Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 710–713, January 1982
Mathematics

Symplectic geometry and positivity of pseudo-differential operators

(canonical imbeddings/standard ℓ-straightenings/final straightenings/spectral decomposition theorem/geometric lemma)

C. FEFFERMAN† AND D. H. PHONG‡

†Department of Mathematics, Princeton University, Princeton, New Jersey 08544-0037; and ‡Department of Mathematics, Columbia University, New York, New York 10027

Contributed by Charles Fefferman, September 17, 1981

ABSTRACT In this paper we establish positivity for pseudo-differential operators under a condition that is essentially also necessary. The proof is based on a microlocalization procedure and a geometric lemma.

The general theme of this article and its predecessors (1–4) is to characterize lower bounds and eigenvalues for a pseudo-differential operator $a(x,D)$ in terms of canonical imbeddings of the unit cube into the sets $S_\lambda = \{(x,\xi) \in T^*(\mathbf{R}^n);\ a(x,\xi) < \lambda\}$. This is intimately related to existence and regularity questions for partial differential equations, because virtually any L^2 *a priori* estimate can be reformulated in terms of bounds for the spectrum of a pseudo-differential operator. For positive symbols, the program has been carried out successfully, and simple consequences include, for example, subellipticity for squares of vector fields and their eigenvalue distributions (5); however, operators can be positive and still admit symbols with negative values. Here we shall extend our earlier work to the general case by providing a condition ensuring positivity that is basically also necessary. The precise statements are given below in the theorem and the following remarks. Our proof relies on a microlocalization, much finer than those previously introduced, to study pseudo-differential operators. The microlocalized pieces may be analyzed by the techniques of our previous papers, and the main difficulty is then to patch them together properly. To accomplish this, we introduce a geometric lemma whose proof is the hardest part of our argument.

Further developments related to the main theorem are sketched at the end of the article.

Let $a(x,\xi) \in C^\infty(\mathbf{R}^n \times \mathbf{R}^n)$ be a real symbol of order 2, i.e.,

$$|D_x^\alpha D_\xi^\beta a(x,\xi)| \le C_{\alpha\beta}(1 + |\xi|)^{2-|\beta|} \qquad (x,\xi) \in \mathbf{R}^n \times \mathbf{R}^n$$

and let A be the corresponding pseudo-differential operator

$$(Au)(x) = \int e^{i\langle x,\xi\rangle} a(x,\xi)\hat{u}(\xi)d\xi \qquad u \in C_0^\infty(\mathbf{R}^n).$$

THEOREM. *Fix* $0 < \delta \ll 1$. *Then there exists a constant* $h > 0$ *depending on* δ *such that*

$$Re\langle Au,u\rangle \ge -C_\delta\|u\|^2_{(\delta)} \qquad u \in C_0^\infty(\mathbf{R}^n) \qquad [1]$$

if the set $S = \{(x,\xi) \in T^*(\mathbf{R}^n);\ a(x,\xi) < 0\}$ *does not contain the image of the cube* $I_h = \{(x,\xi) \in T^*(\mathbf{R}^n);\ |x|, |\xi| < h^{1/2}\}$ *by any canonical transformation.*

Remark 1: The theorem to be proven below is in fact significantly stronger, because not all canonical transformations need to be considered, but only some family F satisfying specific bounds. Also, it is easy to see that it can be extended to symbols of any order.

Remark 2: Conversely, there exists a constant $h^* \gg h$ such that inequality 1 implies that the set $S^0 = \{a(x,\xi) + C(1 + |\xi|^2)^\delta < 0\}$ does not contain the image of $\mathbf{I}_{h^*}$ under any canonical transformation in F.

Outline of the proof: The problem will be microlocalized in several stages. First we may view $a(x,\xi)$ as a symbol on a fixed cube Q with good estimates on a slightly larger cube. In fact, let $\varepsilon = \delta/10$, $T^*(\mathbf{R}^n) = \cup_\nu Q_\nu$ be a partition of phase space into rectangles centered at (x^ν,ξ_ν) of dimensions $|\xi_\nu|^{-\varepsilon} \times |\xi_\nu|^{1-\varepsilon}$, and set $M_\nu = |\xi_\nu|^{1-\varepsilon}$, $R_\nu = |\xi_\nu|^\varepsilon$, $p_\nu = M_\nu^{-2\varepsilon}\, a + M_\nu^{2\varepsilon}$; then symbolic calculus shows that it suffices to establish the inequality

$$\mathrm{Re}\langle(\sigma^2 p_\nu)(x,D)u,u\rangle \ge -C\|u\|^2 \qquad [2]$$

for symbols σ of order 0 on Q_ν. Dropping the index ν, we may thus assume after a suitable dilation and translation that p is defined on a cube Q of sides $M^{1/2} \times M^{1/2}$ and satisfies

$$|D_x^\alpha D_\xi^\beta p(x,\xi)| \le C_{\alpha\beta}M^{2-(|\alpha|+|\beta|)/2} \qquad [3]$$

$$|D_x^\alpha D_\xi^\beta p(x,\xi)| \le C_{\alpha\beta}R^{-2}(M^{1/2}R)^{4-(|\alpha|+|\beta|)}. \qquad [4]$$

To describe the next microlocalization procedures and bounds for the relevant canonical transformations, it is convenient to introduce the following terminology. If $\mathbf{I} = \Pi_{k=1}^n (I_{x_k} \times I_{\xi_k})$ is a rectangle in $\mathbf{R}^{2n}$ with

$$|I_{x_1}| \cdot |I_{\xi_1}| \sim \ldots \sim |I_{x_n}| \cdot |I_{\xi_n}| \sim M, \quad \text{then}$$

(*i*) $I_{x_j}^\lambda$ and $I_{\xi_j}^\lambda$ will denote the intervals I_{x_j} and I_{ξ_j} dilated by a factor of λ.

(*ii*) $\mathbf{I}_\lambda \equiv \Pi_{k=1}^n (I_{x_k}^\lambda \times I_{\xi_k}^\lambda)$.

(*iii*) $S^m(\mathbf{I})$ will be the space of all C^∞ functions p on $\mathbf{I}$ such that

$$|D_x^\alpha D_\xi^\beta p(x,\xi)| \le C_{\alpha\beta}M^m \prod_{k=1}^n |I_{x_k}|^{-\alpha_k} \prod_{k=1}^n |I_{\xi_k}|^{-\beta_k}.$$

We shall sometimes also write $S^m(\lambda \times \Lambda)$ for $S^m(\mathbf{I})$ when $|I_{x_1}| \sim \ldots \sim |I_{x_n}| \sim \lambda$ and $|I_{\xi_1}| \sim \ldots \sim |I_{\xi_n}| \sim \Lambda$ and there is no possibility of confusion.

(*iv*) A function χ on $\mathbf{I}$ will be said to satisfy tame estimates if $\nabla_x\chi$ and $\nabla_\xi\chi$ are of class $S^0(\mathbf{I})$.

(*v*) A testing box will be a set of the form $\chi(\mathbf{I}_{h^{1/2}})$, where $\mathbf{I}$ is a rectangle with $|I_{x_1}| \cdot |I_{\xi_1}| = \ldots = |I_{x_n}| \cdot |I_{\xi_n}| = 1$, $\chi{:}\mathbf{I}_{R^\varepsilon} \to \mathbf{R}^n \times \mathbf{R}^n$ is a canonical transformation with tame estimates on $\mathbf{I}_{R^\varepsilon}$, and h is a suitable positive constant.

With these conventions inequalities 3 and 4 simply say that $p \in S^2(Q)$ and $R^2 p \in S^2(Q_R)$, and we shall prove inequality 2 assuming that

$$\max{}_{\mathsf{B}}\, p \ge M^\varepsilon \qquad [5]$$

for any testing box B.

We shall require the following notions:

NOTION OF ℓ-STRAIGHTENING. An ℓ-straightening of p is a canonical transformation $\Phi{:}\mathbf{I}_R \to \mathbf{R}^n \times \mathbf{R}^n$ such that

(1) $\mathbf{I} = \Pi_{k=1}^n (I_{x_k} \times I_{\xi_k})$ with

$$|I_{x_k}| = M_k^{1/2},\ |I_{\xi_k}| = M_\ell M_k^{-1/2} \quad (k \le \ell)$$
$$|I_{x_k}| = |I_{\xi_k}| = M_\ell^{1/2} \quad (k \ge \ell).$$

(2) Φ satisfies tame estimates on $\mathbf{I}_R$.
(3) For $0 \le \mu < \ell$ set $\hat{p}_\mu = p \circ \Phi$ if $\mu = 0$ and

$$\hat{p}_\mu(x_{\mu+1},\ldots,x_n,\xi_{\mu+1},\ldots,\xi_n) = \frac{1}{M_1^{1/2}\ldots M_\mu^{1/2}} \int_{|x_1| \le M_1^{1/2}} \cdots \int_{|x_\mu| \le M_\mu^{1/2}} (p \circ \Phi)(x_1,\ldots,x_n,0,\ldots,0,\xi_{\mu+1},\ldots,\xi_n)\, dx_1 \ldots dx_\mu$$

if $\mu \ge 1$. Then $\hat{p}_\mu$ has the form

$$\hat{p}_\mu = M_{\mu+1}\,\xi_{\mu+1}^2 + \tilde{p}_\mu(x_{\mu+1},\ldots,x_n,\xi_{\mu+2},\ldots,\xi_n) \text{ on } \prod_{k \ge \mu+1} (I_{xk}^R \times I_{\xi k}^R) \times I_{x_{\mu+1}}^R,$$

for $0 \le \mu < \ell$.

(4) $\hat{p}_\ell \in S^2(M_\ell^{1/2} \times M_\ell^{1/2})$ and $R^2\hat{p}_\ell \in S^2(M_\ell^{1/2}R \times M_\ell^{1/2}R)$. Φ is said to be an ℓ-straightening of p at (z^0,ξ_0) if $(z^0,\xi_0) \in \Phi(\mathbf{I}_2)$.

NOTION OF LEVEL AND STANDARD ℓ-STRAIGHTENING. Fix a point (z^0,ξ_0). By definition p always has level ≥ 0, and the standard 0-straightening of p at (z^0,ξ_0) is $\Phi^0 =$ Identity on $\mathbf{I}_R$, where $\mathbf{I}$ is the dyadic cube of side $M^{1/2}$ that includes (z^0,ξ_0). For $\ell \ge 0$, assume now that p has level $\ge \ell$ at (z_0,ξ^0) and that the standard ℓ-straightening Φ^ℓ has been defined. Given a rectangle $\mathbf{I} \subseteq \mathbf{I}_2$ with $\mathbf{I} = \Pi_k\,(\hat{I}_{xk} \times \hat{I}_{\xi k})$ and

$$|\hat{I}_{xk}| = \hat{M}_k^{1/2},\ |\hat{I}_{\xi k}| = \hat{M}_{\ell+1}\hat{M}_k^{-1/2} \qquad k \le \ell+1$$

$$|\hat{I}_{xk}| = |\hat{I}_{\xi k}| = \hat{M}_{\ell+1}^{1/2} \qquad k \ge \ell+1$$

$$\hat{p}_\ell \in S^2\left(\prod_{k \ge \ell+1} \hat{I}_{xk} \times \hat{I}_{\xi k}\right)$$

we cut $\hat{\mathbf{I}}$ into 2^{2n} congruent smaller rectangles by cutting $\hat{I}_{xk}$ and $\hat{I}_{\xi k}$ into two equal subintervals when $k \ge \ell+1$, $\hat{I}_{\xi k}$ into four equal subintervals, and retaining $\hat{I}_{xk}$ when $k \le \ell$. Starting with $\hat{\mathbf{I}}$ or one of the adjacent congruent rectangles, we repeat the process, stopping at the rectangles $\hat{\mathbf{I}}$ that satisfy one of the following conditions:

(C1) $\min\{\Sigma_{k \le \ell} M_k\xi_k^2;\ \xi_k \in I_{\xi k}(k \le \ell)\} \ge$ (constant) $M_{\ell+1}^2$.
(C2) $\hat{p}_\ell \ge M_{\ell+1}^2$ (constant) on $\Pi_{k \ge \ell+1}\,\hat{I}_{xk} \times \hat{I}_{\xi k}$.
(C3) $M_{\ell+1}^2 \le$ (constant).
(C4) $\min_{\Pi_{k \ge \ell+1}(\hat{I}_{xk} \times \hat{I}_{\xi k})} \hat{p}_\ell \le -$ (constant) $M_{\ell+1}^2$.
(C5) $$\max_{|\alpha|+|\beta|=2} \|D_x^\alpha D_\xi^\beta \hat{p}_\ell\|_{L^\infty(\Pi_{k \ge \ell}\hat{I}_{xk} \times \hat{I}_{\xi k})} \prod_{k \ge \ell+1} |I_{xk}|^{\alpha_k} \prod_{k \ge \ell+1} |I_{\xi k}|^{\beta_k} \ge (\text{constant})\, M_{\ell+1}^2.$$

We obtain in this manner a decomposition $\{\mathbf{I}^\mu\}$ for the union of $\mathbf{I}$ and its adjacent congruent boxes, with $\hat{p}_\ell \in S^2\,(\Pi_{k \ge \ell+1}\,\hat{I}_{xk} \times \hat{I}_{\xi k})$ and $R^2\hat{p}_\ell \in S^2\,(\Pi_{k \ge \ell+1}\,\hat{I}_{xk}^R \times \hat{I}_{\xi k}^R)$ for each of the $\hat{\mathbf{I}}^\mu$.

A symbol p will be said to have level $\ge(\ell+1)$ at (z^0,ξ_0) if $(\Phi^\ell)^{-1}(z^0,\xi_0) \in \hat{\mathbf{I}}^{\mu_0}$ with $\hat{\mathbf{I}}^{\mu_0}$ arising from C5, and to have level exactly ℓ at (z^0,ξ_0) otherwise.

If $\hat{\mathbf{I}}^{\mu_0} = \Pi_k\,\hat{I}_{xk} \times \hat{I}_{\xi k}$ arises from C5, let

$$\psi^\mu\colon \prod_{k \ge \ell+1} \hat{I}_{xk}^R \times \hat{I}_{\xi k}^R \to \prod_{k \ge \ell+1} \hat{I}_{xk}^{2R} \times \hat{I}_{\xi k}^{2R}$$

be a canonical transformation such that

$$\hat{p}_\ell \circ \psi^\mu(x_{\ell+1},\ldots,x_n,\xi_{\ell+1},\ldots,\xi_n) = \hat{M}_{\ell+1}\xi_{\ell+1}^2 + \tilde{p}_{\ell+1}(x_{\ell+1},\ldots,x_n,\xi_{\ell+1},\ldots,\xi_n)$$

with $\tilde{p}_{\ell+1} \in S^2\,(\Pi_{k \ge \ell+1}\,I_{xk} \times I_{\xi k})$ and $R^2\tilde{p}_{\ell+1} \in S^2(\Pi_{k \ge \ell+1}\,I_{xk}^R \times I_{\xi k}^R)$, and define the standard $(\ell+1)$-straightening of p at (z^0,ξ_0) to be

$$\Phi^{\ell+1} = \Phi^\ell \circ (Id \times \psi^\mu)\colon (\hat{\mathbf{I}}^{\mu_0})_R \to \mathbf{R}^n \times \mathbf{R}^n.$$

This completes the description of the notions of level and standard ℓ-straightenings.

NOTION OF FINAL STRAIGHTENING. Assume first that p has level exactly ℓ at (z^0,ξ_0) with $0 \le \ell < n$. If Φ^ℓ is the standard ℓ-straightening of p at (z^0,ξ_0), we know that $(\Phi^\ell)^{-1}(z^0,\xi_0)$ must belong to some $\hat{\mathbf{I}}^{\mu_0}$ arising from C1–4. Define the final straightening Φ of p at (z^0,ξ_0) to be $\Phi = \Phi^\ell|_{(\mathbf{I}^{\mu_0})_R}$.

Next assume that p has level n, and set

$$p_n = \frac{1}{\left|\prod_k I_{xk}\right|} \int_{\Pi_k I_{xk}} (p \circ \Phi^n)(x,0)dx,$$

where $\Phi^n\colon \mathbf{I}_R \to \mathbf{R}^n \times \mathbf{R}^n$ is the standard n-straightening. Decompose $\mathbf{I}$ and its adjacent congruent boxes, by cutting at each step all the $I_{\xi k}$ into two equal subintervals, retaining the I_{xk}s, and stopping when either

$$|I_{xk}|^2\,|I_{\xi k}|^2 \le \max\{100B^*,\hat{p}_n\}$$

or

$$\min_{I_{\xi k}} \left(\sum_k M_k\,\xi_k^2\right) \ge (\text{constant})\ |I_{xk}|^2\,|I_{\xi k}|^2.$$

Here B^* is a constant we shall pick later. Call $\{\hat{\mathbf{I}}^\nu\}$ the resulting rectangles. We can now complete the definition of final straightenings by letting the final straightening of p at (z^0,ξ_0) be $\Phi = \Phi^n|_{(\hat{\mathbf{I}}^{\mu_0})_R}$, where $\mathbf{I}^{\mu_0}$ is the rectangle containing $(\Phi^n)^{-1}(z^0,\xi^0)$.

Note that Φ is a straightening of p—i.e., $\hat{p}_\mu$ is of the form $\hat{p}_\mu = M_{\mu+1}\,\xi_{\mu+1}^2 + \tilde{p}_\mu(x_{\mu+1},\ldots,x_n,\xi_{\mu+2},\ldots,\xi_n)$ and $\mu <$ critical μ^* and $\hat{p}_{\mu^*} \ge$ (constant) $|\hat{I}_{xk}|^2\,|I_{\xi k}|^2$ on $\hat{\mathbf{I}}^{\mu_0}$ unless $|\hat{I}_{xk}|\,|\hat{I}_{\xi k}| \le 10$.

The desired estimate is a simple consequence of the case $\ell = 0$ of the following lemma.

LEMMA 1. *Assume that the symbol* p *has level* $\ge\ell$ *at* (z^0,ξ_0), $|z^0|$, $|\xi_0| \le 2M^{1/2}$, *and let* $\Phi^\ell\colon \mathbf{I}_R^\ell \to \mathbf{R}^n \times \mathbf{R}^n$ *be the standard* ℓ*-straightening of* p *at* (z^0,ξ_0). *If* $\sigma \in S^0\,(\mathbf{I}^\ell)$ *is a symbol supported in* $\mathbf{I}_2^\ell$, *we have the estimate*

$$Re\langle[\sigma^2(p \circ \Phi^\ell)](x,D)u,u\rangle \ge B_*\|\sigma(x,D)u\|^2 - \tfrac{3}{4}(n - \ell + 1)C_*\|u\|^2,$$

where B_* *and* C_* *are suitable constants with* B_* *taken to be large, depending on* C_*.

Lemma 1 can be reduced to

LEMMA 2. *Given* (z^0,ξ_0), $|z^0|$, $|\xi_0| \le 2M^{1/2}$, *let* $\Phi\colon\mathbf{I}_R \to \mathbf{R}^n \times \mathbf{R}^n$ *be the final straightening of* p *at* (z^0,ξ_0), *and let* $\sigma \in S^0(\mathbf{I})$ *be supported on* $\mathbf{I}_2$. *Then we have*

$$Re\langle[\sigma^2(p \circ \Phi)](x,D)u,u\rangle \ge B_*\|\sigma(x,D)u\|^2 - \tfrac{2}{3}C_*\|u\|^2.$$

Reduction of Lemma 1 *to* Lemma 2: Assuming *Lemma 2*, we began by establishing the case $\ell = n$ in *Lemma 1*. Recall that the final straightening is obtained from the standard n-straightening $\Phi^n\colon \mathbf{I}_R^n \to \mathbf{R}^n \times \mathbf{R}^n$ by a decomposition of $\mathbf{I}_2^n$ into rectangles $\{\hat{\mathbf{I}}^\mu\}$. Corresponding to this decomposition is a symbolic calculus. Next *Lemma 2* implies

$$\mathrm{Re}\langle\sigma^2(p \circ \Phi^n)\,(x,D)u,u\rangle \ge B_*\|\sigma(x,D)u\|^2 - \tfrac{3}{4}C_*\|u\|^2$$

as desired.

We shall now show that the case (ℓ) of *Lemma 1* follows from the case $(\ell + 1)$. In fact, to $\mathbf{I}_2^\ell$ is associated a decomposition

$\{\hat{I}^\mu\}$, where each of the rectangles $\{\hat{I}^\mu\}$ arises for one of the reasons C1–5. If $\hat{I}^\mu$ was chosen for one of the reasons C1–4, then $\Phi^\ell|_{\hat{I}^\mu_R}$ is the final straightening of p at a suitable point, and again we have by *Lemma 2*

$$\mathrm{Re}\langle\sigma_\mu^2(p\circ\Phi^\ell)(x,D)u,u\rangle \geq B_*\|\sigma_\mu(x,D)u\|^2 - \tfrac{2}{3}C_*\|u\|^2.$$

Similar estimates also hold for $\hat{I}^\mu$ arising from C5. This follows from the fact that $\Phi^{\ell+1} = \Phi^\ell \circ (Id \times \psi^\mu)$, because by case $(\ell + 1)$ of *Lemma 1*

$$\mathrm{Re}\langle\hat{\sigma}_\mu^2(p\circ\Phi^{\ell+1})(x,D)u,u\rangle \geq B_*\|\hat{\sigma}_\mu(x,D)u\|^2 - \tfrac{3}{4}C_*(n+1-(\ell+1))\|u\|^2$$

for $\hat{\sigma}_\mu = \sigma_\mu(id \times \psi^\mu)$, and thus the previous assertion is a simple consequence of the sharp Egorov principle. We can now use symbolic calculus as before to sum the estimates for different $\hat{I}^\mu$s and thus obtain case (ℓ) of *Lemma 1*.

Lemma 2 in turn can be easily derived from

LEMMA 2_m. *Let* $\Phi: I_R \to \mathbf{R}^n \times \mathbf{R}^n$ *be the final straightening of* $2^{-m}p$ *at* (z^0,ξ_0) $(|z^0|, |\xi_0| \leq 2M^{1/2})$, *and assume that* $I = \Pi_k(I_{x_k} \times I_{\xi_k})$ *with*

$$|I_{x_k}|\cdot|I_{\xi_k}| \geq B_{(min)}. \quad [7]$$

Then for $\sigma \in S^0(I)$ *supported in* I_2, *we have*

$$Re\langle\sigma^2(p\circ\Phi)(x,D)u,u\rangle \geq B_*\|\sigma(x,D)u\|^2 - \tfrac{2}{3}C_*\|u\|^2.$$

Note that the case $m = 0$ yields *Lemma 2* at once because inequality 5 implies $\max_I(p\circ\Phi) \geq M^e$, whereas $p\circ\Phi \in S^2(I)$, so that $|I_{x_k}|\cdot|I_{\xi_k}| \geq M^{e/2} \geq B_{(min)}$. Also *Lemma* 2_m holds vacuously for $2^m \geq M^2$; indeed $2^{-m}p$ is then bounded by a constant in $|z|,|\xi| \leq M^{1/2}$ and hence all the blocks $\{\hat{I}^\mu\}$ of the initial decompositions arise from C3 and are cubes of side ~ 1; in particular, the hypothesis 7 never holds.

The main part of the proof is thus to show that *Lemma* 2_{m+1} implies *Lemma* 2_m. The arguments involved will require two subsidiary theorems, one analytic, the other geometric. The analytic theorem is the spectral decomposition theorem of ref. 4, and the geometric one is the key *Geometric Lemma* to be established in this work.

Proof of Lemma 2_m *assuming Lemma* 2_{m+1}: Let then $\Phi: I \to \mathbf{R}^n \times \mathbf{R}^n$ be the final straightening for $2^{-m}p$ at (z^0,ξ_0), $I = \Pi_k(I_{x_k} \times I_{\xi_k})$, and recall that

$$2^{-m}(p\circ\Phi) \in S^2(I)$$

$$\frac{1}{|I_{x_1}\times\ldots\times I_{x_\mu}|}\int_{I_{x_1}}\cdots\int_{I_{x_\mu}}(2^{-m}p\circ\Phi)$$
$$(x_1,\ldots,x_\mu,x_{\mu+1},\ldots,x_n,0,\ldots,0,\xi_{\mu+1},\ldots,\xi_n)\,dx_1\ldots dx_\mu$$
$$= |I_{x_{\mu+1}}|^2\,\xi_{\mu+1}^2 + \tilde{p}_\mu(x_{\mu+1},\ldots,x_n,\xi_{\mu+2},\ldots,\xi_n) \text{ for } \mu = 0,1,\ldots,\mu_{(cr)}.$$

$$\tilde{p}_{\mu_{(cr)}} \geq (\text{constant})\,|I_{x_k}|^2\,|I_{\xi_k}|^2.$$

Observe also that C4 never arises due to our hypothesis 5 on testing boxes, whereas C3 is ruled out by the hypothesis 7 of *Lemma* 2_m.

Two cases will be considered separately:

First case: $|I_{x_k}|\cdot|I_{\xi_k}| \leq B_{(min)}^{100}$.

We shall show that for $\sigma \in S^0(I)$ supported in I_2

$$\mathrm{Re}\langle\sigma^2[|I_{x_\mu}|^2\,\xi_\mu^2 + \tilde{p}_\mu(x_\mu,\ldots,x_n,\xi_{\mu+1},\ldots,\xi_n)](x,D)u,u\rangle > \lambda_\mu\|\sigma(x,D)u\|^2 - (\text{constant})\,\|u\|^2 \quad [8]$$

for $1 \leq \mu \leq \mu_{(cr)}$, with $\lambda_1 \geq B_*$. This will yield *Lemma* 2_m except when $\mu_{(cr)} = 0$, in which case $p\circ\Phi$ is already elliptic on I_2, and *Lemma* 2_m is obvious.

For $\mu = \mu_{(cr)}$, inequality 8 is trivial with $\lambda_{\mu_{(cr)}} = |I_{x_k}|^2\,|I_{\xi_k}|^2$ because $\tilde{p}_{\mu_{(cr)}} \geq (\text{constant})\,|I_{x_k}|^2\,|I_{\xi_k}|^2$. Next, observe that the assumption of the first case and the hypothesis 5 on testing boxes imply that $p\circ\Phi \geq -1$ on I_2, and consequently $\tilde{p}_\mu \geq -1$ on I_2. To pass from the estimate for $(\mu+1)$ to the estimate for μ, we are thus in position to apply the spectral decomposition theorem. This theorem says that the μth estimate will hold provided that the operators $p_{\mu+1,\ell}$ defined by

$$\tilde{p}_{\mu+1,\ell} = \frac{1}{|I_{\mu\ell}|}\int_{I_{\mu,\ell}}\tilde{p}_\mu\,(x_\mu,x_{\mu+1},\ldots,x_n,\xi_{\mu+1},\ldots,\xi_n)\,dx_\mu$$

where $I_{\mu\ell}$ are subintervals of I_{x_μ} with $|I_{\mu\ell}|/|I_{x_\mu}| = (\lambda_\mu)^{-1/2}$, satisfy the estimates

$$\mathrm{Re}\langle\sigma^2\tilde{p}_{\mu+1,\ell}(x,D)u,u\rangle \geq \lambda_\mu\|\sigma(x,D)u\|^2 - (\text{constant})\,\|u\|^2.$$

(Here we assume that $B_{(min)} \gg 1$ and $\lambda_\mu \geq B_{(min)}^{(\text{small power})}$.) However, inequalities 3 and 4 and tame estimates on $\Phi|_{I_R}$ together show that $p\circ\Phi$, and hence all the $\tilde{p}_\mu$ are essentially polynomials of bounded degrees $\leq d$. As a consequence we may write

$$1 + \tilde{p}_{\mu+1,\ell} \equiv \text{average of } (1+\tilde{p}_\mu) \text{ over } I_{\mu\ell}$$
$$\geq \left(\left|\frac{I_{\mu\ell}}{I_{x_\mu}}\right|\right)^d \text{ average of } 1+\tilde{p}_\mu \text{ over } I_{x_\mu}$$
$$\geq \lambda_\mu^{-d/2}\,[|I_{x_{\mu+1}}|^2\,\xi_{\mu+1}^2 + \tilde{p}_{\mu+1}].$$

So the desired estimate for $\tilde{p}_{\mu+1,\ell}\,(x,D)$ follows from

$$\mathrm{Re}\langle\sigma^2\lambda_\mu^{-d/2}\,[|I_{x_{\mu+1}}|^2\,\xi_{\mu+1}^2 + \tilde{p}_{\mu+1}]\,(x,D)u,u\rangle \geq \lambda_\mu\|\sigma(x,D)u\|^2 - (\text{constant})\,\|u\|^2,$$

which in turn follows from

$$\mathrm{Re}\langle\sigma^2[|I_{x_{\mu+1}}|^2\xi_{\mu+1}^2 + \tilde{p}_{\mu+1}]\,(x,D)u,u\rangle \geq (\lambda_\mu)^{(d+2)/2}\,\|\sigma(x,D)u\|^2 - (\text{constant})\,\|u\|^2.$$

Thus inequality 8 holds as long as $\lambda_\mu = (\lambda_{\mu+1})^{2/(d+2)}$. Observe that each λ_μ is in particular a power of $|I_{x_k}|\cdot|I_{\xi_k}|$ as needed, and that $\lambda_0 = (|I_{x_k}|\,|I_{\xi_k}|)^\gamma$ with $\gamma = 2[2/(d+2)]^{\mu_{cr}}$. This proves *Lemma* 2_m in the first case, provided $B_{(min)} \geq B_*^{\bar{\gamma}}$ with $\bar{\gamma} = [(d+2)/2]^{n+1}$.

Second case: $|I_{x_k}|\cdot|I_{\xi_k}| \geq B_{(min)}^{100}$.

This step in the proof relies heavily on the following:

GEOMETRIC LEMMA. *Let* $\Phi_{[m]}: I_R^{[m]} \to \mathbf{R}^n \times \mathbf{R}^n$ *be the final straightening of* $2^{-m}p$ *at* (z^0,ξ_0). *Assume* $|I_R^{[m]}| \geq B_{(min)}^{100n}$. *Take* $(x^*,\xi^*) \in I_{3/2}^{[m]}$, *and let* $\Phi_{[m+1]}: I_R^{[m+1]} \to \mathbf{R}^n \times \mathbf{R}^n$ *be the final straightening of* $2^{-(m+1)}p$ *at* $\Phi_{[m]}(x^*,\xi^*)$. *Finally, denote by* ι_m:(unit cube) $\to I^{[m]}$, ι_{m+1}:(unit cube) $\to I^{[m+1]}$ *the natural affine changes of scales. Then*

(1) $|I^{[m+1]}| \geq B_{(min)}^n$.

(2) $\iota_{m+1}^{-1}\,\Phi_{[m+1]}^{-1}\,\Phi_{[m]}\iota_m$ *and its inverse map are well defined in a ball of radius* δ *about* $\iota_m^{-1}(x^*,\xi^*)$ *and its image, and they are smooth there with* a priori *bounds on their seminorms.*

Assuming the *Geometric Lemma*, we can complete the proof of *Lemma* 2_m.

Set $1 = \Sigma_j\,\eta_j^2$ in $\mathrm{supp}(\sigma)$ with $\eta\circ\iota_m$ supported in a ball of radius δ about a point in the cube about 0 of side 3. Corresponding to each η_j we have a straightening $\Phi_{[m+1]}$ as above. Now

$$2^{-(m+1)}\eta_j^2(p\circ\Phi_{[m]}) = (2^{-(m+1)}(\eta_j^2\, p\circ\Phi_{[m+1]}))\circ(\Phi_{[m+1]}^{-1}\circ\Phi_{[m]}).$$

Because we are in the second case, the hypothesis $|\mathbf{I}^{[m]}| \geq B_{(\min)}^{100n}$ holds, and the *Geometric Lemma* applies. Thus $\Phi_{[m+1]}^{-1}\circ\Phi_{[m]}$ is a canonical transformation with natural bounds, defined in the double of $\mathrm{supp}(\eta_j^2\circ\Phi_{[m]})$, and $|\mathbf{I}^{[m+1]}| \geq B_{(\min)}^{n}$, so that *Lemma* 2_{m+1} holds. In view of the sharp Egorov principle, we may write

$$\mathrm{Re}\langle 2^{-(m+1)}\eta_j^2\, p\circ\Phi_{[m]}(x,D)u,u\rangle \geq B_*\|\eta_j(x,D)u\|^2 - \tfrac{2}{3}C_*\|u\|^2 - C\|u\|^2,$$

where C is a harmless constant. From symbolic calculus we may then deduce that

$$\mathrm{Re}\langle 2^{-(m+1)}\sigma^2\, p\circ\Phi_{[m]}(x,D)u,u\rangle \geq B_*\|\sigma(x,D)u\|^2 - (\tfrac{2}{3}C_* + C)\delta^{-2n}\|\sigma(x,D)u\|^2 - C\delta^{-2n}\|u\|^2,$$

and hence

$$\mathrm{Re}\langle 2^{-m}\sigma^2\, p\circ\Phi_{[m]}(x,D)u,u\rangle \geq \{2B_* - \tfrac{4}{3}\delta^{-2n}(C_* + C)\}\|\sigma(x,D)u\|^2 - C\delta^{-2n}\|u\|^2.$$

Choosing $B_* \geq \tfrac{4}{3}(C_* + C)\,\delta^{-2n}$ and $C_* \geq C\delta^{-2n}$ yields the desired estimate.

The proof of the *Geometric Lemma* is too lengthy to be described here and will be presented (in detail) in a more complete version of this article.

Applications. Subelliptic estimates for partial differential equations take the form

$$\mathrm{Re}\langle Pu,u\rangle \leq C\mathrm{Re}\langle Qu,u\rangle + C\|u\|_{(\delta)}^2 \qquad [9]$$

or

$$\|Pu\| \leq C\|Qu\| + C\|u\|_{(\delta)} \qquad [10]$$

for pseudodifferential operators $P = p(x,D)$, $Q = q(x,D)$.

Our results suggest the

CONJECTURE. (A) *For* p,q *nonnegative and second-order, inequality* **9** *holds if and only if* $MAX_{\mathsf{B}}\,p \leq C\ MAX_{\mathsf{B}}\{q + (1 + |\xi|^2)^{\delta}\}$ *for each box* $\mathsf{B} \in \mathsf{F}(q)$. (B) *For* P,Q *self-adjoint second-order and* $Re\ q \geq 0$, *inequality* **10** *holds if and only if* $MAX_{\mathsf{B}}\,|Re\,p| \leq C\ MAX_{\mathsf{B}}\{Re\ q + (1 + |\xi|)^{\delta}\}$ *for each box* $\mathsf{B} \in \mathsf{F}(Re\ q)$.

Here $\mathsf{F}(q)$ denotes the family of testing boxes arising from the final straightening of $2^{-m}q$ $(m \geq 0)$. From this conjecture one could immediately read off sharp estimates such as $\|Xu\|_{(\varepsilon-1)}^2 \leq C\Sigma_j\|X_ju\|^2 + C\|u\|^2$ and $\|X_jX_ku\| \leq C\|\Sigma_jX_j^2u\| + C\|u\|$ for non-commuting vector fields X_j,X. See refs. 6–8.

Now *Part A* of the conjecture follows easily from our positivity theorem applied to *q-cp*. Note that *Part B* amounts to the positivity of a fourth-order operator.

It would be very interesting to understand the estimate **10** for non-self-adjoint operators P and Q. In particular, Egorov's subellipticity theorem (9–12) is of the form **10** for non-self-adjoint first-order Q. We hope to return to these matters in a later work.

The authors were partially supported by National Science Foundation Grants MCS80-03072 and MCS78-27119.

1. Fefferman, C. & Phong, D. H. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4673–4674.
2. Fefferman, C. & Phong, D. H. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6055–6056.
3. Fefferman, C. & Phong, D. H. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5622–5625.
4. Fefferman, C. & Phong, D. H. (1981) *Commun. Pure Appl. Math.* **34**, 285–331.
5. Fefferman, C. & Phong, D. H., *Proceedings of the Chicago Conference in Honor of A. Zygmund*, in press.
6. Hörmander, L. (1967) *Acta Math.* **119**, 147–171.
7. Kohn, J. J. (1973) *Proc. Symp. Pure Math. AMS* **23**, 61–69.
8. Rothschild, L. P. & Stein, E. M. (1976) *Acta Math.* **137**, 247–320.
9. Egorov, Y. V. (1975) *Usp. Mat. Nauk.* **30**, 57–114.
10. Egorov, Y. V. (1975) *Uspekhi Mat. Nauk.* **30**, 57–104.
11. Hörmander, L. (1979) *Ann. Math. Stud.* **91**, 127–208.
12. Treves, F. (1971) *Commun. Pure Appl. Math.* **24**, 71–115.