Bollettino UMI s. VIII, v. 1-B, n.3 (ott. 1998) 479-500.

Imbedding theorems of Sobolev spaces into Lorentz spaces

Luc TARTAR

CARNEGIE-MELLON University, Pittsburgh 15213, U.S.A.

Dedicated to Jacques-Louis LIONS

When I was a student at Ecole Polytechnique, which was still in Paris on the "Montagne Sainte Geneviève" at the time (1965 to 1967), I had the chance of having two great teachers in Mathematics, Laurent SCHWARTZ and Jacques-Louis LIONS. Apart from a lecture on Calculus of Variations that he taught in place of Laurent SCHWARTZ, Jacques-Louis LIONS taught the Numerical Analysis course, which then meant mostly classical algorithms; partial differential equations only occured in one dimension, and were treated by finite difference schemes, and it was only in a seminar for interested students that I first heard about Sobolev spaces. Later, I heard Jacques-Louis LIONS teach about various technical properties of Sobolev spaces, but although he often used Sobolev imbedding theorem, I do not remember hearing him give a proof. I had read the original proof of SOBOLEV [So], which I had first seen mentioned in Laurent SCHWARTZ's book on distributions [Sc], and the proof that Jacques-Louis LIONS had taught in Montréal [Li], based on the ideas of Emilio GAGLIARDO [Ga].

While I was working for my thesis under the guidance of Jacques-Louis LIONS, I had the pleasure of being invited a few times in a restaurant near the "Halles aux Vins" (the term "Jussieu" was not yet in use). These dinners usually followed talks by famous mathematicians at the seminar which Jacques-Louis LIONS and Laurent SCHWARTZ were organizing every Friday at the Institut Henri POINCARÉ (abbreviated as IHP). The restaurant, "Chez MOISSONNIER", had a room which Laurent SCHWARTZ had once described as a concrete example of a barreled space, and it was there that I first met Sergei SOBOLEV, although I had not been aware that he had given a talk. My understanding of English was too poor at the time to converse with visitors, but fortunately Sergei SOBOLEV spoke French, perfectly.

Shortly after, I noticed that $u \in W^{1,p}(R^N)$ implies $\frac{u}{r} \in L^p(R^N)$ if $1 \leq p < N$ (Appendix I), a result which is not a consequence of Sobolev imbedding theorem; Jacques-Louis LIONS did not know this result, but it had been found before, as we learned a few weeks after when Pierre GRISVARD mentioned it in a talk at the LIONS-SCHWARTZ seminar; as Jacques-Louis LIONS had mentioned to him that I had just proved that result, Pierre GRISVARD did put my name in a reference of an article that he wrote soon after, with an unusual indication, my phone number (I have moved and changed phone number at least six times since, and I do not remember which one I had at that time; I wonder if anyone ever called this number to ask what my proof was).

When I met Sergei SOBOLEV for the second time, at the International Congress of Mathematicians in Nice in 1970, I did not mention my result because I had already noticed that it followed from an improvement by Jaak PEETRE, using imbedding theorems in Lorentz spaces [Pe]. I met Sergei SOBOLEV a third time, when I traveled to Novosibirsk with a group from INRIA in 1976; he was working on completely different questions, and at that time I did not know yet about the various improvements of his imbedding theorem that I am going to describe below.

In the Fall of 1984, in relation with studying functional spaces adapted to the Fokker-Planck equation, I was trying to use an example of a general hypoellipticity result of Lars HÖRMANDER, which I had heard of around 1969 in the LIONS-SCHWARTZ seminar (the talks were not given by Lars HÖRMANDER, whom I only met in 1976): as an example of a much more general theorem, the space V of functions $f(x, v, t)$ on $R^N \times R^N \times R$ satisfying $f \in L^2$, $\frac{\partial f}{\partial t} + \sum_{j=1}^N v_j \frac{\partial f}{\partial x_j} \in L^2$ and $\frac{\partial f}{\partial v_k} \in L^2$ for $k = 1, \ldots, N$, is continuously imbedded in $H^{1/2}_{loc}$ (in all its $2N + 1$ variables); I had easily found a direct proof of that particular example, using a partial Fourier transform; much later I had learned that ROTHSCHILD & STEIN had obtained an L^p version of Lars HÖRMANDER's result. In 1984 my concern was that if one uses the fact that $H^{1/2}$ is imbedded in some L^q space with $q > p$, the value of q would probably not be the largest possible for the space V, as one had not used entirely the information that the functions considered had one full derivative in each of the v_k variables. I was then led to develop a method which could use a non optimal imbedding result and transform it into a better imbedding theorem into Lorentz spaces, giving in particular the largest exponent q for which V is continuously imbedded into L^q. When applied to the classical Sobolev imbedding, my method actually gives the imbedding theorem of $W^{1,p}$ into the Lorentz space $L^{p^*,p}$ for $1 \le p < N$, as had been noticed by Jaak PEETRE [Pe], and for $p = N$ it gives the improvement by Neil TRUDINGER [Tru] of a classical result that Fritz JOHN and Louis NIRENBERG had derived in their pioneering study of BMO [Jo&Ni]. My method also gives a result for functions having their first derivatives in different L^p spaces, a result which I had first heard in a talk of A. KUFNER in Trento in 1978, before I learned from Carlo SBORDONE that it had been obtained earlier by M. TROISI [Tro]. Until a few years ago, I could not find how to derive in the same way the generalization of Neil TRUDINGER's result that Haïm BRÉZIS and Stephen WAINGER had obtained for functions having their derivatives in the Lorentz space $L^{N,p}$ [Br&Wa]. A few years ago, during a meeting in Cortona, I finally understood how to derive their result from a simple variant of my method, and I could then even treat the case of functions having their first derivatives in different Lorentz spaces, a result that seems inaccessible by any of the methods of proofs that I had heard of before.

The results that I present here, which are related to the work of Sergei SOBOLEV, will use simple ideas from the general theory of interpolation spaces developped by Jacques-Louis LIONS and Jaak PEETRE [Li&Pe], which in the particular case of interpolation between L^1 and L^∞ makes the Lorentz spaces appear, and truncation does appear in a natural way in many constructions. It is then with great pleasure that I dedicate these results to Jacques-Louis LIONS.

I. Why use Lorentz spaces?

For $1 \le p < N$, Sobolev imbedding theorem states that $W^{1,p}(R^N)$ is continuously imbedded in $L^{p^*}(R^N)$ where $p^* = \frac{Np}{N-p}$, or $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{N}$. In order to prove it, Sergei SOBOLEV started from the identity $u = \sum_{j=1}^N \big(\frac{\partial E}{\partial x_j} \star \frac{\partial u}{\partial x_j}\big)$ for an elementary solution E of Δ, he noticed that the derivatives of E behave like $\frac{1}{r^{N-1}}$ and he generalized then the classical YOUNG's inequality for convolution when one function is a power of $\frac{1}{r}$, using radial rearrangements. As Sergei SOBOLEV was only using L^q spaces, he could not find the more precise theorem using the family of spaces $L^{q,r}$ introduced by Lorentz, which appeared later to be interpolation spaces between L^1 and L^∞. I will define them in a moment, but here I only need to know that these spaces are defined for $1 < q < \infty$, $1 \le r \le \infty$, that $L^{q,r} \subset L^{q,s}$ whenever $r \le s$, that $L^{q,q} = L^q$, and that $L^{q,\infty}$ is the weak L^q space of MARCINKIEWICZ, defined as $\big\{f : \int_\omega |f|dx \le c|\omega|^{(q-1)/q}$ for every measurable set $\omega\big\}$ ($|\omega|$ denotes the measure of ω); one can actually also use the spaces $L^{q,r}$ with $1 < q < \infty$, but $0 < r < 1$. Multiplication and convolution act in this family of spaces in the following way: if $f \in L^{a,b}$ and $g \in L^{c,d}$ then $fg \in L^{q,r}$ while $f \star g \in L^{s,r}$ where $\frac{1}{q} = \frac{1}{a} + \frac{1}{c}$, $\frac{1}{r} = \frac{1}{b} + \frac{1}{d}$ and $\frac{1}{s} = \frac{1}{a} + \frac{1}{c} - 1$, assuming that $1 < a, c < \infty$, $1 \le b, d \le \infty$, and that $\frac{1}{a} + \frac{1}{b} > 1$ in the convolution case (of course some limiting cases are true: for $1 < p < \infty$, $L^{p,q} \times L^{p',r}$ is mapped into L^1 by multiplication and into L^∞ by convolution if $\frac{1}{q} + \frac{1}{r} \ge 1$). I knew this from an application of a bilinear interpolation theorem that Jacques-Louis LIONS had taught me, but a classical reference for that result is R. O'NEIL [O'Ne]. I believe that Jacques-Louis LIONS's result is unpublished, as I mentioned in my thesis when I extended it to a nonlinear setting, but O'NEIL's method is the same method in disguise, although the clear underlying idea of Jacques-Louis LIONS's proof does not appear well in his treatment. O'NEIL was interested in obtaining precise estimates, but did not mention much about interpolation, an omission which seems general among specialists of singular integrals! As noticed by Jaak PEETRE [Pe], one can improve Sergei SOBOLEV's proof by applying the convolution theorem for Lorentz spaces, and one finds that $W^{1,p}(R^N)$ is not only continuously imbedded in $L^{p^*}(R^N)$, but in the smaller space $L^{p^*,p}(R^N)$, for $1 \le p < N$ of course.

There are some situations where it is useful to know this refined imbedding theorem. Soon after I had found that for $N > 2$, $u \in H^1(R^N)$ implies $\frac{u}{r} \in L^2(R^N)$, a fact that cannot be deduced from Sobolev imbedding theorem, I noticed that it can be deduced from Jaak PEETRE's refined version $W^{1,p}(R^N) \subset L^{p^*,p}(R^N)$ for $1 \le p < N$, showing that $u \in W^{1,p}(R^N)$ implies $\frac{u}{r} \in L^p(R^N)$: indeed, $\frac{1}{r} \in L^{N,\infty}(R^N)$ and $u \in L^{p^*,p}(R^N)$ imply $\frac{u}{r} \in L^{p,p}(R^N) = L^p(R^N)$. This result is useful for proving that functions which are 0 in a neighbourhood of 0 are dense in $W^{1,p}(R^N)$, by approaching a given $u \in W^{1,p}(R^N)$ by $u(x)v(n\,r)$ where $v(s) = 0$ for $0 \le s \le 1$ and $v(s) = 1$ for $s \ge 2$, and the proof follows easily from applying

Lebesgue dominated convergence theorem if one uses the fact that $\frac{u}{r} \in L^p(R^N)$ in order to bound terms like $n\,u(x)v'(n\,r) = \frac{u(x)}{r} n\,r\,v'(n\,r)$. The case $p = N$ is more technical as it is not $\frac{u}{r}$ which belongs to $L^p(R^N)$.

The use of Lorentz spaces may appear then useful to those who know them, but their use is not crucial for the particular example shown above as one can obtain the same result by a direct proof (Appendix I). However, as we will see later, the use of Lorentz spaces is crucial in other situations.

II. Equivalent definitions for Lorentz spaces.

The various definitions of the Lorentz space $L^{p,q}(\Omega)$ consider (equivalence classes of measurable) functions f and use the nonincreasing rearrangement f^* of $|f|$, which is the nonnegative function defined on the interval $(0, |\Omega|)$ which is nonincreasing and equimeasurable to $|f|$.

Definition 1. For $1 < p < \infty$ and $1 \le q \le \infty$, $L^{p,q}(\Omega)$ is the space of functions f such that

$$\begin{gathered}\Big(\int_0^{|\Omega|} |t^{1/p} f^*(t)|^q \frac{dt}{t}\Big)^{1/q} = ||f||_{L^{p,q}(\Omega)} < \infty \text{ if } 1 \le q < \infty,\\ \sup_{0<t<\infty} |t^{1/p} f^*(t)| = ||f||_{L^{p,\infty}(\Omega)} < \infty.\end{gathered} \tag{II.1}$$

One could extend the definition to the case $p = 1$, giving $L^1(\Omega)$ for $q = 1$ and $L^{1,w}(\Omega)$, the weak L^1 space for $q = \infty$, but for $p = \infty$ only $q = \infty$ would make sense and give $L^\infty(\Omega)$.

A second definition, consists in considering Lorentz spaces as interpolation spaces between $L^1(\Omega)$ and $L^\infty(\Omega)$. For $f \in L^1(\Omega) + L^\infty(\Omega)$ one defines $K(t, f)$ by

$$K(t, f) = \int_0^t f^*(s)ds, \tag{II.2}$$

and the relation with the K-method of interpolation of Jaak PEETRE (which simplifies some of his earlier work with Jacques-Louis LIONS) comes from the equivalent definition

$$K(t, f) = \inf\{||g||_{L^1(\Omega)} + t||h||_{L^\infty(\Omega)} : f = g + h\}, \tag{II.3}$$

and truncation appears in this context because an optimal decomposition of f consists in truncating it at some level λ so that $h(x) = f(x)$ when $|f(x)| \le \lambda$ and $h(x) = \lambda\, sign\big(f(x)\big)$ when $|f(x)| > \lambda$, with $\lambda \in [f^*(t + 0), f^*(t - 0)]$.

Most properties of Lorentz spaces can be derived from general interpolation theorems.

Definition 2. For $1 < p < \infty$ and $1 \le q \le \infty$, $L^{p,q}(\Omega)$ is the space of functions $f \in L^1(\Omega) + L^\infty(\Omega)$ such that

$$t^{-\theta} K(t, f) \in L^q\Big(0, \infty; \frac{dt}{t}\Big) \text{ with } \theta = \frac{1}{p'} = 1 - \frac{1}{p}, \tag{II.4}$$

i.e. the interpolation space $\big(L^1(\Omega), L^\infty(\Omega)\big)_{\theta,q}$ with the norm

$$\begin{gathered}\Big(\int_0^\infty |t^{-1/p'} K(t, f)|^q \frac{dt}{t}\Big)^{1/q} = |||f|||_{L^{p,q}(\Omega)} < \infty \text{ for } 1 \le q < \infty\\ \sup_{0<t<\infty} |t^{-1/p'} K(t, f)| = |||f|||_{L^{p,\infty}(\Omega)} < \infty.\end{gathered} \tag{II.5}$$

The norms $||f||_{L^{p,q}(\Omega)}$ and $|||f|||_{L^{p,q}(\Omega)}$ are equivalent for $1 < p < \infty$; as f^* is nonincreasing, one has $K(t, f) \ge t f^*(t)$ and therefore $|||f||| \ge ||f||$; Hardy inequality gives a reversed inequality. More precisely, let u be a nonnegative smooth function with compact support in $(0, \infty)$, denote $U(t) = \int_0^t u(s)\, ds$ and assume that $1 < m\,q$. Then $0 = \int_0^\infty d(t^{-mq+1} U(t)^q) = -(m\,q - 1) \int_0^\infty t^{-mq} U(t)^q\, dt + q \int_0^\infty t^{-mq+1} U(t)^{q-1} u(t)\, dt$, so that $(m\,q-1)||t^{-m}U||_{L^q(0,\infty)} \leq q||t^{-m+1}u||_{L^q(0,\infty)}$ by Hölder inequality, and by density this inequality is valid for $t^{-m+1}u \in L^q(0,\infty)$ if $-m+1 < 1-\frac{1}{q} = \frac{1}{q'}$. One applies it here to $-m+1 = \frac{1}{p}-\frac{1}{q}$, which is allowed if $p>1$, and one obtains

$$\frac{1}{p'}|||f|||_{L^{p,q}(\Omega)} \leq ||f||_{L^{p,q}(\Omega)} \leq |||f|||_{L^{p,q}(\Omega)}. \tag{II.6}$$

Extending the definitions to $p=1$ gives different spaces, as it is now $q=\infty$ which corresponds to $L^1(\Omega)$. As the interpolation framework only considers functions belonging to $L^1(\Omega)+L^\infty(\Omega)$, $L^{1,w}$ does not appear naturally in this context.

The third definition contained in the following Proposition appeared naturally in the method for generalizing Sobolev imbedding theorem that I devised in 1984. Let $k>1$ be chosen, and for any function v defined on Ω and satisfying

$$meas\big\{x\in\Omega : |v(x)|>\lambda\big\} < \infty, \text{ for every } \lambda>0, \tag{II.7}$$

one chooses the levels $a_n \geq 0$, $n\in Z$, such that

$$a_n \in [v^*(k^{-n}+0), v^*(k^{-n}-0)], n\in Z, \tag{II.8}$$

where v^* is the nonincreasing rearrangement of $|v|$, or equivalently, as it will be used in later applications, such that

$$meas\big\{x\in\Omega : |v(x)|>a_n\big\} \leq k^{-n} \leq meas\big\{x\in\Omega : |v(x)|\geq a_n\big\}. \tag{II.9}$$

If the function $|v|$ avoids an interval, it may well happen that many choices of a_n are possible, so in order to be more precise in some inequalities, one defines a_n^- and a_n^+ by

$$a_n^+ = v^*(k^{-n}-0);\ a_n^- = v^*(k^{-n}+0). \tag{II.10}$$

Proposition 3. For $1<p<\infty$ and $1\leq q\leq\infty$, and v satisfying (II.7) and extended by 0 outside Ω, one has

$$v\in L^{p,q}(R^N) \text{ is equivalent to } k^{-n/p}(a_{n+1}-a_n)\in l^q. \qquad \text{(II.11))}$$

Proof: Using the fact that all the intervals $(k^{-(n+1)}, k^{-n})$ have the same measure $\log(k)$ for $\frac{dt}{t}$, that v^* is nonincreasing and therefore $a_n^- \leq a_n \leq a_n^+$ for all n, one deduces

$$\begin{aligned} &\log(k)\sum_{n\in Z}|k^{-(n+1)/p}a_n|^q \leq \log(k)\sum_{n\in Z}|k^{-(n+1)/p}a_n^+|^q \leq ||v||^q_{L^{p,q}(\Omega)} \\ &||v||^q_{L^{p,q}(\Omega)} \leq \log(k)\sum_{n\in Z}|k^{-n/p}a_n^-|^q \leq \log(k)\sum_{n\in Z}|k^{-n/p}a_n|^q. \end{aligned} \tag{II.12}$$

As $a_{n+1}-a_n \leq a_{n+1}^+$ for all n, (II.12) shows that $v\in L^{p,q}(\Omega)$ implies $k^{-n/p}(a_{n+1}-a_n)\in l^q(Z)$.

Conversely, defining $b_n = |a_n - a_{n-1}|$ and assuming that $k^{-n/p}b_n \in l^q(Z)$, one wants to deduce that $k^{-n/p}a_n \in l^q(Z)$. Indeed, as $a_n = \sum_{m=-\infty}^n b_m$ because $a_m \to 0$ as m tends to $-\infty$ as a consequence of (II.7), one finds that $k^{-n/p}a_n \leq \sum_{m=-\infty}^n k^{(m-n)/p}(k^{-m/p}b_m)$, and by using a classical convolution inequality, one deduces that

$$\Big(\sum_{n\in Z}|k^{-n/p}a_n|^q\Big)^{1/q} \leq \Big(\sum_{m\leq 0}k^{m/p}\Big)\Big(\sum_{n\in Z}|k^{-n/p}(a_n-a_{n-1})|^q\Big)^{1/q}. \tag{II.13}$$

III. Old and new variants of the imbedding theorem of Sergei SOBOLEV

In 1984, dealing with a question that I describe in Appendix II, I was led to consider the following situation, connected to imbedding theorems for Sobolev spaces. Suppose that for $1 \le p < \infty$, one has found an imbedding theorem of the type $W^{1,p}(R^N) \subset L^q(R^N)$ with $q > p$, i.e. one has proved an inequality

$$\Big(\int_{R^N} |v(x)|^q\,dx\Big)^{1/q} \le C_1\Big(\int_{R^N} |grad(v)|^p\,dx\Big)^{1/p} + C_2\Big(\int_{R^N} |v(x)|^p\,dx\Big)^{1/p} \text{ for all } v \in \mathcal{D}(R^N). \tag{III.1}$$

Although mathematicians often write inequalities of this kind, physicists tend to question the unnatural habit of adding terms like $\big(\int_{R^N} |grad(v)|^p\,dx\big)^{1/p}$ and $\big(\int_{R^N} |v(x)|^p\,dx\big)^{1/p}$, which are not expressed in the same unit. In order to correct this small defect, one applies (III.1) to the rescaled function w defined by $w(x) = v(\lambda x)$ with $\lambda \ne 0$, and (III.1) becomes

$$\lambda^{-N/q}\Big(\int_{R^N} |v(x)|^q\,dx\Big)^{1/q} \le C_1\lambda^{1-(N/p)}\Big(\int_{R^N} |grad(v)|^p\,dx\Big)^{1/p} + C_2\lambda^{-N/p}\Big(\int_{R^N} |v(x)|^p\,dx\Big)^{1/p}, \tag{III.2}$$

and then one chooses the best value of λ, in order to deduce the following inequality, which is now invariant by scaling:

$$\Big(\int_{R^N} |v(x)|^q\,dx\Big)^{1/q} \le C_3\Big(\int_{R^N} |grad(v)|^p\,dx\Big)^{\theta/p}\Big(\int_{R^N} |v(x)|^p\,dx\Big)^{(1-\theta)/p} \text{ for all } v \in \mathcal{D}(R^N) \tag{III.3}$$

with θ defined by

$$\frac{1}{q} = \frac{\theta}{p^*} + \frac{1-\theta}{p} = \frac{1}{p} - \frac{\theta}{N} \text{ and } \frac{1}{p^*} = \frac{1}{p} - \frac{1}{N}. \tag{III.4}$$

One has $0 < \theta \le 1$, because in the case $1 \le p < N$, one must have $q \le p^*$, where p^* is the so-called Sobolev exponent of p, as the case $q > p^*$ leads to a contradiction by letting λ tend to ∞.

After having taken advantage of rescaling in $x \in R^N$, I wondered about the effect of rescaling in $v \in R$, but as changing v into $k\,v$ has no effect on (III.3), I considered a nonlinear rescaling, replacing v by $\varphi(v)$ for a list of suitable functions φ. As both the norms of $\varphi(v)$ in $L^p(R^N)$ and in $L^q(R^N)$ appear in (III.3), I had to use functions φ for which these norms could be compared, and I was led to introduce the particular sequence of functions φ_n defined by

$$\varphi_n(v) = \begin{cases} 0 & \text{if } 0 \le |v| \le a_n \\ |v| - a_n & \text{if } a_n \le |v| \le a_{n+1} \\ a_{n+1} - a_n & \text{if } |v| \ge a_{n+1}, \end{cases} \tag{III.5}$$

with the levels a_n defined for $n \in Z$ as in (II.8). As $\varphi_n(v) \ge (a_{n+1} - a_n)\chi_{n+1}^+$ where χ_{n+1}^+ is the characteristic function of the set where $|v| \ge a_{n+1}$ and $\varphi_n(v) \le (a_{n+1} - a_n)\chi_n^-$ where χ_n^- is the characteristic function of the set where $|v| > a_n$, one finds by using (II.9) that

$$k^{-(n+1)/r}(a_{n+1} - a_n) \le \Big(\int_{R^N} |\varphi_n(v)|^r\,dx\Big)^{1/r} \le k^{-n/r}(a_{n+1} - a_n) \text{ for } 0 < r < \infty. \tag{III.6}$$

Using $\varphi_n(v)$ in (III.3), and defining p^* as in (III.4) even for $N \le p < \infty$ (in which case $p^* \le 0$), leads to the following improvement of Sobolev imbedding theorem.

Proposition 4. For $1 \le p < \infty$ and $v \in W^{1,p}(R^N)$, one has

$$(a_{n+1} - a_n)k^{-n/p^*} \le C(k,p,N)\Big(\int_{a_n < |v| < a_{n+1}} |grad(v)|^p\,dx\Big)^{1/p} \in l^p. \tag{III.7}$$

Remark 5. In the case $1 \le p < N$, Proposition 4 gives the variant of Jaak PEETRE, $W^{1,p}(R^N) \subset L^{p^*,p}(R^N)$, because (III.7) implies $v \in L^{p^*,p}(R^N)$ by Proposition 3.

In the case $N < p < \infty$, Proposition 4 gives $W^{1,p}(R^N) \subset L^\infty(R^N)$ as $a_{n+1} - a_n$ is bounded by a convergent geometric series for $n \geq 0$ and a_0 can be estimated from the norm of v in $L^p(R^N)$. It is then easy to deduce that $W^{1,p}(R^N) \subset C^{0,\alpha}(R^N)$ for $\alpha = 1 - \frac{N}{p}$ by applying the result to $v(x-h) - v(x)$ for $h \in R^N$, and noticing that one has $||v(\cdot - h) - v(\cdot)||_{L^p} \leq |h| \, ||grad(v)||_{L^p}$.

In the case $p = N$, one has $p^* = 0$ and Proposition 4 gives $a_{n+1} - a_n \in l^N(Z)$, from which one can deduce the theorem of Neil TRUDINGER: for $v \in W^{1,N}(R^N)$ and for every $\lambda > 0$ one has $e^{\lambda v^{N'}} \in L^1_{loc}(R^N)$, where N' is the conjugate exponent of N [Tru]. Indeed, as $a_{n+1} - a_n \in l^N(Z)$, one deduces that for every $\varepsilon > 0$ one has $a_n^{N'} \leq \varepsilon n + C_\varepsilon(v)$ for all $n \geq 0$ by applying Hölder inequality to the sequence $a_{n+1} - a_n$ for $n \geq m$ with m large enough (and this choice depends upon v). On the set where $|v| < a_n$, which has measure $\leq k^{-n}$, one has $e^{\lambda v^{N'}} \leq e^{\lambda(\varepsilon n + C_\varepsilon(v))}$, and by choosing ε such that $e^{\lambda \varepsilon} < k$ one finds that $e^{\lambda v^{N'}}$ is integrable on any set where $v \geq \alpha > 0$.

The improved version (III.7) of Sobolev imbedding theorem follows then from any crude imbedding theorem (III.1), but one may even start from Sobolev imbedding theorem itself to deduce the improvement of Jaak PEETRE, i.e. $q = p^*$ if $1 \leq p < N$, $N < q < \infty$ if $p = N$ and $q = \infty$ if $p > N$. For more general cases like those described in the following Remark or in Appendix II, a crude estimate can be obtained by the method of Appendix III.

Remark 6. The method extends to spaces of functions v satisfying

$$v \in L^{p_0}(R^N); \; \frac{\partial v}{\partial x_j} \in L^{p_j}(R^N) \text{ for } j = 1, \ldots, N, \tag{III.8}$$

where $1 \leq p_j \leq \infty, j = 0, \ldots, N$. If one knows an analog of (III.1), i.e. an inequality

$$||v||_{L^q} \leq C\Big(||v||_{L^{p_0}} + \sum_{j=0}^N \Big|\Big| \frac{\partial v}{\partial x_j} \Big|\Big|_{L^{p_j}}\Big) \text{ for all } v \in \mathcal{D}(R^N), \tag{III.9}$$

for some $q > p_0$, one applies (III.9) to the rescaled function w defined by

$$w(x) = v(\lambda_1 x_1, \ldots, \lambda_N x_N), \tag{III.10}$$

where $\lambda_1, \ldots, \lambda_N$, are positive parameters. Defining μ by

$$\mu = (\lambda_1 \ldots \lambda_N)^{1/N}, \tag{III.11}$$

the best choice of λ for a given μ corresponds to

$$\lambda_j \mu^{-N/p_j} \Big|\Big| \frac{\partial v}{\partial x_j} \Big|\Big|_{L^{p_j}} = \mu^{1-(N/p)} \Big(\prod_{k=1}^N \Big|\Big| \frac{\partial v}{\partial x_k} \Big|\Big|_{L^{p_k}} \Big)^{1/N}, j = 1, \ldots, N, \tag{III.12}$$

which produces the inequality

$$\mu^{-N/q} ||v||_{L^q} \leq C\Big(\mu^{-N/p_0} ||v||_{L^{p_0}} + N \mu^{1-(N/p)} \Big(\prod_{k=0}^N \Big|\Big| \frac{\partial v}{\partial x_k} \Big|\Big|_{L^{p_k}} \Big)^{1/N}\Big) \text{ for all } v \in \mathcal{D}(R^N), \tag{III.13}$$

and the best choice of μ gives

$$||v||_{L^q} \leq C' \Big(\prod_{k=0}^N \Big|\Big| \frac{\partial v}{\partial x_k} \Big|\Big|_{L^{p_k}} \Big)^{\theta/N} ||v||^{1-\theta}_{L^{p_0}} \text{ for all } v \in \mathcal{D}(R^N), \tag{III.14}$$

where p, θ (and p^*) are defined by

$$\frac{1}{p} = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{p_j};\ \frac{1}{p^*} = \frac{1}{p} - \frac{1}{N};\ \frac{1}{q} = \frac{\theta}{p^*} + \frac{1-\theta}{p}. \tag{III.15}$$

Then the application to $\varphi_n(v)$ of (III.14) gives

$$(a_{n+1} - a_n)k^{-n/p^*} \le C''_p \prod_{j=1}^{N}\Big(\int_{a_n<|v|<a_{n+1}} \Big|\frac{\partial v}{\partial x_j}\Big|^{p_j}\,dx\Big)^{1/Np_j} \in l^p(Z), \tag{III.16}$$

and therefore $v \in L^{p^*,p}(R^N)$ if $p < N$ and $u \in L^\infty(R^N)$ if $p > N$ in particular, generalizing a result of M. Troisi [Tro].

The original method of proof of Sergei Sobolev is not adapted to situations where derivatives belong to different spaces. The previous proof that I had heard in 1978 from A. Kufner relied on a classical method which, I was told, was introduced independently by Emilio Gagliardo and by Louis Nirenberg; I have also heard Jacques-Louis Lions use a similar argument to prove a result that he attributed to Olga Ladyzhenskaya, and I describe a different way to obtain this type of result in Appendix IV.

The preceding result is a particular case of a more general one where one considers functions having their partial derivatives in various Lorentz spaces, but I was not able to prove the theorem that I expected in this general case by using my original argument. I had noticed that (III.7) was a direct consequence of Sobolev imbedding theorem for $W^{1,q}$ with $1 \le q \le p$, and therefore the case for $q = 1$ implied all the other known results by Sergei Sobolev, Jaak Peetre, Neil Trudinger, but not that of Haïm Brézis & Stephen Wainger, and this discrepancy bothered me for a long time.

Sergei Sobolev certainly knew that his imbedding theorem for $p = 1$ was related to the isoperimetric inequality (which had been known for quite a while, and is equivalent to it if one uses the imbedding theorem with the best constant), and I wonder if it was known that one can deduce Sobolev imbedding theorem for $W^{1,p}$ and the improvements that I have mentioned above from it: as I do not read much, I hope that someone will give me a reference if it had been done before.

When inspiration came a few years ago, during a meeting in Cortona, it gave me the way not only to prove the result of Haïm Brézis & Stephen Wainger but the more general case where the derivatives may belong to different Lorentz spaces, and I did not need to change much my original argument. The simple trick which had escaped my attention for so long was that instead of an additive form of the imbedding theorem for $W^{1,1}$ (or the isoperimetric inequality), I should have used the following multiplicative form.

Lemma 7. There exists a constant C such that for every $v \in W^{1,1}(R^N)$ one has

$$||v||_{L^{1^*}} \le C\Big(\prod_{j=1}^{N}\Big|\Big|\frac{\partial v}{\partial x_j}\Big|\Big|_{L^1}\Big)^{1/N}. \tag{III.17}$$

Proof: As in Remark 6, one starts from the classical Sobolev imbedding theorem for $p = 1$,

$$||v||_{L^{1^*}} \le C\sum_{j=1}^{N}\Big|\Big|\frac{\partial v}{\partial x_j}\Big|\Big|_{L^1}, \tag{III.18}$$

and one applies it to the function w defined by (III.10), and (III.17) results from the choice

$$\lambda_i = \prod_{j\ne i}\Big|\Big|\frac{\partial v}{\partial x_j}\Big|\Big|_{L^1}. \tag{III.19}$$

The best constant in (III.17) is N times the best constant in (III.18), which is related to the classical isoperimetric inequality.

In the case where all the derivatives belong to the same Lorentz space $L^{N,p}(R^N)$, Haïm BRÉZIS & Stephen WAINGER [Br&Wa] had shown that

$$e^{C\,|v|^{p'}} \in L^1_{loc}(R^N) \text{ for every } C > 0, \tag{III.20}$$

in the case $1 < p < \infty$, extending the result of Neil TRUDINGER who had considered the case $p = N$ [Tru], and their proof followed the method introduced by Sergei SOBOLEV, but they based their estimate of the convolution product on O'NEIL's formula [O'Ne], which they had to analyze in detail for this limiting case. For $p = 1$ one has $v \in C_0(R^N)$, and for $p = \infty$ one has $v \in BMO(R^N)$ and therefore by a classical result of Fritz JOHN & Louis NIRENBERG [Jo&Ni], $e^{\varepsilon|v|} \in L^1_{loc}(R^N)$ for $\varepsilon > 0$ small enough. The following Theorem will extend these results to the more general situation of functions v satisfying (II.7) and

$$\frac{\partial v}{\partial x_j} \in L^{p_j,q_j}(R^N), \text{ for } j = 1, \ldots, N. \tag{III.21}$$

Theorem 8: Assume that v satisfies (II.7) and (III.21) with $1 \le p_j, q_j \le \infty$ for $j = 1, \ldots, N$ (and $q_j = p_j$ if $p_j = 1$ or $p_j = \infty$), then one has

$$(a_{n+1} - a_n)k^{-n/p^*} \in l^q(Z), \tag{III.22}$$

where

$$\frac{1}{p} = \frac{1}{N}\sum_{j=1}^N \frac{1}{p_j};\ \ \frac{1}{q} = \frac{1}{N}\sum_{j=1}^N \frac{1}{q_j};\ \ \frac{1}{p^*} = \frac{1}{p} - \frac{1}{N}. \tag{III.23}$$

In particular, $v \in L^{p^*,q}(R^N)$ if $p < N$, $v \in L^\infty(R^N)$ if $p > N$ or if $p = N$ and $q = 1$, $e^{C|v|^{q'}} \in L^1_{loc}(R^N)$ for every $C > 0$ if $p = N$ and $1 < q < \infty$, $e^{\varepsilon|v|} \in L^1_{loc}(R^N)$ for $\varepsilon > 0$ small enough if $p = N$ and $q = \infty$.
Proof: One applies Lemma 7 to the sequence of functions $\varphi_n(v)$. If one denotes

$$f_j = \frac{\partial v}{\partial x_j} \text{ for } j = 1, \ldots, N, \tag{III.24}$$

one has

$$\Big|\Big|\frac{\partial \varphi_n(v)}{\partial x_j}\Big|\Big|_1 \le \int_0^{k^{-n}} f_j^*(t)\,dt \text{ for } j = 1, \ldots, N, \tag{III.25}$$

where f_j^* is the nonincreasing rearrangement of f_j, by a classical result of HARDY and LITTLEWOOD [Ha&Li&Po], as $\varphi_n'(v)$ is different from 0 on a subset of measure at most k^{-n}. (III.17) applied to $\varphi_n(v)$ implies then

$$(a_{n+1} - a_n)k^{-n/1^*} \le C\Big(\prod_{j=1}^N \int_0^{k^{-n}} f_j^*(t)\,dt\Big)^{1/N}. \tag{III.26}$$

As $f_j \in L^{p_j,q_j}(R^N)$ means $t^{-\theta_j}\int_0^t f_j^*(s)\,ds \in L^{q_j}\big(0,\infty;\frac{dt}{t}\big)$, or equivalently $k^{n\theta_j}\int_0^{k^{-n}} f_j^*(s)\,ds \in l^{q_j}(Z)$, where $\theta_j = 1 - \frac{1}{p_j}$, one deduces that $k^{n\theta}\big(\prod_{j=1}^N \int_0^{k^{-n}} f_j^*(t)\,dt\big)^{1/N} \in l^q(Z)$ with $\theta = \frac{1}{N}\sum_{j=1}^N \theta_j = 1 - \frac{1}{p}$, and as $k^{-n/1^*}k^{n\theta} = k^{-n/p*}$, (III.26) implies (III.22).

Remark 9: In the case $p = N$ and $q = 1$ (with all $p_j < \infty$), one has actually $v \in C_0(R^N)$. As $a_{n+1} - a_n \in l^1(Z)$ and a_n tends to 0 as n tends to $-\infty$ because of the condition (II.7), one has proved an inequality of the form

$$||v||_{L^\infty} \le C\sum_{j=1}^N \Big|\Big|\frac{\partial v}{\partial x_j}\Big|\Big|_{L^{p_j,1}} \text{ or } ||v||_{L^\infty} \le C'\Big(\prod_{j=1}^N \Big|\Big|\frac{\partial v}{\partial x_j}\Big|\Big|_{L^{p_j,1}}\Big)^{1/N}$$
$$\text{if } v \text{ satisfies (II.7) and (III.21), in the case } \sum_{j=1}^N \frac{1}{p_j} = 1. \tag{III.27}$$

Notice that the hypothesis (II.7) does not allow to add a nonzero constant to v. As the difference of two functions satisfying (II.7) also satisfies (II.7), one can apply the preceding inequality to $v(\cdot+h)-v(\cdot)$, whose derivative in x_j is $f_j(\cdot+h)-f_j(\cdot)$ and has a small norm in $L^{p_j,1}$ when $|h|$ is small (as $p_j<\infty$ implies that smooth functions with compact support are dense in $L^{p_j,1}$). Therefore (III.27) shows that v is uniformly continuous, and as (II.7) holds v must tend to 0 at infinity, so that $v \in C_0(R^N)$.

In the case $p>N$, v is Hölder continuous, but with different orders according to the directions if the p_j are distinct. If v satisfies (III.21) and $v \in L^{r,s}$ with $1 \le r < \infty$ and $s=1$ if $r=1$, one has $a_0 \le C||v||_{L^{r,s}}$ and therefore an inequality of the form

$$||v||_{L^\infty} \le C\Big(||v||_{L^{r,s}} + \sum_{j=1}^N \Big|\Big|\frac{\partial v}{\partial x_j}\Big|\Big|_{L^{p_j,q_j}}\Big), \tag{III.28}$$

which after using the rescaling (III.10) as in Remark 6 gives

$$||v||_{L^\infty} \le C||v||_{L^{r,s}}^{1-\theta}\Big(\prod_{j=1}^N \Big|\Big|\frac{\partial v}{\partial x_j}\Big|\Big|_{L^{p_j,q_j}}\Big)^{\theta/N}, \tag{III.29}$$

and θ, which depends upon r, is the only value which makes both sides vary in the same way under rescaling, i.e. $\frac{1-\theta}{r} + \big(\sum_{j=1}^N \frac{1}{p_j} - 1\big)\frac{\theta}{N} = 0$, or $\frac{1-\theta}{r} + \frac{\theta}{p^*} = 0$, noticing that $p^* < 0$. In order to find the Hölder exponent in the direction x_i of a function v satisfying (III.21) and (II.7), one applies the preceding inequality to $v(\cdot + t\,e_i) - v(\cdot)$, whose norm in L^{p_i,q_i} is bounded by $|t|\,||f_i||_{L^{p_i,q_i}}$, and therefore (III.29) with $r = p_i$ implies

$$\begin{aligned} &|v(x+t\,e_i)-v(x)| \le C\,|t|^{\alpha_i}\Big|\Big|\frac{\partial v}{\partial x_i}\Big|\Big|_{L^{p_i,q_i}}^{\alpha_i}\Big(\prod_{j=1}^N \Big|\Big|\frac{\partial v}{\partial x_j}\Big|\Big|_{L^{p_j,q_j}}\Big)^{(1-\alpha_i)/N} \text{ for } i=1,\ldots,N,\\ &\alpha_i\Big(1-\frac{N}{p}+\frac{N}{p_i}\Big) = 1-\frac{N}{p} \text{ for } i=1,\ldots,N,\\ &\text{in the case } p>N, \text{ i.e. } \sum_{j=1}^N \frac{1}{p_j} < 1. \end{aligned} \tag{III.30}$$

Appendix I.

For $p = 2 < N$, a standard method, which I learned in Hardy & Littlewood & Polya [Ha&Li&Po], consists in developping the inequality

$$\int_{R^N} \sum_{j=1}^N \Big| \frac{\partial u}{\partial x_j} - a(r)\frac{x_j}{r}\Big|^2\,dx \ge 0, \tag{A.I.1}$$

for a smooth function u having compact support, where the function $a(r)$ must be chosen. After integrating by parts the terms in $u\frac{\partial u}{\partial x_j}$, one obtains

$$\int_{R^N} \Big(\sum_{j=1}^N \Big|\frac{\partial u}{\partial x_j}\Big|^2\Big)\,dx \ge \int_{R^N} |u|^2\Big(-a(r)^2 - \sum_{j=1}^N \frac{\partial}{\partial x_j}\Big(a(r)\frac{x_j}{r}\Big)\Big)\,dx. \tag{A.I.2}$$

Choosing $a(r) = \frac{K}{r}$ gives for the coefficient of $|u|^2$: $-a^2 - a' - \frac{(N-1)a}{r} = \frac{-K^2+(2-N)K}{r^2}$ and for $N > 2$ the best choice is $K = \frac{2-N}{2}$ (for $N = 2$ one can choose $a(r) = \frac{1}{2rLog(r/r_0)}$ and one must use an open set where $r \ne r_0$).

For $p \ne 2$ and $p < N$, one can follow the same idea with a suitable convexity inequality: from the convexity of the function $z \mapsto \big(\sum_j |z_j|^2\big)^{p/2}$ on R^N, applied at the point $a(r)u\frac{x}{r}$, one obtains

$$\Big(\sum_{j=1}^N \Big|\frac{\partial u}{\partial x_j}\Big|^2\Big)^{p/2} \ge |a(r)u|^2 + \sum_{j=1}^N p|a(r)u|^{p-2}a(r)\frac{x_j}{r}u\Big(\frac{\partial u}{\partial x_j} - a(r)\frac{x_j}{r}\Big), \tag{A.I.3}$$

which, after integration by parts, gives

$$\int_{R^N} \Big(\sum_{j=1}^N \Big|\frac{\partial u}{\partial x_j}\Big|^2\Big)^{p/2} dx \ge \int_{R^N} |u|^p\Big((1-p)|a|^p - \sum_{j=1}^N \frac{\partial}{\partial x_j}(|a|^{p-2}pa\frac{x_j}{r})\Big)\,dx, \tag{A.I.4}$$

and the choice $a(r) = \frac{\alpha}{r}$ gives the coefficient of $|u|^p$ equal to $\frac{\beta}{r^p}$ with $\beta = (1-p)|\alpha|^p - (N-p)|\alpha|^{p-2}\alpha$, so the best choice is $\alpha = \frac{p-N}{p}$. The preceding computations actually give the best constants in the desired inequalities.

Appendix II.

In 1984, I had developed the method of paragraph III for an academic question related to the Fokker-Planck equation, and a classical imbedding theorem was not available in that case: I assumed that a function f defined on $R^N \times R^N \times R$ (with arguments x, v, t corresponding to position, velocity and time) satisfied

$$f,\ \frac{\partial f}{\partial t} + \sum_{j=1}^N v_j \frac{\partial f}{\partial x_j},\ \frac{\partial f}{\partial v_k}, \in L^p(R^N \times R^N \times R),\ k = 1,\ldots,N, \tag{A.II.1}$$

and I wanted to deduce that f belongs to $L^q(R^N \times R^N \times R)$ for the best possible value of q. I assumed then that for some $q > p$ one already knew an imbedding theorem

$$||f||_q \le C\Big(\Big|\Big|\frac{\partial f}{\partial t} + \sum_{j=1}^N v_j\frac{\partial f}{\partial x_j}\Big|\Big|_p + \sum_{k=1}^N \Big|\Big|\frac{\partial f}{\partial v_k}\Big|\Big|_p + ||f||_p\Big) \text{ for } f \in \mathcal{D}(R^N \times R^N \times R), \tag{A.II.2}$$

where $||\cdot||_r$ denotes the norm in $L^r(R^N \times R^N \times R)$. One uses then the rescaled function g defined by $g(x,v,t) = f(a\,x, b\,v, c\,t)$ with the relation $a = b\,c$ because the units for length, velocity and time are dependent, and (A.II.2) applied to g gives

$$b^{-2N/q}c^{-(N+1)/q}||f||_q \le C\Big(b^{-2N/p}c^{(p-N-1)/p}\Big|\Big|\frac{\partial f}{\partial t} + \sum_{j=1}^N v_j\frac{\partial f}{\partial x_j}\Big|\Big|_p + b^{(p-2N)/p}c^{-(N+1)/p}\sum_{k=1}^N\Big|\Big|\frac{\partial f}{\partial v_k}\Big|\Big|_p + b^{-2N/p}c^{-(N+1)/p}||f||_p\Big), \tag{A.II.3}$$

where one must then choose the best values of b, c in order to deduce an inequality which is invariant by scaling; (A.II.3) has the form

$$||f||_q \leq C\Big(b^{-\alpha}c^{1-\beta}\Big|\Big|\frac{\partial f}{\partial t}+\sum_{j=1}^{N} v_j\frac{\partial f}{\partial x_j}\Big|\Big|_p + b^{1-\alpha}c^{-\beta}\sum_{k=1}^{N}\Big|\Big|\frac{\partial f}{\partial v_k}\Big|\Big|_p + b^{-\alpha}c^{-\beta}||f||_p\Big), \tag{A.II.4}$$

with $\alpha = \frac{2N}{p} - \frac{2N}{q}, \beta = \frac{N+1}{p} - \frac{N+1}{q}$, and $q > p$ implies $\alpha, \beta > 0$, but one must also have $\alpha + \beta \leq 1$, as $\alpha + \beta > 1$ would imply $||f||_q = 0$ by letting $b = c$ tend to ∞. Defining $\theta = \alpha + \beta$, one has $0 < \theta \leq 1$ and

$$\frac{1}{q} = \frac{1}{p} - \frac{\theta}{3N+1}, \tag{A.II.5}$$

and if $p < 3N + 1$ one has $p < q \leq p^{**}$, where

$$\frac{1}{p^{**}} = \frac{1}{p} - \frac{1}{3N+1}. \tag{A.II.6}$$

The choice $b = \frac{||f||_p}{\big|\big|\frac{\partial f}{\partial t}+\sum_{j=1}^{N} v_j\frac{\partial f}{\partial x_j}\big|\big|}, c = \frac{||f||_p}{\sum_{k=1}^{N}\big|\big|\frac{\partial f}{\partial v_k}\big|\big|}$ gives

$$||f||_q \leq C\Big|\Big|\frac{\partial f}{\partial t}+\sum_{j=1}^{N} v_j\frac{\partial f}{\partial x_j}\Big|\Big|_p^{(N+1)\theta/(3N+1)}\Big(\sum_{k=1}^{N}\Big|\Big|\frac{\partial f}{\partial v_k}\Big|\Big|_p^{2N\theta/(3N+1)}\Big)||f||_p^{1-\theta}, \tag{A.II.7}$$

and applying (A.II.7) to the sequence of functions $\varphi_n(f)$ with φ_n defined as in (II.6), one obtains

$$(a_{n+1}-a_n)k^{-n/p^{**}} \leq C\Big|\Big|\varphi_n'(f)\Big(\frac{\partial f}{\partial t}+\sum_{j=1}^{N} v_j\frac{\partial f}{\partial x_j}\Big)\Big|\Big|_p^{(N+1)/(3N+1)}\Big(\sum_{k=1}^{N}\Big|\Big|\varphi_n'(f)\,\frac{\partial f}{\partial v_k}\Big|\Big|_p^{2N/(3N+1)}\Big), \tag{A.II.8}$$

with p^{**} defined by (A.II.6) (even for $3N + 1 \leq p$, in which case $p^{**} \leq 0$).

Of course, (A.II.8) means that the space of functions defined by (A.II.1) is imbedded in $L^{p^{**},p}(R^N \times R^N \times R)$ if $1 \leq p < 3N + 1$, in $L^\infty(R^N \times R^N \times R)$ if $p > 3N + 1$; in the case $p = 3N + 1$ it is imbedded in $L^r(R^N \times R^N \times R)$ for any r such that $p \leq r < \infty$, and moreover for every $\lambda > 0$ one has $e^{\lambda|f|^{p'}} \in L^1_{loc}$, where p' is the conjugate exponent of $3N + 1$.

Appendix III.

The first method described in Section III requires a crude imbedding estimate like (III.1) or (III.10) (or (A.II.2) in Appendix II), while the second method for proving the more general Theorem 8 only uses the classical imbedding theorem for $W^{1,1}(R^N)$. I show here a simple method for obtaining the crude estimates needed. Actually, Lorentz spaces will also appear naturally, at least the Marcinkiewicz spaces $L^{q,\infty}$, but one can easily use a crude estimate involving $L^{q,\infty}$ norms, as these norms scale like L^q norms, and the $L^{q,\infty}$ norm of a characteristic function coincides with its L^q norm.

Let $u \in W^{1,p}(R^N)$. One would like to decompose $u = u_0 + u_1$, with $u_0 \in L^1$ and $u_1 \in L^\infty$, and obtain precise bounds for the norms of u_0, u_1, but when one uses the natural idea of defining the terms u_1 by convolution with a smoothing sequence, it will be the L^p norm of the corresponding term u_0 that will be easy to bound. Let ρ be a bounded function with compact support having integral 1, and for $\varepsilon > 0$ let ρ_ε be defined as

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon^N}\rho\Big(\frac{x}{\varepsilon}\Big), x \in R^N, \tag{A.III.1}$$

and consider the decomposition

$$u = u_0 + u_1 \text{ with } u_0 = u - \rho_\varepsilon \star u, u_1 = \rho_\varepsilon \star u. \tag{A.III.2}$$

Denoting by $||f||_q$ the norm of a function f in $L^q, 1 \leq q \leq +\infty$, Hölder inequality gives

$$||u_1||_\infty \leq ||u||_p ||\rho_\varepsilon||_{p'} = ||u||_p ||\rho||_{p'} \varepsilon^{-N/p}. \tag{A.III.3}$$

In order to bound $||u_0||_p$, one uses the fact that

$$||u - \tau_a u||_p \leq \sum_{j=1}^N |a_j| \, \Big|\Big| \frac{\partial u}{\partial x_j} \Big|\Big|_p \text{ for } a \in R^N \text{ and } u \in W^{1,p}(R^N), \tag{A.III.4}$$

where $\tau_a f(x) = f(x-a)$, and from

$$||u - \rho_\varepsilon \star u||_p \leq \int_{R^N} |\rho_\varepsilon(y)| \, ||u - \tau_y u||_p dy \tag{A.III.5}$$

one deduces that

$$||u_0||_p \leq \sum_{j=1}^N \Big|\Big| \frac{\partial u}{\partial x_j} \Big|\Big|_p \int_{R^N} |\rho_\varepsilon(y)| \, |y_j| \, dy = \varepsilon \sum_{j=1}^N \Big|\Big| \frac{\partial u}{\partial x_j} \Big|\Big|_p \int_{R^N} |\rho(y)| \, |y_j| \, dy. \tag{A.III.6}$$

These estimates show that u belongs to an interpolation space between L^p and L^∞, which is actually a Marcinkiewicz space $L^{q,\infty}$. This can be seen by either using the interpolation theory and the reiteration theorem of Jacques-Louis LIONS & Jaak PEETRE [Li&Pe], or by estimating the integral of $|u|$ on an arbitrary set ω

$$\begin{aligned} \int_\omega |u| \, dx &\leq \int_\omega |u_0| \, dx + \int_\omega |u_1| \, dx \leq ||u_0||_p |\omega|^{1/p'} + ||u_1||_\infty |\omega| \\ &\leq \varepsilon |\omega|^{1/p'} \sum_{j=1}^N \Big|\Big| \frac{\partial u}{\partial x_j} \Big|\Big|_p \int_{R^N} |\rho(y)| \, |y_j| \, dy + \varepsilon^{-N/p} |\omega| ||u||_p ||\rho||_{p'}, \end{aligned} \tag{A.III.7}$$

where $|\omega|$ denotes the measure of ω. Choosing $\varepsilon = |\omega|^{1/(N+p)}$, one deduces

$$\int_\omega |u| \, dx \leq C\Big(||u||_p + \sum_{j=1}^N \Big|\Big| \frac{\partial u}{\partial x_j} \Big|\Big|_p \Big) |\omega|^{1/q'}, \tag{A.III.8}$$

i.e.

$$u \in L^{q,\infty} \text{ with } \frac{1}{q} = \frac{1}{N+p} \frac{N}{p}. \tag{A.III.9}$$

Notice that the crude imbedding theorem obtained by this method is never optimal.

The same idea can be applied to obtain a crude imbedding theorem in the case of functions having derivatives in different L^p spaces,

$$u \in L^p(R^N) \text{ and } \frac{\partial u}{\partial x_j} \in L^{p_j}(R^N) \text{ for } j = 1, \ldots, N. \tag{A.III.10}$$

One decomposes $u - \tau_a u$ into a sum of N functions v_j, $j = 1, \ldots, N$, with $v_j(x) = u(x_1, \ldots, x_j, x_{j+1} - a_{j+1}, \ldots, x_N - a_N) - u(x_1, \ldots, x_{j-1}, x_j - a_j, \ldots, x_N - a_N)$ (and obvious changes of notation for $j = 1$ or $j = N$), from which one deduces that

$$u - \rho \star u = \sum_{j=1}^N w_j \text{ with } ||w_j||_{p_j} \leq \Big|\Big| \frac{\partial u}{\partial x_j} \Big|\Big|_{p_j} \int_{R^N} |\rho(y)| \, |y_j| \, dy, \tag{A.III.11}$$

for any function ρ having integral 1. One deduces

$$\int_\omega |u|\,dx \le \int_\omega |\rho\star u|\,dx + \sum_{j=1}^N \int_\omega |w_j|\,dx \le |\omega| ||u||_p ||\rho||_{p'} + \sum_{j=1}^N |\omega|^{1/p'_j} \Big|\Big|\frac{\partial u}{\partial x_j}\Big|\Big|_{p_j} \int_{R^N} |\rho(y)|\,|y_j|\,dy, \quad \text{(A.III.12)}$$

so that for having all the powers of $|\omega|$ equal, one needs to rescale ρ in a different manner than (A.III.1). A natural choice is then to replace ρ by

$$\rho^\varepsilon(x) = \frac{1}{\varepsilon_1\ldots\varepsilon_N}\rho\Big(\frac{x_1}{\varepsilon_1},\ldots,\frac{x_N}{\varepsilon_N}\Big), x\in R^N, \quad \text{(A.III.13)}$$

so that (A.III.12) becomes

$$\int_\omega |u|\,dx \le (\varepsilon_1\ldots\varepsilon_N)^{-1/p}|\omega| ||u||_p ||\rho||_{p'} + \sum_{j=1}^N \varepsilon_j |\omega|^{1/p'_j} \Big|\Big|\frac{\partial u}{\partial x_j}\Big|\Big|_{p_j} \int_{R^N} |\rho(y)|\,|y_j|\,dy. \quad \text{(A.III.14)}$$

Choosing

$$\varepsilon_j = |\omega|^{m_j} \text{ with } m_j = \frac{1}{p_j} - \frac{1}{N+p}\sum_{k=1}^N \frac{1}{p_k}, \quad \text{(A.III.15)}$$

one obtains

$$\int_\omega |u|\,dx \le C\Big(||u||_p + \sum_{j=1}^N \Big|\Big|\frac{\partial u}{\partial x_j}\Big|\Big|_{p_j}\Big)|\omega|^{1/r'}, \quad \text{(A.III.16)}$$

i.e.

$$u\in L^{r,\infty} \text{ with } \frac{1}{r} = \frac{1}{N+p}\sum_{j=1}^N \frac{1}{p_j}. \quad \text{(A.III.17)}$$

I describe quickly now the case related to the Fokker-Planck equation described in Appendix III. Assuming that u satisfies

$$u,\ \frac{\partial u}{\partial t} + \sum_{j=1}^N v_j \frac{\partial u}{\partial x_j}, \frac{\partial u}{\partial v_k} \in L^p(R^N\times R^N\times R),\ k=1,\ldots,N, \quad \text{(A.III.18)}$$

a crude imbedding theorem is obtained by the preceding method by using the flows generated by the first order differential operators of the list in (A.III.18). Like for Lars HÖRMANDER's hypoellipticity result, the key point is that the commutator of $\frac{\partial}{\partial t} + \sum_{j=1}^N v_j \frac{\partial}{\partial x_j}$ and $\frac{\partial}{\partial v_k}$ is $\frac{\partial}{\partial x_k}$. One defines the group of operators S_a by

$$S_a u(x,v,t) = u(x - v\,a, v, t-a) \quad \text{(A.III.19)}$$

so that $\big(\frac{\partial}{\partial a} + \frac{\partial}{\partial t} + \sum_{j=1}^N v_j \frac{\partial}{\partial x_j}\big)S_a u = 0$. This gives the estimate

$$||S_a u - u||_p \le |a| \Big|\Big|\frac{\partial u}{\partial t} + \sum_{j=1}^N v_j \frac{\partial u}{\partial x_j}\Big|\Big|_p. \quad \text{(A.III.20)}$$

For an index k, one defines the group of operators T_b by

$$T_b u(x,v,t) = u(x, v_1,\ldots,v_{k-1}, v_k - b, v_{k+1},\ldots,v_N,t) \quad \text{(A.III.21)}$$

so that $\big(\frac{\partial}{\partial b} + \frac{\partial}{\partial v_k}\big)T_b u = 0$, and this gives the estimate

$$||T_b u - u||_p \le |b| \Big|\Big|\frac{\partial u}{\partial x_k}\Big|\Big|_p. \quad \text{(A.III.22)}$$

The commutation relation quoted above consists in noticing that the operator $T_{-b}S_{-a}T_bS_a$ is given by

$$T_{-b}S_{-a}T_bS_a u(x,v,t) = u(x_1,\ldots,x_{k-1},x_k - a\,b, x_{k+1},\ldots,x_N,v,t). \tag{A.III.23}$$

In order to simplify the notations, the end of the argument is shown for $N = 1$. One can estimate the L^p norm of $u(x,v,t) - u(x - a\,v, v, t - a)$ by $|a|$, the norm of $u(x,v,t) - u(x, v - b, t)$ by $|b|$ and the norm of $u(x,v,t) - u(x - c, v, t)$ by $|c|^{1/2}$, so the norm of $u(x,v,t) - u(x - a\,v - c, v - b, t - a)$ is estimated by $|a| + |b| + |c|^{1/2}$. One decomposes u into $u_0 + u_1$, with u_0 defined by

$$u_0(x,v,t) = \int \rho(a,b,c)u(x - a\,v - c, v - b, t - a)\,da\,db\,dc, \tag{A.III.24}$$

where ρ has integral 1. Then the norm in L^∞ of u_0 is bounded by $||u||_p||\rho||_{p'}$ and the norm in L^p of u_1 is bounded by $C\int|\rho(a,b,c)|(|a| + |b| + |c|^{1/2})\,da\,db\,dc$, and one concludes as before.

Appendix IV.

In some situations involving L^2, one can deduce imbedding theorems using Lorentz spaces for the Fourier transform, and this can be done even if only some fractional derivatives of u are in L^2. If

$$u \in L^2(R^N),\ \Big(\frac{\partial}{\partial x_j}\Big)^{\alpha_j} u \in L^2(R^N),\ j = 1,\ldots,N \tag{A.IV.1}$$

with $\alpha_j > 0$ for $j = 1,\ldots,N$, it is equivalent that the Fourier transform $\mathcal{F}u$ of u belongs to a weighted L^2 space

$$\mathcal{F}u\,W(\xi) \in L^2(R^N),\ \text{with } W(\xi) = \Big(1 + \sum_{j=1}^N |\xi_j|^{\alpha_j}\Big)^{1/2}. \tag{A.IV.2}$$

A direct computation shows that the weight W is such that $\frac{1}{W}$ belongs to the Marcinkiewicz space $L^{q,\infty}$, with $q = \sum_{j=1}^N \frac{1}{\alpha_j}$, and also to L^∞ of course. If $q > 2$ for example, one deduces that $\mathcal{F}u \in L^{r,2}$ and therefore $u \in L^{r',2}$, with $\frac{1}{r} = \frac{1}{q} + \frac{1}{2}$, $\frac{1}{r'} = \frac{1}{2} - \frac{1}{q}$.

This argument is similar in nature to that of Jack PEETRE using convolution by powers of $\frac{1}{r}$, which shows for example that $H^{1/2}(R^2) \subset L^{4,2}(R^2)$, and his argument follows Sergei SOBOLEV's original idea, while the argument of Olga LADYZHENSKAYA which I learned from Jacques-Louis LIONS shows that $H^{1/2}(R^2) \subset L^4(R^2)$ by an argument similar to that of Emilio GAGLIARDO and Louis NIRENBERG.

I have derived another proof, which I find more natural. As was pointed out by Jacques-Louis LIONS, there would not be much to prove if the limiting Sobolev imbedding theorem was true, i.e. if $H^1(R^2)$ was a subset of $L^\infty(R^2)$, because the imbedding theorem for $H^{1/2}$ would then be deduced by a simple interpolation result. I observe then that

$$X = \Big(H^2(R^2), L^2(R^2)\Big)_{1/2,1} \subset L^\infty(R^2), \tag{A.IV.3}$$

and then, using the reiteration theorem of Jacques-Louis LIONS and Jaak PEETRE [Li&Pe],

$$H^{1/2}(R^2) = \Big(X, L^2(R^2)\Big)_{1/2,2} \subset \Big(L^\infty(R^2), L^2(R^2)\Big)_{1/2,2} = L^{4,2}(R^2). \tag{A.IV.4}$$

Of course, (A.IV.3) follows also from another important observation of Jacques-Louis LIONS and Jaak PEETRE [Li&Pe], as it is equivalent to

$$||u||_{L^\infty} \le C||u||_{H^2}^{1/2}||u||_{L^2}^{1/2} \text{ for all } u \in H^2(R^2), \tag{A.IV.5}$$

which in turn follows from the fact that $H^2(R^2) \subset L^\infty(R^2)$ by a scaling argument.

I have then only added a simple argument of scaling to the power of the theory of interpolation of Banach spaces developed by Jacques-Louis LIONS and Jaak PEETRE [Li&Pe], and the same method shows that if $1 \le p < \infty$ and $0 < s < \frac{N}{p}$, then $W^{s,p}(R^N) \subset L^{q,p}(R^N$ with $\frac{1}{q} = \frac{1}{p} - \frac{s}{N}$.

Acknowledgements. Some of the results described here have been obtained while I was working at Commissariat à l'Energie Atomique in Limeil from 1982 to 1987, thanks to the help of Robert Dautray. My research is now supported by Carnegie-Mellon University and the National Science Foundation (grant DMS-97.04762). It is a pleasure to acknowledge the support of the Max Planck Institute for Mathematics in the Sciences in Leipzig for my sabbatical year 1997-98.

References.

[Br&Wa] Brézis H. & Wainger S., "A note on limiting cases of Sobolev imbeddings and convolution inequalities," Comm. Partial Differential Equations, 5 (1980), 773-789.

[Ga] Gagliardo E., "Proprietà di alcune classi di funzioni in più variabili," Ricerche di Mat., Vol. VII (1958), 102-137.

[Ha&Li&Po] Hardy G. H. & Littlewood L. E. & Polya G., *Inequalities*, Cambridge University Press, Cambridge, 1952.

[Jo&Ni] John F. & Nirenberg L., "On functions of bounded mean oscillations," Comm. Pure Appl. Math. 14 (1960), 93-140.

[Li] Lions J.-L., *Problèmes aux limites dans les équations aux dérivées partielles*, Les Presses de l'Université de Montréal, Montréal, 1965.

[Li&Pe] Lions J.-L. & Peetre J., "Sur une classe d'espaces d'interpolation," Inst. Hautes Etudes Sci. Publi. Math. 19 (1964), 5-68.

[O'Ne] O'Neil R., "Convolution operators and $L^{p,q}$ spaces," Duke Math. J. 30 (1963), 129-142.

[Pe] Peetre J., "Espaces d'interpolation et espaces de Soboleff," Ann. Inst. Fourier (Grenoble) 16 (1966), 279-317.

[Sc] Schwartz L., *Théorie des distributions*, Hermann, Paris, 1966.

[So] Sobolev S., "Sur un théorème d'analyse fonctionnelle," Mat. Sb. 46 (1938), 471-497.

[Tro] Troisi M., "Teoremi di inclusione per spazi di Sobolev non isotropi," Ricerche di Matematica XVIII (1969), 3-24.

[Tru] Trudinger N., "On imbeddings into Orlicz spaces and some applications," J. Math. Mech., 17 (1967), 473-483.