|f\| = 1$ then, using the fact that $\|Tf\| = 1$ also, we obtain

(6.4) $$\frac{1}{2} T \left(f (\log |f|)^2 \right) - Tf (\log |Tf|)^2 - \Omega_1 (T(f \log |f|) - Tf \log |Tf|).$$

If f is the characteristic function of a set E then $f = \chi_E$ is always 0 or 1 and hence

$$f \log |f| = f(\log |f|)^2 \equiv 0.$$

In that case (6.4) simplifies to

$$-\frac{1}{2} Tf(\log|f|)^2 + \Omega_1(Tf \log |Tf|)$$

$$= -\frac{1}{2} Tf(\log |Tf|)^2 + Tf \log |Tf| \log(|Tf \log |Tf|| / \|Tf \log |Tf|\|).$$

Set

(6.5) $\alpha = -\log \|Tf \log |Tf|\|$

so

$$\log(|Tf \log |Tf|| / \|Tf \log |Tf|\|) = \log|Tf| + \log|\log|Tf|| + \alpha$$

and hence we have the following.

Corollary 6.6.   *Suppose T is bounded on L^p for $p_1 \le p \le p_2$ where $p_1 < 2 < p_2$. Suppose T is unitary on L^2. Let $f = \chi_E$ for some set E of measure 1. With α given by (6.5) we have*

$$\left\| \frac{1}{2} Tf (\log |Tf|)^2 + Tf \log |Tf| (\log|\log |Tf|| + \alpha) \right\| \le c.$$

We could continue in this way. It is clear that the dominant term of K_n would involve $Tf(\log |Tf|)^n$. However it is not clear how to organize the rest of the terms effectively.

C. Lattices

In the previous two subsections we saw that for L^p spaces and weighted L^2 we had relatively simple formulas for the Ω_j's. In fact that is true whenever we work with a scale generated by two Banach lattices (which satisfy certain additional technical restrictions). We will summarize that here. Before doing that it is worth noting that many common function spaces which are not ordinarily presented as lattices can be realized as lattice sequence spaces by passage to wavelet coefficients.

We refer to [**K2**] for the technical background and will assume the spaces we consider are admissible in the sense of that paper. That paper also presents a version of complex interpolation which can be used as a context in which to extend our earlier results to infinite dimensional spaces.

Suppose we are given two admissible lattices Y_0 and Y_1 on a common measure space. The complex interpolation spaces in the sense of Calderón are the lattices $\{Y_t\}_{0<t<1}$ on the same measure space with norm

$$(6.6) \qquad \|y\|_t = \inf\{\max(\|u\|_{Y_0}, \|v\|_{Y_1} : |y| = u^{1-t}v^t, \, u, v \geq 0)\}.$$

We construct our spaces on Γ from this family;

$$X_{e^{i\theta}} = Y_{2\cos t - 1}.$$

In this case $X_0 = Y_{1/2}$. Furthermore the u and v which solve the extremal problem in (6.6) allow us to construct $A^{(0)}(y)(z)$. The infimum in (6.6) is attained and if we also suppose that the norms on Y_0 and Y_1 are strictly convex and that $u = v = 0$ off the support of y then u and v will be unique and independent of t. We then set

$$\Lambda(y) = \begin{cases} 0 & \text{off the support of } y \\ \frac{1}{2}(\log v - \log u) & \text{on the support of } y. \end{cases}$$

In this case

$$A^{(1)}(z) = u \exp(\Lambda(y)z).$$

Hence

$$\Omega_n(y) = \frac{1}{n!}(\Lambda(y))^n u.$$

It is interesting to note that we can recover all of the $\Omega_n(y)$ once we know y, $\Omega_1(y)$, and $\Omega_2(y)$.

Finally we point out that the examples of the previous subsections are covered by this analysis. We don't know if the fact that all of the $\Omega_n(y)$ can be recaptured from the first two remains true if the interpolation family is not generated by a pair of lattices.

7. Questions.

There are a number of natural questions related to this work. Here are two that seem central. What, if anything, are the analogs in real interpolation of these results? For $n = 1$ there are analogs of Ω_1 and there is a commutator theorem. For $n > 1$ nothing is known.

The estimates on K_n clearly contain subtle information about operators. However the information is packaged in a complicated way. Is there a way to extract that information effectively?

References

[CCRSW1] R.R. Coifman, M. Cwikel, R. Rochberg, Y. Sagher and G. Weiss, *The complex method of interpolation of operators acting on families of Banach spaces*, Lecture Notes in Mathematics, **779** (1980), 123-153.

[CCRSW2] _____, *A theory of complex interpolation for families of Banach spaces*, Advances in Math., **43** (1982), 203-229.

[CG] R.R. Coifman and L. Grafakos, *Hardy space estimates for multilinear operators I*, Rev. Mat. Iberoamericana, **8** (1992), 45-67.

[CLMS] R. Coifman, P.L. Lions, Y. Meyer and S. Semmes, *Compensated compactness and Hardy spaces*, Jour. Math. Pures et Appl., **72** (1993), 247-286.

[CRW] R. Coifman, R. Rochberg and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. Math., **103** (1976), 611-635.

[CJMR] M. Cwikel, B. Jawerth, M. Milman and R. Rochberg *Analysis at Urbana II*, (Berkson et al, eds.), Cambridge University Press, Cambridge, England , (1989), 170-220.

[G] L. Grafakos, *Hardy space estimates for multilinear operators*, Rev. Mat. Iberoamericana, **8** (1992), 69-92.

[IS] T. Iwaniec and C. Sbordone, *Weak minima of variational integrals*, to appear.

[JRW] B. Jawerth, R. Rochberg and G. Weiss, *Commutator and other second order estimates in real interpolation theory*, Ark. Mat., **24** (1986), 191-219.

[K1] N. Kalton, *Nonlinear commutators in interpolation theory*, Memoirs. Amer. Math. Soc., **385** (1988), Providence R.I.

[K2] _____, *Differentials of complex interpolation processes for Köthe function spaces*, Trans. Amer. Math. Soc., **333** (1992), 479-529.

[M1] M. Milman, *A commutator theorem with applications*, Conference on Function Spaces, Poznan, 1992, to appear.

[M2] _____, *Inequalities for Jacobians: Interpolation techniques*, Proc. 2nd Latin American Analysis Conference, Rev. Mat. Colombiana, to appear.

[M3] _____, *Integrability of the Jacobian of orientation preserving maps: Interpolation techniques*, Comptes Rendus Acad. Sci., to appear.

[MS] M. Milman and T. Schonbek, *Second order estimates in interpolation theory and applications*, Proc. Amer. Math. Soc., **110** (1990), 961-969.

[RW] R. Rochberg and G. Weiss, *Derivatives of analytic families of Banach spaces*, Ann. Math., **118** (1983), 315-347.

WASHINGTON UNIVERSITY
ST. LOUIS, MO 63130
E-mail address: rr@math.wustl.edu