EXPOSITORY NOTES ON DISTRIBUTION THEORY AND OTHER TOPICS, FALL 2018

While these notes are under construction, I expect there will be many typos.

The main reference for this is volume 1 of Hörmander, The analysis of liner partial differential equations. I have picked a few of the most useful and concrete highlights. The references are based on the 1989 hardcover second edition.

1. Generalities (from Ch. 2 and 3)

Definition 1.1. Let U be an open set in $\mathbb{R}^n$. A distribution $u \in \mathcal{D}'(U)$ is a linear function $u : C_0^\infty(U) \to \mathbb{C}$. One can write $u(\phi) =< u, \phi >$ and think of this, informally, as $u(\phi) = \int u\phi$. It is required that u is continuous in the following sense:

For every $K \subset U$ compact there exist C, k such that

$$|u(\phi)| = | < u, \phi > | \leq C \sum_{|\alpha| \leq k} \sup_x |\partial^\alpha \phi| \tag{1}$$

for every $\phi \in C_0^\infty(U)$ supported in K.

If one k works for all K, u is of finite order. The smallest such k is the order of u.

We will need an equivalent formulation of the continuity condition.

Definition 1.2. Let $\phi_j, \phi \in C_0^\infty(U)$. The sequence $\phi_j \to \phi$ in $C_0^\infty(U)$ if there exists a compact subset of U which contains the support of all ϕ_j, ϕ and for every fixed α, $\sup_x |\partial^\alpha (\phi_j(x) - \phi(x))| \to 0$ as $j \to \infty$.

Theorem 1.3. *A linear function $u : C_0^\infty(U) \to \mathbb{C}$ is a distribution if and only if $u(\phi_j) \to u(\phi)$ for every $\phi_j \to \phi$ in $C_0^\infty(U)$.*

Proof. To show that if u is a distribution, then $u(\phi_j) \to u(\phi)$ for every $\phi_j \to \phi$ in $C_0^\infty(U)$ is clear from the definition. The other half is an easy exercise in negations. □

Examples:

(1) If $\tilde{u}$ is a locally integrable function, $u(\phi) := \int \tilde{u}\phi$. This identifies the function $\tilde{u}$ with a distribution u.
(2) Dirac delta function. $\delta_a(\phi) = \phi(a)$

(3) Weak derivatives: If u is a smooth function, and $\phi \in C_0^\infty$ is a test function, $\alpha = (\alpha_1, \cdots, \alpha_n)$ is a multi-index and $\partial^\alpha u = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} u$, then
$< \partial^\alpha u, \phi >:= \int \partial^\alpha u \phi = (-1)^{|\alpha|} \int u \partial^\alpha \phi = (-1)^{|\alpha|} < u, \partial^\alpha \phi >$. (integration by parts). This motivates the **definition** of $\partial^\alpha u$ for any distribution u: $< \partial^\alpha u, \phi >= (-1)^{|\alpha|} < u, \partial^\alpha \phi >$.

(4) It takes some work (thm. 4.4.7 in Hörmander) and we will not prove this, but the above essentially accounts for all possible distributions:

If $u \in \mathcal{D}'(U)$ then there exists a locally finite family of continuous functions f_α (each compact subset of U intersects only finitely many of the supports of the f_αs) such that

$$u = \sum_\alpha \partial^\alpha f_\alpha$$

Definition 1.4. A sequence of distributions u_i converges to u in $\mathcal{D}'(U)$ (or in the sense of distribution theory) if $u_i(\phi) \to u(\phi)$ for every $\phi \in C_0^\infty(U)$

Also, if $u_i \in \mathcal{D}'(U)$ and for each fixed $\phi \in C_0^\infty(U)$ the limit $u_i(\phi)$ exists and is denoted $u(\phi)$, then u is automatically a distribution. See Theorem 2.1.8. We will not prove this.

Definition 1.5. Let $u \in \mathcal{D}(U)$ and $f \in C^\infty(U)$. Then the distributions $\frac{\partial u}{\partial x_k}$ and fu are defined by

$$\left(\frac{\partial u}{\partial x_k}\right)(\phi) = -u\left(\frac{\partial}{\partial x_k}\right)$$

$$(fu)(\phi) = u(f\phi)$$

Unlike classical convergence, if $u_i \to u$ in $\mathcal{D}'(U)$, then $\partial^\alpha u_i \to \partial^\alpha u$ in $\mathcal{D}'(U)$ is trivial.

Example 1: Let H be the Heavyside function. Then $H' = \delta_0$.

The following will be worked out in class:

If E is the fundamental solution of the Laplace operator, ∇E in the sense of distributions agrees with the locally integrable function ∇E defined for $x \neq 0$, but ΔE in the sense of distributions does not agree with the locally integrable function $\Delta E = 0$ defined for $x \neq 0$. In fact $\Delta E = \delta_0$.

Definition 1.6. A distribution u is defined to be 0 in an open set $V \subset U$ if $u(\phi) = 0$ for every $\phi \in C_0^\infty(V)$. The union of all such subsets V is the largest open set where u is 0, and the complement of that is defined to be the support of u.

Thus the support of a distribution $u \in \mathcal{D}(U)$ is always (relatively) closed in U. If the support of u is compact, u is called compactly supported. The set of compactly supported distributions in U is denoted by $\mathcal{E}'(U)$

Recall the support of a function ϕ is the closure of the set $\{\phi(x) \neq 0\}$. If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\phi \in C_0^\infty(\mathbb{R}^n)$, and the support of ϕ and u are disjoint, then $u(\phi) = 0$. However, if ϕ is zero on the support of u, it does not follow that $u(\phi) = 0$. Example: $\delta'(x)$.

If $u \in \mathcal{E}'(U)$, $u(\phi)$ is well defined for $\phi \in C^\infty$: Let K be the support of u, $K \subset V \subset U$ with V open. There exists a smooth cut-off function $\zeta \in C_0^\infty(U)$, and $\zeta = 1$ in V. Then $u(\zeta\phi)$ is well-defined, and is independent of the choice of ζ. Define $u(\phi) = u(\zeta\phi)$ for ζ as above.

Definition 1.7. A distribution u is defined to be smooth in an open set $V \subset U$ if there exists $\tilde{u} \in C^\infty(V)$ such that $u(\phi) = \int \tilde{u}(x)\phi(x)dx$ for all $\phi \in C_0^\infty(V)$ The union of all such subsets V is the largest open set where u is smooth, and the complement of that is defined to be the singular support of u.

2. Distributions supported at one point

Theorem 2.1. *If $u \in \mathcal{D}'(\mathbb{R}^n)$ is supported at a point, say 0, then u is a finite linear combination*

$$u = \sum c_\alpha \partial^\alpha \delta$$

Proof. Assume u is of order k (and prove: any compactly supported distribution is of finite order). Pick a test function ϕ and write $\phi(x) = T(x) + R(x)$ the kth order Taylor polynomial plus remainder. $u(T)$ is what we want (check!), and the point is to show that $u(R) = 0$ where R is the remainder. We know $|R(x)| \leq C|x|^{k+1}$ for $|x| \leq 1$ and in fact $|\partial^\alpha R(x)| \leq C|x|^{k+1-|\alpha|}$ for all $|\alpha| \leq k$. Let $\epsilon > 0$, and let χ be a cut-off function, identically 1 in a neighborhood of 0.
Then $|u(R)| = |u(\chi\left(\frac{x}{\epsilon}\right)R\big) \leq C\sum_{|\alpha|\leq k} \sup_x |\partial^\alpha(\chi(\frac{x}{\epsilon})R)| \leq C\epsilon$. Now let $\epsilon \to 0$. ∎

Application to PDE: Let $E = \frac{1}{|x|^{n-2}}$ $(n \geq 3)$. Then $\Delta E = 0$ for x away from 0 by calculation, thus ΔE is a distribution supported at 0. It is a finite linear combination of the delta function and its derivatives. An additional homogeneity argument shows $\Delta E = c\delta$.

If u is a locally integrable function in $\mathbb{R}^n - \{0\}$, u is homogeneous of degree α if $u(tx) = t^\alpha u(x)$ for all $t > 0$ and $x \neq 0$. Denoting

$\phi_t(x) = t^n\phi(tx)$ this is equivalent to

$$\int u\phi = t^\alpha \int u\phi_t$$

and the definition of a homogeneous distribution in $\mathbb{R}^n$ (or $\mathbb{R}^n - \{0\}$) is

$$u(\phi) = t^\alpha u(\phi_t)$$

for every $\phi \in C_0^\infty(\mathbb{R}^n)$ or $C_0^\infty(\mathbb{R}^n - \{0\})$.

3. Convolutions (Chapter 4 in Hörmander's book)

The major goal of this section is to prove

1) If $f \in \mathcal{E}'(\mathbb{R}^n)$, then there exists $u \in \mathcal{D}'(\mathbb{R}^n)$ such that $\Delta u = f$.

2) If u, f are as above, and $f \in C^\infty(V)$ for some open set V, then $u \in C^\infty(V)$.

Both of these goals follow from the properties of the convolution of a distribution with a compactly supported distribution. Part 1 follows by writing $u = E * f$, $\Delta u = (\Delta E) * f = \delta * f = f$, but we have to assign rigorous meaning to this. Part 2 follows from the fact that the fundamental solution E is C^∞ away from 0. The exact same results hold for $\frac{\partial}{\partial t} - \Delta$ but not $\frac{\partial^2}{\partial t^2} - \Delta$.

Definition 3.1. If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\phi \in C_0^\infty(\mathbb{R}^n)$,

$u * \phi(x) =< u, \phi(x - \cdot) >$ (where $\cdot$ stands for y, and u acts in the y variable)

Check $u * \phi \in C^\infty$, $\partial^\alpha(u * \phi)(x) = (\partial^\alpha u) * \phi = u * (\partial^\alpha \phi)(x)$: We have

$$\phi(x - y + \epsilon e_i) - \phi(x - y) = \epsilon \frac{\partial}{\partial x_i}\phi(x - y) + R(x - y, \epsilon)$$

where

$$R(x - y, \epsilon) = \int_0^1 \frac{d^2}{dt^2}\left(\phi(x - y + t\epsilon e_i)\right)(1 - t)dt$$
$$= \epsilon^2 \int_0^1 \left(\frac{\partial^2 \phi}{\partial x_i^2}\right)(x - y + t\epsilon e_i)(1 - t)dt$$

Fix x. $R(x - y, \epsilon)$ is in C_0^∞, and $\sup_y |\partial_y^\alpha R(x - y, \epsilon)| \le C_\alpha \epsilon^2$. Using the continuity condition (1) we see

$$\lim_{\epsilon \to 0} \langle u, \frac{R(x - \cdot, \epsilon)}{\epsilon} \rangle = 0$$

and

$$\lim_{\epsilon\to 0}\frac{u\left(\phi(x-\cdot+\epsilon e_i)\right)-u\left(\phi(x-\cdot)\right)}{\epsilon}=u(\frac{\partial}{\partial x_i}\phi(x-\cdot))=\frac{\partial u}{\partial x_i}(\phi(x-\cdot)$$

Check *support*$(u*\phi)\subset$ *support* u + *support* ϕ: Fix x. If $\phi(x-\cdot)$ is supported in the complement of *support* u, then $u(\phi(x-\cdot)=0$ by the definition of *support* u. If $u(\phi(x-\cdot)\neq 0$, then $\exists\ y\in$ *support* u and $y\in$ *support* $\phi(x-\cdot)$. Thus $y=\lim y_i$ with $\phi(x-y_i)\neq 0$, and $x-y\in$ *support* ϕ.

Finally, if $x\in$ *supportu* $*\,\phi$, there exist $x_i\to x$ with $u*\phi(x_i)\neq 0$ and

$$\begin{aligned}&x_i\in \mathit{support}\,\phi+\mathit{support}\,u\\&x\in\overline{\mathit{support}\,\phi+\mathit{support}\,u}=\mathit{support}\,\phi+\mathit{support}\,u\end{aligned}$$

because *support* u is compact.

We also have

Theorem 3.2. *Let* $u\in\mathcal{D}'(\mathbb{R}^n)$*, and* $\phi,\psi\in C_0^\infty(\mathbb{R}^n)$*. Then* $(u*\phi)*\psi=u*(\phi*\psi)$*.*

Proof. Before starting the proof, review Definition (1.2). $u*\phi\in C^\infty$. Fix x.

$$\begin{aligned}(u*\phi)*\psi(x)&=\int(u*\phi)(x-y)\psi(y)dy\\&=\lim_{h\to 0+}\sum_{k\in\mathbb{Z}^n}(u*\phi)(x-kh)\psi(kh)h^n\\&=\lim_{h\to 0+}u\left(\sum_{k\in\mathbb{Z}^n}\phi(x-kh-\cdot)\psi(kh)h^n\right)\\&=u\left(\int\phi(x-y-\cdot)\psi(y)dy\right)\end{aligned}$$

In the last line, we used the (obvious) fact that, for x fixed,

$$\sum_{k\in\mathbb{Z}^n}\phi(x-kh-z)\psi(kh)h^n\to\int\phi(x-y-z)\psi(y)dy$$

uniformly in z, and the same is true for after differentiating with respect to z an arbitrary number of times. Also, both LHS and RHS are supported in a fixed compact set. In other words, LHS $\to$ RHS in C_0^∞. □

This implies the important theorem on approximating distributions by C^∞ functions.

Theorem 3.3. *Let $u \in \mathcal{D}'(\mathbb{R}^n)$, and let η_ϵ be the standard mollifier. Then $u * \eta_\epsilon \in C^\infty(\mathbb{R}^n)$ and $u * \eta_\epsilon \to u$ in the sense of distribution theory (as $\epsilon \to 0$).*

Proof. We have to check

$$(u * \eta_\epsilon)(\phi) \to u(\phi)$$

for every $\phi \in C_0^\infty(\mathbb{R}^n)$. The proof is based on the observation that $u(\phi) = u * \phi_-(0)$ where $\phi_-(x) = \phi(-x)$. So it suffices to show $(u*\eta_\epsilon) * \phi(0) \to u * \phi(0)$. But

$$(u * \eta_\epsilon) * \phi(0) = u * (\eta_\epsilon * \phi)(0) \to u * \phi(0)$$

since $\eta_\epsilon * \phi \to \phi$ in C_0^∞. □

Now we **define the convolution of two distribution u_1, u_2, one of which is compactly supported.**

This is defined so that the formula

$$(u_1 * u_2) * \phi = u_1 * (u_2 * \phi)$$

holds for all $\phi \in C_0^\infty(\mathbb{R}^n)$. For simplicity, let's assume u_2 is compactly supported. Instead of defining $< u_1 * u_2, \phi >$ it suffices to define $(u_1 * u_2) * \phi(0)$. This is done in the obvious way:

$$(u_1 * u_2) * \phi(0) = u_1 * (u_2 * \phi)(0)$$

We have to check that $u_1 * u_2$ satisfies the continuity condition. Let $\phi_j \to 0$ in $C_0^\infty(\mathbb{R}^n)$ (see Definition (1.2)). Then so does $u_2 * \phi_j$, and $u_1 * (u_2 * \phi_j)(0) \to 0$.

Also, it τ_h denotes a translation, $(\tau_h \phi)(x) = \phi(x+h)$, then $\tau_h(u*\phi) = u * (\tau_h \phi)$ and

$$\begin{aligned}(u_1 * u_2) * \phi(h) = \tau_h\left((u_1 * u_2) * \phi\right)(0) &= \left((u_1 * u_2) * \tau_h \phi\right)(0) \\ = u_1 * (u_2 * \tau_h \phi)(0) &= u_1 * \left(\tau_h (u_2 * \phi)\right)(0) \\ &= u_1 * (u_2 * \phi)(h)\end{aligned}$$

Proposition 3.4. *Let $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$, one of which is compactly supported. Then*

$$\mathit{support}\,(u_1 * u_2) \subset \mathit{support}\, u_1 + \mathit{support}\, u_2$$

Proof. Let η_ϵ be a standard mollifier supported in a ball or radius ϵ. It suffices to show

$$\mathit{support}\,(u_1 * u_2) \subset \mathit{support}\, u_1 + \mathit{support}\, u_2 + \mathit{support}\, \eta_{\epsilon_0}$$

for all $\epsilon_0 > 0$. We do know

$$\begin{aligned} &\textit{support}\,(u_1 * u_2 * \eta_\epsilon) \subset \textit{support}\,u_1 + \textit{support}\,u_2 + \textit{support}\,\eta_\epsilon \\ &\subset \textit{support}\,u_1 + \textit{support}\,u_2 + \textit{support}\,\eta_{\epsilon_0} \end{aligned}$$

for all $0 < \epsilon < \epsilon_0$. Also remark that if A is closed and u is a distribution such that $\textit{support}\,u * \eta_\epsilon \subset A$ for all $\epsilon_0 > \epsilon > 0$, then $\textit{support}\,u \subset A$. This amounts to showing that if $u * \eta_\epsilon = 0$ in A^c, then $u = 0$ in A^c, which follows from $u * \eta_\epsilon \to u$ in the sense of distributions. □

Theorem 3.5. *Let u_1, u_2, u_3 distributions in $\mathbb{R}^n$, two of which are compactly supported. Then*

$$(u_1 * u_2) * u_3 = u_1 * (u_2 * u_3)$$

Proof. The proof follows by noticing it suffices to check $((u_1 * u_2) * u_3) * \phi = (u_1 * (u_2 * u_3)) * \phi$ for every $\phi \in C_0^\infty(\mathbb{R}^n)$ which follows easily from the defining property of Theorem (3.2). □

Theorem 3.6. *Let $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$, one of which is compactly supported. Then*

$$u_1 * u_2 = u_2 * u_1$$

Proof. The strategy is to show that $(u_1 * u_2) * (\phi * \psi) = (u_2 * u_1) * (\phi * \psi)$ for all test functions ϕ, ψ. This is done using the associativity property Theorem (3.2) together with the fact that convolutions of functions is commutative. We will not prove this □

Theorem 3.7. *Let $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$, one of which is compactly supported. Then*

$$\partial^\alpha(u_1 * u_2) = (\partial^\alpha u_1) * u_2 = u_1 * \partial^\alpha u_2 \tag{2}$$

Proof. We already know $\partial^\alpha(u * \phi) = (\partial^\alpha u) * \phi = u * (\partial^\alpha \phi)$, so the theorem is proved by convolving (2) with ϕ. □

Theorem 3.8. *Let $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$, one of which is compactly supported. Then*

$$\textit{sing support}\,(u_1 * u_2) \subset \textit{sing support}\,u_1 + \textit{sing support}\,u_2$$

Proof. The proof is based on the fact that if one of u_1, u_2 is smooth, so is $u_1 * u_2$. Let χ_1, χ_2 be supported in small neighborhoods of $\textit{sing support}\,u_1$, $\textit{sing support}\,u_2$, so that $(1 - \chi_1)u_1$ and $(1 - \chi_2)u_2$ are smooth. Then

$$\textit{sing support}\,(u_1 * u_2) \subset \textit{sing support}\,(\chi_1 u_1) * (\chi_2 u_2) \subset \textit{support}\,\chi_1 u_1 + \textit{support}\,\chi_2 u_2$$

□

Now we come back to PDEs. Let $P(D)$ be a constant coefficient differential operator. A distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ is called a fundamental solution if $P(D)E = \delta$. We already know formulas for (the) fundamental solution of the Laplace and heat operators. We will write down later several fundamental solutions of the wave operator.

Theorem 3.9. *If sing support* $(E) = \{0\}$*,* U *is open and* $u \in \mathcal{D}'(U)$ *is such that* $P(D)u \in C^\infty(U)$*, then* $u \in C^\infty(U)$

Proof. Let $V \subset\subset U$ an arbitrary open subset. It suffices to show $u \in C^\infty(V)$. Let $\zeta \in C_0^\infty(U)$, $\zeta = 1$ on V. Then $P(D)(\zeta u) = P(D)u$ in V, and in particular is C^∞ there. Finally,

$$\zeta u = \zeta u * \delta = \zeta u * P(D)(E) = (P(D)(\zeta u)) * E$$

and therefore

$$\textit{sing support}\,(\zeta u) \subset \textit{sing support}\,(P(D)(\zeta u)) + \{0\} = \textit{sing support}\,(P(D)(\zeta u))$$

But we know that *sing support* $(P(D)(\zeta u))$ is disjoint from V, so *sing support* (ζu) is also disjoint from V, in other words ζu, which equals u in V, is smooth there. ∎

4. The Fourier transform

Definition 4.1. The space of Schwartz functions $\mathcal{S}$ is defined by the requirement that all semi-norms

$$\sup_x |x^\alpha \partial^\beta f|$$

be finite. Convergence in this space means

$$\sup_x |x^\alpha \partial^\beta (f_n - f)| \to 0$$

for all α, β.

The Fourier transform $\mathcal{F}(f) = \hat{f}$ is defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)dx$$

The following are elementary properties which will be checked in class:

Lemma 4.2. *Let* $f \in \mathcal{S}$*, denote* $f_\lambda(x) = f(\lambda x)$ *(*$\lambda > 0$*),* $\tau_y f(x) = f(x+y)$ *(*$y \in \mathbb{R}^n$ *) and* $D_j = \frac{1}{i}\frac{\partial}{\partial x_j}$*. Then* $\hat{f} \in \mathcal{S}$ *and* $f \to \hat{f}$ *is* *continuous in the topology of $\mathcal{S}$. Also,*

$$\hat{f}_\lambda(\xi) = \frac{1}{\lambda^n}\hat{f}(\frac{\xi}{\lambda})$$
$$\mathcal{F}(\tau_y f)(\xi) = e^{iy\cdot\xi}\hat{f}(\xi)$$
$$\mathcal{F}(D_j f)(\xi) = \xi_j \hat{f}(\xi)$$
$$\mathcal{F}(x_j f)(\xi) = -D_j\hat{f}(\xi)$$
$$\mathcal{F}\left(e^{-\frac{|x|^2}{2}}\right)(\xi) = (2\pi)^{n/2}e^{-\frac{|\xi|^2}{2}}$$
$$\int f\hat{h} = \int \hat{f}g \quad \textit{for all } \hat{f},\, \hat{h} \in \mathcal{S}$$
$$\mathcal{F}\,(f * g) = \hat{f}\hat{g}$$

These easily imply the inversion formula and Plancherel formulas , which will be proved in class.

Theorem 4.3. *Let $f \in \mathcal{S}$. Then*

$$f(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\hat{f}(\xi)d\xi$$

Also,

$$\int_{\mathbb{R}^n} f(x)\overline{g}(x)dx = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{f}(\xi)\overline{\hat{g}}(\xi)d\xi$$

Definition 4.4. The space of continuous linear functionals $u : \mathcal{S} \to \mathbb{C}$ is the space of tempered distributions $\mathcal{S}'$. $u \in \mathcal{S}'$ if and only if there exists N and C such that

$$| < u, \phi > | \leq C \sum_{|\alpha|,|\beta|\leq N} \sup_x |x^\alpha \partial^\beta (f)|$$

for all $\phi \in \mathcal{S}$. If $u \in \mathcal{S}'$, then $\hat{u} \in \mathcal{S}'$ is defined by

$$< \hat{u}, \phi >=< u, \hat{\phi} >$$

for all $\phi \in \mathcal{S}$.

Example: The constant function $1 \in \mathcal{S}$ and $\hat{1} = (2\pi)^n\delta$.

5. Integrating functions on a k-dimensional hypersurface in $\mathbb{R}^n$

This section uses geometric notation: coodinates are written x^i.

Let S be a compact C^1 k-dimensional hypersurface in $\mathbb{R}^n$.
We will integrate $f : S \to \mathbb{R}$, continuous.
Each point in S has a neighborhood (a ball B) such that $S \cap B$ can be parametrized:
There exists

$$P : C \to S \cap B \subset \mathbb{R}^n$$

where C is open in $\mathbb{R}^k$ and P is one-to-one and onto. We assume P is C^1 and the n vectors ∇P_i are linearly independent. This insures S is a C^1 hypersurface. S can be covered by finitely many such balls $B_{r_i}(x_i)$.
Before anything else, we break up f as a finite sum of continuous functions f_i, each supported in one such B.
For convenience and without loss of generality, we assume f has been extended as a continuous function to $\mathbb{R}^n$, and is supported in the union of the $B_{r_i}(x_i)$.

Theorem 5.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^l function supported in a finite union of k balls $\cup_{i=1}^k B_{r_i}(x_i)$. Then there exist C^l functions f_i supported in $B_{r_i}(x_i)$ such that $f = \sum f_i$.*

Proof. First we argue that there are $s_i < r_i$ such that the support of f is covered by $\cup B_{s_i}(x_i)$. Indeed, the infinite union $\cup_{s_i<r_i} B_{s_i}(x_i)$ cover the compact support of f, so finitely many also do.
Let χ_i be C^l functions supported in $B_{r_i}(x_i)$, $\chi_i = 1$ on $B_{s_i}(x_i)$. (It is easy to construct such functions). Write

$$0 = f(x)(1 - \chi_1(x))(1 - \chi_2(x)) \cdots (1 - \chi_k(x))$$

so

$$\begin{aligned}
f(x) &= f(x)\chi_1(x) + f(x)(1 - \chi_1(x))\chi_2(x) + f(x)(1 - \chi_1(x))(1 - \chi_2(x))\chi_3(x) \\
&+ \cdots f(x)(1 - \chi_1(x))(1 - \chi_2(x)) \cdots \chi_k(x) \\
&=: f_1(x) + f_2(x) + \cdots f_k(x)
\end{aligned}$$

□

We proceed to integrate one of the f_is, written as f from now on.
The Euclidean metric in $\mathbb{R}^n$ induces a Riemannian metric on S. With respect to there coordinates, $g_{ij}(x)$ is defined as the Euclidean inner

product of the push-forward of the usual basis vectors in $\mathbb{R}^k$:

$$g_{ij}(x) = \frac{\partial P(x)}{\partial x^i} \cdot \frac{\partial P(x)}{\partial x^j}$$

In matrix notation,

$$(g_{ij}(x)) = (DP(x))^T \, DP(x)$$

The standard Riemannian geometry definition of $\int_S f dVol$ is

$$\int_S f dVol = \int_C f\,(P(x))\,\sqrt{|\det(g_{ij}(x))|}dx^1 \cdots dx^k$$

An explicit calculation shows this is independent of the parametrization. It agrees with familiar formulas in low dimensions:

Example 1: $k = 1$ (curves in $\mathbb{R}^n$). Here $P : (a, b) \to \mathbb{R}^n$, $P'(x) \neq 0$ for all x, and $g_{11}(x) = |P'(x)|^2$, so

$$\int_S f ds = \int_a^b f(P(x)|F'(x)|dx$$

Example 2: $k = 2$, $n = 3$ (surfaces in $\mathbb{R}^3$).
Here $DP(x)$ has two columns, $\frac{\partial P}{\partial x^1}$ and $\frac{\partial P}{\partial x^2}$.
Exercise (using a suitable rotation)

$$|\det(g_{ij})| = \left| \frac{\partial P}{\partial x^1} \times \frac{\partial P}{\partial x^2} \right|^2$$

so this is consistent with the Math 241 formula

$$\int_S f dS = \int_C f(P(x)) \left| \frac{\partial P}{\partial x^1} \times \frac{\partial P}{\partial x^2} \right| dx^1 dx^2$$

Example 3: $k = n$.
Here $(g_{ij}) = (DP(x))^T \, DP(x)$, $\det(g_{ij}) = \det(DP)^2$ so we get the change-of-variables formula

$$\int_S f dx = \int_C f(P(x))|\det DP(x)|dx$$

Example 4 (what we need to prove the divergence theorem)
Let $k = n - 1$ and S given (locally) as the graph of a function r:

$$P(x^1, \cdots x^{n-1}) = (x^1, \cdots x^{n-1}, r(x^1, \cdots x^{n-1}))$$

Then

$$(g_{ij}(x)) = \begin{pmatrix} 1 & 0 & \cdots & \frac{\partial r}{\partial x^1} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & \frac{\partial r}{\partial x^{n-1}} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 1 \\ \frac{\partial r}{\partial x^1} & \frac{\partial r}{\partial x^2} & \cdots & \frac{\partial r}{\partial x^{n-1}} \end{pmatrix}$$
$$= I_{(n-1)\times(n-1)} + \begin{pmatrix} \frac{\partial r}{\partial x^1} \\ \frac{\partial r}{\partial x^2} \\ \cdots \\ \frac{\partial r}{\partial x^{n-1}} \end{pmatrix} \begin{pmatrix} \frac{\partial r}{\partial x^1} & \frac{\partial r}{\partial x^2} & \cdots & \frac{\partial r}{\partial x^{n-1}} \end{pmatrix}$$

Exercise: $\det(g_{ij}) = 1 + |\nabla r|^2$.
So

$$\int_S f dS = \int_C f(x^1, \cdots, r(x^1, \cdots x^{n-1}))\sqrt{1 + |\nabla r|^2} dx^1 \cdots dx^{n-1}$$

5.2. **The length of a curve and the equation for geodesics as an Euler-Lagrange equation.** In the same set-up as before, let $x : [a, b] \to C \subset \mathbb{R}^k$ (C open) be a parametrized C^2 curve, $P : C \to S \subset \mathbb{R}^n$ the parametrization of (a subset of) the surface S (P is C^2, $DP(x)$ has maximal rank for all x). The length of the parametrized curve on S given by $\gamma(s) = P \circ x(s)$ is

$$\int_a^b |(P \circ x(s))'| ds = \int_a^b \sqrt{< (DP(x(s))x'(s), (DP(x(s))x'(s) >} ds$$
$$= \int_a^b \sqrt{\sum g_{ij}(x(s)\dot{x^i}(s)\dot{x^j}(s)} ds \tag{3}$$

We can forget P and S, and, given a Riemannian metric $g_{ij}(x)$ on C (that is, the matrix $g_{ij}(x)$ is positive definite for every x, and C^1, we can define the length of a parametrized curve $x(s)$ by (3).

We will prove the following:

Theorem 5.3. *If $x(s)$ is parametrized by arc-length (that is, $\sum g_{ij}(x(s)\dot{x^i}(s)\dot{x^j}(s) = 1$), and if $x(s)$ is the shortest path from $A = x(a)$ to $B = x(b)$, (A, B fixed) then it satisfies the geodesic equation*

$$\ddot{x^i}(s) + \Gamma^i_{jk}(x(s))\dot{x}^j(s)\dot{x}^k(s) = 0 \tag{4}$$

where the Christoffel symbols Γ^i_{jk} *are defined by*

$$\Gamma^i_{jk}(x) = \sum g^{il}(x)\Gamma_{ljk}(x)$$
where
$$\big(g^{il}(x)\big) = \big(g_{il}(x)\big)^{-1} \text{ matrix inverse}$$
$$\Gamma_{ljk} = \frac{1}{2}\left(\frac{\partial g_{lj}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l}\right)$$

Proof. The proof follows section 31.2 of volume 1 of the books by Dubrovin, Fomenko and Novikov. Let $L_1(\dot{x}, x)$ be the Lagrangian density $L_1(\dot{x}, x) = \sqrt{\sum g_{ij}(x(s)\dot{x^i}(s)\dot{x^j}(s)}$, and $L_2(\dot{x}, x) = (L_1(\dot{x}, x))^2 = \sum g_{ij}(x(s)\dot{x^i}(s)\dot{x^j}(s)$. Easy calculus shows that if x is parametrized by arc-length, then the Euler-Lagrange equations for L_1 are equivalent with the Euler-Lagrange equations for $(L_1)^2$.

It is also easy to see that the Euler-Lagrange equations for

$$L_2(\dot{x}, x) = \sum g_{ij}(x(s)\dot{x^i}(s)\dot{x^j}(s)$$

are exactly (4). We will also show that they satisfy $\sum g_{ij}(x(s))\dot{x^i}(s)\dot{x^j}(s) = const.$. This is "conservation of energy", similar to the conservation formulas for the energy-momentum tensor for PDEs which are Euler-Lagrange equations.

□

6. The gradient of a characteristic function

This is background material (formula 3.1.5 in Hörmander's book).

Theorem 6.1. *Let U be an open set with C^1 boundary. Then*

$$\nabla \chi_U = -\nu dS$$

where dS is surface measure on ∂U and ν is the outward pointing unit normal.

Proof. Let $h : \mathbb{R} \to \mathbb{R}$ be a smoothed out Heaviside function: $h(x) = 0$ if $x \leq 0$, $h(x) = 1$ if $x \geq 1$ and smooth in-between. Using a partition of unity, it suffices the prove the theorem for test functions ϕ supported in a small neighborhood of $x_0 \in \partial U$, where U agrees with $x_n > r(x_1, \cdots x_{n-1})$. Then

$$
\begin{aligned}
&< \nabla\chi_U, \phi >= - < \chi_U, \nabla\phi > \\
&= -\lim_{\epsilon\to 0}\int h(\frac{x_n - r(x_1,\cdots x_{n-1})}{\epsilon})\nabla\phi(x_1,\cdots,x_n)dx_1\cdots dx_n \\
&= \lim_{\epsilon\to 0}\int \nabla\left(h(\frac{x_n - r(x_1,\cdots x_{n-1})}{\epsilon})\right)\phi(x_1,\cdots,x_n) \\
&= \lim_{\epsilon\to 0}\int_{\mathbb{R}^n}\frac{1}{\epsilon}h'(\frac{x_n - r(x_1,\cdots x_{n-1})}{\epsilon})\cdot(-\nabla r(x_1,\cdots,x_{n-1}),1)\phi(x)dx \\
&= \int_{\mathbb{R}^{n-1}}(-\nabla r(x_1,\cdots,x_{n-1}),1)\left(\lim_{\epsilon\to 0}\int_{\mathbb{R}}\frac{1}{\epsilon}h'(\frac{x_n - r(x_1,\cdots x_{n-1})}{\epsilon})\phi(x)dx_n\right)dx_1\cdots dx_{n-1} \\
&= \int_{\mathbb{R}^{n-1}}\phi(x_1,\cdots x_{n-1}, r(x_1,\cdots r_{x-1}))(-\nabla r(x_1,\cdots,x_{n-1}),1)dx_1\cdots dx_{n-1} \\
&= -\int_{\partial U}\phi\nu dS
\end{aligned}
$$

(by the Calculus formulas for ν and dS). We used the fact that $\frac{1}{\epsilon}h'(\frac{x}{\epsilon})$ is an "approximation to the identity".

□

7. Solving the Cauchy problem for the wave equation in 1 and 3 dimensions

To solve (in $n+1$ dimensions, i.e. $x \in \mathbb{R}^n$, $t \in \mathbb{R}$)

$$
\begin{aligned}
&u_{tt} - \Delta u = 0 \ \text{ if } \ t > 0 \qquad (5)\\
&u(0,x) = f(x) \\
&u_t(0,x) = g(x)
\end{aligned}
$$

with $u \in C^2$, $u(t,\cdot) \in \mathcal{S}(\mathbb{R}^n)$ for each fixed $t > 0$, take Fourier transform in x:

$$
\begin{aligned}
&\hat{u}_{tt}(t,\xi) + |\xi|^2\hat{u}(t,\xi) = 0 \ \text{ if } \ t > 0 \\
&\hat{u}(0,\xi) = \hat{f}(\xi) \\
&\hat{u}_t(0,\xi) = \hat{g}(\xi)
\end{aligned}
$$

This ODE has solution

$$
\hat{u}(t,\xi) = \cos(t|\xi|)\hat{f}(\xi) + \frac{\sin(t|\xi|)}{|\xi|}\hat{g}(\xi)
$$

As it is clear from this formulation, it suffices to solve the problem with $f = 0$.

So we want $u(t,x)$ such that

$$\hat{u}(t,\xi) = \frac{\sin(t|\xi|)}{|\xi|}\hat{g}(\xi)$$

We are looking for a compactly supported distribution $E(t) \in \mathcal{E}'(\mathbb{R}^n)$ such that

$$\hat{E}(t) = \frac{\sin(t|\xi|)}{|\xi|}$$

At least in 1 and 3 dimensions, such a distribution is well-known and "elementary" (see the next section for other dimensions).

Then the solution will be

$$u(t,x) = E(t) * g$$

or, equivalently

$$\hat{u}(t,\xi) = \hat{E}(t,\xi)\hat{g}(\xi)$$

(Background facts: if E is a compactly supported distribution and $g \in \mathcal{S}$, $E * g(x)$ is defined as $< E, g(x-\cdot) >$. Its Fourier transform is $\hat{E}\hat{g}$. See Theorem 7.1.5 in Hörmander's book).

Also, for a compactly supported distribution E, $\hat{E}(\xi) =< E, e^{-ix\cdot\xi} >$ (E acts in the x variable). See Theorem 7.1.14 in Hörmander's book.

In one dimension, the Fourier transform of the characteristic function of $[-t,t]$ is

$$\int_{-t}^{t} e^{-ix\cdot\xi}dx = 2\frac{\sin(t\xi)}{\xi}$$

We found that in one space dimension

$$E(t,x) = \frac{1}{2}\chi_{[-t,t]}$$

and the solution to (5) with $f = 0$ is

$$\begin{aligned} u(t,x) &= \int E(t,y)g(x-y)dy \\ &= \frac{1}{2}\int_{-t}^{t} g(x-y)dy \\ &= \frac{1}{2}\int_{x-t}^{x+t} g(y)dy \end{aligned}$$

In 3 space dimensions, we compute the Fourier transform of surface measure on S^2: $\int_{S^2} e^{-ix\cdot\xi}dS_x$. Without loss of generality, $\xi = (0,0,|\xi|)$.

Integrating in spherical coordinates $(x_1, x_2, x_3) = (\sin(\phi)\cos(\theta), \sin(\phi)\sin(\theta), \cos(\phi))$,

$$\int_{S^2} e^{-ix_3|\xi|} dS_x = \int_0^{\pi}\int_0^{2\pi} e^{-i\cos(\phi)|\xi|}\sin(\phi)d\phi d\theta$$
$$= 2\pi \int_0^{\pi} e^{-i\cos(\phi)|\xi|}\sin(\phi)d\phi$$
$$= 2\pi \int_{-1}^{1} e^{-i\lambda|\xi|}d\lambda$$
$$= 4\pi\frac{\sin(|\xi|)}{|\xi|}$$

By the exact same calculation, the Fourier transform of surface measure on the sphere of radius $t > 0$ is $4\pi t\frac{\sin(t|\xi|)}{|\xi|}$.

Thus in 3 dimensions, if $f = 0$,

$$E(t, x) = \frac{1}{4\pi t}\text{surface measure on the sphere of radius } t$$

and

$$u(t, x) = \frac{1}{4\pi t}\int_{\partial B(0,t)} g(x - y)dS_y$$
$$= \frac{1}{4\pi t}\int_{\partial B(x,t)} g(y)dS_y$$

while, in general, the solution to (5) is

$$u(t, x) = \frac{\partial}{\partial t}\left(\frac{1}{4\pi t}\int_{\partial B(x,t)} f(y)dS_y\right)$$
$$+ \frac{1}{4\pi t}\int_{\partial B(x,t)} g(y)dS_y$$

8. The formula for the Cauchy problem for the wave equation in $n + 1$ dimensions

We need the (famous) family of distributions χ_+^{α}. For $\alpha > -1$, these are functions defined by

$$\chi_+^{\alpha} = \frac{x_+^{\alpha}}{\Gamma(\alpha + 1)}$$

Using properties of the Γ function,

$$(\chi_+^{\alpha})' = \chi_+^{\alpha-1}$$

This allows one to define χ_+^α for $\alpha \le -1$. Of special interest to us are

$$\chi_+^0(x) = H(x) \text{ (the Heaviside function)}$$
$$\chi_+^{-\frac{1}{2}}(x) = \frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{x}}H(x)$$
$$\chi_+^{-1} = H' = \delta$$

The general statement (which we will not prove right now) is

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_1^2} - \cdots - \frac{\partial^2}{\partial x_n^2}\right)\chi_+^{\frac{1-n}{2}}\left(t^2 - \cdots - x_n^2\right) = 4\pi^{\frac{n-1}{2}}\delta$$

The definition of the composition of a distribution with a smooth function is explained in the next section.

The above formula provides a fundamental solution for the wave equation (there are others, see Hörmander's book). The solution to the Cauchy problem,

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_1^2} - \cdots - \frac{\partial^2}{\partial x_n^2}\right)u = 0 \text{ if } t > 0$$
$$u(0,x) = 0, \;\; u_t(0,x) = f$$

is

$$u(t,x) = \left(E_+(t,\cdot) * f\right)(x)$$

where E_+ is defined as follows in the open set $t > 0$:

$$E+ = \frac{1}{2\pi^{\frac{n-1}{2}}}\chi_+^{\frac{1-n}{2}}\left(t^2 - \cdots - x_n^2\right)$$

This agrees with the results from the previous section. To see that in 3 dimensions, we need the important formula $\delta(f) = \frac{dS}{|\nabla f|}$ if f is C^1, $\nabla f(x) \neq 0$ if $f(x) = 0$ and dS is surface measure on $f = 0$ (explained in the next section).

9. Compositions with smooth functions and the chain rule

Theorem 9.1. *Let $u \in \mathcal{D}'(\mathbb{R})$, and $f : \mathbb{R}^n \to \mathbb{R}$ C^∞, such that $\nabla f(x) \neq 0$ for all $x \in supp\ u$. Then there exists a unique distribution $u \circ f$ (or f^*u) such that if u_i is a sequence of continuous functions, $u_i \to u$ in the sense of distribution theory, then $u_i \circ f \to u \circ f$. As a consequence, the chain rule is true.*

Proof. Using a partition of unity, it suffices to prove this for test functions supported in a sufficiently small open set. Let U open such that $\frac{\partial f}{\partial x_n}$ is bounded away from 0 on U. Consider the map Φ defined by

$(x_1, \cdots x_n) \to (x_1, \cdots, x_{n-1}, f(x_1, \cdots, x_n))$. By the inverse function theorem, after possibly shrinking U, we have $\Phi : U \to V$, one to one, onto, with a smooth inverse, and V open. Also,

$$f(\Phi^{-1}(x_1, \cdots, x_n)) = x_n \text{ for all } (x_1, \cdots, x_n) \in V$$

Let ϕ be a test function supported in U, and let u_i be a sequence of continuous functions converging to u in the sense of distribution theory. Then

$$< u_i \circ f, \phi >= \int_U u_i(f(x))\phi(x)dx$$

$$= \int_V u_i(x_n)\phi(\Phi^{-1}(x_1, \cdots, x_n))|\det \frac{\partial(\Phi^{-1})}{\partial x}|dx$$

$$= \int_{\mathbb{R}^{n-1}} < u_i, \phi(\Phi^{-1}(x_1, \cdots, \cdot))|\det \frac{\partial(\Phi^{-1})}{\partial x}| > dx_1 \cdots dx_{n-1}$$

$$\to \int_{\mathbb{R}^{n-1}} < u, \phi(\Phi^{-1}(x_1, \cdots, \cdot))|\det \frac{\partial(\Phi^{-1})}{\partial x}| > dx_1 \cdots dx_{n-1}$$

To pass to the limit inside the integral, we need the following lemma:

Then $< u_i, \phi(x_1, \cdots, x_{n-1}, \cdot) > \to < u, \phi(x_1, \cdots, x_{n-1}, \cdot)$ and the sequence is uniformly bounded in $(x_1, \cdots, x_{n-1})$. This follows from the uniform boundedness principle in a Frechet space. As a consequence, the chain rule is true. Indeed, given u we know we can find $u_i \to u$ in $\mathcal{D}'(\mathbb{R})$, $u_i \in C^\infty$. Let $f : \mathbb{R}^n \to \mathbb{R}$, C^∞, such that $\nabla f(x) \neq 0$ for all x. Then $u_i \circ f \to u \circ f$ as above, and $\nabla(u_i \circ f) = (u_i' \circ f)\nabla f \to (u' \circ f)\nabla f = \nabla(u \circ f)$. So $\nabla(u \circ f) = (u' \circ f)\nabla f$. $\square$

Remark 9.2. As an important application, let U be a C^1 bounded domain given by a defining function r. $U = \{r > 0\}$. Then $H \circ r = \chi_U$, and $\nabla(H \circ r) = \nabla(\chi_U) = dS\frac{\nabla r}{|\nabla r|}$, but $\nabla(H \circ r)$ also equals $H'(r)\nabla r = \delta(r)\nabla r$. Here δ stands for the delta function on the real line. As a consequence, $\delta(f) = \frac{dS}{|\nabla f|}$ where dS is surface measure on the surface $f = 0$.

At this stage, $\chi_+^{\frac{1-n}{2}}(t^2 - \cdots - x_n^2)$ is defined in the set $\mathbb{R}^{n+1} - \{0\}$, and homogeneous of degree $-n + 1$.

To extend it to $\mathcal{D}'(\mathbb{R}^{n+1})$ we need the following technical result (Theorem 3.2.3 in Hörmander). We will not prove this in class.

Theorem 9.3. *If $u \in \mathcal{D}'(\mathbb{R}^n - \{0\})$ is homogeneous of degree α and α is not an integer $\leq -n$, then u has a unique extension to $\mathcal{D}'(\mathbb{R}^n)$, and this extension is also homogeneous of degree α.*