ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 25, Number 3, Summer 1995

EIGENVALUE ESTIMATES FOR DEGENERATE PARTIAL DIFFERENTIAL OPERATORS

J. MICHAEL WILSON

1. Introduction. Consider the fourth-order operator $L = \Delta^2 - V$, where Δ is the Laplacian on $\mathbf{R}^d$ and $V \in L^1_{\mathrm{loc}}(\mathbf{R}^d)$ is nonnegative. (For reasons that will be evident, we will assume that $d > 4$.) An integration by parts shows that L will be a nonnegative operator, i.e., have no negative spectrum, if

$$\int_{\mathbf{R}^d} |f|^2 V\, dx \le \int_{\mathbf{R}^d} |\Delta f|^2\, dx \tag{1.1}$$

for all $f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$. The work of Fefferman, Phong and others [**3, 2, 4**] shows that (1.1) is true if V's averages over cubes $Q \subset \mathbf{R}^d$ are suitably small. Specifically, let $p > 1$. Then there is a $\gamma(p,d) > 0$ such that if

$$\ell(Q)^4 \left(\frac{1}{|Q|} \int_Q V^p\, dx \right)^{1/p} \le \gamma(p,d)$$

for all Q ($\ell(Q)$ denotes Q's sidelength) then (1.1) holds. (The L^p norm can be replaced by an Orlicz norm of order $L(\log^{1+\varepsilon} L)$; see [**7**].) Moreover, this condition is close to being necessary, since (1.1) implies trivially that

$$\mathrm{supp}_{\,Q \subset R^d} \frac{\ell(Q)^4}{|Q|} \int_Q V\, dx \le \gamma' < \infty;$$

all one need do is test (1.1) over translates and dilates of a fixed bump function.

What about an operator that has lower-order cross terms? Let $\mathbf{R}^d = \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$, and suppose that $L_\delta = \delta(\Delta^2) + \Delta_1\Delta_2 - V$, where $\delta > 0$ and Δ_i is the Laplacian on $\mathbf{R}^{d_i}$. If $\delta > 1$, then the Δ^2 term

Received by the editors on June 20, 1993.
Research supported under NSF grant DMS 8811775.

dominates, and L_δ can be handled much like L. But if $\delta \to 0$ this is no longer true. Then the cross term $\Delta_1\Delta_2$ becomes important; the Fefferman-Phong machinery does not even tell us whether L_δ remains positive for very small δ's.

In [**8**], we proved that L_δ will stay positive, *even if we set δ equal to zero*, provided that V's Orlicz averages over *rectangles* are controlled. (A rectangle R is a Cartesian product $Q_1 \times Q_2$, where the Q_i are cubes in $\mathbf{R}^{d_i}$.) We proved

Theorem 1. *Let $\varepsilon > 0$. There is a $\gamma(\varepsilon, d_1, d_2) > 0$ such that if*

$$\operatorname{supp}_{R\subset \mathbf{R}^{d_1}\times\mathbf{R}^{d_2}} \frac{\ell(Q_1)^2\ell(Q_2)^2}{|R|} \cdot \int_R V(x_1,x_2)\log^{2+\varepsilon}\left(e+\frac{V(x_1,x_2)}{V_R}\right) dx_1, dx_2 \le \gamma(\varepsilon, d_1, d_2)$$

(V_R denotes V's average over R) then $L_0 \geq 0$; i.e.,

$$\int_{\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}} |f|^2 V\,dx_1\,dx_2 \le \int_{\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}} |\nabla_1\nabla_2 f|^2\,dx_1\,dx_2$$

for all $f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$.

The proof used certain "two-parameter" Littlewood-Paley inequalities, which were themselves derived from one-parameter results contained in [**6**] and [**7**].

The methods of Fefferman and Phong also provide fairly sharp estimates for the bottom of L's spectrum and the number of its negative eigenvalues, in the case where L is *not* nonnegative. In [**8**] we used Theorem 1 to obtain an analogous estimate for the bottom of L_0's spectrum, but we were unable to count eigenvalues at that time.

In the Fefferman-Phong theory, L's negative eigenvalues correspond (more or less) to disjoint cubes for which a) $(\ell(Q)^4/|Q|)\int_Q V\,dx$ or b) $\ell(Q)^4(|Q|^{-1}\int_Q V^p\,dx)^{1/p}$, $p > 1$, is "big enough." Specifically, if there are cubes $Q_1, \ldots, Q_N$ with disjoint doubles for which a) is bigger than some $\gamma(d)$, then L has at least N negative eigenvalues; and if L has N negative eigenvalues, then there are cN disjoint cubes Q_α for which b) is bigger than some positive $\gamma(p, d)$.

Unfortunately, disjointness is not a useful property when one is studying families of rectangles. Put another way, a family of rectangles that is large enough to have interesting "two-parameter" properties is likely to be very far from being pairwise disjoint. In order to do a "Fefferman-Phong" analysis of operators like L_0, it has been necessary to apply their original idea very literally, which was not to consider disjoint cubes, but disjoint regions of *phase space*.

We shall try to make our meaning clear with two examples. We shall work on $\mathbf{R}^1$ first. Let $\psi \in \mathcal{C}_0^\infty(\mathbf{R})$ be real and even, have support contained in $\{|x| \leq 1\}$ and satisfy $\int \psi = 0$. (Of course, we assume that $\psi \not\equiv 0$!) The ψ's Fourier transform, $\hat{\psi}$, is approximately supported in an "annulus," which we will take to be $\{1/2 \leq |x| \leq 3/2\}$. For every dyadic interval $I \subset \mathbf{R}$ (with center x_I and length $\ell(I)$, set $\psi_I(x) = |I|^{-1/2}\psi(2(x - x_I)/\ell(I))$. Each ψ_I has support contained in I and has its Fourier transform approximately supported in $\{\ell(I)^{-1} \leq |\xi| \leq 3 \cdot \ell(I)^{-1}\}$; so its "phase space support" is (roughly) a rectangle of dimensions $\ell(I) \times \ell(I)^{-1}$. According to the Fefferman-Phong theory, the phase space rectangles on which $\xi^4 - V < 0$ (in some averaged sense) correspond to negative eigenvalues of $\Delta^2 - V$, and the associated ψ_I's are the approximate eigenfunctions. (The hard work, which is accomplished by means of Littlewood-Paley theory, comes in making this heuristic remark precise and rigorous.)

Now let ψ_I be as described in the preceding paragraph, and consider the family of functions defined on $\mathbf{R}^2$, $\{\psi_I(x_1) \cdot \psi_J(x_2)\}_{I,J \subset \mathbf{R}}$. Each of these functions has its support contained in a rectangle (*not* a phase space rectangle!) measuring $\ell(I) \times \ell(J)$ and its Fourier transform is approximately supported in a rectangle measuring $\ell(I)^{-1} \times \ell(J)^{-1}$. Nothing unusual so far. But now let us look at those functions such that $\ell(I)/\ell(J) = 2^l$, where l is fixed. These functions are supported on double-dyadic rectangles of fixed eccentricity 2^l. All of them have their Fourier transforms approximately supported in the set $\{(\xi_1, \xi_2) : 2^{l-1} \leq |\xi_2/\xi_1| \leq 3 \cdot 2^{l-1}\}$, which is the union of *four angular sectors*. Different eccentricities, different l's, yield essentially disjoint sectors.

The double-dyadic rectangles of fixed eccentricity 2^l have the same inclusion properties as the dyadic intervals: any two of them are either disjoint, or else one is contained in the other. It turns out that the analysis carried out in [**3**] can be applied to functions of the form

$f_l = \sum_{\ell(I)/\ell(J)=2^l} \lambda_{I,J} \cdot \psi_I(x) \cdot \psi_J(y)$, so that negative eigenvalues of L_0, when restricted to the functions f_l, will correspond to disjoint rectangles R on which $[(1|R|)\int_R V^p]^{1/p} \geq c_p l(I)^{-2} l(J)^{-2}$. This can be done for every $l \in \mathbf{Z}$. If we can decompose a $\mathcal{C}_0^\infty$ f into a sum $\sum_{l\in\mathbf{Z}} f_l$, and if we can show that the pieces corresponding to different l's act more or less independently of each other, then the Fefferman-Phong machinery will yield, almost verbatim, the analogous estimates for L_0. As always, the hard work will come in showing that this actually happens.

As in [**3**], the "hard work" is accomplished via Littlewood-Paley theory; in particular, weighted norm inequalities for (variants of) the Lusin area function. The inequality used in [**3**] had the following general form: Let $\varepsilon > 0$. There is a C_ε such that for every nonnegative weight V and every $f = \sum \lambda_I \cdot \psi_I$ (finite sum),

$$\int |f|^2 V\,dx \leq C_\varepsilon \sum \frac{|\lambda|^2}{|I|} \int_I V(x) \log^{1+\varepsilon}(e + V(x)/V_I)\,dx,$$

where V_I denotes V's average over I [**8**]. (Note: Fefferman used an inequality that was weaker than this.)

In order to treat L_0, we need *two* kinds of inequalities. The first, like the one just described, controls the size of f by the size of some quadratic functional (the square function). The second sort that we need controls the size of the square function by the size of f. We need both kinds because we must control an $f = \sum f_l$ as above by $(\sum |f_l|^2)^{1/2}$, in order to apply the methods of Fefferman-Phong, and doing so requires two steps: first, bounding $\int |f|^2 V\,dx$ by a square function expression which turns out to be a sum of square functions of the f_l's, and then controlling all of these by the f_l's. Hence the need for two kinds of inequalities.

Fortunately, we have these ready to hand in [**2**] and [**8**], and therefore we are able to give a fairly direct proof (modulo a few technicalities) of the analogue of the *hard* part of Fefferman's result on the Schrödinger operator [**3**] for L_0. The difficult part for us is then to find a nearly-equivalent condition on N rectangles R which will imply that L_0 has N negative eigenvalues; recall that this is the easy half of Fefferman's theorem! It is here that our own result is less satisfactory. Our condition does not imply the existence of N eigenvalues for L_0, but instead for the slightly more negative operator $\Delta_1\Delta_2 - M_S V$, where M_S is the "strong" (two-parameter) maximal function.

In Section 2 we review the square function results from [**6, 7, 8, 2**] that we need. In Section 3 we state and prove our theorems.

2. Square function results. We will work on $\mathbf{R}^d = \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$, where we assume that d_1 and d_2 are both bigger than 2. For $i = 1, 2$, let $\psi_i \in \mathcal{C}_0^\infty(\mathbf{R}^{d_i})$ be nontrivial, real, radial, and have support contained in $\{|x| \le 1\}$. We furthermore assume that $\int \psi_i(x)P(x)\,dx = 0$ for all polynomials of degree ≤ 3 (which is easy to arrange) and that ψ_i is normalized to satisfy:

$$\int_0^\infty |\hat{\psi}_i(\xi_i t)|^2 \frac{dt}{t} \equiv 1, \qquad \xi_i \neq 0 \tag{2.1}$$

(here the hat denotes the Fourier transform in the ξ_i variable only). For $y_i > 0$ we let $(\psi_i)_{y_i}(x_i) \equiv y_i^{-d_i}\psi_i(x_i/y)$ denote the usual L^1-dilate of ψ_i. Corresponding to the ordered pairs $t = (t_1, t_2) \in \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$ and $y = (y_1, y_2) \in \mathbf{R}^+ \times \mathbf{R}^+$, we define $\Psi_y(t) = (\psi_1)_{y_1}(t_1) \cdot (\psi_2)_{y_2}(t_2)$.

Let $f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$. Because of our normalization (2.1) and Fourier inversion,

$$f(x) = \int_{\mathbf{R}_+^{d+1}} (f * \Psi_y(t)) \cdot \Psi_y(x - t) \frac{dt\,dy}{y_1 y_2},$$

where the convergence is as nice as we please if $f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$, and the integral converges in $L^2(\mathbf{R}^d)$, at least, if $f \in L^2$. Let Q_i be dyadic cubes in $\mathbf{R}^{d_i}$, and let $R = Q_1 \times Q_2$ be a rectangle. We let $T(R)$ denote the top half of R's "shadow" in $R_+^{d+1} : T(R) = \{(t_1, t_2, y_1, y_2) : t_i \in Q_i, \ell(Q_i)/2 \le y_i \le \ell(Q_i), i = 1, 2\}$, where $\ell(Q_i)$ is Q_i's sidelength. Because of what we said about convergence,

$$\begin{aligned} f &= \sum_R \int_{T(R)} (f * \Psi_y(t)) \cdot \Psi_y(x - t) \frac{dt\,dy}{y_1 y_2} \\ &= \sum_R b_R(x). \end{aligned}$$

Clearly each b_R is supported in $\tilde{R}$, the triple of R. In addition, the function b_R inherits the cancellation properties of the ψ_i's: for each fixed $x_2 \in \mathbf{R}^{d_2}$,

$$\int b_R(t_1, x_2) \cdot P(t_1)\, dt_1 = 0$$

for all polynomials P of degree ≤ 3; and the analogous relation holds for each fixed $x_1 \in \mathbf{R}^{d_1}$. It is easy to see that each b_R is $\mathcal{C}^\infty$, with good bounds on the derivatives. We make this last statement precise by writing $b_R(x) = \lambda_R \cdot a_R(x)$, where λ_R is a constant chosen to make a_R satisfy:

(i) $||\nabla_i a_R||_\infty \leq \ell(Q_i)^{-1}|R|^{-1/2}$ for $i = 1, 2$, where ∇_i is the gradient in the x_i-variables;

(ii) $||\nabla_1 \nabla_2 a_R||_\infty \leq \ell(Q_1)^{-1}\ell(Q_2)^{-1}|R|^{-1/2}$.

(Functions like the a_Rs are sometimes called *adapted functions* or *elementary particles*, see [**1**]. Each a_R is "adapted" to the rectangle $\tilde{R}$.) Conditions i) and ii) are consistent with having

$$|\lambda_R| \leq C\left(\int_{T(R)} |f * \Psi_y(t)|^2 \frac{dt\, dy}{y_1 y_2}\right)^{1/2}, \tag{2.2}$$

where C is a constant that only depends on d_1 and d_2. (It also depends on the ψ_is, but these depend on the dimensions.) Whenever we write $f = \sum_R \lambda_R \cdot a_R$ as in the preceding paragraphs ($f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$), we will assume that (2.2) holds.

The main result of [**8**] is:

Theorem 2.1. *Let $f \in L^2(\mathbf{R}^d)$, and write $f = \sum_R \lambda_R \cdot a_R$ as above. Let $\varepsilon > 0$. There is a constant C, that depends only on ε, d_1 and d_2, such that, for any nonnegative $V \in L^1_{\text{loc}}(\mathbf{R}^d)$,*

$$\int_{\mathbf{R}^d} |f|^2 V\, dx \leq C \sum_R \frac{|\lambda_R|^2}{|R|} \int_R V(x) \log^{2+\varepsilon}\left(e + \frac{V(x)}{V_R}\right) dx.$$

Theorem 2.1 has an immediate consequence via the Plancherel theorem [**8**].

Corollary 2.2. *Let $\varepsilon > 0$. There is a $\gamma(\varepsilon, d_1, d_2) > 0$ such that if*

$$\operatorname{supp}_{R \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \frac{\ell(Q_1)^2 \ell(Q_2)^2}{|R|} \cdot \int_R V(x) \log^{2+\varepsilon}\left(e + \frac{V(x)}{V_R}\right) dx \leq \gamma(\varepsilon, d_1, d_2),$$

then $L_0 \geq 0$, i.e.,

$$\int_{\mathbf{R}^d} |f|^2 V \, dx \leq \int_{\mathbf{R}^d} |\nabla_1 \nabla_2 f|^2 \, dx$$

for all $f \in \mathcal{C}_0^\infty(\mathbf{R}^d)$.

We will need one more result, whose proof is essentially contained in [**2**]. Recall that the *strong maximal function* M_S is defined by

$$M_S V(x_1, x_2) \equiv \operatorname{supp}_{\substack{(x_1,x_2) \in R = Q_1 \times Q_2 \\ Q_1 \times Q_2 \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}}} \frac{1}{|R|} \int_R |V| \, dx.$$

Theorem 2.3. *For $f \in L^1_{\text{loc}}(\mathbf{R}^d)$, let λ_R (and a_R) be defined as above. There is a constant $C(d_1, d_2)$ such that for all such f and all nonnegative $V \in L^1_{\text{loc}}(\mathbf{R}^d)$,*

$$\int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} \left(\sum_{x \in R} \frac{|\lambda_R|^2}{|R|} \right) V \, dx \leq C(d_1, d_2) \int_{\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}} |f|^2 M_S V \, dx. \tag{2.3}$$

Proof of Theorem 2.3. For every integer k, let $E_k = \{x \in \mathbf{R}^{d_1} \times \mathbf{R}^{d_2} : M_S V > 2^k\}$, and let $F_k = \{R \subset \mathbf{R}^{d_1} \times \mathbf{R}^{d_2} : \tilde{R} \subset E_k, \tilde{R} \not\subset E_{k+1}\}$. It is important to notice that each R belongs to one and only one F_k, and that $\cup_{R \in F_k} \tilde{R} \subset E_k$.

Rewrite the lefthand side of (2.3) and get

$$\sum_k \sum_{R \in F_k} |\lambda_R|^2 \frac{1}{|R|} \int_R V \, dx \leq C \sum_k 2^k \sum_{R \in F_k} \int_{T(R)} |f * \Psi_y(t)|^2 \frac{dt \, dy}{y_1 y_2}.$$

Because of the support restriction on the ψ_is, $(t, y) \in T(R)$ implies $f * \Psi_y(t) = (f\chi_{\tilde{R}}) * \Psi_y(t)$. Set $E_k^* = \cup\{\tilde{R} : R \subset E_k\}$. Clearly, there is a fixed n such that $E_k^* \subset E_{k-n}$ for all k. Therefore, we can replace $\chi_{\tilde{R}}$ with $\chi_{E_{k-n}}$. Plancherel's theorem then implies

$$\sum_{R\in F_k} \int_{T(R)} |(f\chi_{E_{k-n}}) * \Psi_y(t)|^2 \frac{dt\,dy}{y_1 y_2} \le C \int_{E_{k-n}} |f|^2\,dx.$$

Plugging this back into the big sum,

$$\begin{aligned}\sum_k \sum_{R\in F_k} |\lambda_R|^2 \frac{1}{|R|} \int_R V\,dx &\le C2^n \sum_k 2^k \int_{E_k} |f|^2\,dx \\ &\le C2^n \int_{\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}} |f|^2 M_S V\,dx. \qquad \square\end{aligned}$$

Before finishing this discussion, we note two easy corollaries of the preceding argument that will be useful in the next section.

First, note that Theorem 2.3 remains true if we replace the compactly-supported ψ_is by functions which are merely in the Schwartz class and have integral 0. The reason for this is that if $\phi \in \mathcal{S}(\mathbf{R}^d)$ and $\int \phi = 0$, then ϕ can be written $\phi = \sum_0^\infty c_k \phi_k$, where each ϕ_k is supported in $\{|x| \le 2^{k+5}\}$, has integral zero, and satisfies $\int |\hat{\phi}_k(\xi t)|^2\,dt/t \equiv 1$, and the $c_k \to 0$ faster than any power of 2^{-k}. This is a well-known construction, which can be found in [**5**]; we describe it in the Appendix to this paper. If we replace the function Ψ by one of the form $\rho_{k,j} = (\phi_1)_k(x_1)\cdot(\phi_2)_j(x_2)$ in the proof of Theorem 2.3, we get the same result as before, but with a constant in front like $C(d_1, d_2)2^{kd_1+jd_2}$. If $\Psi = \psi_1 \cdot \psi_2$, where the ψ_i are Schwartz and have integral 0, then we can write $\Psi = \sum_{0,0}^{\infty,\infty} c_k c_j' \rho_{k,j}$, with the constants rapidly decreasing. Now the Cauchy-Schwarz inequality yields Theorem 2.3 for these "general" Ψs.

Second, notice that the "compact support trick" used in the proof of the theorem can be used to prove it in a slightly different form: Let $\{a_R\}_R$ be an orthonormal family of functions, indexed over a collection of double-dyadic rectangles $\{R\}$, and with $\operatorname{supp} a_R \subset \tilde{R}$ for each R. Let $f = \sum \lambda_R \cdot a_R$ be a finite linear sum from this family. Then, for any weight V,

$$\int_{\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}} \left(\sum_{x\in R} \frac{|\lambda_R|^2}{|R|}\right) V\,dx \le C \int_{\mathbf{R}^{d_1}\times\mathbf{R}^{d_2}} |f|^2 M_S V\,dx.$$

The (easy) proof is left to the interested reader.

3. Eigenvalue estimates. We need a definition: When we say that L_0 has N negative eigenvalues, we mean that there is a subspace $\aleph \subset \mathcal{C}_0^\infty(\mathbf{R}^d)$, of dimension N, such that $\langle L_0\phi, \phi\rangle \le 0$ for all $\phi \in \aleph$. Whenever we write $M_s^k V$, we mean the result of applying M_S (the operator) k times to the function V. If $\omega \ge 0$ is a function, we use ω_R to denote the average of ω on the rectangle R.

In this section we prove

Theorem 3.1. *There exist positive constants* $c_3 = c_3(d_1, d_2)$ *and* $c_4 = c_4(d_1, d_2)$ *such that the following holds. If there are* $c_3 N$ *double-dyadic rectangles* $\{R_i\}$ *such that* $\ell(Q_1^i)^2 \ell(Q_2^i)^2 (1/|R_i|) \int_{R_i} V\,dx \ge c_4$, *then* $\Delta_1\Delta_2 - M_S V$ *has* N *negative eigenvalues.*

Theorem 3.2. *There exist positive constants* $c_1 = c_1(d_1, c_2)$ *and* $c_2 = c_2(d_1, d_2)$ *such that the following holds. Let* $L_0 = \Delta_1\Delta_2 - V$ *have* N *negative eigenvalues. There exists as family of* $c_1 N$ *triples of double-dyadic rectangles* $\{R_i\}$ *with the property that any two* R_i*s with equal eccentricities must be disjoint, and such that*

$$\ell(Q_1^i)^2 \ell(Q_2^i)^2 \frac{1}{|R_i|} \int_{R_i} M_S^4 V(x) \log^{1.5}\left(e + \frac{M_S^4 V(x)}{(M_S^4 V)_R}\right) dx \ge c_2$$

for each $R_i = Q_1^i \times Q_2^i$.

Proof of Theorem 3.1. Let the rectangles be $R_1, \dots, R_N$, and denote the sidelengths of their component cubes by $\ell_1(R_i)$ and $\ell_2(R_i)$. We can partition this collection into four families $\mathcal{F}_k$, $k = 1, \dots, 4$, with the property that if R_i and R_j belong to the same $\mathcal{F}_k$, then $\log_2(\ell_1(R_i)/\ell_1(R_j))$ and $\log_2(\ell_2(R_i)/\ell_2(R_j))$ are both even. One of these $\mathcal{F}_k$ must have cardinality $\ge N/4$. Renumber the rectangles in this family as $R_1, \dots, R_m$ and throw out all the others.

Let ρ be the following function, defined for $x \in [0,1] \subset \mathbf{R}$:

$$\rho(x) = \begin{cases} 0 & x = 0, 1, \\ 1 & x = 1/4, 3/4, \\ -2 & x = 1/2, \\ \text{linear} & \text{in between.} \end{cases}$$

The function ρ is orthogonal (in the ordinary L^2 sense) to all functions of the form $ax + b$. If I is any interval, let p_I be its left endpoint, and set $\rho_I = |I|^{-1/2}\rho((x - p_I)/|I|)$; this is just ρ "fitted" onto I in the usual (L^2-invariant) sense. We've built ρ so that $\rho \perp \rho_I$ if I is any four-adic subinterval of $[0,1]$.

If $Q = I_1 \times \cdots \times I_p$ is a cube in $\mathbf{R}^p$, we set $P_Q(x_1, \dots, x_p) = \prod_1^p \rho_{i_i}(x_i)$. If $R = Q_1 \times Q_2$ is a rectangle in $\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$, we define $\Pi_R(x_1, x_2) = P_{Q_1}(x_1) \cdot P_{Q_2}(x_2)$ (where, of course, the x_i are now points in $\mathbf{R}^{d_i}$).

Because of how we defined ρ and chose our rectangles R_i, the functions $\{\Pi_{R_i}\}_1^m$ are pairwise orthogonal. They each have the same, nonzero L^2 norm. Let us assume that we have normalized them to have norm 1. The only thing that prevents them from being a basis for our eigenspace is their lack of smoothness. They are Lip 1, but not $\mathcal{C}^\infty$. We make them $\mathcal{C}^\infty$ by convolving them with a mollifier. This will mess up their orthogonality, but only a little. The resulting family will be a basis for our "eigenspace."

Let $\eta \in \mathcal{C}_0^\infty(-1,1)$ satisfy $\int \eta = 1$, and let $H(x) = H(x_1, \dots, x_d) = \prod_1^d \eta(x_i)$. For $\varepsilon > 0$, we let $H_\varepsilon(x)$ be the usual L^1-dilate of H (as a function defined on $\mathbf{R}^d = \mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$). Define $\mathcal{P}_{i,\varepsilon}(x) = H_\varepsilon * \Pi_{R_i}(x)$. We wish to show that, for ε sufficiently small, the family $\{\mathcal{P}_{i,\varepsilon}\}_1^m$ is "almost orthonormal." In particular, we claim the following: For all $\delta > 0$ there is an $\varepsilon > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$ and all $f = \sum_{i=1}^m \lambda_i \mathcal{P}_{i,\varepsilon}$,

$$|\lambda_i|^2 \le |\langle f, \mathcal{P}_{i,\varepsilon}\rangle|^2 + \delta \sum_{k=1}^m |\lambda_k|^2.$$

Proof of claim. Fix i. For $0 < \varepsilon \le 1$ and $(\lambda_1, \dots, \lambda_m) \in \mathbf{C}^m$ such that $\sum_1^m |\lambda_k|^2 = 1$, define

$$\Phi(\lambda_1, \dots, \lambda_m, \varepsilon) = |\lambda_i|^2 - |\langle f, \mathcal{P}_{i,\varepsilon}\rangle|^2,$$

where $f = \sum \lambda_k \mathcal{P}_{k,\varepsilon}$, and set $\Phi(\lambda_1, \dots, \lambda_m, 0) \equiv 0$. Then Φ is continuous, hence uniformly continuous, on the compact set $\{(\lambda_1, \dots, \lambda_m) \in \mathbf{C}^m : \sum_1^m |\lambda_k|^2 = 1\} \times [0, 1]$. The claim follows immediately now from compactness and homogeneity.

Now we prove the theorem. Let $f = \sum \lambda_i \mathcal{P}_{i,\varepsilon}$, where $\varepsilon > 0$ will be chosen momentarily. A virtual repetition of the preceding argument shows that, for ε small enough,

$$\int |\nabla_1 \nabla_2 f|^2 \, dx \leq C \sum_1^m |\lambda_i|^2 \ell_1(R_i)^{-2} \ell_2(R_i)^{-2},$$

for a constant C that only depends on d_1 and d_2. In particular, C remains uniformly bounded as $\varepsilon \to 0$.

Using the claim, this last quantity is seen to be less than or equal to a fixed constant times

$$\sum_1^m |\langle f, \mathcal{P}_{i,\varepsilon} \rangle|^2 \ell_1(R_i)^{-2} \ell_2(R_i)^{-2} + \delta \sum_1^m \sum_1^m |\lambda_k|^2 \ell_1(R_i)^{-2} \ell_2(R_i)^{-2},$$

where δ can be made as small as we like. If we make δ small enough, we can throw the last term over onto the lefthand side and conclude that

$$\int |\nabla_1 \nabla_2 f|^2 \leq C \sum_1^m |\langle f, \mathcal{P}_{i,\varepsilon} \rangle|^2 \ell_1(R_i)^{-2} \ell_2(R_i)^{-2},$$

for $0 < \varepsilon < \varepsilon_0$, for all such f. Fix $\varepsilon = \varepsilon_0/2$. Note that, for such ε, the functions $\{\mathcal{P}_{i,\varepsilon}\}_1^m$ are linearly independent (since $f \equiv 0$ implies each $\langle f, \mathcal{P}_{i,\varepsilon} \rangle = 0$ implies $\sum_1^m |\lambda_i|^2 \ell_1(R_i)^{-2} \ell_2(R_i)^{-2} = 0$ implies each $\lambda_i = 0$), so the dimension of span $\{\mathcal{P}_{i,\varepsilon}\}_1^m$ is $\geq N/4$.

Because $V_{R_i} \geq c\ell_1(R_i)^{-2} \ell_2(R_i)^{-2}$ for each i, the last quantity is dominated by

$$\frac{C}{c} \sum_1^m |\langle f, \mathcal{P}_{i,\varepsilon} \rangle|^2 \frac{1}{|R_i|} \int_{R_i} V \, dx;$$

but this, following *precisely* the argument of Theorem 2.3, is less than or equal to

$$C' \frac{C}{c} \int |f|^2 M_S V \, dx \leq \int |f|^2 M_S V \, dx,$$

if c is big enough. Thus, for these f, $\langle(\Delta_1\Delta_2 - M_S V)f, f\rangle \le 0$, and Theorem 3.1 is proved. □

Proof of Theorem 3.2. For $i = 1, 2$, let η_i and ψ_i be nontrivial Schwartz functions which satisfy the following conditions:

a) each is real and radial;

b) $\hat{\eta}_i(\xi_i) \ge 0$, $\hat{\eta}_i(\xi_i) \equiv 1$ in $\{1/2 \le |\xi_i| \le 2\}$ and $\hat{\eta}_i(\xi) \equiv 0$ outside of $\{1/4 \le |\xi_i| \le 4\}$;

c) $\operatorname{supp}\psi_i \subset \{|x| \le 1\}$;

d) $\int \psi_i P(x_i)\,dx_i = 0$ for all polynomials of degree ≤ 1.

e) $\int_0^\infty \hat{\psi}_i(t\xi_i) \cdot \hat{\eta}_i(t\xi) dt/t \equiv 1$ for all $\xi_i \ne 0$.

It is easy to construct such a pair. Begin with a ψ_i that satisfies a), c) and d). Since ψ_i has compact support, its Fourier transform cannot vanish to infinite order at the origin. We may dilate ψ_i (scrunch it up) so that its support remains inside $\{|x| \le 1\}$ while its Fourier transform spreads out. For some dilation, $|\hat{\psi}_i(\xi)|$ will be strictly positive on $\{1/2 \le |\xi_i| \le 2\}$. Now let η_i be taken to be any function satisfying a) and b), such that the integral in e) is nonzero. Replace ψ_i by the dilate and replace η_i with an appropriate scalar multiple of itself (to make e) true).

Let $\Psi_y(t) = (\psi_1)_{y_1}(t_1) \cdot (\psi_2)_{y_2}(t_2)$ be as in the preceding sections and define $E_y(t) = (\eta_1)_{y_1}(t_1) \cdot (\eta_2)_{y_2}(t_2)$ similarly.

By the Fourier inversion theorem,

$$f = \int_{\mathbf{R}_+^{d_1} \times \mathbf{R}_+^{d_2}} (f * E_y(t))\Psi_y(x-t)\frac{dt\,dy}{y_1 y_2},$$

with the same caveats as before (or lack of same) on convergence. The integral converges in L^2, and that's all we need.

With $T(R)$ defined as in Section 2, we can write:

$$f = \sum_R b_R = \sum_R \lambda_R a_R,$$

where

$$b_R(x) = \int_{T(R)} (f * E_y(t))\Psi_y(x-t)\frac{dt\,dy}{y_1 y_2},$$

and the λ_R and a_R are defined analogously, so that each a_R is adapted to $\tilde{R}$.

For $k \in \mathbf{Z}$, let $\mathcal{F}_k = \{R = Q_1 \times Q_2 : \ell(Q_1)/\ell(Q_2) = 2^k\}$. By the results from Section 2,

$$\begin{aligned}
\int |f|^2 V\,dx &\le C \sum_R \frac{|\lambda_R|^2}{|\tilde{R}|} \int_{\tilde{R}} V(x) \log^{2.5}\left(e + \frac{V}{V_{\tilde{R}}}\right) dx \\
&= C \sum_k \sum_{R \in \mathcal{F}_k} \frac{|\lambda_R|^2}{|\tilde{R}|} \int_{\tilde{R}} V(x) \log^{2.5}\left(e + \frac{V}{V_{\tilde{R}}}\right) dx \qquad (3.1) \\
&\le C \sum_k \sum_{R \in \mathcal{F}_k} \frac{|\lambda_R|^2}{|\tilde{R}|} \int_{\tilde{R}} M_S^3 V\,dx,
\end{aligned}$$

where

$$\lambda_R = \left(\int_{T(R)} |f * E_y(t)|^2 \frac{dt\,dy}{y_1 y_2} \right)^{1/2}.$$

It is important to notice that, if $(t, y) \in T(R)$ and $R \in \mathcal{F}_k$, then $2^{-k-1} \le y_1/y_2 \le 2^{-k+1}$, and so the Fourier transform of $f * E_y$, for such y, is supported in the sector $\mathcal{S}_k \equiv \{(\xi_1, \xi_2) : 2^{k-3} \le |\xi_1|/|\xi_2| \le 2^{k+3}\}$. Let f_k denote the inverse Fourier transform of $\hat{f} \cdot \chi_{S_k}$. Then $f * E_y(t) = f_k * E_y(t)$ when $(t, y) \in T(R)$ and $R \in \mathcal{F}_k$. If we plug into (3.1) above, we get

$$\begin{aligned}
\int |f|^2 V\,dx &\le C \sum_k \sum_{R \in \mathcal{F}_k} \left(\int_{T(R)} |f * E_y(t)|^2 \frac{dt\,dy}{y_1 y_2} \right) \\
&\qquad \cdot \frac{1}{|\tilde{R}|} \int_{\tilde{R}} M_S^3 V\,dx \\
&= C \sum_k \sum_{R \in \mathcal{F}_k} \left(\int_{T(R)} |f_k * E_y(t)|^2 \frac{dt\,dy}{y_1 y_2} \right) \\
&\qquad \cdot \frac{1}{|\tilde{R}|} \int_{\tilde{R}} M_S^3 V\,dx \\
&\le C \sum_k \int |f_k|^2 M_S^4 V\,dx,
\end{aligned}$$

where the last line follows from (the noncompactly supported version of) Theorem 2.3.

Notice that the double-dyadic rectangles, the basis of all our constructions, have disappeared. They are about to reappear. For each $k \in \mathbf{Z}$, write

$$f_k(x) = \int_{\mathbf{R}_+^{d_1} \times \mathbf{R}_+^{d_2}} (f_k * E_y(t)) \Psi_y(x-t) \frac{dt\, dy}{y_1 y_2}. \tag{3.2}$$

The important thing to remember here is that, since $\hat{f}_k$ is supported in $\mathcal{S}_k$, and $\hat{E}_y$ is supported in $\{(\xi_1, \xi_2) : 2^{-4} y_2/y_1 \le |\xi_1|/|\xi_2| \le 2^4 y_2/y_1\}$, $f_k * E_y \equiv 0$ unless $2^{-4} y_2/y_1 \le 2^{k+3}$ or $2^4 y_2/y_1 \ge 2^{k-3}$, which means that the convolution will be 0 unless $2^{-k-7} \le y_1/y_2 \le 2^{-k+7}$. Hence, if we decompose the integral in (3.2) (as we have done before) in order to represent f_k as a sum of $\lambda_R a_R$'s, then all of the rectangles R that occur in the sum will have eccentricities bounded between 2^{k+10} and 2^{k-10}.

Let us momentarily fix $k \equiv k_0$ and write

$$f_{k_0} = \sum_{j=k_0-10}^{k_0+10} \sum_{R \in \mathcal{F}_j} \lambda_R a_R = \sum_{j=k_0-10}^{k_0+10} g_j, \tag{3.3}$$

where the $\lambda_R a_R$s are those that occur in the decomposition of f_{k_0} above. (The reader may have noted that f_{k_0} will in general not be a Schwartz function; hence, the sum in (3.3) need not converge pointwise. However, it does converge in L^2, and by the arguments in [**8**], this suffices to let us control $|f_{k_0}|^2$ by the $|\lambda_R|^2$s, which is what we need.)

Now fix j and look at g_j as above. The rectangles $R \in \mathcal{F}_j$ have the same inclusion and disjointness properties as the dyadic cubes in $\mathbf{R}^{d_1} \times \mathbf{R}^{d_2}$. Following the procedure in [**7**, Lemma 2.1], this family can be divided into $3^{d_1+d_2}$ families $\mathcal{G}_l$ such that the *triples* $\tilde{R} \in \mathcal{G}_l \subset \mathcal{F}_j$ have these same properties. Set $g_{j,l} = \sum_{\tilde{R} \in \mathcal{G}_l \subset \mathcal{F}_j} \lambda_R a_R$. For fixed k_0, $f_{k_0} = \sum_{j=k_0-10}^{k_0+10} \sum_l g_{j,l}$, and for general k we may write $f_k = \sum_{j=k-10}^{k+10} \sum_l g_{j,l}^{(k)}$.

We are almost at the point where we may finish the proof with the observation that "we've already solved that problem." For each k, j, $k-10 \le j \le k+10$, l, $1 \le l \le 3^{d_1+d_2}$, let $\tilde{R}_1, \dots, \tilde{R}_{N_{k,j,l}}$ be the minimal rectangles $\tilde{R} \in \mathcal{G}_l$ such that

$$\frac{\ell_1(\tilde{R})^2 \ell_2(\tilde{R})^2}{|\tilde{R}|} \int_{\tilde{R}} M_S^4 V(x) \log^{1.5}\left(e + \frac{M_S^4 V(x)}{(M_S^4 V)_{\tilde{R}}}\right) dx \ge c, \tag{3.4}$$

where the constant c will, as usual, be chosen later, and will be seen to depend on d_1 and d_2, but not on anything else. These rectangles are pairwise disjoint. Now follow the procedure described in [**3**] and [**7**] to find $cN_{k,j,l}$ additional "special" rectangles $\tilde{R} \in \mathcal{G}_l$ that also satisfy (3.4). Call the resulting family (original plus additional rectangles) $\mathcal{T}_{k,j,l}$. We define a subspace $\mathcal{U}$ as follows

$$\mathcal{U} \equiv \bigcap_{k,j,l} \Big\{ f \in \mathcal{C}_0^\infty(\mathbf{R}^d) : \int_{\tilde{R}} g_{l,j}^{(k)} P(x)\, dx = 0$$
$$\text{for all } \tilde{R} \in \mathcal{T}_{k,j,l} \text{ and all polynomials } P \text{ of degree } \leq 1 \Big\}.$$

The codimension of $\mathcal{U}$ is $\leq c \sum_{k,j,l} N_{k,j,l}$. The arguments in [**3, 7, 8**] now apply verbatim (it is here that we apply the proof of Corollary 2.2) to show that, for each k,

$$\int |g_{j,l}^{(k)}|^2 M_S^4 V\, dx \leq c \sum_{\tilde{R} \in \mathcal{G}_l} |\lambda_R|^2 \ell_1(R)^{-2} \ell_2(R)^{-2}$$

whenever $f \in \mathcal{U}$. Hence, if $f \in \mathcal{U}$,

$$\begin{aligned} \int |f_k|^2 M_S^4 V\, dx &\leq c \sum_{k-10 \leq j \leq k+10} \sum_{R \in \mathcal{F}_j} |\lambda_R|^2 \ell_1(R)^{-2} \ell_2(R)^{-2} \\ &\leq c \int |\xi_1|^2 |\xi_2|^2 |\hat{f}_k|^2\, d\xi \end{aligned}$$

for each k. Summing on k, and recalling the definition of f_k, we get

$$\begin{aligned} \int |f|^2 V\, dx &\leq C \sum_k \int |f_k|^2 M_S^4 V\, dx \\ &\leq Cc \sum \int |\xi_1|^2 |\xi_2|^2 |\hat{f}_k|^2\, d\xi \\ &\leq C'c \int |\xi_1^2| |\xi_2|^2 |\hat{f}|^2\, d\xi \\ &\leq \int |\nabla_1 \nabla_2 f|^2\, dx; \end{aligned}$$

for c small enough (depending on d_1 and d_2), i.e., $\langle L_0 f, f\rangle \geq 0$ for $f \in \mathcal{U}$, which implies that L_0 has $\leq c\sum_{k,j,l} N_{k,j,l}$ negative eigenvalues. Theorem 3.2 is proved. □

Appendix

We show how a Schwartz function ϕ that satisfies $\int \phi = 0$ can be decomposed into pieces $\phi = \sum c_k \phi_k$, where the ϕ_k are approximately supported in dyadic annuli, have good bounds on their derivatives, and the constants $c_k \to 0$ very rapidly.

Let ψ be any nonnegative function in $\mathcal{C}_0^\infty(\mathbf{R}^d)$ that is identically equal to 1 in $\{1/2 \leq |x| \leq 2\}$ and vanishes outside $\{1/4 \leq |x| \leq 4\}$. Let $\Psi = \sum_{-\infty}^{\infty} \psi(2^{-k}x)$. Setting $\psi_k(x) = \psi(2^{-k}x)/\Psi(x)$, we note that $\psi_k(x) = \psi_0(2^{-k}x)$. Thus, the properties of ψ_k can be read off from those of ψ_0. Clearly, $\sum_k \psi_k(x) \equiv 1$ for $x \neq 0$, and $\sum_{k>0} \psi_k(x) \equiv 1$ for $|x| > 20$. Set $\eta_k = \psi_k$ for $k > 0$ and $\eta_0 = 1 - \sum_{k>0} \eta_k$. The η_k satisfy $\sum_0^\infty \eta_k \equiv 1$.

Now,

$$\phi = \sum_0^\infty \phi \cdot \eta_k \equiv \sum_0^\infty \rho_k.$$

The functions ρ_k have the support and smoothness conditions we're looking for, but they will usually not have integral zero. We fix this with the help of a telescoping series. For each $k \geq 0$, set

$$\Pi_k(x) = \eta_k(x) \cdot \frac{\int (\sum_{j\leq k} \eta_j)\phi\, dx}{\int \eta_k\, dx}.$$

Then write

$$\sum_k \rho_k = (\rho_0 - \Pi_0) + \sum_{k>0} (\rho_k + \Pi_{k-1} - \Pi_k).$$

Clearly, $\int (\rho_0 - \Pi_0) = \int (\eta_0 \phi) - \int (\eta_0 \phi) = 0$. For the rest,

$$\begin{aligned}\int (\rho_k + \Pi_{k-1} - \Pi_k) &= \int \eta_k \phi + \int \Big(\sum_{j\leq k-1} \eta_j\Big)\phi - \int \Big(\sum_{j\leq k} \eta_j\Big)\phi \\ &= \int \eta_k \phi - \int \eta_k \phi \\ &= 0.\end{aligned}$$

To see that the Π_ks, along with their derivatives, decrease in size rapidly enough (which is all that we have to worry about), we note, first, that $\int \eta_k \geq c2^{kd}$, for some $c > 0$, and second that

$$\int \left(\sum_{j \leq k} \eta_j \right) \phi \, dx = - \int \left(\sum_{j > k} \eta_j \right) \phi \, dx, \tag{A.1}$$

because $\sum \eta_k = 1$ and $\int \phi = 0$. Since ϕ is Schwartz, the righthand side of (A.1) goes to 0 very rapidly.

REFERENCES

1. S.-Y.A. Chang and R. Fefferman, *A continuous version of duality H^1 with* BMO *on the bidisc*, Annals Math. **112** (1980), 179–201.

2. S.-Y.A. Chang, J.M. Wilson and T.H. Wolff, *Some weighted norm inequalities concerning the Schrödinger operator*, Comm. Math. Helv. **60** (1985), 217–246.

3. C.L. Fefferman, *The uncertainty principle*, Bull. Amer. Math. Soc. **9** (1983), 129–206.

4. R. Kerman and E.T. Sawyer, *Weighted norm inequalities for potentials with applications to Schrödinger operators, Fourier transforms and Carleson measures*, preprint 1984.

5. A. Uchiyama, *A constructive proof of the Fefferman-Stein decomposition of* BMO($\mathbf{R}^n$), Acta Math. **148** (1982), 215–241.

6. J.M. Wilson, *A sharp inequality for the square function*, Duke Math. J. **55** (1987), 879–887.

7. ———, *Weighted norm inequalities for the continuous square function*, Trans. Amer. Math. Soc. **314** (1989), 661–692.

8. ———, *Some two-parameter square function inequalities*, Indiana Univ. Math. J. **40** (1991), 419–442.

UNIVERSITY OF VERMONT, BURLINGTON, VERMONT 05405