CHAPTER 2

Distributions

13. The concept of distributions

In theories describing physical phenomena, it is always assumed that the spacetime is a continuum, and measurable quantities are functions that have pointwise values on the continuum. Calculus with classical functions, that is, differentiation and integration is a main tool to model physical phenomena by equations in partial derivatives. For example, electromagnetic waves, their generation and interactions with matter are described by electric and magnetic vector fields satisfying Maxwell's equations. However, no measurement of the field strength, or electric current and charge densities can be made at a point in space and at a precise moment of time. In reality, any measurement gives us some smeared or averaged values of physical quantities in space and time. The very notion of "instant" and "point-like" is a mathematical idealization of the situation in which any measurement of time or position in space is assumed to have no uncertainty at all. The latter is, of course, not true. Anything that happens "instantly" has a duration in time that is not possible to resolve in measurements. Anything that is said to be at a "point", in reality, occupy a portion of space that is determined by uncertainties in position measurements. Thus, only mean values of physical quantities make sense.

From this perspective, the concepts of classical calculus, like derivatives, make no sense as their values require, first, knowledge of physical quantities as functions having pointwise values and, second, taking limits. The former are not available due to the very nature of measurements, and the latter is not possible to do in practice because arbitrary small distances between any two points or any two moments of time cannot be reached.

A *distribution* (or a *generalized function*) is an extension of the concept of a classical function. Distributions are not required to have pointwise values but they are defined by smeared or averaged values in any neighborhood of any point. So, any locally integrable (classical) function is a distribution because it has an integral mean value (possibly with some weight) on any neighborhood of any point. However, mathematical modeling of reality often requires other distributions than those defined by classical functions.

Consider a process in which a force applied to a particle creates a finite momentum change of the particle during an arbitrary small interval of time. This force can be viewed as the limit of a force whose amplitude rises from zero and then decreases back to zero in an arbitrary small interval of time while the integral of the force (the net momentum change) is finite. The precise details of the increase and decrease of the force as a function of time are irrelevant for the process because the time interval during which this happens is too short to be even measured. The limit force cannot be described by a classical function because the latter would be zero everywhere except the time moment when the momentum transfer occurs. The integral of such a function is equal to zero and the limit force cannot create any finite change of the momentum. In contrast, the limit process can well be described by distributions. If the limit force is a distribution then it is defined not by its pointwise values but rather by smeared or averaged values in any (arbitrary small) neighborhood of any time moment. This value can be set to be a given constant for any neighborhood of the time moment at which an instant momentum transfer occurs. This constant may depends on details of smearing or averaging (representing experimental observations), but it does not vanish in the limit when the size of a neighborhood tends to zero.

Similarly, the concept of a point particle is a mathematical idealization of a situation in which the "inner" structure of the particle, such as distributions of mass or electric charge within the particle, either cannot be measured or irrelevant for the process studied. However, such particles create gravitational or electric fields extended throughout space. Moving point-like electric charges create extended magnetic fields. Finding these fields from equations for the fields (e.g., Maxwell's equations) requires mass and electric charge densities. The mass or electric charge density of a point particle can again be viewed as a limit of a density defined in successively smaller volumes occupied by a "real" particle. This limit cannot be described by a classical function as any such function would have zero value everywhere except the point at which the particle is located. Distributions must be used to describe such densities because the value of a distributional mass density is set by its averaged value in any neighborhood of the position of the particle, that is, by the total mass of the particle, regardless how small this neighborhood is.

Algebra, calculus, and solving differential equations with distributions are quite different from their classical analogue but coincide with the latter whenever physical quantities are assumed to be classical functions. For example, every distribution is infinitely many times differentiable (in the sense of distributions) so that many complicated issues of classical analysis about smoothness of solutions to differential equations becomes obsolete. But the price is more complicated techniques in calculations with distributions. The main objective throughout next few chapters is to give a precise meaning to distributions and develop basic calculus with them as well as to extend other technical tools from the classical analysis to distributions, e.g., summations of series, taking Fourier transforms, etc.

13.1. Dirac delta-function. P. Dirac introduced[1] the first distribution into physics as a "function" $\delta(x)$ that is zero everywhere except one point, say, $x = 0$, but its integral with any smooth function $\varphi(x)$ gives the value of φ at $x = 0$:

$$\text{“}\int \delta(x)\,\varphi(x)\,dx\text{”} = \varphi(0)\,.$$

Since then $\delta(x)$ is called *the Dirac delta-function*. The quotation marks around the integral stand for a mathematical fact that there exists no locally integrable function with such a property. This was the reason that the concept for such a "function" was not appreciated by the mathematical community of the time. However, despite not being well mathematically defined, the Dirac delta-function became a wonderful technical tool in quantum mechanics that allowed to physicists to calculate physically observable quantities. The stunning predictive power of quantum mechanics and, later, quantum field theory whose mathematical techniques were based on objects similar to the Dirac delta function (e.g., Feynman's propagators) changed perception of these objects by mathematicians, which eventually led to the theory of distributions.

As it stands, the Dirac delta-function resembles the physical concept of a force that can instantly make a finite momentum change or that of the mass or electric charge density of a point particle. Let us investigate this in detail.

13.1.1. A force making an instant momentum change. Suppose that a particle of unit mass that can only move along a line is subject to a force $f_\tau(t) \geq 0$ that has a finite duration $0 \leq t \leq \tau$ and is continuous

[1] P.A.M. Dirac, Principles of Quantum Mechanics

for all t. Then, according to Newton's second law, the net momentum change of the particle is

$$\Delta p = \int f_\tau(t)\,dt = \int_0^\tau f_\tau(t)\,dt\,.$$

The force has an integral mean value $\Delta p/\tau$ that can be measured by measuring its duration τ and the particle velocity before $t = 0$ and after $t = \tau$. When $\tau \to 0^+$, then τ eventually becomes smaller than a time interval that can possibly be measured, and the very concept of describing the force by a function of time becomes meaningless, whereas the net momentum change is still perfectly measurable. What is the limit force that can create such an instant change of the momentum?

Let us set $\Delta p = 1$, just to have the net momentum change to be 1 (not zero) in momentum units for any $\tau > 0$, and investigate the limit $\tau \to 0^+$. To mimic the fact that any measurement of the force only provides a mean or smeared value of the force, consider the limit of the integral

$$(f_\tau, \varphi) \stackrel{\text{def}}{=} \int f_\tau(t)\,\varphi(t)\,dt\,,$$

where $\varphi(t)$ is a smooth function with a bounded support that represents the averaging process. It will be called a *test function. The symbol* (f, φ) *stands for a "smeared or averaged" value of a distribution* f *on a test function* φ. A support of any continuous non-zero function always has non-zero measure because, if this function is not zero at a point, then by continuity it is not zero in a neighborhood of this point. So, the choice of a smooth (vs arbitrary) function as a test function represents that any measurement can be done only during a finite interval of time, although this interval can be arbitrary small but, most importantly, never zero.

Let us show that

$$\lim_{\tau \to 0^+} \int f_\tau(t)\,\varphi(t)\,dt = \varphi(0)\,,$$

for any test function φ, that is, the limit force has the characteristic property of the Dirac delta-function if the order of taking the limit and integration can formally be interchanged. Since φ is smooth, by the mean value theorem there exists t^* between t and 0 such that

$$\varphi(t) - \varphi(0) = \varphi'(t^*)t\,.$$

The derivative $\varphi'(t)$ of a smooth function is a continuous function and also has a bounded support. Therefore it is bounded:

$$\sup|\varphi'(t)| = M < \infty\,,$$

This implies that for any $t \geq 0$

$$|\varphi(t) - \varphi(0)| \leq Mt\,.$$

Since the integral of f_τ is normalized to 1 and f_τ is non-negative, the following chain of inequalities holds:

$$\begin{aligned}\Big|\int f_\tau(t)\,\varphi(t)\,dt - \varphi(0)\Big| &= \Big|\int f_\tau(t)\,\Big(\varphi(t)-\varphi(0)\Big)\,dt\Big| \\ &\leq \int f_\tau(t)\,\Big|\varphi(t)-\varphi(0)\Big|\,dt \\ &\leq M\int f_\tau(t)\,|t|\,dt = \int_0^\tau f_\tau(t)\,t\,dt \\ &\leq \tau M\int_0^\tau f_\tau(t)\,dt = M\tau \;\to\; 0\,,\end{aligned}$$

as $\tau \to 0^+$, as required.

Thus,

$$\lim_{\tau\to 0^+} f_\tau(t) = \delta(t)\,.$$

Note well that *this limit cannot be interpreted as a pointwise limit* because for any t, $f_\tau(t) \to f(t) = 0$ as $\tau \to 0^+$. Indeed, since $f_\tau(t) = 0$ for any $t \leq 0$, $f_\tau(t) \to 0$ for any $t \leq 0$ as $\tau \to 0^+$. Furthermore, $f_\tau(t) = 0$ if $t \geq \tau$. Therefore for any $t > 0$, $f_\tau(t) = 0$ for all small enough τ. Thus, the pointwise limit of $f_\tau(t)$ is zero. The integral of the zero function $f(t)\varphi(t) = 0$ is zero for any test function φ. There is no contradiction with the above result. It merely refers to the well known fact that the order of integration and taking the limit with respect to a parameter cannot always be interchanged. For this reason, the limit force $\delta(t)$ cannot be defined by an integral of some function with pointwise values, but rather it should be defined by its averaged or smeared values for any test function, that is,

$$(\delta, \varphi) \overset{\text{def}}{=} \varphi(0)\,. \tag{13.1}$$

This rule makes a perfect sense for any smooth φ but cannot be written as an integral average of some locally integrable function $\delta(t)$. Consequently, the limit $f_\tau(t) \to \delta(t)$ must be understood in the sense that the numerical limit

$$\lim_{\tau\to 0^+} (f_\tau, \varphi) = (\delta, \varphi)$$

holds for any test function φ.

13.1.2. Mass density of a point particle. In the simplest case, a particle of mass m can be modeled by a ball of radius $a > 0$ in which the mass is homogeneously distributed. Then the mass density is

$$\rho_a(x) = \begin{cases} m/V_a\,, & |x| < a \\ 0\,, & |x| > a \end{cases}$$

where $x \in \mathbb{R}^3$ and $V_a = \frac{4}{3}\pi a^3$ is the volume of the ball. A point particle corresponds to the limit $a \to 0^+$. The pointwise limit produces a density that vanishes almost everywhere

$$\lim_{a\to 0^+} \rho_a(x) = \begin{cases} 0\,, & x \neq 0 \\ \infty\,, & x = 0 \end{cases}$$

and, hence, its integral over any set is zero but it is supposed to be equal to the total mass m.

As in the case of an instant force, this contradiction can be resolved by interpreting the limit in the distributional sense. Only average values of $\rho_a(x)$ with some smooth test function φ can be an outcome of any measurement. A test function is assumed to be a smooth function in $\mathbb{R}^3$ with bounded support. The boundedness of support represents that any measurement is always curried out in a bounded region of space. If test function is not zero at a point, then it is not zero in a neighborhood of this point by the continuity argument, although this neighborhood can be arbitrary small.

By the integral mean value theorem, there exists a point x_a in a ball $|x| < a$ such that

$$\int_{|x|<a} \varphi(x)\, d^3x = V_a \varphi(x_a)$$

for any test function φ. Therefore

$$\lim_{a\to 0^+} \int \rho_a(x)\, \varphi(x)\, d^3x = m \lim_{a\to 0^+} \varphi(x_a) = m\varphi(0)$$

by continuity of φ. So, if the Dirac delta function of $x \in \mathbb{R}^3$ is defined by (**13.1**) for any test function on $\mathbb{R}^3$, then the limit density is the mass of the particle multiplied by the Dirac delta function in $\mathbb{R}^3$:

$$\lim_{a\to 0^+} (\rho_a, \varphi) = (m\delta, \varphi)\,. \qquad (\mathbf{13.2})$$

The next striking observation is that *the property (***13.2***) of the limit mass density does not depend on peculiarities of the mass distribution within the ball and holds for any non-negative $\rho_a(x)$ supported in the ball $|x| \leq a$ and whose integral is equal to the mass m.* This assertion follows from the mean value theorem for functions of several variables (**1**.**5**). Indeed, using the normalization property of the mass density

$$m = \int \rho_a(x)\, d^3x = \int_{|x|<a} \rho_a(x)\, d^3x\,, \quad a > 0\,,$$

and non-negativity $\rho_a(x) \geq 0$, one infers that

$$\begin{aligned}\left|\int \rho_a(x)\,\varphi(x)\, d^3x - m\varphi(0)\right| &\leq \int_{|x|<a} \rho_a(x)\left|\varphi(x) - \varphi(0)\right| d^3x \\ &\leq M \int_{|x|<a} \rho_a(x)\, |x|\, d^3x \\ &\leq Ma \int_{|x|<a} \rho_a(x)\, d^3x \\ &= Mma \to 0\,,\end{aligned}$$

when $a \to 0$, as required. Thus, the characteristic property (**13**.**2**) is *universal and does not depend on details of the mass distribution within the ball.* In particular, one can take ρ_a to be any integrable function such that $\rho_a(0) = 0$ for all $a > 0$. Then the pointwise limit of $\rho_a(x)$ is the zero function as $a \to 0^+$. In fact, ρ_a can be altered on any set of measure and the conclusion still holds.

13.2. Functionals. Let $\mathcal{D}$ denote a collection of functions of N real variables. Let us define a real-valued function on $\mathcal{D}$:

$$f:\ \mathcal{D} \to \mathbb{R}$$

that is, f *is a rule that assigns a unique real number, denoted by* (f, φ) *for every function* $\varphi \in \mathcal{D}$. A function defined on a set of functions is called a *functional*. For example, for any locally integrable function $f \in \mathcal{L}_{\mathrm{loc}}$ one can define a *functional* f by the rule

$$(f, \varphi) = \int f(x)\varphi(x)\, d^N x\,, \qquad \varphi \in \mathcal{D}\,,$$

if $\mathcal{D}$ consists of smooth functions with bounded support. If support of φ lies in a ball $|x| < R$, then the integration region can be limited to this ball. Since any compactly supported continuous function is bounded, the integral exists by the comparison test

$$|f(x)\varphi(x)| \leq M|f(x)|\,, \quad M = \sup|\varphi(x)| < \infty\,.$$

because the integral of $|f(x)|$ is finite over any ball by local integrability. So, the rule makes sense for any function from $\mathcal{D}$.

Another example of a functional is given by the Dirac delta-function that is defined by the rule (**13**.**1**). Physical examples studied above suggest that distributions can be identified with functionals on space of smooth functions with bounded support.

In classical analysis, two functions are said to be equal if they have equal values at any point. Similarly, two functional f and g are equal if they have equal values on all test functions:

$$f = g \quad \Leftrightarrow \quad (f,\varphi) = (g,\varphi)\,, \quad \varphi \in \mathcal{D}\,.$$

This reflects our general idea that two physical quantities, represented by distributions, are identical if they have the same average values in any measurements or testing.

13.2.1. Differentiation of distributions. As already noted, physical quantities are governed by equations in partial derivatives. Therefore one needs a differentiation rule for distributions. The guidance is provided by distributions defined by locally integrable functions. Let f be locally integrable in $\mathbb{R}$. The classical derivative is defined by the limit

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}\,,$$

provided it exists. However the pointwise values of f cannot be used to investigate the limit because for a generic distribution they either do not exist or are not known. Only the averaged values (f,φ) exist for any test function φ.

The locally integrable function $f_h(x) = f(x+h)$ is also a distribution or a functional on $\mathcal{D}$ whose values on a test function can be expressed via the values of the functional f:

$$(f_h,\varphi) = \int f(x+h)\varphi(x)\,dx = \int f(x)\varphi(x-h)\,dx\,.$$

The function $\varphi_h(x) = \varphi(x-h)$ is smooth and has bounded support for any test function φ and any real h. Therefore $\varphi_h \in \mathcal{D}$ and, hence, the functional f_h can be defined by the rule

$$(f_h,\varphi) = (f,\varphi_h)\,, \quad \varphi_h(x) = \varphi(x-h)\,,$$

for *any functional* f on $\mathcal{D}$ and any real h. The derivative f' must also be a functional on $\mathcal{D}$, that is, it must have a value on any test function. So, the best one can do to define the derivative of a distribution f is to put

$$(f',\varphi) \stackrel{\text{def}}{=} \lim_{h\to 0} \frac{(f_h,\varphi) - (f,\varphi)}{h}\,. \tag{13.3}$$

Let us investigate the limit. Again, the guidance is provided by distributions defined by locally integrable functions. Note that a locally integrable function is not differentiable in general. So, the rule

(**13.3**) already extends beyond the concept of a classical derivative. *By linearity of the integral*, one has

$$(f', \varphi) = -\lim_{h\to 0} \int f(x)\psi_h(x)\,dx\,, \quad \psi_h(x) = \frac{\varphi(x) - \varphi(x-h)}{h}\,.$$

There are two facts about the function ψ_h to be noted. First, for any $h \neq 0$, ψ_h is a test function for any $\varphi \in \mathcal{D}$, that is, ψ_h is a smooth function with bounded support. Second, $\psi_h(x) \to \varphi'(x)$ as $h \to 0$ for any $x \in \mathbb{R}$. Therefore, the limit can be found by the Lebesgue dominated convergence theorem. Indeed, for all small enough h, the support of ψ_h lies in $[-R, R]$ (with R being independent of such h). By (**1.5**)

$$|\psi_h(x)| \leq M < \infty\,, \quad M = \sup|\varphi'(x)|\,,$$

and therefore the integrand has an integrable bound independent of the parameter h

$$|f(x)\psi_h(x)| \leq M|f(x)| \in \mathcal{L}(-R, R)\,.$$

Hence, by the Lebesgue dominated convergence theorem, the order of integration and taking the limit can be interchanged, giving

$$(f', \varphi) = -\int f(x)\varphi'(x)\,dx = -(f, \varphi')\,. \tag{13.4}$$

Even for a locally integrable f that is not differentiable in the classical sense, the derivative f' exists as a functional or as a distribution. The rule (**13.4**) looks like an integration by parts, but this is a false impression because f must be from class C^1 in order to integrate by parts.

Two things can be deduced from this observation. First if f is from class C^1, then its classical and distributional derivatives are equal. Indeed, let $\{f'(x)\}$ denote the classical derivative of $f \in C^1$. Then $\{f'(x)\}$ is continuous and defines a functional by the rule

$$(\{f'\}, \varphi) = \int \{f'(x)\}\varphi(x)\,dx\,.$$

By integrating by parts in this integral one infers that

$$(\{f'\}, \varphi) = -\int f(x)\varphi'(x)\,dx = (f', \varphi)$$

for any test functions. This means that $f' = \{f'\}$ as functionals on $\mathcal{D}$. Second, if f is not differentiable, its distributional derivative still exist! In fact, the functional f' may not even be defined by a locally integrable function as illustrated by the following example.

13.2.2. The distributional derivative of the step function. The Heaviside step function is defined as

$$\theta(x) = \begin{cases} 1 \ , \ x \geq 0 \\ 0 \ , \ x < 0 \end{cases}$$

It is bounded and, hence, is locally integrable. Let $\{\theta'(x)\}$ denote the classical derivative wherever it exists. The step function has a jump discontinuity at $x = 0$. So the classical derivative does not exist at $x = 0$ and vanishes everywhere else, or it vanishes almost everywhere. Therefore it is a functional that has zero value on any test function:

$$\{\theta'(x)\} = 0 \quad \text{a.e.} \quad \Rightarrow \quad (\{\theta'\}, \varphi) = \int \{\theta'(x)\}\varphi(x)\, dx = 0\,.$$

Let us calculate its distributional derivative:

$$(\theta', \varphi) \overset{(1)}{=} -(\theta, \varphi') \overset{(2)}{=} -\int_0^\infty \varphi'(x)\, dx \overset{(3)}{=} \varphi(0) \overset{(4)}{=} (\delta, \varphi)\,.$$

Here (1) is by the rule (**13.4**), (2) is the value of the distribution θ on a test function, (3) is by evaluating the integral using the fundamental theorem of calculus and by that a test function has a bounded support ($\varphi(x) = 0$ if $|x| > R$ for some $R > 0$), and (4) follows from the definition (**13.1**) of the Dirac delta function. Since the equalities hold for all test functions, it is concluded that the distributional derivative of the step function is the Dirac delta function:

$$\theta'(x) = \delta(x)\,.$$

13.2.3. Distributions as linear functionals. Can the rule (**13.4**) be extended to all functionals (or distributions) that are not necessarily defined by locally integrable functions? This question must be answered affirmatively if one wants to develop calculus for general distributions and formulate equations for physical phenomena as equations in partial derivatives.

Let us reexamine the procedure for derivation of (**13.4**) with the purpose to identify steps in which the assumption that a distribution is defined by a locally integrable function was crucial. The goal is to find additional (sufficient) conditions on a general functional f to validate the derivation of (**13.4**).

As already noted $\varphi_h(x) = \varphi(x - h)$ is a test function if φ is such. So, the rule $(f_h, \varphi) = (f, \varphi_h)$ can be extended to any functional on $\mathcal{D}$. To evaluate the limit in (**13.3**), the linearity of the integral has been used to conclude that

$$\frac{(f, \varphi) - (f, \varphi_h)}{h} = (f, \psi_h)\,.$$

This is not true for a general functional on $\mathcal{D}$, unless this functional is *linear*. So, any distribution describing a physical quantity must be a *linear functional* on $\mathcal{D}$.

The space of test functions $\mathcal{D}$ is linear, that is, a linear combination of smooth functions with bounded supports is a smooth function with bounded support. A functional f is called *linear* if for any two numbers $c_{1,2}$ and any two functions $\varphi_{1,2}$ from $\mathcal{D}$

$$(f, c_1\varphi_1 + c_2\varphi_2) = c_1(f, \varphi_1) + c_2(f, \varphi_2)$$

in other words, the value of a linear functional on a linear combination of functions is the corresponding linear combination of values of the functional on each of these functions.

For example, the functional defined by the rule

$$(f, \varphi) = \varphi(0) + 1$$

is not linear. Indeed,

$$\begin{aligned} (f, c_1\varphi_1 + c_2\varphi_2) &= c_1\varphi_1(0) + c_2\varphi_2(0) + 1 \\ c_1(f, \varphi_1) + c_2(f, \varphi_2) &= c_1\varphi_1(0) + 1 + c_2\varphi_2(0) + 1 \neq (f, c_1\varphi_1 + c_2\varphi_2) \end{aligned}$$

The Dirac delta-function provides an example of a linear functional:

$$\begin{aligned} (\delta, c_1\varphi_1 + c_2\varphi_2) &= c_1\varphi_1(0) + c_2\varphi_2(0) \\ &= c_1(\delta, \varphi_1) + c_2(\delta, \varphi_2) \,. \end{aligned}$$

The zero function $\varphi(x) = 0$ is a test function. For any linear functional f on $\mathcal{D}$

$$(f, 0) = 0$$

because $(f, c\varphi) = c(f, \varphi)$, for any number c, and the property follows if $c = 0$.

13.2.4. The space of test functions. The right-hand side of (**13**.**4**) makes sense for an arbitrary linear functional f only if the derivative φ' is a test function. This is not true if $\mathcal{D}$ consists of functions from class C^p with $p < \infty$ because $\varphi' \in C^{p-1}$ and C^p is a subspace of C^{p-1}. So, the differentiation should not throw elements of $\mathcal{D}$ from $\mathcal{D}$ for consistency of (**13**.**4**). Thus, *the space of test functions must consist of function from class C^∞ with bounded support.*

This naturally leads to the conclusion that any distribution can be differentiated any number of times because the rule (**13**.**4**) can be used to calculate derivatives of derivatives:

$$(f^{(n)}, \varphi) = (-1)^n (f, \varphi^{(n)}) \,.$$

In particular, the Dirac delta function can be differentiated any number of times in the distributional sense:

$$(\delta^{(n)},\varphi) = (-1)^n(\delta,\varphi^{(n)}) = (-1)^n\varphi^{(n)}(0)\,. \tag{13.5}$$

13.2.5. Distributions as continuous functionals. The next step in derivation of (**13**.**4**) requires that

$$\lim_{h\to 0}(f,\psi_h) = (f,\lim_{h\to 0}\psi_h) = (f,\varphi')$$

A change of the order of taking the limit $h \to 0$ and calculating the value of f was established by means of the Lebesgue dominated convergence theorem which is not possible to apply for general linear functional on $\mathcal{D}$ that is not defined by a locally integrable function. Functionals for which the said order can be interchanged are called *continuous*.

A continuous functional is defined similarly to a continuous function. *A real-valued functional*

$$f:\ \mathcal{D}\to\mathbb{R}$$

is continuous at $\varphi\in\mathcal{D}$ *if for any sequence* $\{\varphi_n\}$ *converging to* φ *in* $\mathcal{D}$*, the numerical sequence* $\{(f,\varphi_n)\}$ *converges to the number* (f,φ)*:*

$$\{\varphi_n\}:\quad \varphi_n\to\varphi \text{ in } \mathcal{D} \quad\Rightarrow\quad \lim_{n\to\infty}(f,\varphi_n) = (f,\varphi)$$

and f *is continuous on* $\mathcal{D}$*, if it is continuous at every element of* $\mathcal{D}$*.*

Since now "points" in the domain are functions, one has to give a meaning (definition) to "*a sequence* $\{\varphi_n\}$ *converges to* φ *in* $\mathcal{D}$". In mathematical terms, this means that the functional space $\mathcal{D}$ must be equipped with *topology*.

If $\mathcal{D}$ were equipped with a metric or a distance function like spaces C^p (see Sec. **12.6**), then one can give a precise meaning to the convergence by requiring that $d(\varphi_n,\varphi)\to 0$ as $n\to\infty$, where $d(\phi,\varphi)$ is the distance function on $\mathcal{D}$. It is possible to define a distance on the space of functions from C^∞ with bounded support. However, it is also possible to prove that there exists no metric on $\mathcal{D}$ with respect to which $\mathcal{D}$ is a complete space. For example, the space C^p has a metric with respect to which it is complete. The completeness of a functional space guarantees that the limit function of a convergent functional sequence belongs to the space. So, completeness of $\mathcal{D}$ is essential for continuity of functionals on $\mathcal{D}$.

Fortunately, the metric is not the only way to introduce topology into a functional space. It will be done in the next section.

13.3. Distributions as linear continuous functionals. The analysis of basic calculus with distributions leads to the following concept of distributions as a generalization of classical functions. *A linear continuous function on a set of functions $\mathcal{D}$ is called a distribution.* Thus, among all functionals, a particular class is selected whose elements have two characteristic properties:

- linearity
- continuity

These two properties must be verified in order to find out if a given functional (f, φ) is a distribution or not. It is worth noting that *a linear functional is continuous if and only if it maps every null sequence in $\mathcal{D}$ to a numerical null sequence*:

$$\varphi_n \to 0 \text{ in } \mathcal{D} \quad \Rightarrow \quad \lim_{n\to\infty} (f, \varphi_n) = 0 \, .$$

In other words, a linear functional is continuous if and only if it is continuous at the zero function. For any $\{\varphi_n\}$ converging to φ in $\mathcal{D}$, the sequence $\psi_n = \varphi_n - \varphi$ is a null sequence in $\mathcal{D}$ and so is the numerical sequence (f, ψ_n) by linearity of f.

In the next two sections, the concept of a distribution as a linear continuous functional will be rigorously formulated, and the properties of distributions will be analyzed.

13.4. Exercises.

1. Give an example of a continuous mass density $\rho_\varepsilon(x)$ with support $|x| \leq a$ for which property (**13**.**2**) holds but
(i) $\lim_{a\to 0} \rho_a(0) = 0$;
(ii) $\lim_{a\to 0} \rho_a(0)$ does not exist.

2. The electric charge density can be positive and negative. Prove (**13**.**2**) if ρ_a is Lebesgue integrable and has a support in $|x| \leq a$.
Hint: Consider the densities of positive and negative charges. Are these densities integrable? If so, can the line of arguments given in Sec. **13.1.2** be applied to them to extend (**13**.**2**) to all integrable functions?

3. Define a mass density of a infinitely thin wire occupying a line segment of length L in $\mathbb{R}^3$ and having mass m that is distributed uniformly in the segment, as a linear functional on a suitable functional space $\mathcal{D}$.

4. Define a mass density of a infinitely thin plate occupying a rectangle of dimensions $a \times b$ in $\mathbb{R}^3$ and having mass m that is distributed uniformly in the plate, as a linear functional on a suitable functional space $\mathcal{D}$.

5. Define a mass density of a infinitely thin sphere of radius R in $\mathbb{R}^3$ that has mass m which is distributed uniformly over the sphere, as a linear functional on a suitable functional space $\mathcal{D}$.

6. Define an electric charge density of a infinitely thin dielectric sphere of radius R in $\mathbb{R}^3$ whose one hemisphere has a positive charge Q_+ uniformly distributed and other hemisphere has a negative charge Q_- uniformly distributed, as a linear functional on a suitable functional space $\mathcal{D}$.

7. (i) Define the mass density of n particles of masses m_j, $j = 1, 2, ..., n$, moving in $\mathbb{R}^3$ along smooth trajectories $x = x_j(t)$, where t is time, as a family of linear functionals on a suitable functional space $\mathcal{D}$ that are labeled by parameter t. In other words, for each time moment t, the mass density is a linear functional on $\mathcal{D}$.
(ii) Is it possible to define the momentum density of this system as a family of vector-valued distributions labeled by time t? In other words, every component of the momentum density is a linear functional on $\mathcal{D}$ for each fixed moment of time t.
(iii) If particles interacts repulsively in accord with the Coulomb law, each particle having a charge q_j, find the energy density of the system as a linear functional on $\mathcal{D}$ for each fixed moment of time t.
(iv) The same as (iii), but change the repulsive force by the same attractive force.

14. The space of test functions

The objective is to give a precise description of tests functions and show that this class of functions is rich enough to approximate practically any type of functions used in applications. The latter is known as *approximation theorems* for test functions.

14.1. Definition of $\mathcal{D}$. A function φ on an open set $\Omega \subset \mathbb{R}^N$ is called a *test function* if

(i) *φ is from class $C^\infty(\Omega)$;*
(ii) *φ has a bounded support,* $\operatorname{supp}\varphi \subset B_R \cap \Omega$

for some ball B_R. A collection of all test functions on Ω is denoted by $\mathcal{D}(\Omega)$ and called the *space of test functions on* Ω. If $\Omega = \mathbb{R}^N$ or $\Omega = (a,b)$, then $\mathcal{D}(\mathbb{R}^N) = \mathcal{D}$ or $\mathcal{D}((a,b)) = \mathcal{D}(a,b)$ for brevity. Clearly, $\mathcal{D}(\Omega)$ is a linear space:

$$\varphi_{1,2} \in \mathcal{D}(\Omega) \quad \Rightarrow \quad c_1\varphi_1 + c_1\varphi_2 \in \mathcal{D}(\Omega)$$

for any numbers $c_{1,2}$. A second observation is that all partial derivatives of a test function are test functions:

$$\varphi \in \mathcal{D}(\Omega) \quad \Rightarrow \quad D^\alpha\varphi \in \mathcal{D}(\Omega)\,.$$

for any $|\alpha| \geq 0$.

14.1.1. Analytic functions vs test functions. For any $\varphi \in \mathcal{D}(\Omega)$, its support $K = \operatorname{supp}\varphi$ is a closed bounded subset (a compact) in Ω (see Sec. **1.1.10**). Since Ω is open, the distance between K and the boundary $\partial\Omega$ is not zero, and therefore there exists a neighborhood of the boundary $\partial\Omega$ that does not overlap with K. This implies that φ *and all its partial derivatives $D^\alpha\varphi$ vanish in a neighborhood of any point of $\partial\Omega$*:

$$D^\alpha\varphi(x) = 0\,, \quad d(x,\partial\Omega) < \delta\,, \quad \alpha \geq 0\,,$$

for some $\delta > 0$ (δ depends on φ). In particular, all $D^\alpha\varphi$ can be continuously extended by zeros to the boundary $\partial\Omega$. For example, if $\Omega = (a,b)$, then there exists a sufficiently small $\delta > 0$ such that $\operatorname{supp}\varphi \subseteq [a+\delta, b-\delta]$, and all derivatives $D^\alpha\varphi(x)$ vanish for $a < x < a+\delta$ and $b > x > b-\delta$. Therefore φ and all its derivatives can be continuously extended to $[a,b]$ so that $D^\alpha\varphi(a) = D^\alpha\varphi(b) = 0$.

Furthermore despite being from class C^∞, the test function are not analytic in Ω. Suppose first that φ is from $\mathcal{D}(\mathbb{R})$ and $\operatorname{supp}\varphi = [a,b]$. Then $\varphi(x) = 0$ if $x \leq a$ and $\varphi(x) \neq 0$ if $a < x < a+\varepsilon$ for some $\varepsilon > 0$. If φ were analytic at $x = a$, then its values near $x = a$ would be given by a power series about a:

$$\varphi(x) = c_0 + \sum_{n=1}^{\infty} c_n (x-a)^n$$

for all $|x-a| < R$ where $R > 0$ is the radius of convergence. By the Taylor theorem, the coefficients in the power series representation of φ are proportional to the derivatives

$$c_n = \frac{\varphi^{(n)}(a)}{n!}\,.$$

However, by continuity

$$\varphi^{(n)}(a) = \lim_{x\to a^-} \varphi^{(n)}(x) = 0\,.$$

because the derivatives vanish $\varphi^{(n)}(x) = 0$ for $x < a$. It follows from the power series representation that $\varphi(x) = 0$ for all $|x-a| < R$ for some $R > 0$, which cannot be true because $\varphi(x) \neq 0$ for $x > a$. Thus, *the only analytic function in $\mathcal{D}$ is the zero function!*

The conclusion can readily be extended to test functions of several variables from $\mathcal{D}(\Omega)$. If $K = \operatorname{supp}\varphi$, then the boundary ∂K lies in Ω as shown above. The test function is not analytic at any point of ∂K. The support of any non-zero test function is not empty, and, hence any non-zero test function is not analytic in Ω.

14.1.2. Topology in $\mathcal{D}$. *A sequence $\{\varphi_n\}$ is said to converge to φ in $\mathcal{D}(\Omega)$ if*

(i) *There exists a compact $K \subset \Omega$ that contains supports of all elements of the sequence,*

$$\operatorname{supp}\varphi_n \subset K\,;$$

(ii) *Sequences of all partial derivatives, $D^\alpha\varphi_n$, converge uniformly to the corresponding partial derivatives of the limit function,*

$$\lim_{n\to\infty} \sup |D^\alpha\varphi(x) - D^\alpha\varphi_n(x)| = 0\,, \qquad |\alpha| \geq 0\,.$$

and in this case one writes

$$\varphi_n \to \varphi \quad \text{in } \mathcal{D}\,.$$

Clearly, the limit of a convergent sequence is unique because there is only one test function with the property $\sup |D^\alpha\varphi(x)| = 0$ for any α; it is the zero function $\varphi(x) = 0$. A consistency of this definition follows from Theorems **1.6.4** and **1.6.5**. The uniform convergence of sequences $\{D^\alpha\varphi_n\}$ for any α guarantees that the limit function is from the class C^∞. Since supports of all terms in the sequence $\{\varphi_n\}$ lie in K that is a proper subset of Ω, the support of the limit function must also be in

$K \subset \Omega$ (be a proper subset in Ω). So, conditions (i) and (ii) guarantee that the limit function belongs to $\mathcal{D}(\Omega)$.

The condition (i) might seem unnecessary. However if it is lifted, then there are sequences in $\mathcal{D}(\Omega)$ that satisfy (ii) but the limit function is not in $\mathcal{D}(\Omega)$. For example, let $\phi \in \mathcal{D}(-a, 2a)$, $a > 0$, and $\operatorname{supp}\phi = [0, a]$. Put $\varphi_n(x) = \phi(x - a + \frac{a}{n})$ which is a test function with support $[a - \frac{a}{n}, 2a - \frac{a}{n}]$ that is a proper subset of $(-a, 2a)$ for any $n = 1, 2, \ldots$. Then φ_n and all its derivatives converge uniformly on $(-a, 2a)$ to $\varphi(x) = \phi(x - a) \in C^\infty(-a, 2a)$ and the corresponding derivatives of φ. However the support of the limit function is $[a, 2a]$ that is not a subset of $(-a, 2a)$ and, hence, the limit function is not in $\mathcal{D}(-a, 2a)$. The condition (i) is ensures that the limit function is supported in Ω.

14.1.3. Subspaces of the space of test functions. Let Ω' be an open subset in Ω. Then $\mathcal{D}(\Omega')$ is a subspace of $\mathcal{D}(\Omega)$ because $\operatorname{supp}\varphi \subset \Omega' \subset \Omega$ if $\varphi \in \mathcal{D}(\Omega')$ and, hence, $\varphi \in \mathcal{D}(\Omega)$. Moreover, if $\varphi_n \to \varphi$ in $\mathcal{D}(\Omega')$, then the sequence φ_n also converges to φ in topology of the larger space $\mathcal{D}(\Omega)$ because any φ from $\mathcal{D}(\Omega')$ vanishes outside Ω' and hence $\sup_\Omega |D^\alpha\varphi| = \sup_{\Omega'} |D^\alpha\varphi|$ for any α.

14.2. How many elements are in $\mathcal{D}$, anyway? It was shown in Sec.**14.1.1** that analytic functions are not in $\mathcal{D}$ despite being from class C^∞. So, the condition of having a bounded support for a smooth functions looks rather restrictive. It seems natural to ask:

(i) *Do there exist C^∞ functions with bounded support?*
(ii) *If affirmative, how big is the set of such functions?*

14.2.1. The hat function. Put

$$\omega(x) = \begin{cases} e^{-\frac{1}{x}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Evidently, the function ω belongs to class C^∞ for $x < 0$ and $x > 0$. It is continuous at $x = 0$ because $e^{-\frac{1}{x}} \to 0 = \omega(0)$ as $x \to 0^+$. The derivative at $x = 0$ exists and is equal to zero

$$\omega'(0) = \lim_{x\to 0} \frac{\omega(x) - \omega(0)}{x} = 0$$

because the fraction is identically zero for $x < 0$ and for $x > 0$ it is equal to $ye^{-y} \to 0$ where $y = \frac{1}{x} \to \infty$ as $x \to 0^+$. On the other hand, $\omega'(x) = 0$ for $x < 0$ so that $\omega'(x) \to 0$ as $x \to 0^-$, and for $x > 0$ $\omega'(x) = \frac{1}{x^2}e^{-\frac{1}{x}} = y^2e^{-y} \to 0$ where $y = \frac{1}{x} \to \infty$ when $x \to 0^+$. Therefore $\omega'(x) \to \omega'(0) = 0$ as $x \to 0$, which means that the derivative is continuous at $x = 0$. This argument can recursively be extended to show that $\omega^{(n)}(0) = 0$ for any $n > 0$ and $\omega^{(n)}(x) \to 0 = \omega^{(n)}(0)$ as $x \to 0$. First note that

$$\omega^{(n)}(x) = \begin{cases} p_n(\frac{1}{x})e^{-\frac{1}{x}}\,, & x > 0 \\ 0\,, & x < 0 \end{cases}$$

where p_n is a polynomial or degree $2n$. It follows from

$$\lim_{x\to 0^+} x^{-m}e^{-\frac{1}{x}} = \lim_{y\to\infty} y^m e^{-y} = 0\,, \quad m = 0, 1, 2, \dots\,,$$

that

$$\lim_{x\to 0} \omega^{(m)}(x) = 0\,.$$

Then the relation

$$\omega^{(n)}(0) = \lim_{x\to 0} \frac{\omega^{(n-1)}(x) - \omega^{(n-1)}(0)}{x} = 0\,, \quad n = 1, 2, ...,$$

is proved by mathematical induction (it is true for $n = 1$ as shown above). Therefore $\omega \in C^\infty(\mathbb{R})$ but it is not analytic at $x = 0$ because it has no power series representation near $x = 0$ (all derivatives vanish at $x = 0$).

Using the function ω, it is now not difficult to construct a test function with support being the interval $[a, b]$:

$$\varphi_{a,b}(x) = \omega(x - a)\omega(b - x)\,. \tag{14.1}$$

Furthermore, if this function is multiplied by any function from class C^∞, the resulting function is also a test function: $\varphi(x) = a(x)\varphi_{a,b}(x) \in \mathcal{D}$ for any $a \in C^\infty$. This shows that $\mathcal{D}$ is roughly as big as C^∞.

If $x \in \mathbb{R}^N$, then $|x|$ does not have continuous partials at $x = 0$. So, the function $\omega(|x| - a)\omega(b - |x|)$ is not smooth enough to be from $\mathcal{D}(\mathbb{R}^N)$. However when $b = -a$, the product is equal to $e^{-\frac{2a}{a^2-|x|^2}}$ for $|x| < a$ and vanishes otherwise. This function is smooth for $|x| < a$ because it is a smooth function of the polynomial $|x|^2$. Therefore put

$$\omega_a(x) = \begin{cases} c_a \exp\Big(-\frac{a^2}{a^2-|x|^2}\Big)\,, & |x| < a \\ 0\,, & |x| > a \end{cases} \tag{14.2}$$

where c_a is a normalization constant to be defined later. The function ω_a is called a *hat function* because its graph resembles a hat for $x \in \mathbb{R}^2$. It is supported in the ball $|x| \le a$.

If $x \in \mathbb{R}$, then near $x = \pm a$, the function ω_a has a behavior similar to that of the function ω considered above near the point $x = 0$. For example, put $y = x - a$. Then

$$\omega_a(x) = c_a \exp\left(-\frac{a^2}{y(y+2a)}\right) \approx c_a \exp\left(-\frac{a}{2y}\right)$$

for small y, and similarly for $x = -a$. Therefore ω_a is from C^∞ but it is not analytic at $x = \pm a$.

When $x \in \mathbb{R}^N$, ω_a depends only the radial variable $r = |x|$ and exhibits the same behavior in it near $r = a$ as in the one-variable case. Therefore ω_a is continuous across the sphere $|x| = a$. Continuity of partial derivatives is verified similarly to the one-variable case. Since for $|x| \neq a$, ω_a is from class C^∞, its partials are easy to find. If $|x| > a$, then $D_j\omega_a(x) = 0$. If $|x| < a$, then by the chain rule

$$D_j\omega_a(x) = \frac{x_j}{r} D_r\omega_a(x)\,.$$

Therefore $D_j\omega_a(x) \to 0$ as $x \to y$ for any $|y| = a$ because $|D_j\omega_a(x)| \leq |D_r\omega_a(x)| \to 0$ as $r \to a^-$. When $|x| = a$, the partials $D_j\omega_a$ are defined by the limit

$$D_j\omega_a(x) = \lim_{h\to 0} \frac{\omega_a(x + he_j) - \omega_a(x)}{h}\,,$$

where $|x| = a$ and e_j is the standard basis vector parallel to the j^{th} coordinate axis. The partials are continuous at the sphere $|x| = a$ if this limit vanishes for any $|x| = a$. If $h > 0$, then the point $x + he_j$ lies outside the ball $|x| \leq a$ for any $|x| = a$. Therefore the right limit $h \to 0^+$ vanishes because the numerator is identically zero for any $h > 0$. If $h \to 0^-$, then the above limit becomes $\frac{1}{h}e^{c/h} \to 0$ where $c = \frac{a}{2\cos(\theta_j)} > 0$ and θ_j is the angle between x and e_j. Thus, $D_j\omega_a(x) = 0$ for any $|x| = a$ and, hence, $D_j\omega_a$ is continuous across the sphere $|x| = a$. Continuity of $D^\alpha\omega_a$ is established by mathematical induction by repeating the above reasonings for $D_jD^\alpha\omega_a$ assuming that $D^\alpha\omega_a$ is continuous across the sphere $|x| = a$. Note that for $|x| < a$, $D^\alpha\omega_a(x) = r^{-|\alpha|}P(x)D_r^{|\alpha|}\omega_a(x)$, where $P(x)$ is a polynomial of degree $|\alpha|$, and $D^\alpha\omega_a(x) = 0$ for $|x| > a$. Therefore $D_jD^\alpha\omega_a(x) \to 0$ as $x \to y$ for any $|y| = a$. The partials $D_jD^\alpha\omega_a(x)$ for $|x| = a$ are defined by the above limit where ω_a is replaced by $D^\alpha\omega_a$ and is shown to vanish if $D^\alpha\omega_a(x) = 0$ for $|x| = a$ (by the hypothesis of mathematical induction).

So, the hat function ω_a is a test function. It is not analytic on the sphere $|x| = a$. Furthermore, for any $\phi \in C^\infty(\mathbb{R}^N)$, the product $\phi\omega_a \in \mathcal{D}$ is also a test function.

The normalization constant c_a is chosen so that integral of ω_a is equal to one:

$$\int \omega_a(x)\, d^N x = 1\,.$$

If $\sigma_{_N}$ denotes the surface area of a unit sphere in $\mathbb{R}^N$, then using spherical coordinates, where $|x| = r$,

$$\int \omega_a(x)\, d^N x = \sigma_{_N} \int_0^a e^{-\frac{a^2}{a^2-r^2}} r^{N-1} dr = a^N \sigma_{_N} \int_0^1 e^{-\frac{1}{1-u^2}}\, u^{N-1}\, du\,.$$

Therefore

$$c_a = \frac{1}{a^N \sigma_{_N} c_{_N}}\,, \qquad c_{_N} = \int_0^1 e^{-\frac{1}{1-u^2}}\, u^{N-1}\, du\,.$$

14.2.2. Properties of the hat function. The hat function has the following scaling property:

$$\omega_a(x) = \frac{1}{a^N}\, \omega_1\Big(\frac{x}{a}\Big)\,, \quad x \in \mathbb{R}^N\,.$$

When $a \to 0^+$, the support of the hat function is shrinking but its integral remains 1 for any $a > 0$. In this limit the behavior of the hat function resembles a limiting process for the mass density of a point particle of unit mass. Therefore by the Lebesgue dominated convergence theorem and the scaling and normalization properties

$$\begin{aligned} \lim_{a\to 0^+} (\omega_a, \varphi) &= \lim_{a\to 0^+} \int \omega_a(x)\, \varphi(x)\, d^N x = \lim_{a\to 0^+} \int_{|y|\le 1} \omega_1(y)\, \varphi(ay)\, d^N y \\ &= \varphi(0) \int_{|y|\le 1} \omega_1(y)\, d^N y = \varphi(0) = (\delta, \varphi)\,, \end{aligned} \tag{14.3}$$

for any test function φ, where $x = ay$. Note that integrand has an integrable bound independent of a: $|\omega_1(y)\, \varphi(ay)| \le M\omega_1(y)$ where $M = \sup|\varphi|$. The hat function defines a family of tests functions such that, when $a \to 0^+$, it converges to the Dirac delta function in the sense of distributions.

14.2.3. Bump functions. Let $x \in \mathbb{R}$. Consider the function

$$\eta_\delta(x) = \int_{-\infty}^x \omega_\delta(y)\, dy\,.$$

By construction, it has the following properties. First, $\eta_\delta \in C^\infty$. Second, $\eta_\delta(x) = 0$ if $x < -\delta$ and $\eta_\delta(x) = 1$ if $x > \delta$. Since the hat function is non-negative, $0 \le \eta_\delta(x) \le 1$ if $|x| \le \delta$ and η_δ is increasing monotonically. It is a smooth regularization of the step function. In fact, $\eta_\delta(x) \to \theta(x)$ as $\delta \to 0^+$ for any $x \neq 0$ and $\eta_\delta(0) = \frac{1}{2}$. For any $a < b$

$$\varphi(x) = \eta_\delta(x-a) - \eta_\delta(x-b) \in \mathcal{D}$$

is a test function with support $[a-\delta, b+\delta]$. It takes unit value in the interval $[a+\delta, b-\delta]$, assuming that $b - a > 2\delta$.

This shows that for any bounded interval one can construct a test function that takes unit value in a neighborhood of the interval. Its graph looks like a bump (a flat top with smooth transitions at the edges). A smooth function that takes unit value in a neighborhood of a set $\Omega \subset \mathbb{R}$ is called a *bump function* for the set Ω. If Ω is bounded, then its bump function is a test function.

Let us construct multidimensional bump functions. For any set $\Omega \subset \mathbb{R}^N$, a neighborhood Ω_δ of Ω of radius δ is the union of all open balls of radius δ centered at every point of Ω:

$$\Omega_\delta = \bigcup_{x\in\Omega} B_\delta(x)\,.$$

so that the distance between Ω and the boundary of Ω_δ is $\delta > 0$.

THEOREM **14.1**. *Let Ω be a subset in $\mathbb{R}^N$ and Ω_δ be a neighborhood of Ω of radius $\delta > 0$. Then for any positive $a > 0$ there exists a function η_a with the following properties:*

(i) $\eta_a \in C^\infty\,;$

(ii) $0 \le \eta_a(x) \le 1\,;$

(iii) $\eta_a(x) = 1\,, \quad x \in \Omega_a\,;$

(iv) $\eta_a(x) = 0\,, \quad x \notin \overline{\Omega}_{3a}\,;$

(v) $|D^\beta \eta_a(x)| \le M_\beta a^{-\beta}$

for some constant M_β independent of a.

Theorem **14.1** is proved by verifying properties (i)-(v) for the convolution of a hat function with the characteristic function of a neighborhood of Ω of radius $2a$:

$$\eta_a(x) = \int \chi_{\Omega_{2a}}(y)\,\omega_a(x-y)\,d^N y = \int_{\Omega_{2a}} \omega_a(x-y)\,d^N y\,.$$

By Theorem **7.2**, all partial derivatives of the convolution are continuous everywhere. So, $\eta_a \in C^\infty$. The second property follows from that values of the characteristic function are either 0 or 1 and that the hat function is non-negative:

$$0 \le \eta_a(x) \le \int \omega_a(x-y)\, d^N y = \int \omega_a(z)\, d^N z = 1\,.$$

To verify the remaining properties, consider three neighborhoods of Ω:

$$\Omega \subset \Omega_a \subset \Omega_{2a} \subset \Omega_{3a}\,.$$

Let $x \in \Omega_a$. One has

$$\begin{aligned}\eta_a(x) &\overset{(1)}{=} \int_{B_a(x)} \chi_{\Omega_{2a}}(y)\, \omega_a(x-y)\, d^N y \\ &\overset{(2)}{=} \int_{B_a(x)} \omega_a(x-y)\, d^N y \overset{(3)}{=} \int_{B_a} \omega_a(z)\, d^N z \overset{(4)}{=} 1\,.\end{aligned}$$

Here the equality (1) follows from the property

$$\omega_a(x-y) = 0\,, \quad |x-y| > a\,,$$

so that the integration region in the convolution integral can be reduced to a ball of radius a and centered at x, (2) is valid because, if $x \in \Omega_a$, then the ball $B_a(x)$ lies in Ω_{2a} and therefore $\chi_{\Omega_{2a}}(y) = 1$ if $y \in B_a(x)$, (3) is obtained by the shift of the integration variable $z = x - y$, and (4) is by the normalization property of the hat function.

Finally, if x does not belong to the closure $\overline{\Omega}_{3a}$, the open ball $B_a(x)$ has no overlap with Ω_{2a}. This implies that the hat function $\omega_a(x-y)$ vanishes for any $y \in \Omega_{2a}$ so that

$$\eta_a(x) = \int_{\Omega_{2a}} \omega_a(x-y)\, d^N y = 0\,, \quad x \notin \overline{\Omega}_{3a}$$

By Theorem **7.2** one can show that

$$|D^\beta \eta_a(x)| \le \int |D^\beta \omega_a(z)|\, d^N z\,.$$

Then the property (v) follows from from the scaling property of the hat function, $D^\beta \omega_a(z) = a^{-N-\beta} D^\beta \omega_1(z/a)$. The proof is complete.

There is a useful consequence of this theorem.

Corollary **14.1**. *Let K be a compact subset of an open set $\Omega \subseteq \mathbb{R}^N$. Then there exists a test function $\eta_{_K} \in \mathcal{D}(\Omega)$ that takes values in $[0,1]$ and is equal to 1 in a neighborhood of K.*

Since Ω is open, K is a proper subset in Ω, and the distance between the boundary $\partial\Omega$ and K is not zero. In Theorem **14.1**, take $a = \frac{1}{4} d(\partial K, \partial\Omega)$. Then $\eta_{_K} = \eta_a \in \mathcal{D}(\Omega)$ is a test function with required properties.

For any set Ω, a function η with properties stated in Theorem **14.1** will be called a *bump function for a set* Ω. In particular, a bump function for any bounded set in $\mathbb{R}^N$ is a test function from $\mathcal{D}$. Furthermore, for any function u from class C^∞, one can always find a test function that is equal to u in a neighborhood of any open bounded set Ω. Indeed, the test function with required properties is

$$\varphi(x) = u(x)\eta_{\Omega}(x) \in \mathcal{D}\,,$$

where η_Ω is a bump function for Ω.

14.2.4. Regularization of a locally integrable function. Let f be a locally integrable function in $\mathbb{R}^N$. Then the convolution of f and a test function ω,

$$(\omega * f)(x) = \int \omega(x-y)f(y)\,d^Ny\,, \tag{14.4}$$

exists because the integral converges absolutely thanks to local integrability of f and to that the support of ω lies in a ball $|x| \le R$:

$$\int |\omega(x-y)f(y)|\,d^Ny \le M_0 \int_{|x-y|<R} |f(y)|\,d^Ny < \infty\,,$$

where $M_0 = \sup|\omega(x)| < \infty$.

By Theorem **7.2** the convolution $(\omega * f)$ is from class C^∞. Indeed, for all x in a ball $|x| < R_1$, the integrand $\omega(x-y)f(y)$ vanishes for all $|y| > R + R_1$ if the support of ω lies in a ball of radius R. Then for any $\alpha \ge 0$, any partial derivative of the integrand has an integrable bound independent of x:

$$|D_x^\beta \omega(x-y)f(y)| \le M_\beta |f(y)| \in \mathcal{L}(B_{R+R_1})$$

for all $x \in \Omega$, where $M_\beta = \sup|D^\beta\omega(x)| < \infty$, because f is locally integrable. By Theorem **7.2**, $\omega * f$ has continuous partial derivatives of any order in any ball $|x| < R_1$. Since R_1 is arbitrary, the convolution $\omega * f$ is from class C^∞ and

$$D^\beta(\omega * f)(x) = \int D_x^\beta \omega(x-y)f(y)\,d^Ny\,. \tag{14.5}$$

Furthermore, if the support of f is bounded, then the support of $\omega * f$ is also bounded. The convolution vanishes for all $|x| > R + R_f$ if the supports of ω and f lie in balls of radii R and R_f, respectively. In this case, the convolution is a test function.

Thus, with any locally integrable function f that has a bounded support one can associate a test function $f_\omega = \omega * f$. *This shows that* *the space of test functions is roughly as "big" as the space of integrable functions with bonded supports.*

For any non-negative test function ϕ whose integral is normalized to 1,

$$\int \phi(x)\, d^N x = 1\,,$$

and any $a > 0$, put

$$\phi_a(x) = \frac{1}{a^N}\phi\Big(\frac{x}{a}\Big)\,.$$

Then ϕ_a is a test function. If the support of ϕ lies in a ball of radius R, then the support of ϕ is in a ball of radius aR. The convolution

$$f_a(x) = (\phi_a * f)(x) = \int \phi_a(x-y) f(y)\, d^N y \tag{14.6}$$

is called a *regularization* of a locally integrable function f. In particular, one can take $\phi_a = \omega_a$ (a hat function). A regularizing test function ϕ_a has the same characteristic property (**14.3**) as the hat function. A proof is similar to the line of reasoning in (**14.3**).

The regularization of a locally integrable function with bounded support is a test function. It is a smooth function that vanishes outside a neighborhood of support of f. If $K = \operatorname{supp} f$ and $\operatorname{supp}\phi \subset B_R$, then

$$(\phi_a * f)(x) = 0\,, \quad d(x,K) \geq aR\,. \tag{14.7}$$

This follows from (**14.4**) for $\omega(x-y) = \phi_a(x-y)$ because $\phi_a(x-y) = 0$ if $|x-y| \geq aR$ whereas $y \in K$ so that whenever the distance between x and K exceeds aR, the convolution integral vanishes.

14.2.5. Regularization of continuous functions. The term "regularization" implies that the function being regularized is close to its regularization in some sense. This is indeed so. Let us show that *a regularization $\phi_a * f$ of a continuous function f on $\mathbb{R}^N$ converges to f pointwise in the limit $a \to 0^+$*:

$$\lim_{a\to 0^+} (\phi_a * f)(x) = f(x)\,, \quad x \in \mathbb{R}^N\,.$$

Recall from Sec.**1.2.6** that f is uniformly continuous on any compact $K \subset \mathbb{R}^N$. This means that for any K and any $\varepsilon > 0$ there exists δ (that generally depends on K and ε) such that

$$|f(x) - f(y)| < \varepsilon \quad \text{whenever} \quad |x-y| < \delta\,,$$

for all x and y in K. Take K that contains a neighborhood of x, fix $\varepsilon > 0$, and find the corresponding δ. Using the normalization and scaling properties of $\phi_a(x)$ in the regularization (**14.6**)

$$
\begin{aligned}
|(\phi_a * f)(x) - f(x)| &= \Big| \int \phi_a(y)\Big(f(x-y) - f(x)\Big)\, d^N y \Big| \\
&= \int_{|z|<R} \phi(z) \Big| f(x - az) - f(x) \Big|\, d^N z \\
&< \varepsilon \int_{|z|<R} \phi(z)\, d^N z = \varepsilon\,, \quad a < \delta/R\,. \qquad (\mathbf{14.8})
\end{aligned}
$$

Since ε is arbitrary, this implies that $(\phi_a * f)(x)$ converges to $f(x)$ as $a \to 0^+$ for any x. So, $\phi_a * f$ is an *approximation* to f at every point by a smooth function. If f is continuous and has bounded support, then the approximation is a test function.

14.3. Approximation theorems. A regularization of any continuous function converges to the function pointwise in the limit $a \to 0^+$. Can a regularization be used to approximate functions from various metric functional spaces by test functions? The answer is provided by the so called approximation theorems for test functions.

A typical approximation theorem asserts the following. Let $\mathcal{F}$ be a linear functional normed space so that the distance between any two elements $f, g \in \mathcal{F}$ is given by $d(f, g) = \|f - g\|_{\mathcal{F}}$ where $\|\cdot\|_{\mathcal{F}}$ is a norm on $\mathcal{F}$. Let the space of test functions be a subset in $\mathcal{F}$. Then for any $\varepsilon > 0$ and any $f \in \mathcal{F}$ there exists a test function φ such that

$$
d(\varphi, f) = \|\varphi - f\|_{\mathcal{F}} < \varepsilon\,.
$$

In other words, any function from $\mathcal{F}$ can be approximated with any desired accuracy by a test function.

In what follows, it will be shown that for any open $\Omega \subset \mathbb{R}^N$ this is true for $\mathcal{F} = C_0^p(\Omega)$, $p \geq 0$, and $\mathcal{F} = \mathcal{L}_p(\Omega)$, $p \geq 1$, where $C_0^p(\Omega)$ is a subset of compactly supported functions from $C^p(\Omega)$. This means that $\mathcal{D}(\Omega)$ is dense in these spaces and test functions in them are much like rational numbers in reals. Furthermore, a test function φ is obtained by a regularization of f as discussed above. The reader not interested in mathematical details may skip the rest of this section.

14.3.1. $\mathcal{D}$ as a dense subset of C_0^0. The space of test functions is a subspace in the space of continuous bounded functions in which the distance is defined by the supremum norm. Clearly, $\mathcal{D}$ cannot be dense in this space. For example, a unit function on $\mathbb{R}^N$ is continuous and bounded, but the distance between it and any test function cannot be arbitrary small because

$$\|1-\varphi\|_\infty = \sup|1-\varphi(x)| \geq \sup_{|x|>R}|1-\varphi(x)| = 1$$

if the support of φ lies in B_R.

Let $C_0^0 \subset C^0$ denote a subspace of continuous functions on $\mathbb{R}^N$ with bounded support. Then $\mathcal{D} \subset C_0^0$. The first approximation theorem states that *$\mathcal{D}$ is dense in C_0^0 with respect to the supremum norm.* This means that *for any $f \in C_0^0$ there exists a test function $\varphi \in \mathcal{D}$ such that the distance $\|f-\varphi\|_\infty$ can be made arbitrary small with a suitable choice of φ.*

To prove this assertion, note that f and its regularization $f_a \in \mathcal{D}$ have bounded supports, that is, they both vanish outside a ball of large enough radius, $f_a(x) = f(x) = 0$ for $|x| > R_f$. This implies that the inequality (**14.8**) holds for all $x \in \mathbb{R}^N$ because one can take K to be the ball $|x| \leq R_f$. Therefore one can take the supremum in the left-hand side of (**14.8**):

$$\|f_a - f\|_\infty = \sup|f_a(x) - f(x)| \leq \varepsilon\,.$$

This shows that a regularization f_a of a continuous function f with bounded support converges to f *uniformly* as $a \to 0^+$ and, hence, $\mathcal{D}$ is dense in C_0^0.

14.3.2. $\mathcal{D}$ as a dense subset of C_0^p. Let C_0^p, $p \geq 0$, be the subspace of all function from class C^p that have bounded support. Let us show that a regularization f_a of any $f \in C_0^p$ and all partial derivatives $D^\beta f_a$ up order p converge uniformly to f and $D^\beta f$, respectively, as $a \to 0^+$. This comprises a generalization of the approximation theorem proved above.

THEOREM **14.2**. *$\mathcal{D}$ is dense in C_0^p. In particular, for any $f \in C_0^p$, a regularization f_a of f is a test function from $\mathcal{D}$ and*

$$\lim_{a\to 0^+}\|f-f_a\|_{C^p} = \lim_{a\to 0^+}\max_{|\beta|\leq p}\sup_x |D^\beta f(x) - D^\beta f_a(x)| = 0\,.$$

A proof is analogous to the case of C_0^0. Owing to the boundedness of support of f, its partials $D^\beta f$, $|\beta| \leq p$, are uniformly continuous on $\mathbb{R}^N$, that is, for any $\varepsilon > 0$ and every $|\beta| \leq p$ one can find δ_β such that

$$\left|D_x^\beta f(x) - D_y^\beta f(y)\right| < \varepsilon \quad \text{whenever } |x-y| < \delta_\beta\,.$$

In the integral representation (**14.5**) of the convolution with a test function, the integration by parts is permitted up to p times if $f \in C_0^p$ so that

$$D^\beta(\phi_a * f)(x) = \int D_x^\beta \phi_a(x-y) f(y)\, d^N y = \int \phi_a(x-y) D_y^\beta f(y)\, d^N y$$
$$= (\phi_a * D^\beta f)(x)\,, \quad 0 \le \beta \le p\,,$$

because $D_x \phi_a(x-y) = -D_y \phi_a(x-y)$. Then replacing f by $D^\beta f$ in (**14.8**), it is concluded that

$$|D^\beta f_a(x) - D^\beta f(x)| < \varepsilon\,, \quad a < \delta/R\,,$$

for all $x \in \mathbb{R}^N$ and all $|\beta| \le p$ if $\delta = \min_\beta\{\delta_\beta\}$. This implies that $D^\beta f_a$ converges to $D^\beta f$ uniformly and that $\mathcal{D}$ is dense in C_0^p with respect to the C^p norm:

$$\|f_a - f\|_{C^p} \le \varepsilon\,, \quad a < \delta/R\,.$$

14.3.3. $\mathcal{D}(\Omega)$ as a dense subset of $C_0^p(\Omega)$. *A function f is said to belong to $C_0^p(\Omega) \subset C^p(\Omega)$, where Ω is an open set in $\mathbb{R}^N$, if the support of f is a bounded subset of Ω.* Since $\operatorname{supp} f$ is a closed subset in an open set Ω, the distance between it and the boundary $\partial\Omega$ is not zero, similarly to test functions from $\mathcal{D}(\Omega)$. This implies that any function f from class $C_0^p(\Omega)$ and all partial derivatives $D^\alpha f$ up to order p vanish in a neighborhood of the boundary $\partial\Omega$. So, $\mathcal{D}(\Omega) \subset C_0^p(\Omega)$. The same reasonings as in the proof of Theorem **14.2** lead to a consequence that $\mathcal{D}(\Omega)$ is dense in $C_0^p(\Omega)$ relative to the C^p norm.

COROLLARY **14.2**. *For any open set $\Omega \subset \mathbb{R}^N$, the space of test functions $\mathcal{D}(\Omega)$ is dense in $C_0^p(\Omega)$. In particular, a regularization f_a of $f \in C_0^p(\Omega)$ and its partials $D^\beta f_a$, $|\beta| \le p$, converge uniformly to f and the corresponding partials $D^\beta f$ on Ω as $a \to 0^+$.*

It should be noted that the support of a regularization f_a of $f \in C_0^p(\Omega)$ is not in Ω for any $a > 0$. However, $f_a \in \mathcal{D}(\Omega)$ for all small enough a as follows from Corollary **14.1**. If $\delta > 0$ is the distance between the boundary $\partial\Omega$ and the support of f, then the support of f_a lies in Ω if $Ra < \frac{1}{4}\delta$ (support of ϕ_a lies in a ball of radius aR).

14.3.4. $\mathcal{D}(\Omega)$ as a dense subset of $\mathcal{L}_p(\Omega)$. Another family of functional spaces that are often used in applications is $\mathcal{L}_p$ spaces, $p \ge 1$. They are complete metric spaces and test functions form a subset in them. It turns out that the space of test functions is a dense subset in any $\mathcal{L}_p$ space. This assertion comprises another approximation theorem.

THEOREM **14.3**. *For any open $\Omega \subset \mathbb{R}^N$, the space of test functions $\mathcal{D}(\Omega)$ is dense in $\mathcal{L}_p(\Omega)$, that is, for any $\varepsilon > 0$ and any $f \in \mathcal{L}_p(\Omega)$,* *there exists a test function* $\varphi \in \mathcal{D}(\Omega)$ *such that*

$$\|f-\varphi\|_p < \varepsilon\,.$$

Let Ω_R be the largest subset of Ω such that the distance between it and the boundary is equal to $\frac{1}{R}$, $d(\Omega_R, \partial\Omega) = \frac{1}{R}$ and, in addition, if Ω is not bounded, it is required that $|x| \leq R$ for all points of Ω_R. So, Ω_R is a proper subset of Ω. If $f \in \mathcal{L}_p(\Omega)$, then by continuity of the Lebesgue integral,

$$\lim_{R\to\infty}\int_{\Omega_R} |f(x)|^p\, d^N x = \int_{\Omega} |f(x)|^p\, d^N x\,.$$

Put $f_R = \chi_R f$, where χ_R is the characteristic function of Ω_R. For any $\varepsilon > 0$ and all large enough R, $\|f - f_R\|_p < \frac{\varepsilon}{2}$. If there is a test function $\varphi \in \mathcal{D}(\Omega)$ that is arbitrary close to f_R, e.g., $\|f_R - \varphi\|_p < \frac{\varepsilon}{2}$, the conclusion of the theorem follows from the triangle inequality

$$\|f-\varphi\|_p \leq \|f - f_R\|_p + \|f_R - \varphi\|_p < \varepsilon\,.$$

Let us show that φ can be obtained by a regularization of f_R. There is a technical fact to be established first.

PROPOSITION **14.1**. *Let* $f \in \mathcal{L}_p(\Omega)$, $p \geq 1$, *and* f *be extended to the whole* $\mathbb{R}^N$ *by setting it to zero outside* Ω. *If* f_a *is a regularization of* f, *then*

$$\|f_a\|_p \leq \|f\|_p\,.$$

A proof is based on Hölder's inequality (**12.2**). One has

$$\begin{aligned}
\|f_a\|_p^p &= \int_\Omega |f_a(x)|^p\, d^N x \overset{(1)}{\leq} \int_\Omega \Big(\int |f(y)|\phi_a(x-y)\, d^N y\Big)^p d^N x\\
&\overset{(2)}{\leq} \int_\Omega\int_\Omega |f(y)|^p \phi_a(x-y)\, d^N y \Big(\int \phi_a(x-z)\, d^N z\Big)^{p-1} d^N x\\
&\overset{(3)}{=} \int_\Omega\int_\Omega |f(y)|^p \phi_a(x-y)\, d^N y\, d^N x\\
&\overset{(4)}{=} \int_\Omega |f(y)|^p\, d^N y = \|f\|_p^p\,.
\end{aligned}$$

Here (1) is obtained by definition **14.6**, (2) holds by Hölder's inequality and the identity $\phi_a = \phi_a^{1/p}\phi_a^{1/q}$ where $\frac{1}{q} = \frac{p-1}{p}$, (3) follows from the normalization property of ϕ_a, and (4) is established changing the order of integration by Fubini's theorem and using the normalization property of ϕ_a again.

Proposition **14.2**. *Let f be from $\mathcal{L}_p(\Omega)$ and the support of f is a proper subset of Ω. Then a regularization f_a of f belongs to $\mathcal{D}(\Omega)$ for all small enough a and converges to f in $\mathcal{L}_p(\Omega)$ as $a \to 0^+$,*

$$\lim_{a\to 0^+} \|f - f_a\|_p = 0\,.$$

The support $K = \operatorname{supp} f$ is a bounded and closed subset in an open Ω so that the distance between K and the boundary $\partial\Omega$ is not zero, $d(K, \partial\Omega) = a_0 > 0$. Then $f_a \in \mathcal{D}(\Omega)$ if $a < a_0$. In Sec.**12.5.3** it is shown that $C_0^0(\Omega)$ is dense in $\mathcal{L}_p(\Omega)$. This means that for any $\varepsilon > 0$, there exists a continuous function g with bounded support in Ω such that

$$\|f - g\|_p < \varepsilon\,.$$

On the other hand, $\mathcal{D}(\Omega)$ is dense in $C_0^0(\Omega)$ and there exists $a_1 > 0$ such that

$$\|g - g_a\|_\infty < \varepsilon\,, \quad a < a_1$$

where g_a is a regularization of g. This implies that

$$\begin{aligned}\|g - g_a\|_p &= \Big(\int_{K_1} |g(x) - g_a(x)|^p\, d^N x\Big)^{\frac{1}{p}} \\ &\le \sup|g - g_a|\Big(\int_{K_1} d^N x\Big)^{\frac{1}{p}} = M\|g - g_a\|_\infty < M\varepsilon\end{aligned}$$

where M^p is the Lebesgue measure (volume) of any compact K_1 in Ω that contains supports of g and g_a for $a < a_1$.

Let $(f-g)_a$ be a regularization of $f-g$. It follows from the triangle inequality and Proposition **14.2** that

$$\begin{aligned}\|f - f_a\|_p &\le \|f - g\|_p + \|g - g_a\|_p + \|(f-g)_a\| \\ &\le 2\|f - g\|_p + \|g - g_a\|_p < (2 + M)\varepsilon\end{aligned}$$

This shows that the $\mathcal{L}_p$ distance between f and its regularization can be made arbitrary small for all sufficiently small a. This proves the assertion.

14.4. Conclusion. It has been shown that the space of test functions is sufficiently rich. Practically, all functions that are used to describe physical phenomena can be approximated with any desired accuracy (in some topology) by test functions.

14.5. Exercises.

1. Let φ and ψ be test functions. Is the product $\varphi(x)\psi(x)$ a test function? Express the support of the product in terms of supports of

φ and ψ.

2. Let η be a test function of a real variable that is equal to one in a neighborhood of $x = 0$. If $\varphi(x)$ is a test function, show that $\psi(x)$ defined by the equality

$$\varphi(x) = \varphi(0)\eta(x) + x\psi(x)$$

is a test function. Hint: Put

$$\psi(x) = \frac{\varphi(x) - \varphi(0)\eta(x)}{x}, \quad x \neq 0$$

By definition ψ has continuous derivatives of any order for all $x \neq 0$. Show that $\psi^{(n)}(x)$ can be extended continuously to $x = 0$ for all $n = 0, 1, 2, ...$ so that ψ is from class C^∞ (e.g., by using l'Hospital's rule). In particular, prove that

$$\psi^{(n)}(0) = \lim_{x\to 0} \psi^{(n)}(x) = \varphi^{(n+1)}(0), \quad n = 0, 1, 2, ...$$

Show that ψ has a bounded support.

3. Let η be a test function of a real variable that is equal to one in a neighborhood of $x = 0$. If $\varphi(x)$ is a test function and

$$p_{m-1}(x) = \varphi(0) + \varphi'(0)x + \cdots + \frac{x^{m-1}}{(m-1)!}\varphi^{(m-1)}(0)$$

is a Taylor polynomial of φ about $x = 0$, show that $\psi(x)$ defined by the equality

$$\varphi(x) = p_{m-1}(x)\eta(x) + x^m\psi(x)$$

is a test function. Hint: Prove that

$$\psi^{(n)}(0) = \lim_{x\to 0} \psi^{(n)}(x) = \frac{\varphi^{(n+m)}(0)}{m!}, \quad n = 0, 1, 2, ...$$

4. Let $\omega_a(x)$ denote the hat function of a real variable x. Is the product

$$|x^2 - a^2|^\nu \omega_a(x)$$

a test function for any real ν?

5. Put $\varphi_h(x) = \frac{1}{h}(\varphi(x+h) - \varphi)$ where $h \neq 0$ is real and $\varphi \in \mathcal{D}$. Show that $\varphi_h \to \varphi'$ in $\mathcal{D}$ for $h \to 0$.

6. Let ϕ_a be a test function used in the regularization integral (**14.6**). Show that $\lim_{a\to 0^+}(\phi_a, \varphi) = \varphi(0)$ for any test function $\varphi \in \mathcal{D}$. This means ϕ_a converges to the Dirac delta function as $a \to 0^+$, just like the hat function ω_a does (see (**14.3**)).

15. The space of distributions

15.1. The space $\mathcal{D}'$. A linear continuous functional

$$f:\ \mathcal{D}(\Omega) \to \mathbb{R}$$

is called a *distribution*. The value of f on a test function φ is denoted by (f, φ). The collection of all distributions is denoted by $\mathcal{D}'(\Omega)$. If $\Omega = \mathbb{R}^N$, then, for brevity, the set of all distributions is denoted by $\mathcal{D}'$. If $\Omega = (a, b) \subset \mathbb{R}$, then $\mathcal{D}'((a, b)) = \mathcal{D}'(a, b)$.

To show that a functional defined by some rule is a distribution one has to verify three things:

(i) Existence: the rule makes sense for all test functions.
(ii) Linearity: the rule defines a linear functional.
(iii) Continuity: the rule defines a continuous functional.

Two distributions f and g from $\mathcal{D}'(\Omega)$ are said to be equal if they have the same values for all test functions in $\mathcal{D}(\Omega)$:

$$f = g \text{ in } \mathcal{D}'(\Omega) \quad \Leftrightarrow \quad (f, \varphi) = (g, \varphi)\,, \quad \varphi \in \mathcal{D}(\Omega)$$

15.1.1. $\mathcal{D}'(\Omega)$ as a liner space. Let f and g be distributions. A linear combination $h = af + bg$, where a and b are real numbers, is defined by the rule

$$(h, \varphi) \stackrel{\text{def}}{=} a(f, \varphi) + b(g, \varphi)\,.$$

Let us show that a linear combination of distributions is a distribution and, hence, $\mathcal{D}'(\Omega)$ is a liner space.

Clearly, h is defined on any test function. Linearity follows from the linearity of f and g. For any test functions $\varphi_{1,2}$ and any reals $c_{1,2}$

$$\begin{aligned}(h, c_1\varphi_1 + c_2\varphi_2) &= a(f, c_1\varphi_1 + c_2\varphi_2) + b(g, c_1\varphi_1 + c_2\varphi_2)\\ &= a\Big(c_1(f, \varphi_1) + c_2(f, \varphi_2)\Big) + b\Big(c_1(g, \varphi_1) + c_2(g, \varphi_2)\Big)\\ &= c_1\Big(a(f, \varphi_1) + b(g, \varphi_1)\Big) + c_2\Big(a(f, \varphi_2) + b(g, \varphi_2)\Big)\\ &= c_1(h, \varphi_1) + c_2(h, \varphi_2)\,.\end{aligned}$$

The first equality is by definition of h, the second by linearity of f and g, the third is obtained by regrouping the term, and the final equality is again by definition of h.

Let $\varphi_n \to \varphi$ in $\mathcal{D}(\Omega)$. By the limit laws and continuity of f and g

$$\begin{aligned}\lim_{n\to\infty}(h, \varphi_n) &= \lim_{n\to\infty}\Big(a(f, \varphi_n) + b(g, \varphi_n)\Big)\\ &= a\lim_{n\to\infty}(f, \varphi_n) + b\lim_{n\to\infty}(g, \varphi_n)\\ &= a(f, \varphi) + b(g, \varphi) = (h, \varphi)\,.\end{aligned}$$

The first equality is by definition of h, the second by limit laws, the third by continuity of f and g, and the final equality is again by definition of h. The continuity is proved. Thus, *a space of distributions is a linear space.*

15.1.2. Remark. In general, linear functionals are not necessarily continuous. However, no explicit form of a linear non-continuous functional on the space of test functions has ever been constructed. It is only possible to prove the existence of such functionals by using the axiom of choice. So, practically all linear functionals on $\mathcal{D}(\Omega)$ defined explicitly are turned out to be continuous (or distributions).

15.2. Regular distributions. Let f be a locally integrable function. Then the rule

$$(f,\varphi) = \int f(x)\varphi(x)\,d^Nx\,, \quad \varphi \in \mathcal{D}\,, \tag{15.1}$$

defines a distribution from $\mathcal{D}'$. It is called a *regular distribution.*

The existence and linearity of this functional has already been established in Sec.**13.2**. Let us investigate its continuity. Let $\varphi_n \to 0$ in $\mathcal{D}$. One has to show that the numerical sequence (f,φ_n) converges to zero. The convergence in $\mathcal{D}$ implies that the functional sequence φ_n converges uniformly to the zero function and support of all elements of the sequence lies in one ball B_R. Therefore

$$\begin{aligned}
|(f,\varphi_n)| &= \left|\int f(x)\varphi_n(x)\,d^Nx\right| = \left|\int_{B_R} f(x)\varphi_n(x)\,d^Nx\right| \\
&\leq \int_{B_R} |f(x)|\,|\varphi_n(x)|\,d^Nx \leq \sup|\varphi_n| \int_{B_R} |f(x)|\,d^Nx \\
&= M \sup|\varphi_n| \to 0 \quad \text{as } n \to \infty
\end{aligned}$$

because by local integrability of f, $M < \infty$ and $\sup|\varphi_n| \to 0$ as $n \to \infty$ by uniform convergence of φ_n to zero. Thus, *with any locally integrable function one can associate a distribution by the rule* (**15.1**).

15.3. Isomorphism of locally integrable functions and regular distributions. Let f and g be two locally integrable functions. Each of them defines a distribution by the rule (**15.1**). Suppose that

$$(f,\varphi) = (g,\varphi)\,, \quad \varphi \in \mathcal{D}(\Omega)\,.$$

This means that the corresponding distributions are equal. Does this imply that *the functions f and g are equal in* Ω? In other words, do there exist more than one locally integrable function that correspond to the same distribution? Clearly, if the functions are equal almost everywhere, $f(x) = g(x)$ *a.e.*, then they define the same regular distribution. Is this condition also necessary? The answer is given by the Du Bois-Reymond lemma.

15.3.1. The du Bois-Reymond Lemma. *In order for a locally integrable f to vanish in a region Ω in the sense of distributions, it is necessary and sufficient that $f(x) = 0$ a.e. in Ω:*

$$(f, \varphi) = \int f(x)\varphi(x)\, d^N x = 0\,, \quad \varphi \in \mathcal{D}(\Omega) \quad \Leftrightarrow \quad f(x) = 0 \ a.e.$$

It follows from the du Bois-Reymond lemma that *every regular distribution is defined by a unique locally integrable function (modulo adding a function that is equal to zero almost everywhere).*

In what follows, no distinction will be made between regular distributions and locally integrable functions.

15.3.2. Proof of the du Bois-Reymond lemma. A proof is based on the following assertion from the Fourier analysis[2].

PROPOSITION **15.1**. *If the Fourier transform of a Lebesgue integrable function vanishes, then the function is zero almost everywhere:*

$$\int e^{i(k,x)} f(x)\, d^N x = 0 \quad \Rightarrow \quad f(x) = 0 \ \ a.e.$$

Take a point x_0 in an open set Ω. Then there exists an open ball $B_a(x_0)$ that lies in Ω together with its boundary, $\overline{B_a(x_0)} \subset \Omega$. By the hypothesis

$$(f, \varphi) = 0\,, \quad \varphi \in \mathcal{D}(B_a(x_0))$$

Fix $k \in \mathbb{R}^N$ and put

$$\phi_k(x) = e^{i(k,x)}\omega_a(x - x_0) \in \mathcal{D}(\Omega)$$

where ω_a is the hat function. Then

$$0 = (f, \phi_k) = \int f(x)\omega_a(x - x_0)e^{i(k,x)}\, d^N x$$

By Proposition **15.1**

$$f(x)\omega_a(x - x_0) = 0 \ a.e. \quad \Rightarrow \quad f(x) = 0 \ a.e.$$

as required.

[2] A.N. Kolmogorov and S.V. Fomin, Elements of the theory of functions and functional analysis, Chapter VIII, Sec. 3

15.4. Singular distributions. All distributions that are not regular are called *singular* distributions, that is, a singular distribution cannot be defined by an integral of a test function with some locally integrable function.

15.4.1. Dirac delta-function as a distribution. The Dirac delta function is defined by the rule

$$(\delta, \varphi) = \varphi(0)\,, \quad \varphi \in \mathcal{D}(\mathbb{R}^N)\,.$$

It is a linear continuous functional on $\mathcal{D}$. The linearity is obvious. Take a null sequence

$$\varphi_n \to 0 \quad \text{in } \mathcal{D}\,.$$

Then one has to check that the functional δ maps it to a numerical sequence that converges to zero. This is indeed true

$$\lim_{n\to\infty} (\delta, \varphi_n) = \lim_{n\to\infty} \varphi_n(0) = 0$$

because by topology in $\mathcal{D}$, the functional sequence $\varphi_n(x)$ converges to the zero function uniformly, which implies in particular that the sequence of values $\varphi_n(0)$ converges to zero:

$$|\varphi_n(0)| \le \sup |\varphi_n(x)| \to 0 \quad \text{as} \quad n \to \infty\,.$$

Thus, the Dirac delta-function is a linear continuous functional on $\mathcal{D}$ and, hence, it is a distribution.

Let us show that *the Dirac delta function is a singular distribution.* Suppose conversely that there exists a locally integrable function $\delta(x)$ such that

$$\int \delta(x)\varphi(x)\, d^N x = \varphi(0)\,, \quad \varphi \in \mathcal{D}\,.$$

By Theorem **14.1** there exists a test function $0 \le \eta_a(x) \le 1$ supported in a ball B_a for any choice of $a > 0$, and $\eta_a(x) = 1$ in a smaller ball $|x| < a/3$ (a neighborhood of $x = 0$). Then the product $\varphi\eta_a$ is also a test function and

$$\varphi(0) = \varphi(0)\eta_a(0) = (\delta, \varphi\eta_a) = \int \delta(x)\varphi(x)\eta_a(x)\, d^N x$$
$$= \int_{B_a} \delta(x)\varphi(x)\eta_a(x)\, d^N x\,.$$

It follows from this representation and $0 \le \eta_a(x) \le 1$ that

$$|\varphi(0)| \le \int_{B_a} |\delta(x)\varphi(x)\eta_a(x)|\, d^N x \le \sup|\varphi| \int_{B_a} |\delta(x)|\, d^N x\,.$$

Since the measure (volume) of the integration domain B_a can be made arbitrary small (by taking the radius a small enough), the integral can also be arbitrary small for any locally integrable function δ by Theorem **6.2**. But the value $|\varphi(0)|$ is not zero for all test functions, hence, a contradiction. Thus, *the Dirac delta-function is not a regular distribution and its action on a test function cannot be written in the integral form.*

15.4.2. Derivatives of the delta function. Let us show that all derivatives of the delta function, defined by the rule (**13.5**), are distributions. Since $\varphi \in C^\infty$, the rule makes sense for any test function. Linearity follows from the linearity of differentiation on the space of test functions. Finally, let $\varphi_m \to 0$ in $\mathcal{D}$ as $m \to \infty$. Then the functional $\delta^{(n)}$ is continuous because

$$|(\delta^{(n)}, \varphi_m)| = |\varphi_m^{(n)}(0)| \leq \sup|\varphi_m^{(n)}| \to 0 \quad \text{as } m \to \infty$$

by definition of that $\varphi_n \to 0$ in $\mathcal{D}$.

15.4.3. Shifted delta functions. Consider a functional on the space of test function defined by the rule

$$\Big(\delta(x - x_0), \varphi(x)\Big) = \varphi(x_0)\,,$$

for any $x_0 \in \mathbb{R}^N$. The rule makes sense for any test function φ. The functional is linear. If $\varphi_m \to 0$ in $\mathcal{D}$, then the numerical sequence $\varphi_m(x_0)$ converges to 0 for any point x_0 because $|\varphi_m(x_0)| \leq \sup|\varphi_m| \to 0$ as $m \to \infty$. Thus, the functional is a distribution. It is called a *shifted delta function*. It is singular distribution. A proof of this assertion is similar to the proof of singularity of the Dirac delta function and left to the reader as an exercise.

15.4.4. The principal value distribution. Define a functional $\mathcal{P}\frac{1}{x}$ on the space of test function of a real variable x by the rule

$$(\mathbf{15.2}) \qquad \Big(\mathcal{P}\frac{1}{x}, \varphi\Big) \stackrel{\text{def}}{=} p.v. \int \frac{\varphi(x)}{x}\,dx \stackrel{\text{def}}{=} \lim_{a\to 0^+} \int\limits_{|x|>a} \frac{\varphi(x)}{x}\,dx\,,$$

where *p.v.* indicates that the improper integral is understood as the *Cauchy principal value* (defined by the subsequent equality). Note that the function $1/x$ is not locally integrable (its integral does not exist over any bounded interval that contains $x = 0$). Let us show that $\mathcal{P}\frac{1}{x}$ defines a (singular) distribution. It is called the *the principal value distribution.*

Existence. Since the support of φ is bounded, $\operatorname{supp}\varphi \subset [-R, R]$ and

$$p.v. \int\limits_{|x|<R} \frac{dx}{x} = \lim_{a\to 0^+} \left(\int_{-R}^{-a} + \int_{a}^{R} \right) \frac{dx}{x} = 0\,,$$

one infers that

$$\begin{aligned} \left(\mathcal{P}\frac{1}{x}, \varphi\right) &= \lim_{a\to 0^+} \int\limits_{|x|>a} \frac{\varphi(x)}{x}\,dx = \lim_{a\to 0^+} \int\limits_{a<|x|<R} \frac{\varphi(x)}{x}\,dx \\ &= \lim_{a\to 0^+} \int\limits_{a<|x|<R} \frac{\varphi(x)-\varphi(0)}{x}\,dx = \int\limits_{|x|<R} \frac{\varphi(x)-\varphi(0)}{x}\,dx\,. \end{aligned}$$

The latter integral exists because the integrand can be continuously extended to $x = 0$:

$$\lim_{x\to 0} \frac{\varphi(x)-\varphi(0)}{x} = \varphi'(0)$$

so that it is continuous on the interval $[-R, R]$ and, hence, integrable on it.

Linearity. One has

$$\begin{aligned} \left(\mathcal{P}\frac{1}{x}, c_1\varphi_1 + c_2\varphi_2\right) &= \lim_{a\to 0} \int\limits_{a<|x|} \frac{c_1\varphi_1(x) + c_2\varphi_2(x)}{x}\,dx \\ &\overset{(1)}{=} c_1 \lim_{a\to 0} \int\limits_{a<|x|} \frac{\varphi_1(x)}{x}\,dx + c_2 \lim_{a\to 0} \int\limits_{a<|x|} \frac{\varphi_2(x)}{x}\,dx \\ &= c_1\left(\mathcal{P}\frac{1}{x}, \varphi_1\right) + c_2\left(\mathcal{P}\frac{1}{x}, \varphi_2\right); \end{aligned}$$

here (1) follows from linearity of the Lebesgue integral and the limit laws.

Continuity. Let $\varphi_n \to 0$ in $\mathcal{D}$. This implies that the sequence of derivatives $\varphi_n'(x)$ converges uniformly to the zero function:

$$\lim_{n\to\infty} \sup |\varphi_n'(x)| = 0\,.$$

By the mean value theorem there exists a point x^* between 0 and x such that

$$\left| \frac{\varphi_n(x) - \varphi_n(0)}{x} \right| = |\varphi_n'(x^*)| \le \sup |\varphi_n'|\,.$$

Then the following chain of inequalities holds:

$$\begin{aligned}\left|\left(\mathcal{P}\frac{1}{x},\varphi_n\right)\right| &= \left|p.v.\int \frac{\varphi_n(x)}{x}\,dx\right| \overset{(1)}{=} \left|p.v.\int_{|x|<R} \frac{\varphi_n(x)}{x}\,dx\right| \\ &\overset{(2)}{=} \left|p.v.\int_{|x|<R} \frac{\varphi_n(x)-\varphi_n(0)}{x}\,dx\right| \\ &\le p.v.\int_{|x|<R} \left|\frac{\varphi_n(x)-\varphi_n(0)}{x}\right|\,dx \\ &\overset{(3)}{\le} \sup|\varphi_n'| \int_{|x|<R} dx = 2R\sup|\varphi_n'| \to 0\end{aligned}$$

as $n \to \infty$. Here (1) holds because supports of all terms φ_n are in an interval $[-R,R]$ for some R; (2) is true because the Cauchy principal value integral of $\frac{1}{x}$ over the interval $(-R,R)$ vanishes; (3) follows from the inequality derived from the mean value theorem.

15.4.5. General principal value distributions. For any integer $n \ge 1$, the function x^{-n} is not locally integrable. But it is possible to associate a distribution with it by the rule similar to (**15.2**):

$$\left(\mathcal{P}\frac{1}{x^n},\varphi\right) \overset{\text{def}}{=} p.v.\int \frac{\varphi(x)-T_{n-2}(x)}{x^n}\,dx\,, \quad n = 2,3,\dots\,, \tag{15.3}$$

where

$$T_{n-2}(x) = \sum_{k=0}^{n-2} \frac{\varphi^{(k)}(0)}{k!}\,x^k$$

is the Taylor polynomial of φ about $x=0$ of order $n-2$ (see Sec.**1.4.1**).

The existence of this functional follows from the relation

$$\begin{aligned}\left(\mathcal{P}\frac{1}{x^n},\varphi\right) &= \lim_{a\to 0^+}\int_{|x|<a} \frac{\varphi(x)-T_{n-2}(x)}{x^n}\,dx \\ &= \lim_{a\to 0^+}\int\limits_{a<|x|<R} \frac{\varphi(x)-T_{n-2}(x)}{x^n}\,dx + \int\limits_{|x|>R} \frac{T_{n-2}(x)}{x^n}\,dx\end{aligned} \tag{15.4}$$

that holds for any test function φ with support in the interval $[-R,R]$. The last integral exists because $T_{n-2}/x^n \sim 1/x^2$ as $|x| \to \infty$. In the first integral the Taylor polynomial T_{n-2} can be replaced by T_{n-1} because the integral of $1/x$ over the symmetric interval $a<|x|<R$ vanishes. Therefore the first integral converges in the limit $a \to 0^+$ because the integrand has a continuous extension to $x=0$ by Taylor's theorem:

$$\lim_{x\to 0} \frac{\varphi(x)-T_{n-1}(x)}{x^n} = \frac{\varphi^{(n)}(0)}{n!}\,. \tag{15.5}$$

Linearity of the functional follows from linearity of the integral and that a Taylor polynomial of a linear combination of two functions φ_1 and φ_2 is the corresponding linear combination of the Taylor polynomials for φ_1 and φ_2.

Let us show continuity. Let $\varphi_m \to 0$ in $\mathcal{D}$ as $m \to \infty$. Supports of all φ_m lie in an interval $[-R, R]$. Replacing φ by φ_m in (**15.4**), one can see that the integral over the interval $|x| > R$ vanishes in the limit $m \to \infty$

$$\int_{|x|>R} \frac{T_{n-2}(x)}{x^n}\, dx = \sum_{k=0}^{n-2} \frac{\varphi_m^{(k)}(0)}{k!} \int_{|x|>R} \frac{x^k}{x^n}\, dx$$

because $|\varphi_m^{(k)}(0)| \leq \sup|\varphi_m^{(k)}| \to 0$. As already noted, p_{n-2} can be replaced by T_{n-1} in the first integral. By Taylor's theorem, there exists a point x^* between x and 0 such that

$$\frac{\varphi_m(x) - T_{n-1}(x)}{x^n} = \frac{\varphi_m^{(n)}(x^*)}{n!}$$

so that

$$\Big| \int_{|x|<R} \frac{\varphi(x) - T_{n-1}(x)}{x^n}\, dx \Big| \leq \frac{\sup|\varphi_m^{(n)}|}{n!}\, 2R \to 0$$

as $m \to \infty$.

15.5. Spherical delta-function. Consider a distribution of N real variables defined by the surface integral of a test function over the sphere $|x| = a$

$$(\delta_{S_a}, \varphi) = \int_{|x|=a} \varphi(x)\, dS\,.$$

The integral is reduced to an iterated integral using parametric equation of the sphere as shown in Sec. **8.3.1**. For example, for $N = 2$, the integral is evaluated in polar coordinates. If θ is the polar angle, then $dS = a d\theta$ is the arclength on a circle of radius a so that

$$\int_{|x|=a} \varphi(x)\, dS = a \int_0^{2\pi} \varphi\Big(a\cos(\theta), a\sin(\theta)\Big)\, d\theta\,.$$

In three-dimensional space, the surface integral is evaluated in spherical coordinates

$$\int_{|x|=a} \varphi(x)\, dS = a^2 \int_0^{2\pi} \int_0^{\pi} \varphi(x(\phi,\theta))\, \sin(\phi)\, d\phi\, d\theta\,,$$

where ϕ and θ are, respectively, the zenith and polar angles in the spherical coordinates.

The existence and linearity follows from the existence and linearity of the surface integral of a smooth function. To establish continuity let $\varphi_n \to 0$ in $\mathcal{D}$. Then φ_n converges to the zero function uniformly $\lim_{n\to\infty} \sup|\varphi_n| = 0$. Hence,

$$|(\delta_{S_a}, \varphi_n)| \le \int_{|x|=a} |\varphi_n(x)|\, dS \le \sup|\varphi_n| \int_{|x|=a} dS$$
$$= a^{N-1}\sigma_N \sup|\varphi_n| \to 0$$

as $n \to \infty$. Here σ_N is the surface are of the unit sphere $|x| = 1$ in $\mathbb{R}^N$. So, $\delta_{S_a}(x)$ is a distribution of N real variables.

15.5.1. The distribution δ_{S_a} is singular. By Theorem **14.1**, there exists a bump function η_ε for the sphere $|x| = a$ which is a test function with support in a spherical layer $a - \varepsilon \le |x| \le a + \varepsilon$ (denoted by $B_{a\pm\varepsilon}$ for brevity), and η_ε has unit value in a neighborhood of the sphere. Suppose there exists a locally integrable function $\delta_{S_a}(x)$ such that

$$(\delta_{S_a}, \varphi) = \int \delta_{S_a}(x)\varphi(x)\, d^N x$$

for any test function φ. Then $\eta_\varepsilon\varphi$ is also a test function and therefore

$$\int_{|x|=a} \varphi(x)\, dS = \int_{|x|=a} \eta_\varepsilon(x)\varphi(x)\, dS = \Big(\delta_{S_a}, \eta_\varepsilon\varphi\Big)$$
$$= \int \delta_{S_a}(x)\eta_\varepsilon(x)\varphi(x)\, d^N x = \int\limits_{B_{a\pm\varepsilon}} \delta_{S_a}(x)\eta_\varepsilon(x)\varphi(x)\, d^N x$$

This relation implies that

$$\left|\int_{|x|=a} \varphi(x)\, dS\right| \le \sup|\varphi| \int\limits_{B_{a\pm\varepsilon}} |\delta_{S_a}(x)|\, d^N x\,.$$

because $0 \le \eta_\varepsilon(x) \le 1$. The measure (volume) $\mu(B_{a\pm\varepsilon}) = O(\varepsilon)$ tends to 0 as $\varepsilon \to 0^+$ and, hence, the right-hand side of this inequality can be made arbitrary small for any locally integrable function $\delta_{S_a}(x)$ by Theorem **6.2**, but the left-hand side is finite. This contradiction implies that no such locally integrable function $\delta_{S_a}(x)$ exists, and δ_{S_a} is a singular distribution.

15.5.2. Physical significance of δ_{S_a}. In $\mathbb{R}^2$, the distribution $\delta_{S_a}(x)$ can describe a mass density of a thin circular uniform wire occupying the circle $|x| = a$. If m is the total mass of the wire, then its mass density is

$$\rho(x) = \frac{m}{2\pi a}\, \delta_{S_a}(x)\,, \quad x \in \mathbb{R}^2\,.$$

In $\mathbb{R}^3$, this distribution can be used to describe the mass density of a thin uniform spherical shell, $|x| = a$. If the total mass is m, then the mass density is

$$\rho(x) = \frac{m}{4\pi a^2}\,\delta_{S_a}(x)\,, \quad x \in \mathbb{R}^3\,.$$

Similarly, in $\mathbb{R}^N$, the mass density of a thin spherical shell can be modeled as the distribution

$$\rho(x) = \frac{m}{a^{N-1}\sigma_N}\,\delta_{S_a}(x)\,, \quad x \in \mathbb{R}^N\,.$$

15.6. Delta-functions on smooth M-surfaces in $\mathbb{R}^N$. The concept of a spherical delta-function can be extended to general smooth surfaces of dimension M in $\mathbb{R}^N$. It is assumed that a smooth M-surface S has a finite area in any ball $|x| < R$. Let $\nu(x)$ be a function that is continuous on S. Define a functional by the rule

$$(\nu\delta_S, \varphi) = \int_S \nu(x)\varphi(x)\,dS\,, \quad \varphi \in \mathcal{D}\,. \tag{15.6}$$

The surface integral is calculated by means of a parameterization of S (see Sec.**8.3**). Since the support of φ lies in a ball B_R, the integration is reduced to $S_R = S \cap B_R$, the part of S that lies in B_R. Owing to continuity of ν on S, $|\nu(x)\varphi(x)| \le M$ on S_R, and therefore, the integral exists because S_R has a finite area by assumption. Linearity and continuity $\nu\delta_S$ are verified in a similar fashion as in the case of the spherical delta-function. The distribution $\nu\delta_S$ is singular, which is again established by the same line of arguments as for the spherical delta-function. The technical details are left to the reader as an exercise.

The distribution $\nu\delta_S$ can describe a density of some quantity distributed over an M-surface in $\mathbb{R}^N$ with a *surface density* $\nu(x)$ (amount per unit surface area at a point x of the surface). For example, a dielectric wire in space can have a non-uniformly distributed electric charge. In this case, $\nu(x)$ is an electric charge per unit length of a smooth curve (modeling the wire) at a point x of the curve.

15.6.1. Mass density of moving objects. Consider a collection of M particles moving in space along trajectories $x = x_p(t)$, $p = 1, 2, ..., M$. The mass density of the system can be viewed as a distribution in space and time variables, $\rho(x, t)$. It acts on a test function by the rule

$$(\rho, \varphi) = \sum_{p=1}^{M} m_p \int_{-\infty}^{\infty} \varphi(x_p(t), t))\,dt\,,$$

where m_a are masses of the particles. Thanks to the boundedness of support of φ, the integration interval can be reduced to a bounded interval so that the integrals always exist. For a *fixed* t, the rule resembles the definition of a shifted delta-function. For this reason this distribution is often *formally* written as

$$\rho(x,t) = \sum_{p=1}^{M} m_p \delta(x - x_p(t))\,.$$

Consider a one-dimensional object of a finite length L moving in a space, like a string. Then it sweeps a two dimensional surface in spacetime, $x = u(\xi, t)$. This function defines the shape of the object at each moment of time t and $0 \leq \xi \leq L$ is the natural parameter along the string. If $\nu(\xi)$ is the linear mass density of the object ($\nu(\xi)\,d\xi$ is the mass of a portion of the string of length $d\xi$ at a point ξ), then the mass density $\rho(x,t)$ is the 2-surface delta-function, defined by the rule

$$(\rho, \varphi) = \int_0^L \int_{-\infty}^{\infty} \nu(\xi)\varphi(u(\xi,t),t)\,dt\,d\xi\,.$$

A verification of that these mass density functionals are distributions is left to the reader as an exercise.

15.7. Complex-valued distributions. A complex-valued function is a test function from $\mathcal{D}(\Omega)$ if its real and imaginary parts are from $\mathcal{D}(\Omega)$. A linear continuous functional f can take complex values on the space of test functions, that is, $(f, \varphi) \in \mathbb{C}$. In this case, f is said to be a complex-valued distribution. For example, a locally integrable complex-valued function $f(x) = e^{itx}$ of a real variable x defines a complex-valued distribution

$$(f, \varphi) = \int e^{itx}\varphi(x)\,dx$$

for any real or complex parameter t.

A complex-conjugated distribution $\bar{f}$ is defined by the rule

$$(\bar{f}, \varphi) = \overline{(f, \bar{\varphi})}$$

for any complex-valued distribution f. The linear combinations

$$\mathrm{Re}\, f = \frac{1}{2}\Big(f + \bar{f}\Big)\,, \quad \mathrm{Im}\, f = \frac{1}{2i}\Big(f - \bar{f}\Big)$$

are called the real and imaginary parts of the distribution f, respectively. A distribution is said to be real if its imaginary part is the zero distribution.

For example, the delta-function is a real distribution:

$$(\bar{\delta}, \varphi) = \overline{(\delta, \bar{\varphi})} = \overline{\bar{\varphi}(0)} = \varphi(0) = (\delta, \varphi)\,.$$

15.8. Topology in the space of distributions. In the process of modeling a mass density of a point particle, or more generally, a density of some quantity distributed over a set of measure (volume) zero, a limiting process was designed in which a sequence of smooth functions converges to a distribution. Since every smooth function can be viewed as a distribution, this limiting process can be defined as *convergence in the sense of distributions or weak convergence*. It will be shown later that any distribution can always be viewed as a weak limit of a sequence of smooth functions.

A sequence of distributions $\{f_n\} \subset \mathcal{D}'(\Omega)$ *is said to converge to a distribution* f *if for any test function the sequence of values of* f_n *converges to the value of* f*:*

$$\lim_{n\to\infty} (f_n, \varphi) = (f, \varphi)\,, \quad \varphi \in \mathcal{D}(\Omega)\,.$$

In this case, one writes

$$f_n \to f \quad \text{in } \mathcal{D}'(\Omega)\,.$$

This is somewhat similar to pointwise convergence of a sequence of ordinary functions, with just one difference that "points" at which a sequence of values is computed are now test functions.

Convergence of series of distributions is defined via convergence of a sequence of partial sums. The series $\sum_n f_n$ converges in $\mathcal{D}'$ if the limit of $\sum_{|n|<k}(f_n, \varphi)$ as $k \to \infty$ exists for any test function, and the equality

$$\sum_n f_n(x) = f(x) \in \mathcal{D}'$$

means that

$$\lim_{k\to\infty} \sum_{|n|<k} (f_n, \varphi) = (f, \varphi)\,, \quad \varphi \in \mathcal{D}\,.$$

15.8.1. Example. Let us find the limit of smooth integrable functions

$$f_a(x) = \frac{a}{x^2 + a^2}\,, \quad x \in \mathbb{R}\,,$$

as $a \to 0$ in the distributional sense. It is not difficult to see that

$$\int_{-\infty}^{\infty} f_a(x)\,dx = \int_{-\infty}^{\infty} \frac{dy}{y^2 + 1} = \pi\,.$$

The objective is to calculate the limit

$$\lim_{a\to 0}(f_a,\varphi) = \lim_{a\to 0}\int_{-\infty}^{\infty}\frac{a\varphi(x)}{a^2+x^2}\,dx = \lim_{a\to 0}\int_{-\infty}^{\infty}\frac{\varphi(ay)}{y^2+1}\,dy\,.$$

The Lebesgue dominated convergence theorem is a main technical tool to calculate distributional limits of sequences of classical functions.

In this case, note first that

$$\lim_{a\to 0}\frac{\varphi(ay)}{y^2+1} = \frac{\varphi(0)}{y^2+1}$$

for any y. To interchange the order of integration and taking the limit, one has to find an integrable bound for the integrand that is independent of the parameter a:

$$\frac{|\varphi(ay)|}{y^2+1} \le \frac{M}{y^2+1} \ \in\ \mathcal{L}\,,\quad M = \sup|\varphi|\,.$$

Therefore by the Lebesgue dominated convergence theorem,

$$\begin{aligned}\lim_{a\to 0}(f_a,\varphi) &= \int_{-\infty}^{\infty}\lim_{a\to 0}\frac{\varphi(ay)}{y^2+1}\,dy = \varphi(0)\int_{-\infty}^{\infty}\frac{dy}{y^2+1} = \pi\varphi(0)\\ &= \pi(\delta,\varphi)\,.\end{aligned}$$

This means that, in the sense of distributions,

$$f_a(x) \to \pi\delta(x) \quad \text{in } \mathcal{D}'$$

as $a \to 0$.

15.8.2. Sequences of regular distributions. *Let f_n be a sequence of locally integrable functions that converges to a function f uniformly on any compact set K:*

$$\lim_{n\to\infty}\sup_K|f(x)-f_n(x)| = 0\,.$$

It follows from this limit that there exists an integer m such that $\sup_K|f(x)-f_m(x)| \le 1$ so that for all $x \in K$

$$|f(x)| \le |f(x)-f_m(x)| + |f_m(x)| \le 1 + |f_m(x)| \in \mathcal{L}(K)\,,$$

because $f_m \in \mathcal{L}_{\rm loc}$. This implies that f is integrable on any compact, or f is locally integrable and, hence, defines a regular distribution.

Let us show that $f_n \to f$ *in $\mathcal{D}'$, that is, classical and distributional limits coincide in this case.* For any test function φ with support K

$$|(f-f_n,\varphi)| = \Big|\int_K (f-f_n)\varphi\, d^Nx\Big| \le \sup_K|f-f_n|\int_K|\varphi|\,d^Nx \to 0$$

as $n \to \infty$.

15.8.3. Distributional limit of rapidly oscillating functions. The function $f_t(x) = e^{itx}$ is locally integrable and defines a regular distribution for any t:

$$(f_t, \varphi) = \int_{-\infty}^{\infty} e^{itx}\varphi(x)\, dx\,.$$

Its action on a test function gives the Fourier transform of the test function. The function $f_t(x)$ rapidly oscillates with increasing t and has no pointwise limit anywhere expect $x = 0$ as $t \to \infty$. However, the limit exists in the distributional sense. Indeed, using the integration by parts and boundedness of the support of φ,

$$(f_t, \varphi) = \frac{1}{it}\int_{-R}^{R} \varphi(x)\, d\, e^{itx} = -\frac{1}{it}\int_{-R}^{R} e^{itx}\varphi'(x)\, dx\,,$$

where the boundary term vanishes as $\varphi(\pm R) = 0$. Since the absolute value of the derivative $|\varphi'(x)|$ is integrable, it follows that

$$|(f_t, \varphi)| \le \frac{1}{t}\int_{-R}^{R} |\varphi'(x)|\, dx = \frac{M}{t} \to 0$$

as $t \to \infty$. Thus, in the distributional sense

$$\lim_{t\to\infty} e^{itx} = 0 \quad \text{in } \mathcal{D}'\,.$$

This conclusion looks rather odd. Does this limit make any sense in a physical reality? In general, the answer is affirmative. Think of a dielectric rod in which electric charge is distributed by periodically arranged layers carrying opposite charges so that a negatively charged layer is followed by a positively charged layer an so on. All layers have the same total charge (either positive or negative) that is distributed uniformly so that the charge density is periodic along the rod (it resembles a cosine function).

Next, imagine that the thickness of each layer is getting smaller. For example, each layer is cut in half and some of the neighboring layers with opposite charges are swapped so that a positively charged layer is followed by a negatively charged one. The process can be repeated. With every step of the process, the period of oscillations is reduced by factor of 2, while the amplitude remains the same. The charge density begins to rapidly oscillate along the rod and does not converge to any function. However, from the practical point of view, the density at a point is defined by a measured charge of a portion of the rod that has unit length and contains the point. When the period of oscillations becomes much less than the smallest length that can be measured, the rod would appear electrically neutral, that is, the measured electrical charge density is zero. Indeed, the total charge of any interval being exactly an integer multiple of the period is zero. Therefore the total charge over any interval oscillates between zero and the total charge of a single layer. This implies that the charge density at any point is decreasing to zero with increasing the number of oscillations per the minimal length. The argument can be made rigorous if the limit charge density is understood in the sense of distributions (see Problem 8 in Exercises).

So, the distributional interpretation of the charge density is more adequate for mathematical modeling than the picture based on classical functions and their limits because the former takes into account a general concept inherent to our perception and understanding of the physical reality that all quantities distributed in space and time cannot be measured at a point of space or at an exact moment of time but rather their measurements include some averaging procedure of small regions of space or intervals of time.

15.9. Completeness of the space of distributions. If $f_n \to f$ in $\mathcal{D}'$, then the sequence $\{f_n\}$ is a Cauchy sequence in the distributional sense, meaning that the numerical sequence (f_n, φ) is a Cauchy sequence for any test function. Suppose $\{f_n\}$ is a Cauchy sequence in the distributional sense. Then by the Cauchy criterion for numerical sequences every sequence (f_n, φ) has a limit and, hence, this limit defines a functional on $\mathcal{D}$. Is this functional a distribution? Or, in other words, is the space $\mathcal{D}'$ complete? The answer is affirmative.

THEOREM **15.1**. *Let $\{f_n\}$ be a sequence of distributions such that the numerical sequence (f_n, φ) converges for any test function φ. Then the functional f defined by*

$$(f, \varphi) = \lim_{n\to\infty} (f_n, \varphi)\,, \quad \varphi \in \mathcal{D}$$

is linear and continuous, that is, f is a distribution.

A proof requires to verify linearity and continuity of f. Linearity follows from the limit laws. A verification of continuity is a bit technical and omitted here[3].

The completeness theorem also implies that the enlargement of the set of classical functions (or regular distributions) by adding all limits of weakly convergent sequences of classical functions cannot give anything larger than $\mathcal{D}'$. The completeness property of the set of distributions is

[3]A proof can be found, e.g., in: G. Grubb, Distributions and operators, Theorem 3.9

shown to drastically simplify differential calculus for classical functions and their sequences and series if they are treated as distributions. In particular, any functional series or sequence converging in a distributional sense can be differentiated term-by-term infinitely many times to get the corresponding derivatives of the limit distribution!

15.9.1. Example. Let $f_n(x) = \frac{3}{2}n^3 x$ if $|x| \le \frac{1}{n}$ and $f_n(x) = 0$ otherwise. For every n, f_n defines a regular distribution

$$(f_n, \varphi) = \frac{3n^3}{2} \int_{-\frac{1}{n}}^{\frac{1}{n}} x\varphi(x)\, dx\,, \quad \varphi \in \mathcal{D}\,.$$

Let us investigate the convergence of this numerical sequence. Making the substitution $y = nx$ and integrating by parts, one infers that

$$\begin{aligned}(f_n, \varphi) &= \frac{3n}{2} \int_{-1}^{1} y\varphi(y/n)\, dy \\ &= \frac{3}{4} n\Big(\varphi(1/n) - \varphi(-1/n)\Big) - \frac{3}{4}\int_{-1}^{1} y^2 \varphi'(y/n)\, dy\,.\end{aligned}$$

The limit of the first term is not difficult to compute:

$$n\Big(\varphi(1/n) - \varphi(-1/n)\Big) = n\left(2\varphi'(0)\frac{1}{n} + O\Big(\frac{1}{n^2}\Big)\right) = 2\varphi'(0) + O\Big(\frac{1}{n}\Big)\,.$$

The limit of the second term can be found by means of the Lebesgue dominated convergence theorem. Let $g_n(y) = y^2\varphi'(y/n)$. Then $g_n(y) \to y^2\varphi'(0)$ as $n \to \infty$ for any y. To justify interchanging the order of taking the limit and the integral by means of the Lebesgue dominated convergence theorem, one has to find an integrable bound for $g_n(y)$ that is independent of n. Since $|\varphi'(x)| \le M$ for all x (as any continuous function with a bounded support), $|g_n(y)| \le My^2 \in \mathcal{L}(-1, 1)$. Therefore

$$\begin{aligned}\lim_{n\to\infty} \int_{-1}^{1} y^2\varphi'(y/n)\, dy &= \int_{-1}^{1} \lim_{n\to\infty} y^2 \varphi'(y/n)\, dy \\ &= \varphi'(0)\int_{-1}^{1} y^2 dy = \frac{2}{3}\,\varphi'(0)\,,\end{aligned}$$

and

$$\lim_{n\to\infty} (f_n, \varphi) = \frac{3}{2}\varphi'(0) - \frac{1}{2}\,\varphi'(0) = \varphi'(0)\,.$$

By the completeness theorem, the functional

$$(f, \varphi) = \varphi'(0)\,, \quad \varphi \in \mathcal{D}\,,$$

is a distribution, that is, it is a linear continuous functional on $\mathcal{D}$. In Sec.**15.4.2** it was also shown that $f(x) = -\delta'(x)$.

15.10. Exercises.

1. Show that the rule $(f, \varphi) = \varphi^{(n)}(x_0)$ where $\varphi \in \mathcal{D}$ defines a distribution.

2. Show that $\delta_a \to \delta$ in $\mathcal{D}'$ as $a \to 0^+$ for each of the following families of smooth functions:

$$\text{(i)} \qquad \delta_a(x) = \frac{1}{\sqrt{2\pi a}} e^{-\frac{x^2}{2a}}$$
$$\text{(ii)} \qquad \delta_a(x) = \frac{1}{\pi x} \sin\left(\frac{x}{a}\right)$$
$$\text{(iii)} \qquad \delta_a(x) = \frac{a}{\pi x^2} \sin^2\left(\frac{x}{a}\right)$$

3. Let x be a real variable. Consider the sequence of functions $f_n(x) = n - n^2|x|$ if $|x| < 1/n$ and $f_n(x) = 0$ if $|x| > 1/n$. Find the limit of the sequence in the sense of distributions using only the definition of such a limit.

4. Let $f_n(x) = n^2\varepsilon(x)$ if $|x| < \frac{1}{n}$ and $f_n(x) = 0$ otherwise, where $\varepsilon(x)$ is the sign function; it is equal to 1 if $x > 0$ and to -1 if $x < 0$. Show that the sequence $\{f_n\}$ converges in the sense of distributions and find the limit distribution.

5. (i) Find a sequence of locally integrable function $f_n(x)$ in $\mathbb{R}^3$ that converges to the spherical delta-function:

$$f_n \to \delta_{S_a} \text{ in } \mathcal{D}', \quad (\delta_{S_a}, \varphi) = \oint_{|x|=a} \varphi(x)\, dS$$

(ii) Find a sequence of test functions $\varphi_n \in \mathcal{D}(\mathbb{R}^3)$ that converges to the spherical delta function in the distributional sense.
Hint: Use a suitable regularization of a sequence from Part (i).

6. Show that the functional defined in Sec.**15.6** is a singular distribution, that is, show that it is a linear continuous functional on the space of test function $\mathcal{D}$, and there exists no locally integrable function such that the value of this functional on a test function is given by the integral of the product of the locally integrable function and the test function.

7. Let n be a positive integer and $\theta(x)$ is the step function. Find the following limits in the distributional sense or show that the limit does not exist:

$$\text{(i)} \quad \lim_{t\to\infty} t^n e^{itx}\,,$$
$$\text{(ii)} \quad \lim_{t\to\infty} x^n e^{itx}\,,$$
$$\text{(iii)} \quad \lim_{t\to\infty} \sin^n(tx)\,,$$
$$\text{(iv)} \quad \lim_{t\to\infty} e^{itx}\theta(x)\,,$$
$$\text{(v)} \quad \lim_{t\to\infty} t^n e^{itx}\theta(x)\,.$$

that is, if the limit exists, then give an explicit rule how to compute the value of the limit distribution for a test function.

8. Let f be a periodic continuous function such that

$$f(x+T) = f(x)\,, \quad \int_0^T f(x)\,dx = 0\,.$$

Put $f_n(x) = f(nx)$, $n = 1, 2, \ldots$. Show that $f_n \to 0$ in $\mathcal{D}'$. Does the conclusion hold if continuity of f is replaced by local integrability? *Hint*: Show first that for any interval $[a, b]$

$$\Big|\int_a^b f(x)\,dx\Big| \le \int_0^T |f(x)|\,dx = M$$

Find the function F_n such that $(f_n, \varphi) = -(F_n, \varphi')$ and show that $\sup|F_n| \le M/n$. Proceed.

9. (i) Let $\{a_n\}$ be any sequence or real numbers, and $\{x_n\}$ be a sequence that has no limit points. Show the series

$$\sum_n a_n \delta(x - x_n)$$

converges in $\mathcal{D}'$.
(ii) In part (i), assume that $x_n \to x_0$. Does the series converge in the sense of distributions? If not, construct an explicit example of the sequence $\{a_n\}$ for which the series does not converge.

16. Singular functions as distributions

There are functions that are not locally integrable. Can such function be "turned" into distributions? For example, the function $f(x) = \frac{1}{x}$ is not locally integrable because of non-integrable singularity at $x = 0$. Although the product $f(x)\varphi(x)$ is not integrable, it is possible to *regularize* the integral by means of the Cauchy principal value. This turns the singular function $f(x)$ into a singular distribution $\mathcal{P}\frac{1}{x}$. The function $f(x)$ is from class C^∞ outside any neighborhood of $x = 0$. The values of $\mathcal{P}\frac{1}{x}$ at a test function φ whose support lies in $|x| > 0$ is the same as the value of the integral of $\varphi(x)/x$. So, $\mathcal{P}\frac{1}{x}$ and $\frac{1}{x}$ only differ near $x = 0$. In this sense, $\mathcal{P}\frac{1}{x}$ is said to be a *distributional regularization* of a singular function $\frac{1}{x}$. Let us try to extend this idea to other singular functions. To do so, it is necessary to make the above concept of distributions "equal near a point" precise.

16.1. Distributions equal in an open set. A singular distribution cannot be associated with a locally integrable function and, hence, can have no value at some points. In contrast to classical functions, distributions cannot be compared pointwise. However, they can be compared in open sets (or in a neighborhood of any point) by comparing their values on test functions supported in such sets.

16.1.1. Distribution vanishing in an open set. *A distribution f is said to vanish in an open set Ω if its value on any test function with support in Ω is equal to zero*, and in this case one writes

$$f(x) = 0\,, \quad x \in \Omega \quad \Leftrightarrow \quad (f, \varphi) = 0\,, \quad \varphi \in \mathcal{D}(\Omega)\,.$$

For example, the Dirac delta function vanishes in any open set that does not include the origin:

$$\delta(x) = 0\,, \quad x \in \Omega \subset (-\infty, 0) \cup (0, \infty)\,.$$

Indeed, for any test function φ from $\mathcal{D}(\Omega)$

$$(\delta, \varphi) = \varphi(0) = 0$$

because $\operatorname{supp}\varphi \subset \Omega$.

If the difference of two distributions is equal to zero in some open set, then they are equal in this set, that is, *two distributions f and g are said to be equal in an open set Ω if their values on any test function with support in Ω are equal*, and in this case one writes

$$f(x) = g(x)\,, \quad x \in \Omega \quad \Leftrightarrow \quad (f, \varphi) = (g, \varphi)\,, \quad \varphi \in \mathcal{D}(\Omega)\,.$$

For example,

$$\mathcal{P}\frac{1}{x} = \frac{1}{x}\,, \quad x \neq 0\,.$$

16.1.2. Support of a distribution. *Let O_f be the largest open set in which a distribution f vanishes. Then its complement in called the support of the distribution f:*

$$\operatorname{supp} f = \mathbb{R}^N \setminus O_f\,.$$

By construction, the support of a distribution is a closed set.

The support of the Dirac delta-function is the origin:

$$\operatorname{supp} \delta = \{x = 0\}$$

The step-function $\theta(x)$, $x \in \mathbb{R}$, is continuous for $x \neq 0$. Therefore the largest open set on which $\theta(x) = 0$ is the interval $(-\infty, 0)$, and

$$\operatorname{supp} \theta = [0, \infty)\,.$$

The support of the principal value distribution is the whole real axis:

$$\operatorname{supp} \mathcal{P}\,\frac{1}{x} = \mathbb{R}\,.$$

16.1.3. Classical and distributional supports of regular distributions. Let us compare distributional and classical supports (see Sec.**1.2.2**) for regular distributions. The assertion is that *distributional and classical supports do not generally coincide, but they are the same for regular distributions defined by continuous functions.*

Suppose first that a distribution is defined by a continuous function f. Let A_f be a collection of all points where $f(x) \neq 0$. The classical support is the closure of A_f. If $x_0 \in A_f$, then $f(x) \neq 0$ near x_0 by continuity of f so that the integral (f, φ) cannot vanish for all test functions φ supported near x_0. Therefore x_0 is not in $\mathcal{O}_f$, which implies that A_f and $\mathcal{O}_f$ do not intersect. Let x_0 be a limit point of A_f that is not in A_f. Then there exists a sequence $x_n \in A_f$ such that $x_n \to x_0$ and any ball $B_a(x_0)$ contains elements of the sequence $\{x_n\} \subset A_f$. The latter implies that x_0 cannot be in $\mathcal{O}_f$ because $\mathcal{O}_f$ is open and there exists a ball centered at x_0 that lies in $\mathcal{O}_f$ and, hence, cannot contain any points from A_f. Thus, the complement of $\mathcal{O}_f$ contains the classical support of f.

Conversely, let x_0 be in the distributional support of f. If $f(x_0) \neq 0$, then $x_0 \in A_f$. Let $f(x_0) = 0$. Since $x_0 \notin \mathcal{O}_f$, $f(x)$ cannot vanish everywhere in any neighborhood of x_0. This implies that one can find a sequence x_n such that $f(x_n) \neq 0$ and $x_n \to x_0$. Therefore x_0 must be in the closure of A_f, which implies that the distributional support of f

is contained in its classical support. Thus, the distributional and classical supports coincide for continuous functions. In particular, every test function defines a regular distribution, and its distributional and classical supports are the same.

If f is defined by a generic locally integrable function, then the distributional and classical supports can be quite different. For example, let a distribution f be defined by the Dirichlet function. Then f is the zero distribution and the distributional support is empty. On the other hand, A_f is the set of all rational numbers and its closure is the whole $\mathbb{R}$. One can add the Dirichlet function to any locally integrable function so that the classical support can always be made $\mathbb{R}$ for any regular distribution. However, the distribution and, hence, its support cannot be changed by adding the zero distribution defined by any function that vanishes almost everywhere.

It follows from the definition that *if supports of a distribution f and a test function φ have no common points, then f vanishes on φ*:

$$\text{supp}\, f \cap \text{supp}\, \varphi = \emptyset \quad \Rightarrow \quad (f, \varphi) = 0\,. \tag{16.1}$$

16.2. Extensions of a distribution. Let Ω_1 be an open subset of an open set Ω_2 and f be a distribution from $\mathcal{D}'(\Omega_1)$. A distribution $g \in \mathcal{D}'(\Omega_2)$ is called an *extension* of f to Ω_2 if

$$g(x) = f(x)\,, \quad x \in \Omega_1 \subset \Omega_2\,.$$

An extension, if it exists, is not unique because one can always add to it any distribution h with support in the difference $\Omega_2 \backslash \Omega_1$.

For example, put $\Omega_1 = (-\infty, 0) \cup (0, \infty)$ and $\Omega_2 = \mathbb{R}$ and let $f(x) = \frac{1}{x}$ which is a regular distribution from $\mathcal{D}'(\Omega_1)$. Then the distribution

$$g(x) = \mathcal{P}\frac{1}{x} + h(x)\,, \qquad \text{supp}\, h = \{x = 0\}\,,$$

is an extension of f for any distribution h with support consisting of the single point $x = 0$. For example, a linear combination of the delta-function and its derivatives has the point support $x = 0$:

$$h(x) = c_0 \delta(x) + \sum_{k=1}^{m} c_k \delta^{(k)}(x)\,.$$

It will be proved in Sec.**21.7.1** that any distribution with a point support has this form. Therefore any distributional extension of $\frac{1}{x}$ to $\mathbb{R}$ coincides with the sum of $\mathcal{P}\frac{1}{x}$ and a linear combination of the delta-functions and its derivatives.

16.3. Distributional regularization of a singular function. A function f is said to have a non-integrable singularity at a point x_0 if it is not integrable in any neighborhood of x_0:

$$\int_{B_a(x_0)} |f(x)|\, d^N x = \infty\,.$$

So, f does not define a regular distribution in any open set containing x_0. Let S_f be the set of all non-integrable singularities of a function $f(x)$. Then f is a regular distribution on $\mathbb{R}^N \setminus S_f$. Note that S_f is a closed set in $\mathbb{R}^N$. If this distribution can be extended to the whole $\mathbb{R}^N$, then this extension is also called a *distributional regularization* of the function f and denoted by $\mathrm{Reg} f(x)$:

$$\mathrm{Reg}\, f(x) \in \mathcal{D}' \;:\quad \mathrm{Reg}\, f(x) = f(x)\,,\quad x \in \mathbb{R}^N \setminus S_f\,.$$

Clearly, a regularization, if it exists, is a singular distribution that is not unique because it is defined up to an additive distribution with support in S_f.

For example, the function $f(x) = \frac{1}{x^n}$, $x \in \mathbb{R}$, is singular at $x = 0$ if n is a positive integer. The principal value distribution (**15.3**) is its distributional regularization because

$$\mathcal{P}\frac{1}{x^n} = \frac{1}{x^n}\,,\quad x \neq 0\,.$$

Indeed, if support of a test function φ does not contain $x = 0$, then φ and all its derivatives vanish at $x = 0$. It follows from (**15.3**) that

$$\Big(\mathcal{P}\frac{1}{x^n}, \varphi\Big) = \int \frac{\varphi(x)}{x^n}\, dx\,,\quad 0 \notin \mathrm{supp}\,\varphi\,.$$

The existence of the integral is guaranteed by vanishing φ in a neighborhood of $x = 0$. As noted in the previous section, any distributional regularization of $\frac{1}{x^n}$ can be written in the form

$$\mathrm{Reg}\,\frac{1}{x^n} = \mathcal{P}\frac{1}{x^n} + \sum_{k=0}^{m} c_k \delta^{(k)}(x)\,,$$

for some integer m and constants c_k.

There are two basic techniques that are commonly employed to find a distributional regularization of a singular function:

(i) *Principal value regularizations*,
(ii) *Shifting singularities into a complex plane*.

Later, in Sec.**21.7**, it will be shown that all distributions can be obtained as linear combinations of distributional derivatives of continuous functions. Therefore if a singular function coincides with a linear combination classical derivatives of some continuous functions wherever the derivatives exist, then its distributional regularization is the linear combination of the corresponding distributional derivatives of those continuous functions. In practice, however, this general approach is not easy to use for finding a distributional regularization of a given singular function, whereas the above two techniques often lead quickly to a desired result. There are singular functions that cannot be written as a linear combination of derivatives of continuous functions. Such functions do not have a distributional regularization. An example is given below in Sec.**16.5**.

16.4. Principal value regularizations. Suppose a function f has a non-integrable singularity at a single point. Without loss of generality, it is set to be at the origin. Suppose further that

$$f(x) = \frac{g(x)}{|x|^s}\,, \quad x \in \mathbb{R}^N\,, \quad s > 0\,, \quad g \in \mathcal{L}_{\rm loc}\,.$$

Consider the functional on $\mathcal{D}$ defined by the rule

$$(\mathbf{16.2})\quad (\mathcal{P}_r f, \varphi) = \int\limits_{|x|<1} f(x)\Big(\varphi(x) - T_m(x)\Big)\,d^N x + \int\limits_{|x|>1} f(x)\varphi(x)\,d^N x$$

where T_m is the Taylor polynomial of order m for φ about the singular point $x_0 = 0$ (see (**1.6**)). The choice of the unit ball, $|x| < 1$, to isolate the singular point in the rule (**16.2**) is a convention. One can choose a ball of any suitable radius.

Let us show that the rule (**16.2**) makes sense for any test function for some large enough m. By Taylor's theorem (see (**1.7**)), there is a constant M such that

$$|\varphi(x) - T_m(x)| \le M|x|^{m+1}\,,$$

in a neighborhood of $x = 0$. Therefore the integral over the unit ball exists if m is such that

$$g(x)|x|^{m+1-s} \in \mathcal{L}(|x| < 1)\,.$$

For definitiveness, let m be *the smallest positive integer for which this condition holds*. In particular, if g is bounded in a neighborhood of $x = 0$, then the condition holds if $m > s - N - 1$ (see Sec.**4.5.3**). For example,

$$(\mathbf{16.3})\qquad \Big(\mathcal{P}_r \frac{1}{|x|^N}, \varphi\Big) = \int_{|x|<1} \frac{\varphi(x) - \varphi(0)}{|x|^N}\,d^N x + \int_{|x|>1} \frac{\varphi(x)}{|x|^N}\,d^N x\,.$$

Linearity of the functional (**16.2**) is obvious. It remains to show continuity. Take a null sequence in $\mathcal{D}$, $\varphi_n \to 0$ as $n \to \infty$. One has to show that the numerical sequence defined by (**16.2**) where $\varphi = \varphi_n$ converges to zero. Since supports of all φ_n are in one ball $|x| \le R$, the second integral converges to zero because

$$\Big|\int_{|x|>1} f(x)\varphi_n(x)\,d^N x\Big| = \Big|\int_{1<|x|<R} f(x)\varphi_n(x)\,d^N x\Big| \le \sup|\varphi_n|\int_{1<|x|<R} |f(x)|\,d^N x \;\to\; 0$$

as $n \to \infty$ because f is integrable on any compact that does not contain $x = 0$. To estimate the first term, the inequality (**1.7**) is used. Let $T_{mn}(x)$ is a Tyalor polynomial for φ_n about $x_0 = 0$. Then by (**1.7**)

$$|\varphi_n(x) - T_{mn}(0)| \le M_n|x|^{m+1}\,, \quad M_n \sim \max_{|\alpha|=m}\sup|D^\alpha\varphi_n| \to 0\,,$$

as $n \to \infty$. Therefore

$$\Big|\int\limits_{|x|<1} f(x)[\varphi_n(x) - T_{mn}(x)]\,d^N x\Big| \le M_n \int\limits_{|x|<1} |g(x)||x|^{m+1-s}\,d^N x \to 0\,.$$

Thus, $\mathcal{P}_r f(x)$ is a distribution.

It remains to show that $\mathcal{P}_r f(x)$ is a distributional extension of $f(x)$ to $\mathbb{R}^N$. If $\operatorname{supp}\varphi$ does not contain the singular point of f in (**16.2**), then φ and all its derivatives vanish near $x = 0$ so that $T_m(x) = 0$ for any m and

$$(\mathcal{P}_r f, \varphi) = \int f(x)\varphi(x)\,d^N x\,,$$

which means that

$$\mathcal{P}_r f(x) = f(x)\,, \quad x \ne 0\,,$$

as required.

By invoking the result of Sec.**21.7.1**, any distributional regularization (or extension) of f can always be written in the form

$$\operatorname{Reg} f(x) = \mathcal{P}_r f(x) + \sum_{|\alpha|\le k} C_\alpha D^\alpha \delta(x) \tag{16.4}$$

for some integer $k \ge 0$ and constants C_α.

It is also worth noting that the rule (**16.2**) defines a distribution for all large enough m. All these distributions differ from $\mathcal{P}_r f$ by terms containing the delta function and its derivatives. For example, let $s = N$, and $g(x)$ be bounded near $x = 0$. Then $m = 0$ in (**16.2**) as in

(**16**.**3**). However, if $T_0 = \varphi(0)$ is replaced by $T_1(x) = T_0 + (x, \nabla\varphi(0))$ in (**16**.**2**), the new distributional regularization has an extra term

$$-\int_{|x|<1} g(x)\frac{(x,\nabla\varphi(0))}{|x|^N}\,d^N x = -(c,\nabla\varphi(0)) = \Big((c,\nabla)\delta,\varphi\Big)\,,$$

where components of the vector c are given by the integrals

$$c_j = \int_{|x|<1} g(x)\frac{x_j}{|x|^N}\,d^N x\,.$$

So, the distribution $\mathcal{P}_r f(x)$ is changed by adding the term $(c,\nabla)\delta(x)$ in full accord with (**16**.**4**).

16.4.1. Relation to the Cauchy principal value regularization. For singular functions of one real variable, the Cauchy principal value regularization does not generally coincide with the distribution (**16**.**2**). According to (**16**.**4**), they can differ at most by a linear combination of the delta-function and its derivatives. For example,

$$\mathcal{P}\frac{1}{x} = \mathcal{P}_r\frac{1}{x}\,,\qquad \mathcal{P}\frac{1}{x^2} = \mathcal{P}_r\frac{1}{x^2} - 2\delta(x)\,.$$

The last relation is established by the following calculations:

$$\begin{aligned}\Big(\mathcal{P}_r\frac{1}{x^2},\varphi\Big) &= \int_{-1}^{1}\frac{\varphi(x)-\varphi(0)-\varphi'(0)x}{x^2}\,dx + \int_{|x|>1}\frac{\varphi(x)}{x^2}\,dx\\ &= p.v.\int_{-1}^{1}\frac{\varphi(x)-\varphi(0)}{x^2}\,dx + \int_{|x|>1}\frac{\varphi(x)}{x^2}\,dx\\ &= p.v.\int\frac{\varphi(x)-\varphi(0)}{x^2}\,dx + \varphi(0)\int_{|x|>1}\frac{dx}{x^2}\\ &= \Big(\mathcal{P}\frac{1}{x^2},\varphi\Big) + 2(\delta,\varphi)\,.\end{aligned}$$

16.4.2. Functions with many singular points. As a final remark on principal value regularizations, let us define a distributional regularization of a function with countably many singular points $\{x_n\}$. It is assumed that any ball contains at most finitely points from $\{x_n\}$. Suppose that

$$f(x) = \frac{g_n(x)}{|x-x_n|^{\nu_n}}\,,\quad |x-x_n|<a\,,$$

where $g_n(x)$ is locally integrable. Let Ω_a be the complement of the union of balls $|x-x_n|\le a$ in $\mathbb{R}^N$ and $f(x)$ be locally integrable in Ω_a.

Then the principal value regularization of f is defined by the rule

$$(\mathcal{P}_r f, \varphi) = \sum_n \int_{|x-x_n|<a} f(x)\Big(\varphi(x) - T_m(x;x_n)\Big)\, d^N x + \int_{\Omega_a} f(x)\varphi(x)\, d^N x\,,$$

where $T_m(x;x_n)$ is a Taylor polynomial of φ about $x = x_n$ of a minimal order m defined by that $f(\varphi - T_m)$ is integrable on a ball $|x - x_n| < a$. The series converges because there are finitely many points x_n is a ball that contains support of a test function φ (the sum has only finitely many terms for any test function). Continuity of $\mathcal{P}_r f$ is proved in the same way as for the principal value regularization of a function with one singular point. If support of φ contains no singular point, then $\varphi = 0$ near any x_n and $T_m(x;x_n) = 0$ for any $m \geq 0$ so that

$$(\mathcal{P}_r f, \varphi) = \int f(x)\varphi(x)\, d^N x \quad \Rightarrow \quad \mathcal{P}_r f(x) = f(x)\,, \quad x \neq x_n\,.$$

In other words, $\mathcal{P}_r f$ is a distributional regularization of a singular function f.

It should be noted that for singular functions of one real variable, there exists an alternative distributional regularization near x_n if g_n is a smooth function and ν_n is an integer:

$$\operatorname{Reg} f(x) = g_n(x)\mathcal{P}\frac{1}{(x-x_n)^{\nu_n}}\,, \quad |x - x_n| < a\,.$$

As noted earlier, different distributional regularizations differ only by distributions supported at $x = x_n$ (by a linear combination of $\delta(x - x_n)$ and its derivatives).

For example, let $f(x) = \frac{1}{\sin(x)}$. This function has non-integrable singularities at $x = x_n = \pi n$ where n is any integer, and

$$f(x) = \frac{g_n(x)}{x - \pi n}\,, \quad g_n(x) = \frac{x - \pi n}{\sin(x)}\,, \quad x \in I_n^a = (\pi n - a, \pi n + a)\,,$$

where $0 < a \leq \frac{\pi}{2}$. Let I_a be the complement of the union of all I_n^a. Note that I_a is a set of measure zero if $a = \frac{\pi}{2}$. Then the principal value regularization reads

$$\begin{aligned}\Big(\mathcal{P}_r \frac{1}{\sin(x)}, \varphi(x)\Big) &= \sum_n \int_{I_n^a} \frac{\varphi(x) - \varphi(\pi n)}{\sin(x)}\, dx + \int_{I_a} \frac{\varphi(x)}{\sin(x)}\, dx \\ &= \sum_n \int_{I_n} \frac{\varphi(x) - \varphi(\pi n)}{\sin(x)}\, dx\,,\end{aligned}$$

where $I_n = I_n^a$ for $a = \frac{\pi}{2}$. The regularization does not depend on a in this case because

$$\int_{I_n \setminus I_n^a} \frac{dx}{\sin(x)} = 0$$

by the skew-symmetry of the integrand under the reflection of the argument about $x = \pi n$. Since $g_n \in C^\infty(I_n)$ and $\nu_n = 1$, an alternative regularization based on the Cauchy principal value distribution is

$$\Big(\mathcal{P}\frac{1}{\sin(x)}, \varphi(x)\Big) = \sum_n p.v. \int_{I_n^a} \frac{\varphi(x)}{\sin(x)}\, dx + \int_{I_a} \frac{\varphi(x)}{\sin(x)}\, dx\,.$$

For this function, both the regularizations produce the same distribution because

$$p.v. \int_{I_n^a} \frac{dx}{\sin(x)} = 0 \quad \Rightarrow \quad \mathcal{P}_r \frac{1}{\sin(x)} = \mathcal{P}\frac{1}{\sin(x)}\,.$$

16.5. On the existence of a distributional regularization. *There are singular functions that do not admit any distributional regularization near their singular points.* A singular function can "blow up" too fast at a singular point so that the trick with subtracting a Taylor polynomial about the singular point will not work. Furthermore, no other regularization trick will work either. The assertion is illustrated by the following example.

Let

$$f(x) = \exp\left(\frac{1}{x}\right), \quad x \neq 0\,.$$

Clearly, $x = 0$ is the only singular point because f is smooth everywhere but $x = 0$, and $f(x)$ tends to zero faster then any power function when $x \to 0^-$, and $f(x)$ blows up to infinity faster than any reciprocal power function when $x \to 0^+$:

$$\lim_{x\to 0^-} |x|^{-p} e^{\frac{1}{x}} = 0\,, \quad \lim_{x\to 0^+} |x|^{p} e^{\frac{1}{x}} = \infty\,, \quad p > 0\,.$$

Suppose that there exists a distributional extension g of f to $\mathbb{R}$. Then g must be a linear continuous functional on $\mathcal{D}(\mathbb{R})$. Let us show that the latter is false and, hence, by contradiction, a distribution g does not exist.

Let φ be a non-negative test function with support in $(1, b)$. It can always be normalized so that

$$\int \varphi(x)\, dx = \int_1^b \varphi(x)\, dx = 1\,.$$

If g is an extension of f, then

$$(g, \varphi) = (f, \varphi) = \int_1^b e^{\frac{1}{x}} \varphi(x)\, dx\,.$$

Consider a sequence of test functions in $\mathcal{D}$

$$\varphi_n(x) = e^{-\frac{n}{b}} n\, \varphi(nx)\,, \quad n = 1, 2, \ldots$$

Then the sequence is a null sequence $\varphi_n \to 0$ in $\mathcal{D}$ as $n \to \infty$. Indeed, the supports of all φ_n lie in $[1, b]$ and

$$|D^\alpha \varphi_n(x)| \le M_\alpha e^{-\frac{n}{b}} n^{\alpha+1}\,, \quad M_\alpha = \sup |D^\alpha \varphi(x)|$$

so that φ_n and $D^\alpha \varphi_n$ converge uniformly to 0 for any α. If g is a distribution, then

$$(g, \varphi_n) \to 0\,, \quad n \to \infty$$

because g must be a continuous functional on $\mathcal{D}$. On the other hand, the numerical sequence (g, φ_n) is bounded from below by 1 and, hence, cannot converge to 0. Indeed,

$$\begin{aligned}(g, \varphi_n) &\overset{(1)}{=} \int e^{\frac{1}{x}} \varphi_n(x)\, dx \overset{(2)}{=} \int e^{\frac{n}{y} - \frac{n}{b}} \varphi(y)\, dy \\ &\overset{(3)}{=} \int_1^b e^{n(\frac{1}{y} - \frac{1}{b})} \varphi(y) \overset{(4)}{\ge} \int_1^b \varphi(x)\, dx = 1 > 0\,.\end{aligned}$$

Here (1) holds because supports of φ_n do not contain $x = 0$; (2) is obtained by changing variables $y = nx$; (3) holds because the support of φ lies in $(1, b)$; (4) follows from $e^{n(\frac{1}{y} - \frac{1}{b})} > 1$ if $1 < y < b$ and the hypothesis $\varphi(x) \ge 0$. Thus, no distributional extension (or regularization) of f exists.

16.6. Sokhotsky's distributions. A distributional regularization of a singular function, if it exists, can also be obtained by moving singular points into a complex plane. The idea is first illustrated with an example of *Sokhotsky's distributions*.

The function $\frac{1}{x}$ is singular. Consider locally integrable complex-valued functions of a real variable obtained from $\frac{1}{x}$ by shifting the singularity at $x = 0$ to the complex plane:

$$f_{\pm a}(x) = \frac{1}{x \pm ia}\,, \quad a > 0\,.$$

They define regular complex-valued distributions by the rule

$$(f_{\pm a}, \varphi) = \int \frac{\varphi(x)}{x \pm ia}\, dx$$

for every $a > 0$. If the limit of $(f_{\pm a}, \varphi)$ exists for any test function as $a \to 0^+$, then by the completeness theorem it defines a distribution, denoted $f_\pm$. By construction, it is a distributional regularization of $\frac{1}{x}$. Indeed, suppose that the support of a test function φ lies in $0 < \delta \le |x|$. It follows from the inequality

$$\Big|\frac{\varphi(x)}{x \pm ia}\Big| \le \frac{|\varphi(x)|}{\delta}\,, \quad |x| \ge \delta\,,$$

that the order of taking the limit and integration can be interchanged by the Lebesgue dominated convergence theorem so that

$$(f_\pm, \varphi) = \lim_{a\to 0^+} \int \frac{\varphi(x)}{x \pm ia}\, dx = \int_{|x|>\delta} \frac{\varphi(x)}{x}\, dx\,,$$

which means that

$$f_\pm(x) = \frac{1}{x}\,, \quad x \ne 0\,,$$

because $\delta > 0$ is arbitrary.

Let us show that the limit exists for any test function. Since the support of φ lies in some interval $[-R, R]$, the following chain of equalities holds

$$\begin{aligned}(f_{\pm a}, \varphi) = \int_{-R}^{R} \frac{\varphi(x)}{x \pm ia}\, dx &= \int_{-R}^{R} \frac{\varphi(0)}{x \pm ia}\, dx + \int_{-R}^{R} \frac{\varphi(x) - \varphi(0)}{x \pm ia}\, dx \\ &= \varphi(0) \int_{-R}^{R} \frac{x \mp ia}{x^2 + a^2}\, dx + \int_{-R}^{R} \frac{\varphi(x) - \varphi(0)}{x \pm ia}\, dx\end{aligned}$$

The first integral is easy to evaluate

$$\int_{-R}^{R} \frac{x \mp ia}{x^2 + a^2}\, dx = 0 \mp i \int_{-R/a}^{R/a} \frac{dy}{y^2 + 1} = \mp 2i \arctan\Big(\frac{R}{a}\Big) \to \mp i\pi$$

as $a \to 0^+$. To find the limit of the second integral, let us use the Lebesgue dominated convergence theorem. Put

$$g(x, a) = \frac{\varphi(x) - \varphi(0)}{x \pm ia}$$

Then the limit of $g(x, a)$ as $a \to 0^+$ exists for almost every x because

$$\lim_{a\to 0^+} g(x, a) = \frac{\varphi(x) - \varphi(0)}{x}\,, \quad x \ne 0$$

and $|g(x, a)|$ has a Lebesgue integrable bound

$$|g(x, a)| = \frac{|\varphi(x) - \varphi(0)|}{|x \pm ia|} \le \frac{|\varphi(x) - \varphi(0)|}{|x|} \le M\,,$$

where $M = \sup|\varphi'|$ by the mean value theorem and a constant function is integrable on $(-R, R)$. Therefore by the Lebesgue dominated convergence theorem

$$\lim_{a\to 0^+}(f_{\pm a}, \varphi) = \mp i\pi\varphi(0) + \int_{-R}^{R} \frac{\varphi(x) - \varphi(0)}{x}\, dx = (f_\pm, \varphi)\,. \tag{16.5}$$

The limit distributions $f_\pm$ are called *Sokhotsky's distributions* and are denoted as

$$f_\pm(x) = \frac{1}{x \pm i0^+}\,.$$

16.6.1. Sokhotsky's equations. Sokhotsky's distributions and the principal value distribution are distributional regularizations of $\frac{1}{x}$. Therefore there should exists a relation between them and the delta function and its derivatives. Let us find this relation. It is known as *Sokhotsky's equations.*

The integral in (**16**.**5**) can be transformed as follows:

$$\begin{aligned}\int_{-R}^{R} \frac{\varphi(x) - \varphi(0)}{x}\, dx &\overset{(1)}{=} \lim_{a\to 0^+}\left(\int_{-R}^{-a} + \int_{a}^{R}\right)\frac{\varphi(x) - \varphi(0)}{x}\, dx \\ &\overset{(2)}{=} \lim_{a\to 0^+}\left(\int_{-R}^{-a} + \int_{a}^{R}\right)\frac{\varphi(x)}{x}\, dx \\ &\overset{(3)}{=} P.v. \int \frac{\varphi(x)}{x}\, dx = \left(\mathcal{P}\frac{1}{x}, \varphi\right).\end{aligned}$$

Here (1) is by continuity of the integral; (2) holds because the integral of $1/x$ over the symmetric region $a < |x| < R$ vanishes; (3) is the definition of the principal value integral and by that $\varphi(x) = 0$ for all $|x| > R$. Therefore

$$(f_\pm, \varphi) = \mp i\pi(\delta, \varphi) + \left(\mathcal{P}\frac{1}{x}, \varphi\right)$$

for any test function φ, or

$$\frac{1}{x \pm i0^+} = \mp i\pi\,\delta(x) + \mathcal{P}\frac{1}{x}\,, \tag{16.6}$$

which are Sokhotsky's equations.

16.7. A higher dimensional example. The function $f(x) = (|x|^2 - m^2)^{-1}$ is not locally integrable in $\mathbb{R}^N$ because it has non-integrable singularities on the sphere $|x| = m > 0$. One can find a distributional regularization of f by means of shifting singular points into a complex plane. Let us show that

$$\operatorname{Reg} \frac{1}{|x|^2 - m^2} = \frac{1}{|x|^2 - m^2 + i0}\,, \quad x \in \mathbb{R}^N\,,$$

where

$$\Big(\frac{1}{|x|^2 - m^2 + i0}, \varphi\Big) \stackrel{\text{def}}{=} \lim_{a\to 0^+} \int \frac{\varphi(x)}{|x|^2 - m^2 + ia}\, d^N x\,, \quad \varphi \in \mathcal{D}\,,$$

is a distributional regularization of f in the whole $\mathbb{R}^N$. The integral exists for any $a \neq 0$. Therefore, if the limit is proved to exist for any test function, then by the completeness theorem it defines a distribution.

Suppose that the support of a test function φ does not overlap with the sphere $|x| = m$. Therefore there is a non-zero distance $\delta > 0$ between the sphere and $\operatorname{supp}\varphi$. It follows from the inequality

$$\Big|\frac{\varphi(x)}{|x|^2 - m^2 + ia}\Big| \le \frac{|\varphi(x)|}{\delta^2}\,, \quad \delta^2 \le \big||x|^2 - m^2\big|\,,$$

that the order of taking the limit and integration can be interchanged by the Lebesgue dominated convergence theorem so that

$$\Big(\frac{1}{|x|^2 - m^2 + i0}, \varphi\Big) = \int \frac{\varphi(x)}{|x|^2 - m^2}\, d^N x\,,$$

which means that

$$\frac{1}{|x|^2 - m^2 + i0} = \frac{1}{|x|^2 - m^2}\,, \quad |x| \neq m\,.$$

So, the distribution is indeed a distributional regularization of the singular function $(|x|^2 - m^2)^{-1}$. This regularization is unique up an additive distribution supported on the sphere $|x| = m$. For example, the distribution

$$\operatorname{Reg} \frac{1}{|x|^2 - m^2} = \frac{1}{|x|^2 - m^2 + i0} + \nu(x)\delta_{S_m}(x)$$

is also a regularization where $\nu\delta_{S_m}$ is a delta function on the sphere S_m with density $\nu(x)$ (see Sec.**15.6**).

It remains to show that the limit exists for any test function. First, for any test function $\varphi \in \mathcal{D}(\mathbb{R}^N)$, the function

$$\psi(r) = \int_{|y|=1} \varphi(ry)\, dS_y \;\in\; \mathcal{D}(\mathbb{R})\,,$$

is a test function of one real variable r. Indeed, if the support of φ lies in a ball of radius R, then the support of ψ lies in the interval $[-R, R]$. The function is even $\psi(-r) = \psi(r)$ because the measure dS_y is invariant under the parity transformation $y \to -y$. Any partial derivative of the integrand with respect to r has a bound independent of r that is integrable on the unit sphere:

$$\Big|D_r^\alpha \varphi(ry)\Big| = \Big|D_{j_1}\cdots D_{j_\alpha}\varphi(x)\Big|_{x=ry} y_{j_1}\cdots y_{j_\alpha}\Big| \le \sup|D^\alpha\varphi|$$

where $|y_j| \le |y| = 1$ was used. Any constant function is integrable on a unit sphere. Therefore by Theorem **7.2**, ψ is a smooth function and, hence, $\psi \in \mathcal{D}(\mathbb{R})$.

By the partial fraction decomposition,

$$\frac{1}{|x|^2 - z^2} = \frac{1}{2z}\Big(\frac{1}{|x|-z} - \frac{1}{|x|+z}\Big)$$

where $z = m(1-\frac{ia}{m})^{1/2} = m - i\xi + O(\xi^2)$, $\xi = 2a/m$, one infers that

$$\Big(\frac{1}{|x|^2 - m^2 + i0}, \varphi\Big) = \frac{1}{2m}\lim_{\xi\to 0^+}\int_{B_R}\frac{\varphi(x)\,d^Nx}{|x|-m+i\xi} + \frac{1}{2m}\int_{B_R}\frac{\varphi(x)\,d^Nx}{|x|+m}$$

if the support of φ lies in a ball B_R. Converting the first integral into spherical coordinates,

$$\int_{B_R}\frac{\varphi(x)\,d^Nx}{|x|-m+i\xi} = \int_0^R \frac{\psi(r)r^{N-1}dr}{r-m+i\xi}\,.$$

To evaluate the limit $\xi \to 0^+$, put $\phi(r) = \psi(r)r^{N-1}$ which is a smooth function near $r = m > 0$. Therefore

$$\int_0^R \frac{\phi(r)\,dr}{r-m+i\xi} = \int_0^R \frac{\phi(r)-\phi(m)}{r-m+i\xi}\,dr + \phi(m)\int_0^R\frac{dr}{r-m+i\xi}\,.$$

The limit of the first integral exists because

$$\phi(r) - \phi(m) = \phi'(m)(r-m) + O((r-m)^2)$$

so that the integrand has no singularity at $\xi = 0$. The limit of the second integral is evaluated directly:

$$\begin{aligned}\int_0^R \frac{dr}{r-m+i\xi} &= \int_{-m}^{R-m}\frac{s\,ds}{s^2+\xi^2} - i\xi\int_{-m}^{R-m}\frac{ds}{s^2+\xi^2}\\ &= \frac{1}{2}\ln(s^2+\xi^2)\Big|_{-m}^{R-m} - i\arctan\Big(\frac{s}{\xi}\Big)\Big|_{-m}^{R-m}\\ &\to \ln\Big(\frac{R-m}{m}\Big) - i\pi \quad \text{as } \xi\to 0^+\,, \end{aligned} \tag{16.7}$$

where it was assumed that $R > m$ (otherwise the original integral exists without regularization).

16.8. Exercises.

1. Use the du Bois-Reymond lemma to show that any distribution whose support has measure zero is a singular distribution.

2. Find the support of a regular distribution defined by the function $f(x) = \sin(x)$.

3. Find the support of a regular distribution defined by a locally integrable function that vanishes only on a set measure zero.

4. Show that the support of $\delta(x)$ and all partials $D^\beta\delta(x)$, $x \in \mathbb{R}^N$, is the point $x = 0$.

5. Find the support of the spherical delta function in $\mathbb{R}^3$

$$(\delta_{S_a}, \varphi) = \int_{|x|=a} \varphi(x)\, dS$$

6. Let $\theta(y)$ be the step function. Then the locally integrable function

$$f(t,x) = \theta(c^2t^2 - |x|^2)\,, \quad t \in \mathbb{R}\,, \quad x \in \mathbb{R}^3$$

defines a distribution in $\mathbb{R}^4$, where $c > 0$ is a constant. Find its support.

7. Let $t \in \mathbb{R}$ and $x \in \mathbb{R}^2$. Find the support of the distribution defined by the rule

$$(f, \varphi) = \int_0^\infty \int_{|x|=ct} \varphi(x,t)\, dS\, dt$$

where dS stands for the line integral over the circle $|x| = ct$, and $c > 0$ is a constant.

8. Put

$$\Big(\mathrm{Reg}\frac{1}{x}, \varphi\Big) = \int \frac{\varphi(x) - \varphi(-x)}{2x}\, dx$$

for any test function $\varphi \in \mathcal{D}(\mathbb{R})$.
(i) Show that this rule defines a distribution and
(ii) this distribution is an extension of a singular function $\frac{1}{x}$ to $x = 0$, that is,

$$\mathrm{Reg}\frac{1}{x} = \frac{1}{x}\,, \quad x \neq 0\,,$$

in the distributional sense in any open interval that does not contain $x = 0$;
(iii) Find a relation between this distribution and the principal value

distribution $\mathcal{P}\frac{1}{x}$.

9. Show that the functional defined by the rule (**16.2**) is a distribution (a linear continuous functional on $\mathcal{D}$) if
(i) $x \in \mathbb{R}$
(ii) $x \in \mathbb{R}^N$.
Hints: Let $\varphi_n \to 0$ in $\mathcal{D}$. Let p_{nm} be the Taylor polynomial for φ_n of order m about $x = 0$. Use Taylor's theorem to show that

$$|\varphi_n(x) - p_{nm}(x)| \leq M_n |x|^{m+1}\,, \quad |x| < a\,.$$

for some $a \leq 1$ and $M_n \to 0$ as $n \to \infty$. Use the above inequality to show that $(\mathcal{P}_r f, \varphi_n) \to 0$ as $n \to \infty$.

10. Let

$$f_a(x) = \frac{1}{2a}\Big(\delta_{S_{R+a}}(x) - \delta_{S_{R-a}}(x)\Big)$$

where $R > a > 0$ and δ_{S_R} is a spherical delta-function in $\mathbb{R}^N$. Find the distributional limit of f_a as $a \to 0^+$. Give an explicit rule for the value of the limit distribution on a test function.

11. Let $f(x) = \cot(x)$. It is integrable on any interval $[a, b]$ that does not contain points $x_n = \pi n$, $n = 0, \pm 1, \pm 2, \ldots$.
(i) Show that $f(x)$ does not define a regular distribution from $\mathcal{D}'(-\frac{\pi}{2}, \frac{\pi}{2})$ but it has a distributional regularization in $(-\frac{\pi}{2}, \frac{\pi}{2})$, that is, there exists a distribution $g \in \mathcal{D}'(-\frac{\pi}{2}, \frac{\pi}{2})$ such that in the sense of distributions

$$g(x) = \cot(x)\,, \quad x \in (-\tfrac{\pi}{2}, 0) \cup (0, \tfrac{\pi}{2})$$

Hint: Consider the principal value integral

$$p.v. \int_{-R}^{R} \cot(x)\varphi(x)\, dx\,, \quad 0 < R < \tfrac{\pi}{2}\,, \quad \varphi \in \mathcal{D}(-\tfrac{\pi}{2}, \tfrac{\pi}{2})\,.$$

(ii) Use the periodicity of $\cot(x)$ to show that it has a distributional regularization in the whole $\mathbb{R}$, and find it.

12. Let $f(x) = |\cot(x)|$.
(i) Use the rule (**16.2**) to obtain a distributional regularization $\mathcal{P}_r f$ in $\mathcal{D}'(-\frac{\pi}{2}, \frac{\pi}{2})$.
(ii) Use periodicity of f to extend $\mathcal{P}_r f$ to the whole $\mathbb{R}$.

13. Show that the functional

$$f(x_0, x) = \frac{1}{x_0^2 - |x|^2 - m^2 + i0^+}$$

defined by the rule

$$(f, \varphi) = \lim_{a \to 0^+} \int \int_{-\infty}^{\infty} \frac{\varphi(x_0, x)}{x_0^2 - |x|^2 - m^2 + ia} \, dx_0 \, d^N x \,,$$

for any test function φ, where $x \in \mathbb{R}^N$, is a distribution on $N + 1$ variables.

17. Transformations of distributions

Classical functions are included into a space of distributions. There are transformations that allows us to make new functions out of the existing ones. For example, one can multiply two functions, or make a composition of two functions, or take a derivative or antiderivative, etc. In this regard, two basic questions arise. *First, can standard transformations of classical functions be extended to distributions? Second, what is a general principle for constructing transformations of a space of distributions to another space of distributions?*

17.1. Adjoint transformations of distributions. Suppose a transformation T^* maps a locally integrable function f on $\mathbb{R}^N$ to a locally integrable function $T^*(f)$ on $\mathbb{R}^M$. The objective is to investigate whether or not it is possible to extend T^* to singular distributions:

$$T^*:\ f \in \mathcal{D}_1' \subseteq \mathcal{D}'(\mathbb{R}^N) \to T^*(f) \in \mathcal{D}_2' \subseteq \mathcal{D}'(\mathbb{R}^M)\,.$$

Distributions from the space $\mathcal{D}_1'$ are defined on a space of test functions $\mathcal{D}_1 \subseteq \mathcal{D}(\mathbb{R}^N)$ and, similarly, distributions from the space $\mathcal{D}_2'$ are defined on a space of test functions $\mathcal{D}_2 \subseteq \mathcal{D}(\mathbb{R}^M)$. For brevity, put $T^*(f)(y) = f_T(y)$.

Let $f(x)$, $x \in \mathbb{R}^N$, and its transformation $f_T(y)$, $y \in \mathbb{R}^M$, be locally integrable. For any test function $\varphi \in \mathcal{D}_2$

$$\Big(T^*(f), \varphi\Big) = \int f_T(y)\varphi(y)\, d^M y\,.$$

Suppose that one can manipulate this integral in some way to reduce it to the form

$$\int f_T(y)\varphi(y)\, d^M y = \int f(x)\varphi_T(x)\, d^N x = (f, T(\varphi))\,, \tag{17.1}$$

where the function $\varphi_T = T(\varphi)$ is a transformation T of a test function. Thus, for any regular distribution one has the rule

$$(T^*(f), \varphi) = (f, T(\varphi))\,. \tag{17.2}$$

If this rule is to be extended to any $f \in \mathcal{D}_1'$, then it is necessary that T maps a space of test functions to another space of test functions:

$$T:\ \varphi \in \mathcal{D}_2 \subset \mathcal{D}(\mathbb{R}^M) \to T(\varphi) = \varphi_T \in \mathcal{D}_1 \subseteq \mathcal{D}(\mathbb{R}^N)\,. \tag{17.3}$$

However, not any such transformation is suitable. The functional $T^*(f)$ must be linear and continuous on $\mathcal{D}_2$. Since f is linear and continuous on $\mathcal{D}_1$, $T^*(f)$ is linear and continuous, *provided* T is a linear and continuous transformation:

$$\begin{aligned}\textsf{linearity}: &\quad T(c_1\varphi_1 + c_2\varphi_2) = c_1T(\varphi_1) + c_2T(\varphi_2)\\ \textsf{continuity}: &\quad \varphi_n \to \varphi \text{ in } \mathcal{D}_2 \quad \Rightarrow \quad T(\varphi_n) \to T(\varphi) \text{ in } \mathcal{D}_1\end{aligned}$$

for any numbers $c_{1,2}$ and any test functions $\varphi_{1,2} \in \mathcal{D}_2$.

Let us summarize our findings. For any linear and continuous transformation (**17.3**) on a space of test functions, one can define a transformation of distributions by the rule (**17.2**). In this case, the transformation T^* is called the *adjoint* of T. Conversely, any transformation that maps a locally integrable function to a locally integrable function can be extended to all distributions, provided this transformation can be interpreted as the adjoint transformation of some linear and continuous transformation on the space of test functions (**17.3**).

17.1.1. Continuity of the adjoint transformation. Let us show that *the adjoint transformation T^* of a space of distributions is continuous*, that is, if a sequence of distributions $\{f_n\}$ converges to f in $\mathcal{D}'_1$, then the image of the sequence $\{T^*(f_n)\}$ converges to the image of the limit distribution $T^*(f)$ in $\mathcal{D}'_2$:

$$f_n \to f \text{ in } \mathcal{D}'_1 \quad \Rightarrow \quad T^*(f_n) \to T^*(f) \text{ in } \mathcal{D}'_2\,.$$

For any test function $\varphi \in \mathcal{D}_2$ one has

$$\lim_{n\to\infty}(T^*(f_n), \varphi) = \lim_{n\to\infty}(f_n, T(\varphi)) = (f, T(\varphi)) = (T^*(f), \varphi)\,.$$

because, the hypotheses, the numerical sequence (f_n, ψ) converges to (f, ψ) for any test function $\psi = T(\varphi) \in \mathcal{D}_1$.

17.2. Linear change of variables. Consider a general linear change of variables in $\mathbb{R}^N$:

$$x = Ay + b\,, \quad \det A \neq 0\,,$$

where b is a constant vector. The Jacobian reads

$$d^N x = |\det A|\, d^N y\,.$$

Let $f(x)$ be a locally integrable function. Then $f_{A,b}(y) = f(Ay+b)$ is also locally integrable function and, hence, defines a regular distribution by the rule

$$\begin{aligned}(f_{A,b}, \varphi) &= \int f_{A,b}(y)\varphi(y)\, d^N y = \int f(Ay+b)\varphi(y)\, d^N y\\ &= \frac{1}{|\det A|}\int f(x)\varphi\Big(A^{-1}(x-b)\Big)\, d^N x\\ &= (f, \varphi_{A,b})\,, \qquad \varphi_{A,b}(x) = \frac{1}{|\det A|}\,\varphi\Big(A^{-1}(x-b)\Big) \qquad (17.4)\end{aligned}$$

The latter equality establishes a relation between the action of $f_{A,b}$ on a test function and the action of f on a test function.

Can a linear change of variables be done in any distribution? To answer this question, let us try to interpret this change of variables as the adjoint transformation of some linear and continuous transformation on the space of test functions. Consider a transformation T of $\mathcal{D}$ into itself defined by

$$T:\ \varphi(x)\ \to\ T(\varphi)(x)=\varphi_{A,b}(x)\,.$$

The transformation is obviously linear. Let $\varphi_n \to 0$ in $\mathcal{D}$. To prove continuity, one has to show that $T(\varphi_n)\to 0$ in $\mathcal{D}$. It is not difficult to see that

$$|T(\varphi_n)(x)| \le \frac{1}{|\det A|}\sup|\varphi_n|\,,$$

from which it follows that $T(\varphi_n)\to 0$ uniformly. Let a_j be the j^{th} column of A^{-1T}. Then by the chain rule and the Schwartz inequality for the dot product

$$|D_jT(\varphi_n)(x)| = \frac{|a_j|}{|\det A|}\sup|\nabla\varphi_n|\,.$$

Therefore $DT(\varphi_n)\to 0$ uniformly. Using a similar idea, one can show that there exists a constant $C_\alpha(A)$ such that (see Exercises)

$$|D^\alpha T(\varphi_n)(x)| \le \frac{C_\alpha(A)}{|\det A|}\max_{|\beta|=|\alpha|}\sup|D^\beta\varphi_n|\,. \tag{17.5}$$

This implies that $\sup|D^\alpha T(\varphi_n)|\to 0$ as $n\to\infty$ for any α, and, hence, T is continuous.

Therefore the rule (**17**.**4**) can be interpreted as the adjoint transformation of any $f\in\mathcal{D}'$ to $\mathcal{D}'$, that is,

$$\Big(f(Ax+b),\varphi(x)\Big) = (T^*(f),\varphi) = (f,T(\varphi)) = (f,\varphi_{A,b})\,. \tag{17.6}$$

By continuity of the adjoint, if $f_n(x)\to f(x)$ in $\mathcal{D}'$, then $f_n(Ax+b)\to f(Ax+b)$ in $\mathcal{D}'$.

17.2.1. Linear change of variables in a delta function. For a general linear change of variables in the delta-function, one infers from (**17**.**6**) that

$$\Big(\delta(Ax+b),\varphi(x)\Big) = \frac{1}{|\det A|}\Big(\delta(x),\varphi(A^{-1}(x-b)\Big) = \frac{\varphi(-A^{-1}b)}{|\det A|}\,.$$

Comparing this relation with the action of the shifted delta-function on a test function it is concluded that

$$\delta(Ax-b) = \frac{1}{|\det A|}\,\delta(x+A^{-1}b)\,.$$

17.2.2. Distributions invariant under linear transformations. In applications one often deals with functions that are invariant under rotations of the arguments, or periodic functions, or similar. This concept can be extended to distributions. *A distribution f is said to be invariant under linear transformation $x \to Ax$ if $f(Ax) = f(x)$.*

Recall that orthogonal transformations in $\mathbb{R}^N$ preserve the quadratic form $|x|^2 = |Ax|^2$. Any such transformation is uniquely defined by an orthogonal matrix A, $A^{-1} = A^T$. For example, the principal value distribution $\mathcal{P}_r \frac{1}{|x|^s}$ is invariant under rotations in $\mathbb{R}^N$. The delta function is also invariant under orthogonal transformations

$$\delta(Ax) = |\det A|^{-1}\delta(x) = \delta(x)$$

because $\det A = \pm 1$ for any orthogonal transformation.

Lorenz transformations preserve the quadratic form $x_0^2 - |x|^2$, where $x_0 \in \mathbb{R}$ and $x \in \mathbb{R}^N$. In special relativity, $x_0 = ct$ where c is the speed of light in the vacuum and t is time, whereas x is a position in space. A distribution is said to be Lorenz invariant if it is invariant under Lorenz transformations of the argument. For example the distribution

$$f(x_0, x) = \frac{1}{x_0^2 - |x|^2 - m^2 + i0^+}$$

is Lorenz invariant for any parameter $m^2 \geq 0$ (see Exercises in the previous section). It will be shown later that this distribution defines the Feynman propagator in the scalar quantum field theory.

17.2.3. Periodic distributions. Let b be a vector in $\mathbb{R}^N$. *A distribution $f(x)$ is said to be periodic in the direction of b if $f(x+b) = f(x)$*, that is

$$\Big(f(x), \varphi(x-b)\Big) = \Big(f(x), \varphi(x)\Big), \quad \varphi \in \mathcal{D}.$$

For example, put

$$f(x) = \sum_{n=-\infty}^{\infty} \delta(x-n), \quad x \in \mathbb{R}.$$

The series converges in the distributional sense because for any test function φ supported in $[-R, R]$, the series

$$(f, \varphi) = \lim_{m\to\infty} \sum_{|n|\leq m} \varphi(n) = \sum_{|n|<R} \varphi(n)$$

is a finite sum and, hence, converges. The distribution f is periodic because $f(x+1) = f(x)$. Indeed,

$$(f(x+1), \varphi(x)) = \sum_n \varphi(n-1) = \sum_m \varphi(m) = (f(x), \varphi(x)),$$

where the shift of summation index has been made, $m = n - 1$.

17.2.4. Parity transformations of distributions. A distribution f is said to be *even* if

$$f(-x) = f(x)$$

and f is called *odd* if

$$f(-x) = -f(x)\,.$$

For example, $\delta(x)$ is even, but $\delta'(x)$ is odd. The latter assertion is proved by

$$\begin{aligned}(\delta'(-x), \varphi(x)) &= (\delta'(x), \varphi(-x)) = -(\delta(x), D_x\varphi(-x)) \\ &= (\delta(x), \varphi'(-x)) = \varphi'(0) = -(\delta'(x), \varphi(x))\end{aligned}$$

that holds for any test function φ. Similarly, the distributions $\mathcal{P}\frac{1}{x}$ and $\mathcal{P}_r\frac{1}{|x|}$ are odd and even, respectively.

Any distribution can written as the sum of even and odd distributions:

$$f(x) = f_+(x) + f_-(x)\,, \quad f_\pm(x) = \frac{1}{2}\Big(f(x) \pm f(-x)\Big)\,,$$

where f_+ and f_- are even and odd distributions, respectively. For example, if f is a Sokhotsky distribution, then, by Sokhotsky's equation, its odd part is $\mathcal{P}\frac{1}{x}$, while the even part is proportional to the delta function.

17.3. Distributions independent of some of the variables. Let $f(x, y)$ be a regular distribution of two variables $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^M$. If $f(x, y) = g(x)$, that is, f is independent of y, then for any test function $\varphi(x, y)$,

$$(f, \varphi) = \int g(x) \int \varphi(x, y)\, d^M y\, d^N y = (g, \psi)\,, \quad \psi(x) = \int \varphi(x, y)\, d^M y\,.$$

Can this rule be extended to all singular distributions? In other words, *a distribution $f(x, y)$ is said to be independent of the variable y if there exists a distribution $g(x)$ such that*

$$(f(x, y), \varphi(x, y)) = \Big(g(x), \int \varphi(x, y)\, d^N y\Big)\,. \tag{17.7}$$

The answer is affirmative because the rule (**17.7**) can be interpreted as the adjoint of some linear and continuous transformation on the space of test functions. Consider a transformation of $\mathcal{D}(\mathbb{R}^{M+N})$ defined by the rule

$$T(\varphi)(x) = \int \varphi(x, y)\, d^M y\,.$$

One has to show that

$$T:\ \mathcal{D}(\mathbb{R}^{M+N}) \to \mathcal{D}(\mathbb{R}^N)$$

and T is linear and continuous. Then the adjoint

$$T^*:\ \mathcal{D}'(\mathbb{R}^N) \to \mathcal{D}'(\mathbb{R}^{N+M})$$

defines the distribution $g = T^*(f)$ in Eq. (**17.7**).

Let us show first that $T(\varphi)$ is a test function. The support of φ is bounded and, hence, it lies in a ball $|x|^2 + |y|^2 < R^2$. Therefore the support of $T(\varphi)$ also lies in a ball $|x| < R$. The integration range in $T(\varphi)$ can be reduced to the ball $|y| < R$ and any constant function is integrable on this ball. Therefore partial derivatives of the integrand with respect to parameters x have integrable bounds independent of x,

$$|D_x^\alpha \varphi(x,y)| \leq \sup |D_x^\alpha \varphi(x,y)| = M_\alpha \in \mathcal{L}(|y| < R)$$

and, by Theorem **7.2**,

$$D^\alpha T(\varphi)(x) = \int D_x^\alpha \varphi(x,y)\, d^M y\,.$$

for any α. So, $T(\varphi) \in \mathcal{D}(\mathbb{R}^N)$.

Let us show that the transformation

$$T:\ \mathcal{D}(\mathbb{R}^{N+M})\ \to\ \mathcal{D}(\mathbb{R}^N)$$

is linear and continuous. The linearity follows from the linearity of the integral. Let $\varphi_n(x,y) \to 0$ in $\mathcal{D}(\mathbb{R}^{N+M})$. Then supports of all φ_n lie in a ball $|x|^2 + |y|^2 < R^2$ so that

$$|D^\alpha T(\varphi_n)(x)| \leq \int_{|y|<R} |D_x^\alpha \varphi_n(x,y)|\, d^M y \leq \sup |D_x^\alpha \varphi_n(x,y)| \int_{|y|<R} d^M y$$

for all x. Therefore

$$\sup |D^\alpha T(\varphi_n)(x)| \leq V_M(R) \sup |D_x^\alpha \varphi_n(x,y)|\,,$$

where $V_M(R)$ is the volume of the integration ball. This inequality shows that the convergence $\varphi_n \to 0$ in $\mathcal{D}(\mathbb{R}^{N+M})$ implies that $T(\varphi_n) \to 0$ in $\mathcal{D}(\mathbb{R}^N)$. The proof is complete

For example, let $f(x,y) = \delta(x)$. The distribution f does not depend on y. In this case

$$\Big(\delta(x), \varphi(x,y)\Big) = \Big(\delta(x), \int \varphi(x,y)\, d^M y\Big) = \int \varphi(0,y)\, d^M y\,.$$

Let $1(y)$ be the unit function (it has unit value everywhere). It defines a regular distribution in the variable y so that

$$T(\varphi)(x) = \Big(1(y), \varphi(x,y)\Big) \in \mathcal{D}(\mathbb{R}^N)\,.$$

Any distribution $g(x)$ can be viewed as a distribution of two variables $f(x,y) = g(x)\cdot 1(y)$ where the product of two distributions of *different* variables is defined by the rule

$$\Big(g(x)\cdot 1(y), \varphi(x,y)\Big) = \Big(g(x), \Big(1(y), \varphi(x,y)\Big)\Big)\,.$$

This product is known as *the direct or tensor product* of distributions (see Sec.**29.1**)

17.4. Exercises.

1. Prove the inequality (**17.5**) and use to show the transformation $\varphi \to \varphi_{A,b}$ is continuous. *Hint*: Show first that if $D^\alpha = D_j^s$ for fixed j and an integer $s \geq 0$, then $C_\alpha(A) = (\sum_i |a_{ji}|)^s$ where a_{ji} is the i^{th} component of the vector a_j.

18. Multiplication by a smooth function

Let f be a regular distribution. Let $a(x)$ be a function such that the product $a(x)f(x)$ is still locally integrable and, hence, defines a regular distribution af. Then

$$(af,\varphi)=\int a(x)f(x)\varphi(x)\,d^Nx=(f,a\varphi)\,,\quad \varphi\in\mathcal{D}(\Omega)\,.$$

Let us extend this rule for multiplication of a regular distribution by a function to all distributions. According to our general principle given in Sec.**17.1**, this operation should be interpreted as the adjoint transformation of some linear and continuous transformation on the space of test functions.

Consider the transformation of $\mathcal{D}(\Omega)$

$$T:\ \varphi(x)\to T(\varphi)(x)=a(x)\varphi(x)\,.$$

The product $a\varphi$ must be a test function for any $\varphi\in\mathcal{D}(\Omega)$. Therefore a must necessarily be from class C^∞. Next, one should verify that T is linear and continuous. The linearity is obvious. Let $\varphi_n\to 0$ in $\mathcal{D}(\Omega)$. One has to show that $T(\varphi_n)\to 0$ in $\mathcal{D}(\Omega)$. The supports of $T(\varphi_n)(x)=a(x)\varphi_n(x)$ lie in a compact set $K\subset\Omega$ if $\operatorname{supp}\varphi_n\subset K$. Since a and all its derivatives are bounded on any compact, put

$$\max_{|\beta|=s}\sup_K|D^\beta a|=M_s<\infty\,.$$

By the product rule

$$\Big|D\Big(a(x)\varphi_n(x)\Big)\Big|\le M_1\sup|\varphi_n|+M_0\sup|D\varphi_n|\,.$$

Since the inequality holds for all x in the left-hand side, one can take the supremum in it:

$$\sup_K\Big|D\Big(a(x)\varphi_n(x)\Big)\Big|\le M_1\sup|\varphi_n|+M_0\sup|D\varphi_n|\,.$$

Therefore $DT(\varphi_n)\to 0$ uniformly on Ω. Using the binomial expansion of high-order partials of the product (see Sec.**1.3**), one infers that

$$\sup|D^\beta(a\varphi_n)|\le\sum_{\alpha\le\beta}\binom{\beta}{\alpha}M_{|\beta|-|\alpha|}\sup|D^\alpha\varphi_n|\,.$$

Therefore $D^\beta T(\varphi_n)\to 0$ uniformly for any β because $D^\alpha\varphi_n\to 0$ uniformly for any α.

Thus, *for any distribution f and any C^∞ function a, the product af is a distribution defined by the rule*

$$(af,\varphi)=(f,a\varphi)\,,\quad \varphi\in\mathcal{D}(\Omega)\,,\quad a\in C^\infty\,. \tag{18.1}$$

18.0.1. Properties of multiplication by a smooth function. It follows from Definition **18.1** that the multiplication of a distribution by a smooth function is linear,

$$a(c_1f_1 + c_2f_1) = c_1af_1 + c_2af_2\,, \quad f_{1,2} \in \mathcal{D}'\,, \quad a \in C^\infty\,, \tag{18.2}$$

where constants $c_{1,2}$ are constants, and it is also distributive and commutative,

$$(ab)f = a(bf) = b(af)\,, \quad f \in \mathcal{D}'\,, \quad a, b \in C^\infty\,. \tag{18.3}$$

18.1. Multiplication of the delta-function and its derivatives. Let us find a distribution obtained by multiplication of the delta-function by a smooth function:

$$(a\delta, \varphi) = (\delta, a\varphi) = a(0)\varphi(0) = a(0)(\delta, \varphi) = \Big(a(0)\delta, \varphi\Big)\,.$$

Since this relation holds for any test function,

$$a(x)\delta(x) = a(0)\delta(x)\,, \quad x \in \mathbb{R}^N\,.$$

In particular,

$$x\delta(x) = 0\,, \quad x \in \mathbb{R}\,.$$

Similarly, one can calculate the product of a smooth function with the derivative of the delta function defined by the rule (**13.5**). For any test function φ

$$(aD\delta, \varphi) = (D\delta, a\varphi) = -(\delta, D(a\varphi)) = -a(0)D\varphi(0) - Da(0)\varphi(0)\,.$$

Since $D\varphi(0) = -(D\delta, \varphi)$, it is concluded that

$$a(x)D\delta(x) = a(0)D\delta(x) - Da(0)\delta(x)\,, \quad x \in \mathbb{R}^N\,. \tag{18.4}$$

Using the binomial expansion of $D^\alpha(a\varphi)$, it is not difficult to show that

$$a(x)D^\alpha\delta(x) = \sum_{\beta\le\alpha}(-1)^{|\beta|}\binom{\alpha}{\beta}D^\beta a(0)D^{\alpha-\beta}\delta(x) \tag{18.5}$$

In particular,

$$x\delta'(x) = -\delta(x)\,, \quad x \in \mathbb{R}\,.$$

and

$$x^n\delta^{(k)} = \sum_{m=0}^{k}(-1)^m\binom{k}{m}D^m(x^n)\Big|_{x=0}\delta^{(k-m)}(x)\,.$$

The derivatives $D^m(x^n)$ vanish at $x = 0$ if $m \le k < n$. When $k \le n$, only the term $m = n$ contributes. Therefore

$$x^n\delta^{(k)}(x) = 0\,, \quad n > k\,, \quad x^n\delta^{(k)}(x) = \frac{(-1)^n k!}{(k-n)!}\,\delta^{(k-n)}(x)\,, \quad n \le k\,.$$

18.2. Multiplication of the principal value distribution. Let us show that

$$x\mathcal{P}\frac{1}{x} = 1\,.$$

For any test function φ,

$$\Big(x\mathcal{P}\frac{1}{x}, \varphi\Big) = \Big(\mathcal{P}\frac{1}{x}, x\varphi\Big) = \lim_{a\to 0}\int_{|x|>a} \varphi(x)\,dx = (1,\varphi)$$

by continuity of the Lebesgue integral. It also follows from Sokhotsky's equations (**16**.**6**) and linearity of multiplication (**18**.**2**) that

$$x\,\frac{1}{x\pm i0^+} = 1$$

because $x\delta(x) = 0$. Furthermore, by the distributive law (**18**.**3**)

$$x^n\mathcal{P}\frac{1}{x} = x^{n-1}\Big(x\mathcal{P}\frac{1}{x}\Big) = x^{n-1}\,,$$
$$x^n\,\frac{1}{x\pm i0^+} = x^{n-1}\Big(x\,\frac{1}{x\pm i0^+}\Big) = x^{n-1}\,.$$

Let a be a C^∞ function. Define the function b by relation $a(x) = a(0) + b(x)x$. Then b is also a C^∞ function by the Taylor theorem. Therefore

$$a(x)\mathcal{P}\frac{1}{x} = a(0)\mathcal{P}\frac{1}{x} + b(x)x\mathcal{P}\frac{1}{x} = a(0)\mathcal{P}\frac{1}{x} + \frac{a(x)-a(0)}{x}\,.$$

18.3. Multiplication by a bump function. It follows from (**16**.**1**) that *if smooth functions $a(x)$ and $b(x)$ are equal in a neighborhood of support of a distribution f, then*

$$a(x)f(x) = b(x)f(x)\,. \qquad (\mathbf{18.6})$$

Indeed, for any test function φ, the supports of f and $(a-b)\varphi$ have no common points and, hence, by (**16**.**1**)

$$(af - bf, \varphi) = ((a-b)f, \varphi) = (f, (a-b)\varphi) = 0\,.$$

It should be pointed out that the assertion may be false if $a = b$ only on $\operatorname{supp} f$. Let us illustrate this with an example.

Let $f = \delta^{(n)}$ (see Sec. **15.4.2**). The support of f consists of the single point $x = 0$ for any n. Let smooth functions a and b be non-vanishing and equal on the support of δ', that is, $a(0) = b(0)$. Then it follows from (**18**.**4**) that

$$a(x)\delta'(x) - b(x)\delta'(x) = -(a'(0) - b'(0))\delta(x) \neq 0\,.$$

unless $a'(0) = b'(0)$. Furthermore, by Eq. (**18.5**)

$$a(x)\delta^{(n)}(x) - b(x)\delta^{(n)} = \sum_{k=0}^{n}(-1)^k \binom{n}{k}[a^{(k)}(0) - b^{(k)}(0)]\delta^{(n-k)}(x)\,,$$

for any integer $n > 0$. This shows that in order for the equality (**18.6**) to hold for any derivative of the delta function, the derivatives of a and b of any order must be equal at $x = 0$. This is guaranteed if $a(x) = b(x)$ in a neighborhood of $x = 0$, not just at $x = 0$. However, it is not necessary because the difference $a(x) - b(x)$ can be a smooth function that vanishes at $x = 0$ and is not analytic at $x = 0$. For example, $\omega_\epsilon(x - \epsilon)\delta^{(n)}(x) = 0$ for any integer $n \geq 0$, where ω_ϵ is the hat function.

Let η_f be a function from class C^∞ that has unit value in a neighborhood of the support of a distribution f. By Theorem **14.1**, a function η_f with the said properties exists and can be constructed by means of the convolution of the hat function ω_a and the characteristic function of $\operatorname{supp} f$ (η_f is a bump function for $\operatorname{supp} f$). Then by setting $a = \eta_f$ and $b = 1$ in (**18.6**), one infers that

$$\eta_f(x) f(x) = f(x)\,. \tag{18.7}$$

This rule for multiplication of a distribution by a bump function for its support will often be used later.

18.4. General solution to $x^n f(x) = 0$. Let us find all solutions to the following distributional equation:

$$x^n f(x) = 0\,, \quad f \in \mathcal{D}'(\mathbb{R})\,. \tag{18.8}$$

It follows from the result of Sec.**18.1** that the equation has a non-trivial solution of the form

$$f(x) = \sum_{k=0}^{n-1} c_k \delta^{(k)}(x)\,, \tag{18.9}$$

for any choice of constants c_k. If the support of a test function φ does not contain $x = 0$, then $\psi(x) = \varphi(x)/x^n$ is also a test function so that for any solution to (**18.8**) $(x^n f, \psi) = (f, x^n\psi) = (f, \varphi) = 0$ which means that $f(x) = 0$, $x \neq 0$. So, the support of f consists of one point $x = 0$. Let us show that *any solution has the form (***18.9***).*

To prove the assertion, the following fact is useful. *For any test function φ, the function*

$$\psi(x) = \frac{\varphi(x) - \eta(x)T_{n-1}(x)}{x^n} \in \mathcal{D}\,,$$

is also a test function, where $T_{n-1}(x)$ *is the Taylor polynomial of order* $n-1$ *for* φ *about* $x=0$ *and* η *is a test function that has unit value in a neighborhood of* $x=0$. Indeed, this function is from C^∞ near any point except possibly $x=0$. By Taylor's theorem $\varphi(x)-\eta(x)T_{n-1}(x) = O(x^n)$ as $x\to 0$ because $\eta(x)=1$ near $x=0$. Therefore ψ is from class C^∞ near $x=0$. It also has a bounded support because η and φ have bounded supports. Thus, $\psi\in\mathcal{D}$.

Let $f(x)$ be a solution to (**18.8**). Then

$$(f,\varphi) = (f,\eta T_{n-1}) + (f, x^n\psi) = (f,\eta T_{n-1}) + (x^n f,\psi) = (f,\eta T_{n-1})$$
$$= \sum_{k=0}^{n-1}\frac{(f,x^k\eta)}{k!}\,\varphi^{(k)}(0) = \sum_{k=0}^{n-1}\frac{(f,x^k\eta)}{k!}\,(-1)^k(\delta^{(k)},\varphi)$$

This shows that

$$c_k = \frac{(-1)^k(f,x^k\eta)}{k!}\,.$$

Note that c_k do not depend on the choice of η because the support of any distributional solution f is $x=0$. So, the action of f on a test function from $\mathcal{D}(\mathbb{R})$ is determined by properties of the test function in a neighborhood of $x=0$ where $\eta(x)=1$. The coefficients c_k are determined by the action of f on a test function that looks like x^k near $x=0$.

18.5. Limits of rapidly oscillating functions multiplied by a distribution. In Sec.**15.8.3** it was shown that a smooth periodic function with period tending to zero converges to the zero distribution. Consider a similar problem when a smooth periodic function is multiplied by a distribution. Here it is proved that

$$\lim_{t\to\infty} e^{itx}\mathcal{P}\frac{1}{x} = i\pi\delta(x)\,. \tag{18.10}$$

and as a consequence of Sokhotsky's equations

$$\lim_{t\to+\infty}\frac{e^{itx}}{x-i0^+} = 2\pi i\delta(x)\,,\quad \lim_{t\to+\infty}\frac{e^{itx}}{x+i0^+} = 0\,. \tag{18.11}$$

For any test function φ with support in $[-R,R]$

$$\Big(e^{itx}\mathcal{P}\frac{1}{x},\varphi\Big) = \Big(\mathcal{P}\frac{1}{x}, e^{itx}\varphi\Big) = \lim_{a\to 0}\int_{a<|x|<R}\frac{e^{itx}\varphi(x)}{x}\,dx$$
$$= \lim_{a\to 0}\left(\int\limits_{a<|x|<R}\frac{e^{itx}[\varphi(x)-\varphi(0)]}{x}\,dx + \varphi(0)\int\limits_{a<|x|<R}\frac{e^{itx}}{x}\,dx\right)$$

The function

$$\psi(x) = \frac{\varphi(x) - \varphi(0)}{x}$$

is from class C^1 because it is smooth for $x \neq 0$ and $\psi(x) = \varphi'(0) + \frac{1}{2}\,\varphi''(0)x + O(x^2)$ near $x = 0$. This implies that a can be set to 0 in the first integral and by integration by parts

$$\int_{-R}^{R} e^{itx}\psi(x)\,dx = \frac{e^{itx}}{it}\,\psi(x)\Big|_{-R}^{R} - \frac{1}{it}\int_{-R}^{R} \psi'(x)e^{itx}\,dx\,.$$

Put

$$M_0 = \sup_{[-R,R]} |\psi(x)|\,, \quad M_1 = \sup_{[-R,R]} |\psi'(x)|\,.$$

The integral can be estimated as

$$\left|\int_{-R}^{R} e^{itx}\psi(x)\,dx\right| \leq \frac{2M_0}{t} + \frac{2M_1R}{t}\,.$$

This shows that the integral vanishes in the limit $t \to \infty$.

The second integral can be evaluated by means of the Cauchy theorem. If C_a and C_R denote semi-circles in the upper part of the complex plane of radii a and R, respectively, and oriented positively, then by the Cauchy theorem applied to the function e^{itz}/z that is analytic in the upper part of the complex plane bounded by the semi-circles it follows that

$$\begin{aligned}\int\limits_{a<|x|<R} \frac{e^{itx}}{x}\,dx &= \int_{C_a} \frac{e^{itz}}{z}\,dz - \int_{C_R} \frac{e^{itz}}{z}\,dz \\ &= i\int_0^{\pi} e^{itae^{i\theta}}\,d\theta - i\int_0^{\pi} e^{itRe^{i\theta}}d\theta\,.\end{aligned}$$

Let us evaluate the limits of these integrals as $a \to 0^+$ and $t \to \infty$. Clearly, $e^{itae^{i\theta}} \to 1$ as $a \to 0$ for any t and θ, and $|e^{itRe^{i\theta}}| = e^{-tR\sin(\theta)} \to 0$ as $t \to \infty$ for any $0 < \theta < \pi$ (or almost everywhere in $[0,\pi]$). Therefore one should justify interchanging the order of integration and taking the limit. The integrand in the first integral has a Lebesgue integrable bound independent of parameters a and t

$$\left|e^{itae^{i\theta}}\right| \leq e^{-ta\sin(\theta)} \leq 1 \in \mathcal{L}(0,\pi)$$

for all $t > 0$ and $a > 0$. By the Lebesgue dominated convergence theorem

$$\lim_{a\to 0^+}\int_0^{\pi} e^{itae^{i\theta}}\,d\theta = \int_0^{\pi} \lim_{a\to 0} e^{iate^{i\theta}}\,d\theta = \int_0^{\pi} d\theta = \pi\,.$$

Similarly, the integrand in the integral over C_R has the same Lebesgue integrable bound for all $t > 0$ and converges to zero almost everywhere so that

$$\lim_{t\to+\infty} \int_0^\pi e^{itRe^{i\theta}} d\theta = \int_0^\pi \lim_{t\to+\infty} e^{itRe^{i\theta}} d\theta = \int_0^\pi 0 \, d\theta = 0 \, .$$

Thus,

$$\lim_{t\to+\infty} \lim_{a\to 0^+} \int\limits_{a<|x|<R} \frac{e^{itx}}{x} \, dx = i\pi$$

and

$$\lim_{t\to+\infty} \left(e^{itx} \mathcal{P}\frac{1}{x}, \varphi \right) = \pi i \varphi(0) = \pi i (\delta, \varphi)$$

for all test functions and (**18.10**) follows.

18.6. Exercises.

1. Show that

$$\text{(i)} \qquad x\mathcal{P}\frac{1}{x^n} = \mathcal{P}\frac{1}{x^{n-1}} \, ,$$

$$\text{(ii)} \qquad a(x)\mathcal{P}\frac{1}{x^n} = \frac{a(x) - p_{n-1}(x)}{x^n} + \sum_{k=0}^{n-1} a^{(k)}(0)\mathcal{P}\frac{1}{x^{n-k}} \, ,$$

where $a \in C^\infty$ and p_{n-1} is the Taylor polynomial of order $n-1$ for a about $x = 0$.

1. Show that

$$x\mathcal{P}_r\frac{1}{|x|} = \varepsilon(x) \, , \quad x^2\mathcal{P}_r\frac{1}{|x|} = |x|$$

where $\varepsilon(x)$ is the sign function.

2. Prove each of the following distributional limits

$$\text{(i)} \qquad \lim_{t\to+\infty} \frac{e^{-itx}}{x - i0^+} = 0$$

$$\text{(ii)} \qquad \lim_{t\to+\infty} \frac{e^{-itx}}{x + i0^+} = -2\pi i \delta(x)$$

$$\text{(iii)} \qquad \lim_{t\to+\infty} \cos(tx) \, \mathcal{P}\frac{1}{x} = 0$$

3. Find all distributions f such that

$$\text{(i)} \qquad x^n f(x) = 1 \, ,$$

$$\text{(ii)} \qquad x^n f(x) = a(x) \, , \quad a \in C^\infty \, ,$$

where n is a positive integer. *Hint*: Any of equations is linear. What is the structure of its general solution? Can a particular solution to (i) be used to construct a particular solution to (ii)?

4. Show that

$$(|x|^2 - m^2)\,\frac{1}{|x|^2 - m^2 + i0^+} = 1\,, \quad x \in \mathbb{R}^N\,.$$

19. Regularization of distributions

It was shown earlier that there are sequences of smooth functions that converge to singular distributions. Can any distribution be obtained as the weak limit of a sequence of smooth functions? It turns out that the answer is affirmative.

THEOREM **19.1**. (Regularization of a distribution)
For any distribution f, there exists a family of test functions f_a, $a > 0$, such that $f_a \to f$ in $\mathcal{D}'$ as $a \to 0^+$. In other words, the space of test functions $\mathcal{D}$ is dense in the space of distributions $\mathcal{D}'$.

This theorem has a paramount significance in physics. Many calculations in physics are carried out *formally*, that is, a little attention, if not at all, is paid to questions like smoothness of functions that are to be differentiated, interchanging the order of summations and differentiation or integration in a series, etc. Nonetheless, such formal calculations lead to correct answers supported by experimental evidence. Why? The regularization theorem for distributions provides a justification for many formal calculations in physics.

First, physical quantities are distributions, not classical functions by the very nature of measuring them. Second, most calculus operations are linear and continuous on the space of distributions as is shown in the next chapter. Therefore any such operation can be carried out for regularizations of distributions (that is, on tests functions) and, by continuity, the result will also be valid after removing the regularization.

19.1. The strategy for constructing a regularization of a distribution. Let us first verify that Theorem **19.1** holds for regular distributions. Recall that the regularization f_a, defined in (**14.6**), for a locally integrable function f is a smooth function. The functions f_a and f are also regular distributions and for any test function φ it follows from Fubini's theorem that

$$\begin{aligned}(\phi_a * f, \varphi) &= \int\int \phi_a(x-y) f(y)\varphi(x)\, d^N y d^N x \\ &= \int f(y) \int \phi_a(x-y)\varphi(x)\, d^N x d^N y = (f, \phi_a^- * \varphi)\end{aligned}$$

where $\phi_a^-(x) = \phi_a(-x)$ is the parity transformation of ϕ_a. Recall also that the characteristic property of ϕ_a is that $\phi_a \to \delta$ in $\mathcal{D}'$. By continuity of the parity transformation on $\mathcal{D}'$ (see Sec.**17.1.1** where T (or T^*) is a parity transformation on $\mathcal{D}$ (or $\mathcal{D}'$)), ϕ_a^- is also a delta sequence,

$\phi_a(-x) \to \delta(-x) = \delta(x)$ in $\mathcal{D}'$. It turns out that the convolution of test functions with such delta sequences has the following property.

PROPOSITION **19.1**. *Let ω_a be a sequence of test functions that converges to the delta function as $a \to 0$ in the distributional sense, then the sequence of test functions $\omega_a * \varphi$ converges to φ in $\mathcal{D}$ as $a \to 0$ for any test function φ.*

A proof will be given shortly. Let us first observe its simple consequence: $\phi_a * f \to f$ in $\mathcal{D}'$ for any regular distribution f. Indeed, by continuity of the functional f

$$\lim_{a\to 0}(\phi_a * f, \varphi) = \lim_{a\to 0}(f, \phi_a^- * \varphi) = (f, \varphi)\,, \quad \varphi \in \mathcal{D}\,.$$

Furthermore, let $\eta(x)$ be a bump function for the unit ball $|x| < 1$. Then $\eta_a(x) = \eta(ax)$ is a bump function for the ball $|x| < \frac{1}{a}$. Therefore $\eta_a(x)f_a(x)$ is a test function and

$$\eta_a f_a \to f \quad \text{in } \mathcal{D}'$$

as $a \to 0^+$. Indeed, for all small enough a and any test function φ, $\eta_a(x)\varphi(x) = \varphi(x)$ because the support of φ is bounded, and, hence

$$\lim_{a\to 0^+}(\eta_a f_a, \varphi) = \lim_{a\to 0^+}(f_a, \eta_a\varphi) = \lim_{a\to 0^+}(f_a, \varphi) = (f, \varphi)\,.$$

This construction proves Theorem **19.1** for regular distributions.

The next step is to extend the above construction to any distribution f. This implies that one has to show:

(i) *for any test function ϕ and any distribution f, the functional $\phi * f$ defined by the rule*

$$(\phi * f, \varphi) = (f, \phi^- * \varphi)\,, \quad \varphi \in \mathcal{D}\,, \tag{19.1}$$

is a distribution (a linear and continuous functional)

(ii) *$\phi * f$ is a regular distribution defined by a smooth function.*

These two assertions will be proved below. It follows from them that the sequence of test functions $f_a = \eta_a(\phi_a * f) \to f$ in $\mathcal{D}'$ as $a \to 0^+$ for any sequence of test functions $\phi_a \to \delta$ in $\mathcal{D}'$ as $a \to 0^+$. Indeed, by continuity of the functional f and Proposition **19.1**, one infers that

$$\begin{aligned}\lim_{a\to 0^+}(f_a, \varphi) &= \lim_{a\to 0^+}(\phi_a * f, \eta_a\varphi) = \lim_{a\to 0^+}(\phi_a * f, \varphi)\\ &= \lim_{a\to 0^+}(f, \phi_a^- * \varphi) = (f, \varphi)\,.\end{aligned}$$

For example, one can always take $\phi_a = \omega_a$ (a hat function) construct an approximation of any distribution by a test function.

Let us turn to establishing the aforementioned facts to fill out the gaps in the proof of Theorem **19.1**.

19.2. Proof of Proposition 19.1. By hypotheses

$$\lim_{a\to 0}(\omega_a,\psi) = \lim_{a\to 0}\int \omega_a(x)\psi(x)\,d^N x = \psi(0)\,, \quad \psi\in\mathcal{D}\,.$$

By Fubini's theorem (which applies as supports of all functions are bounded)

$$\begin{aligned}(\omega_a*\varphi,\psi) &= \int\int \omega_a(x-y)\varphi(y)\,d^N y\,\psi(x)\,d^N x\\ &= \int \omega_a(z)\int \varphi(-y)\psi(z-y)\,d^N y\,d^N x = (\omega_a,\varphi^- * \psi)\,.\end{aligned}$$

Here two changes of variables have been used $y\to -y$ and then $z = x+y$. By Sec.**14.2.4**, the convolution of two test functions is a test function, $\varphi^- * \psi\in\mathcal{D}$. Therefore

$$\lim_{a\to 0}(\omega_a*\varphi,\psi) = (\varphi^- * \psi)(0) = \int \varphi(-y)\psi(-y)\,d^N y = (\varphi,\psi)\,,$$

where the change of variables $y\to -y$ is used again. This means that $\omega_a*\varphi\to\varphi$ in $\mathcal{D}'$ as $a\to 0$.

19.3. Convolution of a distribution and a test function. Fix $\omega\in\mathcal{D}$ and consider the transformation of $\mathcal{D}$ into itself defined by the rule

$$T:\quad \varphi\to T(\varphi) = \omega^- * \varphi\,.$$

If T is linear and continuous, then for any distribution f, the rule (**19.1**) defines the adjoint transformation, $T^*(f) = \omega * f$, of $\mathcal{D}'$ to itself.

Linearity of T follows from linearity of the convolution. By the analysis in Sec.**14.2.4**,

$$D^\alpha(\omega*\varphi) = D^{\alpha-\beta}\omega * D^\beta\varphi\,,\quad 0\le\beta\le\alpha\,. \tag{19.2}$$

Let $\varphi_n\to 0$ in $\mathcal{D}$. Then by the above property

$$\begin{aligned}|D^\alpha(\omega^- * \varphi_n)(x)| &\le \int |\omega^-(x-y)||D^\alpha_y\varphi_n(y)|\,d^N y\\ &\le M\sup|D^\alpha\varphi_n|\,,\quad M=\int|\omega(y)|\,d^N y\,,\end{aligned}$$

Since the inequality holds for any x, $\sup|D^\alpha T(\varphi_n)|\le M\sup|D^\alpha\varphi_n|\to 0$ as $n\to\infty$ for any α. This means that $T(\varphi_n)\to 0$ in $\mathcal{D}$. Continuity of T is established. Thus, the rule (**19.1**) defines a distribution $\omega * f$.

Let us investigate if $\omega * f$ is a regular distribution defined by a smooth function for any $f\in\mathcal{D}'$. This is far from obvious because for singular distributions basic theorems about smoothness of functions defined by integrals like (**14.6**) cannot be used. Let us first check if the assertion holds for some examples of singular distributions.

19.3.1. Convolutions with delta functions. Let is calculate $\omega*\delta$. For any test function, one infers that

$$(\omega*\delta,\varphi)=(\delta,\omega_-*\varphi)=(\omega_-*\varphi)(0)$$
$$=\int\omega(x)\varphi(x)\,d^Nx=(\omega,\varphi)$$

Therefore

$$\omega*\delta=\omega\,.$$

The convolution is a test function.

Let $f=\nu\delta_S$ be a surface delta function with density ν and supported on a smooth M-surface S in $\mathbb{R}^N$. It is defined in (**15.6**). Let us calculate its convolution with a test function:

$$(\omega*(\nu\delta_S),\varphi)=(\nu\delta_S,\omega_-*\varphi)=\int_S\nu(x)\int\omega(y-x)\varphi(y)\,d^Ny\,dS_x$$
$$=\int\varphi(y)\int_S\nu(x)\omega(y-x)\,dS_x\,d^Ny$$

where the order of integration is changed by Fubini's theorem which applies because the integrand is an integrable function on $\mathbb{R}^N\times S$. Indeed, if φ is supported in a ball $|y|<R$ and ω is supported in a ball $|y-x|<R_\omega$. Then

$$\int\int_S|\nu(x)\omega(y-x)\varphi(y)|\,dS_x\,d^Ny\le M\int_{S_R}|\nu(x)|dS_x<\infty\,,$$

where $M=\sup|\varphi|\sup|\omega|V(R)$, $V(R)$ is the volume of a ball of radius R, S_R is the part of S that lies in a ball of radius $R+R_\omega$, and the surface integral is finite because the density ν is continuous on S and S_R has a finite area (see Sec.**15.6**). Therefore, the convolution in question is given by the surface integral

$$\omega*(\nu\delta_S)(y)=\int_S\nu(x)\omega(y-x)\,dS_x\,.$$

This function is smooth in any ball $|y|<R$, hence, from class C^∞. Indeed, if support of ω is in a ball of radius R_ω, then the integrand vanishes for any $|x|>R+R_\omega$ if $|y|<R$ so that the partials are bounded: $|\nu(x)D^\beta_y\omega(y-x)|\le\sup|D^\beta\omega||\nu(x)|$, the bound is independent of y and is integrable on the part of S in the ball $|x|<R+R_\omega$ as noted above. By Theorem **7.2**, the surface integral has continuous partials of any order in any ball $|y|<R$. So, the convolution $\omega*(\nu\delta_S)$ is a smooth function.

In what follows, the assertion will be deduced from a more fact proved in the next section. This fact is also essential for several other concepts in the theory of distributions such as, e.g., convolution.

19.4. Test functions generated by distributions. Let $\varphi(x, y)$ be a test function of two variables $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^M$. For every fixed y, $\varphi(x, y)$ is a test function in the variable x. Therefore one can define a function

$$\psi(y) = \Big(f(x), \varphi(x, y)\Big)\,, \tag{19.3}$$

where the value of the distribution f is calculated for every (fixed) y. For example, if f is a regular distribution, then

$$\psi(y) = \int f(x)\varphi(x, y)\, d^N x\,.$$

If $f(x) = D\delta(x)$, then

$$\psi(y) = (D\delta(x), \varphi(x, y)) = -(\delta(x), D_x\varphi(y, x)) = -D_x\varphi(y, 0)\,.$$

The function (**19.3**) has remarkable properties.

PROPOSITION **19.2**. *For any distribution $f \in \mathcal{D}'(\mathbb{R}^N)$ and any test function of two variables, $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^M$, the function defined by Eq. (***19.3***) is a test function, and*

$$\int \Big(f(x), \varphi(x, y)\Big)\, d^M y = \Big(f(x), \int \varphi(x, y)\, d^M y\Big) \tag{19.4}$$

$$D_y^\alpha\Big(f(x), \varphi(x, y)\Big) = \Big(f(x), D_y^\alpha\varphi(x, y)\Big) \tag{19.5}$$

Support of ψ. Let support of φ be in a ball $|x|^2 + |y|^2 < R^2$, then $\psi(y) = 0$ if $|y| > R$ because $\varphi(x, y) = 0$ for all x if $|y| > R$. Thus, the support of ψ is bounded.

Continuity of ψ. Take a sequence $y_n \to y$. Then the sequence of test functions $\varphi_n(x) = \varphi(x, y_n)$ converges to $\varphi(x, y)$ in $\mathcal{D}(\mathbb{R}^N)$ for every y. Indeed, supports of $D_x^\alpha\varphi_n$ lie in a ball $|x| \leq R$. By Sec.**1.2.6**, φ is uniformly continuous and, hence, for any $\varepsilon > 0$ one can find $\delta > 0$ such that

$$|D_x^\alpha\varphi_n(x) - D_x^\alpha\varphi(x, y)| < \varepsilon \quad \text{whenever} \quad |y_n - y| < \delta$$

which holds for all x. Therefore

$$\sup_x |D_x^\alpha\varphi_n(x) - D_x^\alpha\varphi(x, y)| \leq \varepsilon$$

for any y. Since $y_n \to y$, for all large enough n the distance between y_n and y can be made smaller than δ. This implies that $D_x^\alpha \varphi_n \to D_x^\alpha \varphi$ uniformly in the variable x. This means that $\varphi_n \to \varphi$ in $\mathcal{D}(\mathbb{R}^N)$. Continuity of ψ follows from the continuity of the functional f:

$$\lim_{n\to\infty} \psi(y_n) = \lim_{n\to\infty} \Big(f(x), \varphi_n(x)\Big) = \Big(f(x), \varphi(x,y)\Big) = \psi(y)\,.$$

ψ is a test function. Let e_j denote the j^{th} unit vector in the standard basis in $\mathbb{R}^M$. Then by definition

$$\frac{\partial \psi(y)}{\partial y_j} = \lim_{\delta\to 0} \frac{\psi(y+\delta e_j) - \psi(y)}{\delta}$$

Then for every (fixed) y, the test functions

$$\phi_\delta(x) = \frac{\varphi(x, y+\delta e_j) - \varphi(x,y)}{\delta} \to \frac{\partial \varphi(x,y)}{\partial y_j} \quad \text{in } \mathcal{D}(\mathbb{R}^N)$$

as $\delta \to 0$. A proof of this assertion is analogous to the proof of continuity of ψ and left to the reader as an exercise. Then the existence of partial derivatives of ψ follows from the continuity of the functional f:

$$\frac{\partial \psi(y)}{\partial y_j} = \lim_{\delta\to 0} \Big(f(x), \phi_\delta(x)\Big) = \Big(f(x), \frac{\partial \varphi(x,y)}{\partial y_j}\Big)$$

Next continuity of partial derivatives $D\psi$ is established in the same way as the continuity of ψ. Thus, ψ is from class C^1. Repeating the argument for partial derivatives of partial derivatives, $D^2\psi = D(D\psi)$, and so on for $D^\beta \psi$, it is concluded that ψ is from class C^∞ and Eq. (**19.5**) holds for computing any partial derivative of ψ.

Integration of ψ. Consider a sequence of Riemann sums for the integral of φ with respect to y

$$\phi_n(x) = \sum_{p\in P_n} \varphi(x, y_p)\Delta V_p$$

where it is assumed that each partition box R_p lies in a ball of radius $1/n$, $n = 1, 2, \ldots$. Let us show that

$$\phi_n(x) \to \phi(x) = \int \varphi(x,y)\, d^M y \quad \text{in} \quad \mathcal{D}(\mathbb{R}^N)\,.$$

The idea is again based on the uniform continuity of test functions. Fix $\varepsilon > 0$ and find δ such that

$$|D_x^\alpha \varphi(x,y) - D_x^\alpha \varphi(x, y_p)| < \varepsilon \quad \text{whenever} \quad |y = y_p| < \delta\,.$$

By the integral mean value theorem

$$D^\alpha \phi(x) = \sum_{p\in P_n} \int_{R_p} D_x^\alpha \varphi(x,y)\, d^M y = \sum_{p\in P_n} D_x^\alpha \varphi(x, y_p^*)\Delta V_p$$

for some points $y_p^* \in R_p$ that generally depend on α and x. Note that φ vanishes outside some large enough box and the sum has finitely many terms. Therefore for all n such that $\frac{1}{n} < \delta$,

$$|D^\alpha \phi_n(x) - D^\alpha \phi(x)| \le \sum_{p\in P_n} |D_x^\alpha \varphi(x, y_p^*) - D_x^\alpha \varphi(x, y_p)|\Delta V_p < \varepsilon V_R$$

where V_R is the volume of a rectangular box that contains the ball $|y| < R$ if the support of φ is in a ball of radius R. The inequality holds for all x and, hence,

$$\sup |D^\alpha \phi_n(x) - D^\alpha \phi(x)| \le \varepsilon$$

which means that all partial derivatives of ϕ_n converge uniformly to the corresponding partial derivatives of ϕ, or $\phi_n \to \phi$ in $\mathcal{D}$.

Equation (**19.4**) follows from continuity and linearity of the functional f and integrability of ψ:

$$\Big(f(x), \phi(x)\Big) = \lim_{n\to\infty} \Big(f(x), \phi_n(x)\Big) = \lim_{n\to\infty} \sum_{p\in P_n} \Big(f(x), \varphi(x, y_p)\Big)\Delta V_p$$
$$= \lim_{n\to\infty} \sum_{p\in P_n} \psi(y_p)\Delta V_p = \int \psi(y)\, d^M y\,.$$

19.5. Smoothness of the convolution. If f is regular distribution, then the convolution $(\omega * f)(y) = (f(x), \omega(x-y))$ where $\omega(x-y)$ is a test function in the variable x for each y. It turns out that this representation can be extended to singular distributions and, with the help of Proposition **19.2**, this function is proved to be smooth.

PROPOSITION **19.3**. *The convolution of a test function and a distribution is a smooth function:*

$$f \in \mathcal{D}'\,, \quad \omega \in \mathcal{D} \quad \Rightarrow \quad \omega * f \in C^\infty$$

that can be computed by the rule

$$(\mathbf{19.6}) \qquad (\omega * f)(y) = \Big(f(x), \omega(y-x)\Big)$$

and its derivatives are

$$(\mathbf{19.7}) \qquad D^\beta(\omega * f)(y) = \Big(f(x), D_y^\beta \omega(y-x)\Big)$$

If, in addition, the support of f is bounded, then the convolution is a test function:

$$f \in \mathcal{D}', \quad \operatorname{supp} f \subset B_R, \quad \omega \in \mathcal{D} \quad \Rightarrow \quad \omega * f \in \mathcal{D}.$$

Let ϕ be a test function. Then

$$\begin{aligned}(\omega * f, \phi) &\overset{(1)}{=} (f, \omega_- * \phi) \overset{(2)}{=} \Big(f(x), \int \omega(y-x)\phi(y)\, d^N y\Big)\\ &\overset{(3)}{=} \int \Big(f(x), \omega(y-x)\phi(y)\Big)\, d^N y\\ &\overset{(4)}{=} \int \Big(f(x), \omega(y-x)\Big)\phi(y)\, d^N y\end{aligned}$$

Here (1) and (2) are by definition of the convolutions, (3) follows from (**19.4**) where

$$\varphi(x,y) = \omega(y-x)\phi(y) \in \mathcal{D}(\mathbb{R}^{2N})$$

for any test functions ω and ϕ, and (4) by linearity of the functional f. The rule (**19.6**) follows from the last equality. By Proposition **19.2**, $(f(x), \varphi(x,y)) = \phi(y)(\omega * f)(y)$ is a test function. This implies that $\omega * f$ is from class C^∞, and partial derivatives $D^\beta(\omega * f)$ are given by (**19.7**).

Suppose that the support of f is bounded. Let η_f be a bump function for $\operatorname{supp} f$. Then η_f is a test function and $\eta_f(x)f(x) = f(x)$ (see (**18.7**)). It follows from (**19.6**) that

$$(\omega * f)(y) = \Big(\eta_f(x)f(x), \omega(x-y)\Big) = \Big(f(x), \eta_f(x)\omega(x-y)\Big)$$

The test function $\eta_f(x)\omega(x-y) = 0$ vanishes for all $|y| > R_f + R_\omega$ if supports of f and ω are balls of radii R_f and R_ω respectively. This implies that the support of $\omega * f$ lies in a ball of radius $R_f + R_\omega$ so that $\omega * f \in \mathcal{D}$. The proof of Proposition **19.6** is complete and so is the proof of Theorem **19.1**.

19.6. Summary. The proof of Theorem **19.1** offers an explicit method for constructing an approximation of any distribution by a regular distribution defined by either a test function or a smooth function. For example, for any distribution f, the convolution $\omega_a * f$, where ω_a is a hat function, is a regular distribution defined by a C^∞ function, and $\omega_a * f \to f$ in $\mathcal{D}'$ as $a \to 0^+$. If the distribution f is compactly supported, then $\omega_a * f$ is a test function. If f is not compactly supported, then $f_a = \eta_a(\omega_a * f)$ is a test function for any bump function η_a for the ball $|x| < \frac{1}{a}$, and $f_a \to f$ in $\mathcal{D}'$ as $a \to 0^+$.