Math 212a lecture 5.

We proved that the functions e^{inx} constitute an orthonormal basis of L_2(T) by elementary means. In this lecture we will give a much more complicated proof of this fact. But our proof will extend to much more general situations.

Relation to the operator $\frac{d}{dx}$.

Each of the functions e^{inx} is an eigenvector of the operator

$$D = \frac{d}{dx}$$

in that

$$D\left(e^{inx}\right) = ine^{inx}.$$

So they are also eigenvalues of the operator D^2 with eigenvalues $-n^2$. Also, on the space of twice differentiable periodic functions the operator D^2 satisfies

$$(D^2 f, g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f''(x)\overline{g(x)}dx = f'(x)\overline{g(x)}\Big|_{-\pi}^{\pi} - \frac{1}{2\pi}\int_{-\pi}^{\pi} f'(x)\overline{g'(x)}dx$$

by integration by parts. Since f' and g are assumed to be periodic, the end point terms cancel, and integration by parts once more shows that

$$(D^2 f, g) = (f, D^2 g) = -(f', g').$$

Domain of definition.

But of course D and certainly D^2 is not defined on $\mathcal{C}(\mathbf{T})$ since some of the functions belonging to this space are not differentiable. Furthermore, the eigenvalues of D^2 are tending to infinity rather than to zero. So somehow the operator D^2 must be replaced with something like its inverse. In fact, we will work with the inverse of $D^2 - 1$, but first some preliminaries.

We will let $\mathcal{C}^2([-\pi, \pi])$ denote the functions defined on $[-\pi.\pi]$ and twice differentiable there, with continuous second derivatives up to the boundary. We denote by $\mathcal{C}([-\pi, \pi])$ the space of functions defined on $[-\pi, \pi]$ which are continuous up to the boundary. We can regard $\mathcal{C}(\mathbf{T})$ as the subspace of $\mathcal{C}([-\pi, \pi])$ consisting of those functions which satisfy the boundary conditions $f(\pi) = f(-\pi)$ (and then extended to the whole line by periodicity).

We regard $\mathcal{C}([-\pi,\pi])$ as a pre-Hilbert space with the same scalar product that we have been using:

$$(f,g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx.$$

If we can show that every element of $\mathcal{C}([-\pi,\pi])$ is a sum of its Fourier series (in the pre-Hilbert space sense) then the same will be true for $\mathcal{C}(\mathbf{T})$. So we will work with $\mathcal{C}([-\pi,\pi])$.

We can consider the operator $D^2 - 1$ as a linear map

$$D^2 - 1 : \mathcal{C}^2([-\pi,\pi]) \to \mathcal{C}([-\pi,\pi]).$$

This map is surjective, meaning that given any continuous function g we can find a twice differentiable function f satisfying the differential equation

$$f'' - f = g.$$

$$D^2 - 1 : \mathcal{C}^2([-\pi, \pi]) \to \mathcal{C}([-\pi, \pi]).$$

This map is surjective, meaning that given any continuous function g we can find a twice differentiable function f satisfying the differential equation

$$f'' - f = g.$$

In fact we can find a whole two dimensional family of solutions because we can add any solution of the homogeneous equation

$$h'' - h = 0$$

to f and still obtain a solution. We could write down an explicit solution for the equation $f'' - f = g$, but we will not need to. It is enough for us to know that the solution exists, which follows from the general theory of ordinary differential equations.

The general solution of the homogeneous equation is given by

$$h(x) = ae^x + be^{-x}.$$

The space M .

Let

$$M \subset \mathcal{C}^2([-\pi, \pi])$$

be the subspace consisting of those functions which satisfy the "periodic boundary conditions"

$$f(\pi) = f(-\pi), \quad f'(\pi) = f'(-\pi).$$

Given any f we can always find a solution of the homogeneous equation such that $f - h \in M$. Indeed, we need to choose the complex numbers a and b such that if h is as given above, then

$$h(\pi) - h(-\pi) = f(\pi) - f(-\pi), \text{ and } h'(\pi) - h'(-\pi) = f'(\pi) - f'(-\pi).$$

Collecting coefficients and denoting the right hand side of these equations by c and d we get the linear equations

$$(e^{\pi} - e^{-\pi})(a - b) = c, (e^{\pi} - e^{-\pi})(a + b) = d$$

which has a unique solution.

The operator *T* .

So there exists a unique operator

$$T : \mathcal{C}([-\pi, \pi]) \to M$$

with the property that

$$(D^2 - I) \circ T = I.$$

We will prove that

$$T \quad \text{is self adjoint and compact.} \tag{20}$$

The eigenvectors of *T* .

Once we will have proved this fact, then we know every element of M can be expanded in terms of a series consisting of eigenvectors of T with non-zero eigenvalues. But if

$$Tw = \lambda w$$

then

$$D^2 w = (D^2 - I)w + w = \frac{1}{\lambda}[(D^2 - I) \circ T]w + w = \left(\frac{1}{\lambda} + 1\right) w.$$

So w must be an eigenvector of D^2; it must satisfy

$$w'' = \mu w.$$

The eigenvectors of T, continued.

If $Tw = \lambda w$

then

$$w'' = \mu w.$$

So if $\mu = 0$ then $w =$ a constant is a solution. If $\mu = r^2 > 0$ then w is a linear combination of e^{rx} and e^{-rx} and as we showed above, no non-zero such combination can belong to M. If $\mu = -r^2$ then

the solution is a linear combination of e^{irx} and e^{-irx} and the above argument shows that r must be such that $e^{ir\pi} = e^{-ir\pi}$ so $r = n$ is an integer.

We will prove that

$$T \quad \text{is self adjoint and compact.} \tag{20}$$

Thus (20) will show that the e^{inx} are a basis of M, and a little more work that we will do at the end will show that they are in fact also a basis of $\mathcal{C}([-\pi,\pi])$. But first let us work on (20).

It is easy to see that T is self adjoint. Indeed, let $f = Tu$ and $g = Tv$ so that f and g are in M and

$$(u, Tv) = ([D^2-1]f, g) = -(f', g') - (f, g) = (f, [D^2-1]g) = (Tu, v)$$

where we have used integration by parts and the boundary conditions defining M for the two middle equalities.

Gårding's inequality, special case.

We now turn to the compactness. We have already verified that for any $f \in M$ we have

$$([D^2 - 1]f, f) = -(f', f') - (f, f).$$

Taking absolute values we get

$$\|f'\|^2 + \|f\|^2 \leq |([D^2 - 1]f, f)|. \tag{21}$$

(We actually get equality here, the more general version of this that we will develop later will be an inequality.)

Let $u = [D^2 - 1]f$ and use the Cauchy-Schwarz inequality

$$|([D^2 - 1]f, f)| = |(u, f)| \leq \|u\|\|f\|$$

on the right hand side of (21) to conclude that

$$\|f\|^2 \leq \|u\|\|f\|$$

or

$$\|f\| \leq \|u\|.$$

$$\|f'\|^2 + \|f\|^2 \leq |([D^2 - 1]f, f)|. \qquad (21)$$

$$\|f\| \leq \|u\|.$$

Use (21) again to conclude that

$$\|f'\|^2 \leq \|u\|\|f\| \leq \|u\|^2$$

by the preceding inequality. We have $f = Tu$, and let us now suppose that u lies on the unit sphere i.e. that $\|u\| = 1$. Then we have proved that

$$\|f\| \leq 1, \quad \text{and} \quad \|f'\| \leq 1. \qquad (22)$$

We wish to show that from any sequence of functions satisfying these two conditions we can extract a subsequence which converges. Here convergence means, of course, with respect to the norm given by

$$\|f\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx.$$

$$\|f\| \leq 1, \quad \text{and} \quad \|f'\| \leq 1. \tag{22}$$

In fact, we will prove something stronger: that given any sequence of functions satisfying (22) we can find a subsequence which converges in the uniform norm

$$\|f\|_\infty := \max_{x\in[-\pi,\pi]} |f(x)|.$$

Notice that

$$\|f\| = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx\right)^{\frac{1}{2}} \leq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} (\|f\|_\infty)^2 dx\right)^{\frac{1}{2}} = \|f\|_\infty$$

so convergence in the uniform norm implies convergence in the norm we have been using.

To prove our result, notice that for any $\pi \leq a < b \leq \pi$ we have

$$|f(b) - f(a)| = \left| \int_a^b f'(x) dx \right| \leq \int_a^b |f'(x)| dx = 2\pi(|f'|, \mathbf{1}_{[a,b]})$$

where $\mathbf{1}_{[a,b]}$ is the function which is one on $[a, b]$ and zero elsewhere. Apply Cauchy-Schwarz to conclude that

$$|(|f'|, \mathbf{1}_{[a,b]})| \leq \| \ |f'| \ \| \cdot \|\mathbf{1}_{[a,b]}\|.$$

But

$$\|\mathbf{1}_{[a,b]}\|^2 = \frac{1}{2\pi}|b - a|$$

and

$$\| \ |f'|\| = \|f'\| \leq 1.$$

We conclude that

$$|f(b) - f(a)| \leq (2\pi)^{\frac{1}{2}} |b - a|^{\frac{1}{2}}. \tag{23}$$

$$|f(b) - f(a)| \le (2\pi)^{\frac{1}{2}} |b - a|^{\frac{1}{2}}. \tag{23}$$

In this inequality, let us take b to be a point where $|f|$ takes on its maximum value, so that $|f(b)| = \|f\|_\infty$. Let a be a point where $|f|$ takes on its minimum value. (If necessary interchange the role of a and b to arrange that $a < b$ or observe that the condition $a < b$ was not needed in the above proof.) Then (23) implies that

$$\|f\|_\infty - \min|f| \le (2\pi)^{\frac{1}{2}} |b - a|^{\frac{1}{2}}.$$

But

$$1 \ge \|f\| = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f|^2(x)dx\right)^{\frac{1}{2}} \ge \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} (\min|f|)^2 dx\right)^{\frac{1}{2}} = \min|f|$$

and $|b - a| \le 2\pi$ so

$$\|f\|_\infty \le 1 + 2\pi.$$

Thus the values of all the $f \in T[S]$ are all uniformly bounded - (they take values in a circle of radius $1 + 2\pi$) and they are equicontinuous in that (23) holds. This is enough to guarantee that out of every sequence of such f we can choose a uniformly convergent subsequence.

(We recall how the proof of this goes: Since all the values of all the f are bounded, at any point we can choose a subsequence so that the values of the f at that point converge, and, by passing to a succession of subsequences (and passing to a diagonal), we can arrange that this holds at any countable set of points. In particular, we may choose say the rational points in $[-\pi, \pi]$. Suppose that f_n is this subsequence. We claim that (23) then implies that the f_n form a Cauchy sequence in the uniform norm and hence converge in the uniform norm to some continuous function. Indeed, for any ϵ choose δ such that

$$(2\pi)^{\frac{1}{2}}\delta^{\frac{1}{2}} < \frac{1}{3}\epsilon,$$

choose a finite number of rational points which are within δ distance of any point of $[-\pi, \pi]$ and choose N sufficiently large that $|f_i - f_j| < \frac{1}{3}\epsilon$ at each of these points, r. when i and j are $\geq N$. Then at any $x \in [-\pi, \pi]$

$$|f_i(x) - f_j(x)| \leq |f_i(x) - f_i(r)| + |f_j(x) - f_j(x)| + |f_i(r) - f_j(r)| \leq \epsilon$$

since we can choose r such that that the first two and hence all of the three terms is $\leq \frac{1}{3}\epsilon$.)

The purpose of the next few sections is to provide a vast generalization of the results we obtained for the operator D^2. We will prove the corresponding results for any "elliptic" differential operator (definitions below).

I plan to study differential operators acting on vector bundles over manifolds. But it requires some effort to set things up, and I want to get to the key analytic ideas which are essentially repeated applications of integration by parts. So I will start with elliptic operators L acting on functions on the torus $\mathbf{T} = \mathbf{T}^n$, where there are no boundary terms when we integrate by parts. Then an immediate extension gives the result for elliptic operators on functions on manifolds, and also for boundary value problems such as the Dirichlet problem.

The treatment here rather slavishly follows the treatment by Bers and Schechter in *Partial Differential Equations* by Bers, John and Schechter AMS (1964).

What are currently known as **Sobolev spaces** were first introduced by Hans Lewy in his work on the initial value problem for the wave equation. This work was described by J. Hadamard in an appendix devoted to Lewy's work in Hadamard's well known book on the Cauchy problem published in 1932.

Hans Lewy
1904 - 1988

Sergei Sobolev

Born: **6 Oct 1908 in St Petersburg, Russia**
Died: **3 Jan 1989 in Leningrad (now St Petersburg), Russia**

The Sobolev spaces.

Recall that $\mathbf{T}$ now stands for the n-dimensional torus. Let $\mathbf{P} = \mathbf{P}(\mathbf{T})$ denote the space of trigonometric polynomials. These are functions on the torus of the form

$$u(x) = \sum a_\ell e^{i\ell \cdot x}$$

where

$$\ell = (\ell_1, \ldots, \ell_n)$$

is an n-tuplet of integers and the sum is finite. For each integer t (positive, zero or negative) we introduce the scalar product

$$(u, v)_t := \sum_\ell (1 + \ell \cdot \ell)^t a_\ell \overline{b}_\ell. \tag{24}$$

For $t = 0$ this is the scalar product

$$(u, v)_0 = \frac{1}{(2\pi)^n} \int_{\mathbf{T}} u(x) \overline{v(x)} dx.$$

This differs by a factor of $(2\pi)^{-n}$ from the scalar product that is used by Bers and Schecter. We will denote the norm corresponding to the scalar product $(\ ,\ )_s$ by $\|\ \|_s$.

The Laplacian.

If

$$\Delta := -\left(\frac{\partial^2}{\partial(x^1)^2} + \cdots + \frac{\partial^2}{\partial(x^n)^2}\right)$$

the operator $(1+\Delta)$ satisfies

$$(1+\Delta)u = \sum (1 + \ell \cdot \ell) a_\ell e^{i\ell \cdot x}$$

and so

$$((1+\Delta)^t u, v)_s = (u, (1+\Delta)^t v)_s = (u, v)_{s+t}$$

and

$$\|(1+\Delta)^t u\|_s = \|u\|_{s+2t}. \tag{25}$$

The generalized Cauchy-Schwarz inequality.

$$\|(1+\Delta)^t u\|_s = \|u\|_{s+2t}. \tag{25}$$

We then get the "generalized Cauchy-Schwarz inequality"

$$|(u,v)_s| \leq \|u\|_{s+t}\|v\|_{s-t} \tag{26}$$

for any t, as a consequence of the usual Cauchy-Schwarz inequality. Indeed,

$$\begin{aligned}
\sum_\ell (1+\ell\cdot\ell)^s a_\ell \overline{b}_\ell &= \sum_\ell (1+\ell\cdot\ell)^{\frac{s+t}{2}} a_\ell (1+\ell\cdot\ell)^{\frac{s-t}{2}} \overline{b}_\ell \\
&= ((1+\Delta)^{\frac{s+t}{2}} u, (1+\Delta)^{\frac{s-t}{2}} v)_0 \\
&\leq \|(1+\Delta)^{\frac{s+t}{2}} u\|_0 \|(1+\Delta)^{\frac{s-t}{2}} v\|_0 \\
&= \|u\|_{s+t}\|v\|_{s-t}.
\end{aligned}$$

The generalized Cauchy-Schwarz inequality reduces to the usual Cauchy-Schwarz inequality when $t=0$.

Clearly we have

$$\|u\|_s \leq \|u\|_t \quad \text{if } s \leq t.$$

If D^p denotes a partial derivative,

$$D^p = \frac{\partial^{|p|}}{\partial(x^1)^{p_1} \cdots \partial(x^n)^{p_m}}$$

then

$$D^p u = \sum (i\ell)^p a_\ell e^{i\ell \cdot x}.$$

In these equations we are using the following notations:

- If $p = (p_1, \ldots, p_n)$ is a vector with non-negative integer entries we set

$$|p| := p_1 + \cdots + p_n.$$

- If $\xi = (\xi_1, \ldots, \xi_n)$ is a (row) vector we set

$$\xi^p := \xi_1^{p_1} \cdot \xi_2^{p_2} \cdots \xi_n^{p_n}$$

It is then clear that

$$\|D^p u\|_t \leq \|u\|_{t+|p|} \tag{27}$$

and similarly

$$\|u\|_t \leq (\text{constant depending on } t) \sum_{|p| \leq t} \|D^p u\|_0 \quad \text{if } t \geq 0. \tag{28}$$

In particular,

Proposition 2 *The norms*

$$u \mapsto \|u\|_t$$

$t \geq 0$ *and*

$$u \mapsto \sum_{|p| \leq t} \|D^p u\|_0$$

are equivalent.

We let $\mathbf{H}_t$ denote the completion of the space $\mathbf{P}$ with respect to the norm $\| \ \|_t$. Each $\mathbf{H}_t$ is a Hilbert space, and we have natural embeddings

$$\mathbf{H}_t \hookrightarrow \mathbf{H}_s \ \text{ if } s < t.$$

$$\|(1+\Delta)^t u\|_s = \|u\|_{s+2t}. \tag{25}$$

Equation (25) says that

$$(1+\Delta)^t : \mathbf{H}_{s+2t} \to \mathbf{H}_s$$

and is an isometry.

We have proved the "generalized Cauchy-Schwarz inequality"

$$|(u,v)_s| \leq \|u\|_{s+t}\|v\|_{s-t} \tag{26}$$

From the generalized Schwarz inequality we also have a natural pairing of $\mathbf{H}_t$ with $\mathbf{H}_{-t}$ given by the extension of $(\ ,\ )_0$, so

$$|(u,v)_0| \leq \|u\|_t\|v\|_{-t}. \tag{29}$$

In fact, this pairing allows us to identify $\mathbf{H}_{-t}$ with the space of continuous linear functions on $\mathbf{H}_t$. Indeed, if ϕ is a continuous linear function on $\mathbf{H}_t$ the Riesz representation theorem tells us that there is a $w \in \mathbf{H}_t$ such that $\phi(u) = (u,w)_t$. Set

$$v := (1+\Delta)^t w.$$

Then

$$v \in \mathbf{H}_{-t}$$

and

$$(u,v)_0 = (u,(1+\Delta)^t w)_0 = (u,w)_t = \phi(u).$$

We record this fact as

$$\mathbf{H}_{-t} = (\mathbf{H}_t)^*. \tag{30}$$

$$\mathbf{H}_{-t} = (\mathbf{H}_t)^* . \tag{30}$$

As an illustration of (30), observe that the series

$$\sum_\ell (1 + \ell \cdot \ell)^s$$

converges for

$$s < -\frac{n}{2}.$$

This means that if define v by taking

$$b_\ell \equiv 1$$

then $v \in \mathbf{H}_s$ for $s < -\frac{n}{2}$. If u is given by $u(x) = \sum_\ell a_\ell e^{i\ell \cdot x}$ is any trigonometric polynomial, then

$$(u, v)_0 = \sum a_\ell = u(0).$$

So the natural pairing (29) allows us to extend the linear function sending $u \mapsto u(0)$ to all of $\mathbf{H}_t$ if $t > \frac{n}{2}$. We can now give v its "true name": it is the Dirac "delta function" δ (on the torus) where

$$(u, \delta)_0 = u(0).$$

This means that if define v by taking

$$b_\ell \equiv 1$$

then $v \in \mathbf{H}_s$ for $s < -\frac{n}{2}$. If u is given by $u(x) = \sum_\ell a_\ell e^{i\ell \cdot x}$ is any trigonometric polynomial, then

$$(u, v)_0 = \sum a_\ell = u(0).$$

So the natural pairing (29) allows us to extend the linear function sending $u \mapsto u(0)$ to all of $\mathbf{H}_t$ if $t > \frac{n}{2}$. We can now give v its "true name": it is the Dirac "delta function" δ (on the torus) where

$$(u, \delta)_0 = u(0).$$

So $\delta \in H_{-t}$ for $t > \frac{n}{2}$, and the preceding equation is usually written symbolically as

$$\frac{1}{(2\pi)^n} \int_{\mathbf{T}} u(x)\delta(x)dx = u(0);$$

but the true mathematical interpretation is as given above.

Sobolev's Lemma.

The space $\mathbf{H}_0$ is just $L_2(\mathbf{T})$, and we can think of the space $\mathbf{H}_t$, $t > 0$ as consisting of those functions having "generalized L_2 derivatives up to order t". Certainly a function of class C^t belongs to $\mathbf{H}_t$. With a loss of degree of differentiability the converse is true:

Lemma 1 [Sobolev.] *If $u \in \mathbf{H}_t$ and*

$$t \geq \left[\frac{n}{2}\right] + k + 1$$

then $u \in C^k(\mathbf{T})$ and

$$\sup_{x \in \mathbf{T}} |D^p u(x)| \leq \text{const.}\|u\|_t \quad \textit{for } |p| \leq k. \tag{31}$$

Proof of Sobolev's Lemma.

By applying the lemma to $D^p u$ it is enough to prove the lemma for $k = 0$. So we assume that $u \in \mathbf{H}_t$ with $t \geq [n/2] + 1$. Then

$$(\sum |a_\ell|)^2 \leq \left(\sum (1 + \ell \cdot \ell)^t |a_\ell|^2\right) \sum (1 + \ell \cdot \ell)^{-t} < \infty,$$

since the series $\sum (1 + \ell \cdot \ell)^{-t}$ converges for $t \geq [n/2] + 1$. So for this range of t, the Fourier series for u converges absolutely and uniformly. The right hand side of the above inequality gives the desired bound. QED

Distributions aka generalized functions.

A **distribution** on $\mathbf{T}^n$ is a linear function T on $C^\infty(\mathbf{T}^n)$ with the continuity condition that

$$\langle T, \phi_k \rangle \to 0$$

whenever

$$D^p \phi_k \to 0$$

uniformly for each fixed p. If $u \in \mathbf{H}_{-t}$ we may define

$$\langle u, \phi \rangle := (\phi, \overline{u})_0$$

and since $C^\infty(\mathbf{T})$ is dense in $\mathbf{H}_t$ we may conclude

Schwartz's theorem.

Lemma 2 $\mathbf{H}_{-t}$ *is the space of those distributions* T *which are continuous in the* $\| \ \|_t$ *norm, i.e. which satisfy*

$$\|\phi_k\|_t \to 0 \quad \Rightarrow \quad \langle T, \phi_k \rangle \to 0.$$

Theorem 4 [Laurent Schwartz.] $\mathbf{H}_{-\infty}$ *is the space of all distributions. In other words, any distribution belongs to* $\mathbf{H}_{-t}$ *for some* t.

Proof. Suppose that T is a distribution that does not belong to any $\mathbf{H}_{-t}$. This means that for any $k > 0$ we can find a C^∞ function ϕ_k with

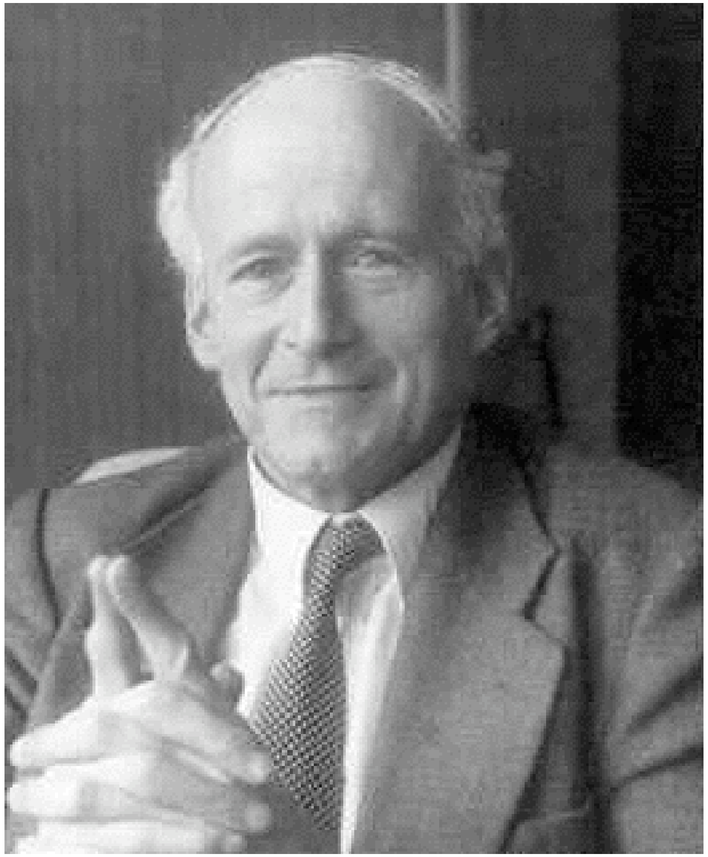
$$\|\phi_k\|_k < \frac{1}{k}$$

and

$$|\langle T, \phi_k \rangle| \geq 1.$$

But by Lemma 1 we know that $\|\phi_k\|_k < \frac{1}{k}$ implies that $D^p \phi_k \to 0$ uniformly for any fixed p contradicting the continuity property of T. QED

Laurent Schwartz



Born: **5 March 1915 in Paris, France**
Died: **4 July 2002**