DISTRIBUTIONS

CHRISTIAN REMLING

1. Basic properties

The infamous Dirac δ-function is an object much beloved by theoretical physicists. It cannot be given a mathematically sound definition (that is, as a *function*), but rather is usually imagined as having the values

$$\delta(x) = \begin{cases} \infty & x = 0 \\ 0 & x \neq 0 \end{cases},$$

and then one assumes that the singularity at $x = 0$ can be fine tuned to make $\int_{-\infty}^{\infty} \delta(x)\, dx = 1$.

Despite these inauspicious beginnings, the δ-function often seems a convenient tool at least in formal calculations, and so it is tempting to try to build a rigorous theory around the idea.

To do this, we take seriously the frequently heard excuse that the "definition" is not to be taken literally, but that everything will make sense when put under an integral sign. We can then in fact hope, slightly more ambitiously, that

$$\int_{-\infty}^{\infty} f(x)\delta(x)\, dx = f(0)$$

(since the integration seems to take place at $x = 0$ exclusively).

So now the δ-function really does something on other functions, and this process, whose exact workings are left unexplained, outputs a number. We now build a precise definition around this idea. For reason that will become clear later (perhaps), we want our *test functions* to be very nice functions.

Definition 1.1. Let $U \subseteq \mathbb{R}^d$ be a non-empty open set. A *distribution* $u \in \mathcal{D}'(U)$ is a bounded linear functional on $\mathcal{D} = C_0^\infty(U)$.

More specifically, $u : \mathcal{D} \to \mathbb{C}$ is a linear map, and if $K \subseteq U$ is any compact set, then there are $C = C(K) \geq 0$ and $N = N(K) \geq 0$ such

Date: February 28, 2020.

that

$$|u(\varphi)| \le C \sup_{x\in U, |\alpha|\le N} \left| \frac{\partial^\alpha}{\partial x^\alpha} \varphi(x) \right| \tag{1.1}$$

for all $\varphi \in \mathcal{D}$ with $\operatorname{supp}\varphi \subseteq K$.

We will be mostly interested in the case $U = \mathbb{R}^d$, in fact most of the time with $d = 1$, but sometimes the extra flexibility in the definition will be useful.

I have used multi-index notation for partial derivatives, so $\alpha = (\alpha_1, \ldots, \alpha_d)$, and

$$\frac{\partial^\alpha f}{\partial x^\alpha} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}} f(x_1, \ldots, x_d).$$

Recall that for well behaved functions such as $f \in \mathcal{D}$, it does not matter in which order these derivatives are taken, so this is well defined. We also set $|\alpha| = \alpha_1 + \ldots + \alpha_d$.

A *functional* on a vector space X over $\mathbb{C}$ is, by definition, a map $u : X \to \mathbb{C}$. The notation X' or X^* for the space of linear functionals (often with extra continuity conditions imposed) is common. The action of $u \in \mathcal{D}'$ on a *test function* $\varphi \in \mathcal{D}$ is often written as (u, φ) instead of $u(\varphi)$. This suggests the viewpoint of a pairing between $\mathcal{D}$ and its dual space $\mathcal{D}'$.

The supremum (really: maximum) on the right-hand side of (1.1) is also denoted by $\|\varphi\|_N$. Condition (1.1) can reformulated as follows, and this indeed has the feel of a continuity condition. (In fact, it is possible to put a topology $\mathcal{T}$ on $\mathcal{D}$ in such a way that (1.1) becomes equivalent to continuity with respect to $\mathcal{T}$ and the standard topology on $\mathbb{C}$, but this $\mathcal{T}$ is rather inconvenient to work with, and it is not useful at all for our purposes.)

Exercise 1.1. Let $u : \mathcal{D} \to \mathbb{C}$ be a linear functional. Show that the following are equivalent: (a) u satisfies the condition from Definition 1.1; (b) If $\varphi_n \in \mathcal{D}$, $\operatorname{supp}\varphi_n \subseteq K$ for some fixed compact set $K \subseteq \mathbb{R}^d$, and $\|\varphi_n\|_N \to 0$ for all $N \ge 0$, then $u(\varphi_n) \to 0$.

Example 1.1. The motivating example, $\delta(\varphi) = \varphi(0)$, is a distribution in this sense. Clearly, the map $\varphi \mapsto \varphi(0)$ is linear, and $|\varphi(0)| \le \sup|\varphi(x)| = \|\varphi\|_0$, as required.

Example 1.2. The functional $u(\varphi) = \varphi'(0)$ (and here $d = 1$) also defines a distribution, since u is obviously linear and $|u(\varphi)| \le \sup|\varphi'(x)| \le \|\varphi\|_1$.

Example 1.3. $u(\varphi) = \sum_{n\in\mathbb{Z}} \varphi(n)$ defines a distribution: Note, first of all, that there are no convergence issues because $\varphi(n) \neq 0$ for only finitely many n for any given $\varphi \in \mathcal{D}$. Moreover, if $\operatorname{supp}\varphi \subseteq [-N, N]$, then $|u(\varphi)| \le (2N+1)\|\varphi\|_0$.

Note that in the last example, the constant in estimate (1.1) does depend on the support of the test function, as anticipated as a possibility in the definition.

Exercise 1.2. Find an example of a distribution u (perhaps similar to the one from Example 1.3) for which N from (1.1) also necessarily depends on the support of test function.

If $f \in L^1_{\text{loc}}(U)$, then

$$u(\varphi) = \int_U f(x)\varphi(x)\, dx$$

defines a distribution since if $\operatorname{supp}\varphi \subseteq K$, then $|u(\varphi)| \le \|\varphi\|_0 \int_K |f|$. Moreover, f can be recovered from the distribution it generates, up to a change of its values on a null set. So we can identify f with the distribution $u = u_f$ and think of a locally integrable function as a distribution when this is convenient.

Theorem 1.2. *Let $f, g \in L^1_{\text{loc}}(U)$ and suppose that $u_f = u_g$. Then $f = g$ a.e.*

Exercise 1.3. Prove Theorem 1.2. If desired, you can focus on the case $U = \mathbb{R}^d$ exclusively, though the general case isn't much different. *Suggestion:* It suffices to show that if $u_f = 0$, then $f = 0$ a.e. Take convolutions of f with suitable functions and interpret these as applications of u_f to test functions.

Exercise 1.4. Show that $L^p \subseteq L^1_{\text{loc}}$ for all $p \ge 1$.

Given a general $u \in \mathcal{D}'$, it is sometimes interesting to ask if $u = u_f$ for some $f \in L^1_{\text{loc}}$ or, as this is often put, if $u \in \mathcal{D}'$ is a function.

Exercise 1.5. Show that $\delta \in \mathcal{D}'$ is not a (locally integrable) function in this sense.

More generally, if μ is a Borel measure on $\mathbb{R}^d$ (and this, as usual, includes the requirement that $\mu(K) < \infty$ for all compact sets $K \subseteq \mathbb{R}^d$), then $u(\varphi) = \int f\, d\mu$ is another distribution. Again, μ can be recovered from u, and thus we can think of measures as distributions if desired.

Example 1.4. Define

$$\left(\operatorname{PV}\frac{1}{x}, \varphi\right) = \lim_{h\to 0+} \int_{|x|>h} \frac{\varphi(x)}{x}\, dx. \tag{1.2}$$

Here, PV stands for *principal value,* and what is meant by this is the regularization of the integral by omitting the neighborhood $(-h,h)$ of the singularity at $x=0$ and then sending $h\to 0$. Some such device is needed since $\varphi(x)/x$ is not integrable if $\varphi(0)\not=0$.

Before we can use (1.2) as a definition, we must in fact establish that the limit exists. By the mean value theorem, $\varphi(x)=\varphi(0)+\varphi'(\xi)x$, for some $\xi=\xi(x)$ between 0 and x. If we also fix an $L>0$ with $\operatorname{supp}\varphi\subseteq[-L,L]$, then we can write the integral as

$$\varphi(0)\int_{h<|x|<L}\frac{dx}{x}+\int_{h<|x|<L}\varphi'(\xi)\,dx.$$

The first integral equals zero for all $h>0$, and the second one converges to $\int_{-L}^{L}\varphi'(\xi)\,dx$, by DC. Thus PV$(1/x)$ is well defined, and we also conclude that $|(\text{PV}(1/x),\varphi)|\le 2L\|\varphi\|_1$. So we have indeed defined a distribution.

Exercise 1.6. Show that PV$(1/x)$ is not a function. *Suggestion:* If it were, what would this function have to be equal to away from $x=0$?

2. Operations on distributions

If $f\in C^1(\mathbb{R})$, then its derivative f', being continuous, is a locally integrable function, so may be viewed as a distribution, which acts as $(f',\varphi)=\int_{-\infty}^{\infty}f'(x)\varphi(x)\,dx$. By an integration by parts, we can rewrite this as

$$(f',\varphi)=-\int_{-\infty}^{\infty}f(x)\varphi'(x)\,dx.$$

Exercise 2.1. Prove this in more detail. Why are there no boundary terms?

This motivates:

Definition 2.1. Let $u\in\mathcal{D}'(\mathbb{R})$. The *distributional derivative* $u'\in\mathcal{D}'$ is defined as the distribution $(u',\varphi)=-(u,\varphi')$.

We must in fact show that this indeed defines a new distribution, but this is easy: if $K\subseteq\mathbb{R}$ is compact and $|(u,\varphi)|\le C\|\varphi\|_N$ for $\varphi\in\mathcal{D}$ with $\operatorname{supp}\varphi\subseteq K$, then $|(u',\varphi)|\le C\|\varphi\|_{N+1}$.

In the same way, we can more generally define partial derivatives of distributions $u\in\mathcal{D}'(U)$, $U\subseteq\mathbb{R}^d$ as $(\partial u/\partial x_j,\varphi)=-(u,\partial\varphi/\partial x_j)$.

Note that any distribution has (distributional) derivatives, in fact of any order, since the derivatives are themselves distributions to which the definition can be applied. For example, $(\delta',\varphi)=-(\delta,\varphi')=-\varphi'(0)$ and, more generally, $\delta^{(n)}(\varphi)=(-1)^n\varphi^{(n)}(0)$.

Exercise 2.2. Show that $u \in \mathcal{D}'(U)$ satisfies $\frac{\partial^2 u}{\partial x_j \partial x_k} = \frac{\partial^2 u}{\partial x_k \partial x_j}$ (that is, higher order partial distributional derivatives can be taken in any order).

In particular, this holds for arbitrary locally integrable functions. We must keep in mind, though, that the distributional derivative could in principle differ from the classical derivative in cases where the latter also exists (in some sense); we will see examples of this later. However, if $f \in C^1$, then the distributional derivative of f agrees with its classical derivative, and in fact this was the case that motivated the definition.

This option of being able to take arbitrary derivatives is one of the main attractions of the theory of distributions. It is often useful even when in the end it turns out no distributions were involved. A typical situation would be that of a function $f(x)$ that we defined in a complicated way, and we would now like to differentiate it except that we don't know at this point if our function is actually differentiable. We can then always take the distributional derivative without having to worry about how to justify steps we would like to take.

Let's now enjoy this new freedom by looking at a few (but relatively harmless still) functions that are not differentiable everywhere in the classical sense.

Example 2.1. The standard calculus example for a function that is not differentiable at a point is $f(x) = |x|$. Clearly, $f \in L^1_{\text{loc}}$, so f does have a distributional derivative, which we'll denote simply by f' (but perhaps more circumspect notation would have been u'_f). We compute

$$\begin{aligned}
(f', \varphi) = -(f, \varphi') &= -\int_{-\infty}^{\infty} |x|\varphi'(x)\, dx \\
&= \int_{-L}^{0} x\varphi'(x)\, dx - \int_{0}^{L} x\varphi'(x)\, dx \\
&= x\varphi(x)\big|_{-L}^{0} - \int_{-L}^{0} \varphi(x)\, dx - x\varphi(x)\big|_{0}^{L} + \int_{0}^{L} \varphi(x)\, dx \\
&= \int_{-\infty}^{\infty} \operatorname{sgn}(x)\varphi(x)\, dx,
\end{aligned}$$

with $\operatorname{sgn}(x) = 1$ for $x > 0$ and $\operatorname{sgn}(x) = -1$ for $x < 0$. We also took $L > 0$ large enough so that $\operatorname{supp}\varphi \subseteq [-L, L]$ in this calculation. We have found, unsurprisingly, that $f' \in \mathcal{D}'$ is in fact a function, and it is equal almost everywhere to the pointwise derivative $f'(x) = \operatorname{sgn}(x)$ of f.

Now let's continue and compute $f''=\text{sgn}'\in\mathcal{D}'$. From the definition of the distributional derivative,

$$(\text{sgn}',\varphi)=-\int_{-\infty}^{\infty}\text{sgn}(x)\varphi'(x)\,dx=\int_{-\infty}^{0}\varphi'(x)\,dx-\int_{0}^{\infty}\varphi'(x)\,dx$$
$$=2\varphi(0)=2(\delta,\varphi).$$

In other words, $f''=2\delta$. This is not a function (compare Exercise 1.5).

Example 2.2. Now let's look at $f(x)=\log|x|$. This is a locally integrable function. What is its distributional derivative $f'\in\mathcal{D}'$? Away from $x=0$, f is smooth and has the classical derivative $f'(x)=1/x$, but this cannot be the answer to our question because this function is not locally integrable, so does not generate a distribution in an obvious way. Let's look at this more closely. By DC,

$$(f',\varphi)=-\int_{-\infty}^{\infty}\log|x|\varphi'(x)\,dx=-\lim_{h\to 0+}\int_{|x|>h}\log|x|\varphi'(x)\,dx.$$

We now again split this last integral into the two parts $x<-h$ and $x>h$ and then integrate by parts in those. This produces

$$(f',\varphi)=\lim_{h\to 0+}\left(\int_{|x|>h}\frac{\varphi(x)}{x}\,dx+\log h(\varphi(h)-\varphi(-h))\right),$$

and here the second term goes to zero because $\varphi(\pm h)=\varphi(0)+O(h)$ and $h\log h\to 0$.

Let's summarize:

Theorem 2.2. *In* $\mathcal{D}'$, $|x|'=\text{sgn}(x)$, $|x|''=2\delta$, $(\log|x|)'=\text{PV}(1/x)$.

As our next example, let's take an increasing function $F:\mathbb{R}\to\mathbb{R}$. Such an F is bounded on any compact set, so is clearly locally integrable and thus generates a distribution. What is its distributional derivative? Before we answer this, recall that a monotone F is differentiable, in the classical sense, at almost all $x\in\mathbb{R}$. Moreover, F generates a Borel measure $\mu=m_F$, by setting $\mu((a,b])=F(b)-F(a)$, if we also make F right-continuous here. This changes F at at most countably many points, so will not affect the *distribution* $F\in\mathcal{D}'$, and we may thus indeed insist on this normalization at no cost. It will also be convenient to make $F(0)=0$, by adding a constant to the original function. It seems intuitively clear that this should not affect F', but let's establish this carefully before we proceed.

Theorem 2.3. *Let* $u,v\in\mathcal{D}'$, $c\in\mathbb{C}$. *Then* $(u+v)'=u'+v'$, $(cu)'=cu'$, *and* $c'=0$.

We of course add distributions and multiply them by constants in the obvious way: for example $(u+v)(\varphi) := u(\varphi) + v(\varphi)$.

Exercise 2.3. Prove Theorem 2.3.

So indeed $(F+c)' = F'$, as anticipated.

With this further normalization, $F(0) = 0$, we have $F(x) = \mu((0,x])$ for $x \geq 0$ and $F(x) = -\mu((x,0])$ for $x < 0$. We are now ready to compute

$$\begin{aligned}(F', \varphi) &= -\int_{-\infty}^{\infty} F(x)\varphi'(x)\,dx \\ &= \int_{-\infty}^{0} \mu((x,0])\varphi'(x)\,dx - \int_{0}^{\infty} \mu((0,x])\varphi'(x)\,dx. \end{aligned} \tag{2.1}$$

Let me look at the last integral in more detail: we can write $\mu((0,x]) = \int_{(0,x]} d\mu(t)$. We obtain an iterated (double) integral, to which we apply Fubini-Tonelli. Note that the integration is over the set $0 < t \leq x$, so this produces

$$-\int_{(0,\infty)} d\mu(t) \int_{t}^{\infty} dx\,\varphi'(x) = \int_{(0,\infty)} \varphi(t)\,d\mu(t).$$

Of course, the first integral from (2.1) can be given a similar treatment: it is equal to $\int_{(-\infty,0]} \varphi(t)\,d\mu(t)$. Putting things together, we thus see that $(F', \varphi) = \int_{\mathbb{R}} \varphi(x)\,d\mu(x)$. In other words, $F' = \mu$, the distribution generated by the measure that F induces. Let's state this formally.

Theorem 2.4. *If $F : \mathbb{R} \to \mathbb{R}$ is increasing, then $F' = \mu$, with $\mu((a,b]) = F(b+) - F(a+)$.*

In particular, note that the distributional derivative need not equal the pointwise derivative. This holds only if μ is absolutely continuous. If μ has a singular part, then F' is not a function. (You may wonder how a measure can ever be a function, but in fact it happens, with our use of terminology, since the measure $f\,dx$ and the function f are the same distribution, so in that sense the measure $f\,dx$ is a function also.) We in fact already saw this principle in action in the simple example $F(x) = \text{sgn}(x)$, which gave $F' = 2\delta$ (and this of course is a measure, though we didn't emphasize this originally.)

This condition, of having a positive measure as its distributional derivative, *characterizes* the increasing functions. This criterion is not particularly useful (usually there will be easier ways to check that a given function is increasing than computing its distributional derivative), but it is elegant and satisfying.

Theorem 2.5. *Let $F : \mathbb{R} \to \mathbb{R}$ be a locally integrable function. Then F is increasing (more precisely, F is almost everywhere equal to an increasing function) if and only if $F' \in \mathcal{D}'$ is a Borel measure on $\mathbb{R}$.*

The qualification spelled out in parentheses is unavoidable since changing F on a null set will not affect the *distribution* $F \in \mathcal{D}'$, so a criterion involving the distributional derivative is clearly insensitive to such a change.

Proof. We already showed, and stated as Theorem 2.4, that an increasing function has a measure as its derivative. Conversely, suppose now that $F \in L^1_{\text{loc}}$ has distributional derivative $F' = \mu$, for some Borel measure μ. Let G be an increasing function that corresponds to μ. In more concrete style, we can set

$$G(x) = \begin{cases} \mu((0,x]) & x \geq 0 \\ -\mu((x,0]) & x < 0 \end{cases}.$$

By Theorem 2.4, $G' = \mu$ in $\mathcal{D}'$. I now refer to the general fact, stated as Theorem 2.6 below, that a distribution can be recovered from its derivative, up to a constant, to conclude that $F(x) = G(x) + c$ a.e., for some $c \in \mathbb{R}$; of course, I also use Theorem 1.2 in this step. So F is almost everywhere equal to an increasing function, as claimed. □

Theorem 2.6. *Let $u \in \mathcal{D}'(\mathbb{R})$ and suppose that $u' = 0$. Then $u = c$.*

More explicitly, the claim is that u is the constant function c, for some $c \in \mathbb{C}$, that is, $(u, \varphi) = \int c\varphi$.

Lemma 2.7. *Let $\varphi \in C_0^\infty(\mathbb{R})$. Then $\varphi = \psi'$ for some $\psi \in C_0^\infty(\mathbb{R})$ if and only if $\int_{-\infty}^{\infty} \varphi(x)\, dx = 0$.*

Proof. It is of course clear that $\int \psi' = 0$. Conversely, suppose that $\int \varphi = 0$. Let $\psi(x) = \int_{-\infty}^{x} \varphi(t)\, dt$. It is then clear that $\psi \in C^\infty$, $\psi' = \varphi$, and $\psi(x) = 0$ for all small x. Our assumption on φ makes sure that $\psi(x) = 0$ also for all large x. So ψ has compact support. □

Proof of Theorem 2.6. Fix once and for all a $\varphi_0 \in C_0^\infty$ with $\int \varphi_0 = 1$. For arbitrary $\varphi \in C_0^\infty$, write $\varphi = a\varphi_0 + (\varphi - a\varphi_0)$, with $a = \int \varphi$. Then $\int (\varphi - a\varphi_0) = 0$, so $\varphi - a\varphi_0 = \psi'$ for some $\psi \in C_0^\infty$, by the lemma. However, $(u, \psi') = -(u', \psi) = 0$, and thus

$$(u, \varphi) = a(u, \varphi_0) = (u, \varphi_0) \int_{-\infty}^{\infty} \varphi(x)\, dx.$$

This says that $u = c$, as claimed, with $c = (u, \varphi_0)$. □

Monotonicity can of course also be characterized in terms of the derivative in a more classical setting, and in fact this is very easy: if $F : \mathbb{R} \to \mathbb{R}$ is differentiable, then F is increasing if and only if $F'(x) \geq 0$. Theorem 2.5 is actually a distributional analog of this statement, though this is not immediately apparent. To make this connection explicit, we need one more definition and a major result.

Definition 2.8. A distribution $u \in \mathcal{D}'$ is called *positive* (notation: $u \geq 0$) if $u(\varphi) \geq 0$ for all test functions $\varphi(x) \geq 0$.

Theorem 2.9. *A distribution $u \in \mathcal{D}'(U)$ is positive if and only if it is a (positive) Borel measure μ on U.*

One direction is obvious: if $\varphi \geq 0$, then also $\int_U \varphi(x)\, d\mu(x) \geq 0$, so a measure defines a positive contribution. The converse is one version of the *Riesz representation theorem.* Its proof is long and technical, and I don't want to discuss it here. By combining the Riesz representation theorem with Theorem 2.5, we obtain, as promised, the following.

Corollary 2.10. *A locally integrable function $F : \mathbb{R} \to \mathbb{R}$ is increasing if and only if $F' \geq 0$ in $\mathcal{D}'$.*

These ideas produce even more satisfying results when applied to other classes of functions. We call $f : \mathbb{R} \to \mathbb{C}$ (locally) *absolutely continuous* and write $f \in AC$ if

$$f(x) = f(0) + \int_0^x g(t)\, dt \tag{2.2}$$

for some $g \in L^1_{\text{loc}}$. In much the same way as above, we can then prove:

Theorem 2.11. *Let $f \in L^1_{\text{loc}}(\mathbb{R})$. Then $f \in AC$ if and only if the distributional derivative of f is a function. In this case, the distributional derivative agrees with the pointwise derivative, which will exist almost everywhere.*

Proof. If $f \in AC$, so (2.2) holds, then we can compute $f' \in \mathcal{D}'$ exactly as above as

$$\begin{aligned}(f', \varphi) &= -\int_{-\infty}^{\infty} f(x)\varphi'(x)\, dx \\ &= \int_{-\infty}^{0} dx\, \varphi'(x) \int_x^0 dt\, g(t) - \int_0^{\infty} dx\, \varphi'(x) \int_0^x dt\, g(t).\end{aligned}$$

The term $f(0)$ from (2.2) doesn't contribute because $\int \varphi' = 0$. We change the order of integration, so for example the first integral becomes

$$\int_{-\infty}^{0} dt\, g(t) \int_{-\infty}^{t} dx\, \varphi'(x) = \int_{-\infty}^{0} g(t)\varphi(t)\, dt,$$

and after putting everything back together we see that $f' = g \in L^1_{\text{loc}}$, as claimed. From general facts about absolutely continuous functions we also know that $g(x) = f'(x)$, the pointwise derivative, almost everywhere.

Conversely, if $f' = g \in L^1_{\text{loc}}$ in $\mathcal{D}'$, then we introduce $h(x) = \int_0^x g(t)\, dt$. By what we just showed, we then know that also $h' = g$ in $\mathcal{D}'$, so $f = h + c$ by Theorem 2.6. □

As a preparation for our discussion of the analogous result for BV functions, we make one more general definition.

Definition 2.12. Let $u \in \mathcal{D}'(U)$ and let $V \subseteq U$ be open. Then the *restriction* $v = u|_V \in \mathcal{D}'(V)$ of u to V is defined as $(v, \varphi) = (u, \varphi_0)$, with $\varphi_0(x) = \varphi(x)$ for $x \in V$ and $\varphi_0(x) = 0$ for $x \in U \setminus V$.

Or, to say this in somewhat less formal style, we restrict a distribution to V by only applying it to test functions φ with $\operatorname{supp} \varphi \subseteq V$. Of course, a distribution cannot be restricted in literally the same way as a function since it cannot be evaluated at individual points $x \in U$.

Example 2.3. $\delta|_{\mathbb{R}\setminus\{0\}} = 0$, and this follows at once from the definition: if $0 \notin \operatorname{supp} \varphi$, then $\varphi(0) = (\delta, \varphi) = 0$, so δ indeed acts as the zero distribution on the smaller test function space $C_0^\infty(\mathbb{R} \setminus \{0\})$.

As one would have hoped, distributions can be recovered from their local restrictions. This is an interesting general fact, and we will also need it soon when we discuss BV functions.

Theorem 2.13. *Let $u \in \mathcal{D}'(U)$, and assume that every $x \in U$ has a neigborhood $V = V_x$ such that $u|_V = 0$. Then $u = 0$.*

Lemma 2.14. *Let $W_\alpha \subseteq U$ be an open cover of U. Then there are $\varphi_n \in C_0^\infty(U)$, $0 \le \varphi_n \le 1$, such that: (1) $\sum_{n\ge1} \varphi_n(x) = 1$; (2) $\operatorname{supp} \varphi_n \subseteq W_{\alpha_n}$ for some α_n, for each $n \ge 1$; (3) if $K \subseteq U$ is compact, then $\sum_{n=1}^N \varphi_n(x) = 1$ for $x \in K$ for some $N = N(K) \ge 1$.*

Such a collection of functions is often called a *partition of unity,* subordinate to $\{W_\alpha\}$.

Sketch of proof. We can find countably many open balls B_n such that each of them is contained in some W_α, and $\bigcup (1/2)B_n = U$. Here cB denotes the ball with the same center as B and c times the radius. Pick functions $\psi_n \in C_0^\infty(U)$ with $0 \le \psi_n \le 1$, $\psi_n = 1$ on $(1/2)B_n$ and $\psi_n = 0$ on $((2/3)B_n)^c$. Then define

$$\varphi_1 = \psi_1,\ \varphi_2 = (1-\psi_1)\psi_2, \ldots, \varphi_n = (1-\psi_1)\cdots(1-\psi_{n-1})\psi_n.$$

Property (2) is then clear: $\operatorname{supp}\varphi_n \subseteq B_n \subseteq W_{\alpha_n}$. Next, I claim that

$$\varphi_1 + \dots \varphi_N = 1 - (1-\psi_1)\cdots(1-\psi_N),$$

and it is in fact straightforward to establish this by an induction on N.

So in particular $\varphi_1 + \ldots + \varphi_N = 1$ on $\bigcup_{n=1}^N (1/2)B_n$. This implies (1) and also (3) since any compact $K \subseteq U$ is contained in such a union for some N. □

Proof of Theorem 2.13. Apply the lemma to $\{W_\alpha\} = \{V_x : x \in U\}$. Let $\varphi \in C_0^\infty(U)$ be an arbitrary test function. We can then write $\varphi = \sum \varphi\varphi_n$. In fact, property (3), applied to $K = \operatorname{supp}\varphi$, shows that $\varphi = \sum_{n=1}^N \varphi\varphi_n$ for some $N = N(\varphi)$. The products $\varphi\varphi_n$ are test functions themselves, and each of them has its support contained in some V_x, by (2). Thus

$$(u, \varphi) = \sum_{n=1}^N (u, \varphi\varphi_n) = 0,$$

as claimed. □

We say that a function $f : \mathbb{R} \to \mathbb{R}$ is (locally) of *bounded variation* if $f \in BV[a,b]$ for all $[a,b] \subseteq \mathbb{R}$, and we write $f \in BV$ in this case. Recall also that $f \in BV$ if and only if f is the difference of two increasing functions. In particular, BV functions are locally bounded and thus also locally integrable.

On bounded intervals, BV functions correspond to signed measures in the same way increasing functions correspond to positive measures. On the real line, this is not quite true as examples such as $f(x) = |x|$ demonstrate: This function is monotone on both half lines, and the associated positive measures are $\mu_1 = \chi_{(-\infty,0)}\,dx$, $\mu_2 = \chi_{(0,\infty)}\,dx$, but $\mu_2 - \mu_1$ is not a signed measure, as both the positive and the negative parts are infinite.

Theorem 2.15. *Let $f : \mathbb{R} \to \mathbb{R}$ be locally integrable. Then $f \in BV$ (more precisely: $f = g$ a.e. for some $g \in BV$) if and only if the restriction of $f' \in \mathcal{D}'(\mathbb{R})$ to $(-L, L)$ is a finite signed measure for every $L > 0$.*

Proof. If $f \in BV$, then we can recognize its derivative f' (locally) as the signed measure $\mu((a,b]) = f(b+) - f(a+)$ by the same argument that we already used a number of times, for example in the proof of Theorem 2.4. Or we could write $f = f_1 - f_2$, with f_j increasing, and then refer to this theorem directly (not its proof).

The converse is also proved by the same arguments as above: If $f' = \mu$ on $(-L, L)$, then we define a BV function as

$$g(x) = \begin{cases} \mu((0,x]) & x \ge 0 \\ -\mu((x,0]) & x < 0 \end{cases};$$

note that we can in fact consistently define g on all of $\mathbb{R}$ in this way since the $\mu = \mu_L$ will be compatible in the sense that if $L_2 > L_1$ and $A \subseteq (-L_1, L_1)$, then $\mu_{L_1}(A) = \mu_{L_2}(A)$, and this is true because these measures were obtained as restrictions of one and the same distribution $f' \in \mathcal{D}'(\mathbb{R})$.

Exercise 2.4. Strictly speaking, I am using the following (obvious looking) fact here: If $W \subseteq V \subseteq U$ are open sets and $u \in \mathcal{D}'(U)$, then $u|_W = (u|_V)\,|_W$. Prove this.

Now g has the same distributional derivative as f, after restricting to $(-L, L)$. By Theorem 2.13, $f' - g' = 0$. By Theorem 2.6, this means that $f = g + c \in BV$. □

The philosophy that led to Definition 2.1 lets us define several other operations on distributions in a natural way. The guiding principle is: move the operation over to the test function, in such a way that it gives the right answer for distributions that are nice functions.

We write $(\tau_a f)(x) = f(x-a)$ for the translation of a function $f : \mathbb{R}^d \to \mathbb{C}$ by $a \in \mathbb{R}^d$. For a well behaved f and a test function φ we then have $\int f(x-a)\varphi(x)\,dx = \int f(x)\varphi(x+a)\,dx$. This suggests to define $\tau_a u$, for $u \in \mathcal{D}'(\mathbb{R}^d)$ and $a \in \mathbb{R}^d$ as $(\tau_a u, \varphi) = (u, \tau_{-a}\varphi)$.

Exercise 2.5. Check that indeed $\tau_a u \in \mathcal{D}'$. Then give an equally natural definition of the dilation $\delta_a u$ of a distribution that is based on the action $(\delta_a f)(x) = f(ax)$ $(a > 0)$ on functions.

Definition 2.16. Let $u \in \mathcal{D}'(\mathbb{R})$, $\varphi \in C_0^\infty(\mathbb{R})$. Then we define the *convolution* of u and φ as $(u * \varphi)(x) = (u, \tau_x R\varphi)$, with $(R\varphi)(x) = \varphi(-x)$.

If $u \in L^1_{\text{loc}}$, then this becomes $(u*\varphi)(x) = \int u(t)\varphi(x-t)\,dt$, so reduces to the usual definition of the convolution. Note that the convolution of a distribution and a test function is a *function.*

This convolution has the same basic properties as the convolution of two functions:

Theorem 2.17. *Let $u \in \mathcal{D}'$, $\varphi \in C_0^\infty$. Then $u * \varphi \in C^\infty$, and*

$$(u * \varphi)' = u' * \varphi = u * \varphi'.$$

Proof. We have

$$\frac{1}{h}\left((u*\varphi)(x+h)-(u*\varphi)(x)\right)=\left(u,\frac{\varphi(x+h-t)-\varphi(x-t)}{h}\right),$$

and here we apply u to the difference quotient as a function of t, for fixed x, h. If we now send $h \to 0$, then this expression will converge to $\varphi'(x-t)$. This is clear in pointwise sense, but this isn't quite good enough here. Rather, what we need is that the functions

$$g_h(t)=\frac{\varphi(x+h-t)-\varphi(x-t)}{h}-\varphi'(x-t)\in C_0^\infty$$

together with all their derivatives converge to zero uniformly in t and their supports are contained in one fixed compact set. This is condition (b) from Exercise 1.1, and by the result of that Exercise, we can then conclude that $(u, g_h) \to 0$ as $h \to 0$.

Exercise 2.6. Prove this claim about g_h in more detail.

It follows that $u*\varphi$ is indeed differentiable, and $(u*\varphi)' = u*\varphi'$. Since this is still of the same general form (convolution of a distribution with a test function), we can repeat this argument, and we have in fact shown that $u*\varphi \in C^\infty$.

Finally,

$$u'*\varphi=(u',\tau_x R\varphi)=-(u,(\tau_x R\varphi)')=(u,\tau_x R\varphi')=u*\varphi'.$$

□

Definition 2.18. Let $g \in C^\infty(U)$, $u \in \mathcal{D}'(U)$. Then $gu \in \mathcal{D}'$ is defined as the distribution $(gu, \varphi) = (u, g\varphi)$.

This definition is of course motivated by the trivial formula $\int (gu)\varphi = \int u(g\varphi)$, which is valid for nice *functions* u. Note that we have defined the product of a distribution and a (smooth) *function;* there is no natural notion of a product of two distributions. This is not surprising. For example, it is indeed hard to come up with a reasonable interpretation of what $\delta \cdot \delta$ might be.

Exercise 2.7. Show that gu indeed is a distribution.

Exercise 2.8. Show that the product rule is valid in the (limited) context of Definition 2.18: if $g \in C^\infty$, $u \in \mathcal{D}'$, then $(gu)' = g'u + gu'$.

Example 2.4. What is $g\delta$? We compute $(g\delta, \varphi) = (\delta, g\varphi) = g(0)\varphi(0)$. This is what the (constant) multiple $g(0)\delta$ would have done on φ, so $g\delta = g(0)\delta$.

Exercise 2.9. Compute similarly $g\delta'$, for $g \in C^\infty(\mathbb{R})$.

3. Convergence of distributions

Definition 3.1. Let $u_n, u \in \mathcal{D}'$. We say that $u_n \to u$ (in $\mathcal{D}'$, or in the sense of distributions) if $(u_n, \varphi) \to (u, \varphi)$ for all $\varphi \in \mathcal{D}$.

This is a rather weak requirement. For example, $e^{inx} \to 0$ in $\mathcal{D}'$, if we view these functions as distributions.

Exercise 3.1. Prove this. *Suggestion:* Riemann-Lebesgue lemma
Can you in fact also show that $n^N e^{inx} \to 0$ in $\mathcal{D}'$ for any $N \geq 1$?

Similarly, for example $n\delta_n \to 0$, where $(\delta_n, \varphi) = \varphi(n)$. To confirm this, simply observe that $(n\delta_n, \varphi) = n\varphi(n)$ will be zero once n is outside the support of φ.

Example 3.1. Let $f_n(x) = n\chi_{(-1/2n,1/2n)}(x)$. Then $f_n \to \delta$. To prove this, we use the mean value theorem to write $\varphi(x) = \varphi(0) + \varphi'(\xi)x = \varphi(0) + O(x)$ and compute

$$(f_n, \varphi) = n\int_{-1/2n}^{1/2n} \varphi(x)\, dx = \varphi(0) + n\int_{-1/2n}^{1/2n} O(1/n)\, dx = \varphi(0) + O(1/n),$$

so this converges to $\varphi(0)$ as $n \to \infty$, as claimed.

Exercise 3.2. Prove, more ambitiously, that for any $f \in L^1(\mathbb{R})$, we have $f_n \to (\int f)\delta$, with $f_n(x) = nf(nx)$.

Theorem 3.2. *If $u_n \to u$ in $\mathcal{D}'$, then also $u_n' \to u'$.*

So we can say that the derivative is a continuous operation on $\mathcal{D}'$. Despite the slightly surprising character of this statement, its *proof* is extremely simple: $(u_n', \varphi) = -(u_n, \varphi') \to -(u, \varphi') = (u', \varphi)$ □

(Differentiation is also continuous on $\mathcal{D}$, if this test function space is endowed with the topology $\mathcal{T}$ mentioned before Exercise 1.1, so Theorem 3.2 is not so very surprising after all.)

Example 3.2. Let's return to the functions $f_n \to \delta$ from the previous Example. We have $f_n' = n(\delta_{-1/2n} - \delta_{1/2n})$, and this can be established by a straightforward calculation, similar to the one that showed that $\text{sgn}' = 2\delta$. Or simply notice that $f_n \in BV$ and recall that the distributional derivative of such a function is the associated (signed) measure.
Theorem 3.2 now shows that

$$n(\delta_{-1/2n} + \delta_{1/2n}) \to \delta'.$$

This can easily be checked directly: $n(\varphi(-1/2n) - \varphi(1/2n)) \to -\varphi'(0)$ since $\varphi(x) = \varphi(0) + \varphi'(0)x + O(x^2)$, by Taylor's theorem.

After these easy introductory examples, we are now ready for a more substantial convergence statement.

Theorem 3.3 (Sokhotski-Plemelj formula)**.** *In $\mathcal{D}'(\mathbb{R})$,*

$$\lim_{h\to 0+} \frac{1}{x-ih} = \mathrm{PV}\,\frac{1}{x} + i\pi\delta.$$

Proof. We write

$$\frac{1}{x-ih} = \frac{x}{x^2+h^2} + i\frac{h}{x^2+h^2},$$

and we'll treat the real and imaginary parts separately. The imaginary part is actually covered by (the continuous version of) Exercise 3.2 because, in the notation used there, $h/(x^2+h^2) = f_{1/h}(x)$, with $f(x) = 1/(1+x^2)$, and (by calculus) $\int f = \pi$. But a direct argument is also easy:

$$\int_{-\infty}^{\infty} \frac{h}{x^2+h^2}\varphi(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{1+t^2}\varphi(ht)\,dt \to \varphi(0)\int_{-\infty}^{\infty} \frac{dt}{1+t^2} = \pi\varphi(0)$$

The convergence follows from DC, and in this step we use that φ is continuous and bounded.

Next, we look at the real part. Since we anticipate the limit being the principal value distribution, we first establish that we can remove a small interval about $x = 0$. More precisely,

$$\int_{-h}^{h} \frac{x}{x^2+h^2}\varphi(x)\,dx = \varphi(0)\int_{-h}^{h} \frac{x}{x^2+h^2}\,dx + \int_{-h}^{h} \frac{x}{x^2+h^2}O(x)\,dx,$$

and this goes to zero because the first integral on the right-hand side has this value, and the second one is $O(h)$.

So it now suffices to show that

$$\lim_{h\to 0+} \int_{|x|>h} \left(\frac{x}{x^2+h^2} - \frac{1}{x}\right)\varphi(x)\,dx = 0.$$

To do this, we write

$$\frac{x}{x^2+h^2} - \frac{1}{x} = \frac{-h^2}{(x^2+h^2)x}$$

and again use the Taylor expansion $\varphi(x) = \varphi(0) + O(x)$ of the test function. The term with $\varphi(0)$ doesn't contribute since the integrand is odd, and what is then left is of order $O(h^2)\int_{|x|>h} dx/(x^2+h^2)$. Since $\int_{\mathbb{R}} h/(x^2+h^2)\,dx = \pi$, this goes to zero. □

Exercise 3.3. Show that $\log(x^2+h^2) \to 2\log|x|$ in $\mathcal{D}'$.

A slightly slicker proof of the second part can be based on this Exercise and Theorem 3.2: Take derivatives to obtain

$$(\log(x^2+h^2))' \to 2(\log|x|)'.$$

Now recall that this latter distribution equals $2(\mathrm{PV}(1/x))$; see Theorem 2.2. Moreover, $\log(x^2+h^2)$ is a smooth function for fixed $h>0$, so its distributional derivative is its classical derivative, and thus

$$(\log(x^2+h^2))' = \frac{2x}{x^2+h^2}.$$

We earlier defined convolutions, and one of their main uses (in the classical setting) is the approximation of general functions by nice ones. The convolution from Definition 2.16 can be put to similar use in the realm of distributions.

Theorem 3.4. *Let $u \in \mathcal{D}'(\mathbb{R})$, and fix a $\varphi \in C_0^\infty(\mathbb{R})$ with $\int \varphi = 1$. Put $\varphi_n(x) = n\varphi(nx)$. Then $u * \varphi_n \to u$ in $\mathcal{D}'$.*

We originally defined $u * \varphi$ as a *function* (not a distribution), but we then saw that this is a smooth function, so is in particular locally integrable and can be viewed as a distribution after all. This is of course the interpretation of $u * \varphi_n$ that we need to adopt here, for the statement to make sense.

So we see that any distribution can be approximated by smooth functions. We can do even better:

Exercise 3.4. (a) Fix a $\psi \in C_0^\infty(\mathbb{R})$ with $\psi(x) = 1$ in a neighborhood of $x = 0$, and let $\psi_n(x) = \psi(x/n)$. Show that $\psi_n u \to u$ for any $u \in \mathcal{D}'$.

(b) Combine this with Theorem 3.4 to conclude that for any $u \in \mathcal{D}'(\mathbb{R})$, there are $\varphi_n \in C_0^\infty(\mathbb{R})$, $\varphi_n \to u$ in $\mathcal{D}'$.

Sketch of proof. Let ψ be a test function. We want to show that $(u * \varphi_n, \psi) \to (u, \psi)$ or, if we write this out,

$$n \int_{-\infty}^{\infty} (u, \varphi(n(x-t))\psi(x)\, dx \to (u, \psi).$$

Here it is again understood that u is applied to $\varphi(n(x-t))$ as a function of t (not x). I would now like to rewrite the left-hand side as

$$\left(u, n \int_{-\infty}^{\infty} \varphi(n(x-t))\psi(x)\, dx \right); \tag{3.1}$$

this looks plausible, based on the linearity of u. To justify this step, we would have to write the integral as a limit of Riemann sums, and the limit must take place in $\mathcal{D}$, in the sense of Exercise 1.1, so that we can then refer to the continuity of u. I'll leave the matter at that.

The final step then is to note that $n\int \varphi(n(x-t))\psi(x)\,dx \to \psi(x)$. Again, this looks exceedingly plausible; for example, it is clear that we will have pointwise (in fact: uniform) convergence, but even that isn't quite good enough here because we want to refer to Exercise 1.1 one more time and so need to establish convergence in the sense of condition (b) from that Exercise. I will again skip all the details.

If all this is granted, then it indeed follows that (3.1) converges to (u,ψ), and we are done.

(Note that even if this seemed unconvincing, the general plan of the proof was simple and straightforward: move the operation of convolving with the approximate identity φ_n over to the test function, and use that convergence on that level is clear and that u is continuous.) □

Let's try this out for $u=\delta$: We obtain approximations $u_n = \delta * \varphi_n$, and what are these equal to? In fact, what is $\delta * \psi$? By the definition of the convolution, $(\delta * \psi)(x) = (\delta, \psi(x-t)) = \psi(x)$. So we simply recover the approximations $\varphi_n \to \delta$ that we already discussed above, in Exercise 3.2.

We have also inadvertently proved the interesting fact that $\delta * \varphi = \varphi$, that is, δ acts as the identity element for the convolution product.

Exercise 3.5. Recall the definition of the translation $\tau_h u$ of a distribution from Section 2. Then show that for any $u \in \mathcal{D}'(\mathbb{R})$,

$$\frac{1}{h}(\tau_{-h}u - u) \to u'$$

as $h \to 0$ in $\mathcal{D}'$. (So distributional derivatives can also be computed as limits of difference quotients.)

4. Tempered distributions

The *Fourier transform* of an $f \in L^1(\mathbb{R}^d)$ is defined as

$$\widehat{f}(t) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i t\cdot x}\,dx. \tag{4.1}$$

The *Fourier inversion theorem* says that if $f, \widehat{f} \in L^1$, then

$$f(x) = \int_{\mathbb{R}^d} \widehat{f}(t)e^{2\pi i t\cdot x}\,dt.$$

We would now like to define the Fourier transform $\widehat{u}$ also for distributions $u \in \mathcal{D}'$. We follow the usual strategy of moving the operation over to the test function, taking the case of a nice function u as our guideline. In this situation,

$$\int_{\mathbb{R}^d} \widehat{u}(x)\varphi(x)\, dx = \int_{\mathbb{R}^d} dx\, \varphi(x) \int_{\mathbb{R}^d} dt\, u(t)e^{-2\pi i t\cdot x}$$
$$= \int_{\mathbb{R}^d} dt\, u(t) \int_{\mathbb{R}^d} dx\, \varphi(x)e^{-2\pi i t\cdot x} = (u, \widehat{\varphi}),$$

so the definition we want to make is $(\widehat{u}, \varphi) = (u, \widehat{\varphi})$. Unfortunately, this doesn't work: $\widehat{\varphi}$ is not guaranteed to be a test function. It will be smooth, but in fact $\widehat{\varphi}$ will never have compact support, unless $\varphi \equiv 0$.

In other words, the Fourier transform is not an operation on $\mathcal{D}$, and this suggests the path we are going to take: we need a new test function space, and the appropriate one for our current purposes is the *Schwartz space* $\mathcal{S}(\mathbb{R}^d)$. It is defined as follows:

$$\mathcal{S}(\mathbb{R}^d) = \left\{ f : \mathbb{R}^d \to \mathbb{C} : f \in C^\infty, \|f\|_{j,n} < \infty \text{ for all } j, n \geq 0 \right\},$$
$$\|f\|_{j,n} = \sup_{x\in\mathbb{R}^d, |\alpha|\leq n} (1 + |x|^j)|\partial^\alpha f(x)|$$

So $\mathcal{S}$ is the space of all smooth functions that together with all their derivatives decay faster than any power. Clearly, $C_0^\infty \subseteq \mathcal{S}$, but $\mathcal{S}$ also contains (many) not compactly supported functions such as $f(x) = e^{-x^2}$.

The norms $\|\cdot\|_{j,n}$ can be combined into a metric d, for example as follows:

$$d(f, g) = \sum_{j,n\geq 0} 2^{-j-n} \frac{\|f - g\|_{j,n}}{1 + \|f - g\|_{j,n}}$$

Exercise 4.1. Show that $d(f_k, f) \to 0$ if and only if $\|f_k - f\|_{j,n} \to 0$ for all $j, n \geq 0$.

That we have made the right choice is confirmed by

Theorem 4.1. *The Fourier transform is a continuous bijection on* $\mathcal{S}$.

I don't want to prove this in detail here, so will just make a few general remarks. The proof is straightforward, but somewhat tedious to write down. Essentially, we have to show that the Fourier transform maps $\mathcal{S}$ continuously back into itself; the rest will then follow quickly from Fourier inversion. The two defining properties, rapid decay and smoothness, are dual to each other in the sense that one will become the other after taking Fourier transforms, so the combination of both is preserved. If this is done carefully, we will then also obtain control on $\|\widehat{f}\|_{j,n}$ in terms of the $\|f\|_{j',n'}$, and this gives the asserted continuity of the Fourier transform.

Definition 4.2. A *tempered distribution* $u \in \mathcal{S}'$ is a continuous linear functional $u : \mathcal{S} \to \mathbb{C}$. Here, continuity means that there are $C \geq 0$ and $j, n \geq 0$ such that $|u(\varphi)| \leq C\|\varphi\|_{j,n}$.

Exercise 4.2. Let $u : \mathcal{S} \to \mathbb{C}$ be a linear functional on $\mathcal{S}$. Show that the following are equivalent: (a) $u \in \mathcal{S}'$; (b) If $\varphi_n \in \mathcal{S}$ and $\|\varphi\|_{j,N} \to 0$ for all $j, N \geq 0$, then $u(\varphi_n) \to 0$; (c) u is continuous with respect to d.

Of course, the reason for doing all this was to have the following definition available.

Definition 4.3. Let $u \in \mathcal{S}'$. Then the *Fourier transform* $\widehat{u} \in \mathcal{S}'$ is defined as $(\widehat{u}, \varphi) = (u, \widehat{\varphi})$.

Note that this does define a tempered distribution: the map $\varphi \mapsto u(\widehat{\varphi})$ is continuous, being the composition of the two continuous maps $\varphi \mapsto \widehat{\varphi}$ and u.

As before, a function $f \in L^1$ generates a tempered distribution (which we'll frequently identify with f) $(f, \varphi) = \int f\varphi$. Indeed, this map is obviously linear, and $|(f, \varphi)| \leq (\int |f|) \sup |\varphi| \leq (\int |f|)\|\varphi\|_{0,0}$. We repeat the motivating calculation we already did at the beginning of this section to confirm that the distributional Fourier transform $\widehat{f} \in \mathcal{S}'$ satisfies

$$\begin{aligned}(\widehat{f}, \varphi) &= \int f(t)\widehat{\varphi}(t)\, dt = \int dt\, f(t) \int dx\, \varphi(x) e^{-2\pi i t \cdot x} \\ &= \int dx\, \varphi(x) \int dt\, f(t) e^{-2\pi i t \cdot x}.\end{aligned}$$

In other words, $\widehat{f}$ is a function, and this function is still given by the original definition (4.1).

Next, let's look at $u = \delta$. First of all, observe that $|\delta(\varphi)| = |\varphi(0)| \leq \|\varphi\|_{0,0}$, so $\delta \in \mathcal{S}'$. (Strictly speaking, we have two δ's at this point: the original distribution $\delta \in \mathcal{D}'$, and its extension to $\mathcal{S}$, which is the tempered distribution $\delta \in \mathcal{S}'$. This confusing situation will be rectified in a moment; let's not get distracted by it right now.) From the definition of the Fourier transform, we have $\widehat{\delta}(\varphi) = \delta(\widehat{\varphi}) = \widehat{\varphi}(0)$, but this is $\int \varphi(x)\, dx$, by (4.1). In other words, $\widehat{\delta} = 1$ (the function $f(x) \equiv 1$).

Similarly, $\widehat{1}(\varphi) = \int \widehat{\varphi} = \varphi(0)$, by Fourier inversion. We summarize:

Theorem 4.4. *We have* $\widehat{\delta} = 1$, $\widehat{1} = \delta$ *in* $\mathcal{S}'(\mathbb{R}^d)$.

We could also have tried to argue based on (4.1) that

$$\widehat{\delta}(t) = \int \delta(x) e^{-2\pi i t x}\, dx = e^{-2\pi i t x}\big|_{x=0} = 1,$$

and this gives the correct answer, but is not a formally correct manipulation because we tried to apply δ to $e^{-2\pi itx}$, which is not a test function.

Now let's take a closer look at the relation between $\mathcal{S}'$ and $\mathcal{D}'$. We already observed that $\mathcal{D} \subsetneq \mathcal{S}$. Because of this, any tempered distribution $u : \mathcal{S} \to \mathbb{C}$ can be restricted to $\mathcal{D}$ (and don't confuse this type of restriction with the one from Definition 2.12). Then $u|_{\mathcal{D}}$ will lie in $\mathcal{D}'$: By assumption, we have $|u(\varphi)| \le C\|\varphi\|_{j,n}$ for some j, n. If now also $\varphi \in \mathcal{D}$, $\operatorname{supp}\varphi \subseteq B(0,R)$, then $\|\varphi\|_{j,n} \le (1 + R^j)\|\varphi\|_n$, and this latter norm is the one from Definition 1.1. This verifies that $u \in \mathcal{D}'$, as claimed; more precisely, it is the *restriction* $u|_{\mathcal{D}}$ that is in $\mathcal{D}'$, but we will often not make this distinction explicit in the notation.

Moreover, the original *tempered* distribution $u \in \mathcal{S}'$ can be recovered from its restriction $u|_{\mathcal{D}} \in \mathcal{D}'$. This follows from the fact that $\mathcal{D} \subseteq \mathcal{S}$ is dense in $\mathcal{S}$ with respect to d.

Exercise 4.3. Prove this. More explicitly, suppose that $\varphi \in \mathcal{S}$, and then show that there are $\varphi_k \in C_0^\infty$ such that $\|\varphi_k - \varphi\|_{j,n} \to 0$ for each $j, n \ge 0$. *Suggestion:* Fix a $\psi \in C_0^\infty$ with $\psi = 1$ near $x = 0$ and then try $\varphi_k(x) = \psi(x/k)\varphi(x)$.

However, if we conversely start out with a $u \in \mathcal{D}'$, then this may or may not have an extension to a tempered distribution $u_0 \in \mathcal{S}'$; if it does, then u_0 is unique, as we just discussed. As an example for such a $u \in \mathcal{D}'$ with no continuous linear extension, we can consider the function $u(x) = e^x$. So $u(\varphi) = \int e^x\varphi(x)\,dx$, and it is already clear intuitively what will go wrong here: there seems to be only one natural way to attempt an extension to $\mathcal{S}$, namely, try to use this formula for general $\varphi \in \mathcal{S}$ also, but then such a function is not guaranteed to have enough decay to make the integral convergent.

To make a proof out of this intuition, we need to proceed differently. Fix a $\psi \in C_0^\infty(\mathbb{R})$ with $\int \psi \ne 0$, and let $\varphi_n(x) = \psi(x-n)e^{-x} \in C_0^\infty \subseteq \mathcal{S}$. Then $\|\varphi_n\|_{j,N} \to 0$.

Exercise 4.4. Prove this in detail.

However, $u(\varphi_n) = \int \psi(x-n)\,dx = \int \psi(x)\,dx$ does not converge to zero. So u has no continuous extension to $\mathcal{S}$.

What we actually showed is that u is already discontinuous on $\mathcal{D}$, at $\varphi = 0$, when we use the metric d on $\mathcal{D}$. So the general message is that the continuity requirement on a tempered distribution $u \in \mathcal{S}'$ is stronger than the one on a $u \in \mathcal{D}'$ (when only test functions $\varphi \in \mathcal{D}$ are considered). Let's summarize:

Theorem 4.5. *The restriction of a tempered distribution to $\mathcal{D}$ is a distribution. Conversely, a distribution may or may not have a continuous linear extension to $\mathcal{S}$; if it does, then the extension is unique.*

It is best to think of this as saying that every tempered distribution is a distribution, but, conversely, only some distributions are also tempered (in this version, we don't distinguish between a distribution and its restriction or extension, which is justified by the uniqueness of these). Symbolically, we can write $\mathcal{S}' \subsetneq \mathcal{D}'$, which is a good way to commit this whole discussion to memory, as long as you don't take this inclusion literally.

The obstacle illustrated by the example $u(x) = e^x \in \mathcal{D}'$, $u \notin \mathcal{S}'$ is the only one that can prevent a distribution from being tempered: too rapid growth near infinity. For example, if $f \in L^1_{\text{loc}}$ has at most polynomial growth, $|f(x)| \le C(1+|x|^N)$ for some $N \ge 0$, then $f \in \mathcal{S}'$.

Exercise 4.5. Prove this.

Exercise 4.6. Let $f \in L^p(\mathbb{R}^d)$, for some $1 \le p \le \infty$. Show that $f \in \mathcal{S}'$.

Another statement along these lines can be established with the help of the notion of the support of a distribution.

Definition 4.6. Let $u \in \mathcal{D}'(U)$. The *support* of u is defined as

$$\operatorname{supp} u = \left(\bigcup V \right)^c ,$$

where the union is over all open $V \subseteq U$ with $u|_V = 0$.

Exercise 4.7. Show that if $f \in C(\mathbb{R}^d)$, then this agrees with the usual notion of the support of a function as the closure of the set $\{x : f(x) \ne 0\}$.

By Theorem 2.13, if $W = (\operatorname{supp} u)^c$, then $u|_W = 0$. Conversely, if $u|_W = 0$ for an open set W, then $W \cap \operatorname{supp} u = \emptyset$, by the definition of the support. This gives a description of the complement of the support as the largest open set on which u is zero.

Example 4.1. $\operatorname{supp} \delta = \{0\}$: indeed, $\delta\big|_{\mathbb{R}\setminus\{0\}} = 0$, as we observed earlier. On the other hand, if U is open and $0 \in U$, then there are $\varphi \in C_0^\infty(U)$ with $\varphi(0) \ne 0$, so $0 \in \operatorname{supp} \delta$.

Theorem 4.7. *Suppose that $u \in \mathcal{D}'(\mathbb{R}^d)$ has compact support. Then $u \in \mathcal{S}'$.*

Here, I have again applied the convenient convention of not distinguishing between $u \in \mathcal{D}'$ and its unique extension to $\mathcal{S}'$. A more explicit version of the statement would be: $u \in \mathcal{D}'$ has a continuous linear extension to $\mathcal{S}$.

Sketch of proof. Fix a $\psi \in C_0^\infty$ with $\psi = 1$ on an open set containing $\operatorname{supp} u$. Then $u = \psi u$ since for any $\varphi \in \mathcal{D}$, we can write

$$\varphi = \psi\varphi + (1-\psi)\varphi,$$

and the second function on the right-hand side is annihilated by u because its support is contained in $(\operatorname{supp} u)^c$. Now $\operatorname{supp}(\psi\varphi) \subseteq K$ for a fixed compact K, for all $\varphi \in \mathcal{D}$, and thus $|u(\varphi)| \le C\|\psi\varphi\|_N$ (the point here is that we have one C and one N that work for all $\varphi \in \mathcal{D}$).

Now ψ and all its derivatives are bounded, so $\|\psi\varphi\|_N \le D\|\varphi\|_{0,N}$, and then the inequality $|u(\varphi)| \le B\|\varphi\|_{0,N}$ allows us to extend u continuously to $\mathcal{S}$, by approximating a general $\theta \in \mathcal{S}$ by $\varphi_n \in \mathcal{D}$, so $\|\varphi_n - \theta\|_{j,k} \to 0$. We can then define $u(\theta) = \lim u(\varphi_n)$. (It needs to be checked that this limit exists and is independent of the approximating sequence, but I don't want to go into these details here.) □

Theorem 4.8. *Suppose that* $u \in \mathcal{S}'(\mathbb{R})$, $g \in C^\infty(\mathbb{R})$, $|g^{(j)}(x)| \le C(1+|x|)^{N_j}$ $(j \ge 0)$. *Then also* $u', gu \in \mathcal{S}'$.

Strictly speaking, we haven't even defined these operations for *tempered* distributions yet, but of course there are two obvious answers to this complaint. We can just mimic the old definitions (and then Theorem 4.8 is claiming that everything is well defined in $\mathcal{S}'$ also), or we refer to Theorem 4.5 and proceed as follows: consider the restriction $u \in \mathcal{D}'$ of $u \in \mathcal{S}'$, and take its distributional derivative (or multiply it by g). Then the claim of Theorem 4.8 is that the new distributions obtained in this way are also tempered.

Proof. To establish the (important) claim about u', we only need to observe that $\|\varphi'\|_{j,N} \le \|\varphi\|_{j,N+1}$. So if $|(u,\varphi)| \le C\|\varphi\|_{j,N}$, then also

$$|(u', \varphi_n)| = |(u, \varphi_n')| \le C\|\varphi\|_{j,N+1},$$

as required.

As for the product gu, we can similarly observe that $\|g\varphi\|_{j,n} \le D\|\varphi\|_{j+M,n}$, with $M = \max_{0\le k\le n} N_k$, and then argue as above. □

While distributions that do not grow too fast near infinity will be tempered, the converse is not true. Consider the function $f(x) = e^x \cos e^x$. This does become large (some of the time, at least) for large x. However, we have $f \in \mathcal{S}'$ anyway, and this follows from Theorem 4.8 because $f = g'$, with $g = \sin e^x \in \mathcal{S}'$.

Definition 4.9. Let $u_n, u \in \mathcal{S}'$. We say that $u_n \to u$ in $\mathcal{S}'$ if $(u_n, \varphi) \to (u, \varphi)$ for all $\varphi \in \mathcal{S}$.

Of course, this is the expected analog of Definition 3.1. As a consequence, if $u_n \to u$ in $\mathcal{S}'$, then also $u_n \to u$ in $\mathcal{D}'$ because this is the same condition, except that we now only test on $\varphi \in \mathcal{D}$. There is a pitfall that must be avoided in this context: if $u_n, u \in \mathcal{S}'$ and $u_n \to u$ in $\mathcal{D}'$, then it does *not* follow that $u_n \to u$ also in $\mathcal{S}'$. In other words, convergence in $\mathcal{S}'$ is a stronger condition than convergence in $\mathcal{D}'$, and it does not simply follow from the fact that all the distributions involved are tempered.

For a concrete example, consider the tempered distributions $u_n = e^{n^2}\delta_n$. As elements of $\mathcal{D}'$, these converge to (the tempered distribution) $u = 0$ in $\mathcal{D}'$ (why?), but $u_n \not\to 0$ in $\mathcal{S}'$ as $(u_n, e^{-x^2}) = 1$.

Let's now develop the theory of the Fourier transform in $\mathcal{S}'$ a bit. We define the *convolution* of a tempered distribution $u \in \mathcal{S}'$ with a test function $\varphi \in \mathcal{S}$ as expected as

$$(u * \varphi)(x) = (u, \tau_x R\varphi).$$

This has the same basic properties as before: $u*\varphi \in C^\infty$ and $(u*\varphi)' = u' * \varphi = u * \varphi'$. Another important observation is that $|(u * \varphi)(x)| \le C(1 + |x|)^N$ for some C, N, so $u * \varphi$ defines a tempered distribution itself. To prove this estimate, observe that

$$1 + |t| = 1 + |t - x + x| \le (1 + |t - x|)(1 + |x|) \tag{4.2}$$

and thus

$$|(u * \varphi)(x)| \le C\|\varphi(x - t)\|_{j,N} \le C(1 + |x|)^j\|\varphi\|_{j,N},$$

since the first norm (which must be taken of $\varphi(x - t)$ as a function of t, for fixed x) involves suprema over t of expressions of the form $(1 + |t|)^j|\varphi^{(n)}(x - t)|$, with $0 \le n \le N$, and we then use (4.2).

Proposition 4.10. *Let $\varphi, \psi \in \mathcal{S}(\mathbb{R})$, $a \in \mathbb{R}$. Then: (a) $(\widehat{\varphi})\widehat{\ } = R\varphi$; (b) $\widehat{\varphi'(t)} = 2\pi i t\widehat{\varphi}(t)$; (c) $(-2\pi i x\varphi(x))\widehat{\ }(t) = \widehat{\varphi}'(t)$; (d) $(\tau_a\varphi)\widehat{\ }(t) = e^{-2\pi i a t}\widehat{\varphi}(t)$; (e) $(e^{2\pi i a x}\varphi)\widehat{\ } = \tau_a\widehat{\varphi}$; (f) $(\varphi * \psi)\widehat{\ } = \widehat{\varphi}\widehat{\psi}$*

Part (a) is Fourier inversion, and the other parts follow from quick calculations.

Theorem 4.11. *Let $u \in \mathcal{S}'$, $\varphi \in \mathcal{S}$. Then u also satisfies (a)-(e), and $(u * \varphi)\widehat{\ } = \widehat{\varphi}\widehat{u}$.*

We have not yet formally introduced all the operations on tempered distributions that are involved here, but it is of course clear how to proceed. We define $(Ru, \varphi) = (u, R\varphi)$ and, as before, $(\tau_a u, \varphi) = (u, \tau_{-a}\varphi)$.

Parts (a)-(e) prove themselves if we just move the operations over to the test function and then refer to Proposition 4.10. Let me do part

(d) as an illustration:

$$((\tau_a u)\widehat{\ }, \varphi) = (\tau_a u, \widehat{\varphi}) = (u, \tau_{-a}\widehat{\varphi})$$

This test function equals $(e^{-2\pi iax}\varphi)\widehat{\ }$, by Proposition 4.1(e), and unpacking this, we then obtain

$$(u, (e^{-2\pi iax}\varphi)\widehat{\ }) = (\widehat{u}, e^{-2\pi iax}\varphi) = (e^{-2\pi iax}\widehat{u}, \varphi).$$

Since this is true for any $\varphi \in \mathcal{S}$, we have $(\tau_a u)\widehat{\ } = e^{-2\pi iax}\widehat{u}$, as claimed.

Only part (f) presents some technical challenges, similar actually to the ones from the proof of Theorem 3.4. Let $\psi \in \mathcal{S}$ be an arbitrary test function. Then

$$((u * \varphi)\widehat{\ }, \psi) = (u * \varphi, \widehat{\psi}) = \int \widehat{\psi}(x)(u, \varphi(x-t))\, dx.$$

I will again rewrite this as $(u, \int \widehat{\psi}(x)\varphi(x-t)\, dx)$ without fully justifying this step. We can then interpret $\int \widehat{\psi}(x)\varphi(x-t)\, dx = (\widehat{\psi} * R\varphi)(t)$, and Fourier inversion says that the inverse operation of $\widehat{\ }$ is $R\widehat{\ }$. Thus $\widehat{\psi} * R\varphi = \widehat{f}$, with

$$f = R(\widehat{\psi} * R\varphi)\widehat{\ } = R(\widehat{\widehat{\psi}} \cdot (R\varphi)\widehat{\ }) = \psi R(R\varphi)\widehat{\ }.$$

Putting things together, we deduce that

$$((u * \varphi)\widehat{\ }, \psi) = (\widehat{u}, \psi R(R\varphi)\widehat{\ }) = (R(R\varphi)\widehat{\ } u, \psi),$$

and since $R(R\varphi)\widehat{\ } = \widehat{\varphi}$, by a calculation, this finally gives (f).

Let's run a quick check on this result. We already know that $\delta * \varphi = \varphi$. Taking Fourier transforms, we obtain $\widehat{\varphi}\widehat{\delta} = \widehat{\varphi}$, and since $\widehat{\delta} = 1$, everything is indeed in perfect order.

Theorem 4.12. *Suppose that $u_n \to u$ in $\mathcal{S}'$. Then also $\widehat{u_n} \to \widehat{u}$ in $\mathcal{S}'$.*

Exercise 4.8. Prove Theorem 4.12.

Recall now Theorem 3.3. What do we obtain from the identity

$$\lim_{h\to 0+} \frac{1}{x - ih} = \text{PV}\frac{1}{x} + i\pi\delta \qquad (4.3)$$

after taking Fourier transforms?

Exercise 4.9. Show that all distributions are tempered and that (4.3) also holds in $\mathcal{S}'$.

The distribution $1/(x - ih)$ is a function for fixed $h > 0$, so it would be nice if we could compute its Fourier transform from (4.1), but this fails because $1/(x - ih) \notin L^1$. We can actually make good use of

Theorem 4.12 right away and get around this difficulty by observing that

$$\lim_{L\to\infty} \chi_{(-L,L)}(x)\frac{1}{x-ih} = \frac{1}{x-ih}$$

in $\mathcal{S}'$; this is a routine application of DC, but do it in more detail please if it's not immediately clear to you. Now this cut off function is integrable, so its Fourier transform is given by

$$\int_{-L}^{L} \frac{e^{-2\pi itx}}{x-ih}\, dx. \tag{4.4}$$

This integral can be evaluated in the limit $L\to\infty$, most conveniently by a typical application of the residue calculus. Let's say $t<0$. We then close the contour $[-L,L]$ by a semicircle in the upper half plane, which can be parametrized as $\gamma(s)=Le^{is}$, $0\le s\le \pi$. This part of the integral can thus be estimated by

$$\frac{L}{L-h}\int_0^{\pi} e^{2\pi tL\sin s}\, ds, \tag{4.5}$$

and this goes to zero as $L\to\infty$, by DC (recall that $t<0$ currently). The only singularity of the integrand from (4.4) occurs at $x=ih$, thus

$$\lim_{L\to\infty}\int_{-L}^{L} \frac{e^{-2\pi itx}}{x-ih}\, dx = 2\pi i e^{2\pi ht}.$$

A similar calculation is possible for $t>0$, and in this case we find that

$$\lim_{L\to\infty}\int_{-L}^{L} \frac{e^{-2\pi itx}}{x-ih}\, dx = 0$$

(because we now close the contour in the lower half plane, to make the exponential small, and there will be no singularities inside our region this time).

So, summing up, we have shown that

$$\lim_{L\to\infty}\int_{-L}^{L} \frac{e^{-2\pi itx}}{x-ih}\, dx = \begin{cases} 2\pi i e^{2\pi ht} & t<0 \\ 0 & t>0 \end{cases}.$$

On reflection, we have actually shown this as a pointwise limit, but this is not quite what we need here because what we know is that the cut off integrals will converge to the Fourier transform of $1/(x-ih)$ in $\mathcal{S}'$. However, it is easy to go over the argument one more time and confirm that this follows, too. We have shown that

$$\widehat{\left(\frac{1}{x-ih}\right)}(t) = 2\pi i\chi_{(-\infty,0)}(t)e^{2\pi ht}.$$

When $h \to 0+$, this converges (in $\mathcal{S}'$!) to

$$2\pi i \chi_{(-\infty,0)}(t) = i\pi(1 - \operatorname{sgn}(t)).$$

When we compare this with (4.3), we obtain the following interesting result.

Theorem 4.13. $(\mathrm{PV}(1/x))\widehat{} = -i\pi \operatorname{sgn}(t)$

The operation of convolving with $\mathrm{PV}(1/x)$ is called the *Hilbert transform* of a function, so

$$(Hf)(x) = \lim_{h\to 0+} \int_{|t|>h} \frac{f(x-t)}{t}\, dt = \lim_{h\to 0+} \int_{|t-x|>h} \frac{f(t)}{x-t}\, dt,$$

and in our current context, this is defined for $f \in \mathcal{S}$. Theorem 4.13 shows that $(Hf)\widehat{}(t) = -i\pi \operatorname{sgn}(t)\widehat{f}(t)$.

By Theorem 4.11(c),

$$(x\mathrm{PV}(1/x))\widehat{} = \frac{1}{-2\pi i}(\mathrm{PV}(1/x))\widehat{}' = \frac{1}{2}\operatorname{sgn}' = \delta.$$

This can be confirmed directly since

$$(x\mathrm{PV}(1/x), \varphi) = \lim_{h\to 0+} \int_{|x|>h} \frac{x\varphi(x)}{x}\, dx = \int_{-\infty}^{\infty} \varphi(x)\, dx,$$

so $x\mathrm{PV}(1/x) = 1$, and indeed $\widehat{1} = \delta$, as we saw earlier in Theorem 4.4.

Theorem 4.14. *Let* $u \in \mathcal{D}'(\mathbb{R})$ *and suppose that* $\operatorname{supp} u = \{a\}$. *Then* $u = \sum_{j=0}^{N} c_j \delta_a^{(j)}$.

Proof. For ease of notation, let's assume that $a = 0$. We showed earlier, in the proof of Theorem 4.7, that a compactly supported distribution satisfies

$$|u(\varphi)| \le C\|\varphi\|_N$$

(the point is that a fixed N and C work for all $\varphi \in \mathcal{D}$).

I now claim that if $\varphi \in \mathcal{D}$, $\varphi(0) = \varphi'(0) = \ldots = \varphi^{(N)}(0) = 0$, then $u(\varphi) = 0$. This we can prove as follows. Given such a φ and $\epsilon > 0$, it will of course be true for all sufficiently small $b > 0$ that $|\varphi^{(N)}(x)| < \epsilon$ on $|x| < b$. By successively integrating these derivatives, we then find that also

$$\left|\varphi^{(j)}(x)\right| < \epsilon b^{N-j} \quad \text{for } j = 0, 1, \ldots, N. \tag{4.6}$$

Then we pick a $\psi \in C_0^\infty$ with $0 \le \psi \le 1$, $\operatorname{supp}\psi \subseteq [-1, 1]$ and $\psi = 1$ on a neighborhood of $x = 0$, and we consider $\varphi_n(x) = \psi(nx)\varphi(x)$. These functions will have their supports inside $[-1/n, 1/n]$. We now take

$b = 1/n$ and assume that n is large so that (4.6) becomes available. From the $j = 0$ case, we conclude that $|\varphi_n| \le \epsilon b^N = \epsilon n^{-N}$. Next,

$$\varphi_n'(x) = n\psi'(nx)\varphi(x) + \psi(nx)\varphi'(x),$$

and this gives us a bound of the type

$$|\varphi_n'| \le C\epsilon(nb^N + b^{N-1}) = (C+1)\epsilon n^{1-N}.$$

We can continue in this style and bound the first N derivatives by

$$\left|\varphi_n^{(j)}\right| \le C_j \epsilon n^{j-N}, \quad j = 0, 1, \ldots, N. \tag{4.7}$$

Here, it is important that the constants C_j only depend on ψ (more precisely, on bounds on ψ and its derivatives) and not on ϵ. Stated more succinctly, (4.7) says that $\|\varphi_n\|_N \le C\epsilon$. Since $\varphi_n = \varphi$ near $x = 0$, we have $u(\varphi) = u(\varphi_n)$ and, as we showed, $|u(\varphi_n)| \le D\epsilon$. Since $\epsilon > 0$ was arbitrary, we conclude that $u(\varphi) = 0$, as claimed.

Exercise 4.10. Explain in more detail why $u(\varphi) = u(\varphi_n)$. This is perhaps best done by establishing the following general fact: If $u \in \mathcal{D}'$ has compact support and $\varphi = \psi$ on an open set $V \supseteq \operatorname{supp} u$, then $u(\varphi) = u(\psi)$.

Now pick $\varphi_j \in C_0^\infty$ with $\varphi_j^{(n)}(0) = \delta_{jn}$ for $0 \le j, n \le N$. For example, we could take $\varphi_j(x) = x^j/j!$ near $x = 0$. For arbitrary $\varphi \in \mathcal{D}$, write

$$\varphi = \sum_{j=0}^{N} a_j\varphi_j + \varphi - \sum_{j=0}^{n} a_j\varphi_j \equiv \sum_{j=0}^{N} a_j\varphi_j + \psi,$$

with $a_j = \varphi^{(j)}(0)$ (this is similar to a Taylor expansion of φ, except that the compactly supported functions φ_j take over the role of the polynomials $x^j/j!$). Then $\psi(0) = \psi'(0) = \ldots = \psi^{(N)}(0) = 0$, thus, by what we just proved, $u(\psi) = 0$ and hence

$$u(\varphi) = \sum_{j=0}^{N} u(\varphi_j)a_j.$$

Since $a_j = \varphi^{(j)}(0) = (-1)^j(\delta^{(j)}, \varphi)$ this is what we claimed, with $c_j = (-1)^j(u, \varphi_j)$. □

Theorem 4.15 (Poisson summation formula)**.** *$u = \sum_{n\in\mathbb{Z}} \delta_n \in \mathcal{S}'$ and $\widehat{u} = u$.*

The name *Poisson summation formula* is usually not given to this rather highbrow version, but instead to

Corollary 4.16. *Let $\varphi \in \mathcal{S}(\mathbb{R})$. Then $\sum_{n\in\mathbb{Z}} \varphi(n) = \sum_{n\in\mathbb{Z}} \widehat{\varphi}(n)$.*

This or also more general versions of the statement can be and usually are proved directly.

Exercise 4.11. Prove that $\sum \delta_n \in \mathcal{S}'$.

Proof of Theorem 4.15. Clearly, $u = \sum \delta_n$ has the following two invariance properties: $\tau_n u = u$ and $e^{2\pi i n x} u = u$, for all $n \in \mathbb{Z}$. By taking Fourier transforms, this implies that $v = \widehat{u}$ has the same two properties. In particular, $(e^{2\pi i x} - 1)v = 0$, and this implies that $\operatorname{supp} v \subseteq \mathbb{Z}$, by arguing as follows: Let $V \subseteq \mathbb{R}$ be an open set with $V \cap \mathbb{Z} = \emptyset$, and let $\varphi \in \mathcal{D}$ be arbitrary with $\operatorname{supp} \varphi \subseteq V$. Since $\operatorname{supp} \varphi$ is a *compact* subset of V, it has positive distance to V^c, and thus $\psi = \varphi/(e^{2\pi i x} - 1) \in \mathcal{D}$ also. Thus

$$(v, \varphi) = (v, (e^{2\pi i x} - 1)\psi) = ((e^{2\pi i x} - 1)v, \psi) = 0.$$

We have shown that $V \cap \operatorname{supp} v = \emptyset$, and this holds for any open V with $V \cap \mathbb{Z} = \emptyset$, thus $\operatorname{supp} v \subseteq \mathbb{Z}$, as claimed.

Consider now the restriction v_0 of v to $(-3/4, 3/4)$, say. Actually, we want an element of $\mathcal{D}'(\mathbb{R})$, not of $\mathcal{D}'(-3/4, 3/4)$, so we really take $v_0 = \psi v$ with a $v \in C_0^\infty$ with $\operatorname{supp} \psi \subseteq (-1, 1)$, $\psi = 1$ on $(-3/4, 3/4)$. (The purpose is to make Theorem 4.14 applicable, but this is a technical point and not essential; the argument would also work with the actual restriction of v, if Theorem 4.14 is slightly adapted instead.)

We have $\operatorname{supp} v_0 \subseteq \{0\}$, so

$$v_0 = \sum_{j=0}^{N} c_j \delta^{(j)}, \tag{4.8}$$

by Theorem 4.14. We can say more here: in fact, $v_0 = c\delta$. To see this, assume that $N \geq 1$ and consider a trigonometric polynomial $p(x) = \sum_{|n| \leq L} a_n e^{2\pi i n x}$, and here we want to choose $L \geq 1$ and then the $a_n \in \mathbb{C}$ such that $p(0) = 1$, $p^{(j)}(0) = 0$ for $1 \leq j \leq N - 1$, and $p^{(N)}(0) \neq 0$.

Exercise 4.12. Show that this is possible.

Finally, we build a test function by multiplying p by a $\varphi \in C_0^\infty$. Then

$$(\varphi p)^{(j)}(0) = \varphi^{(j)}(0) \quad (j < N), \quad (\varphi p)^{(N)}(0) = \varphi^{(N)}(0) + p^{(N)}(0)\varphi(0),$$

and thus, if v_0 is as in (4.8) and invariant under multiplication by $e^{2\pi i n x}$, as we know it is, then

$$\begin{aligned}(v_0, \varphi) &= (p v_0, \varphi) = (v_0, p\varphi) \\ &= \sum_{j=0}^{N} c_j (-1)^j \varphi^{(j)}(0) + c_N (-1)^N p^{(N)}(0)\varphi(0).\end{aligned}$$

On the other hand, if we simply apply v_0 to φ directly, then we see that this must also be equal to $\sum_{j=0}^{N} c_j(-1)^j\varphi^{(j)}(0)$. It follows that $c_N = 0$. The whole argument can then be repeated to see that also $c_{N-1} = 0$ and so on.

So, reconstructing the distribution from its restrictions, based on Theorem 2.13, we conclude that $v = \sum c_n\delta_n$. Since $\tau_n v = v$, the c_n must actually be independent of n, so $v = c\sum \delta_n = cu$. There are $\varphi \in \mathcal{S}$ with $\widehat{\varphi} = \varphi$, for example $\varphi(x) = e^{-\pi x^2}$, and applying both u and $v = \widehat{u}$ to such a φ shows that $c = 1$. □

Theorem 4.17. *Let $0 < p < d$. Then $u = |x|^{-p} \in \mathcal{S}'(\mathbb{R}^d)$ and $\widehat{u} = c_{p,d}|t|^{p-d}$.*

Formally, this looks easy, as

$$\int_{\mathbb{R}^d} |x|^{-p} e^{-2\pi i t\cdot x}\, dx = \int_0^\infty dr\, r^{d-1-p} \int_{v\in S^{d-1}} d\omega(v)\, e^{-2\pi i r v\cdot t},$$

if the integration is done in spherical coordinates, with $d\omega$ denoting the surface measure on the unit sphere. By the spherical symmetry of this measure, the second integral, which we could denote by $\widehat{\omega}$, does not depend on the *direction* of t, so $\widehat{\omega} = \widehat{\omega}(r|t|)$ and thus, after the substitution $r|t| = s$,

$$\int_{\mathbb{R}^d} |x|^{-p} e^{-2\pi i t\cdot x}\, dx = |t|^{p-d} \int_0^\infty \widehat{\omega}(s) s^{d-1-p}\, ds,$$

as claimed (with $c_{p,d}$ apparently given by the integral).

Of course, none of this makes rigorous sense in this form since $|x|^{-p} \notin L^1(\mathbb{R}^d)$, so the very first integral was undefined, and we cannot use (4.1) here. Also, the last integral is not guaranteed to converge. The Fourier transform of the surface measure has asymptotics $|\widehat{\omega}(s)| \simeq s^{(1-d)/2}$, so our answer only makes sense if $p > (d+1)/2$. (The decay of $\widehat{\omega}$ in dimension $d > 1$, by the way, is an initially rather surprising, but well studied phenomenon.)

Still, the effort was not completely wasted because one useful general conclusion can be drawn: the computation of $\widehat{u}$ seems to depend mainly on the symmetry of u (more precisely, its spherical symmetry and homogeneity), and indeed this is exactly how the rigorous argument will proceed.

Sketch of proof. The function $u(x) = |x|^{-p}$ is locally integrable and bounded near infinity, so $u \in \mathcal{S}'$.

In Exercise 2.5, you defined the dilation $\delta_a u$ of a distribution u as $(\delta_a u, \varphi) = a^{-d}(u, \delta_{1/a}\varphi)$, and $(\delta_b\varphi)(x) = \varphi(bx)$. This is done in such a way that if u is a function, then $\delta_a u = u(ax)$. We now call $u \in \mathcal{D}'$

homogeneous of degree r if $\delta_a u = a^r u$ for all $a > 0$. So if u is a function, then this means that $u(ax) = a^r u(x)$, which is the familiar definition of homogeneity for functions.

In now claim that $\widehat{u} \in \mathcal{S}'$ is homogeneous of degree $p-d$. We confirm this by a calculation, which will actually establish the general fact that if $v \in \mathcal{S}'$ is homogeneous of degree r, then $\widehat{v}$ will be homogeneous of degree $-r-d$.

We have

$$(\delta_a \widehat{u}, \varphi) = a^{-d}(\widehat{u}, \delta_{1/a}\varphi) = a^{-d}(u, (\delta_{1/a}\varphi)\widehat{\ }).$$

It is easy to see, along the lines of what we did in Proposition 4.1, that $a^{-d}(\delta_{1/a}\varphi)\widehat{\ } = \delta_a \widehat{\varphi}$.

Exercise 4.13. Prove this in detail.

Thus

$$(\delta_a \widehat{u}, \varphi) = a^{-d}(\delta_{1/a} u, \widehat{\varphi}) = a^{p-d}(u, \widehat{\varphi}) = a^{p-d}(\widehat{u}, \varphi)$$

Here, the second equality follows because the function $u(x) = |x|^{-p}$ is clearly homogeneous of degree $-p$.

In the remainder of the argument, I will discuss explicitly only the case $d = 1$. As our next step, we then show that $\widehat{u}$ is also invariant under reflection in the sense that $R\widehat{u} = \widehat{u}$, with Rv defined as $(Rv, \varphi) = (v, R\varphi)$ and, as above, $(R\varphi)(x) = \varphi(-x)$.

Exercise 4.14. Prove that $R\widehat{u} = \widehat{u}$, by a calculation similar to the one just given.

Now if we already knew that $\widehat{u} = f$ is a function $f \in L^1_{\text{loc}}$, then everything seems clear: homogeneity and reflection invariance say that $f(x) = |x|^{p-1} f(1)$, which is our claim, with $c_{p,1} = f(1)$.

Exercise 4.15. Actually, things are not quite as easy because a distribution that is a function determines its function only almost everywhere, so homogeneity for example only says that $f(ax) = f(x)$ off a null set, which could depend on $a > 0$. Provide the missing details.

We are in this situation for $1/2 < p < 1$ because we can then write $|x|^{-p} = \chi_{\{|x|<1\}}|x|^{-p} + \chi_{\{|x|>1\}}|x|^{-p}$ as the sum of two functions, one in L^1 and one in L^2. The Fourier transform of an L^1 is a function (in C_0), and so is the Fourier transform of an L^2 function (this is in L^2 also).

If, on the other hand, $0 < p < 1/2$, then $1-p$ is in the range just discussed, so $(|x|^{1-p})\widehat{\ } = c|x|^{-p}$ and then Fourier inversion or rather the fact that a tempered distribution is uniquely determined by its Fourier transform gives the claim in this case also.

This leaves only the case $p = 1/2$, which we can handle by a limiting argument, sending $p \to 1/2$, $p \neq 1/2$.

Finally, this argument also works in higher dimensions if we establish one additional symmetry before entering the final part of the argument: $\widehat{u}$ is spherically symmetric. □

Exercise 4.16. Give a precise definition of what is meant by this final claim, and then prove it.

With additional trickery, the constants $c_{p,d}$ can actually be identified. One finds

$$c_{p,d} = \pi^{-d/2+p} \frac{\Gamma((d-p)/2)}{\Gamma(p/2)}. \tag{4.9}$$

The differential operator

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_d^2}$$

is called the *Laplacian.* The corresponding partial differential equation $\Delta u = 0$ is called the *Laplace equation,* and its solutions are also called *harmonic* functions or distributions, if we admit distributional solutions.

Theorem 4.18. *Let $u \in \mathcal{S}'$, $\Delta u = 0$. Then u is a (harmonic) polynomial.*

The assumption that the distribution is *tempered* is crucial here. The theorem does not hold for general distributions, and in fact there are many counterexamples that are smooth functions such as $u(x, y) = e^x \cos y$.

Exercise 4.17. Find all $u \in \mathcal{D}'(\mathbb{R})$ with $\Delta u = 0$.

Corollary 4.19 (Liouville)**.** *A bounded harmonic function is constant.*

Proof. Such a function u can be viewed as a tempered distribution, so is a polynomial by Theorem 4.18, but the only bounded polynomials are the constants. □

In the same way we can establish the more general statement that a harmonic function u of at most polynomial growth, $|u(x)| \lesssim (1+|x|)^N$, actually is a polynomial of degree at most N.

Proof of Theorem 4.18 (sketch). Take Fourier transforms in $\Delta u = 0$ and use Theorem 4.11 to conclude that

$$-4\pi^2 |t|^2 \widehat{u} = 0. \tag{4.10}$$

Note that $|t|^2 = t_1^2 + \ldots + t_d^2$ is a smooth function satisfying the assumptions of Theorem 4.8, so $|t|^2\widehat{u} \in \mathcal{S}'$ is well defined.

Since $1/|t|^2$ is smooth away from $t = 0$, (4.10) implies that $\operatorname{supp} u \subseteq \{0\}$. We now need the multi-dimensional analog of Theorem 4.14. Assuming this, we obtain

$$\widehat{u} = \sum_{|\alpha| \le N} c_\alpha \frac{\partial^\alpha}{\partial t^\alpha}\delta,$$

and then we take (inverse) Fourier transforms to deduce that $u = \sum c_\alpha (-2\pi i)^{|\alpha|} x^\alpha$. □

A *Green function* is a $G \in \mathcal{S}'$ with $\Delta G = \delta$. Usually, one imposes additional conditions on the asymptotics to make G unique and can then meaningfully speak of *the* Green function (assuming for now its existence, which we will establish in a moment). The G I will construct below is the usual choice. Without such extra requirements, G is clearly not unique since we can add an arbitrary harmonic $u \in \mathcal{S}'$ to a given G.

Such a G is an interesting object. For example, it allows us to construct fairly explicit solutions to the *Poisson equation* $\Delta u = f$, at least for $f \in \mathcal{S}$, as $u = G * f$. Indeed, $\Delta(G * f) = (\Delta G) * f = \delta * f = f$, as desired.

Theorem 4.20. *For $d \ge 3$, we can take $G(x) = -c_d|x|^{2-d}$.*

In particular, G indeed is a function (not just a tempered distribution). The constants can be identified as

$$c_d = \frac{\Gamma(-1 + d/2)}{4\pi^{d/2}}.$$

In particular, since $\Gamma(1/2) = \pi^{1/2}$, we have $c_3 = 1/(4\pi)$.

Proof. To motivate what the theorem says is true, we can again take Fourier transforms in $\Delta G = \delta$ as in the proof of Theorem 4.18, to see that (in $\mathcal{S}'$) $-4\pi^2|t|^2\widehat{G} = 1$. This suggests to try $\widehat{G}(t) = -1/(4\pi^2|t|^2)$. Theorem 4.17 then shows that indeed $G(x) = -c_d|x|^{2-d}$, with $c_d = 1/(4\pi^2 c_{d-2,d})$. □

Exercise 4.18. The way I wrote it up, the argument doesn't quite follow the required logic. We should really *define* $G(x)$ as in the theorem and then check that $\Delta G = \delta$. Clean up the argument by reorganizing it along these lines, and also derive the formula for c_d from (4.9).

Theorem 4.20 also works for $d = 1$ and is in fact much easier to obtain in this case. Since $c_1 = -1/2$, it then says that $G(x) = |x|/2$, and we verified long ago that indeed $\Delta G = G'' = \delta$, as required.

If we argue as in the proof of Theorem 4.20 for $d = 1$, then we learn that $u = \widehat{|x|}$ satisfies $-2\pi^2 t^2 u = 1$ and thus

$$u\big|_{\mathbb{R}\setminus\{0\}} = \frac{-1}{2\pi^2 t^2}.$$

This is a smooth function away from $t = 0$ and presents no problems there, but clearly it cannot be the case that u itself is this function because $1/t^2$ is much too singular near $t = 0$ and fails to be integrable there. So what is u? To answer this, recall from Theorem 4.13 that

$$i\pi \operatorname{sgn}\widehat{\ } = \mathrm{PV}\,\frac{1}{x};$$

to write the result in this form, I have also used that $\widehat{\ }\widehat{\ } = R$. So, by Theorem 4.11,

$$2\pi^2 (t \operatorname{sgn}(t))\widehat{\ } = \left(\mathrm{PV}\,\frac{1}{x}\right)'.$$

Now $t \operatorname{sgn}(t)$ is simply $|t|$, so we have already obtained the interesting result that

$$\widehat{|t|} = \frac{1}{2\pi^2}\left(\mathrm{PV}\,\frac{1}{x}\right)'.$$

Finally, let's look at this derivative more closely. We have

$$\left(\left(\mathrm{PV}\,\frac{1}{x}\right)', \varphi\right) = -\lim_{h\to 0+} \int_{|x|>h} \frac{\varphi'(x)}{x}\, dx.$$

We integrate by parts in the two integrals $\int_{-\infty}^{-h} \ldots$ and $\int_h^{\infty} \ldots$ and then put things back together to obtain

$$\left(\left(\mathrm{PV}\,\frac{1}{x}\right)', \varphi\right) = \lim_{h\to 0+} \left(\frac{\varphi(-h)+\varphi(h)}{h} - \int_{|x|>h} \frac{\varphi(x)}{x^2}\, dx\right).$$

Since

$$\frac{\varphi(-h)+\varphi(h)-2\varphi(0)}{h} \to 0$$

and $\int_{|x|>h} dx/x^2 = 2/h$, we can rewrite this as

$$\left(\left(\mathrm{PV}\,\frac{1}{x}\right)', \varphi\right) = \lim_{h\to 0+} \int_{|x|>h} \frac{\varphi(0)-\varphi(x)}{x^2}\, dx.$$

We summarize:

Theorem 4.21.

$$(u, \varphi) = \lim_{h\to 0+} \int_{|x|>h} \frac{\varphi(x) - \varphi(0)}{x^2}\, dx$$

defines a (tempered) distribution. We have $u = -2\pi^2 \widehat{|t|}$ *and also* $u = -(\mathrm{PV}(1/x))'$.