1. Statement of result

Let \( R_1, \ldots, R_d \) be the Riesz transforms of \( \mathbb{R}^d \) (thus \( R_j \) is a Fourier multiplier with symbol \( i\xi_j/|\xi_j| \)). The famous Fefferman-Stein decomposition \([1]\) asserts:

**Theorem 1.1** (Fefferman-Stein theorem). Suppose that \( f \in \text{BMO}(\mathbb{R}^d) \) goes to zero at infinity (thus \( \lim_{x \to \infty} f(x) = 0 \)). Then there exists \( g_0, g_1, \ldots, g_d \in L^\infty(\mathbb{R}^d) \) such that

\[
f = g_0 + R_1 g_1 + \ldots + R_d g_d
\]

and furthermore

\[
\|g_0\|_{L^\infty(\mathbb{R}^d)}, \ldots, \|g_d\|_{L^\infty(\mathbb{R}^d)} \lesssim \|f\|_{\text{BMO}(\mathbb{R}^d)};
\]

here and in the sequel we allow implied constants to depend on \( d \).

The proof in \([1]\) was indirect, using the Hahn-Banach theorem followed by a harmonic majorisation argument. In \([3]\) Uchiyama gave a constructive proof of Theorem 1.1 avoiding both the Hahn-Banach theorem and harmonic majorisation, in particular extending the decomposition to more general Fourier multipliers than \( R_1, \ldots, R_d \). We present Uchiyama’s proof here (but not in full generality).

2. Preliminary reductions

We may take \( f \) (and the \( g_j \)) to be real-valued. Write \( \vec{R} \) for the vector-valued operator \( \vec{R} := (1, R_1, \ldots, R_d) \), and \( \vec{g} \) for \( \vec{g} = (g_0, \ldots, g_d) \). Thus we are asserting for each \( f \in \text{BMO}(\mathbb{R}^d) \) going to zero at infinity, there exists a solution to the vector-valued problem

\[
\vec{R} \cdot \vec{g} = f;
\]

\[
\|\vec{g}\|_{L^\infty(\mathbb{R}^d \to \mathbb{R}^{d+1})} \lesssim \|f\|_{\text{BMO}(\mathbb{R}^d)}.
\]

This is a linear problem, but interestingly it is not possible to express \( \vec{g} \) as a linear operator of \( f \); if we had \( \vec{g} = \vec{T} f \) for some linear operator \( \vec{T} : \text{BMO}(\mathbb{R}^d) \to L^\infty(\mathbb{R}^d \to \mathbb{R}^{d+1}) \), then by averaging over translations one can make \( \vec{T} \) translation invariant; similarly one can make it dilation invariant, and reflection and rotation invariant (rotating and reflecting the Riesz transform vector along with the underlying space). This makes \( \vec{T} \) a combination of the Riesz transforms themselves, but these are not bounded from BMO to \( L^\infty \).

On the other hand, suppose \( f \) was localised in frequency to an annulus \( \{ \xi : |\xi| \sim 2^j \} \), thus one has a reproducing formula \( f = P_j f \) for some Littlewood-Paley type operator \( P_j \). Then we can write

\[
f = \sum_{k=1}^d R_k (R_k P_j f).
\]

It is easy to see that \( R_k P_j \) is bounded from \( \text{BMO}(\mathbb{R}^d) \) to \( L^\infty(\mathbb{R}^d) \), and so we can solve the problem \( (2.1) \) by linear means at each individual frequency scale. The remaining problem is how to concatenate all the frequency scales together (in some nonlinear fashion) without losing control of the solution.

We first perform a standard reduction to an apparently weaker assertion in which a small error in the equation is permitted.
Theorem 2.1 (First reduction). Let $0 < \varepsilon \ll 1$ be small. Then for any $f \in \text{BMO}(\mathbb{R}^d)$ going to zero at infinity with $\|f\|_{\text{BMO}(\mathbb{R}^d)} \leq \varepsilon$ there exists $\vec{g} \in L^\infty(\mathbb{R}^d \to \mathbb{R}^{d+1})$ going to a constant at infinity with $\|\vec{g}\|_{L^\infty(\mathbb{R}^d \to \mathbb{R}^{d+1})} = 1$ such that

$$\|f - \vec{R} \cdot \vec{g}\|_{\text{BMO}(\mathbb{R}^d)} \lesssim \varepsilon^2 \log \frac{1}{\varepsilon}. $$

Indeed, by taking $\varepsilon$ sufficiently small we conclude

$$\|f - \vec{R} \cdot \vec{g}\|_{\text{BMO}(\mathbb{R}^d)} \leq \varepsilon/2$$

and $f - \vec{R} \cdot \vec{g}$ goes to zero at infinity, at which point one can easily iterate the theorem and use the completeness of BMO and $L^\infty$ to obtain the result.

Now for the nonlinear trick: the function $\vec{g}$ is not just bounded in $L^\infty$, but will in fact take values in the unit sphere $S^d := \{v \in \mathbb{R}^{d+1} : |v| = 1\}$, thus $\vec{g} \cdot \vec{g} = 1$. We are thus reduced to

Theorem 2.2 (Second reduction). Let $0 < \varepsilon \ll 1$ be small. Then for any $f \in \text{BMO}(\mathbb{R}^d)$ going to zero at infinity with $\|f\|_{\text{BMO}(\mathbb{R}^d)} \leq \varepsilon$ there exists $\vec{g} \in L^\infty(\mathbb{R}^d \to \mathbb{R}^{d+1})$ with $\vec{g} \to e_1 = (1, 0, \ldots, 0)$ at infinity such that $\vec{g} \cdot \vec{g} = 1$ and

$$\|f - \vec{R} \cdot \vec{g}\|_{\text{BMO}(\mathbb{R}^d)} \lesssim \varepsilon^2 \log \frac{1}{\varepsilon}. $$

The key point will be that in the $\varepsilon$-neighbourhood of any $w \in S^d$, the sphere $S^d$ is approximately equal to a hyperplane $\{v : v \cdot w = 1\}$, up to errors of $O(\varepsilon^2 \log \frac{1}{\varepsilon})$. This will allow us to concatenate the various frequency scales together while only losing errors of $O(\varepsilon^2 \log \frac{1}{\varepsilon})$ in the final error.

3. The linearised problem

In order to solve this problem we will first need to solve the linearised problem

$$\vec{w} \cdot \vec{g} = 0$$
$$\vec{R} \cdot \vec{g} = f$$

for each unit vector $\vec{w} \in S^d$. The key lemma here is that this is solvable (with good estimates) in every direction $\vec{w}$:

Lemma 3.1 (Linear solvability). For each $\vec{w} \in S^d$ there exists a vector-valued Hörmander-Mikhlin multiplier $\vec{T}_\vec{w}$ (i.e. the symbol is a homogeneous symbol of order 0, uniformly in $\vec{w}$) such that

$$\vec{R} \cdot \vec{T}_\vec{w} f = f \quad \text{(3.1)}$$

and

$$\vec{w} \cdot \vec{T}_\vec{w} f = 0 \quad \text{(3.2)}$$

for all $f$. Also $\vec{T}_\vec{w}$ maps real scalar functions to real vector-valued functions.
Proof. Taking Fourier transforms, the task is to obtain a vector-valued symbol \( \vec{m}_{\vec{w}} \) of order 0 such that

\[
(1, \frac{i\xi}{|\xi|}) \cdot \vec{m}_{\vec{w}}(\xi) = 0
\]

for all non-zero \( \xi \), and also

\[
\vec{w} \cdot \vec{m}_{\vec{w}}(\xi) = 0.
\]

To map real scalar functions to real vector-valued functions, we also need \( \vec{m}_{\vec{w}}(-\xi) = \overline{\vec{m}_{\vec{w}}(\xi)} \). If we write \( \vec{w} = (w_0, w) \) and \( \vec{m}_{\vec{w}} = (m_0, m) \) then these equations become

\[
m_0(\xi) = -i \frac{\xi}{|\xi|} \cdot m(\xi)
\]

and

\[
w_0 m_0(\xi) = -w \cdot m(\xi)
\]

Solving for \( m_0 \) in terms of \( m \), we just need to find a symbol \( m \) of order 0 which solves the equation

\[
(w - w_0 i\frac{\xi}{|\xi|}) \cdot m(\xi) = 0
\]

and

\[
m(-\xi) = \overline{m(\xi)}.
\]

The explicit choice

\[
m(\xi) := w + w_0 \frac{i\xi}{|\xi|}
\]

will work here. \( \square \)

4. THE NONLINEAR PROBLEM

Now we return to the nonlinear problem. We use the standard Littlewood-Paley decomposition \( 1 = \sum_j P_j \), where \( P_j \) is frequency-localised to the region \( \{ \xi : 2^{j-1} \leq |\xi| \leq 2^{j+1} \} \), and write \( P_{<j} := \sum_{k<j} P_k \) and \( P_{\geq j} := \sum_{k\geq j} P_k \). Since \( f \) decays at infinity, we see that \( P_{\geq j} f \) converges to \( f \) in BMO norm as \( j \to -\infty \). Thus we have \( \|f - P_{\geq-M}f\|_{\text{BMO}(\mathbb{R}^d)} \lesssim \varepsilon^2 \) for some sufficiently large \( M \). By rescaling we may take \( M = 0 \). Our task is then to solve the problem

\[
\vec{g} \cdot \vec{g} = 1
\]

\[
\|P_{\geq0}f - \vec{R} \cdot \vec{g}\|_{\text{BMO}(\mathbb{R}^d)} \lesssim \varepsilon^2.
\]

It will in fact suffice to achieve the following recursive construction:

**Theorem 4.1** (Third reduction). Let \( 0 < \varepsilon \ll 1 \) be small, and let \( f \in \text{BMO}(\mathbb{R}^d) \) be such that

\[
\|f\|_{\text{BMO}(\mathbb{R}^d)} \lesssim \varepsilon
\]

(4.1)
and such that \( f \) goes to zero at infinity. Then there exists a sequence \( \tilde{g}_{<j}, \tilde{h}_j \) for \( j \geq 0 \) of functions obeying the conditions

\[
\tilde{g}_{<0} = e_1, \quad \tilde{g}_{<j} \to e_1 \text{ at infinity,} \quad \tilde{g}_{<j} \cdot \tilde{g}_{<j} = 1, \quad \tilde{R} \cdot \tilde{h}_j = P_j f,
\]

\[
\| \sum_{0 \leq j \leq J} (\tilde{g}_{<j+1} - \tilde{g}_{<j} - \tilde{h}_j) \|_{\text{BMO}(\mathbb{R}^d)} \lesssim \varepsilon^2 \log \frac{1}{\varepsilon} \text{ for all } J \geq 0
\]

and such that \( \tilde{g}_{<j} \) is locally convergent in \( L^2(\mathbb{R}^d) \).

If we can find such functions, then by setting \( \tilde{g} \) to be the limit of the \( \tilde{g}_{<j} \) we easily conclude the claim (note that \( \tilde{R} \) is bounded on \( \text{BMO}(\mathbb{R}^d) \)). Roughly speaking, the idea is to construct the sphere-valued function \( \tilde{g} \) one Littlewood-Paley component at a time, so that at the \( j \)-th frequency scale one creates a perturbation \( \tilde{h}_j \) to capture the effect of \( P_j f \), and then sets \( \tilde{g}_{<j+1} \) equal to \( \tilde{g}_{<j} + \tilde{h}_j \) plus a small adjustment (required to keep \( \tilde{g}_{<j+1} \) valued on the sphere). (By a strange coincidence, I stumbled much later across a similar method for an unrelated problem tao:wavemaps[2].)

5. Proof of Theorem \( \text{third} 4.1 \)

It remains to prove Theorem \( \text{third} 4.1 \). We set \( \tilde{g}_{<0} \) to be an arbitrary unit vector in \( S^d \), say \((1, 0, \ldots, 0)\), thus establishing (4.2). Now suppose inductively that \( j \geq 0 \), and that \( \tilde{g}_{<j} \) has already been chosen. To build \( \tilde{h}_j \) and \( \tilde{g}_{<j+1} \), what we do first is we perform a wavelet decomposition

\[
P_j f = \sum_{k \in \mathbb{Z}^d} c_{j,k} \psi_{j,k}
\]

where \( \psi_{j,k} \) is a standard wavelet-type object concentrated around the ball \( B(2^{-j}k, 2^{-j}) \) with \( L^\infty \) normalisation (thus the height of this wavelet is \( O(1) \)), and \( c_{j,k} \) is a wavelet coefficient obeying a pointwise bound of the form

\[
|c_{j,k}| \lesssim M P_j f(2^{-j}k),
\]

where \( M \) is the Hardy-Littlewood maximal function; there are many ways to obtain such a decomposition, and the precise form is not important for us here. It is worth noting that the band-limited nature of \( P_j f \) (which in particular implies a reproducing formula of the form \( P_j f = \hat{P}_j f \) for some variant \( \hat{P}_j \) of \( P_j \)) ensures that the maximal function \( M P_j f \) is essentially constant at scales \( 2^{-j} \) (thus \( M P_j f \) is \( O(1) \)) whenever \( |x - y| \lesssim 2^{-j} \). We shall implicitly take advantage of this local constancy at a number of places in the sequel. It is slightly convenient to take \( \psi_{j,k} \) to be band-limited to the annulus \( \{ |\xi| \sim 2^j \} \), and thus allow the wavelets to decay (rapidly) in space.

Let \( \tilde{w}_{j,k} := \tilde{g}_{<j}(2^{-j}k) \in S^d \). We use Lemma \( \text{linear} \) and define

\[
\tilde{h}_j := \sum_{k \in \mathbb{Z}^d} c_{j,k} T_{\tilde{w}_{j,k}} \psi_{j,k},
\]

thus we invert the Riesz transform on each wavelet component of \( P_j f \) while trying to stay orthogonal to \( \tilde{g}_{<j} \) at the centre \( 2^{-j}k \) of the support of that wavelet. The claim
(4.5) is then clear from (3.1). (Note that \( T_{\alpha j,k} \psi_{j,k} \) is a “molecue” which is still bounded by \( O(1) \) and rapidly decreasing away from \( B(2^{-j}k, 2^{-j}) \), which makes it relatively easy to sum the series.)

From (4.1) we have

\[ \| P_j f \|_{L^\infty} \lesssim \varepsilon \]

and hence by (4.1)

\[ |c_{j,k}| \lesssim \varepsilon \]

and hence by (4.2)

\[ \| h_j \|_{L^\infty} \lesssim \varepsilon. \]

In particular, we see that \( \tilde{g}_{<j} + \tilde{h}_j \) takes values in the \( O(\varepsilon) \)-neighbourhood of the sphere \( S^d \). We thus define \( \tilde{g}_{<j+1} \) to be the radial projection of \( \tilde{g}_{<j} + \tilde{h}_j \) to the sphere, thus

\[ \tilde{g}_{<j+1} = \frac{\tilde{g}_{<j} + \tilde{h}_j}{|\tilde{g}_{<j} + \tilde{h}_j|}. \]

This gives (4.4), and with a little more work (noting that \( P_j f \) goes to zero at infinity, hence \( h_j \) does also) one also obtains (4.5).

It remains to show (4.6) and the \( L^2 \) convergence. The first step is to establish some more regularity on the approximate solutions \( \tilde{g}_{<j} \). It is convenient to introduce the local sizes

\[ a_{j,k} := \sum_{m \geq 0} (2/3)^m M_{j-m} P_{j-m} f(2^{-j}k) \]

where \( M_j f \) is the slightly damped maximal function

\[ M_j f(x) := \sup_{r>0} (1 + 2^j r)^{-1} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)| \, dy. \]

The BMO bound (4.1) translates to a Carleson measure bound

\[ \sum_{j: 2^{-j} \leq r} 2^{-dj} \sum_{k: 2^{-1-j} \in B(x,r)} |a_{j,k}|^2 \lesssim \varepsilon^2 r^d \]  (5.3) carl

for any ball \( B(x,r) \). (The exponential decay factor \( (2/3)^m \) ensures that the summation in \( m \) does not cause a difficulty. As pointed out to me by Dmitriy Stolyarov, this bound fails if one does not damp the maximal function.)

**Lemma 5.1** (Lipschitz control on \( \tilde{g}_{<j} \)). For any \( j \geq 0, k \in \mathbb{Z}^d \), and \( x, y \in B(2^{-j}k, O(2^{-j})) \) we have

\[ \tilde{g}_{<j}(x) - \tilde{g}_{<j}(y) = O(a_{j,k} 2^j |x - y|) \]

and

\[ \tilde{h}_j(x) - \tilde{h}_j(y) = O(a_{j,k} 2^j |x - y|) \]

**Proof.** (Sketch) This is achieved by a straightforward induction, noting that \( \tilde{h}_j = O(c_{j,k}) = O(a_{j,k}) \) and \( \nabla \tilde{h}_j = O(2^j a_{j,k}) \) on the ball \( B(2^{-j}k, O(2^{-j})) \), and also observing that radial projection to the sphere does not significantly increase the Lipschitz norm. The fact that \( 2/3 > 1/2 \) will allow one to keep the implied constants in the \( O() \) notation from blowing up. \( \square \)
Let \( \vec{e}_j := \vec{g}_{<j+1} - \vec{g}_{<j} - \vec{h}_j \). On the ball \( B(2^{-j}k, O(2^{-j})) \), \( \vec{g}_{<j} \) differs from \( \vec{w}_{j,k} \) by only \( O(a_{j,k}) \), while \( \vec{h}_j \) has magnitude \( O(a_{j,k}) \). Using (5.2) and (6.2) one then sees that \( \vec{h}_j \) is mostly orthogonal to \( \vec{g}_{<j} \), in that

\[
\vec{h}_j \cdot \vec{g}_{<j} = O(a_{j,k}^2)
\]
on \( B(2^{-j}k, 2^{-j}) \). The key here is the quadratic behaviour in \( a_{j,k} \); the linear behaviour has been cancelled using (6.2). As a consequence one concludes that the error function \( e_j \) is small:

\[
\vec{e}_j = O(a_{j,k}^2).
\]

Also, from the Lipschitz bounds on \( \vec{g}_{<j} \) and \( \vec{h}_j \) one can establish the relatively crude bound

\[
\vec{e}_j(x) - \vec{e}_j(y) = O(2^j|x - y|).
\]

Putting these together with the Carleson bounds (5.3) one can quickly verify

\[
\| \sum_{0 \leq j \leq J} \vec{e}_j \|_{BMO(\mathbb{R}^d)} \lesssim \varepsilon^2 \log \frac{1}{\varepsilon}
\]

for all \( J \geq 0 \), which is (5.4).

Finally, we show local \( L^2 \) convergence. From (5.3) and (5.4), and dominated convergence, one can establish that \( \sum_{j=0}^\infty \vec{e}_j \) is conditionally convergent in \( L^2_{loc} \), while from the frequency support and magnitude bounds on \( \vec{h}_j \), and (5.3) again, one can also obtain conditional local convergence of \( \sum_{j=0}^\infty \vec{h}_j \) in \( L^2_{loc} \). The claim follows.

**References**

1. C. Fefferman, E. M. Stein, \( H^p \) spaces of several variables, Acta Math. 129 (1972), 137–193

Department of Mathematics, UCLA, Los Angeles CA 90095-1555, USA.

E-mail address: tao@math.ucla.edu