Some notes on distributions

January 15, 2006

These are course notes for Mathematics 80600, Musagei Yisod B'analiza. They are meant to complement the material in the handwritten notes which I have already distributed. One should also refer to Paul Garrett's material at `http://www.math.umn.edu/~garrett/m/fun/`.

1 The Dirac δ-function

Formally, distributions are usually defined as continuous linear functions on a certain topological vector space of "test functions". However, it is not surprising that useful mathematical objects such as distributions also have a natural, concrete description. We shall start from this viewpoint, and later explain the abstraction to topological vector spaces, and explain why the latter is important also.

We begin with the example of the most famous distribution: δ_0, the Dirac δ-function on $\mathbb{R}$. This is colloquially described as a nonnegative function from $\mathbb{R}$ to $\mathbb{R}$ with the contradictory properties that it equals zero at all nonzero arguments, but yet that its integral over $\mathbb{R}$ equals 1. Of course no such function can exist.

Nevertheless, there are two complementary ways to make sense out of δ_0. One is to relate it to the linear functional $\Lambda : C(\mathbb{R}) \to \mathbb{C}$ which sends a function $\psi(x)$ to $\Lambda\psi := \psi(0)$. Another is to take a fixed, smooth nonnegative function with rapid decay[1] and integral 1, such as $f(x) = e^{-\pi x^2}$, and study the sequence $f_T(x) = Tf(xT)$ as $T \to \infty$. Each of the functions f_T also

[1] A perfect example is an element of the Schwartz space $\mathcal{S}(\mathbb{R})$, which consists of all smooth functions g on $\mathbb{R}$ such that $\lim_{|x|\to\infty} |x|^n g^{(m)}(x) = 0$ for all $n, m \geq 0$.

has integral 1, and $\lim_{T\to\infty} f_T(x) = 0$ for all $x \neq 0$. Of course f_T does not converge to a function, but one has that

$$\lim_{T\to\infty} \int_{\mathbb{R}} f_T(x)\psi(x)\,dx \quad = \quad \psi(0)\,, \tag{1.1}$$

which agrees with $\Lambda\psi$. Here we can already see a couple of important properties of distributions:

- They are linear functionals on spaces of functions.
- They can be viewed as limits of sequences of smooth functions, meaning that distributions are a type of completion of functions.
- The definition of distribution is *localized*: in this example, the properties of ψ away from zero are irrelevant to δ_0.

With a bit more work we can see some other general features of distributions in this example. The first is that we may formally view distributions as transforming like functions, although to do this technically we must insist that distributions not act on functions like $\psi(x)$, but instead on *measures* like $\psi(x)dx$. Formally we write

$$\int_{\mathbb{R}} \delta_0(x)\psi(x)\,dx \tag{1.2}$$

for the expression in (1.1). This is purely a notational convenience; it really means the action of the linear functional Λ acting on ψ, but we write it this way to remind us that reasonably integrable functions give linear functionals when integrated against functions this way (like the terms in the limit in (1.1)).

With the mantra *"distributions transform like functions"* we may apply some operations to distributions that are valid on functions. The first which comes to mind is that *distributions can be differentiated*: we define

$$\int_{\mathbb{R}} \delta_0'(x)\psi(x)\,dx \quad = \quad -\int_{\mathbb{R}} \delta_0(x)\psi'(x)\,dx \quad = \quad -\psi'(0)\,, \tag{1.3}$$

because the first equality is what we expect from integration by parts. Note that we have no boundary terms in this integration by parts formula because the distribution $\delta_0(x)$ has no association with the properties of functions

$\psi(x)$ away from $x = 0$. We might as well even assume that ψ has compact support. However, it is crucial that ψ be differentiable now, or else of course the righthand side cannot make sense. We see that the regularity of the functions on which distributions may act depends on the distribution itself.

By repeated differentiation, one obtains the n-th derivatives $\delta_0^{(n)}(x)$, which have the property that $\int_{\mathbb{R}} \delta_0^{(n)}(x)\psi(x) = (-1)^n \psi^{(n)}(x)$ for $\psi \in C_c^\infty(\mathbb{R})$, the sign of course coming from an n-fold application of integration by parts. We can also attempt to integrate distributions, or at least take their antiderivatives. Since δ_0 can be integrated against any continuous function, we can make sense of its antiderivative as

$$\delta_0^{(-1)}(x) \;=\; \int_{-\infty}^{x} \delta_0(t)\, dt \;=\; \int_{\mathbb{R}} \chi_{(-\infty,x)}(t)\, d_0(t)\, dt \;=\; \begin{cases} 0, & x < 0 \\ 1, & x > 0 \end{cases} \tag{1.4}$$

Of course we have not specified $\delta_0^{(-1)}(0)$; perhaps giving it the value of $1/2$ is the fairest solution. That means that δ_0 does not have a canonical antiderivative (of course we could also shift it by constants), but any of these definitions satisfies the property that

$$\int_{\mathbb{R}} \delta_0^{(-1)}(x)\psi'(x)\, dx \;=\; \int_0^\infty \psi'(x)dx \;=\; -\psi(0) \;=\; -\int_{\mathbb{R}} \delta_0(x)\psi(x)dx\,, \tag{1.5}$$

which is what we expect from integration by parts. Taking a second antiderivative gives (up to linear polynomials) the continuous function $\delta_0^{(-2)}(x) = \frac{|x|}{2}$. We see here the important principle that *distributions are locally expressible as higher order derivatives of continuous functions.*

We said "locally" expressible here because actually, by the local nature of the action of distributions, these antiderivatives can still be antiderivatives if they are multiplied by a smooth function which vanishes off a neighborhood of $\{0\}$, naturally termed the *support of the distribution* δ_0. The example

$$\tau(x) \;=\; \sum_{n\geq 0} \delta^{(n)}(x-n) \tag{1.6}$$

shows that distributions do not always have *global* expressions as a finite number of derivatives of a continuous function. This leads us to the final general property we wish to discuss in this example, namely that *distributions can be multiplied by smooth functions.* If ϕ is a smooth function, then $\phi(x)\delta_0(x)$ can be made sense of, again through the mnemonic that distributions transform like functions:

$$\int_{\mathbb{R}} \phi(x)\delta_0(x)\psi(x) \quad = \quad \phi(0)\psi(0)\,. \tag{1.7}$$

In other words, $\phi(x)\delta_0(x) = \phi(0)\delta_0(x)$, which in this special case is just a scalar multiple of $\delta_0(x)$.

2 Distributions on an open subset of $\mathbb{R}^n$

We have now seen enough features of the δ-function to give a general definition of distribution. We begin with a few other examples which help phrase the eventual definition compactly. Our general context will be an open subset Ω of $\mathbb{R}^n$, and our integrations will in general be against smooth measures with compact support in Ω, i.e. measures of the form $\psi(x)dx$ for $\psi \in C_c^\infty(\Omega)$.

We start with the observation that any function $f \in C(\Omega)$ can be integrated against $\psi \in C_c^\infty(\Omega)$ by the usual pairing

$$\int_\Omega f(x)\psi(x)\,dx\,, \tag{2.1}$$

which defines a linear functional on $C_c^\infty(\Omega)$. To talk about derivatives in $\mathbb{R}^n$ we need some notation. If $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$ is a *multi-index*, then D^α represents the differential operator $\frac{d^{\alpha_1}}{dx_1^{\alpha_1}}\frac{d^{\alpha_1 2}}{dx_2^{\alpha_2}}\cdots\frac{d^{\alpha_n}}{dx_n^{\alpha_n}}$ of order $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n$. We can think of the derivative $D^\alpha f$ of f (as above) as distribution defined by the following rule:

$$\int_\Omega (D^\alpha f)(x)\psi(x)\,dx \quad = \quad (-1)^{|\alpha|}\int_\Omega f(x)(D^\alpha\psi)(x)\,dx\,. \tag{2.2}$$

Of course if f happens to be smooth, then this last identity is simply integration by parts, and holds anyway. We use it as a mechanism to extend the notion of derivative to nondifferentiable functions.

Definition 2.1. *A* distribution τ *on an open subset* Ω *of* $\mathbb{R}^n$ *is a sum*

$$\tau \quad = \quad \sum_{\alpha \geq 0} D^\alpha f_\alpha \tag{2.3}$$

where the $f_\alpha \in C(\Omega)$ *have the property that all but a finite number of* f_α *vanish on any given compact subset of* Ω *(i.e. the sum is finite on any compact set).*

We use the notation $\mathcal{D}'(\Omega)$ to denote the vector space of distributions τ on Ω.

3 Distributions as linear functionals

Now, having given a concrete, workable definition of distributions, we give a more abstract characterization in terms of topological vector spaces. Again, our setting is an open subset Ω or $\mathbb{R}^n$. Let us recall the Frechet spaces $\mathcal{D}_K = C_c^\infty(K)$, for K a compact subset of Ω, which have topologies given by the seminorms

$$\|f\|_N \quad = \quad \max\{|(D^\alpha f)(x)| \ \mid \ x \in \Omega \ , \ \ |\alpha| \le N\} \, . \tag{3.1}$$

The space of *test functions* $\mathcal{D}(\Omega)$, so called because distributions are determined by their action on them, is the vector space union

$$\mathcal{D}(\Omega) \quad = \quad \cup_{K\subset\Omega,\text{compact}} \mathcal{D}_K \, . \tag{3.2}$$

We make $\mathcal{D}(\Omega)$ into a topological vector space by giving in the following local base:

$$\begin{aligned} \mathcal{B} \quad = \quad \{ & \text{balanced convex subset } W \subset \mathcal{D}(\Omega) \text{ such that } W \cap \mathcal{D}_K \\ & \qquad \text{is open in } \mathcal{D}_K \text{ for all compact subsets } K \subset \Omega\} \, . \end{aligned} \tag{3.3}$$

Of course we need to justify that this gives a TVS topology. This topology is also known as the *direct limit* or *ascending union* of the Frechet spaces $\mathcal{D}_K$, where K is a compact subset of Ω. This is the finest locally convex topology on $\mathcal{D}(\Omega)$ such that all the inclusion maps $\mathcal{D}_K \hookrightarrow \mathcal{D}(\Omega)$ are continuous.

Proposition 3.1. *The translates $f+W$, for $f \in \mathcal{D}(\Omega)$ and $W \in \mathcal{B}$, are closed under finite intersection (unless, of course, the intersection is empty). Thus the topology generated by $\mathcal{B}$ consists of unions of translates of set $W \in \mathcal{B}$.*

Proof. Suppose that f_1, $f_2 \in \mathcal{D}(\Omega)$ and $W_1, W_2 \in \mathcal{B}$, and that f is a member of both f_1+W_1 and f_2+W_2. We need to show the existence of a subset $W \in \mathcal{B}$ such that $f+W \subset (f_1+W_1)\cap(f_2+W_2)$. Since functions in $\mathcal{D}(\Omega)$ have compact support, there exists a compact subset K of Ω containing the support of f, f_1, and f_2, i.e. $f, f_1, f_2 \in \mathcal{D}_K$. By definition of the putative local base $\mathcal{B}$, $\mathcal{D}_K \cap W_1$ and $\mathcal{D}_K \cap W_1$ are open subsets of $\mathcal{D}_K$. The former contains $f - f_1$, and the latter, $f - f_2$. As scalar multiplication is continuous in the topological vectors space $\mathcal{D}_K$, there exists some $\varepsilon > 0$ such that $(1-\varepsilon)^{-1}(f - f_1) \in W_1$ and $(1-\varepsilon)^{-1}(f - f_2) \in W_2$, i.e. $f - f_1 \in (1-\varepsilon)W_1$ and $f - f_2 \in (1-\varepsilon)W_2$. The sets W_1 and W_2 are convex, so

$$f \ - \ f_j \ + \ \varepsilon\, W_j \quad \subset \quad (1-\varepsilon)W_j \ + \ \varepsilon W_j \quad \subset \quad W_j \ \ , \qquad j = 1, 2 \, . \tag{3.4}$$

So, choosing W to be the intersection of εW_1 and εW_2, we see that f indeed is a member of both $f_1 + W$ and $f_2 + W$. □

Theorem 3.2. *The topology on $\mathcal{D}(\Omega)$ with local base $\mathcal{B}$ makes it into a locally convex topological vector space.*

Proof. First we show that singletons are closed by demonstrating that for any two functions $f, g \in \mathcal{D}(\Omega)$ there exists some $W \in \mathcal{B}$ such that $f \notin g + W$. This can be accomplished by choosing $W = \{\phi \in \mathcal{D}(\Omega) \mid \|\phi\|_0 < \|f - g\|_0\}$ (the norms here are just sup-norms). This set is obviously balanced and convex, and its intersection with any $\mathcal{D}_K$ is of course open, so it is indeed a member of the local base $\mathcal{B}$.

Next we demonstrate addition is continuous, which follows immediately from the fact that $\frac{1}{2}W + \frac{1}{2}W \subset W$ for any $W \in \mathcal{B}$ (these sets are all convex). For continuity of scalar multiplication, let α be a scalar and $f \in \mathcal{D}(\Omega)$. Given any $W \in \mathcal{B}$ there exists some small constant $\delta > 0$ such that $\delta f \in \frac{1}{2}W$. Then for all β such that $|\beta - \alpha| < \delta$ and all $\phi \in f + \frac{1}{2(|\alpha|+|\delta|)}W$ we have that $\beta\phi - \alpha f = \beta(\phi - f) + (\beta - \alpha)f \in \frac{1}{2}W + \frac{1}{2}W \subset W$. Thus scalar multiplication is indeed continuous. □

We have said all this to give another definition of distribution, whose equivalence is not obvious: $\mathcal{D}'(\Omega)$ is the space of continuous linear functionals on $\mathcal{D}(\Omega)$ under the above topology.

4 Interlude: Fourier theory

Let us quickly recall some facts about Fourier series and integrals, restricted to the convenient setting of smooth functions and Schwartz functions, respectively. If $f : \mathbb{R}^n/\mathbb{Z}^n \to \mathbb{C}$ is a smooth periodic function, then it equals its Fourier series $f(x) = \sum_{m\in\mathbb{Z}} \widehat{f}(n)e(mx)$, where as usual $e(x) = e^{2\pi inx}$ and mx is shorthand for $m_1x_1 + \cdots + m_nx_m$. The Fourier coefficients $\widehat{f}(m) = \int_{\mathbb{R}/\mathbb{Z}} f(x)e(-mx)\,dx$ have rapid decay as $|m| \to \infty$ and thus the Fourier series converges absolutely.

If τ is a periodic *distribution*, meaning one for which

$$\int_{\mathbb{R}} \tau(x)\psi(x)\,dx \quad = \quad \int_{\mathbb{R}} \tau(x)\psi(x+m)\,dx \tag{4.1}$$

for all $\psi \in C_c^\infty(\mathbb{R})$ and $m \in \mathbb{Z}^n$, then in fact it can be written as a Fourier series – but one that converges only as a distribution (i.e. in the weak* topology). It makes sense to integrate $\tau(x)$ against $e(-mx)$ over any fundamental period of $\mathbb{R}^n/\mathbb{Z}^n$, such as $[0,1)^n$. It is not hard to see that these integrals, whose values we denote by a_m, must be bounded by some polynomial in $|m|$ (this comes from the description of distributions as finite derivatives of continuous functions on compact sets, such as $[0,1]$). Then $\tau(x)$ in fact must equal the infinite sum $\sum_{m\in\mathbb{Z}} a_m e(mx)$ – viewed as a distribution – because the exponentials $e(mx)$ span $C^\infty(\mathbb{R}^n/\mathbb{Z}^n)$. It is clear from the Fourier series that some finite antiderivative of τ is continuous, but not necessarily periodic; if, however, $a_0 = 0$ and $n = 1$ then τ has canonical antiderivatives of the form

$$\tau^{(-k)}(x) \quad = \quad \sum_{m\neq 0} a_n (2\pi i m)^{-k} e(mx)\,. \tag{4.2}$$

Similar formulas can be written when $n > 1$, but one must separate out more types of terms (i.e. those in which a fixed number of the m_i vanish). These are crucial in the next section for showing the equivalence of the two definitions of distribution which we have put forward thus far.

Let $f \in \mathcal{S}(\mathbb{R}^n)$ be a Schwartz function. The Fourier transform of f is defined by the absolutely convergent integral

$$\widehat{f}(r) \quad = \quad \int_{\mathbb{R}^n} f(x) e(-rx)\,dy \tag{4.3}$$

and is also a Schwartz function. One has the inversion formula

$$f(x) \quad = \quad \int_{\mathbb{R}^n} \widehat{f}(r) e(rx)\,dr\,. \tag{4.4}$$

The following formula is a very important connection between Fourier series and transforms.

Theorem 4.1. *(Poisson Summation Formula). Given $f \in \mathcal{S}(\mathbb{R}^n)$,*

$$\sum_{m\in\mathbb{Z}^n} f(m) \quad = \quad \sum_{m\in\mathbb{Z}^n} \widehat{f}(m)\,. \tag{4.5}$$

Proof. Let $F(x) = \sum_{m\in\mathbb{Z}^n} f(x+m)$, which is an absolutely convergent sum and hence defines an element of $C^\infty(\mathbb{R}^n/\mathbb{Z}^n)$. It has a Fourier series whose coefficients $\widehat{F}(m) = \int_{\mathbb{R}^n/\mathbb{Z}^n} F(x)e(-mx) = \widehat{f}(m)$ (combine the sum over m to remove the quotient of $\mathbb{R}^n$ by $\mathbb{Z}^n$ in the integral). Then the Fourier series for at $x = 0$ is $F(0) = \sum f(m) = \sum \widehat{f}(m)$. ☐

Finally, we come to the following important inner product relation involving the Fourier transform:

Theorem 4.2. *(Parseval's Theorem). For any functions $f, g \in \mathcal{S}(\mathbb{R}^n)$ one has that*

$$\int_{\mathbb{R}^n} f(x)\widehat{g}(x)\,dx \quad = \quad \int_{\mathbb{R}^n} \widehat{f}(x)g(x)\,dx\,. \tag{4.6}$$

Proof. Substitute the definition $\widehat{g}(x) = \int_{\mathbb{R}^n} g(y)e(-xy)dy$ into the lefthand side. Since both f and g have rapid decay, Fubini's theorem allows us to write it as the double integral $\iint_{\mathbb{R}^n \times \mathbb{R}^n} f(x)g(y)e(-xy)dxdy$, which is obviously symmetric in f and g. □

5 Equivalence of the two definitions

It is not hard to see that the definition of distribution at the end of section 2 implies that they are continuous linear functionals on $\mathcal{D}(\Omega)$ with its direct limit topology. It is harder to see the converse, namely that all continuous linear functionals on $\mathcal{D}(\Omega)$ are given as sums of derivatives of continuous functions. Since the definition of derivative can be localized, it suffices to prove this for distributions which are supported in the interior of the unit cube $[0,1]^n \subset \mathbb{R}^n$. From this stage, it is no loss of generality to assume it is periodic. We shall prove this converse, then, by showing that any periodic distribution on $\mathbb{R}^n$ is a finite sum of derivatives of continuous periodic functions.

One of the equivalent criteria for a continuous linear functional Λ on $\mathcal{D}(\Omega)$ to be continuous is that on each compact subset K of Ω there exists constants $C > 0$ and $N \in \mathbb{Z}_{\geq 0}$ such that $|\Lambda\phi| \leq C\|\phi\|_N$ for all $\phi \in \mathcal{D}_K$. Since we are speaking of periodic distributions, which are determined on a compact set, we may assume this estimate holds uniformly over all $\phi \in \mathcal{D}_K$. If ϕ is a periodic exponential $e(mx)$, then all its derivatives of order k are multiplies by polynomials of degree k in m, so that $\|\phi\|_N = O(|m|^N)$. Thus, if $\tau \in \mathcal{D}'(\Omega)$ is a distribution, then its Fourier coefficients are polynomially bounded:

$$a_m \quad := \quad \int_{\mathbb{R}^n/\mathbb{Z}^n} \tau(x)e(-mx)\,dx \quad = \quad O(|m|^N)\,. \tag{5.1}$$

We may formally write $\tau(x) = \sum_{m\in\mathbb{Z}^n} a_m e(mx)$. It is clear that, by separating out the terms for which some of the indices of $m = (m_1, \ldots, m_n)$ are 0, that $\tau(x)$ may be written as a finite sum of derivatives of continuous periodic functions. For example, when $n = 1$, we write

$$\tau(x) \quad = \quad a_0 \; + \; \frac{d^{N+2}}{dx^{N+2}} \sum_{m \neq 0} a_m e(mx)/(2\pi i m)^{N+2} \, .$$

6 Tempered distributions

The Schwartz space $\mathcal{S}(\mathbb{R})$ properly contains $C_c^\infty(\mathbb{R})$, so not every distribution can be evaluated on Schwartz functions. Those which can are called *tempered distributions* $\mathcal{S}'(\mathbb{R})$. They have the following characterization:

Theorem 6.1. *The tempered distributions are precisely those elements of* $\mathcal{D}'(\mathbb{R})$ *of the form* $\frac{d^k}{dx^k} f$*, where* f *is a continuous function of at most polynomial growth at infinity.*

The Parseval Theorem 4.2 allows us to extend the notion of Fourier transform to a tempered distribution τ by the rule

$$\int_{\mathbb{R}} \widehat{\tau}(x) \psi(x) \, dx \quad = \quad \int_{\mathbb{R}} \tau(x) \widehat{\psi}(x) \, dx \qquad \text{for } \psi \, \in \, \mathcal{S}(\mathbb{R}) \, . \tag{6.1}$$

As usual, this definition of Fourier transform agrees with the usual definition for tempered distributions which are already Schwartz functions.

As an example, the Poisson summation formula (Theorem 4.1) can be compactly restated in terms of the distribution $\delta_{\mathbb{Z}}(x) := \sum_{n \in \mathbb{Z}} \delta_0(x - n)$:

$$\delta_{\mathbb{Z}} \quad = \quad \widehat{\delta_{\mathbb{Z}}} \, , \tag{6.2}$$

i.e., the tempered distribution $\delta_{\mathbb{Z}}$ is its own Fourier transform.

7 Application: the Riemann ζ-function

As an example of how these techniques can be applied, we will now quickly go through a proof of the functional equation and analytic continuation of the Riemann ζ-function $\zeta(s) = \sum_{n \geq 1} n^{-s}$. Fix a smooth cutoff function $\phi(x)$ which is supported in $(-1, 1)$ and which is identically equal to 1 in a neighborhood of the origin.

Consider the distribution

$$\tau(x) \;=\; \sum_{n\neq 0}\delta_n(x) \,-\, 1 \,+\, \phi(x) \;=\; \sum_{n\neq 0} e(nx) \,-\, \delta_0(x) \,+\, \phi(x)\,. \qquad (7.1)$$

The first expression exhibits $\tau(x)$ as a distribution which vanishes in a neighborhood of the origin, while the second as one which has a Fourier expansion without constant term away from the origin. We would like to integrate $\tau(x)$ against the "Mellin kernel" $|x|^{s-1}$ despite that fact that it does not have compact support (nor is it defined at $x = 0$). The Mellin kernel is, however, smooth on any compact subset away from the origin, and so it locally makes sense to integrate $\tau(x)$ against it away from the potential singularities of $\tau(x)$ at $x = 0$ and ∞. Since the distribution $\tau(x)$ is identically zero in a neighborhood of the origin, the integration is justified near there because $|x|^{s-1}$ is smooth on the support of $\tau(x)$. Near $x = \infty$, $\tau(x)$ reduces to the periodic distribution $\sigma(x) = \sum_{n\neq 0} e(nx)$, which has bounded antiderivatives $\sigma^{(-N)}(x) = \sum_{n\neq 0} e(nx)/(2\pi i n)^N$ of arbitrarily high order. The integration by parts relation says

$$\int_{\mathbb{R}} \sigma(x)|x|^{s-1}\,dx \;=\; (-1)^N \int_{\mathbb{R}} \sigma^{(-N)}(x)\left(\frac{d^N}{dx^N}|x|^{s-1}\right)dx\,. \qquad (7.2)$$

The derivative on the righthand side is a polynomial in s times $|x|^{s-1-N}$, and so the pairing integral on the righthand side is in fact valid in the much larger range Re $s < N$. We use this to extend the definition of the lefthand side to all values of s, so it is hence entire (it is easy to see this defines the same extension since both are holomorphic and agree in a right half plane). Thus $\int_{\mathbb{R}} \tau(x)|x|^{s-1}dx$ makes sense as an entire function of s.

For Re $s < 0$, one may calculate this integral term by term to obtain the identity

$$\int_{\mathbb{R}} \tau(x)|x|^{s-1}\,dx \;=\; 2\zeta(1-s) \,+\, \int_{\mathbb{R}} (\phi(x)-1)|x|^{s-1}\,dx\,. \qquad (7.3)$$

This operation is valid because the sum for $\zeta(1-s)$ converges absolutely here. The integrand in the last term is supported away from 0, and thus makes sense as an ordinary integral. In fact it is not difficult to see that

$$\int_{\mathbb{R}} (\phi(x)-1)|x|^{s-1}\,dx \qquad \text{has a holomorphic continuation to } \mathbb{C}-\{0\} \quad (7.4)$$

with a simple pole at $s = 0$ of residue 2. Thus $\tau(x)$ has an entire Mellin transform, and by (7.3) and (7.4), $\zeta(s)$ has a holomorphic continuation to $\mathbb{C} - \{1\}$ and a simple pole with residue 1 at $s = 1$. This is the analytic continuation of the Riemann ζ-function.

The functional equation of $\zeta(s)$ can also be seen using this argument. We may also calculate $\int_{\mathbb{R}} \tau(x)|x|^{s-1}dx$ using the righthand side of (7.1), but for Re s large:

$$\int_{\mathbb{R}} \tau(x)|x|^{s-1}\,dx \;=\; \sum_{n\neq 0} \int_{\mathbb{R}} e(nx)|x|^{s-1}\,dx + \int_{\mathbb{R}} \phi(x)|x|^{s-1}\,dx\,. \qquad (7.5)$$

The term for $\delta_0(x)$ drops out since when Re s is large, $|x|^{s-1}$ is a continuous function equal to 0 at $x = 0$. Using elementary calculus one sees that the second terms on the righthand sides of (7.3) and (7.5) cancel. Changing variables $x \to x/n$, the first term becomes $\sum_{n\neq 0} |n|^{-s} \int_{\mathbb{R}} e(x)|x|^{s-1}$. Of course this integral must be interpreted using the fact that $e(x)$ has antiderivatives of arbitrarily high order; it is a standard result about the Γ-function that it equals $\frac{\pi^{-s/2}\Gamma(s/2)}{\pi^{-(1-s)/2}\Gamma((1-s)/2)}$, while the infinite sum over n of course gives $2\zeta(s)$. We have now proved the following important theorem:

Theorem 7.1. *(Functional equation and analytic continuation of the Riemann ζ-function). The Riemann ξ-function $\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$ has a holomorphic continuation to $\mathbb{C} - \{0, 1\}$ with simple poles at $s = 0$ and 1 with residues -1 and 1, respectively. It also obeys the functional equation*

$$\xi(s) \;=\; \xi(1-s)\,. \qquad (7.6)$$

8 Application: Venkatesan's model of vision

Ramarathnam Venkatesan has proposed a model of human vision which is inspired by the theory of distributions. He notes that many important signal processing operations involve treating an image as a 2- or 3-dimensional grid of data, integrated against a smooth function. Though the image of course makes sense as a function, Venkatesan posits that the human brain actually views it as a distribution, in the sense that it measures its integral against localized test functions. This is supported by experimental observations; for example, the human eye constantly shift its angle of focus, and the brain reassembles these pieces together.

Empirically, Venkatesan has developed this into a creating a metric for how close two images are to human perception, even though they might have other statistics which distinguish them. His randomized algorithms integrate the image distributions against a variety of randomly chosen test functions, and compare the outcomes. His randomized basis of wavelets, or “randlets”, performs up to 3 times better than Daubechies’ wavelets in some instances. More importantly, the randomization is crucial for cryptographic purposes because repeatedly using a known, canonical basis of wavelets can be exploited by an adversary.