The material here can be found in Hörmander's Volume 1, Chapter VII – but he has already done almost all of distribution theory by this point(!) – Joshi and Friedlander Chapter 8.

- Recall that $\mathcal{S}(\mathbb{R}^n)$ is a complete metric space.
- We know that convergence with respect to this metric represents convergence with respect to each of the underlying norms

$$\|\phi\|_N = \sup_{x\in\mathbb{R}^n,\ |\alpha|\le N} |(1+|x|)^N D_x^\alpha \phi|; \tag{1}$$

  note that these norms increase with N.
- Consider the same sort of thing for continuity of maps

Lemma 1. (1) *A linear map $A : \mathcal{S}(\mathbb{R}^n) \longrightarrow B$ to a normed space is continuous iff and only if there exist C and N such that*

$$\|A\phi\|_B \le C\|\phi\|_N \ \forall\ \phi \in \mathcal{S}(\mathbb{R}^n). \tag{2}$$

(2) *A linear map $P : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^m)$ is continous if and only if for each N' there exists $C = C(N')$ and $N = N(N')$ such that*

$$\|P\phi\|_{N'} \le C\|\phi\|_N \ \forall\ \phi \in \mathcal{S}(\mathbb{R}^n). \tag{3}$$

(3) *A bilinear map $B : \mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^m) \longrightarrow B$ is continuous (for the product topology) if and only if there exist C, N and M such that*

$$\|B(\phi,\psi)\|_B \le C\|\phi\|_N\|\psi\|_M. \tag{4}$$

(4) *A bilinear map $G : \mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^m) \longrightarrow \mathcal{S}(\mathbb{R}^k)$ is continuous if and only if for each N' there exist C, N and M such that*

$$\|G(\phi,\psi)\|_{N'} \le C\|\phi\|_N\|\psi\|_M \tag{5}$$

Proof. Let me quickly go through the proof for B, the others are similar. First check that (4) implies continuity – we are in a metric space setting so sequential continuity is the same thing. If $(\phi_n,\psi_n) \to (\phi,\psi)$ in the product metric space then $\phi_n \to \phi$ and $\psi_n \to \psi$ (and conversely). Moreover, convergence in $\mathcal{S}(\mathbb{R}^n)$ implies (is equivalent to) convergence in *each norm* $\|\cdot\|_N$ since if $\epsilon > 0$ is small, $\epsilon < 1$, choosing $\delta = 2^{-N-1}\epsilon$, then

$$d(\phi_n,\phi) < \delta \Longrightarrow \|\phi_n - \phi\|_N \le \frac{1}{2}\epsilon(1 + \|\phi_n - \phi\|_N) \Longrightarrow \|\phi_n - \phi\|_N \le \epsilon.$$

In the other direction, use the topological definition. Continuity at 0 (which is equivalent to continuity everywhere for a bilinear form) implies that there exists $\delta > 0$ such that

$$d(\psi,0) < \delta,\ d(\phi,0) < \delta \Longrightarrow \|B(\phi,\psi)\|_B < 1.$$

Since $2^{-k+1} < \delta$ for some k we can arrange the inequalities on the left by demanding

$$\|\psi|_k < \frac{1}{4}\delta \Longrightarrow \sum_N 2^{-N}\frac{\|\psi\|_N}{1+\|\psi\|_N} < \frac{1}{2}\delta + \frac{1}{2}\delta \tag{6}$$

splitting the sum in two at k. By scaling using the bilinearity it follows that

$$\|B(\phi,\psi)\|_B \le C\|\phi\| + k\|\psi\|_k,\ C = 8\delta^{-2}. \tag{7}$$

□

- So continuity for Fréchet spaces (or countably normed ones for that matter) is just like continuity on normed spaces with extra qualifiers!
- (Didn't do this) Even though we do not need to use it at the moment I want to recall – because it is significant later – one thing completelness does for us:

Proposition 1. *A bilinear map $B : \mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^m) \longrightarrow \mathbb{C}$ (or more generally as above) is continuous if and only if it is* separately continuous *– the maps $B(\cdot, \psi) : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathbb{C}$ and $B(\phi, \cdot) : \mathcal{S}(\mathbb{R}^m) \longrightarrow \mathbb{C}$ are continous for each fixed ϕ and ψ.*

Proof. I will not do it in class. The point is Baire's Theorem (what used to be called Baire Catergory argument) just like the uniform boundedness principle. For each N look at the set

$$D(N) = \{\phi \in \mathcal{S}(\mathbb{R}^n); |B(\phi, \psi)| \le N\|\phi\|_N\|\psi\|_N \ \forall \ \psi \in \mathcal{S}(\mathbb{R}^m)\}. \tag{8}$$

Continuity in the first variable shows this is closed and continuity in the second shows that the $D(N)$ cover $\mathcal{S}(\mathbb{R}^n)$. Baire's Theorem then shows that one at least of the $D(N)$ has non-empty interior and translating and scaling around an interior point gives the continuity estimate. □

- Now, we know what tempered distributions are, they are the continuous linear maps $u : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathbb{C}$, such that for some N,

$$C = \sup_{\|\phi\|_N = 1} |u(\phi)| < \infty. \tag{9}$$

We denote the linear space – the dual space – as $\mathcal{S}'(\mathbb{R}^n)$. They are called 'tempered' or 'temperate' distributions (meaning in some sense they have polynomial bounds but beware of this).
- The first point is that there is an injection

$$I : L^2(\mathbb{R}^n) \longrightarrow \mathcal{S}'(\mathbb{R}^n) \tag{10}$$

which is actually the transpose of the inclusion $\mathcal{S}(\mathbb{R}^n) \longrightarrow L^2(\mathbb{R}^n)$. The latter is continuous as follows from

$$\|\phi\|_{L^2} \le \sup\left((1+|x|)^{\frac{1}{2}(n+1)}|\phi(x)|\right)\left(\int_{\mathbb{R}^n}(1+|x|)^{-n-1}dx\right)^{\frac{1}{2}} \tag{11}$$

since this integral is finite. Another way of putting this is that the L^2 norm is continuous on $\mathcal{S}(\mathbb{R}^n)$

$$\|\phi\|_{L^2} \le C\|\phi\|_k, \ k > \frac{1}{2}n.$$

Thus if $u \in L^2(\mathbb{R}^n)$ we *define*

$$I(u) \in \mathcal{S}'(\mathbb{R}^n), \ I(u)(\phi) = \int_{\mathbb{R}^n} u(x)\phi(x). \tag{12}$$

Not only is this well-defined but the map I is an injection. In fact we will quite soon drop the I from the notation altogether and regard this map as an identification.
- **Lemma 2.** *The map I in* (11), (12) *is an injection.*

Proof. I will give a direct proof a bit later on using convolution. However, you probably know that in one dimension the eigenfunctions of the harmonic oscillator give an orthonormal basis of $L^2(\mathbb{R})$. They are all the products of polynomials and a gaussian, so are in $\mathcal{S}(\mathbb{R})$ which is therefore dense. Moreover, their products in n variable give an orthonormal basis of $L^2(\mathbb{R}^n)$ proving the density in higher dimensions. $\square$

- You should go through a similar argument to see that $L^1(\mathbb{R}^n) \longrightarrow \mathcal{S}'(\mathbb{R}^n)$ is injective using the same formula (12) – maybe leaving density of $\mathcal{S}(\mathbb{R}^n)$ in $L^1(\mathbb{R}^n)$ until later!
- Now we can specialize to

$$I : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}'(\mathbb{R}^n),\ I(\psi)(\phi) = \int \psi(x)\phi(x) \tag{13}$$

and this is really the heart of the matter.

Lemma 3. *If $\psi \in \mathcal{S}(\mathbb{R}^n)$ then for all $\phi \in \mathcal{S}(\mathbb{R}^n)$*

$$\begin{gathered} I(x^\alpha\psi)(\phi) = I(\psi)(x^\alpha\phi), \\ I(\partial_x^\beta\psi)(\phi) = I(\psi)((-1)^{|\alpha|}\partial_x^\beta\phi), \\ I(\mu\psi)(\phi) = I(\psi)(\mu\phi),\ \mu \in \mathcal{S}(\mathbb{R}^n). \end{gathered} \tag{14}$$

Proof. This is just manipulation under the integral, in the second case involving integration by parts $|\alpha|$ times. In fact it suffices to take $\alpha = e_j$ and iterate and in that case

$$\int_{\mathbb{R}^n} (\partial_j\psi)\phi = \lim_{R\to\infty} \int_{[-R,R]^n} (\partial_j\psi)\phi \tag{15}$$

$$= -\lim_{R\to\infty} \int_{[-R,R]^n} \psi(\partial_j\phi) + \int_{[-R,R]^{n-1}} \left(\psi\phi|_{x_j=R} - \psi\phi|_{x_j=-R}\right) = \int_{\mathbb{R}^n} \psi(-\partial_j\phi)$$

since the integrand, and hence the $n-1$ fold integral, vanishes rapidly at infinity. $\square$

- The second of these identities is what is called the 'weak formulation of differentiation'. We certainly know the map (13) is injective (because we can evaluate at $\phi = \overline{\psi}$) and then $\partial^\beta\psi \in \mathcal{S}'(\mathbb{R}^n)$ is the unique point in the image of $\mathcal{S}(\mathbb{R}^n)$ which satisfies this identity – the right hand side determines $\partial^\beta\psi$ 'as a distibution'.

  This is the fundamental point of distribution theory. We can use the identities in (14) as *definitions.*

Definition 1. If $u \in \mathcal{S}'(\mathbb{R}^n)$ then there are uniquely defined elements $\partial^\beta u$, $x^\alpha u$ and μu defined by the identities

$$\begin{gathered} (x^\alpha u)(\phi) = u(x^\alpha\phi), \\ (\partial_x^\beta u)(\phi) = u((-1)^{|\alpha|}\partial_x^\beta\phi), \\ (\mu u)(\phi) = u(\mu\phi),\ \mu \in \mathcal{S}(\mathbb{R}^n) \end{gathered} \tag{16}$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$.

Of course we need to note that these are indeed distributions, however this follows from the fact that the maps

$$\begin{gathered} \mathcal{S}(\mathbb{R}^n) \ni \phi \longmapsto x^\alpha \phi, \\ \mathcal{S}(\mathbb{R}^n) \ni \phi \longmapsto \partial^\beta \phi, \\ \mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n) \ni (\mu, \phi) \longrightarrow \mu\phi \in \mathcal{S}(\mathbb{R}^n) \end{gathered} \tag{17}$$

are all continuous – and this follows readily especially if you do this week's homework.

- The important point here is that this definition is consistent with the 'pointwise' notions:

$$\begin{aligned} x^\alpha I(\psi) &= I(x^\alpha \psi) \\ \partial_x^\beta I(\psi) &= I(\partial^\beta \psi), \\ \mu I(\psi) &= I(\mu\psi) \end{aligned} \tag{18}$$

where on the left we use the distributional notions and on the right the 'classical' ones. This is why we can safely drop the 'I' a little bit later.

- So, we have now defined the differentiation of an arbitrary tempered distribution. For instance the derivatives of an element of L^2 are well-defined, they cannot in general be functions but they are distributions. We need to improve on the embedding I to show that if a function has 'classical' derivatives in an appropriate sense then these are equal to its distributional derivatives – so far this has only been shown for elements of $\mathcal{S}(\mathbb{R}^n)$.

On Thursday I will prove the Fourier inversion formula and then come back to properties of distributions.

- We define the Fourier transform of a function $u \in L^1(\mathbb{R}^n)$ as

$$\mathcal{F}(u)(\xi) = \hat{u}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x) dx. \tag{19}$$

The boundedness and continuity of the oscillating exponential means that this exists as a Lebesgue integral for each $\xi \in \mathbb{R}^n$ and defines a bounded function which is continuous and vanishes at infinity

$$\mathcal{F} : L^1(\mathbb{R}) \longrightarrow \{\hat{u} \in \mathcal{C}^0(\mathbb{R}^n); \sup_{|x| \le R} |u(x)| \to 0 \text{ as } R \to \infty\}. \tag{20}$$

The continuity of $\hat{u}$ follows from continuity-in-the-mean of L^1 functions – that $\lim_{|t|\to 0} \int |f(x+t) - f(x)| dx = 0$ – and the vanishing at infinity from a density argument below.

- For the moment we are interested rather in showing that

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n) \tag{21}$$

is continuous, then that it is in fact an isomorphism with a continuous inverse given by 'the Fourier inversion formula'

$$u = \mathcal{G}\hat{u}, \ \mathcal{G}(v)(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} v(\xi) d\xi. \tag{22}$$

- First then, continuity of (21). We know that $\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ since $(1 + |x|)^{-n-1} \in L^1(\mathbb{R}^)$ and

$$|u(x)| \le \|u\|_{n+1} (1 + |x|)^{-n-1}. \tag{23}$$

- If we formally differentiate through the integral we find that

$$\partial_{\xi_j}\hat{u}(\xi) = \int \partial_{\xi_j} e^{ix\cdot\xi} u(x) = \mathcal{F}(ix_j u)$$

and since $x_j u \in \mathcal{S}(\mathbb{R}^n)$ the function on the right exists and is at least bounded. So, how to justify the 'exchange of limits' involved in differentiating under the integral? We can appeal to standard theorems (either for the Riemann integral over big rectangles or directly for the Lebesgue integral) or we can just do it. Namely, look at the difference quotient

$$\frac{\hat{u}(\xi + se_j) - u(x)}{s} = \int \frac{e^{ix\cdot(\xi+se_j)} - e^{ix\cdot\xi}}{s} u(x) = \int \frac{e^{isx_j} - 1}{s} e^{ix\cdot\xi} u(x). \tag{24}$$

Taylor's formula with (Legendre's?) remainder or an appropriate application of the Fundamental Theorem of Calculus gives

$$e^{isx_j} - 1 - isx_j = \int_0^1 \int_0^r \frac{d^2}{dt^2} e^{istx_j} dt dr$$

and the integrand of the RHS is $-s^2 x_j^2 e^{istx_j}$ so the integral is globally (in x) bounded by $|s|^2|x|^2$ and hence

$$|\frac{e^{isx_j} - 1 - isx_j}{s}| \le s|x|^2. \tag{25}$$

Since $|x|^2 u(x) \in L^1$ we can pass to the limit and justify

$$\partial_{\xi_j}\hat{u}(\xi) = \mathcal{F}(ix_j u) \ \forall \ u \in \mathcal{S}(\mathbb{R}^n) \tag{26}$$

meaning that the Fouier transform has partial derivatives, they are given by this formula and hence are globally bounded. Now we can iterate and conclude that derivatives of all orders exist, are continuous and are all bounded:-

$$u \in \mathcal{S}(\mathbb{R}^n) \Longrightarrow \hat{u} \in \mathcal{C}^\infty(\mathbb{R}^n),\ \partial_\xi^\alpha \hat{u} = \mathcal{F}(i^\alpha x^\alpha u). \tag{27}$$

- Next we need to do the same thing for $\xi^\beta \hat{u}$ but that is quite a lot easier since it just involves integration by parts.
- So now we can see that $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(\mathbb{R}^n)$ is a continuous linear map.