SOME RECENT DEVELOPMENTS IN FOURIER ANALYSIS AND H^p-THEORY ON PRODUCT DOMAINS[1]

BY SUN-YUNG A. CHANG AND ROBERT FEFFERMAN

Introduction. In this article we wish to discuss a theory which is still developing very rapidly. It is only quite recently that many of the aspects of Fourier analysis of several parameters have been discovered, even though much of the corresponding one-parameter theory has been well known for some time. The topics to be covered include differentiation theory, singular integrals, Littlewood–Paley theory, weighted norm inequalities, Hardy spaces, and functions of bounded mean oscillation, as well as many other related topics. We shall begin in Part I by attempting to give a broad overview of some of the one-parameter results about these topics. The discussion here is, however, anything but encyclopedic. (For more detailed treatments of these matters in the one-parameter setting, the reader can consult such excellent treatments as E. M. Stein, *Singular integrals and differentiability properties of functions* [**75**], R. R. Coifman and G. Weiss, *Extensions of Hardy spaces and their use in analysis* [**30**], and, in the classical domain of the disc, D. Sarason, *Function theory on the unit circle* [**72**], and J. Garnett, *Bounded analytic functions* [**46**].) In Part II we take up these same areas in the two-parameter setting. Since this theory is less well known than the material of Part I, we go into greater detail and devote separate sections to each of several of the above topics.

PART I. THE ONE-PARAMETER THEORY

To begin with the one-parameter theory, perhaps the most basic part is the differentiation of integrals and the maximal function of Hardy–Littlewood. If f is a function on R^n which is Lebesgue integrable, and if

$$A_r(f)(x) = \frac{1}{m(B_r(x))} \int_{B_r(x)} f(y)\, dy$$

denotes the average value of f over the ball with center x and radius r, then

$$\lim_{r \to 0} A_r(f)(x) = f(x) \quad \text{for a.e. } x \in R^n.$$

This fundamental result of Lebesgue, proved in the earlier years of the century, was applied immediately in a number of contexts. For example, Lebesgue saw that it could be used to show that for integrable functions of one variable, the arithmetic means of the partial sums of the Fourier series converge pointwise almost everywhere.

1

For the development of Fourier analysis, the most far reaching result connected with Lebesgue's theorem was that of Hardy–Littlewood in the early thirties–the Hardy–Littlewood Maximal Theorem. It said that if we consider the operator

$$Mf(x) = \sup_{r>0} A_r(|f|)(x),$$

then

$$\|Mf\|_p \leq C_p\|f\|_p \quad \text{for } p > 1,$$

and for L^1-functions,

$$m\{Mf > \alpha\} \leq (C/\alpha)\|f\|_1 \quad \text{for all } \alpha > 0$$

(where $\|f\|_p$ denotes the L^p-norm for $p > 1$). This maximal theorem is easily seen to imply Lebesgue's theorem, and the maximal function and its variants have played a leading role in many areas of analysis, including singular integral operators and Hardy spaces.

The deepest part of the maximal theorem is the estimate

$$m\{Mf > \alpha\} \leq (C/\alpha)\|f\|_1,$$

and this, in turn, depends on a geometric covering lemma. The covering lemma says, roughly, that from an arbitrary collection of balls in R^n we may select a disjoint subcollection whose total volume is at least a fixed fraction of the volume in the whole collection. It is interesting to note that if, in R^n, we denote by $A_{r_1,r_2,\ldots,r_n}(f)(x)$ the average of f over the rectangle centered at x with sides parallel to the axes of lengths $r_1, r_2, \ldots, r_n$, and if $f \in L^1(R^n)$, then for every $x \in R^n$

$$\lim_{r_1,r_2,\ldots,r_n \to 0} A_{r_1,r_2,\ldots,r_n}(f)(x)$$

may not exist when $n > 1$! On the other hand, if $\alpha_1(r), \alpha_2(r), \ldots, \alpha_n(r)$ are increasing functions of $r > 0$, and if $A_r(f)(x)$ denotes the average over the rectangle centered at x with sides parallel to the axes of lengths $\alpha_1(r), \alpha_2(r), \ldots, \alpha_n(r)$, then, if $f \in L^1(R^n)$, $\lim_{r\to 0} A_r(f)(x)$ again exists a.e. [85].

What these results tell us is that if we are interested in differentiating the integral of an integrable function in R^n, then, very roughly speaking, it is not the number of dimensions n that is important, but rather the number of parameters indexing the sets we are averaging over: only one-parameter families of sets can be expected to differentiate the integrals of Lebesgue integrable functions in R^n.

The next topic we discuss in the classical theory is interpolation. This notion is already used in proving the Maximal Theorem. We said above that the basic estimate for the Hardy–Littlewood operator is

(1) $$m\{Mf > \alpha\} \leq (C/\alpha)\|f\|_1.$$

Since, naturally, any average of a bounded function does not exceed the bound of that function, we also have

(2) $$\|Mf\|_\infty \leq \|f\|_\infty.$$

It turns out, according to a celebrated theorem of Marcinkiewicz, that *any* linear, or even sublinear, operator T satisfying the L^1-estimate (1) and L^∞-estimate (2) is bounded on $L^p(R^n)$ for all $1 < p < \infty$. There are a great many theorems these days of this same form–namely, if a linear operator T is bounded between spaces X_0 and Y_0 and also bounded between another pair of spaces X_1 and Y_1, then T is automatically bounded as an operator from X to Y for some appropriate intermediate pair of spaces. To give just one other example, if a linear operator T is bounded on the Hardy space $H^1(R^1)$ and also bounded on $L^2(R^1)$, then it must be bounded on $L^p(R^1)$ for all p between 1 and 2 [44]. There are many, many more examples of this general technique of interpolation.

In this setting of the maximal function and interpolation, another area of real variables and Fourier analysis developed–singular integrals of Calderón–Zygmund. These singular integral operators are generalizations, to the setting of R^n, of the Hilbert transform H on R^1. H is defined by the nonabsolutely convergent integral

$$Hf(x) = \int_{-\infty}^{\infty} f(x-t)\frac{dt}{t}.$$

It turns out that this operator is enormously important for several reasons. Here we shall content ourselves with two of them.

First, there is the connection with complex analytic functions. Suppose $f(x)$ is real valued and $U(x, y)$ is the harmonic extension (Poisson integral) of $f(x)$ to the upper half-plane R_+^2. Let $V(x, y)$ be the unique harmonic function vanishing as $y \to \infty$ so that $U + iV$ is analytic in R_+^2. Then the boundary values of V are none other than $Hf(x)$. Thus, if we identify functions on R^1 with their harmonic extensions to R_+^2, then H is the map which sends the real part of a complex analytic function to its imaginary part.

The other important reason for considering H is the connection with Fourier analysis of functions on R^1. If f is a "nice" function on R^1 and

$$\hat{f}(\xi) = \int_{R^1} f(x)e^{-i\xi x}\, dx$$

is the Fourier transform of f, we wish to know in what sense the Fourier integral

$$\int_{R^1} \hat{f}(\xi)e^{i\xi x}\, d\xi$$

represents $f(x)$. It turns out that for $f \in L^p(R^1)$, $1 < p < \infty$, the integrals

$$\int_{-R}^{+R} \hat{f}(\xi)e^{i\xi x}\, d\xi$$

converge to $f(x)$ in the L^p-norm, and it is easy to see that this is equivalent to H being a bounded operator on $L^p(R^1)$ for $1 < p < \infty$. Originally, the proof of Marcel Riesz that H preserves L^p used Cauchy's theorem in complex analysis. Somewhat later, real-variable proofs were developed, culminating in the Calderón–Zygmund work of the 1950s [14].

In their investigation Calderón and Zygmund considered convolution operators

$$Tf(x) = \int_{R^n} f(y)K(x-y)\,dy,$$

where the kernel $K(x)$ defined on R^n "looks like" $1/x$ does on R^1. They assumed that

$$|K(x)| \le C/|x|^n, \qquad |\nabla K(x)| \le C/|x|^{n+1},$$

$K(x)$ is C^1 away from the origin, and

$$\int_{\alpha < |x| < \beta} K(x)\,dx = 0 \quad \text{for all } 0 < \alpha < \beta.$$

Under these assumptions they proved that, for $1 < p < \infty$, $\alpha > 0$,

$$\|Tf\|_p \le C_p\|f\|_p \quad \text{and} \quad m\{|Tf| > \alpha\} \le (C/\alpha)\|f\|_1.$$

The techniques they developed in their argument set the tone for real-variable theory for many years.

Calderón and Zygmund begin by observing that the assumptions on $K(x)$ imply that $\hat{K}(\xi)$ is bounded. Hence, by the Plancherel theorem,

$$\|Tf\|_2 = \|\widehat{Tf}\|_2 = \|\hat{K} \cdot \hat{f}\|_2 \le \|\hat{K}\|_\infty \|\hat{f}\|_2 \le C\|f\|_2,$$

so T is bounded on L^2. Next they prove the estimate

$$(*) \qquad\qquad m\{|Tf| > \alpha\} \le (C/\alpha)\|f\|_1$$

as follows: Let $f \in L^1(R^n)$ and $\alpha > 0$. Calderón and Zygmund show how to replace f by an L^2-function g by averaging f over certain disjoint cubes Q_k where the average of f is $\le 2^n\alpha$. Of course, Tg is easy to handle, since $g \in L^2$ and in the previous step the boundedness of T on L^2 was proven. What remains in the proof of $(*)$ is an argument to handle the error $b = f - g$ in order to show that

$$m\{|T(b)| > \alpha\} \le (C/\alpha)\|f\|_1.$$

b has mean value zero over each of the Q_k and lives on $\bigcup_k Q_k$. It turns out to be not difficult to show that Tb is negligible outside $\bigcup Q_k$. All that remains is to show that $m(\bigcup Q_k)$ is small enough, i.e., $\le (C/\alpha)\|f\|_1$.

Now we come to an important feature of the argument: the set $\bigcup Q_k$ is precisely $\{Mf > \alpha\}$! In other words, the Hardy–Littlewood maximal function has been introduced to produce the desired L^2-function g, which, as far as T is concerned, is about the same as f itself. Notice that now the desired estimate of $m(\bigcup Q_k)$ is now just the main part of the Hardy–Littlewood Maximal Theorem. Once the L^1-estimate is obtained, interpolation shows that T is bounded on L^p for $1 < p \le 2$. If we notice that the adjoint operator to T is again a singular integral of the same form, we see that T is also bounded on L^p when $2 \le p < \infty$.

There have been a great many applications of the Calderón–Zygmund theorem, and we shall present one of them here. Let $f(\theta)$ be a function on

$[0, 2\pi)$ with Fourier series $\sum_{n=-\infty}^{\infty} \hat{f}(n)e^{in\theta}$. Let $\{\lambda_n\}_{n=-\infty}^{+\infty}$ be a given sequence of complex numbers. It is an important question in Fourier analysis to ask whether, if $f \in L^p(0, 2\pi)$, the same will be true of $\sum_{n=-\infty}^{+\infty} \lambda_n \hat{f}(n)e^{in\theta}$. There are many interesting examples of $\{\lambda_n\}$ where the answer is "yes" when $1 < p < \infty$. For instance, when $\lambda_n = -i\,\mathrm{sgn}(n)$, then $\sum \lambda_n \hat{f}(n)e^{in\theta}$ is the Hilbert transform of f. Another fundamental example is the class of sequences θ_n such that $\lambda_k = \varepsilon_n$ for all k with $2^n \leq |k| < 2^{n+1}$ and where ε_n is either $+1$ or -1. Then we consider

$$\Delta_n(f)(\theta) = \sum_{2^n \leq |k| < 2^{n+1}} \hat{f}(k)e^{ik\theta}$$

and, finally, the so-called square function,

$$\left(\sum_{k=1}^{\infty} |\Delta_k(f)|^2 \right)^{1/2} = S(f).$$

Of course, the function $\sum \lambda_k \hat{f}(k)e^{ik\theta}$ for the sequence λ_k under consideration has exactly the same square function as f. Therefore, the question at hand is answered by the following theorem of Littlewood and Paley:

If $1 < p < \infty$ then $\|S(f)\|_p \approx \|f\|_p$ ($a \approx b$ means a/b is bounded above and below by a quantity depending only on p). In order to better understand this, let us consider a similar operator acting on functions on R^1. To do this, notice that the operators $\Delta_k(f)$ are convolutions of f with functions of integral zero whose Fourier transforms are dilates of each other. We consider

$$g_\psi(f)(x) = \left(\int_0^\infty |f * \psi_t(x)|^2 \, \frac{dt}{t} \right)^{1/2},$$

where $\psi \in C_c^\infty(R^1)$, ψ is odd, and $\psi_t(x) = t^{-1}\psi(x/t)$ (so $\hat{\psi}_t(\xi) = \hat{\psi}(t\xi)$). Then g_ψ is roughly the same kind of operator as S, and the point now is that g_ψ is a singular operator of Calderón–Zygmund (see [2]).

The kernel is Hilbert-space-valued, but the Calderón–Zygmund proof goes over without change to such kernels. To be specific, if

$$K: \ R^1 \to L^2((0, \infty); dt/t)$$

is given by $K(x)(t) = \psi_t(x)$, then $|K(x)| \leq C/|x|$, $|\nabla K(x)| \leq C/|x|^2$, and $\int_{\alpha < |x| < \beta} K(x)\,dx = 0 \ \forall 0 < \alpha < \beta$, while $g_\psi(f)(x) = |f * K(x)|$.

Now again, we wish to make an important point. If C denotes the class of Calderón–Zygmund singular integral kernels on R^n, and if

$$K_\delta(x) = \delta^{-n}K(x/\delta),$$

then we have the following invariance:

If $K \in C$ and $\delta > 0$ then $K_\delta \in C$.

That is, the class C is invariant with respect to the one-parameter class of dilations $x \to \delta x$ on R^n. Again, just as was the case for differentiation of integrals, the theory seems to be more or less the same independent of the

dimension n. What is important is that the operators involved are invariant under a one-parameter class of dilations. So, for instance, whatever the dimension n, we always have for a Calderón–Zygmund operator T,

(X) $m\{x \in R^n \mid |Tf(x)| > \alpha\} \leq (C/\alpha)\|f\|_1.$

Later, we consider classes of kernels invariant with respect to several-parameter classes of dilations, and for these (X) will be false!

The next feature of one-parameter theory which we take up is that of inequalities with respect to measures other than Lebesgue. In fact, there is a single, very simple, necessary and sufficient condition on a locally integrable function $w(x) \geq 0$ on R^n so that

$$\int_{R^n} Mf(x)^p w(x)\, dx \leq C_p \int_{R^n} f^p(x) w(x)\, dx \quad \text{for } f \in L^y(w)$$

(here $1 < p < \infty$). This is the Muckenhoupt A^p condition [66]: (A^p)

$$\left(\frac{1}{|Q|}\int_Q w(x)\, dx\right)\left(\frac{1}{|Q|}\int_Q \left(\frac{1}{w(x)}\right)^{1/(p-1)} dx\right)^{p-1} \leq C \quad \forall \text{ cubes } Q \text{ in } R^n.$$

According to a theorem of Hunt–Muckenhoupt–Wheeden [51], we also have $w \in A^p$ iff

$$\int |Hf|^p w\, dx \leq C_p \int |f|^p \omega\, dx.$$

In fact, Coifman and C. Fefferman [28] extended this to the class of all Calderón–Zygmund operators in R^n. These so called weighted norm inequalities have proven to be of very great value in recent years.

As the reader has no doubt noticed, it seems that the operators T we have considered are bounded on $L^p(R^n)$ for $1 < p < \infty$, are unbounded on $L^1(R^n)$, and satisfy only the weaker estimate

$$m\{m \in R^n \mid |Tf(x)| > \alpha\} \leq (C/\alpha)\|f\|_1.$$

Something must be done in order to have a satisfactory "L^p-theory" of maximal functions and singular integrals when $0 < p \leq 1$. It is for this reason that one considers the Hardy spaces $H^p(R^{n+1}_+)$. First, let us consider $n = 1$.

According to Hardy, for $p > 0$, an H^p-function is a complex analytic function $F(z)$ in the upper half-plane R^2_+ such that the L^p-norms

$$\left(\int_{-\infty}^{+\infty} |F(x+iy)|^p\, dx\right)^{1/p}$$

are bounded independent of $y > 0$. It turned out that when $p > 1$, the H^p theory was very similar to the L^p-theory. So, for example, one of the main theorems of the subject is that H^p-functions, $p > 0$, have boundary values, i.e., when $F(z) \in H^p(R^2_+)$ then $\lim_{y \to 0} F(x+iy)$ exists for a.e. $x \in R^1$. This can be reduced to the theorem on differentiation of integrals of functions in $L^p(R^1)$ *when $p > 1$* (see [77, 85, 71]). The ideas that were originally used to study the Hardy spaces when $p \leq 1$ are much less along the lines of real

variables and are rather a part of the theory of analytic functions. One studied the zeros of these H^p-functions and showed that any function $F(z) \in H^p$ could be factored as $F = BG$, where $B(z)$ is bounded and analytic in R^2_+, and where $G \in H^p(R^2_+)$ and never vanishes. Then since $G(z)$ is never 0, one can form $G(z)^{p/2}$–an analytic function easily seen to be in $H^2(R^2_+)$. Since H^p-functions are known to have boundary values when $p > 1$, $G(z)^{p/2}$ and $B(z)$ will have boundary values, and, hence, so will $F(z)$.

Next, we wish to mention an extension of the theory of Hardy spaces due to E. M. Stein and Guido Weiss. Suppose we denote by R^{n+1}_+ the upper half-space in R^{n+1}, that is, $\{(x,y)|x \in R^n,\ y > 0\}$. Whereas Hardy spaces in R^2_+ are just analytic functions, or pairs of conjugate harmonic functions, Stein and Weiss [**77**] considered $H^p(R^{n+1}_+)$ functions as systems of $n + 1$ harmonic functions, $F(x,y) = \{u_i(x,y)\}$, $i = 0, 1, \ldots, n$, defined on R^{n+1}_+, which are conjugate in the sense that they satisfy the *generalized Cauchy–Riemann equations*

$$\frac{\partial u_i}{\partial x_j} \equiv \frac{\partial u_j}{\partial x_i} \quad \text{and} \quad \sum_{i=0}^{n} \frac{\partial u_i}{\partial x_i} \equiv 0 \qquad (y = x_0)$$

and such that

$$\sup_{y>0} \left(\int_{R^n} |F(x,y)|^p \, dx \right)^{1/p} < \infty.$$

(Here

$$|F(x,y)| = \left(\sum_{i=0}^{n} |u_i(x,y)|^2 \right)^{1/2} . \ )$$

We should point out that these spaces have an interpretation in terms of singular integrals which was alluded to above. Suppose we have a system of harmonic functions $u_0, u_1, \ldots, u_n$ defined in R^{n+1}_+ which are suitably smooth up to the boundary and vanish rapidly at infinity. Then it is not hard to show that the u_i satisfy the generalized Cauchy–Riemann equations if and only if their restrictions on the boundary are related by singular integrals known as Riesz transforms.

More specifically, if $f_i(x) = u_i(x,0)$ then u_i are a conjugate system iff

$$f_i = f_0 * c_n x_i / |x|^{n+1}.$$

The convolution operator

$$R_i(f) = f * x_i / |x|^{n+1}, \qquad i = 1, 2, \ldots, n,$$

the *Riesz transforms*, plays in R^n very much the same role as the Hilbert transform plays in R^1. Identifying a function $f(x)$ on R^n with the harmonic function $u(x,y)$ on R^{n+1}_+ having boundary values equal to $f(x)$, we see that $H^1(R^{n+1}_+)$ can be identified with the space of all real valued $L^1(R^n)$-functions all of whose Riesz transforms are also in $L^1(R^n)$. It can also be shown that for any such function whose Riesz transforms are in $L^1(R^n)$, *any* reasonable singular integral $T(f)$ will belong to $L^1(R^n)$. So it really is the case that

$H^p(R_+^{n+1})$ serves to alleviate the problem of the bad behavior of singular integrals on L^p when $p \leq 1$.

In order to prove theorems about these H^p-functions analogous to those which had been proven for the classical one, Stein and Weiss introduced a number of new ideas. To illustrate some of these, let us consider their theorem on boundary values of $H^p(R_+^{n+1})$-functions: If $F(x,y) \in H^p(R_+^{n+1})$ and $p \geq (n-1)/n$, then $\lim_{y \to 0} F(x,y)$ exists for a.e. $x \in R^n$. Again, for $p > 1$ the theorem had been well known for a long time since it again boils down to Lebesgue's theorem on differentiation of integrals of functions in $L^p(R^n)$. Now, in trying to pass from the case $p > 1$ to $p \leq 1$ we quickly see that the classical approach is not possible. A study of the zeros of F analogous to the classical case is obstructed by the fact that the zero set is no longer discrete, but may be higher dimensional. Also, factorization has no meaning since H^p functions cannot be multiplied meaningfully.

How do we get around these difficulties? Let us sketch the method of Stein and Weiss to do this, pointing out the key features of their argument:

(1) *The equivalence of nontangential boundedness and nontangential convergence almost everywhere for harmonic functions*: Suppose $\Gamma^h(x)$ denotes the cone $\{(t,y) \in R_+^{n+1} | 0 < y < h, |x - t| < y\}$. Then we say that a function $F(t,y)$ on R_+^{n+1} is nontangentially bounded at $x \in R^n$ provided that for some $h > 0$, F is bounded on $\Gamma^h(x)$. We call F nontangentially convergent at x provided $\lim_{(t,y) \to (x,0);(t,y) \in \Gamma(x)} F(t,y)$ exists. Then there is the following basic fact: For a harmonic function u on R_+^{n+1} which is nontangentially bounded at each point x of a set $E \subseteq R^n$, u has nontangential limits at a.e. $x \in E$. This is due to Privalov [68] in R_+^2 and to Calderón [10] in R_+^{n+1} for $n > 1$.

(2) *The subharmonicity of powers of $|F|$*: By an ingenious calculation Stein and Weiss showed that if $\alpha \geq (n-1)/n$ and $|F(x,y)| > 0$, then $\Delta(|F|^\alpha)(x,y) \geq 0$. This means that $|F|^\alpha$ is subharmonic and allows us to pass from H^p-theory when $p \leq 1$ to H^p-theory when $p > 1$, as follows. Take $F \in H^p(R_+^{n+1})$ and assume that $p > (n-1)/n$ (the case $p = (n-1)/n$ works with only slight modifications). Then let $(n-1)/n < \alpha < p$ and consider $G = |F|^\alpha$. This function is subharmonic and has

$$\sup_{y > 0} \int_{R^n} G^r(x,y)\, dx < \infty \quad \text{where } r = \frac{p}{\alpha} > 1.$$

For such a function G it is not hard to show that there is a function $g \in L^r(\mathbf{R}^n)$ such that G at any point of R_+^{n+1} is dominated by the appropriate weighted average of the values of g (the weighting depends, of course, on the point of R_+^{n+1}).

(3) *The introduction of the nontangential maximal function*: If f is any function on R_+^{n+1} we set $f^*(x) = \sup_{(t,y) \in \Gamma(x)} |f(t,y)|$ for every $x \in R^n$. f^* is called the nontangential maximal function of f. In the case of the present theorem it will suffice, in view of (1), to show that $F^*(x) < \infty$ for a.e. $x \in R^n$. This is seen as follows: Since G is dominated by averages of the function g as in (2), it turns out that $G^*(x) \leq M(g)(x)$, where M is the Hardy–Littlewood

maximal function of g. But then

$$\int_{R^n} F^*(x)^p\, dx = \int_{R^n} G^*(x)^{p/\alpha}\, dx \le \int M(g)(x)^r$$

$$\le C \int_{R^n} |g(x)|^r\, dx < \infty.$$

This shows that $F^*(x) < \infty$ a.e. on R^n and finishes the proof.

The last step in the above proof, that is, the introduction of the non-tangential maximal function, is one of crucial importance, and we should mention another result related to it. Recall our mentioning the Littlewood–Paley g function. This is defined on R^n by starting with a function $\psi(x)$ which is sufficiently smooth, decays sufficiently rapidly at infinity, and has $\int_{R^n} \psi(x)\, dx = 0$. Then letting

$$\psi_y(x) = y^{-n}\psi(x/y) \quad \text{for } y > 0,$$

we set, for f a function on R^n,

$$g_\psi(f)(x) = \left(\int_0^\infty |f * \psi_y(x)|^2\, \frac{dy}{y} \right)^{1/2}$$

and

$$S_\psi(f)(x) = \left(\int\int_{\Gamma(x)} |f * \psi_y(t)|^2\, \frac{dt\, dy}{y^{n+1}} \right)^{1/2}.$$

Classically, the most basic example occurs when ψ is the gradient of the Poisson kernel for R_+^{n+1} and then

$$S^2(f)(x) = \int\int_{\Gamma(x)} |\nabla u|^2(t, y)y^{1-n}\, dt\, dy,$$

where u is the Poisson integral of f, that is, the function, harmonic in R_+^{n+1}, which has f as its boundary values. For a harmonic function u on R_+^{n+1} we may also define

$$S^2(u)(x) = \int\int_{\Gamma(x)} |\nabla u|^2(t, y)y^{1-n}\, dt\, dy.$$

Now comes the main point. According to a theorem of A. P. Calderón [11] and E. M. Stein [74], if $u(x, y)$ is a harmonic function on R_+^{n+1}, then, except for a set of points $x \in R^n$ of measure zero, $S(u)(x) < \infty$ if and only if $u^*(x) < \infty$. What is the meaning of this result? In the case of $n = 1$ in R_+^2 [85] part of its meaning is given in the following corollary: *If u and v are conjugate harmonic functions, then the set of x for which u and v approach nontangential limits at x differs only by a set of measure 0.* This remarkable result is a consequence of the Calderón–Stein theorem and the fact that, by the Cauchy–Riemann equations, $|\nabla u| = |\nabla v|$, so $S(u) \equiv S(v)$. (In higher dimensions a similar result holds for Stein–Weiss systems of conjugate harmonic functions, and the proof is along similar lines; see [7].)

There is another meaning of this similar behavior of $S(u)$ and u^*, and this was revealed in a result about harmonic functions in R_+^2 ($n = 1$) due

to Burkholder, Gundy, and Silverstein [7]. Their theorem says that for a harmonic u in R_+^2, and for all $p > 0$,

$$c_p \leq \|S(u)\|_{L^p}/\|u^*\|_{L^p} \leq C_p,$$

where the positive constants c_p and C_p are independent of u. This is the global variant of the Calderón–Stein result on finiteness of $S(u)(x)$ and $u^*(x)$, and it has the following interpretation:

If $F(z) \in H^p(R_+^2)$ we saw above that $F^* \in L^p(R^1)$. Clearly, if $F^* \in L^p(R^1)$ and $F(z)$ is holomorphic in R_+^2, then $F \in H^p(R_+^2)$. Now according to Burkholder, Gundy, and Silverstein, if $F(z)$ is holomorphic and $F = u + iv$, then $F \in H^p$ iff $u^* \in L^p(R^1)$. (This is because $S(u) = S(v)$.) In other words, one can tell just by looking at u^* whether or not $F \in H^p$. We need not worry about v. This is a major step in the direction of freeing the theory of H^p from a dependence on the theory of holomorphic functions. (Incidently the proof of this theorem is by arguments involving Brownian motion, so it is important for its method as well as for the end result.)

The last set of results we wish to mention here, due to Charles Fefferman and E. M. Stein [38], showed that we may think of H^p-spaces entirely in terms of real variables with no dependence whatever on harmonic or holomorphic functions. For C. Fefferman–Stein an H^p-"function" is defined by first considering a Schwartz function $\phi(x)$ on R^n such that $\int \phi \neq 0$ and saying that a distribution f on R^n is in H^p provided that the maximal function $f^*(x) = \sup_{y>0} |f * \phi_y(x)|$ belongs to $L^p(R^n)$ (here, $\phi_y(x) = y^{-n}\phi(x/y)$). The class of distributions so defined is proven to be independent of ϕ. Also, if ψ is a suitably nontrivial function in the Schwartz class such that $\int_{R^n} \psi(x)\,dx = 0$, then

$$f \text{ is in } H^p \quad \text{iff} \quad S_\psi(f) \in L^p(R^n),$$

where

$$S_\psi^2(f)(x) = \int\int_{\Gamma(x)} |f * \psi_y(t)|^2 \frac{dt\,dy}{y^{n+1}},$$

and again the choice of ψ is irrelevant. If u is harmonic in R_+^{n+1} and suitably nice (smooth at the boundary and small at infinity), then

$$S(u) = \left(\int\int_{\Gamma(x)} |\nabla u|^2(t,y)y^{1-n}\,dt\,dy\right)^{1/2} \in L^p \quad \text{iff} \quad S_\psi(f) \in L^p,$$

where

$$S_\psi^2(f)(x) = \int\int_{\Gamma(x)} |f * \psi_y(t)|^2 \frac{dy\,dy}{y^{n+1}}$$

and $f(x) = u(x, 0)$. Similarly, if

$$u_\phi^*(x) = \sup_{(t,y)\in\Gamma(x)} |f * \phi_y(t)|,$$

where ϕ is Schwartz and $\int_{R^n} \phi = 1$, and if $u^*(x)$ is the usual nontangential maximal function of u, then

$$u_\phi^* \in L^p \quad \text{iff} \quad u^* \in L^p.$$

At least when $p \geq (n-1)/n$, this newer notion of H^p coincides with the Stein–Weiss notion. Fefferman and Stein also show that Calderón–Zygmund singular integrals preserve these H^p-spaces, so we are justified once again in regarding H^p-spaces as being the right replacement for L^p such that maximal and singular integral operators map the spaces to L^p when $p \leq 1$.

Thus, the class of H^1-functions is a space near L^1 which is invariant under singular integrals. There is also a class of functions near L^∞ invariant under the Calderón–Zygmund operators, namely, BMO(R^n). This is the space of functions, introduced by John and Nirenberg [56], satisfying

$$\frac{1}{m(Q)} \int_Q |\phi(x) - \phi_Q|\, dx \leq C,$$

where ϕ_Q denotes the mean value of ϕ over the cube Q, and C is independent of Q. These functions of *bounded mean oscillation* are a priori only assumed to be locally integrable, but in fact are locally in the exponential class, as expressed by the John–Nirenberg inequality

$$\frac{1}{|Q|} \int_Q \exp\frac{c|\phi(x) - \phi_Q|}{\|\phi\|_*} \leq C,$$

where

$$\|\phi\|_* = \sup_Q \frac{1}{|Q|} \int_Q |\phi(x) - \phi_Q|\, dx$$

is the BMO norm of ϕ.

There are a number of very useful characterizations of BMO, and in order to discuss the one we have in mind it will be helpful to first consider a basic result of Lennart Carleson [15]. A positive measure μ on R^{n+1}_+ is called a Carleson measure provided that $\mu(S(Q)) \leq Cm(Q)$ for all cubes Q in R^n, where

$$S(Q) = \{(x,y)|x \in Q,\ 0 < y < \text{side length}(Q)\}.$$

Carleson proved that these measures are exactly the ones for which

$$\int\int_{R^{n+1}_+} |u(x,y)|^p\, d\mu \leq C \int_{R^n} |f(x)|^p\, dx, \qquad p > 1,$$

if u is the Poisson integral of f.

Charles Fefferman [35] was able to prove that if u is the Poisson integral of a function $\phi(x)$ on R^n, then

$$\phi(x) \in \text{BMO}(R^n) \text{ iff } |\nabla u|^2(x,t)t\, dt\, dx \text{ is a Carleson measure in } R^{n+1}_+.$$

He showed, using this characterization, that the dual space of H^1 was BMO. Somewhat later, R. R. Coifman [27] found a particularly striking proof of this duality by using his constructive proof of a decomposition theorem for the space H^1.

This decomposition provides an enormously powerful tool for attacking problems relating to H^1. It says that any H^1-function can be written as $f = \sum \lambda_k a_k$, where λ_k are scalars such that $\sum |\lambda_k| \leq C\|f\|_{H^1}$, and where the a_k are H^1-atoms, i.e., a_k is supported in a cube Q_k, has mean value 0 over Q_k, and satisfies $\|a_k\|_{L^\infty} \leq 1/|Q_k|$. From this it is clear that a BMO function

ϕ acts on H^1, since, if $f \in H^1$, $f = \sum \lambda_k a_k$ is an atomic decomposition of f, then

$$\left| \int f\phi \right| = \left| \sum \lambda_k \int_{Q_k} a_k \phi \right| = \left| \sum \lambda_k \int_{Q_k} a_k(\phi - \phi_{Q_k}) \right|$$

$$\leq \sum |\lambda_k| \frac{1}{|Q_k|} \int_{Q_k} |\phi - \phi_{Q_k}| \, dx$$

$$\leq \|\phi\|_* \sum |\lambda_k| \leq C\|\phi\|_* \|f\|_{H^1}.$$

The space BMO has been under intensive study in the past ten years or so. It turns out, even in the classical domain (i.e., the unit disk), that properties of BMO, its relation to Carleson measures and A_p weights, etc., become very powerful tools in dealing with problems arising in function algebras (e.g., Chang [20], Marshall [63], and Sarason [73] for the Douglas Problem), univalent functions (e.g. Baernstein [1], Pommerenke [67], quasi-conformal maps [69]), and many other topics. For these and other developments about BMO, the reader is referred to [72, 46], and also the survey article of L. Carleson [19].

We have mentioned many topics here from classical real variables and Fourier analysis. These have one common thread running through them. That is, they all deal with operators indexed by one parameter or are invariant with respect to a one-parameter family of dilations on R^n. All the results therefore have a "one-dimensional" quality, since the dimension n seems to play no role at all. In the rest of this article we deal with the theory in several parameters, treating the maximal function, singular integrals, Littlewood–Paley theory, A^p-spaces, etc., in this new context.

PART II. THE THEORY FOR THE CASE OF SEVERAL PARAMETERS

1. Differentiation theory and the maximal function. In what follows we are usually concerned with operators acting on functions on $R^{n_1} \times R^{n_2} \times \cdots \times R^{n_k}$ invariant under the full k-parameter family of dilations

$$(x_1, x_2, \ldots, x_k) \xrightarrow{\tau_{\delta_1, \delta_2, \ldots, \delta_k}} (\delta_1 x_1, \delta_2 x_2, \ldots, \delta_k x_k).$$

We sometimes call the theory of these operators "a product theory". Naturally there are many other interesting families of dilations that we could consider. For example, in R^3, we briefly consider the two-parameter family

$$\sigma_{\delta_1, \delta_2}(x_1, x_2, x_3) = (\delta_1 x_1, \delta_2 x_2, \delta_1 \delta_2 x_3).$$

Most of the time we treat the product theory with respect to the full k-parameter family $\tau_{\delta_1, \delta_2, \ldots, \delta_k}$.

The maximal function invariant under the action of the dilations $\tau_{\delta_1, \delta_2, \ldots, \delta_k}$ is the "strong maximal function" M_s of Jessen–Marcinkiewicz–Zygmund. To define it let $B_1, B_2, \ldots, B_k$ denote the unit balls of $R^{n_1}, R^{n_2}, \ldots, R^{n_k}$, respectively. Then

$$M_s(f)(x_1, x_2, \ldots, x_k)$$

$$= \sup_{\delta_1, \delta_2, \ldots, \delta_k > 0} \frac{1}{m(\Pi_{i=1}^k \delta_i B_i)} \int_{\Pi_{i=1}^k (\delta_i B_i)} |f(x_1 + t_1, x_2 + t_2, \ldots, x_k + t_k)|$$

$$\cdot \, dt_1 \, dt_2 \cdots dt_k.$$

This maximal function with respect to products of balls, or "rectangles", behaves very differently from the Hardy–Littlewood maximal function. It is quite possible for a function $f \in L^1(R^{n_1} \times R^{n_2} \times \cdots \times R^{n_k})$ to have $M_s(f)(x) = \infty$ everywhere. The natural question to ask is then "What is the least stringent restriction on the size of a function which guarantees that $M_s(f)(x) < \infty$ for a.e. $x \in \prod_{i=1}^{k} R^{n_i}$?" The answer, which depends on k, is that $f \in L(\log L)^{k-1}$, i.e.,

$$\int_{\Pi R^{n_i}} |f(x)|(1 + \log^+ |f(x)|)^{k-1} \, dx < \infty.$$

In this case we also have

$$m \left\{ x \in \prod_{i=1}^{k} R^{n_i}, \ |x| < 1 \mid M_s(f)(x) > \alpha \right\} \leq \frac{C}{\alpha} \|f\|_{L(\log L)^{k-1}}.$$

This is the basic result for the operator M_s proven in the 1930s by Jessen, Marcinkiewicz, and Zygmund [54]. The proof is quite simple and proceeds as follows: we let M_i be the Hardy–Littlewood maximal operator in the ith factor space R^{n_i}, i.e.,

$$M_i f(x) = \sup_{r>0} \frac{1}{m(rB_i)} \int_{rB_i} |f(x_1, x_2, \ldots, x_{i-1}, x_i + t_i, x_{i+1}, \ldots, x_k)| \, dt_i.$$

Then a simple application of Fubini's theorem applied to the average of $|f|$ over a rectangle shows immediately that

$$M_s f(x) \leq M_1 \circ M_2 \circ \cdots \circ M_k f(x),$$

where $\circ$ denotes composition of operators. Now we know from the one-parameter theory that each M_i is bounded on L^p, $p > 1$, and immediately this gives the boundedness of M_s on L^p for $p > 1$. The sharp result is obtained by using the fact that the one-parameter maximal operator maps $L(\log L)^j$ boundedly to $L(\log L)^{j-1}$, a result obtained through interpolating between the estimates

(1) $$m\{M_i f > \alpha\} \leq (C/\alpha)\|f\|_1$$

and

(2) $$\|M_i f\|_\infty \leq \|f\|_\infty.$$

Once we know this, each application of one of the M_i "loses a log", so that starting with $f \in L(\log L)^{k-1}$,

$$M_2 \circ M_3 \circ \cdots \circ M_k f \in L^1(|x| < 1),$$

and now M_1 satisfies estimate (1) on L^1 so that

$$m\{|x| < 1 \mid M_1(M_2 \circ M_3 \circ \cdots \circ M_k f)(x) > \alpha\} \leq (C/\alpha)\|f\|_{L(\log L)^{k-1}}.$$

This is a very short argument, and the result that it obtains is sharp, so that one might suspect that this is the end of the story of M_s. This is not the case for several reasons. For example, if we change the operator M_s slightly, we

may find that the preceding argument no longer applies to the new operator. Let us give two such perturbations of M_s for which this is the case.

First, imagine that we have replaced Lebesgue measure on $\prod_{i=1}^{k} R^{n_i} = R^N$ with some measure μ and we now wish to view all operators in terms of this measure, as though μ were the only existing measure on R^N. Then the strong maximal function would be

$$M_s^\mu f(x) = \sup_{x \in R} \frac{1}{\mu(R)} \int_R |f| \, d\mu,$$

where the sup is taken over all rectangles R (i.e., products of balls in the spaces R^{n_i}) containing the point $x \in R^N$. If μ is a product of measures arising from the factor spaces R^{n_i}, then the argument of Jessen–Marcinkiewicz–Zygmund works to prove estimates for M_s^μ. If μ is a measure which "looks very much" like Lebesgue measure, but is not a product measure, then the old methods no longer work.

Next, suppose we return to averages taken with respect to Lebesgue measure, but we change the operator M_s by allowing the rectangles to tilt a little. To be precise, consider the case in R^2 where $B = \{$all rectangles with longest side making an angle of 2^{-k}, for some integer $k > 0$, with the positive x-axis$\}$ (in other words, we allow rectangles to tilt at the angles $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}, \ldots$ only, and the side lengths are arbitrary). If we set

$$M f(x) = \sup_{x \in R \in B} \frac{1}{m(R)} \int_R |f| \, dy,$$

then M, we might feel, should act quite a bit like M_s since the directions of the allowed rectangles are converging so rapidly to one fixed direction. Again, here, the iteration argument will not suffice, and something more is needed.

There is another situation which cannot be handled by the Jessen–Marcinkiewicz–Zygmund method, and this is illustrated most simply as follows: In R^3 consider the family of rectangles S whose sides are parallel to the axes and whose side lengths are δ_1, δ_2, and $\delta_1\delta_2$. This is a two-parameter family of rectangles in R^3. Hence, if our philosophy is right, and it is the number of parameters that is important, rather than the dimension of the underlying space, the corresponding maximal function

$$M_z f(x) = \sup_{x \in R \in S} \frac{1}{m(R)} \int_R |f| \, dy$$

should satisfy

$(\sim)$ $m\{x \in R^3 \mid |x| < 1, \ M_z(f)(x) > \alpha\} \leq (C/\alpha)\|f\|_{L \log L(R^3)};$

the method of iteration yields only the estimate $(\sim)$ with $\|f\|_{L(\log L)^2}$ on the right side.

Even as far as L^p-estimates are concerned, sometimes a simple iteration argument will not work. A good example of this is the weighted norm inequalities for multiparameter maximal functions. If M_s is the strong maximal operator, and we are interested in those weights $w(x) \geq 0$ for which

$$\int [M_s f(x)]^p w(x) \, dx \leq \int |f(x)|^p \, w(x) \, dx,$$

where $p > 1$, then the simple iteration argument given above shows that these are exactly those weights which are uniformly in the Muckenhoupt class $A^p(R^{n_i})$ in the ith variable for each $1 \leq i \leq k$. Equivalently, w satisfies an A^p-condition with respect to rectangles:

$$\left(\frac{1}{m(R)} \int_R w\right) \left(\frac{1}{m(R)} \int_R w^{-1/(p-1)}\right)^{p-1} \leq C$$

for each R a product of balls in R^{n_i}.

In other cases of multiparameter maximal functions, iteration does not yield the desired answers. For instance, if again B is the two-parameter family of rectangles in R^3 of sides δ_1, δ_2, and $\delta_1 \delta_2$, where $\delta_1, \delta_2 > 0$ are arbitrary, and M_z is the corresponding maximal operator, then one would expect a weighted inequality

$$\int_{R^3} (M_z f)^p w \leq C \int_{R^3} |f|^p w$$

if and only if w satisfies the A^p-condition, but only over the relevant rectangles in B:

$$\left(\frac{1}{m(R)} \int_R w\right) \left(\frac{1}{m(R)} \int_R w^{-1/(p-1)}\right)^{p-1} \leq C \quad \forall R \in B.$$

The proofs of the above results depend on a certain method having to do with covering lemmas for sets more general than balls. The basic result in this direction is the following:

THE COVERING LEMMA FOR RECTANGLES. *Let $\{R_j\}$ be an arbitrary collection of rectangles in $R^N = \prod_{i=1}^k R^{n_i}$. Then there exists a subcollection $\{\tilde{R}_j\}$ of $\{R_j\}$ satisfying*

(1) $$m\left(\bigcup \tilde{R}_j\right) \geq cm\left(\bigcup R_j\right)$$

and

(2) $$\left\|\sum \chi_{\tilde{R}_j}\right\|_{\exp(L)^{1/(k-1)}} \leq Cm\left(\bigcup R_j\right).$$

See Cordoba–R. Fefferman [33]. The meaning of this result is that, just as in the one-parameter case of balls, we have extracted a subcollection of at least a fixed fraction of the volume of the original rectangles in such a way that the chosen rectangles are sparse. Of course, they are not disjoint as in the one-parameter case, but rather merely sparse of varying degrees depending on the number of parameters k.

To see why the Orlicz norm $\exp(L)^{1/(k-1)}$ should appear, let us observe that balls $\tilde{B}_j$ are disjoint iff $\|\sum \chi_{\tilde{B}_j}\|_{L^\infty} \leq 1$, and L^∞ is the dual of L^1. Since in the k-parameter case the basic estimate on the maximal operator M_s is on $L(\log L)^{k-1}$ functions (L^1 when $k = 1$), we would expect the norm applied to $\sum \chi_{\tilde{R}_j}$ to be that of the dual class of $L(\log L)^{k-1}$, which is $\exp(L)^{1/(k-1)}$. The method of proof of this covering lemma is induction on the number of parameters k, and the important feature of this is that it provides a general method for controlling higher-parameter maximal functions by simpler lower-parameter ones. This is the way the first results on maximal functions with

respect to tilting rectangles, such as M above, were proven. (See Stromberg [**78**] and Cordoba–R. Fefferman [**33**]. See also the exciting article of E. Stein and S. Wainger, *Problems in harmonic analysis related to curvature*, Bull. Amer. Math. Soc. **82** (1978), where another approach to these problems is discussed in detail.) In this case the maximal operator (three-parameter) M is controlled by the classical (two-parameter) operator M_s.

Nowadays many operators are known to be controlled by operators involving fewer parameters. This is also the method used to handle the operator M_s^μ. In fact, if $\mu = w(x)\, dx$, and if w satisfies an A^∞-condition in each of the factor spaces R^{n_i} uniformly (a weight $w(x)$ on R^n is said to be A^∞ provided that for subsets $E \subset Q$, Q a cube, such that $m(E)/m(Q) > 1/2$, then $\int_E w / \int_Q w > \varepsilon$, for some ε independent of E and Q), then M_s^μ is bounded on $L^p(d\mu)$ for $p > 1$. (See R. Fefferman [**41**] and Jawerth–Torchinsky [**53**].) Because of the control of a k-parameter operator by a $(k-1)$-parameter operator, we need only assume that w is uniformly in $A^\infty(R^{n_i})$ for $k-1$ of the i's. This is the key observation that enables us to show [**43**] that for w on R^3 satisfying an A^p-condition for all rectangles of side lengths δ_1, δ_2 and $\delta_1\delta_2$, then we have the corresponding weighted norm inequality

$$(*) \qquad \int (M_z f)^p w \le Cp \int |f|^p\, w, \qquad p > 1.$$

This follows easily from the boundedness of

$$M_z^\mu(f)(x) = \sup_{x \in R \in B} \frac{1}{\mu(R)} \int_R |f|\, d\mu \quad \text{on } L^p(d\mu).$$

In turn, this follows because our $w \in A^p(B)$ is uniformly in the class $A^p(R^1)$ in the x and y variables so that $M_s^\mu f \ge M_z^\mu f$ satisfies

$$\int (M_s^\mu f)^p\, d\mu \le Cp \int |f|^p\, d\mu.$$

This is also Cordoba's method of proof [**31**] of the estimate

$$m\{x \in R^3, \ |x| < 1 \mid M_z f(x) > \alpha\} \le (C/\alpha)\|f\|_{L \log L(R^3)}.$$

Thus, although there is still quite a number of unsettled questions in differentiation theory of several parameters, the method of controlling k-parameter operators by $(k-1)$-parameter operators started by the covering lemma for rectangles is quite a useful machine.

2. Singular integrals. The reader will recall that in the introduction we discussed the one-parameter theory of singular integrals due to Calderón–Zygmund [**14**]. This involved convolution operators $Tf = f * K$ on R^n, where K satisfied

(1) $$|K(x)| \le C/|x|^n$$

(2) $$|\nabla K(x)| \le C/|x|^{n+1},$$

and

(3) $$\int_{\alpha < |x| < \beta} K(x)\, dx = 0 \quad \text{for all } 0 < \alpha < \beta.$$

We wish to formulate here the generalization of the Calderón–Zygmund theory to the two-parameter setting for functions on $R^n \times R^m$ introduced in R. Fefferman–E. M. Stein [**44**]. The basic example is the so-called double Hilbert transform when $n = m = 1$. In this case the kernel is $K(x, y) = 1/xy$, and we are, of course, in the product case, where the kernel is invariant with respect to the dilations $(x, y) \rightarrow (\delta_1 x, \delta_2 y)$ for all $\delta_1, \delta_2 > 0$; i.e.,

$$K(x/\delta_1, y/\delta_2)\delta_1^{-1}\delta_2^{-1} = K(x, y).$$

Notice in this example that the kernel K splits into a product of the x and y variables separately:

$$K(x, y) = K_1(x) \cdot K_2(y),$$

where K_1 and K_2 are Calderón–Zygmund kernels in the x or y variable. This is a great simplification for most of the problems we consider. In this case a simple iteration argument analogous to that used in handling the strong maximal operator often suffices. When $K(x, y)$ does not split into such a product, things are trickier, and we shall concentrate on this general case.

If we look at the integral defining the double Hilbert transform,

$$Hf(x, y) = \int \int_{R^2} f(x - s, y - t)\, \frac{ds\, dt}{st},$$

then it will not be absolutely convergent, even for the simplest of functions $f(x, y)$. So there is initially a problem of defining H and other singular integrals rigorously. This is done by principal value integrals, i.e., we consider the truncated integral

$$H_{\varepsilon_1, \varepsilon_2} f(x, y) = \int \int_{\substack{|s| > \varepsilon_1 \\ |t| > \varepsilon_2}} f(x - s, y - t)\, \frac{ds\, dt}{st}$$

and then prove that, as $\varepsilon_1, \varepsilon_2$ tend to zero independently, $\lim_{\varepsilon_1, \varepsilon_2 \to 0} H_{\varepsilon_1 \varepsilon_2} f$ exists in L^p-norm or pointwise almost everywhere, and the limit Hf yields a bounded operator on $L^p(R^n \times R^m)$.

Let us begin by asking for the right way to formulate conditions on a kernel $K(x, y)$ generalizing (1)–(3) in the one-parameter case for which we have the boundedness of the operators involved.

The best way to understand these conditions is as follows: On $R^n \times R^m$, $K(x, y)$ will be given so that if we view $K(x, y)$ as a kernel on R^n of the x variable taking on as values functions of the y variable, then K is a one-parameter Calderón–Zygmund kernel in x, taking on values in the space of all Calderón–Zygmund kernels in y. We give the Calderón–Zygmund kernels a norm, $\| \ \|_{CZ}$: If $K(y)$ is a kernel on R^m satisfying

$$\int_{\alpha < |y| < \beta} K(y)\, dy = 0 \quad \forall\, 0 < \alpha < \beta, \qquad |K(y)| \leq \frac{C}{|y|^m},$$

and

$$|K(y + h) - K(y)| \leq C(|h|/|y|)^\eta (1/|y|^m)$$

for some $\eta > 0$, whenever $2|h| < |y|$, then the smallest C which makes all of the above inequalities valid will be called the Calderón–Zygmund norm of K, $\|K\|_{CZ}$.

Now given $K(x,y)$, set $K_x(y) = K(x,y)$ and consider the class C of those K for which

$$(1') \qquad \int_{\alpha < |x| < \beta} K_x \, dx = 0,$$

$$(2') \qquad \|K_x\|_{\mathrm{CZ}} \le C/|x|^n,$$

and

$$(3') \qquad \|K_{x+h} - K_x\|_{\mathrm{CZ}} \le C(|h|/|x|)^\eta (1/|x|^n),$$

and try to show that convolution with these kernels is bounded on L^p. Decoding $(1')$–$(3')$ so that we have conditions defined directly on $K(x,y)$, we find that K must satisfy

$$(\mathrm{i}) \qquad \int_{\alpha < |x| < \beta} K(x,y) \, dx = 0 \quad \text{for each fixed } y \in R^m;$$

$$(\mathrm{ii}) \qquad \int_{\alpha < |y| < \beta} K(x,y) \, dy = 0 \quad \text{for each fixed } x \in R^n;$$

$$(\mathrm{iii}) \qquad |K(x,y)| \le C/|x|^n |y|^m \quad \text{for } x \in R^n, \ y \in R^m;$$

$$(\mathrm{iv}) \qquad |K(x+h,y) - K(x,y)| \le \frac{C}{|y|^m} \left(\frac{|h|}{|x|} \right)^\eta \frac{1}{|x|^n}$$

whenever $2|h| < |x|$, and where $\eta > 0$ is fixed;

$$(\mathrm{v}) \qquad |K(x,y+h) - K(x,y)| \le \frac{C}{|x|^n} \left(\frac{|h|}{|y|} \right)^\eta \frac{1}{|y|^m}$$

whenever $2|h| < |y|$;

(vi) if we define $\Delta_h^{(x)} K(x,y) = K(x+h,y) - K(x,y)$ and $\Delta_k^{(y)} K(x,y) = K(x,y+k) - K(x,y)$ then

$$|\Delta_k^{(y)}[\Delta_h^{(x)} K(x,y)]| \le \frac{C}{|x|^n |y|^m} \left(\frac{|h|}{|x|} \frac{|k|}{|y|} \right)^\eta .$$

For kernels satisfying (i)–(vi), i.e., for $K \in C$, it is proven in [44] that the operators

$$T_{\varepsilon_1, \varepsilon_2} f = f * K_{\varepsilon_1, \varepsilon_2} \quad \text{for } K_{\varepsilon_1, \varepsilon_2}(x,y) = \chi_{|x| > \varepsilon_1}(x) \cdot \chi_{|y| > \varepsilon_2}(y) K(x,y)$$

are uniformly bounded on $L^p(R^n \times R^m)$ for $1 < p < \infty$ and converge in the L^p-norm to an operator T which is therefore bounded on L^p. The proof does not follow the usual Calderón–Zygmund program as outlined in the introduction. The reason for this is that, although there is a procedure for modifying an $L \log L(R^n \times R^m)$ function so that it becomes L^2, the error term, rather than consisting of functions of mean value zero living on disjoint cubes, consists

of functions which have the appropriate cancellation, but are supported on rectangles which have an enormous amount of overlap.

The method of proof used compares the Littlewood–Paley functions of Tf and f, a method which, in the one-parameter case, goes back to Stein [74]. If ψ is some suitably nontrivial function with $\psi \in C_c^\infty(R^n \times R^m)$, $\int_{R^n} \psi(x, y)\, dx = 0$ for each $y \in R^m$, and $\int_{R^m} \psi(x, y)\, dy = 0$ for each $x \in R^n$, then we let

$$S_\psi^2(f)(x, y) = \int\int_{\Gamma_1(x) \times \Gamma_2(y)} |f * \psi_{t_1, t_2}(u, v)|^2 t_1^{-1-n} t_2^{-1-m}\, du\, dv\, dt_1\, dt_2,$$

and, for $\lambda > 1$,

$$g_{\lambda, \psi}^2(f)(x, y) = \int\int_{R_+^{n+1} \times R_+^{m+1}} |f * \psi_{t_1, t_2}(u, v)|^2 \left(\frac{1}{1 + |u - x|/t_1}\right)^{\lambda n}$$
$$\cdot \left(\frac{1}{1 + |v - y|/t_2}\right)^{\lambda m} \cdot t_1^{-1-n} t_2^{-1-m}\, du\, dv\, dt_1\, dt_2,$$

where

$$\psi_{t_1, t_2}(x, y) = t_1^{-n} t_2^{-m} \psi(x/t_1, y/t_2).$$

In [44] it is shown that under assumptions (i)–(vi) on K, we have

$$(\#) \qquad\qquad S_{\psi^* \psi}(Tf)(x, y) \le C g_{\lambda, \psi}^*(f)(x, y)$$

for all $(x, y) \in R^n \times R^m$.

Now we come to the main point. The operators S and g_λ^* are every bit as bad as T is; they are Hilbert-space-valued two-parameter singular integrals. But the above inequalities are valid for virtually any choice of $\psi(x, y)$. If we choose $\psi(x, y)$ to be of the form $\psi_1(x) \cdot \psi_2(y)$, then these singular integrals S and g_λ^* have kernels with some product structure and can be handled by iteration methods involving vector valued functions (see [50 and 44]) to give

$$\|S_\psi(f)\|_{L^p} \sim \|f\|_{L^p} \quad \text{and} \quad \|g_{\lambda, \psi}(f)\|_{L^p} \sim \|f\|_{L^p}$$

for $1 < p < \infty$ and for λ large enough.

It is interesting to note that $(\#)$ needs to be replaced by a different inequality in case we desire weighted norm inequalities,

$$\int\int_{R^n \times R^m} |Tf|^p\, w\, dx\, dy \le Cp \int\int_{R^n \times R^m} |f|^p\, w\, dx\, dy,$$

where w satisfies a uniform A^p-condition in each variable separately. To do this, it turns out that $(\#)$ must be replaced by an inequality with the same left side, but whose right side becomes

$$C \left(\int_0^\infty \int_0^\infty M_s^2(|f * \psi_{t_1, t_2}(x, y)|) \frac{dt_1\, dt_2}{t_1 t_2}\right)^{1/2},$$

a vector maximal operator, which can be studied by iterating the techniques of Charles Fefferman and E. M. Stein in [38].

The deepest of the estimates for singular integrals in the two-parameter setting has to do with the existence, pointwise almost everywhere, of the principal value integral in question,

$$\lim_{\varepsilon_1, \varepsilon_2 \to 0} T_{\varepsilon_1, \varepsilon_2}(f)(x, y),$$

where

$$T_{\varepsilon_1, \varepsilon_2}(f)(x, y) = \int \int_{\substack{|x'| > \varepsilon_1 \\ |y'| > \varepsilon_2}} f(x - x', y - y') K(x', y') \, dx' \, dy'.$$

To prove that for $f \in L^p(R^n \times R^n)$ this limit exists a.e., it turns out to be sufficient to get an estimate of the form

$$(*) \qquad \left\| \sup_{\varepsilon_1, \varepsilon_2 > 0} |T_{\varepsilon_1, \varepsilon_2}(f)| \right\|_{L^p} \le C_p \|f\|_{L^p}, \qquad 1 < p < \infty.$$

The proof of $(*)$ turns out to be related to an interesting piece of work on differentiation of integrals due to Stein [**74**], and this is the following: Suppose on R^n we set, for $f \in C_c^\infty(R^n)$,

$$A_r(f)(x) = \int_{S^{n-1}} |f(x + rt)| \, d\sigma(t),$$

where S^{n-1} is the unit sphere in R^n and $d\sigma$ is the element of surface measure on S^{n-1}. Let

$$m(f)(x) = \sup_{r > 0} A_r(f)(x).$$

Then according to Stein we have the a priori estimate

$$\|m(f)\|_{L^p(R^n)} \le C_{p,n} \|f\|_{L^p(R^n)}$$

whenever $p > n/(n-1)$ and $n > 2$. The tools necessary for this proof are a type of interpolation using so-called analytic families of operators and quadratic functionals resembling the Littlewood–Paley g function. The same techniques are required to obtain $(*)$. The details are too complicated to discuss here, but we shall be content to discuss briefly the correct Littlewood–Paley g function in this setting.

The way to understand it best is to return to the one-parameter setting of Calderón–Zygmund kernels $K(x)$ on R^n. If we set $K_\varepsilon(x) = \chi_{|x| > \varepsilon}(x) K(x)$ for $\varepsilon > 0$ and $T_\varepsilon f = f * K_\varepsilon$, then the usual proof that the operator $T^* f(x) = \sup_{\varepsilon > 0} |T_\varepsilon f(x)|$ is bounded on $L^p(R^n)$ depends on the inequality

$$T^* f(x) \le C\{M(Tf)(x) + Mf(x)\},$$

where $Tf = f * K$, which is known to exist as the L^p-norm limit, as $\varepsilon \to 0$, of $T_\varepsilon f$. This, in turn, is proven by picking a positive bump function $\phi(x) \in C_c^\infty(R^n)$ with $\int_{R^n} \phi = 1$ and noting that K_ε looks very much like $K * \phi_\varepsilon = T(\phi_\varepsilon)$, where $\phi_\varepsilon(x) = \varepsilon^{-n} \phi(x/\varepsilon)$. In fact,

$$(**) \qquad |K * \phi_\varepsilon(x) - K_\varepsilon(x)| \le C\Phi_\varepsilon(x),$$

where Φ_ε is a radial function of x which decreases as $|x|$ increases and has the property that $\int \Phi_\varepsilon = 1 \ \forall \varepsilon > 0$. It is trivial to see that convolution with such functions as Φ_ε is dominated by the maximal Hardy–Littlewood operator:

$$|f * \Phi_\varepsilon(x)| \leq Mf(x).$$

It follows from (∗∗) that

$$|(f * K) * \phi_\varepsilon(x) - f * K_\varepsilon(x)| \leq CMf(x),$$

so

$$|T_\varepsilon f(x)| \leq |\phi_\varepsilon * (Tf)(x)| + CMf(x) \leq C\{M(Tf)(x) + M(f)(x)\}.$$

It turns out that the main estimate above—that $\sup_{\varepsilon>0} f * |(K * \phi_\varepsilon) - K_\varepsilon|$ is bounded on L^p—can be improved by replacing the $\sup_\varepsilon$ by a quadratic expression, our desired Littlewood–Paley type function:

$$\left(\int_0^\infty |f * \{(K * \phi_\varepsilon) - K_\varepsilon\}(x)|^2 \, \frac{d\varepsilon}{\varepsilon} \right)^{1/2}.$$

Even this is bounded on $L^p(R^n)$, as can be seen by a little computation, which shows that it is a convolution operator whose kernel $L(x)$ has values in $L^2((0, \infty), d\varepsilon/\varepsilon)$ and satisfies the Calderón–Zygmund assumptions. It is this type of quadratic functional whose L^p-boundedness in the two-parameter setting is responsible for the L^p-boundedness of $\sup_{\varepsilon_1, \varepsilon_2} |T_{\varepsilon_1, \varepsilon_2} f(x, y)|$.

We should also point out just in the case of the strong maximal operator on $R^n \times R^m$, there are also weak type inequalities for $f \in L(\log L)$:

$$m\{(x, y) \in R^n \times R^m \, 1 \big| \ |x|, |y| \leq 1, \ |T^* f(x, y)| > \alpha\}$$

$$\leq (C/\alpha)\|f\|_{L(\log L)(R^n \times R^m)}.$$

These are along the same lines as the L^p inequalities for $1 < p < \infty$; for more details we refer the reader to [43] (also see the deep work of C. Fefferman [36] for the case of the double Hilbert transform).

3. H^p-theory on the polydisc. In this section we very briefly mention some results of M. P. and P. Malliavin [62] and Gundy-Stein [50] about H^p-theory on product domains.

Recall the notation we have used in the introduction: Suppose u is a function defined on R^n. Let u^* denote the nontangential maximal function of u, and let $S(u)$ denote the Lusin area integral function of u. Then we have the qualitative statement about the relation between u^* and $S(u)$ (theorem of Calderón–Stein):

(1) $\{x \in R^n \colon u^*(x) < \infty\} \overset{\text{a.e.}}{=} \{x \in R^n \colon S(u)(x) < \infty\},$

and also the quantitative statement (theorem of Burkholder–Gundy–Silverstein [7] and C. Fefferman–Stein [44]):

(2) $u^* \in L^p$ iff $S(u) \in L^p$ with $\|u^*\|_p \sim \|S(u)\|_p$ for all $0 < p < \infty$.

Because of (2) we have adopted $u^* \in L^p$ as the definition for the space $H^p(R^{n+1}_+)$.

In the late 1970s, both (1) and (2) have been generalized to product do-
mains. To make the statements on product domains more explicit, we intro-
duce more notation. Let D denote the unit disc $\{z \in C: |z| < 1\}$, $D^2 = D_1 \times D_2$, the bidisc with variables $z_1, z_2 \in D$, respectively. For each point θ in the
distinguished boundary T^2 of D^2, $\theta = (e^{i\theta_1}, e^{i\theta_2})$, we set $\Gamma(\theta) = \Gamma(\theta_1) \times \Gamma(\theta_2)$,
the product cone with vertex at θ. For a function u defined on T^2, if we let
$u(z_1, z_2)$ denote its biharmonic extension (i.e., harmonic in both z_1, z_2 in-
dependently) to D^2, then we can define, similar to the situation R_+^{n+1}, the
nontangential maximal function of u as

$$u^*(\theta) = \sup_{(z_1, z_2) \in \Gamma(\theta)} |u(z_1, z_2)|.$$

The "area integral" of u will be the sum

$$S^2(u) = S_{12}^2(u) + S_1^2(u) + S_2^2(u) + |u(0,0)|^2,$$

where

$$S_{12}^2(u)(\theta) = \int_{\Gamma(\theta)} |\nabla_1 \nabla_2 u|^2 \, dm$$

and

$$S_1^2(u)(\theta) = \int_{\Gamma(\theta_1)} |\nabla_1 u(\cdot, 0)|^2 \, dm_1$$

(here dm_j is Lebesgue measure on the disc D_j and $dm = dm_1 \, dm_2$).

Suppose we restrict ourselves to the one-dimensional case again. One way
to establish the part $\{e^{i\theta} \in T: u^*(\theta) < \infty\} \subset \{S(u)(\theta) < \infty\}$ (up to a set of
measure zero) of statement (1) is to establish the following inequality (cf. [39],
[50] and also the expository article [47] from where the following explanation
is taken);

$$(3) \qquad m_1(S(u) > \lambda) \le \frac{C}{\lambda^2} \int_{u^* \le \lambda} |u^*(\theta)|^2 \, dm_1 + cm_1(u^*(\theta) > \lambda).$$

The original C. Fefferman–Stein strategy for proving (3) is based on the fol-
lowing observation: Consider the set

$$G = \{e^{i\theta}: u^*(\theta) \le \lambda\},$$

and the region

$$G^+ = \bigcup_{\theta \in G} \Gamma(\theta).$$

To establish (3) we need only estimate the measure of the set $\{S(u)(\theta) > \lambda, e^{i\theta} \in G\}$. Now observe that, on G^+, we have $|u| \le \lambda$. The boundary
of G^+ consists of a "sawtooth" type region which can be approximated by
Lispchitz regions and on which Green's Theorem can be applied (based on the
identity $\Delta u^2 = 2|\nabla u|^2$). Thus, we can relate $\int_G S^2(u) \, dm_1$ to $\int_{\partial G^+} |u|^2 \, dm_1$
and obtain (3). (For more details see [39, 50].)

On the bidisc the corresponding region G^+ has quite a complicated bound-
ary, and it is not clear how to apply Green's Theorem in this domain. In [62]
M. P. and P. Malliavin overcame this geometric difficulty of the proof by some
delicate and complicated algebraic arguments.

Essentially what they did is this: instead of applying Green's Theorem in G^+, they considered some function $u^2 \tilde{\chi}_{G^+}$, where $\tilde{\chi}_{G^+}$ is a smooth version of the characteristic function χ_{G^+}, and applied Stokes' theorem to $u^2 \tilde{x}_{G^+}$ on the entire domain D^2. They established that

$$\{u^*(\theta) < \infty\} \underset{\text{a.e.}}{\subseteq} \{S(u)(\theta) < \infty\}$$

on the bidisc. Their techniques were later generalized and simplified by Gundy and Stein ([**50**], see also [**49**]) to establish (3) and the full scope of (1) and (2). Statement (2) was actually verified in [**52**] in a much more general setting. Namely, suppose we take any $\phi \in S(R^2)$ (S denotes the Schwartz class) with $\int \phi \, dx \neq 0$ and let

$$u^*(\theta) = \sup_{\varepsilon > 0} |u * \phi_\varepsilon(\theta)|.$$

Then

$$\|u^*\|_p \sim \|S(u)\|_p \quad \text{for all } 0 < p < \infty.$$

Based on this we may henceforth identify the class of functions u with $u^* \in L^p$ as the "real-variable" version of the definition for functions in H^p for product domains.

4. BMO on product domains. A locally integrable function ϕ is of *bounded mean oscillation* (BMO) on R^n if

$$\|\phi\|_* = \sup_Q \frac{1}{m(Q)} \int_Q |\phi(x) - m_Q(\phi)| \, dx,$$

where the Q are cubes in R^n, is finite, where $m_Q(\phi)$ denotes the mean value of ϕ over Q. The space BMO was introduced by John and Nirenberg [**56**] and has been used in many different contexts (e.g., John [**55**], Moser [**65**]). For our purposes we only mention C. Fefferman's fundamental duality theorem on BMO(R^n) (i.e., BMO(R^n) is the dual space of $H^1(R^{n+1}_+)$) and report some of our efforts to generalize this theorem to the setting of product domains. As the reader may clearly see, our generalizations are so far incomplete. (The main deficiency is that the characterization we had for BMO functions on product domains lacks the clear and clean geometrical description of the original definition of BMO(R^n) as given above). Nevertheless, there are some positive results which indicate that our approach is in the right direction. We list some of these results in §7.

Before we restate C. Fefferman's theorem, we remark that, as a consequence of the John–Nirenberg inequality (see the introduction) on BMO, we have

$$\|\phi\|_*^p \approx \sup_Q \frac{1}{m(Q)} \int_Q |\phi(x) - m_Q(\phi)|^p \, dx$$

for all $p > 0$. As we shall see, the case $p = 2$ is especially interesting, since it relates the local behavior of functions in BMO more closely to other "square" functions in H^p-theory.

If ϕ is any locally integrable function satisfying

$$\int_{R^n} \frac{|\phi(x)|}{1 + |x|^{n+1}} \, dx < \infty,$$

let $\phi(x,t)$ denote its Poisson extension to $R_+^{n+1} = \{(x,t): x \in R^n, \ t > 0\}$ ($\phi(x,t) = (P_t * \phi)(x)$). Recall that a measure μ is a *Carleson measure* if $\mu(S(Q)) \leq C|Q|$ for all cubes Q, where

$$S(Q) = \left\{ (x,t) \in R_+^{n+1}, \ \prod_{i=1}^{n}(x_i - t, x_i + t) \subset Q \right\}.$$

THEOREM. *The following three conditions on ϕ are equivalent:*
(a) ϕ *is in* BMO.
(b) $\phi = \phi_0 + \sum_{j=1}^{n} R_j(\phi_j)$, *where* $\phi_0, \phi_1, \ldots, \phi_n \in L^\infty$, *where* R_j *are the Riesz transforms defined by*

$$(\widehat{R_j f})(\xi) = (i\xi_j/|\xi|)\hat{f}(\xi).$$

(b') *The linear functional* $f \to \int_{R^n} f(x)\phi(x)\,dx$ *is bounded on* H^1.
(c) $t|\nabla\phi(x,t)|^2\,dx\,dt$ *is a Carleson measure on* R_+^{n+1}, *where*

$$|\nabla\phi(x,t)|^2 = \left|\frac{\partial\phi}{\partial t}\right|^2 + \sum_{j=1}^{n}\left|\frac{\partial\phi}{\partial x_j}\right|^2.$$

In the above theorem, (b) and (b') are equivalent by a function-theoretical argument.

A direct generalization of this theorem to two-parameter product domains may be stated as follows. (For simplicity we will state the result in $R_+^2 \times R_+^2$. One may state the same result in $R_+^{n+1} \times R_+^{m+1}$ by changing intervals in R to cubes in R^n and R^m.)

In the following we suppose ϕ to be a locally integrable function on R^2 satisfying

$$\int_{R^2} \frac{|\phi(x)|\,dx}{(1 + |x_1|^2)(1 + |x_2|^2)} < \infty,$$

where $x = (x_1, x_2) \in R^2$, and let $\phi(x,t)$ denote the biharmonic extension of ϕ to $R_2^+ \times R_2^+$ ($\phi(x,t) = (P_{t_2} * P_{t_1} * \phi)(x)$). Then we have:

PROPOSITION. *The following conditions on ϕ are equivalent:*
(a) ϕ *satisfies*

$$\frac{1}{m(R)} \int_R |\phi(x) - \phi_I(x_2) - \phi_J(x_1) + \phi_R|^2\,dx \leq C$$

for every rectangle $R = I \times J$ on R^2, where

$$\phi_I(x_2) = \frac{1}{m(I)} \int_I \phi(x_1, x_2)\,dx_1,$$

$$\phi_J(x_1) = \frac{1}{m(J)} \int_J \phi(x_1, x_2)\,dx_2,$$

and ϕ_R is the mean value of ϕ over R.
(c) *The measure*

$$d\mu_\phi = t_1 t_2 |\nabla_1 \nabla_2 \phi(x,t)|^2\,dx_1\,dx_2\,dt_1\,dt_2$$

satisfies

$$\int\int_{S(I)\times S(J)} d\mu_\phi \leq Cm(I \times J)$$

for all I, J intervals in R ($S(I)$ is the Carleson region associated with I).

An immediate question arising from this proposition is this: Does the space of functions described by (a) characterize the dual space of $H^1(R_+^2 \times R_+^2)$? As we have mentioned before, a dual form of C. Fefferman's duality theorem, $\mathrm{BMO}(R) = (H^1(R_+^2))^*$, is the so-called "atomic decomposition" of H^1. (More details of this paper in the next section.) If we formally "analyze" the function space in statement (a), it is not hard to see that it is the dual space of $H_{\mathrm{Rect}}^1(R_+^2 \times R_+^2)$ where $H_{\mathrm{Rect}}^1(R_+^2 \times R_+^2)$ is defined via atoms supported on rectangles as follows:

DEFINITION. A rectangle atom is a function $a(x)$ supported on a rectangle $R = I \times J$ having the property $\|a\|_2 \leq 1/(m(R))^{1/2}$,

$$\int_I a(x_1, x_2)\, dx_1 = 0 = \int_J a(x_1, x_2)\, dx_2 \quad \text{for every } (x_1, x_2) \in R.$$

DEFINITION. $H_{\mathrm{Rect}}^1(R_+^2 \times R_+^2)$ is the space of functions $\sum \lambda_k a_k$ with each a_k a rectangle atom and $\sum_k |\lambda_k| < \infty$.

In other words, we may add the following equivalent statement to the above proposition:

(b) *The linear functional $f \rightarrow \int_{R^2} f(x)\phi(x)\, dx$ is bounded on*

$$H_{\mathrm{Rect}}^1(R_+^2 \times R_+^2).$$

The immediate question is then: Is $H_{\mathrm{Rect}}^1(R_2^+ \times R_2^+)$ the same as the space $H^1(R_+^2 \times R_+^2)$ defined in §3? The question lingered for awhile until L. Carleson [17] constructed an example of a measure μ satisfying the product form of the Carleson measure condition (as in (c)) but is not bounded on H^1. Based on his example, R. Fefferman [6] then constructed a ϕ which satisfies condition (a) but is not even on $L^4(R^2)$ (hence, not in the dual of $H^1(R_+^2 \times R_+^2)$). However, one can still strive to say something positive.

DEFINITION. For each open set $\Omega \subset R^2$, define $S(\Omega)$ to be the (generalized) Carleson region

$$\{(x,t) \in R_2^+ \times R_2^+, \text{ with } (x_1 - t_1, x_1 + t_1) \times (x_2 - t_2, x_2 + t_2) \subset \Omega\}.$$

Then call a positive measure μ defined on $R_2^+ \times R_2^+$ a Carleson measure if it satisfies the condition $\mu(S(\Omega)) \leq C|\Omega|$ for all open sets $\Omega \subset R^2$.

Notice that this definition coincides with the original one-dimensional definition if $\Omega \subset R$. The reason is that: in the real line R each Ω can be written as a disjoint union of open intervals I_j and $S(\Omega) = \bigcup_j S(I_j)$. The situation changes quite a lot from one to two dimensions, mainly because one must consider overlapping rectangles of different lengths and widths contained in Ω.

THEOREM [22, 40, 23, 7]. *The following conditions on ϕ are equivalent:*

(b) $$\phi = \phi_0 + H_{x_1}(\phi_1) + H_{x_2}(\phi_2) + H_{x_1} \circ H_{x_2}(\phi_3),$$

where $\phi_0, \phi_1, \phi_2, \phi_3 \in L^\infty(R^2)$ and H_{x_j} denotes the Hilbert transform in the x_j direction.

(b') *ϕ is in the dual of $H^1(R_+^2 \times R_+^2)$.*

(c) *The measure $d\mu_\phi = t_1 t_2 |\nabla_1 \nabla_2 \phi(x,t)|^2\, dx\, dt$ satisfies the Carleson measure condition for open sets as defined above.*

Here again the equivalence between (b) *and* (b') *is function theoretic.*

The major deficiency of this theorem is that (c) is, in practice, difficult to check. Also, it does not give much insight about the geometric properties of ϕ. To overcome these difficulties, we tried in [**23**] to adopt other square functions to ϕ which are more easily accessible than the gradients of the biharmonic extensions of ϕ. To motivate the reader, we first formulate our result in the one-dimensional "dyadic" form. (The "dyadic" version described below is a special case of the martingale theory. As we have mentioned before, H^p-theory was first developed by Burkholder–Gundy–Silverstein [**7, 8**] in its martingale form. Thus, it is natural to understand BMO through its dyadic analogue.)

Define $\mathrm{BMO}_d(R)$ to be the space of functions satisfying the bounded mean oscillation properties with respect to dyadic intervals $I = [(k-1)/2^n, k/2^n]$, $1 \le k \le 2^n$, n an integer only. Then it is easy to see that $\mathrm{BMO}(R) \subsetneq \mathrm{BMO}_d(R)$ (for example, the function $\phi(x) = 0$ for $x \le 0$, $\phi(x) = \log(1/|x|)$ for $x > 0$ is in $\mathrm{BMO}_d(R) \backslash \mathrm{BMO}(R)$). We also have the duality result that $\mathrm{BMO}_d(R) = (H_d^1(R_+^2))^*$ (where $H_d^1(R_+^2)$ can be defined as the space of functions with

$$h^*(x) = \sup_{\substack{I \text{ dyadic} \\ x \in I}} \left| \frac{1}{m(I)} \int_I h(t)\, dt \right|$$

in L^1; actually H_d^1 has several equivalent definitions), as H^1 does (cf. [**7, 8**], and the section on unconditional bases in this article). There is an alternative way to describe functions in BMO_d, that is, through its Haar series expansion. Fix a dyadic interval I. The Haar function h_I associated with I is

$$h_I(x) = \begin{cases} 1/(m(I))^{1/2} & \text{on left-half of } I, \\ -1/(m(I))^{1/2} & \text{on right-half of } I, \\ 0 & \text{otherwise.} \end{cases}$$

The constant function, together with $\{h_I\}_{I \text{ dyadic}}$, forms an orthogonal basis of $L^2(R)$. Notice that for a fixed dyadic interval I we have

$$(\phi(x) - m_I(\phi))\chi_I(x) = \sum_{J \subset I} C_J h_J(x),$$

where C_J is the Haar coefficient of ϕ w.r.t. h_J. Hence, we may reformulate the definition of $\mathrm{BMO}_d(R)$ (through its L^2-form) as follows.

PROPOSITION. *$\phi \in \mathrm{BMO}_d$ if and only if*

$$\sum_{\substack{J \subset I \\ J \text{ dyadic}}} |C_J|^2 \le C m(I)$$

for all dyadic intervals I.

Thus, if we try to understand BMO via BMO_d, it amounts to finding a "continuous" analogue of the Haar expansion of a function. We will describe two methods to do this. One is via the S_ψ-function (cf. Part I); the other (and maybe the more original) is via the work of B. Maurey [64] and [18, 83, 23] on the unconditional basis of H^1 (see §5). Before we do this we would like to remark that the dyadic version for the bidisc of this proposition was carried out in Bernard [4]. Also there is another point of view to understanding BMO via BMO_d (or H^1 via H_d^1)–that is, very roughly speaking, characterizing BMO functions f through averaging of the translates $(f_t = f(x - t))$ of f which are in BMO_d. For details and more precise information on this latter point of view see [34, 47].

We now begin to describe $BMO(R_+^2 \times R_+^2)$ through the square function S_ψ. If we choose a function $\psi \in C^\infty$ with compact support in $[-1, 1]$ and mean value zero for $x = (x_1, x_2) \in R^2$, $y = (y_1, y_2)$ with $y_i > 0$, denote

$$\psi_y(x) = \frac{1}{y_1} \frac{1}{y_2} \psi\left(\frac{x_1}{y_1}\right) \psi\left(\frac{x_2}{y_2}\right),$$

and normalize ψ so that $\int_0^\infty |\hat{\psi}(\xi)|^2 \, d\xi/\xi = 1$, then, for $f \in C_0^\infty(R^2)$ with $\int f(x_1, x_2) \, dx_1 = \int f(x_1, x_2) \, dx_2 = 0$, we have

$$f(x) = \int \int_{(t,y) \in R_+^2 \times R_+^2} (f * \psi_y)(t) \psi_y(x - t) \frac{dt \, dy}{y_1 y_2}$$

(recall also that

$$S_\psi^2 f(x) = \int \int_{\Gamma(x_1) \times \Gamma(x_2)} |f * \psi_y(t)|^2 \left(\frac{dt \, dy}{y_1^2 y_2^2}\right)).$$

This expression, which can easily be proved by taking Fourier transforms on both sides, can be thought of as a continuous analogue of the Haar expansion of f as follows: For each dyadic rectangle $R = I \times J$, let R_+ denote the region in the bi-upper-half-plane as follows:

$$R_+ = \left\{ (t, y) \in R_+^2 \times R_+^2 \colon t \in R, \ |I|/2 \leq y_1 < |I|, \ |J|/2 \leq y_2 < |J| \right\}.$$

Notice as R runs through all dyadic rectangles in R^2, $\{R_+\}$ forms a pairwise disjoint union of $R_+^2 \times R_+^2$. Thus, if we let

$$f_R(x) = \int \int_{(\ ,y) \in R_+} (f * \psi_y)(t) \psi_y(x - t) \frac{dt \, dy}{y_1 y_2},$$

then

$$f(x) = \sum_{\substack{R \text{ dyadic} \\ \text{rectangles}}} f_R(x)$$

(cf. also A. P. Calderón [12] for this decomposition).

Notice that the properties of f_R can be compared to the Haar function $C_R h_R$ (C_R = Haar coefficient of f w.r.t. h_R) as follows: Each f_R is supported on $\tilde{R} = \tilde{I} \times \tilde{J}$, where $\tilde{I}$ is the interval with the same center and three

times the length of I, and f_R like h_R has the property $\int f_R(x_1, x_2)\, dx_1 = \int f_R(x_1, x_2)\, dx_2 = 0$ for all $(x_1, x_2) \in R^2$. Each f_R is a C^1 function (unlike h_R) with L^2-norm of f_R bounded by

$$
S_R(f) = \left(\int \int_{R_+} |f * \psi_y(t)|^2 \frac{dt\, dy}{y_1 y_2} \right)^{1/2}.
$$

The f_R's are not orthonormal to each other (unlike h_R), but their "almost" orthogonal property is expressed by the simple

LEMMA. *Suppose $R = I \times J$, $R_1 = I_1 \times J_1$ are two dyadic rectangles with $R \cap R_1 \neq \emptyset$. Let*

$$
r(R_1, R) = \left(\frac{\min |I|, |I_1|}{\max |I|, |I_1|}, \frac{\min |J|, |J_1|}{\max |J|, |J_1|} \right)^{3/2}.
$$

Then

$$
S_R^2(f_{R_1}) \leq C r(R, R_1) S_{R_1}^2(f),
$$

where C is a constant depending only on $\|\psi\|_\infty, \|\psi'\|_\infty$.

The continuous analogue of the proposition on BMO_d is the following:

THEOREM. *Suppose $\phi \in L^2(R^2)$ satisfies*

$$
\int \phi(x_1, x_1)\, dx_1 = \int \phi(x_1, x_2)\, dx_2 = 0
$$

for all $(x_1, x_2) \in R^2$. Then the following conditions on ϕ are equivalent.

(a)
$$
\sup_{\Omega \text{ open}} \left\| \sum_{\substack{R \subset \Omega \\ R \text{ dyadic rectangle}}} \phi_R \right\|_2^2 = \|\phi\|_*^2 < \infty;
$$

(b) *ϕ is in the dual of $H^1(R_+^2 \times R_+^2)$;*

(c)
$$
\sup_{\Omega \text{ open}} \frac{1}{m(\Omega)} \sum_{R \subset \Omega} S_R^2(\phi) < \infty.
$$

An immediate corollary is the following representation theorem for functions in BMO.

COROLLARY. *A function ϕ is in $\mathrm{BMO}(R_2^+ \times R_2^+)$ (in the sense of the dual space of $H^1(R_+^2 \times R_+^2)$) if and only if there exist functions $\{b_R\}_R$ and nonnegative real numbers $\{\lambda_R\}_R$, where R is taken over all dyadic rectangles in R^2, such that $\phi = \sum_R \lambda_R b_R$ with support $b_R \subset R$, $\int b_R(x_1, x_2)\, dx_1 = \int b_R(x_1, x_2)\, dx_2 = 0$, for all $(x_1, x_2) \in R^2$, and each b_R is in C^1 with*

$$
\|b_R\|_\infty \leq C, \quad \left\| \frac{\partial b_R}{\partial x_1} \right\|_\infty \leq \frac{C}{|I|}, \quad \left\| \frac{\partial b_R}{\partial x_2} \right\|_\infty \leq \frac{C}{|J|}, \quad \left\| \frac{\partial^2 b_R}{\partial x_1 \partial x_2} \right\|_\infty \leq \frac{C}{|R|}
$$

for $R = I \times J$, and with λ_R satisfying $\sum_{R \subset \Omega} \lambda_R^2 |R| \leq C|\Omega|$ for all open set $\Omega \subset R^2$.

The one-parameter form this corollary has been adopted by A. Uchiyama [80] as the first (small) step in a complicated constructive proof of the C. Fefferman–Stein decomposition of BMO(R^n) (cf. also P. Jones [57] for the decomposition of BMO(R) using methods of complex analysis). We would like also to point out that it is not immediately clear how to apply Uchiyama's method to obtain constructive decompositions of BMO functions in product domains.

5. The atomic decomposition of $H^1(R_+^2 \times R_+^2)$. In this section we wish to give the simplest possible treatment of an atomic decomposition of H^p-functions in the multiparameter setting. To do this we shall restrict our attention to the case $p = 1$ and to the two-parameter space $H^1(R_+^2 \times R_+^2)$ defined previously. We have already seen that there are several equivalent ways to define H^p-spaces (see §3). As in the case of BMO, on the product domain $(R_+^2 \times R_+^2)$, we found out it is somewhat easier to adopt the "square function" (in L^1) definition of H^1. When we are trying to decompose functions in $H^1(R_+^2 \times R_+^2)$, our natural approach is to try to decompose it into L^2-atoms. (See [30] for the equivalence of L^p-atoms $1 < p \leq \infty$ on R_+^{n+1}.) What should an L^2-atom look like on $R_+^2 \times R_+^2$? Our best hope is this: An atom $a(x,y)$ on R^2 is a function supported on a rectangle R such that

$$\int a(x_1, x_2)\, dx_1 = \int a(x_1, x_2)\, dx_2 = 0 \quad \forall (x_1, x_2) \in R^2$$

and

$$\|a\|_{L^2} \leq 1/|R|^{1/2}.$$

Unfortunately, as we have seen in the previous section, convergent sums of such atoms form the space $H^1_{\text{rectangle}}$ which is a proper subspace of

$$H^1(R_+^2 \times R_+^2).$$

In addition, as suggested by our definition of Carleson measure on product domains, the role played by rectangles in the definition of atoms should be replaced by an arbitrary open set of finite measure in R^2. Let us begin to describe the appropriate definition which we will use for our atoms.

To do this we require some notation. Let $\Omega \subseteq R^2$ be an open set of finite measure. Then we set

$$\tilde{\Omega} = \{(x_1, x_2) \in R^2 \mid M_s(\chi_\Omega)(x_1, x_2) > 1/2\}$$

and, for $t_1, t_2 > 0$, $R_{t_1, t_2}(x_1, x_2) = $ the rectangle of side lengths t_1 and t_2 centered at (x_1, x_2). It will be convenient to define

$$A = \{\text{all } \phi \in C_c^\infty(R^2) | \phi \text{ is supported in } (|x_1| < 1) \times (|x_2| < 1) \text{ and satisfies}$$
$$\|\partial^{\alpha+\beta}\phi/\partial x_1^\alpha\, \partial x_2^\beta\|_\infty \leq 10^{10} \text{ for all } \alpha, \beta \text{ such that } |\alpha| + |\beta| \leq 20\}.$$

Finally, we remind the reader that we are using the notation

$$\phi_{t_1, t_2}(x_1, x_2) = t_1^{-1} t_2^{-1} \phi(x_1/t_1, x_2/t_2) \quad \text{for } t_1, t_2 > 0.$$

Below, in the statement of the atomic decomposition for $H^1(R_+^2 \times R_+^2)$, we give three ways of thinking about atoms. First, there is the definition which is dual to our notion of $\mathrm{BMO}(R_+^2 \times R_+^2)$ from the previous section. The second definition says, roughly, that an atom is an L^2-function $a(x)$ supported in an open set $\Omega \subseteq R^2$ which has enough cancellation so that the averages of $a(x)$ over rectangles far away from Ω are all suitably small (much smaller than if $a(x)$ had no cancellation). The third says, more or less, that the square function of $a(x)$ is very small far from Ω.

In order to understand the meanings of these definitions, let us illustrate the second in some detail. Consider, in R^1, an H^1-atom $a(x)$ supported on an interval I centered at the origin. Take a point x far from I, say $|x| > 2|I|$. Then for a bump function ϕ supported in $[-1, +1]$ we have

$$(\sim) \qquad\qquad |a * \phi_t(x)| \le C\frac{|I|}{t^2} = \frac{C}{|I|}\left(\frac{|I|}{t}\right)^2.$$

It turns out that, in applications, the fact that $a(x)$ has mean value 0 over I is not that crucial. What we need is that $\|a\|_2 \le 1/|I|^{1/2}$ and that "averages" of a with respect to suitably dilated bump functions are very small, as given in $(\sim)$.

In fact these two properties of $a(x)$ easily imply that $a \in H^1$. Letting

$$a^*(x) = \sup_{t>0} |a * \phi_t(x)|,$$

we have

$$\int_{R^1} a^*(x)\,dx = \int_I a^*(x)\,dx + \int_{I^c} a^*(x)\,dx$$

$$\le \left(\int_I M^2(|a|)\,dx\right)^{1/2} |I|^{1/2} + C|I| \int_{2|I|}^{\infty} \frac{dx}{x^2} \le C'.$$

Finally, before stating the atomic decomposition, there is one last definition: A pre-atom $b_R(x_1, x_2)$ is a function supported on the double of a dyadic rectangle $R = I \times J$ so that

$$\int_I b_R(x_1, x_1)\,dx_1 = 0 \quad \text{for all } x_1 \in R^1,$$

$$\int_J b_R(x_1, x_2)\,dx_2 = 0 \quad \text{for all } x_2 \in R^1, \text{ and}$$

$$\left\|\frac{\partial^{\alpha+\beta} b_R}{\partial x_1^\alpha\,\partial x_2^\beta}(x_1, x_2)\right\|_\infty \le \frac{1}{|I|^\alpha |J|^\beta} \quad \text{for } \alpha + \beta \le 2.$$

THE ATOMIC DECOMPOSITION FOR $H^1(R_+^2 \times R_+^2)$ [**23**, **24**, **42**]. *Any function $f \in H^1(R_+^2 \times R_+^2)$ can be written as $f = \sum \lambda_k a_k$, where λ_k are scalars with $\sum |\lambda_k| \le C\|f\|_{H^1(R_+^2 \times R_+^2)}$, and where the a_k are H^1-atoms. Conversely, any sum $\sum \lambda_k a_k$, where $\sum |\lambda_k| < \infty$ and a_k are H^1-atoms, defines an element $f \in H^1(R_+^2 \times R_+^2)$ with $\|f\|_{H^1(R_+^2 \times R_+^2)} \le C \sum |\lambda_k|$. In the above we may take as the definition of H^1-atom any of the following:*

(1) *An atom is a function $a(x_1, x_2)$, supported in an open set $\Omega \subseteq R^2$ of finite measure, which can be written as $a = \sum_{R \subset \Omega} C_R b_R$ for scalars C_R and pre-atoms b_R such that $\sum |C_R|^2 |R| \le 1/|\Omega|^{1/2}$.*

(2) *An atom is a function $a(x_1, x_2)$ supported in an open set Ω of finite measure such that $\|a\|_2 \le 1/|\Omega|^{1/2}$ and*

$$|a * \phi_{t_1, t_2}(x_1, x_2)| \le \frac{1}{|\Omega|} \left(\frac{|R_{2t_1, 2t_2}(x_1, x_2) \cap \tilde{\Omega}|}{|R_{t_1, t_2}|} \right)^{10}$$

$\forall \phi \in A, t_1, t_2 > 0$ *and* $(x_1, x_2) \notin \tilde{\Omega}$.

(3) *An atom is a function $a(x_1, x_2)$ supported in an open set Ω of finite measure such that for every $(x_1, x_2) \in R^2$,*

$$S(a)(x_1, x_2) \le M_s^{10}(\chi_{\tilde{\Omega}})(x_1, x_2) \cdot A(x_1, x_2)$$

for some function A with $\|A\|_2 \le 1/|\Omega|^{1/2}$.

Which definition is most useful depends on the problem at hand. We should point out that it is immediately clear from (2) or (3) that an atom belongs to $H^1(R_+^2 \times R_+^2)$. For instance, consider (2) and let

$$a^*(x_1, x_2) = \sup_{t_1, t_2 > 0} |a * \phi_{t_1, t_2}(x_1, x_2)|.$$

Then

$$\int_{R^2} a^* \, dx_1 \, dx_2 = \int_{\tilde{\Omega}} a^* \, dx_1 \, dx_2 + \int_{c_{\tilde{\Omega}}} a^* \, dx_1 \, dx_2$$

$$\le \left(\int_{\tilde{\Omega}} M_s^2(|a|) \, dx_1 \, dx_2 \right)^{1/2} |\Omega|^{1/2} + \frac{1}{|\Omega|} \int_{c_{\tilde{\Omega}}} M_s^{10}(\chi_\Omega) \, dx_1 \, dx_2$$

$$\le C \left[\left(\int_{R^2} |a|^2 \, dx_1 \, dx_2 \right)^{1/2} |\Omega|^{1/2} + \frac{1}{|\Omega|} \int_{R^2} \chi_\Omega^{10} \, dx_1 \, dx_2 \right] \le C'.$$

We should like to close this section by briefly mentioning some applications.

First, the decomposition implies, almost immediately, that $\mathrm{BMO}(R_+^2 \times R_+^2)$ is the dual space of $H^1(R_+^2 \times R_+^2)$. (See [23] and the section of this article on $\mathrm{BMO}(R_+^2 \times R_+^2)$.)

Second, this kind of atomic decomposition allows us to prove that we can interpolate between $H^1(R_+^2 \times R_+^2)$ and $L^2(R^2)$ in order to get $L^p(R^2)$, $1 < p < 2$, as an intermediate space. In other words, if T is a sublinear operator bounded from $H^1(R_+^2 \times R_+^2)$ to $L^1(R^2)$ and on $L^2(R^2)$, then T is bounded on $L^p(R^2)$ for $1 < p < 2$. For more details, see §8 on interpolation and also [24].

Third, the method of proof (especially the use of the sets where $S(f)$ is a certain size in order to produce the atoms of f) of the atomic decomposition is quite useful in the more classical one-parameter setting as well. In their work on H^p-theory on the Heisenberg group, Folland and Stein [45] make use of these methods. In the work of Coifman–Meyer–Stein [29] on the application of the "tent spaces" to commutator integrals, the methods are also used.

Finally, there is one last application which we wish to mention. As has already been pointed out, it is one of the most basic principles of the real-variable theory of H^1 that H^1 can be defined in an equivalent manner by area integrals or by maximal functions. For $H^1(R_+^2 \times R_+^2)$ this is due to Gundy and Stein [**50**]. However, their proof that $S(f) \in L^1(R^2)$ implies $f^* \in L^1(R^2)$ makes use of complex analytic functions. If we think about the atomic decomposition, it is clear that it provides an immediate real-variable proof of this theorem. Indeed, it shows how to take a function whose area integral is in L^1 and cut it up into an absolutely covergent linear combination of pieces (the atoms), each of which obviously has its maximal function of bounded $L^1(R^2)$-norm.

6. Unconditional bases of H^1. We would like to discuss another approach to obtain a "continuous" version of the Haar expansion of functions. It begins with the work of Maurey [**64**] on the existence of an unconditional basis of H^1 of the unit disc. Recall that a basis $\{b_\nu\}_\nu$ is called an unconditional basis for a Banach space B if, whenever an element $v = \sum_\nu a_\nu b_\nu$ is in B, so are the elements $v_\varepsilon = \sum_\nu \varepsilon_\nu a_\nu b_\nu$, where $\varepsilon = (\varepsilon_\nu)$ is a vector with components $\varepsilon_\nu = \pm 1$ for all ν. For example, it is not hard to see that the Haar system $\{h_I\}_{I \text{ dyadic in } R}$, together with constant functions, forms an unconditional basis for $L^p(R)$ for all $p > 1$. Yet when the exponent $p \to 1$, the system becomes an unconditional basis for H_d^1 instead of L^1. (This latter fact may be seen through the square-function characterization of H_d^1, i.e., a function h with Haar expansion $\sum_I C_I h_I$ is in H_d^1 if and only if its dyadic square function

$$S_d(h)(x) = \left(\sum_{x \in I} |C_I|^2/|I| \right)^{1/2}$$

is in L^1; for more details see [**6, 7**].)

The main result in [**16**] states that there exists a linear isomorphism between H^1 of the unit disc and H_d^1 of the unit disk. Since H_d^1 has Haar functions as unconditional basis, by a theorem in Banach space, Maurey concluded the existence of an unconditional basis for H^1 of the unit disc. The first explicit unconditional basis for H^1 was constructed by L. Carleson in [**18**]. Later, P. Wojtaszczyk [**83**] discovered that the orthogonal Franklin system also forms an unconditional basis of $H^1[0, 1]$. (We may define $H^1[0, 1]$ as the subspace of functions in $H^1(R)$ which are supported in $[0, 1]$.) The Franklin system has been investigated in the earlier literature (cf. [**26, 4**]). For example, S. V. Bockarev [**5**] has used the system to construct a basis for the disk algebra. It is the system of orthonormal piecewise linear functions defined on $[0, 1]$ which is obtained through the Schmidt orthonormalization procedure from the functions

$$\phi_0(x) = 1, \quad \phi_1(x) = x, \quad \phi_I(x) = \int_0^x h_I(t)\, dt, \qquad x \in [0, 1],$$

where each

$$I = I_{n,k} = \left[\frac{k-1}{2^n}, \frac{k}{2^n} \right], \qquad 1 \le k \le 2^n, \quad n = 0, 1, 2, \ldots,$$

is a dyadic interval of $[0,1]$. (The order in the procedure is first ϕ_0, ϕ_1, then $\phi_{I_{n,k}}$ for fixed n, $1 \le k \le 2^n$ for $n = 0, 1, 2, \ldots$.)

Only the following specific properties of the system are used in establishing the main result in [83] (cf. [26] for proofs of these properties).

(1) The system $\{f_0, f_1, f_I\}_I$ dyadic is an orthonormal set of functions which is a basis for $L^2[0,1]$.

(2) Each fixed f_I satisfies

    (a) $\int_0^1 f_I(t) \, dt = 0$.

    (b) There exist a constant C and some $q > 1$ independent of I such that $|f_I(t)| \le (C/|I|^{1/2}) q^{d(t,I)/|I|}$, where d is the Lebesgue distance.

$$
(\text{c}) \qquad |f_2(t_1) - f_1(t_2)| \le \frac{C|t_1 - t_2|}{|I|^{3/2}} q^{d([t_1,t_2],I)/|I|}.
$$

In establishing an unconditional basis for H^1, the proof in both [18] and [83] depended on an explicit description of the dual space of H^1 of the disc, namely BMO. As we have mentioned in the last section, such explicit geometric description for the dual space of H^1 of the product domain is still unknown. Instead we only have Carleson measure type characerizations of the space. Yet unlike the situation in Carleson's counterexample on the bidisc ([17], namely a measure which satisfies the product form of the Carleson measure condition may not be in the dual of H^1 of the bidisc), the product form of the Franklin system does form an unconditional basis of $H^1([0,1] \times [0,1])$. We may state this result again in its dual form (cf. [22, 24]):

THEOREM. *If $f \in L^2([0,1] \times [0,1])$ satisfies*

$$
(3) \qquad \int_0^1 f(x_1, x_2) \, dx_1 = \int_0^1 f(x_1, x_1) \, dx_2 = 0,
$$

$$
(4) \qquad \int_0^1 x_1 f(x_1, x_2) \, dx_1 = \int_0^1 x_2 f(x_1, x_2) \, dx_2 = 0,
$$

for all $x_1, x_2 \in [0,1]$ with expansion

$$
f(x_1, x_2) = \sum_{I,J \text{ dyadic}} C_{IJ} f_I(x_1) f_J(x_2),
$$

then f is in the dual of $H^1([0,1] \times [0,1])$ if and only if

$$
\sum_{I \times J \subset \Omega} |C_{I,J}|^2 \le C|\Omega|
$$

for all open sets Ω contained in $[0,1] \times [0,1]$.

The proof of this theorem depends on a careful comparison of $S_R(f_I f_J)$ (same notation as in §4) to C_{IJ} in terms of the relative position between the dyadic rectangle R and $I \times J$ and is mainly technical. But we may draw from it the desired conclusion that the double Franklin system

$$
\{\{f_0, f_1, f_I\} \times \{f_0, f_1, f_J\}\}_{I,J \text{ dyadic intervals}}
$$

forms an unconditional basis for functions in $H^1([0,1] \times [0,1])$. Actually, we can say a little more (cf. [23]):

COROLLARY. $H^1[0,1] = \{f \text{ in } L^1[0,1] \text{ with the series } C_0 f_0 + C_1 f_1 + \sum_I C_I f_I \text{ converges unconditionally, where}$

$$C_i = \int_0^1 f(x) f_i(x)\, dx \quad (i = 0, 1), \qquad C_I = \int_0^1 f(x) f_I(x)\, dx \Big\}.$$

Furthermore, if, for every $\varepsilon = (\varepsilon_0, \varepsilon_1, \varepsilon_I)_I$ $(\varepsilon_i = \pm 1, \ i = 0, 1, \ \varepsilon_I = \pm 1 \ \forall I)$,

$$f_\varepsilon = \sum_{i=0}^1 \varepsilon_i C_i f_i + \sum_{I \text{ dyadic}} \varepsilon_I C_I f_I,$$

then

$$\|f\|_{H^1} \approx \sup_\varepsilon \|f_\varepsilon\|_{L^1}.$$

The same result also is true for $H^1([0,1] \times [0,1])$ with the product Franklin system substituted for the Franklin system.

We finish this section with some remarks:

(1) One can also think of the above results as statements about the dual space of $H^1(T)$ and $H^1(T^2)$ (where T denotes the unit circle, T^2 the torus) after the usual identification of T as $[0,1]$ and T^2 as $[0,1] \times [0,1]$ with functions extending periodically at the endpoints 0 and 1. In doing so we need to modify the Franklin system as defined above. One possible way to do this is to apply the Schmidt orthonormalization procedure to $\{\phi_0 \equiv 1, \ \phi_I\}_{I \text{ dyadic}}$ and observe that for each dyadic I, $\phi_I(0) = \phi_I(1) = 0$; thus, the new system $\{f_0, f_I\}_{I \text{ dyadic}}$ forms an orthonormal basis for functions f in $L^2[0,1]$, $f(0) = f(1)$. One can check (as in [**26**]) to see that $\{f_I\}$ still satisfies properties (a)–(c) earlier. See [**4**] also for other methods to modify the Franklin system from $[0,1]$ to T.

(2) Although it is relatively difficult to compute the coefficient with respect to Carleson's [**8**] basis of $H^1(T)$, the dual form of his basis gives very clear geometric insight about the structure of BMO(T). It remains open whether the product form of Carleson's basis also forms an unconditional basis for $H^1(T^2)$. A positive answer to the problem may help to form a clearer geometric description of BMO(T^2).

(3) A very complicated "ordering" problem arises when one tries to generalize the Franklin system as above to get an unconditional basis for $H^1(R^n)$. For the construction of such systems on R^n, $n \geq 1$, and actually constructions of spline systems of higher orders (which is necessary in order to obtain an unconditional basis for H^p, $0 < p \leq 1$), the reader is referred to the article of J. O. Stromberg [**79**].

7. Some further results. In this section we briefly report on some recent results on product domains which are applications of the techniques we have described. As the reader will see, the results are mainly incomplete, since the whole field is still in its developing stage.

(1) *The corona problem.* Suppose $f_1, f_2, \ldots, f_n$ are bounded analytic functions defined on the open disc D with $|f_1(z)| + |f_2(z)| + \cdots + |f_n(z)| \geq \delta$ for all $z \in D$. The corona problem is this: Do there also exist bounded analytic

functions $g_1, g_2, \ldots, g_n$ in D with $\sum_{j=1}^{n} f_j(z)g_j(z) \equiv 1$ on D? The problem was answered affirmatively by L. Carleson [15] in the early 1960s. Actually, the terminology "Carleson measure" was invented as a device to solve this problem. One of the major difficulties in the solution is the construction of certain Carleson measures to solve some $\bar{\partial}$-equation (cf. [16]). As we have now seen, one way to understand BMO functions is through the connection with associated Carleson measures; namely, if $\phi \in$ BMO(T) then

$$|\nabla \phi(z)|^2 \, \log(1/|z|) \, dz \, d\bar{z}$$

is a Carleson measure with norm comparable to $\|\phi\|_*^2$, where $\phi(z)$ is the harmonic extension of ϕ to D evaluated at z in D. In particular, if ϕ is a bounded analytic function in D (hence its boundary value on the unit circle is in $L^\infty(T)$, hence in BMO(T)), then

$$|\phi'(z)|^2 \, \log(1/|z|) \, dz \, d\bar{z}$$

is a Carleson measure (with norm $\approx \|\phi\|_\infty^2$). It was the brilliant idea of T. Wolff to apply this latter fact to solve the following $\bar{\partial}$-equation and obtain an alternative proof of the corona problem on D.

If μ is any Carleson measure, we say $\mu \in C$ and denote its measure norm by $\|\mu\|_c$.

MAIN LEMMA (T. WOLFF; SEE [46, CHAPTER VIII]). *Suppose g is a smooth function defined on $\overline{D}$ and satisfies*
  (a) $|\partial g(z)/\partial z| \, \log(1/|z|) \, dz \, d\bar{z} \in C$ *with measure norm* $\leq M^2$.
  (b) $|g(z)|^2 \, \log(1/|z|) \, dz \, d\bar{z} \in C$ *with measure norm* $\leq M^2$.
  Then there exists a solution u of the equation $\partial u/\partial z = g$ on D, such that the radial limit of u exists on T with $\|u\|_{L^\infty}(T) \leq M$.

A straightforward generalization of Wolff's lemma to the bidisc $D^2 = D_1 \times D_2$, with $D_1 = D_2 = D$, takes the following form. (We use the notation $\partial_i g = \partial g/\partial z_i$, $\bar{\partial}_i g = \partial g/\partial z_i$ for $i = 1, 2$, with $(z_1, z_2) \in D^2$, and $dA = dz_1 \, d\bar{z}_1 \, dz_2 \, d\bar{z}_2$.)

LEMMA 1 [22]. *Suppose g is a smooth function on $\overline{D}^2$ satisfying*

(a) $$|\partial_1 \partial_2 g(z_1, z_2)| \log \frac{1}{|z_1|} \log \frac{1}{|z_2|} \, dA \in C \quad \text{on } D^2,$$

(b) $$\int\int_{D^2} |\partial_1 g(z_1, z_2)| \log \frac{1}{|z_1|} \log \frac{1}{|z_2|} \, dA \leq C\|h\|_{H^1}$$

for all $h \in H^1(D^2)$,
  (c) *Same condition as in* (b) *with z_1, z_2 interchanged,*

(d) $$|g^2(z_1, z_2)| \log \frac{1}{|z_1|} \log \frac{1}{|z_2|} \, dA \in C \text{ on } D^2.$$

Then there exists a solution v on D^2 with $\bar{\partial}_1 \bar{\partial}_2 v = g$, the radial limit of v exists on T^2, and $\|v\|_{L^\infty(T^2)}$ is bounded by a constant depending only on the Carleson measure constants in (a)–(d).

Combining Wolff's lemma and Lemma 2 we get

LEMMA 2. *Suppose h_1, h_2 are smooth functions on $\overline{D}^2$ satisfying*

$(a)'$ $\qquad |h_1(z_1, z_2)|^2 \log \dfrac{1}{|z_1|} \, dA_1, \qquad |\partial_1 h_1(z_1, z_2)| \log \dfrac{1}{|z_1|} \, dA_1,$

$(b)'$ $\qquad |h_2(z_1, z_2)|^2 \log \dfrac{1}{|z_2|} \, dA_2, \qquad |\partial_2 h_2(z_1, z_2)| \log \dfrac{1}{|z_2|} \, dA_2$

are all Carleson measures in D with measure constants uniformly bounded for all $(z_1, z_2) \in D^2$. Suppose $g = \overline{\partial}_1 h_2 = \overline{\partial}_2 h_1$ satisfies (a)–(d) of Lemma (1). Then there exists a solution u of the $\overline{\partial}$-equation $\overline{\partial}_1 u = h_1$, $\overline{\partial}_2 u = h_2$ such that the radial limit of u exists on T^2 and $u \in \mathrm{BMO}(T^2)$ with $\|u\|_{\mathrm{BMO}(T^2)}$ bounded by Carleson measure constants in $(a)'$, $(b)'$ and (a)–(d) for the function g.

We would like to point out that a BMO (instead of L^∞-) solution is the best one may hope for under the given conditions in Lemma 2. This is similar to the situation that occurs in the unit ball in C^n (cf. [**81**]). The above lemma can be applied to obtain H^p-solutions for the corona problem on the bidisc:

PROPOSITION [**22**]. *Suppose f_1, f_2 are bounded analytic functions defined on the bidisc D^2 satisfying $\sum_{j=1}^{2} |f_j(z)| \geq \delta$ for all $z \in D^2$. Then there exist g_1, g_2 analytic in D^2 and $\bigcap_{p<\infty} H^p(D^2)$ with $\sum_{j=1}^{2} f_j(z) g_j(z) = 1$ in D^2.*

One would think that the number n of functions $f_1, f_2, \ldots, f_n$ is not important for the existence of solutions for the corona problem. That this is not the case on D^d for $d \geq 2$ was pointed out to the first author by N. Varopoulous. This is due to the fact that, when $d \geq 2$, the $(0,1)$-forms on D^d may not be automatically $\overline{\partial}$-closed (when $n = 2$ some special trick may be applied, and this difficulty does not show up). To overcome this difficulty when $d \geq 2$, Varopoulous [**82**] developed and applied some machinery in stochastic integration. Recently K. C. Lin [**60**] was able to handle the extra terms which occurred in the Koszul complex in solving $\overline{\partial}$-equations when $d \geq 2$ by estimates similar (but somewhat easier) to those in Lemma 2. Based on this and some combinatorial arguments, Lin also obtained a real-analysis proof of the following result.

THEOREM ([**60**]; ALSO [**82**]). *Suppose $f_1, f_2, \ldots, f_n$ are bounded analytic functions defined on the bidisc D^d satisfying $\sum_{j=1}^{n} |f_j(z)| \geq \delta$ for all $z \in D^d$. Then there exists $g_1, \ldots, g_n$ analytic in D^d and $\bigcap_{p<\infty} H^p(D^d)$ with $\sum_{j=1}^{n} f_j(z) g_j(z) = 1$.*

The challenging problem of whether or not one can obtain bounded analytic g_j's remains open. (The same problem also remains open for unit balls in C^d.)

(2) *Interpolation between H^p-spaces on the bi-upper-half-plane.* Methods similar to those used in obtaining the atomic decomposition of $H^1(R_+^2 \times R_+^2)$ also yield the following "decomposition" result for L^p-functions $(1 < p < 2)$, which can be viewed as a generalization of the (now) classical Calderón–Zygmund decomposition lemma (cf. [**75**]) to the product domain $(R_+^2 \times R_+^2)$:

LEMMA 3 [25]. *Let $\alpha > 0$ be given and $f \in L^p(R^2)$, $1 < p < 2$. Then we may write $f = g + b$, where $g \in L^2(R^2)$ and $b \in H^1(R^2_+ \times R^2_+)$ with*

$$(*) \qquad \|g\|_2^2 \le \alpha^{2-p}\|f\|_p^p \quad \text{and} \quad \|b\|_{H^1} \le C\alpha^{1-p}\|f\|_p^p,$$

where C is a universal constant.

Just as the Calderón–Zygmund lemma served as a basic tool for various interpolation results between $H^p(R^{n+1}_+)$ [9, 38, 13, 70], the above lemma can be applied for interpolation between $H^p(R^2_+ \times R^2_+)$. An immediate corollary of Lemma 3 is

COROLLARY. *Suppose T is a linear operator bounded from $H^1(R^2_+ \times R^2_+)$ to $L^1(R^2)$ and on $L^2(R^2)$. Then T is bounded on $L^p(R^2)$ for all $1 < p < 2$.*

PROOF OF THE COROLLARY. Let $f \in L^p(R^2)$ and $\alpha > 0$. Applying Lemma 3 we may write $f = g + b$, where $(*)$ holds. Then

$$m\{|Tf| > \alpha\} \le m\left\{|Tg| > \frac{\alpha}{2}\right\} + m\left\{|Tb| > \frac{\alpha}{2}\right\} \le c\left(\frac{1}{\alpha^2}\|Tg\|_2^2 + \frac{1}{\alpha}\|Tb\|_1\right)$$

$$\le c\left(\frac{1}{\alpha^2}\|g\|_2^2 + \frac{1}{\alpha}\|b\|_{H^1}\right) \le c\frac{1}{\alpha^p}\|f\|_p^p.$$

T is therefore weak type (p,p) for all $1 < p < 2$. Hence, by the Marcinkiewicz interpolation theorem, T is bounded on L^p for p in the same range.

A typical T which satisfies the assumptions in the corollary is the double Hilbert transform $Tf = H_{x_1}H_{x_2}f$. Thus, the above result is a generalization of the classical M. Riesz theorem to the setting of product domains.

One can ask questions about interpolating between different H^p-spaces (particularly when $p \le 1$). There are many interpolation methods which apply to H^p. Among them, two suffice for most applications: namely, the real interpolation method $(\ )_{\theta,q}$ $(0 < \theta < 1, 0 < q \le \infty)$ and the Calderón complex interpolation method $(\ )_{\theta}$ $(0 < \theta < 1)$ (cf. [9, 3] for definitions). For H^p on product domains $R^2_+ \times R^2_+$, Lin [61] has the following results:

THEOREM.

(a) $\quad\begin{aligned}&(H^1, L^{p_1})_{\theta,q} = L^{p,q}, \qquad 1 < p_1 < \infty, \ 1/p = 1 - \theta + \theta/p_1,\\ &(H^1, L^\infty)_{\theta,q} = (H^1, \text{BMO})_{\theta,q} = L^{p,q}, \qquad 1/p = 1 - \theta.\end{aligned}$

($L^{p,q}$ *denotes Lorentz space*)

(b) $\quad\begin{aligned}&(H^1, L^{p_1}) = L^p, \qquad 1 < p_1 < \infty, \ 1/p = 1 - \theta + \theta/p_1,\\ &(H^1, L^\infty)_\theta = (H^1, \text{BMO})_\theta = L^p, \qquad 1/p = 1 - \theta.\end{aligned}$

It should be noted that for H^1 on the unit disc or H^1 on the domain R^{n+1}_+, all statements in this theorem are known. (For the real method part (a), see [52, 39, and 70]; for the complex method part (b), see [38, 13 and 58]. For a general survey and references, see [39].) But the methods used in the proof for the classical domains (e.g., unit disk or R^{n+1}_+)—especially the part concerning BMO functions in [38]—apparently do not apply in the product domain $R^2_+ \times R^2_+$. To obtain endpoint results as BMO in the results above,

Lin used the constructive method for interpolation used in [13], and he also applied the elegant four interpolation theorem of Wolff [84].

(3) *Singular integrals of commutator type.* About twenty years ago Calderón [86] proved that a certain singular integral operator was bounded on L^p-spaces. This operator, connected with the study of partial differential equations and complex variables, is known as the "first commutator integral" and is given by

$$Tf(x) = \int_{R^1} \frac{A(x) - A(y)}{(x-y)^2} f(y)\, dy,$$

where $A(x)$ is a differentiable function on R^1 and $A'(x) \in L^\infty$. By the classical Calderón–Zygmund techniques, once it is proven that T is a bounded operator on L^2, it follows immediately that T is also bounded on L^p for all $1 < p < \infty$. Somewhat later, the so called "higher commutators", T_k $(k > 1)$, given by

$$T_k f(x) = \int_{R^1} \left[\frac{A(x) - A(y)}{x-y} \right]^k \frac{1}{(x-y)} f(y)\, dy,$$

were proven to be bounded on L^2 (hence, on L^p for $1 < p < \infty$) by Coifman and Meyer [87]. These operators T_k are different from the classical ones of Calderón–Zygmund discussed in Part I for two reasons. Of course, all these operators are integral operators of the form

$$Tf(x) = \int_{R^1} K(x,y) f(y)\, dy$$

for some kernel $K(x,y)$. The classical Calderón–Zygmund singular integrals have kernels of the form $K(x-y)$, i.e., they are convolution operators, while the commutators are not. The cancellation possessed by the kernels of Calderón–Zygmund is quite obvious:

$$\int_{\alpha < |x-y| < \beta} K(x,y)\, dy = 0 \quad \forall 0 < \alpha < \beta.$$

The cancellation of the kernels for T_k,

$$K_k(x,y) = (A(x) - A(y))/(x-y)^{k+1},$$

is much more subtle. Quite recently, G. David and J. L. Journé have obtained beautiful, simple, and general theorems [88] concering the boundedness on L^2 of operators like T_k. One such theorem [88] is the following:

Suppose $Tf(x) = \int_{R^1} K(x,y) f(y)\, dy$ *and* K *satisfies*

(1) $K(x,y) = -K(y,x),$

(2) $|K(x,y)| \leq \dfrac{C}{|x-y|};$ $\left| \dfrac{\partial K}{\partial x} \right| + \left| \dfrac{\partial K}{\partial y} \right| \leq \dfrac{C}{(x-y)^2},$

and

(3) $T(1) \in \text{BMO}(R^1)$.
Then T *is bounded on* $L^2(R^1)$.

Even though this is an L^2-theorem, just as in the work of Calderón, Coifman, and Meyer, the spaces H^1 (and BMO) play a vital role here. For example, in order to see that, for the kth commutator T_k, we have $T_k(1) \in$ BMO, David and Journé proceed inductively as follows: For $k = 1$,

$$T_1(1)(x) = \int_{R^1} \frac{A(x) - A(y)}{(x-y)^2} \, dy = \int \frac{A'(y)}{x-y} \, dy,$$

as is immediate upon integrating by parts. This last expression is just $H(A')$, the Hilbert transform of a bounded function, which is well known to be in BMO. For $k > 1$ observe that integration by parts shows that

$$\begin{aligned}
T_k(1)(x) &= \int_{R^1} \frac{[A(x) - A(y)]^k}{(x-y)^{k+1}} \, dy \\
&= \int_{R^1} \frac{[A(x) - A(y)]^{k-1} A'(y)}{(x-y)^k} \, dy \\
&= T_{k-1}(A')(x).
\end{aligned}$$

Now the point is that according to a theorem of C. Fefferman–Stein [**38**], if an integral operator $Tf(x) = \int K(x,y)f(y) \, dy$ is bounded on L^2 and satisfies

$$|\partial K(x,y)/\partial x| \leq C/(x-y)^2,$$

it is automatically bounded from L^∞ to BMO. By induction we may assume that $T_{k-1}(1) \in$ BMO, and, hence, by the above theorem T_{k-1} is bounded on L^2. Applying the C. Fefferman–Stein result, we see that $T_{k-1}(A') \in$ BMO, whereupon $T_k(1) \in$ BMO. In order to prove that T_k is bounded on L^2, it was important to understand the action of the operators T_k as operators from L^∞ to BMO.

More recently, Journé has extended some of these results to the setting of product spaces. The model operator here is the "product commutator" T_k, given as follows: Let $A(x_1, x_2)$ be a C^2-function on R^2 such that $(\partial^2 A/\partial x_1 \partial x_2) \in L^\infty$, and let

$$\tilde{\tilde{A}}(x_1, x_2, y_1, y_2) = \frac{A(x_1, x_2) + A(y_1, y_2) - \tilde{A}(y_1, x_2) - A(x_1, y_2)}{(x_1 - y_1)(x_2 - y_2)}.$$

Then

$$T_k f(x_1, x_2) = \int \int_{R^2} [\tilde{\tilde{A}}(x_1, x_2, y_1, y_2)]^k \frac{f(y_1, y_2)}{(x_1 - y_1)(x_2 - y_2)} \, dy_1 \, dy_2.$$

Journé proves in [**89**] a general theorem on product singular integrals which, in particular, implies that the T_k are bounded on $L^p(R^2)$ for $1 < p < \infty$. Rather than go into detail concerning the general class of operators involved, let us make a few observations about the method of proof. A basic property of T_k which must be proved, just as in the one-parameter case, is that $T_k(1) \in$ BMO$(R_+^2 \times R_+^2)$. This is done inductively by proving that T_{k-1} is bounded from L^∞ to BMO$(R_+^2 \times R_+^2)$. Once this is done, the T_k will then be bounded on L^2. The product domain theorems on H^1–BMO duality and interpolation (between H^1 and L^2) are then applied to obtain boundedness on L^p, $1 < p < \infty$.

The arguments of Journé proving that the operators map L^∞ to

$$\mathrm{BMO}(R_+^2 \times R_+^2)$$

are ingenious and along the lines of the proof of Chang [21] that bounded functions have Poisson integrals which give rise to Carleson measures. We should also point out that the interpolation results of Lin, mentioned earlier in this section, are used in the arguments to prove L^2-boundedness of the T_k. The reader interested in applications of the methods described throughout this article is urged to see [89].

REFERENCES

1. A. Baernstein, II, *Univalence and bounded mean oscillation*, Michigan Math. J. **23** (1976), 217–223.

2. A. Benedek, A. P. Calderón and R. Panzone, *Convolution operators on Banach space valued functions*, Proc. Nat. Acad. Sci. U.S.A. **48** (1972).

3. J. Bergh and J. Lofstrom, *Interpolation spaces*, Springer-Verlag, 1976.

4. A. Bernard, *Espaces H^1 de martingales a deux indices, Dualité avec les martingales de type "B.M.O."*, Bull. Sci. Math. (2) **103** (1979), 297–303.

5. S. V. Bockarev, *On a basis in the space of functions continuous in the closed disc and analytic inside it*, Soviet Math. Dokl. **15** (1974), 1195–1216.

6. D. L. Burkholder, *Martingale theory and harmonic analysis in Euclidean spaces*, Proc. Symposia in Pure Math. **35** (2) (1979), 283–302.

7. D. L. Burkholder, R. F. Gundy and M. L. Silverstein, *A maximal function characterization of the class H^p*, Trans. Amer. Math. Soc. **157** (1971), 137–153.

8. D. L. Burkholder and R. F. Gundy, *Extrapolation and interpolation of quasi-linear operators on martingales*, Acta. Math. **124** (1970), 249–304.

9. A. P. Calderón, *Intermediate spaces and interpolation, the complex method*, Studia Math. **24** (1964), 113–190.

10. _____, *On the behaviour of harmonic functions at the boundary*, Trans. Amer. Math. Soc. **68** (1950), 47–54.

11. _____, *On a theorem of Marcinkiewicz and Zygmund*, Trans. Amer. Math. Soc. **68** (1950), 55–61.

12. _____, *An atomic decomposition of distributions in parabolic H^p spaces*, Adv. in Math. **25** (1977), 216–225.

13. A. P. Calderón and A. Torchinsky, *Parabolic maximal functions associated with a distribution*. II, Adv. in Math. **24** (1977), 101–171.

14 A. P. Calderón and A. Zygmund, *On the existence of certain singular integrals*, Acta Math. **88** (1952), 85–139.

15. L. Carleson *Interpolations by bounded analytic functions and the Corona problem*, Ann. of Math. (2) **76** (1962), 547–559.

16. _____, *The Corona problem* (Proc. 15th Scandinavian Congress, 968 (1970). Lecture Notes in Math., Springer, 118.

17. _____, *A counterexample for measures bounded on H^p for the bi-disc*, Mittag-Leffler Report No. 7, 1974.

18. _____, *An explicit unconditional basis in H^1*, Bull. Sci. Math. (2) **104** (1980), 405–416.

19. _____, *BMO in 10 years*, Proc. Scandinavian Congress Lecture Notes, 1981.

20. S. Y. A. Chang, *A characterization of Douglas algebras*, Acta Math. **137** (1976), 81–89.

21. _____, *Carleson measure on the bi-disc*, Ann. of Math. (2) **109** (1979), 613–620.

22. _____, *Two remarks of H^1 and BMO on the bi-disc*, Conf. on Harmonic Analysis in honor of A. Zygmund, vol. II, pp. 373–393.

23. S. Y. A. Chang and Z. Ciesielski, *Spline characterization of H^1*, Studia Math. (to appear)

24. S. Y. A. Chang and R. Fefferman, *A continuous version of the duality of H^1 and BMO on the bi-disc*, Ann. of Math. (2) **112** (1980), 179–201.

25. _____, *The Calderón–Zygmund decompositions on product domains,* Amer. J. Math. **104** (1982), 445–468.

26. Z. Ciesielski, *Properties of the orthonormal Franklin system.* I, II, Studia Math. **73** (1963), 141–157; ibid. **27** (1966), 289–323.

27. R. R. Coifman, *A real variable characterization of H^p,* Studia Math. **51** (1974), 269–274.

28. R. R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals,* Studia Math. **51** (1974), 241–250.

29. R. R. Coifman, Y. Meyer and E. M. Stein *Un nouvel espace functionnel adapte a l'etude des operateurs definis par des integrales singulieres* (preprint).

30. R. R. Coifman and G. Weiss, *Extensions of Hardy spaces and their use in analysis,* Bull. Amer. Math. Soc. **83** (1977), 569–645.

31. A. Cordoba, *Maximal functions, covering lemmas and Fourier multipliers,* Proc. Sympos. Pure Math., vol. 35, Part I, Amer. Math. Soc., Providence, R.I., pp. 29–50.

32. A. Cordoba and R. Fefferman, *A geometric proof of the strong maximal theorem,* Ann. of Math. (2) **102** (1975), 95–100.

33. _____, *On differentiation of integrals,* Proc. Nat. Acad. Sci. U.S.A. **74** (1977), 2211–2213.

34. B. Davis, *Hardy spaces and rearrangements,* Trans. Amer. Math. Soc. **261** (1980), 211–233.

35. C. Fefferman, *Characterization of bounded mean oscillation,* Bull. Amer. Math. Soc. **77** (1971), 587–588.

36. _____, *Estimates for double Hilbert transform,* Studia Math. **44** (1972), 1–15.

37. C. Fefferman and E. M. Stein, *Some maximal inequalities,* Amer. J. Math. **93** (1971), 107–115.

38. _____, *H^p spaces of several variables,* Acta Math. **129** (1972), 137–193.

39. C. Fefferman, N. M. Riviere and Y. Sagher, *Interpolations between H^p-spaces, the real method,* Trans. Amer. Math. Soc. **191** (1974), 75–81.

40. R. Fefferman, *Functions of bounded mean oscillations on the bi-disc,* Ann. of Math. (2) **10** (1979), 395–406.

41. _____, *Strong differentiation with respect to measures,* Amer. J. Math. **103** (1981), 33–40.

42. _____, *The atomic decomposition of H^1 on product spaces,* Adv. in Math. (to appear).

43. _____, *Some weighted norm inequalities for Cordoba's maximal function,* Amer. J. Math. (to appear).

44. R. Fefferman and E. M. Stein, *Singular integrals on product spaces,* Adv. in Math. **45** (1982), 117–143.

45. G. Folland and E. M. Stein, *Hardy spaces on homogeneous groups,* Princeton Univ. Press, Princeton, N.J., 1982.

46. J. Garnett, *Bounded analytic functions,* Academic Press, 1981.

47. J. Garnett and P. Jones, *BMO from dyadic BMO,* Pacific J. Math. **99** (1982), 351–371.

48. R. Gundy, *Maximal function characterizations of H^p for the bi-disc,* Lecture Notes in Math., Springer-Verlag, 1978.

49. _____, *Inégalités pour martingales à un et deux indices l'espace H^p,* L'Ecole d'ete de St. Flour, 1978.

50. R. Gundy and E. M. Stein, *H^p-theory for the polydisk,* Proc. Nat. Acad. Sci. U.S.A. **76** (1979), 1026–1029.

51. R. Hunt, B. Muckenhoupt and R. Wheeden, *Weighted norm inequalities for the conjugate function and Hilbert transform,* Trans. Amer. Math. Soc. **176** (1973), 227–251.

52. S. Igari, *An extension of the interpolation theorem of Marcinkiewicz,* Proc. Japan Acad. **38** (1962), 731–734.

53. B. Jawerth and A. Torchinsky, *The strong maximal function with respect to measures* (preprint).

54. B. Jessen, J. Marcinkiewicz and A. Zygmund, *Note on the differentiability of multiple integrals,* Fund. Math. **24** (1953), 217–234.

55. F. John, *Rotation and strain,* Comm. Pure Appl. Math. **14** (1961), 319–413.

56. F. John and L. Nirenberg, *On functions of bounded mean oscillation,* Comm. Pure Appl. Math. **14** (1961), 415–426.

57. P. Jones, *Carleson measures and the Fefferman–Stein decomposition of $BMO(R)$*, Ann. of Math. (2) **111** (1980), 197–208.

58. ____, *L^∞-estimate for $\bar\partial$-equations*, Acta Math. **150** (1982), 136–152.

59. ____, *Interpolation between H^p-spaces*, Conf. in Harmonic Analysis in Honor of A. Zygmund, Vol. I, pp. 437–451.

60. K. C. Lin, *Thesis*, Univ. of California, Los Angeles, 1984.

61. ____, *H^p interpolation on the bi-disc* (to appear).

62. M. P. Malliavin and P. Malliavin, *Intégrales de Lusin–Calderón pour les fonctions biharmoniques*, Bull. Sci. Math. **101** (1977), 357–384.

63. D. Marshall, *Subalgebras of L^∞ containing H^∞*, Acta Math. **137** (1976), 91–98.

64. B. Maurey, *Isomorphismes entre espaces H^1*, Acta Math. **145** (1980), 79–120.

65. J. Moser, *On Harnack's theorem of elliptic differential equations*, Comm. Pure Appl. Math. **14** (1961), 577–591.

66. B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. **165** (1972), 207–226.

67. Ch. Pommerenke, *Schlichte Funktionen und analytische funktionen von beschrankter mittler oszillation*, Comment. Math. Helv. **52** (1977), 591–602.

68. J. Privalov, *Sur les fonctions conjuguées*, Bull. Soc. Math. France **44** (1916), 100–103.

69. H. M. Reimann, *Functions of bounded mean oscillation and quasi-conformal mappings*, Comment. Math. Helv. **49** (1974), 260–276.

70. N. M. Riviere and Y. Sagher, *Interpolation between L^∞ and H^1, the real method*, J. Funct. Anal. **14** (1973), 401–409.

71. W. Rudin, *Real and complex analysis*, McGraw-Hill, 1966.

72. D. Sarason, *Function theory on the unit circle*, Lecture Notes (Conf. Virginia Polytechnic and State Univ., 1978).

73. ____, *Algebras between L^∞ and H^∞*, Lecture Notes in Math., vol. 512, Springer-Verlag, 1976, pp. 117–129.

74. E. Stein, *On the theory of harmonic functions of several variables. II*, Acta Math. **106** (1961), 137–174.

75. ____, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, N.J., 1970.

76. ____, *A variant of the area integral*, Bull. Sci. Math. (2) **103** (1979), 449–460.

77. E. Stein and G. Weiss, *On the theory of H^p spaces*, Acta Math. **103** (1960), 25–62.

78. J. O. Stromberg, *Weak estimates on maximal function with rectangles in certain directions*, Ark. Math. **15** (1977), 229–240.

79. ____, *A modified Franklin system and higher order spline systems on R^n as unconditional bases for Hardy spaces*, Proc. Conf. on Harmonic Analysis in Honor of A. Zygmund, Vol. II, pp. 475–494.

80. A. Uchiyama, *A constructive proof of the Fefferman–Stein decomposition of* BMO(R^n), Acta Math. **148** (1982), 215–241.

81. N. Varopoulous, *BMO functions and the $\bar\partial$-equation*, Pacific J. Math. **71** (1977), 221–273.

82. ____, *Probabilistic approach to some problems in complex analysis*, Bull. Sci. Math. (2) **105** (1981), 181–224.

83. P. Wojtaszczyk, *The Franklin system is an unconditional basis in H^1*, Ark. Math. **20** (1980), 293–300.

84. T. Wolff, *A note on interpolation space*, (Proc. Conf. Harmonic Analysis, Minneapolis, 1981), Lecture Notes in Math., vol. 908, Springer-Verlag, pp. 199–204.

85. A. Zygmund, *Trigonometric series*, 2nd ed., Cambridge Univ. Press., London and New York, 1968.

86. A. P. Calderon, *Commutators of singular integral operators*, Proc. Nat. Acad. Sci. U.S.A. **53** (1965), 1092–1099.

87. R. Coifman and Y. Meyer, *Commutateurs d'integrales singulieres et operateurs multilineaires*, Ann. Inst. Fourier (Grenoble) **28** (1978), 177–202.

88. G. David and J. L. Journé, *A boundedness criterion for generalized Calderón–Zygmund operators* (preprint).

89. J. L. Journé, *Vector valued singular integral operators and singular integral operators on product spaces* (preprint).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTHERN CALIFORNIA, DRB 306, UNIVERSITY PARK, LOS ANGELES, CALIFORNIA 90089

Current address (S.-Y. A. Chang): Department of Mathematics, University of California, Los Angeles, California 90024