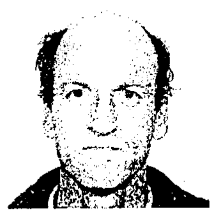
A Very Elementary Proof of the Malgrange-Ehrenpreis Theorem

Jean-Pierre Rosay, *University of Wisconsin, Madison*



Jean-Pierre Rosay: Ph.D. in 1970 at Grenoble (France), under the supervision of Alain Bernard, in functions algebras.
Postdoctoral positions: University of Kentucky-Lexington, Université Paris-Sud Orsay.
Professor: at Marseilles 1973–86, at the University of Wisconsin-Madison since 1986.
Places visited for a semester: UCLA, University of North Carolina at Chapel-Hill.
Field: Several complex variables.

The aim of this note is to present an easy, self-contained, and totally elementary proof of the fundamental theorem of Malgrange and Ehrenpreis ([1], [6]), on the existence of fundamental solutions for linear partial differential operators with constant coefficients.

By totally elementary the following is meant:

- we will not use any Fourier transform argument
- only few integrations by parts will be used, together with the Hilbert space structure of L^2, and the Riesz representation theorem. Convolutions are used, twice, in a very simple way.
- Basically no knowledge of distribution theory is needed. We first show that for every $g \in L^2_{\text{loc}}(\mathbf{R}^n)$ one can solve the equation $P(D)u = g$, with $u \in L^2_{\text{loc}}(\mathbf{R}^n)$, if $P(D)$ is a (non-zero) linear differential operator on $\mathbf{R}^n$ with constant coefficients.

This is only to say ("by integration by parts") that for all $\varphi \in \mathscr{C}_0^\infty(\mathbf{R}^n)$: $\langle u, P^*(D)\varphi\rangle = \langle g, \varphi\rangle$, where,

$$P(D) = \sum a_J \frac{\partial^{|J|}}{\partial x^J},$$

$$P^*(D) = \sum (-1)^{|J|}\overline{a_J}\frac{\partial^{|J|}}{\partial x^J},$$

$$J = (j_1, \ldots, j_n), \qquad |J| = j_1 + \cdots + j_n,$$

$$\frac{\partial^{|J|}}{\partial x^J} = \frac{\partial^{j_1 + \cdots + j_n}}{\partial x_1^{j_1} \cdots \partial x_n^{j_n}}.$$

Also $\langle\ ,\ \rangle$ denotes the usual scalar product in $L^2(\mathbf{R}^n)$, $\langle u, v\rangle = \int_{\mathbf{R}^n} u\bar{v}$.

- Only the last step, which uses a trivial and well known trick (whose origin, I do not know) requires one to know the definition of a distribution.

I wish to make clear that there is but little novelty in this proof. The basic and almost unique ingredient is Hörmander's inequality ([3] 2.6) (whose proof is so elementary and such a masterpiece). There is an allusion in [8] to an "immediate" proof of the existence of fundamental solutions, using Hörmander's inequality. Here we give the full detail in a way which seems somewhat different. Sections 1 and 2 are taken from Malgrange [7]. The only possibly new point is that we prove

the approximation property in §3, using the same tools, whereas it has usually been done using Fourier transform arguments and using polynomial-exponential solutions ([4] 3.4, [5]). It may be argued that all this was implicit in the literature. Let us then make it explicit!

We do not assume any knowledge of the theory of PDE. But, for instance, it should be noticed that the results on supports in §2 follow immediately from Hölmgren's uniqueness theorem (cf, e.g. [4] 5.3). In fact, they could also be obtained immediately from the uniqueness of solutions of O.D.E., by partial Fourier transform, since only the case of functions with compact support is under consideration. A relatively simple proof of the Malgrange Ehrenpreis theorem, using Fourier transforms, can be found in [2].

Notations. If Ω is an open set in $\mathbf{R}^n$, $\mathscr{C}_0^\infty(\Omega)$ will denote the space of $\mathscr{C}^\infty$ functions with compact support in Ω. The norm and the scalar product in $L^2(\Omega)$ will be respectively denoted by $\|\ \|$ and $\langle\ ,\ \rangle$, or $\|\ \|_\Omega$ and $\langle\ ,\ \rangle_\Omega$ if some ambiguity is possible.

The adjoint of an operator $P(D)$ is defined as above and denoted by $P^*(D)$.

1. L^2 solvability on bounded open sets in $\mathbf{R}^n$.

1.1. Hörmander's inequality. Let $P(D)$ be a (non-zero) linear differential operator with constant coefficients, of order (or degree) m.

Theorem. *For every bounded open set Ω in $\mathbf{R}^n$, there exists a constant $C > 0$ such that for every $\varphi \in \mathscr{C}_0^\infty(\Omega)$:*

$$\|P(D)\varphi\| \geqslant C\|\varphi\|.$$

As constant C, one can take $C = |P|_m K_{m,\Omega}$, where

$$|P|_m = \max\{|a_J|, |J| = m\}\left(P(D) = \sum a_J \frac{\partial^{|J|}}{\partial x_J}\right),$$

and $K_{m,\Omega}$ depends only on m (the order of $P(D)$) and the diameter of Ω.

Remark. The proof will use a trick that it is nice to look at first in the case $n = 1$, $\Omega = (0, 1)$ and $P(D) = d/dx$! We have to show that $\|\varphi'\| \geqslant C\|\varphi\|$ for all $\varphi \in \mathscr{C}_0^\infty((0,1))$, for some positive constant C.

One has $\langle(x\varphi)', \varphi\rangle = \langle x\varphi', \varphi\rangle + \langle\varphi, \varphi\rangle$. Further, $\langle(x\varphi)', \varphi\rangle = -\langle x\varphi, \varphi'\rangle$, by integration by parts. Hence: $\langle\varphi, \varphi\rangle = -\langle x\varphi', \varphi\rangle - \langle x\varphi, \varphi'\rangle$. Since $|x| < 1$, one gets $\|\varphi\|^2 \leqslant 2\|\varphi'\|\|\varphi\|$, by the Schwarz inequality. So $\|\varphi'\| \geqslant \frac{1}{2}\|\varphi\|$.

The general proof goes by induction on the degree of P.

Proof. If $P(D)$ is a differential operator of order $m \geqslant 0$ on $\mathbf{R}^n$, then for $j = 1, \ldots, n$ define $P_j(D)$ by

$$P(D)(x_j\varphi) = x_j P(D)\varphi + P_j(D)\varphi.$$

The operator $P_j(D)$ is zero if and only if $P(D)$ does not involve any differentiation with respect to x_j, and if non-zero, $P_j(D)$ is of order $< m$. Let $A = \sup_\Omega |x|$. By induction on m, we will show that for every $\varphi \in \mathscr{C}_0^\infty(\Omega)$

$$\|P_j(D)\varphi\| \leqslant 2mA\|P(D)\varphi\|. \qquad (**)$$

Let us first notice that the inequality $(**)$ and the definition of P_j yield: $\|P(D)(x_j\varphi)\| \leqslant (2m+1)A\|P(D)\varphi\|$. Since differential operators with constant coefficients commute, we have, for all $\varphi \in \mathscr{C}_0^\infty(\Omega)$

$$\|P(D)\varphi\|^2 = \langle P(D)\varphi, P(D)\varphi\rangle = \langle \varphi, P^*(D)P(D)\varphi\rangle = \langle \varphi, P(D)P^*(D)\varphi\rangle$$
$$= \langle P^*(D)\varphi, P^*(D)\varphi = \|P^*(D)\varphi\|^2.$$

This is of course just folklore from the theory of normal operators!

The inequality $(**)$ is trivial for $m = 0$, since then $P_j(D) = 0$. Let us assume the inequality $(**)$ has been verified for operators of order at most $m-1$, and let $P(D)$ be an operator of order m. Compute in two different ways $\langle P(D)(x_j\varphi), P_j(D)\varphi\rangle$. From the definition of $P_j(D)$

$$\langle P(D)(x_j\varphi), P_j(D)\varphi\rangle = \langle x_jP(D)\varphi, P_j(D)\varphi\rangle + \|P_j(D)\varphi\|^2.$$

By integration by parts (i.e., the definition of the adjoint) and using the commutation of $P^*(D)$ and $P_j(D)$

$$\langle P(D)(x_j\varphi), P_j(D)\varphi\rangle = \langle P_j^*(D)(x_j\varphi), P^*(D)\varphi\rangle.$$

Therefore,

$$\|P_j(D)\varphi\|^2 = \langle P_j^*(D)(x_j\varphi), P^*(D)\varphi\rangle - \langle x_jP(D)\varphi, P_j(D)\varphi\rangle.$$

As noticed above, the induction hypothesis yields

$$\|P_j^*(D)(x_j\varphi)\| \leqslant (2m-1)A\|P_j(D)\varphi\|.$$

Since

$$|\langle x_jP(D)\varphi, P_j(D)\varphi\rangle| \leqslant A\|P(D)\varphi\|\|P_j(D)\varphi\|,$$

one gets

$$\|P_j(D)\varphi\|^2 \leqslant 2mA\|P_j(D)\varphi\|\|P(D)\varphi\|$$

and this establishes $(**)$.

If $P(D)$ is an operator of order $m \geqslant 1$, there exists $j \in \{1, \ldots, n\}$ so that $P_j(D)$ is of order $m-1$, and $|P_j|_{m-1} \geqslant |P|_m$. The theorem follows then immediately by induction on m.

1.2. L^2 solvability

Corollary. *If Ω is a bounded set in $\mathbf{R}^n$, then for every $g \in L^2(\Omega)$ there exists $u \in L^2(\Omega)$ so that $P(D)u = g$.*

Proof. This follows immediately from the inequality on the adjoint $\|P^*(D)\varphi\| \geqslant C\|\varphi\|, \varphi \in \mathscr{C}_0^\infty(\Omega)$. Indeed $P(D)u = g$ means that for all $\varphi \in \mathscr{C}_0^\infty(\Omega)$

$$\langle g, \varphi\rangle = \langle u, P^*(D)\varphi\rangle. \qquad (*)$$

Let $E = \{\Psi \in \mathscr{C}_0^\infty(\Omega), \Psi = P^*(D)\varphi$ for some $\varphi \in \mathscr{C}_0^\infty(\Omega)\}$. Hörmander's inequality shows that the antilinear form $\Psi \mapsto \langle g, \varphi\rangle$ is well defined and continuous in the L^2 norm. It can therefore be extended to $\bar{E}$, the closure of E in $L^2(\Omega)$. Then **the Riesz representation theorem gives the existence of $u \in \bar{E}$ so that $(*)$ holds.**

2. Hörmander's inequality with weight and support.

Theorem. *$P(D)$ and Ω (bounded) are given as previously. There exists $C' > 0$ so that for all $\eta \in R$ and $\varphi \in \mathscr{C}_0^\infty(\Omega)$*

$$\int_\Omega e^{\eta x_1}|P(D)\varphi|^2 \geqslant C'\int_\Omega e^{\eta x_1}|\varphi|^2.$$

Note that C' does not depend on η.

Proof. Apply Hörmander's inequality to $\Psi = e^{(\eta/2)x_1}\varphi$ and the operator $Q(D)$ defined by

$$Q(D)(\Psi) = e^{(\eta/2)x_1}Q(D)[e^{-(\eta/2)x_1}\Psi],$$

which is indeed a constant coefficient operator with the same terms of higher degree as $P(D)$.

Corollary 1. *If $\varphi \in \mathscr{C}_0^\infty(\mathbf{R}^n)$ and $P(D)\varphi = 0$ in the half space $\{x_1 > 0\}$, then $\varphi = 0$ in this half space.*

Proof. Take Ω an open set containing the support of φ, and let η tend to $+\infty$ in the inequality.

Corollary 2. *Let $\varphi \in \mathscr{C}_0^\infty(\mathbf{R}^n)$, or more generally $\varphi \in L^2(\mathbf{R}^n)$ and assume that φ has compact support. If $P(D)\varphi$ is supported by the ball of radius r around 0, then so is φ.*

In the application $P(D)\varphi$ is given by a function in $L^2(\mathbf{R}^n)$.

Proof. By using translations and rotations, the smooth case follows immediately from Corollary 1, by writing the ball as an intersection of half spaces. To treat the non-smooth case, consider $\chi \in \mathscr{C}_0^\infty(\mathbf{R}^n)$, χ supported by the ball of radius 1 around the origin, and satisfying $\int\chi = 1$. For $\varepsilon > 0$ set

$$\varphi_\varepsilon = \varphi * \chi_\varepsilon = \frac{1}{\varepsilon^n}\int_{\mathbf{R}^n}\varphi(x-t)\chi\left(\frac{t}{\varepsilon}\right)dt.$$

Then $\varphi_\varepsilon \in \mathscr{C}_0^\infty(\mathbf{R}^n)$, $P(D)\varphi_\varepsilon (= [P(D)\varphi]_*\chi_\varepsilon)$ is supported by the ball of radius $r+\varepsilon$, and $\varphi_\varepsilon \to \varphi$ in L^2 as $\varepsilon \to 0$. This reduces the problem to the smooth case.

Now that we are at the end of section 2, let us remind the reader that sections 1 and 2 are taken from Malgrange [7].

3. Approximation Theorem. For $\rho > 0$, let B_ρ be the ball of radius ρ around 0 in $\mathbf{R}^n$.

Theorem. *Let $0 < r < r' < R$. If $v \in L^2(B_{r'})$ and satisfies $P(D)v = 0$ on $B_{r'}$, there exists (v_j), a sequence in $L^2(B_R)$ so that $P(D)v_j = 0$ on B_R and (v_j) tends to v in $L^2(B_r)$ as $j \to \infty$.*

Proof. By smoothing by convolution we can assume v to be smooth, possibly shrinking r' slightly (to preserve $P(D)v = 0$). We have to show that if $g \in L^2(B_R)$ and satisfies $\langle \alpha, g\rangle_{B_r} = 0$ for all $\alpha \in L^2(B_R)$ such that $P(D)\alpha = 0$, then $\langle v, g\rangle_{B_r} = 0$.

Claim. There exists $w \in L^2(B_R)$ so that for all $\varphi \in \mathscr{C}_0^\infty(\mathbf{R}^n)\langle \varphi, g\rangle_{B_r} = \langle P(D)\varphi, w\rangle_{B_R}$.

Proof of the claim. This is just to say that for some constant C

$$|\langle \varphi, g \rangle_{B_r}| \leqslant C \| P(D)\varphi \|_{B_R}.$$

Notice that if $P(D)\varphi = 0$, we indeed have $\langle \varphi, g \rangle = 0$. If $P(D)\varphi \neq 0$, by §1 we can find $\Psi \in L^2(B_R)$ so that $P(D)\Psi = P(D)\varphi$ and $\|\Psi\|_{B_R} \leqslant C_1 \|P(D)\varphi\|_{B_R}$ (C_1 a constant). Then $\langle \varphi, g \rangle_{B_r} = \langle \varphi - \Psi, g \rangle_{B_r} + \langle \Psi, g \rangle_{B_r} = \langle \Psi, g \rangle_{B_r}$. Hence $|\langle \varphi, g \rangle_{B_r}| \leqslant C \|P(D)\varphi\|_{B_R}$ with $C = C_1 \|g\|$. And the claim is proved.

Pick w as given by the claim. Extend g and w on $\mathbf{R}^n$ to $\tilde{g}$ and $\tilde{w}$, by setting $\tilde{g} = 0$ on $R^n - B_r$ and $\tilde{w} = 0$ on $\mathbf{R}^n - B_R$. We then have (by definition) $\tilde{g} = P^*(D)\tilde{w}$. Since $\tilde{w}$ has compact support, and $P^*(D)\tilde{w}$ is supported by the ball B_r, we conclude from §2 that $w = 0$ on $B_R - B_r$.

Then, take v as in the beginning of the proof, and extend it to be a smooth function on $\mathbf{R}^n$ (but no longer satisfying $P(D)v = 0$ off B_r). One has

$$\langle v, g \rangle_{B_r} = \langle P(D)v, w \rangle_{B_R} = \langle P(D)v, w \rangle_{B_r} = 0. \qquad \text{Q.E.D}$$

4. Solvability of $P(D)u = g$ in $L^2_{\mathrm{loc}}(\mathbf{R}^n)$. The following theorem results immediately from §1 and §3 by a standard procedure (Mittag-Leffler):

THEOREM. *Let $P(D)$ be a (nonzero) constant coefficient linear differential operator on $\mathbf{R}^n$. Then for every $g \in L^2_{\mathrm{loc}}(\mathbf{R}^n)$ there exists $u \in L^2_{\mathrm{loc}}(\mathbf{R}^n)$ such that $P(D)u = g$.*

Proof. As previously, B_p denotes the ball of radius p around 0 in $\mathbf{R}^n$. By §1 there exists $u_1 \in L^2(B_2)$ so that $P(D)u_1 = g$ on B_2. Then, inductively, assuming that u_p has been chosen in $L^2(B_{p+1})$ so that $P(D)u_p = g$, one chooses u_{p+1} in $L^2(B_{p+2})$ in the following way. Let w be an arbitrary solution of $P(D)w = g$, in $L^2(B_{p+2})$. On B_{p+1} one has $P(D)(u_p - w) = 0$. By §3 there exists $v \in L^2(B_{p+2})$ such that: $P(D)v = 0$, and $\|v - (u_p - w)\| B_p \leqslant 1/2^p$. Set $u_{p+1} = v + w$. Then $P(D)u_{p+1} = g$ on B_{p+2}, and $\|u_{p+1} - u_p\|_{B_p} \leqslant 1/2^p$. The sequence (u_p) is obviously convergent in $L^2_{\mathrm{loc}}(\mathbf{R}^n)$, and its limit satisfies $P(D)u = g$.

5. Fundamental solution.

THEOREM (Malgrange-Ehrenpreis [1], [6]). *Every nonzero linear differential operator with constant coefficients on $\mathbf{R}^n$ has a fundamental solution (i.e., a distribution E such that $P(D)E = \delta_0$, δ_0 the Dirac mass at 0).*

At last, we have to assume that the reader knows the definition of a distribution, and the definition of the derivatives of a distribution.

Proof. Let H be the function (product of Heaviside functions) defined on $\mathbf{R}^n$ by: $H(x_1, \ldots, x_n) = 1$ if all the x_j are positive and 0 otherwise. Then $(\partial^n/\partial x_1 \cdots \partial x_n)H = \delta_0$. By §4 there exists $u \in L^2_{\mathrm{loc}}(\mathbf{R}^n)$ so that $P(D)u = H$. Set $E = \partial^n u/(\partial x_1 \cdots \partial x_n)$ (a distribution). Then

$$P(D)E = P(D)\frac{\partial^n u}{\partial x_1 \cdots \partial x_n} = \frac{\partial^n}{\partial x_1 \cdots \partial x_n}(P(D)u) = \delta_0.$$

This ends the proof!

Remark. In the proof one can replace the operator $\partial^n/(\partial x_1 \cdots \partial x_n)$ by any operator which has a fundamental solution that belongs to $L^2_{\mathrm{loc}}(\mathbf{R}^n)$, for example: The Laplacian Δ, if $n = 1, 2, 3$, or $(I - \Delta)^p$ for general n, if $p > n/4$.

REFERENCES

1. L. Ehrenpreis, Solutions of some problems of division I, *Am. J. Math.*, 76 (1954) 883–903.
2. G. B. Folland, Lectures on partial differential equations, *Tata. Inst.*, 1983.
3. L. Hörmander, On the theory of general partial differential operators, *Acta Math.*, 94 (1955) 161–248.
4. ______, Linear partial differential operators, *Grund. Math. Wis.*, 116 (1964) Springer-Verlag.
5. ______, The analysis of linear partial differential operators II, *Grund. Math. Wis.*, 257 (1983) Springer-Verlag.
6. B. Malgrange, Existence et approximation des solutions des équations aux dérivées partielles et des équations de convolution, *Ann. Inst. Fourier*, 6 (1955) 271–355.
7. ______, Equations aux dérivées partielles (conférences aux carrés, rédigées par A. Cerezo), Ecole Normale Supérieure, Paris, 1966.
8. F. Treves, Thèse d' Hörmander I, *Séminaire Bourbaki*, 130 (1956).