CHAPTER 3

Calculus with distributions

20. Change of variables in distributions

20.1. A general change of variables in a distribution. Let us try to extend a linear charge of variables in a distribution to a general change of variables. Let $x = F(y)$ be a change of variables in a Lebesgue integral (as defined in Sec.**6.14**):

$$F:\ \Omega \subseteq \mathbb{R}^N \to \mathbb{R}^N\,, \quad \Omega_F = F(\Omega)\,,$$

for some open set Ω. The strategy is to find the associated transformation transformation T_F of the space of test functions into itself using a change of variables in regular distributions and then to analyze conditions under which T_F is linear and continuous. Then the adjoint T_F^* will map $\mathcal{D}'$ to a space of distributions.

Let $f(x)$ be locally integrable on Ω_F. Then f defines a regular distribution from $\mathcal{D}'(\Omega_F)$. Put

$$f_F(y) = f(F(y))\,.$$

The function f_F is locally integrable on Ω and defines a regular distribution from $\mathcal{D}'(\Omega)$ by the rule

$$(f_F, \varphi) = \int_\Omega f(F(y))\varphi(y)\, d^N y\,, \quad \varphi \in \mathcal{D}(\Omega)\,.$$

Let us change the integration variable $y = F^{-1}(x)$,

$$d^N x = J(y)\, d^N y\,, \qquad J(y) = \left|\det\left(\frac{\partial F_j}{\partial y_k}\right)\right| \neq 0\,,$$

where the Jacobian $J(y)$ does not vanish anywhere in Ω. Therefore

$$(f_F, \varphi) = \int_{\Omega_F} f(x)\varphi_F(x)\, d^N x = (f, \varphi_F)\,,$$

where

$$\varphi_F(x) = \frac{1}{J(F^{-1}(x))}\,\varphi(F^{-1}(x))\,. \qquad (\mathbf{20}.1)$$

If $F(y) = Ay + b$, then this equation is reduced to (**17**.4).

Define a transformation on the space of test functions

$$T_F:\quad \varphi(x) \to T_F(\varphi)(x) = \varphi_F(x)\,.$$

Then, if φ_F is a test function from $\mathcal{D}(\Omega_F)$ and T_F is linear and continuous, the adjoint transformation $T_F^* : \mathcal{D}'(\Omega_F) \to \mathcal{D}'(\Omega)$ is defined by the rule

$$(T_F^*(f), \varphi) = (f, T_F(\varphi)), \quad f \in \mathcal{D}'(\Omega_F), \quad \varphi \in \mathcal{D}(\Omega) \tag{20.2}$$

In what follows,

$$T_F^*(f)(y) = f(F(y))$$

for brevity, just like for regular distributions.

It is noted first that the function φ_F, defined in (**20.1**), is not smooth enough to be a test function because F^{-1} is from class C^1. So the change of variables must be from class C^∞. The following fact[1] from mathematical analysis is invoked to find sufficient conditions for the outlined strategy to work.

PROPOSITION **20.4**. *Let Ω be open in $\mathbb{R}^N$ and $F : \Omega \to \mathbb{R}^N$ be a one-to-one transformation from class C^∞ whose Jacobian does not vanish anywhere on Ω. Then the inverse transformation $F^{-1} : F(\Omega) \to \Omega$ is from class C^∞ that is one-to-one and whose Jacobian does not vanish anywhere in $F(\Omega)$. The converse is also true.*

A transformation F satisfying the hypotheses in Proposition **20.4** is called a *diffeomorphism* of Ω onto $F(\Omega)$. The inverse of a diffeomorphism is also a diffeomorphism.

COROLLARY **20.1**. *Let F be a diffeomorphism of an open set $\Omega \subseteq \mathbb{R}^N$ onto $\Omega_F = F(\Omega)$. Then the transformation T_F defined by the rule (**20.1**) maps $\mathcal{D}(\Omega)$ to $\mathcal{D}(\Omega_F)$, and is linear and continuous. The adjoint transformation of any distribution $f \in \mathcal{D}'(\Omega_F)$ defined by the rule (**20.2**) is a distribution from $\mathcal{D}'(\Omega)$.*

To prove the assertion, it is sufficient to show that T_F exists and is linear and continuous. The conclusion follows from Sec.**17.1**. The existence is established by showing that the function φ_F defined by (**20.1**) is a test function from $\mathcal{D}(\Omega_F)$ for any $\varphi \in \mathcal{D}(\Omega)$. By Proposition **20.4**, φ_F is from class C^∞ for any $\varphi \in C^\infty$ because F^{-1} is from class C^∞ and the Jacobian does not vanish anywhere. Now recall a basic fact from mathematical analysis[2] that *a continuous transformation of $\mathbb{R}^N$ maps a compact set into a compact set.* A test function φ is compactly supported and so is φ_F:

$$\operatorname{supp} \varphi_F = F(\operatorname{supp} \varphi) \subset \Omega_F .$$

[1]R.E. Kass and P.W. Vos, Geometrical foundations of asymptotic interference, John Willey & Sons, 1997, Appendix A

[2]See, e.g., W. Rudin, Principles of mathematical analysis, Chapter 4.

Thus, $\varphi_F \in \mathcal{D}(\Omega_F)$.

The map T_F is obviously linear. Let $\varphi_n \to 0$ in $\mathcal{D}(\Omega)$ as $n \to \infty$. Then supports of all φ_n lie in one compact $K \subset \Omega$. Therefore the supports of $T_F(\varphi_n)$ are in the compact $F(K)$. It remains to show that $\sup_\Omega |D^\alpha \varphi_n| \to 0$ for any $\alpha \geq 0$ implies that $\sup_{\Omega_F} |D^\alpha T_F(\varphi_n)| \to 0$ for all $\alpha \geq 0$. The reciprocal of $J(F^{-1}(x))$ is from class C^∞ and multiplication by a smooth function is a linear and continuous transformation of the space of test functions (see Sec.**18**), it is then sufficient to establish uniform convergence for derivatives of $\varphi_n(F^{-1}(x))$. By the chain rule

$$\sup_{F(K)} |D_x \varphi_n(F^{-1}(x))| \leq M_1 \sup_K |D_y \varphi_n(y)|\,, \quad M_1 = \sum_{i,j=1}^N \sup_{F(K)} \left|\frac{\partial y_i}{\partial x_j}\right|$$

because all partial derivatives of $y = F^{-1}(x)$ are bounded on any compact subset of Ω_F as $F^{-1} \in C^\infty(\Omega_F)$. This shows that $D_x\varphi_n(F^{-1}(x)) \to 0$ uniformly if $D_y \varphi_n(y) \to 0$ uniformly. Similarly, for other partial derivatives. Technical details are left to the reader as an exercise.

20.1.1. Change of variables in delta functions. Let $x \in \mathbb{R}$ and F be a diffeomorphism such that

$$F:\ \mathbb{R} \to F(\mathbb{R}) = \mathbb{R}\,.$$

Since the derivative F' does not vanish anywhere, it is either strictly positive or strictly negative everywhere. Therefore the function F is strictly monotonic (either increasing or decreasing monotonically) and, hence, one-to-one. So, $x = F(y)$ is a change of variables. In this case

$$\varphi_F(x) = \frac{\varphi(y)}{|F'(y)|}\Big|_{y=F^{-1}(x)}\,.$$

Note that if $\operatorname{supp}\varphi \subset [-R, R]$, then the support of φ_F lies in $[A_-, A_+]$ where $F(A_\pm) = \pm R$ or $F_\pm = F^{-1}(\pm R)$. Owing to the monotonicity of F, the numbers $A_\pm$ are unique. The absolute value does not produce any problems with differentiability because F' never changes its sign so that $\varphi_F \in C^\infty$ if $F^{-1} \in C^\infty$.

Let $\delta_a(x) = \delta(x - a)$ be a shifted delta-function. Then

$$\Big(\delta_a(F(y)), \varphi(y)\Big) = (\delta_a, \varphi_F) = \varphi_F(a) = \frac{\varphi(b)}{|F'(b)|} = \frac{1}{|F'(b)|}\Big(\delta_b(y), \varphi(y)\Big)$$

where b is the root of the equation $F(b) = a$ or $b = F^{-1}(a)$. Owing to the monotonicity of F, there is only one such root for any $a \in \mathbb{R} = F(\mathbb{R})$. Thus,

$$\delta\Big(F(y) - a\Big) = \frac{1}{|F'(b)|}\,\delta(y - b)\,, \quad b = T^{-1}(a)\,.$$

For example, put $F(y) = \sinh(y)$, then

$$y = F^{-1}(x) = \ln\Big(x + \sqrt{x^2+1}\Big) \in C^\infty, \quad x \in \mathbb{R},$$
$$\varphi_F(x) = \frac{1}{\sqrt{1+x^2}}\,\varphi\Big(\ln\Big(x+\sqrt{x^2+1}\Big)\Big) \in \mathcal{D}(\mathbb{R}),$$
$$\delta(\sinh(y) - a) = \frac{1}{\sqrt{1+a^2}}\,\delta(y-b), \quad b = \ln(a+\sqrt{a^2+1}).$$

Let F be a diffeomorphism of $\mathbb{R}^N$ onto $\mathbb{R}^N$. Then for any test function $\varphi \in \mathcal{D}$,

$$\Big(\delta(F(y)-a)), \varphi(y)\Big) = (\delta(x-a), \varphi_F(x)) = \varphi_F(a) = \frac{\varphi(F^{-1}(a))}{|J(F^{-1}(a))|}.$$

Therefore

$$\delta\Big(F(y) - a\Big) = \frac{1}{|J(F^{-1}(a))|}\,\delta\Big(y - F^{-1}(a)\Big).$$

20.2. Transformations with a reduced range. Consider a diffeomorphism of $\mathbb{R}^N$ whose range is a subset of $\mathbb{R}^N$:

$$F:\ \mathbb{R}^N \ \to\ \Omega_F \subset \mathbb{R}^N.$$

A peculiarity of this case is that T_F maps $\mathcal{D}$ into its subspace $\mathcal{D}(\Omega_F)$. Therefore the adjoint transformation

$$\Big(f(F(y)), \varphi(y)\Big) = \Big(f(x), \varphi_F(x)\Big), \quad \varphi_F \in \mathcal{D}(\Omega_F),$$

applies only to distributions $f \in \mathcal{D}'(\Omega_F) \subset \mathcal{D}'$, not to every distribution in $\mathcal{D}'$. If one formally extends the above relation to any $f \in \mathcal{D}'$, then the right-hand side vanishes if $f = 0$ in Ω_F. This implies that if the support of $f \in \mathcal{D}'$ does not overlap with Ω_T, then $f(T(y)) = 0$:

$$f \in \mathcal{D}', \quad \operatorname{supp} f \cap F(\mathbb{R}^N) = \emptyset \quad \Rightarrow \quad f(F(y)) = 0.$$

This observation naturally agrees with a change of variable for regular distributions

$$(f(F(y)), \varphi(y)) = \int f(F(y))\varphi(y)\,d^N y = \int_{\Omega_F} f(x)\varphi_F(x)\,d^N x,$$

The integral vanishes if $f(x) = 0$ a.e. in Ω_F even though $f(x) \neq 0$ in $\mathbb{R}^N$.

20.2.1. Example: Shifted delta functions. Let $x = F(y)$ be diffeomorphism of $\mathbb{R}$ as in Sec.**20.1.1** but its range is not the whole $\mathbb{R}$. Then the equation $F(b) = a$ has no root if a does not belong to the range of F. In this case,

$$\delta(F(y) - a) = 0\,, \quad a \notin F(\mathbb{R})\,.$$

For example,

$$\begin{aligned} x &= F(y) = \arctan(y)\,, \quad F'(y) = \frac{1}{1+y^2} > 0\,, \quad y \in \mathbb{R}\,, \\ y &= F^{-1}(x) = \tan(x)\,, \quad x \in \Omega_F = F(\mathbb{R}) = (\tfrac{\pi}{2}, \tfrac{\pi}{2})\,, \\ \varphi_F(x) &= \frac{1}{\cos^2(x)}\,\varphi\Big(\tan(x)\Big) \in \mathcal{D}(\tfrac{\pi}{2}, \tfrac{\pi}{2})\,. \end{aligned}$$

The test function φ_F and all its derivatives vanish in a neighborhood of the points $x = \pm\frac{\pi}{2}$ because φ has a bounded support and, hence, vanishes for all $|\tan(x)| > R$ and some $R > 0$ (see Sec.**14.1.1**). In this case,

$$\begin{aligned} \delta(\arctan(y) - a) &= 0\,, \quad |a| \geq \frac{\pi}{2}\,, \\ \delta(\arctan(y) - a) &= \frac{1}{1+a^2}\,\delta(y - \tan(a))\,, \quad |a| < \frac{\pi}{2}\,. \end{aligned}$$

The same holds for a multi-variable delta function, $\delta(F(y) - a) = 0$ if a is not in the range of a diffeomorphism F of $\mathbb{R}^N$ into itself.

20.3. A general smooth transformation of the argument of a distribution. Let $x = F(y)$ be a general transformation of $\mathbb{R}^N$ from class C^∞. What can go wrong with the rule (**20.2**)? Can the distribution $f(F(y))$ be defined in some sense? It turns out that if a smooth transformation is not a diffeomorphism, then the change of arguments can still be done for some distributions.

20.3.1. Zeros of the Jacobian. The Jacobian of a smooth one-to-one transformation, $x = F(y)$, can have zeros. For example, put $x = F(y) = y^3$ so that $dx = 3y^2 dy$ and the inverse is $x = y^{1/3}$. Note that the inverse transformation is not smooth at $y = 0$. The Jacobian $J = 3y^2$ also vanishes at $y = 0$. Then the rule (**20.2**) does not make any sense for the delta-function because

$$\Big(\delta(y^3), \varphi(y)\Big) = \Big(\delta(x), \varphi_F(x)\Big) = \frac{\varphi(x^{1/3})}{3x^{2/3}}\Big|_{x=0} \text{ does not exist}$$

for any test function $\varphi(y)$ that does not vanish at $y = 0$.

Clearly, the reason for the failure of (**20.2**) is that the transformation F does not induce a map of $\mathcal{D}(\Omega_F)$ to $\mathcal{D}(\Omega)$. The rule (**20.1**) fails to produce a test function. However, if one takes any open subset of Ω where the Jacobian has no zeros, then the transformation *restricted* to this subset satisfies the hypotheses of Proposition **20.4**. Therefore the change of variables (**20.2**) in any distribution whose support lies in the largest open set that does not contain zeros of the Jacobian should work just fine.

Let us formalize this observation. Suppose that the support of a distribution $f(x)$ lies in an open set that does not contain zeros of the Jacobian. This means that the distance between $\operatorname{supp} f$ and the set of zeros of the Jacobian is not zero. Therefore there exists a function $\eta(x)$ from class C^∞ such that $\eta(x) = 1$ in a neighborhood of support of a distribution f, η vanishes in a neighborhood of the set of zeros of the Jacobian (see Sec.**18.3**), and $f(x) = \eta(x)f(x)$. Then the rule (**20.2**) yields

$$\Big(f(F(y)), \varphi(y)\Big) = \Big(\eta(x)f(x), \varphi_{F}(x)\Big) = \Big(f(x), \eta(x)\varphi_{F}(x)\Big) .$$

The function $\eta(x)\varphi_{F}(x)$ is a test function and the rule makes sense and defines a linear continuous functional.

For example, let $x = F(y) = y^3$ where x and y are real. Take a shifted Dirac delta function, $f(x) = \delta(x-a)$, $a \neq 0$. The support of this distribution is $x = a$. Let $\eta(x)$ be a bump function for the point set $\{x = a\}$ that vanishes for $|x-a| > \delta$ and some $\delta < |a|$. Then

$$\eta(x)\delta(x-a) = \eta(a)\delta(x-a) = \delta(x-a) ,$$

$$\varphi_{F}(x) = \frac{\eta(x)}{2x^{2/3}}\,\varphi(x^{1/3}) \in \mathcal{D}$$

because the singular point of the transformation $x = 0$ lies outside the support of φ_{F}. Therefore

$$\Big(\delta(y^3-a), \varphi(y)\Big) = \Big(\delta(x-a), \varphi_{T}(x)\Big) = \varphi_{F}(a) = \frac{1}{2a^{2/3}}\varphi(a^{1/3})$$

for any test function $\varphi(y)$. Thus,

$$\delta(y^3-a) = \frac{1}{2a^{2/3}}\,\delta(y-a^{1/3}) , \quad a \neq 0 .$$

20.3.2. Smooth transformations that are not one-to-one. Let F be a C^∞ transformation of $\mathbb{R}^N$ into itself. Let $\Omega_J \subset F(\mathbb{R}^N)$ be the largest open subset where the Jacobian is not zero. The construction from the previous section needs additional amendments because F is not one-to-one. The idea is to identify all largest open sets in the domain of F such that when F is restricted on any of them, say Ω_n, it becomes a diffeomorphism of Ω_n onto Ω_J so that a change of variables can be carried out for any distribution supported in Ω_n.

For example, consider $F:\ \mathbb{R}\to\mathbb{R}$ defined by a smooth function:

$$x = F(y) = \sin(y)\,.$$

Then $\Omega_J = (-1,1)$ and

$$\Omega_n = \Big(-\frac{\pi}{2}+\pi n, \frac{\pi}{2}+\pi n\Big)\,, \quad F(\Omega_n)=\Omega_J\,,$$

where n is any integer. Then the rule (**20.2**) can be used for any distribution supported in Ω_J if one figures out how to deal with the fact that F^{-1} has now many branches, like $\arcsin(x)$, in the rule (**20.1**). Let us investigate this.

The following assumptions are made about properties of a transformation F:

(i) *There are at most countably many open sets Ω_n that are not intersecting and F is a diffeomorphism of Ω_n onto Ω_J which is an open set in $F(\mathbb{R}^N)$.*

(ii) *Any ball intersects only finitely many sets Ω_n.*

Put

$$F_n^{-1}:\ \Omega_J \to \Omega_n\,.$$

So, the functions $y = F_n^{-1}(x)$ are different branches of the inverse, like the branches of the arcsine function. They are solutions to the equation

$$(\mathbf{20.3})\qquad F(y) = x\in\Omega_J \quad\Rightarrow\quad y = F_n^{-1}(x)\in\Omega_n\,.$$

Let $f(x)$ be a regular distribution with support in Ω_J and η is a bump function for a neighborhood of $\operatorname{supp} f$ so that $\eta(x)f(x)=f(x)$ and η vanishes in a neighborhood of all zeros of the Jacobian. Then $\eta(F(y))$ vanishes in a neighborhood of the complement of the union of Ω_n, and the following equalities hold for any test function $\varphi\in\mathcal{D}$:

$$(\mathbf{20.4})\qquad \begin{aligned}\Big(f(F(y)),\varphi(y)\Big) &= \Big(\eta(F(y))f(F(y)),\varphi(y)\Big)\\ &= \sum_n\int_{\Omega_n}\eta(F(y))f(F(y))\varphi(y)\,d^Ny\\ &= \sum_n\int_{\Omega_J} f(x)\eta(x)\frac{\varphi(F_n^{-1}(x))}{J(F_n^{-1}(x))}\,d^Nx\\ &= \int_{\Omega_J} f(x)\varphi_F(x)\,d^Nx = (f,\varphi_F)\,,\end{aligned}$$

$$(\mathbf{20.5})\qquad \varphi_F(x) = \eta(x)\sum_n\frac{\varphi(F_n^{-1}(x))}{J(F_n^{-1}(x))}\in\mathcal{D}(\Omega_F)\,.$$

The support of any test function φ lies in a ball, and any such ball overlaps with finitely many Ω_n. Therefore, the sum in (**20**.**5**) has finitely many terms and, hence, exists for any test function. Every term in the sum is a test function and therefore φ_F is a test function.

Moreover, the transformation $T_F:\ \mathcal{D}\to\mathcal{D}(\Omega_J)$, defined by (**20**.**5**), is linear and continuous. Indeed, if $\{\varphi_m\}$ is a null sequence in $\mathcal{D}$, then the sequence $\{\varphi_{m_F}\}$ obtained from $\{\varphi_m\}$ by the rule (**20**.**5**) is a null sequence in $\mathcal{D}(\Omega_J)$. Supports of all φ_m lie in one ball B_R. Then the series (**20**.**5**) for each $\varphi=\varphi_m$ has finitely many terms, defined by the condition $|F_n^{-1}(x)|<R$ that is independent of m. Since each term in the sum is a null sequence in $\mathcal{D}(\Omega_J)$ (see Sec.**20.1**), $\{\varphi_{m_F}\}$ is also a null sequence in $\mathcal{D}(\Omega_J)$. Therefore the adjoint transformation $T_F^*:\ \mathcal{D}'(\Omega_J)\to\mathcal{D}'$ defines a distribution by the rule

$$\text{(20.6)}\qquad \Big(f(F(y)),\varphi(y)\Big)=\Big(f(x),\varphi_F(x)\Big)\,.$$

20.3.3. An alternative derivation. Consider a regularization f_a of a distribution $f\in\mathcal{D}'(\Omega_J)$ by test functions, $f_a\to f$ in $\mathcal{D}'(\Omega_J)$ as $a\to 0$. Then all the equalities in (**20**.**4**) and (**20**.**5**) hold if f is replaced by test functions f_a. Since $f_a\to f$ in $\mathcal{D}(\Omega_J)$ as $a\to 0$, by continuity of the adjoint transformation $f_a(F(y))$ converges to $f(F(y))$ in $\mathcal{D}'$. Therefore taking the limit $a\to 0$ in $(f_a(F(y)),\varphi(y))=(f_a,\varphi_F)$ the rule (**20**.**6**) is obtained.

20.3.4. Example. Let $F(y)=\sin(y)$ and $y=\arcsin(x)\in(-\frac{\pi}{2},\frac{\pi}{2})$ denote the main branch of F^{-1}. Then for any $\varphi\in\mathcal{D}$ by (**20**.**5**),

$$\varphi_F(x)=\frac{\eta(x)}{\sqrt{1-x^2}}\sum_n\varphi\Big(\pi n+(-1)^n\arcsin(x)\Big)\ \in\mathcal{D}(-1,1)\,,$$

for any smooth function $\eta(x)$ with support in $(-1,1)$. The series has only finitely many non-zero terms defined by the condition $|\pi n+(-1)^n\arcsin(x)|<R$ if the support of $\varphi(x)$ lies in $[-R,R]$.

Let $f(x)=\delta(x)$. Then $\eta(x)$ is a test function with support in $(-1,1)$ and $\eta(x)=1$ near $x=0$. It follows from the rule (**20**.**6**) that

$$\Big(\delta(\sin(y)),\varphi(y)\Big)=\varphi_F(0)=\sum_n\varphi(\pi n)=\Big(\sum_n\delta(y-\pi n),\varphi(y)\Big)\,.$$

Therefore

$$\text{(20.7)}\qquad \delta(\sin(y))=\sum_n\delta(y-\pi n)\,.$$

Note that the series of shifted delta-functions converges in the distributional sense because

$$\lim_{n\to\infty}\Big(\sum_{|k|<n}\delta(y-\pi k),\varphi(y)\Big)=\sum_{|k|<R/\pi}\varphi(\pi k)\,.$$

if the support of φ lies in $[-R,R]$.

Let $\omega_a(x)$ be a hat function and $a<1$. Then $\omega_a\to\delta$ in $\mathcal{D}'(-1,1)$ as $a\to 0^+$. It follows from the continuity of the adjoint transformation that

$$\lim_{a\to 0^+}\omega_a(\sin(y))=\sum_n\delta(y-\pi n)\,.$$

20.3.5. Remark. It is possible that Ω_J is the union of two or more disjoint open sets. Let $\Omega_J=\Omega_{J1}\cup\Omega_{J2}$ and $\Omega_{J1,2}$ do not overlap. The stated rule for changing variables holds only for distributions with supports in Ω_J. This implies that the support of any such distribution f contains (disjoint) closed sets that lie in non-overlapping open sets so that f is uniquely defined by its reductions to Ω_{Jk}, $k=1,2$. Let

$$f(x)=f_k(x)\,,\quad x\in\Omega_{Jk}\,.$$

Then the change of variables can be carried out for each $f_k(x)$ by the rules (**20**.**6**) and (**20**.**5**) so that

$$f(F(y))=f_k(F(y))\,,\quad y\in F^{-1}(\Omega_{Jk})\,.$$

In this case, Eq. (**20**.**3**) should be solved in each Ω_{Jk}. Moreover, the number of solution (branches of the inverse) may be different for each Ω_{Jk}. The procedure is readily extended to Ω_J being the countable union of non-overlapping open sets such that any ball intersects finitely many of them.

For example, let $F:\ \mathbb{R}\to\mathbb{R}$ defined by

$$x=F(y)=y(y-a)(y-b)\,,\quad 0<a<b\,.$$

The largest open set that contains no zeros of the Jacobian has three disjoint open intervals

$$\Omega_J=(-\infty,x_a)\cup(x_a,x_b)\cup(x_b,\infty)=\Omega_{J1}\cup\Omega_{J2}\cup\Omega_{J3}\,,$$

where $x_a=F(y_a)$ and $x_b=F(y_b)$ and $y_{a,b}$ are roots of the equation $F'(y)=0$. Equation (**20**.**3**) has just one solution for $x\in\Omega_{J1}$ or $x\in\Omega_{J3}$, but three solutions for $x\in\Omega_{J2}$. The stated change of variables can be carried out for any distribution whose support does not contain the singular points $x=x_{a,b}$. For example, if $f(x)=\delta(x-c)$, then

$$\delta(F(y)-c)=\sum_n\frac{1}{|F'(y_n)|}\delta(y-y_n)\,.$$

where y_n are the roots of the equation $F(y) = c$. This equation has three roots if c lies between the singular points x_a and x_b and just one root otherwise. The change of variables cannot be defined if $c = x_a$ or $c = x_b$. Technical details are left to the reader as an exercise.

20.4. Transformations from class C^1. If a change of variables is not from class C^∞ or the support of a distribution includes zeros of the Jacobian, the above construction does not work. However, the alternative approach based on the continuity of the adjoint transformation introduced in Sec.**20.3.3** in combination with the completeness theorem can still work. Let f_a be a family of regular distributions such that $f_a \to f$ in $\mathcal{D}'$ as $a \to 0$. For example, f_a is a regularization of f. Suppose a C^1 transformation F of $\mathbb{R}^N$ has the following properties. There exist at most countably many open sets Ω_n with piecewise smooth boundaries that are not intersecting but the union of the closures Ω_n coincides with $\mathbb{R}^N$, any ball intersects only finitely many Ω_n, and $F:\ \Omega_n \to \Omega_J$ is a change of variables in the Lebesgue integral. Under these assumptions

$$\Big(f_a(F(y)), \varphi(y)\Big) = \int f_a(F(y))\varphi(y)\, d^N y = \sum_n \int_{\Omega_n} f_a(F(y))\varphi(y)\, d^N y$$
$$= \int_{\Omega_J} f_a(x)\varphi_F(x)\, d^N x\,,$$

where φ_F is given by (**20**.**5**) (with $\eta(x) = 1$) but it is not a test function because φ_T is not smooth enough. It can now have integrable singularities at zeros of the Jacobian. Since the integral exists for any $a \neq 0$, put, by definition,

$$(\mathbf{20.8}) \qquad \Big(f(F(y)), \varphi(y)\Big) \stackrel{\text{def}}{=} \lim_{a\to 0^+} \Big(f_a(F(y)), \varphi(y)\Big)\,,$$

provided the limit exists for any test function φ. If F is a diffeomorphism of $\mathbb{R}^N$ into Ω_J, then φ_F is a test function and the limit is (f, φ_F), where $f \in \mathcal{D}'(\Omega_J)$ (see Sec.**20.2**). For a general change of variables, the argument fails because φ_F is not a test function, and *the limit must be investigated by other means*. However, if it exists, then by the completeness theorem $f(F(y))$ is a distribution from $\mathcal{D}'$.

20.4.1. General change of variables in a delta function. Let $x = F(y) \in C^1(\mathbb{R}^N)$ such that the equation $F(y) = 0$ have at most countably many roots $y = y_n$, the sequence $\{y_n\}$ has no limit points (in other words, any ball contains only finitely many roots), and the Jacobian does not vanish at the roots, $J(y_n) \neq 0$. Then

$$(\mathbf{20.9}) \qquad \delta(F(y)) = \sum_n \frac{1}{J(y_n)}\, \delta(y - y_n)\,, \quad J(y) = \Big|\det\Big(\frac{\partial F_j(y)}{\partial y_k}\Big)\Big|\,.$$

In particular, for any C^1 function $F(y)$, $y \in \mathbb{R}$, with simple isolated zeros y_n,

$$\delta(F(y)) = \sum_n \frac{1}{|F'(y_n)|}\, \delta(y - y_n)\,.$$

Note that the convergence of the series of shifted delta-functions follows from that only finitely many zeros y_n lie in support of any test function:

$$\Big(\delta(F(y)), \varphi(y)\Big) = \sum_{|y_n|<R} \frac{\varphi(y_n)}{J(y_n)}\,, \quad \operatorname{supp}\varphi \subset B_R\,.$$

To prove the assertion, consider a regularization of the delta function by a hat function $\omega_a(x)$, $\omega_a \to \delta$ in $\mathcal{D}'$ as $a \to 0^+$. The support of ω_a is the ball $|x| \leq a$. By hypothesis, the Jacobian does not vanish in a neighborhood of each root y_n of the equation $F(y) = 0$. Therefore the transformation is invertible in this neighborhood by the implicit function theorem. For a small enough a_0, there exists $\delta_n > 0$ such that $y = F_n^{-1}(x)$ for $|y - y_n| < \delta_n$ and $F(F_n^{-1}(x)) = x$ for all $|x| < a_0$. The local inverse transformations F_n^{-1} are from class C^1 (by the implicit function theorem). For all $a < a_0$, the images $F_n^{-1}(B_a)$ are bounded and not overlapping sets, and the support of the function $\omega_a(F(y))$ lies in their union. The chain of equalities given (**20**.**4**) and (**20**.**5**) holds in this case because $\omega_a(F(y))$ has properties similar to $\eta(F(y))$ where $\Omega_n = F_n^{-1}(B_a)$ and $\Omega_J = B_a$. Therefore

$$\Big(\delta(F(y)), \varphi(y)\Big) \stackrel{\text{def}}{=} \lim_{a\to 0^+} \int_{|x|<a} \omega_a(x)\varphi_{F}(x)\, d^N x\,,$$

$$\varphi_{F}(x) = \sum_n \frac{1}{J(F_n^{-1}(x))}\, \varphi(F_n^{-1}(x))\,, \quad |x| < a\,.$$

The series converges because it has only finitely many terms owing to boundedness of the support of φ, $\varphi(F_n^{-1}(x)) = 0$ for all large enough n. By the assumption any ball B_R contains only finitely roots y_n. Therefore the support of φ overlaps only with finitely many neighborhoods $F_n^{-1}(B_a)$.

Let us find the limit of the integral (ω_a, φ_{F}) (it does not follow from $\omega_a \to \delta$ because φ_{F} is merely a continuous function). To this end, using the normalization properties of the hat function the integral can be transformed as follows

$$\int \omega_a(x)\varphi_{_F}(x)\,d^Nx = \varphi_{_F}(0) + \int \omega_a(x)\Big(\varphi_{_F}(x) - \varphi_{_F}(0)\Big)\,d^Nx$$

The integral in the right-hand side vanishes in the limit $a \to 0^+$. Indeed, $|\varphi_{_F}(x) - \varphi_{_F}(0)|$ is continuous on the closed ball $|x| \le a$ which is the support of ω_a. Therefore it attains its maximal value on it. Put

$$M_a = \max_{|x|\le a}\Big|\varphi_{_F}(x) - \varphi_{_F}(0)\Big|\,.$$

Then by the normalization property of the hat function

$$\Big|\int \omega_a(x)\Big(\varphi_{_F}(x) - \varphi_{_F}(0)\Big)\,d^Nx\Big| \le M_a\int_{|x|<a}\omega_a(x)\,d^Nx = M_a\,.$$

But $M_a \to 0$ as $a \to 0^+$ by continuity of $\varphi_{_F}$. Therefore

$$\Big(\delta(F(y)),\varphi(y)\Big) = \varphi_{_F}(0) = \sum_n \frac{1}{J(y_n)}\,\varphi(y_n)\,,$$

which proves (**20**.**9**).

20.4.2. Change of variables in the principal value distribution. Consider a transformation $F:\ \mathbb{R}\to\mathbb{R}$ defined in the example of Sec.**20.3.5** as a change of variables in $\mathcal{P}\frac{1}{x}$. The support of $\mathcal{P}\frac{1}{x}$ is $\mathbb{R}$ and, hence, contains the singular points of F. Let us attempt to make such a change of variables using the limit method (**20**.**8**). It is not difficult to verify that

$$f_\varepsilon(x) = \frac{\theta(|x|-\varepsilon)}{x} \to f(x) = \mathcal{P}\frac{1}{x} \quad \text{in } \mathcal{D}'$$

as $\varepsilon \to 0^+$. Therefore by (**20**.**8**)

$$\Big(\mathcal{P}\frac{1}{F(y)},\varphi\Big) = \lim_{\varepsilon\to 0^+}\int_{|F(y)|>\varepsilon}\frac{\varphi(y)}{F(y)}\,dy\,,$$

provided the limit exists. Using the partial fraction decomposition

$$\frac{1}{F(y)} = \frac{1}{ab}\cdot\frac{1}{y} + \frac{1}{a(a-b)}\cdot\frac{1}{y-a} + \frac{1}{b(b-a)}\cdot\frac{1}{y-b}\,,$$

one can guess that

$$\mathcal{P}\frac{1}{F(y)} = \frac{1}{ab}\mathcal{P}\frac{1}{y} + \frac{1}{a(a-b)}\mathcal{P}\frac{1}{y-a} + \frac{1}{b(b-a)}\mathcal{P}\frac{1}{y-b}\,.$$

This is indeed so, provided the integral over the set $|F(y)| > \varepsilon$ is *equivalent* to the Cauchy principal value regularization of the three integrals in the partial fraction decomposition when $\varepsilon \to 0^+$.

All roots of $F(y)$ are simple. Suppose y_0 is a root of $F(y)$. Support of φ is bounded and, hence, lies in an interval $|y-y_0| < R$ for some $R > 0$. By linearity of the integral, it is sufficient to show that

$$\lim_{\varepsilon\to 0^+}\int_{|F(y)|>\varepsilon}\frac{\varphi(y)}{y-y_0} = \Big(\mathcal{P}\frac{1}{y-y_0},\varphi(y)\Big) = \int_{|y-y_0|<R}\frac{\varphi(y)-\varphi(y_0)}{y-y_0}\,dy\,.$$

The equation $F(y) = \pm\varepsilon$ has the roots $y = y_\pm = y_0 \pm \varepsilon/g(y_0) + O(\varepsilon^2)$ where $F(y) = g(y)(y-y_0)$. Therefore

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\int_{|F(y)|>\varepsilon}\frac{\varphi(y)}{y-y_0} &= \int_{|y-y_0|<R}\frac{\varphi(y)-\varphi(y_0)}{y-y_0}\,dy\\ &\quad+\varphi(y_0)\lim_{\varepsilon\to 0^+}\Big(\int_{y_0-R}^{y_-}+\int_{y_+}^{y_0+R}\Big)\frac{dy}{y-y_0}\,,\end{aligned}$$

where it is assumed that $y_- < y_0 < y_+$ (or $g(y_0) > 0$). The limit in the right-hand vanishes. Indeed, by a direct evaluation of the integral

$$\lim_{\varepsilon\to 0^+}\ln\left|\frac{y_- - y_0}{y_+ - y_0}\right| = \lim_{\varepsilon\to 0^+}\ln|1+O(\varepsilon)| = 0$$

as required. The analysis can be extended to any smooth $F(y)$ with at most countably many simple roots if the sequence of the roots has no limit points.

20.5. Exercises.

1. Express each of the following distributions in terms of shifted delta-functions

(i) $\delta(3x-4)\,,\quad x\in\mathbb{R}$

(ii) $\delta(y(x))\,,\quad y\in\mathbb{R}^2\,,\quad y_1 = 2x_1+3x_2-3\,,\ y_2 = x_1 - x_2 + 2$

(iii) $\sin(x)\delta(2x+5\pi)\,,\quad x\in\mathbb{R}$

2. Complete the proof of Corollary **20.1**.

3. Find the distributional limit of the following sequence of functions

$$f_n(x) = \frac{\sin(n\tanh(x))}{\pi x}\,,$$

where $n = 1, 2, \ldots.$

4. Express $\delta(x^2-a^2)$ in terms of shifted delta functions if $a > 0$.

5. For the example of $x = F(y)$ given in Sec.**20.3.5**,
(i) justify the expression given for $\delta(F(y)-c)$,

(ii) give the explicit form of $\delta(F(y))$ as a linear combination of (shifted) delta functions.

6. Let
$$f_a(x) = \frac{\theta(|x|-a)}{x}, \quad x \in \mathbb{R}.$$
(i) Show that $f_a(x) \to \mathcal{P}\frac{1}{x}$ in $\mathcal{D}'$ as $a \to 0^+$.
(ii) Use the rule (**20**.**8**) to find $\mathcal{P}\frac{1}{\sin(y)}$. Give an explicit rule to find the value of $\mathcal{P}\frac{1}{\sin(y)}$ on a test function.

7. Use the rule (**20**.**8**) and the limit
$$\frac{1}{x \pm i\varepsilon} \to \frac{1}{x \pm i0^+}$$
for $\varepsilon \to 0^+$ to change variables in the Sokhotsky distributions:
(i) $x = y(y-a)(y-b)$;
(ii) $x = \sin(y)$.
Express the answer in terms of Sokhotsky distributions and delta functions (if necessary).

8. Put $f_k(x) = a(x)\delta^{(k)}(\sin(x))$, $k = 0, 1, 2, ...$, where $a \in C^\infty$. Express f_k in terms of shifted delta functions and its derivatives or show that f_k is not a distribution.

21. Differentiation of distributions

21.1. Derivatives of a distribution. The rule (**13**.5) can be extended to partial derivatives of any order for any distribution f:

$$(D^\beta f, \varphi) = (-1)^\beta (f, D^\beta \varphi)\,. \tag{21.1}$$

Of course, one has to show that $D^\beta f$ is a linear and continuous functional on the space of test functions. This follows from the linearity and continuity of f. If $\varphi_n \to 0$ in $\mathcal{D}$, then $D^\beta \varphi_n \to 0$ in $\mathcal{D}$ for any β because the sequence of *any* partials $D^\alpha \varphi_n$ converges to zero uniformly. So, the functional $D^\beta f$ is continuous for any continuous f. Thus, all distributions are infinitely many times differentiable. If a distribution is defined by a smooth function, the classical and distributional derivatives are equal

$$\{D^\beta f(x)\} = D^\beta f(x)\,, \quad \beta \le p\,, \quad f \in C^p\,. \tag{21.2}$$

This follows from the integration by parts and the du Bois-Raymond lemma:

$$\begin{aligned}(D^\beta f, \varphi) &= (-1)^\beta (f, D^\beta \varphi) = (-1)^\beta \int f(x) D^\beta \varphi(x)\, d^N x \\ &= \int \{D^\beta f(x)\}, \varphi(x)\, d^N x = (\{D^\beta f\}, \varphi)\,,\end{aligned}$$

for any test function φ. Owing to the boundedness of support of φ, the integration by parts does not produce any boundary terms.

Definition (**21**.1) implies that derivatives of a singular distribution are singular distributions. In particular, the Dirac delta-function can be differentiated any number of times so that

$$(D^\beta \delta, \varphi) = (-1)^\beta D^\beta \varphi(0)\,. \tag{21.3}$$

The rule (**21**.1) also implies that *any locally integrable function can be differentiated any number of times if derivatives are viewed in the distributional sense!*

21.1.1. Distributional derivatives of the absolute value. Let $f(x) = |x|$. It is continuous on $\mathbb{R}$ but not differentiable at $x = 0$. The distributional derivatives are

$$|x|' = \varepsilon(x)\,, \quad |x|'' = 2\delta(x)\,, \quad |x|^{(n)} = 2\delta^{(n-2)}(x)\,, \quad n \ge 2\,.$$

where $\varepsilon(x)$ is the sign function. Indeed, the third relation follows from (**21**.3) if the second one holds. The first and second relations are proved by the following chain of equalities for any test function φ:

$$\begin{aligned}(|x|', \varphi) &= -(|x|, \varphi') = \int_{-\infty}^{0} x\varphi'(x)\,dx - \int_{0}^{\infty} x\varphi'(x)\,dx \\ &= -\int_{-\infty}^{0} \varphi(x)\,dx + \int_{0}^{\infty} \varphi(x)\,dx = (\varepsilon, \varphi)\,, \\ (|x|'', \varphi) &= -(\varepsilon, \varphi') = \int_{-\infty}^{0} \varphi'(x)\,dx - \int_{0}^{\infty} \varphi'(x)\,dx \\ &= \varphi(0) + \varphi(0) = (2\delta, \varphi)\,.\end{aligned}$$

21.1.2. Distributional derivatives of the log function. The function $f(x) = \ln(|x|)$ is locally integrable on $\mathbb{R}$. It defines a regular distribution. Its classical derivative $\{f'(x)\} = \frac{1}{x}$, $x \neq 0$, is not locally integrable and, hence, does not define a distribution. However, the distributional derivative f exists:

$$\frac{d}{dx}\ln(|x|) = \mathcal{P}\frac{1}{x}\,.$$

For any test function φ, one has

$$\begin{aligned}(f', \varphi) \;&=\; -(f, \varphi') = -\int \ln(|x|)\varphi(x)\,dx \\ &\overset{(1)}{=} -\lim_{a\to 0^+}\int_{|x|>a} \ln(|x|)\varphi'(x)\,dx = \lim_{a\to 0^+}\int_{|x|>a} \frac{\varphi(x)}{x}\,dx \\ &\overset{(2)}{=} \left(\mathcal{P}\frac{1}{x}, \varphi\right),\end{aligned}$$

where (1) and (2) hold by continuity of the Lebesgue integral and by integration by parts, respectively. The higher-order derivatives are obtained from the relation

$$(\mathbf{21}.4) \qquad \frac{d}{dx}\mathcal{P}\frac{1}{x^n} = -n\mathcal{P}\frac{1}{x^{n+1}}\,, \quad n = 1, 2, \ldots$$

It is proved in a similar fashion. Let p_n and $\tilde{p}_n$ be Taylor polynomials of order n for a test function φ and its derivative φ' about $x = 0$, respectively. Then $\tilde{p}_{n-2}(x) = p'_{n-1}(x)$, and one infers that

$$\begin{aligned}&\left(\frac{d}{dx}\mathcal{P}\frac{1}{x^n}, \varphi\right) = -\left(\mathcal{P}\frac{1}{x^n}, \varphi'\right) = -\lim_{a\to 0^+}\int_{|x|>a} \frac{\varphi'(x) - \tilde{p}_{n-2}(x)}{x^n}\,dx \\ &\overset{(1)}{=} \lim_{a\to 0^+}\left(\frac{\varphi(x) - p_{n-1}(x)}{x^n}\Big|_{-a}^{a} - n\int_{|x|>a} \frac{\varphi(x) - p_{n-1}(x)}{x^{n+1}}\,dx\right) \\ &\overset{(2)}{=} -n\left(\mathcal{P}\frac{1}{x^{n+1}}, \varphi\right),\end{aligned}$$

where (1) is obtained by integration by parts and (2) holds because the boundary term vanishes in the limit by property (**15**.**4**).

It follows from (**21**.**4**) that

$$\frac{d^n}{dx^n}\mathcal{P}\frac{1}{x} = (-1)^n n!\,\mathcal{P}\frac{1}{x^{n+1}}\,. \tag{21.5}$$

21.2. Properties of distributional derivatives. Here basic properties of classical derivatives are extended to distributions and applied to calculate distributional derivatives of some commonly used distributions. A continuous function is not generally differentiable in the classical sense, but it defines a regular distribution and, hence, can be differentiated infinitely many times in the distributional sense. It turns out that *any distribution can be viewed as a linear combination of distributional derivatives of some continuous functions.*

21.2.1. Clairaut's theorem for distributions. The order of classical partial derivatives does not matter if they are continuous (Clairaut's theorem). Distributional derivatives can be taken in any order, the result is the same derivative:

$$D^\alpha(D^\beta f) = D^\beta(D^\alpha f) = D^{\alpha+\beta} f$$

for any distribution f. This follows from Clairaut's theorem for test functions

$$D^\alpha(D^\beta \varphi) = D^\beta(D^\alpha \varphi) = D^{\alpha+\beta}\varphi\,, \quad \varphi \in \mathcal{D}$$

because partial derivatives of any order of a test function are continuous. Indeed,

$$\begin{aligned}(D^\alpha(D^\beta f),\varphi) &= (-1)^\alpha(D^\beta f, D^\alpha\varphi) = (-1)^{\alpha+\beta}(f, D^\beta(D^\alpha\varphi))\\ &= (-1)^{\alpha+\beta}(f, D^\alpha(D^\beta\varphi)) = (D^\beta(D^\alpha f),\varphi)\\ (D^\alpha(D^\beta f),\varphi) &= (-1)^{\alpha+\beta}(f, D^{\alpha+\beta}\varphi) = (D^{\alpha+\beta} f,\varphi)\,.\end{aligned}$$

21.2.2. Leibniz rule. If a is a smooth function and f is a distribution, then the product af is a distribution and its derivative can be found by the Leibniz rule (known also as the product rule for classically differentiable functions):

$$D\Big(a(x)f(x)\Big) = Da(x)\,f(x) + a(x)\,Df(x)\,.$$

Note that $Da(x) \in C^\infty$ if $a \in C^\infty$ so that the right-hand side makes sense as a distribution. A proof of this rule is elementary and follows from the product rule for smooth functions:

$$\begin{aligned}\Big(D(af),\varphi\Big) &= -\Big(af, D\varphi\Big) = -(f, aD\varphi)\\ &= -\Big(f, D(a\varphi)\Big) + \Big(f, \varphi Da\Big)\\ &= \Big(Df, a\varphi) + \Big(Da\, f, \varphi\Big)\\ &= \Big(a\, Df + Da\, f, \varphi\Big)\end{aligned}$$

for any test function φ.

It is also clear that the Leibniz rule can be used any number of times by means of the binomial expansion:

$$D^\alpha(af) = \sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} a\, D^\beta f\,,$$

where notations from Sec.**1.3** are used

21.2.3. Simple examples. It has been shown that the derivative of a step function is the delta function, $\theta' = \delta$. Consider a continuous function $f(x) = x\theta(x)$. By the Leibniz rule

$$\begin{aligned}f'(x) &= \theta(x) + x\delta(x) = \theta(x)\,,\\ f''(x) &= \delta(x)\,.\end{aligned}$$

By differentiating the identity $a(x)\delta(x) = a(0)\delta(x)$, one infers that

$$a(x)\delta'(x) = a(0)\delta'(x) - a'(0)\delta(x)\,.$$

A similar identity for the second derivative of the delta function is obtained by taking the second derivative of the said identity:

$$\begin{aligned}a(x)\delta''(x) &= a(0)\delta''(x) - 2a'(x)\delta'(x) - a''(0)\delta(x)\\ &= a(0)\delta''(x) - 2a'(0)\delta'(x) + a''(0)\delta(x)\,.\end{aligned}$$

21.2.4. Continuity of differentiation on the space of distributions. Differentiation of distributions can be viewed as the adjoint of differentiation on the space of test functions:

$$T = (-1)^\beta D^\beta : \quad \varphi \to T(\varphi) = (-1)^\beta D^\beta \varphi$$

As shown in the beginning of this section, differentiation is a linear and continuous transformation of the space of test functions into itself. Therefore the adjoint transformation $T^*(f) = D^\beta f$, $f \in \mathcal{D}'$, is a linear and continuous transformation of the space of distributions into itself:

$$(\mathbf{21.6}) \qquad f_n \to f \quad \text{in } \mathcal{D}' \quad \Rightarrow \quad D^\beta f_n \to D^\beta f \quad \text{in } \mathcal{D}'$$

In particular, this implies that *any converging series of distributions can be differentiated term-by-term infinitely many times and the series of the derivatives converges to the corresponding derivative of the sum of the series*:

$$f = \sum_n f_n \quad \Rightarrow \quad D^\beta f = \sum_n D^\beta f_n \,, \quad \{f_n\} \subset \mathcal{D}' \,, \quad f \in \mathcal{D}' \,.$$

This is quite an improvement of the classical analysis where the uniform convergence of the series of derivatives is usually required!

21.2.5. Differentiation and regularization of distributions. Let $f_a = \phi_a * f$ be a regularization of a distribution f by a smooth function such that $f_a \to f$ in $\mathcal{D}'$ as $a \to 0$ where the convolution is defined in (**19**.**1**). Then by continuity of differentiation $D^\beta f_a \to D^\beta f$ in $\mathcal{D}'$ as $a \to 0$. Moreover, the following relations hold

$$D^\beta f_a = D^\beta \phi_a * f = \phi_a * D^\beta f \,. \tag{21.7}$$

They follows from property (**19**.**2**) for the convolution of test functions. Indeed, for any test function φ one infers that

$$\begin{aligned}(D^\beta f_a, \varphi) &= (-1)^\beta (f_a, D^\beta \varphi) = (-1)^\beta (f, \phi_a^- * D^\beta \varphi) \\ &= (-1)^\beta (f, D^\beta \phi_a^- * \varphi) = (f, (D^\beta \phi_a)^- * \varphi) \\ &= (D^\beta \phi_a * f, \varphi)\end{aligned}$$

and

$$\begin{aligned}(D^\beta f_a, \varphi) &= (-1)^\beta (f, \phi_a^- * D^\beta \varphi) = (-1)^\beta (f, D^\beta (\phi_a^- * \varphi)) \\ &= (D^\beta f, \phi_a^- * \varphi) = (\phi_a * D^\beta f, \varphi) \,.\end{aligned}$$

For example, a hat function $\omega_a = \omega_a * \delta$ converges to the delta function as $a \to 0^+$. By continuity of differentiation in $\mathcal{D}'$,

$$D^\beta \omega_a(x) \to D^\beta \delta(x) \quad \text{in } \mathcal{D}'(\mathbb{R}^N) \,,$$

as $a \to 0^+$.

The significance of (**21**.**7**) is that *a regularization and differentiation are commutative operations* on the space of distributions. This property will be useful for physical interpretation of distributional solutions to partial differential equations.

21.3. Classical vs distributional derivatives. For smooth functions classical and distributional derivatives coincide. If a locally integrable function is not differentiable at some points, then this relation does not generally holds near these points. Suppose that a function f that is locally integrable on $\mathbb{R}^N$ has continuous partials in some open $\Omega \subset \mathbb{R}^N$, then the distributional and classical partials of f coincide in Ω:

$$f \in \mathcal{L}_{\text{loc}} \cap C^1(\Omega) \quad \Rightarrow \quad Df(x) = \{Df(x)\}\,, \quad x \in \Omega \subset \mathbb{R}^N\,.$$

A proof is elementary. The classical derivative $\{Df(x)\}$ is a regular distribution from $\mathcal{D}'(\Omega)$. By integration by parts in the integral $(\{Df\}, \varphi)$ the later is shown to be equal to $-(f, D\varphi) = (Df, \varphi)$ for any $\varphi \in \mathcal{D}(\Omega)$, which means that the distribution Df is equal to $\{Df\}$ in Ω.

Suppose $\{Df\}$ does not exist at a point $x = x_0$. Then near x_0, a relation between classical and distributional derivatives becomes more complicated. If the classical derivative $\{Df\}$ remains locally integrable, then Df and $\{Df\}$ can be equal in the distributional sense or differ by a distribution supported at the single point $x = x_0$. If $\{Df\}$ is singular at x_0, then the distributional derivative is some distributional regularization of it, $Df = \text{Reg}\,\{Df\}$.

This relation is first investigated in the case of one real variable. Suppose f is locally integrable and the derivative $\{f'\}$ does not exist at $x = 0$ and is continuous otherwise. Let us find the distributional derivative. For any test function φ, using continuity of the Lebesgue integral and integration by parts, one has

$$(f', \varphi) = -(f, \varphi') = -\int f(x)\varphi'(x)\,dx = \lim_{a\to 0^+} \int_{|x|>a} f(x)\varphi'(x)\,dx$$

$$= \lim_{a\to 0^+} \left(f(a)\varphi(a) - f(-a)\varphi(-a) + \int_{|x|>a} \{f'(x)\}\varphi(x)\,dx \right). \tag{21.8}$$

The limit exists by the existence of $f' \in \mathcal{D}'$. However, its calculation depends on the behavior of f and $\{f'\}$ near $x = 0$.

Suppose that f has a jump discontinuity at $x = 0$ so that $f(\pm a) \to f_\pm$ as $a \to 0^+$. Then the boundary term converges to $(f_+ - f_-)\varphi(0)$ and, hence, the integral must also converge and defines a principal value regularization of $\{f'\}$ so that

$$f'(x) = (f_+ - f_-)\delta(x) + \mathcal{P}\{f'(x)\}\,. \tag{21.9}$$

It should be noted that even if f is continuous at $x = 0$, $f_+ = f_-$, the existence of the limit of the integral does not imply that $\{f'\}$ is locally integrable and defines a regular distribution. The integral can only converge *conditionally*! For example, the derivative of a continuous function $f(x) = x\sin(\frac{1}{x})$ is not locally integrable (see Exercises). However, if f is absolutely continuous, then the boundary term vanishes, the classical derivative $\{f'\}$ exists almost everywhere and is locally integrable so that by the du Bois-Raymond lemma:

$$f'(x) = \{f'(x)\}\,, \quad f \in AC^0\,.$$

Let us extend the analysis to the multi-variable case. Suppose that classical derivatives $\{Df\}$ of a locally integrable function f are singular on a set $S_f \subset \mathbb{R}^N$ and otherwise smooth enough so that the integration by parts is permitted on any compact that does not intersect S_f. Let N_a be a neighborhood of S_f of radius a and $\hat{n}_x$ be the outward unit normal on ∂N_a. Then

$$\begin{aligned}
(Df, \varphi) &= -(f, D\varphi) = -\int f(x) D\varphi(x)\, d^N x \\
&= -\lim_{a\to 0^+} \int_{\mathbb{R}^{\mathbb{N}}\setminus S_a} f(x) D\varphi(x)\, d^N x \\
&= \lim_{a\to 0^+} \left(\int_{\partial N_a} f(x)\varphi(x)\hat{n}_x\, dS + \int_{\mathbb{R}^{\mathbb{N}}\setminus N_a} \{Df(x)\}\varphi(x)\, d^N x \right). \qquad (\mathbf{21.10})
\end{aligned}$$

The limit exists because Df exists as a distribution. Calculation of the limit depends on the behavior of f and $\{Df\}$ near S_f and is illustrated with examples in the next section. If f is continuous at S_f and $\{Df\}$ is locally integrable, then $Df = \{Df\}$. The same approach can be used to calculate high-order distributional derivatives of functions that are not smooth everywhere.

21.4. Distributional derivatives. Examples. Here the technique for calculating distributional derivatives outlined in Sec.**21.3** is illustrated with examples.

21.4.1. Distributional derivatives vs singular classical derivatives. Suppose that a locally integrable function f is smooth everywhere on $\mathbb{R}$ but $x = 0$, and the classical derivative $\{f'(x)\}$ is singular at $x = 0$. Then the relation (**21.8**) still applies and, as argued, the distributional derivative is a distributional regularization of $\{f'(x)\}$. As an example, let $x \in \mathbb{R}$ and $f(x) = |x|^{-q}$ where $0 < q < 1$. Then $f \in \mathcal{L}_{\rm loc}$ but its classical derivative has a non-integrable singularity at $x = 0$ and is continuous otherwise:

$$\{f'(x)\} = -q\,\frac{x}{|x|^{2+q}}\,, \quad x \neq 0\,.$$

Then for any test function φ with support in $[-R, R]$

$$(f', \varphi) = -(f, \varphi') = -\lim_{a\to 0^+} \int_{a<|x|<R} f(x)\varphi'(x)\,dx$$
$$= \lim_{a\to 0^+} \Big(-f(x)\varphi(x)\Big|_{-R}^{-a} - f(x)\varphi(x)\Big|_{a}^{R} + \int\limits_{a<|x|<R} \{f'(x)\}\varphi(x)\,dx \Big)$$

The boundary term vanishes in the limit because

$$\frac{\varphi(a) - \varphi(-a)}{a^q} = O(a^{1-q})\,.$$

Since the integral of $x/|x|^{q+2}$ over the symmetric interval $a < |x| < 1$ vanishes, one has

$$(f', \varphi) = -q \lim_{a\to 0^+} \int_{a<|x|<1} \frac{x\varphi(x) - x\varphi(0)}{|x|^{q+2}}\,dx - q \int_{|x|>1} \frac{x\varphi(x)}{|x|^{q+2}}\,dx$$

The integrand in the first integral is locally integrable and the regularization can be removed. It follows from (**16**.**2**) that

$$f'(x) = -qx\mathcal{P}_r \frac{1}{|x|^{2+q}} = \mathcal{P}_r\{f'(x)\}\,,$$

which is a distributional regularization of the classical derivative.

Let

$$f(x) = \begin{cases} \frac{1}{x^q} & , x > 0 \\ 0 & , x < 0 \end{cases} \quad \Rightarrow \quad \{f'(x)\} = \begin{cases} -\frac{q}{x^{q+1}} & , x > 0 \\ 0 & , x < 0 \end{cases}\,,$$

where $0 < q < 1$. Then $f \in \mathcal{L}_{\text{loc}}$ but $\{f'\}$ is singular at $x = 0$. Then the relation (**21**.**8**) leads to

$$(f', \varphi) = \lim_{a\to 0^+} \Big(\frac{\varphi(a)}{a^q} + \int_{|x|>a} \{f'(x)\}\varphi(x)\,dx \Big)$$
$$= \lim_{a\to 0^+} \Big(\frac{\varphi(a) - \varphi(0)}{a^q} + \varphi(0) + \int_{a<|x|<1} \{f'(x)\}[\varphi(x) - \varphi(0)]\,dx$$
$$+ \int_{|x|>1} \{f'(x)\}\varphi(x)\,dx \Big)\,.$$

By taking the limit, it is concluded that

$$f'(x) = \delta(x) + \mathcal{P}_r\{f'(x)\}\,,$$

which is again a distributional regularization of the singular classical derivative.

21.4.2. The distribution $\mathcal{P}_r\frac{1}{|x|}$. The function $f(x) = |x|\ln(|x|)$ (extended so that $f(0) = 0$) is continuous everywhere and its classical first derivative is locally integrable. Therefore

$$\frac{d}{dx}|x|\ln(|x|) = \left\{\frac{d}{dx}|x|\ln(|x|)\right\} = \varepsilon(x)\ln(|x|) + 1\,,$$

where $\varepsilon(x)$ is the sign function. The second derivative is found as follows. For any test function φ:

$$\begin{aligned}\Big((|x|\ln(|x|))'', \varphi\Big) &= -\Big((\varepsilon(x)\ln(|x|))', \varphi'\Big)\\ &= \lim_{a\to 0^+}\Big[\int_{-\infty}^{-a}\ln(|x|)\varphi'(x)\,dx - \int_a^{\infty}\ln(|x|)\varphi'(x)\,dx\Big]\\ &= \lim_{a\to 0^+}\Big[\ln(a)\Big(\varphi(-a)+\varphi(a)-2\varphi(0)\Big)\\ &\quad + \int_{a<|x|<1}\frac{\varphi(x)-\varphi(0)}{|x|}\,dx + \int_{|x|>1}\frac{\varphi(x)}{|x|}\,dx\Big]\,,\end{aligned}$$

where the second equality is by continuity of the integral and the third one is obtained by integration by parts and by evaluation of the integral of $\frac{1}{|x|}$ over $a < |x| < 1$. Therefore

$$\frac{d^2}{dx^2}|x|\ln(|x|) = \mathcal{P}_r\frac{1}{|x|}\,.$$

Since $\mathcal{P}_r\frac{1}{|x|} = \frac{1}{|x|}$ in any open interval that does not contain $x = 0$ and $\frac{1}{|x|}$ is smooth in it, the classical and distributional derivatives must be equal in any such interval. Let us calculate the distributional derivative. For any test function φ,

$$\begin{aligned}\Big((\mathcal{P}_r\frac{1}{|x|})', \varphi(x)\Big) &= -\int_{|x|<1}\frac{\varphi'(x)-\varphi'(0)}{|x|}\,dx - \int_{|x|>1}\frac{\varphi'(x)}{|x|}\,dx\\ &= -\lim_{a\to 0^+}\int_{a<|x|<1}\frac{\varphi'(x)-\varphi'(0)}{|x|}\,dx - \int_{|x|>1}\frac{\varphi'(x)}{|x|}\,dx\\ &= \lim_{a\to 0^+}\left(\frac{\varphi(a)-\varphi(-a)}{a} - \int\limits_{a<|x|<1}\frac{x\varphi(x)-x^2\varphi'(0)}{|x|^3}\,dx\right)\\ &\quad - \int_{|x|>1}\frac{x\varphi(x)}{|x|^3}\,dx\,.\end{aligned}$$

The first term in the right-hand side of the last equality converges to $2\varphi(0)$. The numerator in the first integral can be replaced by $\psi(x) - p_2(x)$ where p_2 is the Taylor polynomial of order 2 about $x = 0$ for the function $\psi(x) = x\varphi(x)$. Note that $\psi(0) = 0$, $\psi'(0) = \varphi(0)$, but the integral of $x/|x|^3$ over the symmetric interval $a < |x| < 1$ vanishes, and only the term with $\psi''(0) = 2\varphi'(0)$ contributes. It follows from (**16**.**2**) that

$$\frac{d}{dx}\mathcal{P}_r\frac{1}{|x|} = 2\delta(x) - x\mathcal{P}_r\frac{1}{|x|^3}\,.$$

This is a distributional regularization of a singular classical derivative.

21.4.3. Coulomb and Newton potentials. The locally integrable function

$$f(x) = \frac{1}{|x|}\,, \quad x \in \mathbb{R}^3\,.$$

is, up to a constant factor, the Newton gravitational potential of a point particle at the origin, or the Coulomb electric potential of a point charge. By gravity and electrostatic laws the Laplacian of the potential is proportional to the mass or charge density, but such a density was shown to be proportional to the delta-function. Let us verify that the laws of gravity or electrostatic hold in the distributional sense.

The potential f is smooth for $x \neq 0$ and its classical gradient is locally integrable in $\mathbb{R}^3$. The limit of the surface integral in (**21**.**10**) vanishes:

$$\Big|\int_{|x|=a}\frac{\hat{n}_x\varphi(x)}{|x|}\,dS\Big| \leq \frac{\sup|\varphi|}{a}\int_{|x|=a} dS = 4\pi a\sup|\varphi| \to 0$$

as $a \to 0^+$. Therefore the classical and distributional gradients of f are equal

$$(\mathbf{21.11}) \qquad \partial_j\frac{1}{|x|} = \Big\{\partial_j\frac{1}{|x|}\Big\} = -\frac{x_j}{|x|^3}\,.$$

The classical second derivatives of the Coulomb potential in $\mathbb{R}^3$ are singular at $x = 0$ and smooth otherwise:

$$\Big\{\partial_j\partial_k\frac{1}{|x|}\Big\} = \frac{|x|^2\delta_{jk} - 3x_jx_k}{|x|^5}\,, \quad x \neq 0\,.$$

So, the corresponding distributional derivatives must be a distributional regularization of them near $x = 0$ and coincide with them on any open set that does not contain $x = 0$. For any test function φ with support in a ball $|x| < R$, one infers by continuity of the integral and by integrating by parts twice that

$$\begin{aligned}(\partial_j\partial_k f,\varphi) &= (f,\partial_j\partial_k\varphi) = \lim_{a\to 0^+}\int_{a<|x|<R}\frac{1}{|x|}\,\partial_j\partial_k\varphi(x)\,d^3x \\ &= \lim_{a\to 0^+}\Big(\oint\limits_{|x|=a}\frac{n_j}{|x|}\partial_k\varphi(x)\,dS - \oint\limits_{|x|=a} n_k\partial_j\frac{1}{|x|}\varphi(x)\,dS \\ &\quad + \int_{a<|x|<R}\{\partial_j\partial_k f(x)\}\varphi(x)\,d^3x\Big)\end{aligned}$$

where $n_j = -x_j/a$ is the unit normal on the sphere $|x| = a$, the surface integral over the sphere $|x| = R$ vanishes because φ is zero for $|x| = R$ together with all its partials. Evaluating the necessary derivatives and substituting n_j, one obtains

$$\begin{aligned}(\partial_j\partial_k f,\varphi) &= \lim_{a\to 0^+}\Big(-\frac{1}{a^2}\oint\limits_{|x|=a} x_j\partial_k\varphi(x)\,dS - \frac{1}{a^4}\oint\limits_{|x|=a} x_jx_k\varphi(x)\,dS \\ &\quad + \int_{a<|x|<R}\{\partial_j\partial_k f(x)\}\varphi(x)\,d^3x\Big)\,.\end{aligned}$$

The limit of the first surface integral vanishes. Indeed, since $|x_j| \le |x| = a$ and the gradient of a test function is bounded by $M = \sup|D\varphi|$, the integral can be estimated as

$$\frac{1}{a^2}\Big|\int\limits_{|x|=a} x_j\partial_k\varphi(x)\,dS\Big| \le \frac{M}{a}\int_{|x|=a} dS = 4\pi Ma \to 0\,.$$

The limit of the second surface integral is evaluated by changing variables $x = ay$ so that $dS_x = a^2dS_y$ and the integration is reduced to the unit sphere $|y| = 1$ so that

$$\frac{1}{a^4}\int\limits_{|x|=a} x_jx_k\varphi(x)\,dS = \int\limits_{|y|=1} y_jy_k\varphi(ay)\,dS\,.$$

The integrand has an integrable bound that is independent of a because $|\varphi(ay)| \le \sup|\varphi|$ and by the Lebesgue dominated convergence theorem the order of taking the limit and integration can be interchanged. Since $\varphi(ay) \to \varphi(0)$ for any $|y| = 1$, it remains to calculate the average of y_iy_j over the unit sphere.

In spherical coordinates $y = (\sin(\phi)\cos(\theta), \sin(\phi)\sin(\theta), \cos(\phi))$. Therefore if $i \ne j$ the average vanishes because the integral over the polar angle is zero. If $i = j$, then the average should be same for any

i because the result should not depend on the choice of the coordinate system. Taking $i = 3$, one has

$$\int_{|y|=1} y_i y_j \, dS = \delta_{ij} \int_{|y|=1} y_3^2 \, dS = \int_0^{2\pi} \int_0^{\pi} \cos^2(\phi) \sin(\phi) \, d\theta \, d\phi = \frac{4\pi}{3} .$$

Therefore

$$\lim_{a \to 0^+} \frac{1}{a^4} \int\limits_{|x|=a} x_j x_k \varphi(x) \, dS = \frac{4\pi}{3} \, \delta_{jk} \varphi(0) .$$

and

$$\partial_j \partial_k \frac{1}{|x|} = -\frac{4\pi}{3} \, \delta_{jk} \delta(x) + \mathcal{P} \left\{ \partial_j \partial_k \frac{1}{|x|} \right\} , \tag{21.12}$$

where $\mathcal{P}$ stands for the Cauchy spherical principal value regularization.

It was shown earlier that f is a harmonic function for any $x \neq 0$:

$$\left\{ \Delta \frac{1}{|x|} \right\} = 0 \quad a.e.$$

This agrees that with the laws of physics that the potential of static gravitational or electric fields is a harmonic function in any region where no sources (masses or charges) are present. By contracting the indices in (**21.12**), $\partial_j \partial_j = \Delta$, one infers that

$$\Delta \, \frac{1}{|x|} = -4\pi \delta(x) \, , \quad x \in \mathbb{R}^3 . \tag{21.13}$$

Thus, the Newton and Coulomb potentials are indeed potentials created by a point mass or a point charge if the density and the potential are understood in the distributional sense.

21.5. Distributional derivatives of a piecewise smooth function. A function is said to be piecewise from class C^p if all its partials, $D^\beta f$, $\beta \leq p$, up to order p are piecewise continuous. Clearly, the classical derivatives $\{D^\beta f\}$ exist almost everywhere and are locally integrable. Let us investigate distributional derivatives of such functions.

Consider first the case of a single real variable. Let $f'(x)$ be continuous for $x \neq a$ and $f(x)$ have a jump discontinuity at $x = a$. For brevity, put

$$\operatorname*{disc}_a [f] = \lim_{x \to a^+} f(x) - \lim_{x \to a^-} f(x) .$$

It follows from (**21.9**) that

$$f'(x) = \operatorname*{disc}_a [f] \, \delta(x - a) + \{f'(x)\} .$$

Since the classical derivative is piecewise continuous and, hence, locally integrable, the principal value regularization is not needed. The analysis is readily extended to any a piecewise C^1 function f,

$$f'(x) = \sum_n \operatorname*{disc}_{a_n}[f]\,\delta(x-a_n) + \{f'(x)\}\,. \tag{21.14}$$

The series of shifted delta-functions converges in the sense of distributions. Indeed, if $\operatorname{supp}\varphi \subset [-R,R]$, then the summation in (f',φ) has only finitely many terms for which $|a_n| < R$ and therefore exists for any φ.

Higher-order derivatives of piecewise smooth functions are obtained by applying (**21.14**) to piecewise continuous classical derivatives of f. For example,

$$f''(x) = \sum_n \operatorname*{disc}_{a_n}[f]\,\delta'(x-a_n) + \sum_n \operatorname*{disc}_{b_n}[\{f'\}]\,\delta(x-b_n) + \{f''(x)\}\,.$$

The set of discontinuities of $\{f'\}$ can be larger than that of $\{f\}$. For example, if $f(x) = \varepsilon(x)x^2 + 2\theta(x) + |x-1|$, where $\varepsilon(x)$ is the sign function. Then

$$\begin{aligned} f'(x) &= \operatorname*{disc}_{x=0}[f]\delta(x) + \{f'(x)\} = 2\delta(x) + 2|x| + \varepsilon(x-1)\,, \\ f''(x) &= 2\delta'(x) + \operatorname*{disc}_{x=1}[\{f'\}]\delta(x-1) + \{f''(x)\} \\ &= 2\delta'(x) + 2\delta(x-1) + 2\varepsilon(x)\,. \end{aligned}$$

21.5.1. Simple and double layer distributions. Let S be an $N-1$ dimensional (piecewise) smooth surface in $\mathbb{R}^N$ oriented by a unit normal vector n. Let $\nu(x)$ and $\mu(x)$ be continuous functions on S. Define the distributions

$$\begin{aligned} \Big((\mu\delta_S),\varphi\Big) &= \int_S \mu(x)\varphi(x)\,dS\,, \\ \Big(\frac{\partial}{\partial n}(\nu\delta_S),\varphi\Big) &= -\int_S \nu(x)\frac{\partial\varphi(x)}{\partial n}\,dS = -\int_S \nu(x)(\nabla\varphi, d\Sigma)\,. \end{aligned}$$

They are called a *simple layer with density* μ and a *double layer with density* ν. In physics, the simple layer distribution describes the density of electric charges distributed over a surface S with the surface density μ. The double layer distribution describes the density of electric charges created by point-like electric dipoles distributed over the surface S and dipole moments directed along the normal $\hat{n}$, and ν is the surface density of dipoles (the dipole moment created by surface area dS at a point $x \in S$ is equal to $\nu(x)d\Sigma(x)$ where $d\Sigma(x) = \hat{n}(x)dS$ is the oriented surface area element).

21.5.2. The gradient of a piecewise smooth function. Let Ω be an open set $\Omega \subset \mathbb{R}^N$ with a smooth boundary $\partial\Omega$ oriented outward by a unit normal $\hat{n}$. Suppose that $f \in C^1(\bar{\Omega})$ and f is does not vanish at the boundary $\partial\Omega$. Let f be extended to the whole $\mathbb{R}^N$ by zeros, $f(x) = 0$ for all $x \notin \bar{\Omega}$ so so that f is piecewise C^1. The classical gradient $\{\nabla f\}$ exists everywhere but the boundary $\partial\Omega$ and is locally integrable. Let us find the distributional gradient of f.

A support of any test function φ lies in an open ball B_R. Then the functions f and φ are from class $C^1(\bar{\Omega} \cap \bar{B}_R)$ and by Theorem **8.1**, the integration by parts is permitted

$$\begin{aligned}(\nabla f, \varphi) &= -(f, \nabla\varphi) = -\int_{\Omega\cap B_R} f(x)\nabla\varphi(x)\, d^N x \\ &= -\oint_{\partial\Omega_R} \hat{n}(x) f(x)\varphi(x)\, dS + \int_{\Omega\cap B_R} \{\nabla f\}\varphi(x)\, d^N x\,,\end{aligned}$$

where $\partial\Omega_R$ is the part of $\partial\Omega$ that lies in the ball B_R. The surface integral over the sphere $|x| = R$ vanishes because φ and its partials are equal to zero on it. Therefore

$$\nabla f(x) = \Big\{\nabla f(x)\Big\} - \mu(x)\delta_{\partial\Omega}(x)\,, \tag{21.15}$$

where the surface density $\mu(x)$ is defined by

$$\mu(x) = \hat{n}(x) f(x)\,, \quad x \in \partial\Omega\,.$$

There is a simple generalization of (**21.15**). Let a smooth surface S separate $\mathbb{R}^N$ into two non-connected sets Ω_1 and Ω_2 and be oriented by a unit normal $\hat{n}$. Suppose that f is piecewise C^1, that is, f and Df are continuous everywhere but S and have jump discontinuities on S. Put

$$[f]_S(x) = \lim_{a\to 0^+} f(x + a\hat{n}) - \lim_{a\to 0^+} f(x - a\hat{n})\,, \quad x \in S\,.$$

This function defines the jump discontinuity of f at any $x \in S$ relative to the orientation of S by $\hat{n}$. Then

$$\nabla f(x) = \Big\{\nabla f(x)\Big\} + \hat{n}(x)[f]_S(x)\delta_S(x)\,. \tag{21.16}$$

This is a multi-variable generalization of (**21.14**). A proof of (**21.16**) is left to the reader as an exercise.

21.5.3. Distributional Green's formula. Suppose that $f \in C^2(\bar{\Omega})$ where Ω is an open set in $\mathbb{R}^N$. Let us extend f to $\mathbb{R}^N$ by setting $f(x) = 0$ for all $x \notin \bar{\Omega}$. Then f and its first partial derivatives generally have jump discontinuities at $\partial\Omega$. Let us compute the distributional Laplacian Δf and compare it with the classical one $\{\Delta f\}$. Since f is from class C^2

in $\mathbb{R}^N \setminus \partial\Omega$, the classical and distributional Laplacians are equal in this open set:

$$\Delta f(x) = \{\Delta f(x)\}\,, \quad x \in \mathbb{R}^N \setminus \partial\Omega\,.$$

A test function φ has support in an open ball B_R. Let $\Omega_R = \Omega \cap B_R$. Then

$$(\Delta f, \varphi) = (f, \Delta\varphi) = \int_{\Omega_R} f(x)\Delta\varphi(x)\, d^N x\,.$$

The function f is from class $C^2(\bar{\Omega}_R)$ so that the classical Green's identity holds

$$\int_{\Omega_R} \Big(\{\Delta f\}\varphi - f\Delta\varphi\Big)\, d^N x = \oint_{\partial\Omega_R} \left(\frac{\partial f(x)}{\partial n}\,\varphi(x) - f(x)\frac{\partial\varphi(x)}{\partial n}\right) dS\,,$$

where n is the unit outward normal on $\partial\Omega_R$. The boundary $\partial\Omega_R$ contains a part of the boundary $\partial\Omega$ that lies in B_R and a portion of the sphere $|x| = R$ that lies in Ω. Since φ and all its partials vanish for all $|x| \geq R$ and all classical partials $\{D^\beta f\}$, $\beta \leq 2$, are locally integrable, the integration region can be extended to the whole space in the left-hand side and to the whole boundary $\partial\Omega$ in the right-hand side:

$$\int \Big(\{\Delta f\}\varphi - f\Delta\varphi\Big)\, d^N x = \oint_{\partial\Omega} \left(\frac{\partial f}{\partial n}\varphi(x) - f(x)\frac{\partial\varphi(x)}{\partial n}\right) dS\,,$$

and, hence, in the distributional sense

$$\Delta f = \{\Delta f\} - \frac{\partial f}{\partial n}\,\delta_{\partial\Omega} - \frac{\partial}{\partial n}\Big(f\delta_{\partial\Omega}\Big)\,. \tag{21.17}$$

This is the distributional Green's formula.

A relation between second distributional and classical partials of a piecewise smooth function are obtained by differentiation of (**21.16**) and using the latter to find distributional derivatives of $\{\partial_j f\}$:

$$\begin{aligned}\frac{\partial^2 f}{\partial x_i \partial x_j} &= \frac{\partial}{\partial x_i}\Big\{\frac{\partial f}{\partial x_j}\Big\} + \frac{\partial}{\partial x_i}\Big(\hat{n}_j(x)[f]_S(x)\delta_S(x)\Big)\\ &= \Big\{\frac{\partial^2 f}{\partial x_i \partial x_j}\Big\} + \hat{n}_i\Big[\frac{\partial f}{\partial x_j}\Big]_S \delta_S(x) + \frac{\partial}{\partial x_i}\Big(\hat{n}_j[f]_S\delta_S(x)\Big)\,,\end{aligned} \tag{21.18}$$

where $\hat{n}_i = (\hat{e}_i, \hat{n})$ is the cosine of the angle between the unit normal $\hat{n}$ at a point $x \in S$ and the standard basis vector $\hat{e}_i$. Formula (**21.17**) can also be obtained from (**21.18**) by setting $i = j$ and taking the sum over i. Technical details are left to the reader as an exercise.

21.6. Chain rule for distributions. Suppose a distribution $f(x)$ is defined by a C^1 function and $x = F(y)$ is a change of variables. Then the composition $f(F(y))$ is also a C^1 function and ist partial derivatives are obtained by the chain rule

$$\frac{\partial}{\partial y_j} f(F(y)) = \frac{\partial x_k}{\partial y_j} \frac{\partial}{\partial x_k} f(x)\Big|_{x=F(y)}, \tag{21.19}$$

where the summation over repeated indices is assumed. In Sec.**20** a change of variables in distributions has been defined. Can the chain rule be extended to distributions?

Let $x = F(y)$ be a change of variables such that there exists a linear and continuous transformation $T_F : \mathcal{D}(\Omega) \to \mathcal{D}(\Omega_F)$, where $\Omega_F = F(\Omega)$, as defined in Sec.**20**. Then its adjoint T^* is a linear continuous transformation of $\mathcal{D}'(\Omega_F)$ to $\mathcal{D}'(\Omega)$ that defines a change of variables in a distribution $f(F(y)) = T_F^*(f)(y)$. If $f(x)$ is a regular distribution defined by a C^1 function, then $T_F^*(f)(y)$ is just the composition of two functions and the above chain rule holds for its derivatives. Put $\mathbf{J}_{jk} = \partial F_k/\partial y_j$ which is the Jacobian matrix. The chain rule for regular distributions can be cast in the form suitable for an extension to all distributions:

$$\frac{\partial}{\partial y_j} T_F^*(f)(y) = \mathbf{J}_{jk}(y) T_F^*\Big(\frac{\partial f}{\partial x_k}\Big)(y)\,. \tag{21.20}$$

It says that the gradient of a distribution obtained by a change of variables is the Jacobian matrix multiplied by the distribution obtained by a change of variables in the gradient of the distribution.

To prove (**21.20**), recall that the space of test functions is dense in the space of distributions. Therefore for any distribution $f(x) \in \mathcal{D}'(\Omega_F)$ one can find a sequence of test functions $f_n(x) \in \mathcal{D}(\Omega_F)$ that converges to $f(x)$ in $\mathcal{D}'(\Omega_F)$. By continuity of differentiation on $\mathcal{D}'(\Omega_F)$,

$$f_n \to f \quad \Rightarrow \quad Df_n \to Df \quad \text{in } \mathcal{D}'(\Omega_F)\,.$$

By continuity of the adjoint transformation T_F^*:

$$T_F^*(f_n) \to T_F^*(f) \quad \text{and} \quad T_F^*(Df_n) \to T_F^*(Df) \quad \text{in } \mathcal{D}'(\Omega)\,.$$

Therefore by continuity of differentiation on $\mathcal{D}'(\Omega)$

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

On the other hand, the chain rule holds for distributions defined by smooth functions so that

$$\frac{\partial}{\partial y_j} T_F^*(f_n)(y) = \frac{\partial}{\partial y_j} f_n(F(y)) = \frac{\partial x_k}{\partial y_j} \frac{\partial}{\partial x_k} f_n(x)\Big|_{x=F(y)} = \mathbf{J}_{jk} T_F^*\Big(\frac{\partial f_n}{\partial x_k}\Big)\,.$$

The left-hand side of this equality converges to the left-hand side of (**21**.**20**). By construction, $F(y)$ is from class C^∞ and multiplication by a smooth function is a linear and continuous transformation of $\mathcal{D}'(\Omega)$. Therefore the right-hand side of this equality converges to the right-hand side of (**21**.**20**).

In particular, for shifted distributions

$$Df(x+a) = (Df)(x+a)\,. \tag{21.21}$$

So, shifting the argument of a distribution and differentiation commute.

In what follows, the chain rule for distributions will often be written in the form (**21**.**19**) for brevity.

21.6.1. Example. Consider a one-dimensional example from Sec. **20.3.4**. In this case, $x = F(y) = \sin(y)$, $f(x) = \delta(x)$, and $f'(x) = \delta'(x)$ so that $T_F^*(f)(y) = \delta(\sin(y))$ and $T_F^*(f')(y) = \delta'(\sin(y))$. Let us verify the chain rule (**21**.**20**). In this case, it states that

$$\frac{d}{dy}\delta(\sin(y)) = \cos(y)\delta'(\sin(y))\,.$$

To compare both sides of this relation, let us express them in terms of shifted delta functions and their derivatives. By (**20**.**7**) and (**21**.**21**), the left-hand side is transformed as

$$\frac{d}{dy}\delta(\sin(y)) = \frac{d}{dy}\sum_n \delta(y-\pi n) = \sum_n \delta'(y-\pi n)\,.$$

Note that the order of summation and differentiation can be interchanged by continuity of differentiation on $\mathcal{D}'$. Let us express $\delta'(\sin(y))$ in terms of derivatives of shifted delta functions. In this example

$$T_F(\varphi)(x) = \frac{\eta(x)}{\sqrt{1-x^2}}\sum_n \varphi\Big(\pi n + (-1)^n \arcsin(x)\Big)\,,$$

where η is a test function from $\mathcal{D}(-1,1)$ that is equal to 1 in some interval $(-1+\delta, 1-\delta)$. Then by definition of the adjoint

$$\begin{aligned}(T_F^*(\delta'),\varphi) &= (\delta', T_F(\varphi)) = -\Big(\delta, (T_F(\varphi))'\Big) = -(T_F(\varphi))'(0)\\ &= -\sum_n (-1)^n\varphi'(\pi n) = -\sum_n (-1)^n\Big(\delta(y-\pi n), \varphi'(y)\Big)\,,\end{aligned}$$

where it was used that $\eta(0)=1$ and $\eta'(0)=0$. This shows that

$$T_F^*(\delta')(y) = \delta'(\sin(y)) = \sum_n (-1)^n\delta'(y-\pi n)\,.$$

By differentiating the identity $\cos(y)\delta(y-\pi n) = (-1)^n\delta(y-\pi n)$ one infers that

$$\sin(y)\delta(y-\pi n) + \cos(y)\delta'(y-\pi n) = (-1)^n\delta'(y-\pi n)$$

Since $\sin(y)\delta(y-\pi n) = 0$, the chain rule holds:

$$\cos(y)\delta'(\sin(y)) = \cos(y)\sum_n (-1)^n\delta'(y-\pi n) = \sum_n \delta'(y-\pi n)\,.$$

21.6.2. Example. Let $f(x) = \theta(x)$ and $x = F(y) = \sin(y)$. In this case, $\theta(\sin(y))$ is a piecewise constant function that has jump discontinuities at $x = \pi n$, where n is any integer (it vanishes in $(\pi(2n-1), 2\pi n)$ and is equal to 1 in $(2n\pi, \pi(2n-1))$. The derivative of this distribution is easy to find by means of (**21**.**14**):

$$\frac{d}{dy}\theta(\sin(y)) = \sum_n (-1)^n\delta(y-\pi n)\,.$$

Since $f'(x) = \delta(x)$ and $F'(y) = \cos(y)$, the chain rule (**21**.**20**) yields

$$\frac{d}{dy}\theta(\sin(y)) = \cos(y)\delta(\sin(y)) = \cos(y)\sum_n \delta(y-\pi n)$$
$$= = \sum_n (-1)^n\delta(y-\pi n)\,,$$

as required, where the properties of the delta function stated in Secs. **20.3.4** and **18.1** are used for reducing the result to its final form.

21.6.3. General distributional solution to the 2D wave equation. Let f and g be distributions from $\mathcal{D}'(\mathbb{R})$. Then the distribution of two variables

$$u(x,t) = f(x+ct) + g(x-ct) \qquad (\mathbf{21}.\mathbf{22})$$

satisfies the 2D wave equation in the distributional sense

$$\Box_2 u(x,t) \overset{\text{def}}{=} \frac{\partial^2}{\partial t^2}\,u(x,t) - c^2\frac{\partial^2}{\partial x^2}\,u(x,t) = 0\,.$$

Consider the distribution of two real variables $x_\pm$

$$v(x_+, x_-) = f(x_+) + g(x_-) \in \mathcal{D}'(\mathbb{R}^2)\,.$$

Here f and g are distributions from $\mathcal{D}'(\mathbb{R}^2)$ that are independent of one of the variables (see Sec.**17.3**). Then

$$u(x,t) = v(x+ct, x-ct)\,.$$

The transformation $x_\pm = x \pm ct$ is a change of variables in $\mathbb{R}^2$ from class C^∞ whose Jacobian does not vanish anywhere. Therefore the distributional chain rule applies for computing distributional derivatives:

$$\frac{\partial}{\partial t}u(x,t) = \frac{\partial x_+}{\partial t}\,f' + \frac{\partial x_-}{\partial t}\,g' = cf'(x+ct) - cg'(x-ct)\,,$$
$$\frac{\partial^2}{\partial t^2}u(x,t) = \frac{\partial}{\partial t}\Big(\frac{\partial}{\partial t}u(x,t)\Big) = c^2f''(x+ct) + c^2g''(x-ct)\,,$$
$$\frac{\partial^2}{\partial x^2}u(x,t) = \frac{\partial}{\partial x}\Big(\frac{\partial}{\partial x}u(x,t)\Big) = f''(x+ct) + g''(x-ct)\,,$$

and the assertion follows. The converse is also true, that is, *any distributional solution to the 2D wave equation can written in the form (**21.22**).* A proof of this assertion is postponed until antiderivatives (indefinite integrals) of distributions are introduced.

21.7. The structure of distributions. Examples from Sec.**21.4** suggest many distributions can be obtained by differentiating continuous functions in the distributional sense. Even singular distributions are derivatives of continuous functions. For example, $|x|'' = \delta(x)$ if $x \in \mathbb{R}$ or $\Delta^2|x| = -8\pi\delta(x)$ if $x \in \mathbb{R}^3$. So, it is natural to ask if any distribution can be obtained in this way? It turns out that all compactly supported distributions can be written as linear combinations of distributional derivatives of some continuous functions, and any distribution has this structure *locally*, that is, on any bounded open set. This assertion is known as the *structure theorem for distributions* [3].

Theorem **21.1**. *For any $f \in \mathcal{D}'$ and any compact $K \subset \mathbb{R}^N$, there exists a finite collection of continuous functions g_α supported in a neighborhood of K such that*

$$(f,\varphi) = \sum_{\alpha\le n}(g_\alpha, D^\alpha\varphi)\,,\quad \varphi\in\mathcal{D}\,,\quad \operatorname{supp}\varphi \subseteq K\,,$$

for some integer $n \ge 0$ (that depends on K). If, in addition, f is compactly supported, $\operatorname{supp} f = K$*, then the above relation holds for all test functions in $\mathcal{D}$.*

This theorem is not so much about practical importance, but of some theoretical interest. It states that *there is a limit to how bad the singularities of distributions can be, they are not worse than repeated derivatives of continuous functions.* In the next section, this assertion

[3] G.Grubb, Distributions and operators, Chapter 3;
A.E. Kinani and M. Oudadess, Distribution theory and applications, Chapter 7

will be proved for periodic distributions by an explicit construction of g.

21.7.1. Distributions with a point support. Distributions with a point support have a remarkable structure.

THEOREM **21.2**. *If support of a distribution f consists of a single point $x = 0$, then f is given by a unique linear combination of the delta function and its derivatives:*

$$f(x) = \sum_{\beta=0}^{m} c_\beta D^\beta \delta(x)\,. \tag{21.23}$$

The assertion is a corollary of the following property of compactly supported distributions. *For any compactly supported distribution $f \in \mathcal{D}'$, there exist a constant M and an integer m such that*

$$|(f,\varphi)| \le M\|\varphi\|_{C^m}\,, \quad \|\varphi\|_{C^m} = \sup_{\beta\le m,\,x} |D^\beta\varphi(x)|\,, \tag{21.24}$$

for all test functions $\varphi \in \mathcal{D}$. It is established by contradiction. If no such M and m exist, then one can find a sequence of test functions $\{\varphi_n\}_1^\infty$ such that

$$|(f,\varphi_n)| \ge n\|\varphi_n\|_{C^n}\,.$$

Let η be a bump function for the support of f. Then $\eta \in \mathcal{D}$ because f is compactly supported, and $(f,\varphi_n) = (f,\eta\varphi_n)$. The sequence $\psi_n = \eta\varphi_n/(\sqrt{n}\|\varphi\|_{C^n})$ is a null sequence in $\mathcal{D}$. Indeed, supports of all ψ_n lie in $\operatorname{supp}\eta$. Since $\sup|D^\alpha\eta| \le C_\alpha < \infty$ for any $\alpha \ge 0$, using the binomial expansion of repeated derivatives of the product, one infers that

$$\sup|D^\beta\psi_n| \le \sum_{\alpha=0}^{\beta}\binom{\beta}{\alpha} C_\alpha \frac{\sup|D^\alpha\varphi_n|}{\sqrt{n}\|\varphi_n\|_{C^n}} \le \frac{C}{\sqrt{n}} \to 0$$

as $n\to\infty$ because $\sup|D^\alpha\varphi_n| \le \|\varphi_n\|_{C^n}$ for all $n \ge \alpha$. By continuity of f, $(f,\psi_n)\to 0$ as $n\to\infty$. However, by linearity of f

$$|(f,\psi_n)| = \frac{|(f,\varphi_n)|}{\sqrt{n}\|\varphi_n\|_{C^n}} \ge \sqrt{n}\,,$$

and, hence, a contradiction.

Let us show first that if a distribution f supported at $x = 0$ has a representation (**21.23**), then the coefficients c_β are unique. Let η be a bump function for $\{x = 0\} = \operatorname{supp} f$ such that $\eta(x) = 0$ when $|x| > 1$.

Suppose that there exists a different set of coefficients c'_β for which (**21.23**) holds. Then for a test function $\varphi = x^\alpha \eta$

$$0 = (f - f, \varphi) = \sum_\beta^m (c_\beta - c'_\beta)(D^\beta \delta, \varphi) = \sum_\beta^m (-1)^\beta (c_\beta - c'_\beta)(\delta, \eta D^\beta x^\alpha)$$
$$= (-1)^\alpha \alpha! (c_\alpha - c'_\alpha)$$

because $D^\alpha \eta(0) = 0$ for any $\alpha > 0$. Therefore $c'_\beta = c_\beta$.

Next, for any test function φ, there exist a constant M and an integer m such that (**21.24**) holds. The constant M and the integer m are specific for f but independent of φ. Consider a sequence of test functions $\psi_n(x) = \eta(nx)[\varphi(x) - p_m(x)]$ where $p_m(x)$ is the Taylor polynomial of order m for φ about $x = 0$. One has

$$(f, \varphi) = (\eta f, \varphi) = (f, \eta\varphi) = (f, \psi_1 + \eta p_m)$$
$$= (f, \psi_1) + \sum_{\beta=0}^m \frac{D^\beta \varphi(0)}{\beta!}(f, x^\beta \eta)$$

The conclusion of the theorem follows with $c_\beta = (-1)^\beta (f, x^\beta \eta)/\beta!$ if one can show that $(f, \psi_1) = 0$. To prove the latter, note that $(f, \psi_1) = (f, \psi_n)$ for any $n \geq 1$. Therefore

$$(f, \psi_1) = \lim_{n\to\infty} (f, \psi_n)\,.$$

It remains to show that the above limit vanishes. This is indeed so. One has

$$D^\alpha[\varphi(x) - p_m(x)] = O(|x|^{m+1-\alpha})\,, \quad |x| \to 0\,,$$
$$\sup |D^\gamma \eta(nx)| = n^\gamma \sup |D^\gamma \eta|\,, \quad n \to \infty\,.$$

Since $\eta(nx)$ is supported in the ball $|x| \leq \frac{1}{n}$, it follows from the above estimates and (**21.24**) that

$$|(f, \psi_n)| \leq M \sup_{\beta \leq m, |x| \leq \frac{1}{n}} |D^\beta(\eta(nx)[\varphi(x) - p_m(x)])|$$
$$\leq M \sup_{\beta \leq m, |x| \leq \frac{1}{n}} \sum_{\alpha=0}^\beta \binom{\beta}{\alpha} |D^{\beta-\alpha}\eta(nx)||D^\alpha[\varphi(x) - p_m(x)]|$$
$$\overset{(1)}{\leq} \max_{\beta \leq m} M_\beta n^{\beta - m - 1} = \frac{M_m}{n} \to 0\,,$$

as $n \to \infty$, as required. Here to obtain (1), note that $\alpha \leq \beta \leq m$ so that $\alpha < m+1$ and the supremum of the asymptotic form of $D^\alpha(\varphi - p_m)$ is reached at $|x| = \frac{1}{n}$ as $n \to \infty$ so that the asymptotic behavior of the product in the binomial expansion does not depend on α because

$n^{\beta-\alpha-m-1+\alpha} = n^{\beta-m-1}$ and summation over α produces a constant that depends on β. As $n \to \infty$, the maximum over β is reached for $\beta = m$.

21.8. Exercises.

1. Find the distributional derivative of

$$f(x) = \frac{x}{|x|} + x\sin\left(\frac{1}{x}\right), \quad x \neq 0\,,$$

and compare it with the classical derivative. Is $\{f'\}$ locally integrable?

2. Prove that

$$\frac{d}{dx}\frac{1}{x \pm i0^+} = \mp i\pi\delta'(x) - \mathcal{P}\frac{1}{x^2}$$

3. Express each of the following distributions in terms of linear combinations (or series) of the delta-function and its derivatives:

$$\text{(i)} \quad \sin(x)\delta'(x)$$
$$\text{(ii)} \quad x^2\delta''(x)$$

4. Use properties of the principal value distribution to find a distributional solution to each of the following distributional equations

$$\text{(i)} \quad xf'(x) = 1$$
$$\text{(ii)} \quad xf'(x) = \mathcal{P}\frac{1}{x}$$
$$\text{(iii)} \quad x^2f'(x) = 1$$

5. Let $\delta_{S_a}(x)$ be the spherical delta-function. If $r = |x|$, show that the partial derivative of $\delta_{S_a}(x)$ with respect to r is a distribution and express

$$(|x|^2 - a^2)\frac{\partial}{\partial r}\delta_{S_a}(x)$$

in terms of δ_{S_a}.

6. If $\chi_{B_a}(x)$ is the characteristic function of the ball of radius a centered at the origin show that its distributional gradient is given by

$$\nabla\chi_{B_a}(x) = -\frac{x}{a}\,\delta_{S_a}(x)\,.$$

7. Let $f(x) = \ln(|x|)$ where $x \in \mathbb{R}^2$.
(i) Show that $f(x)$ is a harmonic function wherever it is twice continuously differentiable (that is, outside of any neighborhood of $x = 0$):

$$\Big\{\Delta \ln(|x|)\Big\} = 0$$

(ii) Use the same method as for the Laplacian of the Coulomb potential in $\mathbb{R}^3$ to find $\Delta \ln(|x|)$ in $\mathbb{R}^2$ in the distributional sense.

8. (i) Show that that the following classical Laplacian vanishes

$$\left\{\Delta \frac{1}{|x|^{N-2}}\right\} = 0 \quad \text{a.e.}$$

in $\mathbb{R}^N$.
(ii) Show that

$$\Delta \frac{1}{|x|^{N-2}} = C\,\delta(x) \quad \text{in } \mathcal{D}'$$

and find the constant C.

(iii) Find continuous functions $g_\alpha(x)$ such that

$$\sum_{\alpha \le n} D^\alpha g_\alpha(x) = \delta(x)\,, \quad x \in \mathbb{R}^N\,.$$

9. Prove (**21**.**16**).

10. Prove (**21**.**18**).

11. Let Ω be a disk in $\mathbb{R}^2$, $|x| < a$. Consider a transformation $x = F(y)$ from class C^∞ defined by $x_1 = r\cos(\phi)$, $x_2 = r\sin(\phi)$ where $y = (r, \phi) \in \mathbb{R}^2$.
(i) If $\theta_\Omega(x)$ is the characteristic function of Ω, find partial derivatives of $\theta_\Omega(F(y))$ with respect to r and ϕ.
(ii) Find partial derivatives of $\delta(F(y) - x_0)$, where $x_0 \neq 0$, with respect to r and ϕ. Express the answer in terms shifted delta functions in the variables r and ϕ.
12. Show that $\delta^{(n)}(-x) = (-1)^n \delta^{(n)}(x)$, $n = 0, 1, ...$

13. Let f be from class $C^1(\mathbb{R})$. Find the distributional gradient of the function $g(x) = f(|x|)\theta(R - |x|)$, $x \in \mathbb{R}^N$.

14. Let $x_0 \in \mathbb{R}$ and $x \in \mathbb{R}^N$. Find the first partials of the distributions:
(i) $f(x_0, x) = \theta(x_0)\theta(x_0^2 - |x|^2)$
(ii) $g(x_0, x) = h(x_0, x)\theta(x_0^2 - |x|^2)$ where h is a smooth function.

22. Series with distributions

22.1. Distributional derivatives of a functional series. In applications, one often deals with functional series whose terms are locally integrable function (or regular distributions). Differentiability of the sum of such series is difficult to verify. Typically, uniform convergence of the series of derivatives of terms is to be verified which is sufficient for differentiability. Distributional derivatives exist under much weaker conditions. All it really takes is just to show that the series converges in the distribution sense because differentiation is continuous on $\mathcal{D}'$ and any convergent series can be differentiated term-by-term to obtain the derivative of the sum (see Sec.**21.2.4**). The following theorem is quite useful for this task.

THEOREM **22.1**. *Let $\{u_n\}$ be a sequence of locally integrable functions, and the series*

$$\sum_{n=1}^{\infty} u_n(x) = f(x)$$

converges uniformly on any compact $K \subset \mathbb{R}^N$, that is,

$$\lim_{m\to\infty} \sup_K \Big| \sum_{n=1}^{m} u_n(x) - f(x) \Big| = 0\,.$$

Then the sum $f(x)$ is a distribution and its distributional derivatives are given by

$$D^\beta f(x) = \sum_{n=1}^{\infty} D^\beta u_n(x)\,.$$

To prove the assertion, put

$$f_m(x) = \sum_{n=1}^{m} u_n(x) \ \in \mathcal{L}_{\text{loc}}$$

Each term of the sequence of partial sums is a locally integrable function as a finite sum of locally integrable functions. Hence, $f_m(x)$ is a regular distribution and for any test function supported in a ball B_R

$$(f_m, \varphi) = \sum_{n=1}^{m} \int_{B_R} u_n(x)\varphi(x)\, d^N x\,.$$

Let us show that this numerical sequence converges and, by the completeness theorem, the sum $f(x)$ is a distribution defined by the rule

$$(f, \varphi) = \lim_{m\to\infty} (f_m, \varphi) = \sum_{n=1}^{\infty} \int u_n(x)\varphi(x)\ d^N x\,.$$

The sequence (f_m, φ) is a Cauchy sequence. Indeed, one has

$$|(f_m, \varphi) - (f_k, \varphi)| \le \sup_{B_R} |f_m(x) - f_k(x)| \int_{B_R} |\varphi(x)|\, d^N x \,.$$

By the hypothesis of uniform convergence of the series and by the Cauchy criterion for uniform convergence (Theorem **1.5.2**), the factor at the integral can be made arbitrary small for large enough k and m and stays arbitrary small for all large enough k and m:

$$\sup_{B_R} |f_m(x) - f_k(x)| \to 0 \quad \text{as} \quad m, k \to \infty \,.$$

Therefore, (f_m, φ) is a Cauchy sequence for any test function and, hence, converges.

Since differentiation D^β is linear and continuous on the space of distributions:

$$D^\beta f = D^\beta \lim_{m\to\infty} \sum_{n=1}^{m} u_n(x) = \lim_{m\to\infty} \sum_{n=1}^{m} D^\beta u_n(x) = \sum_{n=1}^{\infty} D^\beta u_n(x)$$

where $D^\beta u_n$ is the distributional derivative of a locally integrable function.

In contrast to the classical theorem of differentiation of a functional series, the uniform convergence of the series of derivatives of terms is not required to change the order of summation and differentiation. In fact, classical derivatives $D^\beta u_n$ may not even exist.

COROLLARY **22.2**. *Let $\{u_k(x)\}$ be a sequence of functions that are bounded on an open set $\Omega \subseteq \mathbb{R}^N$ almost everywhere and the series of bounds converges,*

$$|u_k(x)| \le M_k \quad \text{a.e.}\,, \qquad \sum_k M_k < \infty \,.$$

Then the series $\sum_k u(x)$ converges to a regular distribution $f(x)$ in $\mathcal{D}'(\Omega)$ such that

$$(f, \varphi) = \int f(x)\varphi(x)\, d^N x = \sum_{k=1}^{\infty} \int u_k(x)\varphi(x)\, d^N x \,, \quad \varphi \in \mathcal{D}(\Omega)$$

and

$$D^\alpha f(x) = \sum_k D^\alpha u_k(x) \quad \text{in } \mathcal{D}'(\Omega) \,.$$

Under the hypothesis, the series converges uniformly because

$$\Big| \sum_{n=1}^{m} u_n(x) - f(x) \Big| \le \sum_{n=m+1}^{\infty} |u_n(x)| \le \sum_{n=m+1}^{\infty} M_n \to 0$$

as $m \to \infty$ for all x, and the conclusion follows from Theorem **22.1**. Corollary **22.2** is often easier to apply than Theorem **22.1**.

22.1.1. Example. Put

$$u_n(x) = \frac{e^{inx}}{n^p}\,.$$

The series $\sum_n u_n(x)$ converges uniformly on $\mathbb{R}$ for any $p > 1$ because the series of upper bounds of terms converges:

$$M_n = \sup|u_n(x)| = \frac{1}{n^p}\,, \quad \sum_{n=1}^{\infty} M_n < \infty\,, \quad p > 1\,.$$

Since $u_n(x)$ are continuous, there exists a continuous $f(x)$ such that

$$f(x) = \sum_{n=1}^{\infty} \frac{e^{inx}}{n^p}\,.$$

It is locally integrable on $\mathbb{R}$ and defines a distribution by the rule

$$(f,\varphi) = \sum_{n=1}^{\infty} \frac{1}{n^p} \int e^{inx}\varphi(x)\,dx\,, \quad \varphi \in \mathcal{D}(\mathbb{R})\,.$$

Its distributional derivative exists for any order β and is given by the series

$$f^{(\beta)}(x) = \sum_{n=1}^{\infty} n^{\beta-p} i^{\beta} e^{inx}$$

Note that for $\beta \geq p$, the terms of the series are not decreasing with increasing n:

$$\left| n^{\beta-p} i^{\beta} e^{inx} \right| = n^{\beta-p}\,.$$

The series does not converge in the classical sense, but it converges in the distributional sense and its sum is the distributional derivative of $f(x)$.

22.2. Distributional Fourier series. Here the classical and distributional convergence of trigonometric Fourier series is compared. First, some basic classical results are reviewed[4].

[4]see, e.g., G.H. Hardy and W.W. Rogosinski, Fourier series, Cambridge Univ. Press, 1950 (2nd Edition).

22.2.1. Classical Fourier series. Recall the classical theorem about trigonometric Fourier series. Let $f(x)$ be a periodic function

$$f(x+2\pi) = f(x)$$

that is integrable on $(0, 2\pi)$. Then the series

$$f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}\,, \qquad a_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\,dx$$

is called the *Fourier trigonometric series* of f. The sign $\sim$ is to indicate that the series is formal because nothing is known about its convergence. If the series converges, then $a_n \to 0$ as $|n| \to \infty$.

22.2.2. Convergence of a Fourier series. Suppose that for a given x, there are constants $\delta > 0$ and $M < \infty$ such that

$$|f(x+y) - f(x)| \le M|y|$$

for all $|y| < \delta$. Then its trigonometric Fourier series converges to $f(x)$:

$$\lim_{m\to\infty} \sum_{|n|\le m} a_n e^{inx} = f(x)\,.$$

Note that a sufficient condition for the pointwise convergence of the series is stronger than a mere continuity of f. In fact, continuous functions whose Fourier series diverges at a particular point form a dense subset in periodic function from class C^0. However, the Fourier series of any periodic continuous function f is proved to converge to f almost everywhere (*Carleson's theorem*).

If f is continuous and the derivative f' is locally integrable on $\mathbb{R}$ (or f is absolutely continuous), then the Fourier series converges uniformly to f. If f is piecewise continuous and $f' \in \mathcal{L}_{\rm loc}$, then the Fourier series of f converges to f uniformly on any closed interval of continuity of f, and for any x

$$\textbf{(22.1)} \qquad \lim_{m\to\infty} \sum_{|n|\le m} a_n e^{inx} = \frac{1}{2}\Big(\lim_{y\to x^+} f(y) + \lim_{y\to x^-} f(y)\Big)\,.$$

22.2.3. Convergence in the mean. Let $f \in \mathcal{L}_2(0, 2\pi)$. Then it is proved that the Fourier series of f converges to f in the $\mathcal{L}_2$ norm:

$$\lim_{m\to\infty} \int_0^{2\pi} |f(x) - S_m(x)|^2 dx = 0\,, \quad S_m(x) = \sum_{|n|\le m} a_n e^{inx}\,,$$

so that

$$\|f\|_2^2 = \int_0^{2\pi} |f(x)|^2\,dx = \sum_n |a_n|^2$$

Conversely, if a complex sequence $\{a_n\}$ is square summable

$$\sum_{n=-\infty}^{\infty} |a_n|^2 < \infty$$

then by Bessel's theorem there exists a function $f(x)$ that is square integrable on $(0, 2\pi)$ such that

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx} \quad a.e.$$

and a_n are the Fourier coefficients of f.

22.2.4. Differentiation of trigonometric Fourier series. For any convergent Fourier series $a_n \to 0$ as $|n| \to \infty$. The smoother is the function, the faster its Fourier coefficients vanish with increasing $|n|$.

Suppose that f is a periodic function from class C^p. Then by integration by parts

$$a_n = \frac{1}{2\pi (in)^p} \int_0^{2\pi} f^{(p)}(x)\, e^{-inx}\, dx$$

It follows from this relation that

$$|a_n| \le \frac{M}{n^p}, \qquad M = \frac{1}{2\pi} \int_0^{2\pi} |f^{(p)}(x)|\, dx\,.$$

Let $a_n^{(p)}$ be the Fourier coefficients of $f^{(p)}$. Then $a_n^{(p)} = (in)^p a_n$. The Fourier series of f cannot be differentiated p times term-by-term to obtain $f^{(p)}$ because the series of derivatives

$$\sum_n a_n D^p e^{inx} = \sum_n a_n^{(p)} e^{inx}\,, \quad D = \frac{d}{dx}\,,$$

may not even converge everywhere (it converges to $f^{(p)}$ almost everywhere). To ensure the validity of the term-by-term differentiation, one can demand, in addition, that $f^{(p+1)}$ is locally integrable. Then the above series of derivatives converges uniformly to $f^{(p)}$ and the order of summation and differentiation can be interchanged:

$$D^p \sum_n a_n e^{inx} = \sum_n a_n D^p e^{inx} = \sum_n (in)^p a_n e^{inx} = f^{(p)}(x)\,.$$

In particular, this relation holds for any $p > 0$ if f is from C^∞ and periodic.

22.2.5. Distributional convergence of Fourier series. In contrast to the classical case, the distributional convergence of Fourier series does not even require any decay of a_n for large $|n|$ but imposes a restriction on the growth of the coefficients with increasing $|n|$.

THEOREM **22.2**. (Fourier series as a distribution)
In order for a trigonometric Fourier series to converge in $\mathcal{D}'$,

$$\sum_n a_n e^{inx} = f(x) \ \in \ \mathcal{D}'(\mathbb{R})\,, \tag{22.2}$$

it is necessary and sufficient that

$$|a_n| \leq A|n|^k + B$$

for some constants A, B, and k. In this case,

$$D^\beta f(x) = \sum_n (in)^\beta a_n e^{inx}$$

for any β.

Suppose that a_n grows at most by a power law with increasing $|n|$, then a trigonometric Fourier series with coefficients $a_n/|n|^p$ would converge uniformly for large enough p. Hence, its sum is a regular distribution defined by a continuous function. By continuity of differentiation, the series can be differentiated p times to obtain a distribution defined by the Fourier series with coefficients a_n. Thus, given k, consider the function

$$g(x) = \frac{a_0 x^{k+2}}{(k+2)!} + \sum_{n\neq 0} \frac{a_n}{(in)^{k+2}}\, e^{inx}\,. \tag{22.3}$$

The series converges uniformly because the series of upper bounds of terms converges

$$\sum_{n\neq 0} \left|\frac{a_n}{(in)^{k+2}}\, e^{inx}\right| = \sum_{n\neq 0} \frac{|a_n|}{|n|^{k+2}} \leq \sum_{n\neq 0} \left(\frac{A}{n^2} + \frac{B}{|n|^{k+2}}\right) < \infty\,.$$

By Corollary **22.2**, g is a regular distribution defined by a continuous function, and its distributional derivative $D^{k+2}g$ reads

$$\frac{d^{k+2}}{dx^{k+2}}\, g(x) = \sum_n a_n e^{inx} = f(x)\,. \tag{22.4}$$

By continuity of differentiation on the space of distributions

$$D^\beta f(x) = \sum_n (in)^\beta a_n e^{inx}\,.$$

To show that the converse is true, let us construct a test function for which the series $\sum_n a_n(e^{inx},\varphi)$ does not converge if c_n grow faster than a power law. Since $|n|^k|a_n|\to\infty$ for any $k>0$ as $|n|\to\infty$, there exists a subsequence of integers $\{n_k\}_1^\infty$ such that $|a_{n_k}|>|n_k|^k$. Consider the function defined by the Fourier series

$$h(x)=\sum_{k=1}^{\infty}\frac{e^{-in_kx}}{n_k^k}\,.$$

This function is from class C^∞ because the series of derivatives of terms converges uniformly:

$$\sum_k\Big|\frac{(-in_k)^m e^{-in_kx}}{n_k^k}\Big|=\sum_k\frac{1}{|n_k|^{k-m}}<\infty$$

for any $m=1,2,....$ If $\eta(x)$ is a bump function for the interval $[0,2\pi]$ supported in $[-\varepsilon,2\pi+\varepsilon]$, where $\varepsilon>0$ can be arbitrary small, then $\varphi(x)=h(x)\eta(x)$ is a test function. Therefore

$$\begin{aligned}\sum_{|n|\le m}a_n(e^{inx},\varphi)&=\sum_{|n|\le m}a_n\Big(\int_0^{2\pi}+\int_{-\varepsilon}^{0}+\int_{2\pi}^{2\pi+\varepsilon}\Big)e^{inx}\eta(x)h(x)\,dx\\&=\sum_{|n|\le m}a_n\Big(\int_0^{2\pi}e^{inx}h(x)\,dx+O(\varepsilon)\Big)\\&=\sum_{|n|\le m}a_n\Big(\sum_{k=1}^{\infty}\frac{2\pi}{n_k^k}\,\delta_{nn_k}+O(\varepsilon)\Big)\\&=2\pi\sum_{|n_k|\le m}\Big(\frac{a_{n_k}}{n_k^k}+a_{n_k}O(\varepsilon)\Big)\,.\end{aligned}$$

The third equality is obtained by interchanging the order of summation and integration, which is possible because the series for $h(x)$ converges uniformly. The limit $m\to\infty$ does not exist because $|a_{n_k}|>|n_k|^k$, and the series diverges. This completes the proof.

22.2.6. Poisson summation formula. Let us show that in the sense of distributions

$$\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}e^{inx}=\sum_{n=-\infty}^{\infty}\delta(x-2\pi n)\,.$$

This relation is known as the distributional Poisson summation formula.

To prove it, consider the function

$$f(x) = \frac{1}{2} - \frac{x}{2\pi}\,, \quad 0 < x < 2\pi\,,$$

that is extended periodically

$$f(x + 2\pi) = f(x)$$

to the whole real axis. The function f is continuous and has jump discontinuities at $x = 2\pi n$ where n is an integer. Since the integral of f over $(0, 2\pi)$ vanishes, the antiderivative of f

$$g(x) = \int_0^x f(y)\,dy = \frac{x}{2} - \frac{x^2}{4\pi}\,, \quad 0 \le x \le 2\pi\,,$$

is continuous and 2π periodic on $\mathbb{R}$:

$$g(x + 2\pi) = g(x)\,.$$

Its second distributional derivative can be found by relation (**21**.**14**):

$$\begin{aligned} g'(x) &= \{g'(x)\} = f(x) \\ g''(x) &= \{f'(x)\} + \sum_n \operatorname*{disc}_{2\pi n}[f]\,\delta(x - 2\pi n) \\ &= -\frac{1}{2\pi} + \sum_n \delta(x - 2\pi n) \end{aligned}$$

because the classical derivative of f is equal to $-\frac{1}{2\pi}$ and does not exists at $x = 2\pi n$ where f has jump discontinuities such that $\operatorname{disc}[f] = 1$ at $x = 2\pi n$.

On the other hand, $g(x)$ is continuous and periodic, and $g'(x)$ is locally integrable, therefore its Fourier series converges to $g(x)$ uniformly:

$$g(x) = \frac{\pi}{6} - \frac{1}{2\pi}\sum_{n\neq 0}\frac{1}{n^2}\,e^{inx}$$

and its sum defines a regular distribution. Hence, $g''(x)$ can be found by means of Theorem **22**.**2**:

$$g''(x) = \frac{1}{2\pi}\sum_{n\neq 0} e^{inx}$$

in the distributional sense. The Poisson summation formula is obtained by equating the two expressions for $g''(x)$.

22.3. Periodic extension of a compactly supported distribution. Any distribution defined by a convergent trigonometric Fourier series is periodic. For example, the Fourier series in the Poisson summation formula defines a 2π periodic extension of $\delta(x) \in \mathcal{D}'(-\pi, \pi)$ to $\mathbb{R}$.

Let us show that for any compactly supported distribution $f \in \mathcal{D}'$ one can define a periodic distribution with a period $T > 0$ by the rule

$$f_T(x) = \sum_n f(x + nT)\,. \tag{22.5}$$

One has to show first that the series converges in $\mathcal{D}'$. Then one must verify periodicity of the distribution f_T. For any test function φ, put

$$\phi_m(x) = \sum_{|n|\leq m} \varphi(x - nT)\,.$$

If support of φ lies in $[-R, R]$, then $\varphi(x - nT) = 0$ for all n such that $|x - nT| > R$. Therefore the series converges for any x, $\phi_m(x) \to \phi(x)$ as $m \to \infty$. On any bounded interval, $\phi(x)$ is also defined by a finite sum of test functions and hence ϕ is a C^∞ function. Furthermore, it follows from this observation that $D^\beta \phi_m \to D^\beta \phi$ uniformly on any compact as $m \to \infty$ for any $\beta > 0$. Therefore for any test function $\eta \in \mathcal{D}$, the sequence of test functions $\eta\phi_m$ converges to a test function $\eta\phi$ in topology of $\mathcal{D}$ because $D^\alpha \eta D^\beta \phi_m \to D^\alpha \eta D^\beta \phi$ uniformly for any α and β and supports of all $\eta\phi_m$ lie in support of η.

Now the existence of the distribution f_T follows from continuity and linearity of the functional f. Indeed, let $\eta_f \in \mathcal{D}$ and $\eta_f(x) = 1$ in a neighborhood of $\operatorname{supp} f$ so that $\eta_f f = f$. Then

$$\sum_{|n|\leq m} \Big(f(x{+}nT), \varphi(x)\Big) = \sum_{|n|\leq m} \Big(f(x), \varphi(x{-}nT)\Big) = (f, \phi_m) = (f, \eta_f \phi_m)\,.$$

Since $\eta_f \phi_m \to \eta_f \phi$ in $\mathcal{D}$, the limit $m \to \infty$ exists by continuity of f and

$$(f_T, \varphi) = (f, \eta_f \phi)\,, \quad \phi(x) = \sum_n \varphi(x - nT)\,, \varphi \in \mathcal{D}\,. \tag{22.6}$$

By periodicity $\phi(x - T) = \phi(x)$,

$$\begin{aligned}(f_T(x + T), \varphi(x)) &= (f_T(x), \varphi(x - T)) = (f(x), \eta_f(x)\phi(x - T)) \\ &= (f(x), \eta_f(x)\phi(x)) = (f_T(x), \varphi(x))\,,\end{aligned}$$

for any $\varphi \in \mathcal{D}$, which means that f_T is a periodic distribution with period T. One should point out that the series (**22.5**) generally diverges for distributions that are not compactly supported, e.g., $f(x) = e^x$.

Suppose that $\operatorname{supp} f \subset (0, T)$. In this case, the rule (**22.6**) defines a periodic *extension* of $f \in \mathcal{D}'(0, T)$ to $\mathbb{R}$. Indeed, if $\varphi \in \mathcal{D}(0, T)$ in (**22.6**), then $(f_T, \varphi) = (f, \varphi)$. This means that f_T is an extension of f:

$$\operatorname{supp} f \subset (0, T) \quad \Rightarrow \quad f_T(x) = f(x)\,, \quad x \in (0, T)\,.$$

Suppose $f \in \mathcal{L}(0, b)$ and $f(x) = 0$ for all $x \notin [0, b]$, then f_T defined by the series (**22.5**) generally does not coincide with f in $(0, b)$ if $T < b$ even in the distributional sense. But f_T is a periodic extension of f if $T > b$. Even when $T = b$, f_T is a periodic extension of f in the distributional sense because two locally integrable functions that are equal almost everywhere define the same distribution (the distribution f_T does not depend on the values $f(0)$ and $f(T)$). However, singular distributions can have point supports. For this reason, f_T may not be an extension of f, unless $\operatorname{supp} f$ lies in an *open* interval of length T. For example, let $f(x) = \delta(x) - \delta(x - T)$. Then $f_T(x) = \sum_n f(x + nT) = 0$.

22.4. Expansion of a periodic distribution into a Fourier series. Any distribution defined by a convergent trigonometric Fourier series is periodic. How about the converse? Can a periodic distribution be expanded into a convergent trigonometric Fourier series? The answer is affirmative.

To simplify the discussion, any periodic distribution is assumed to have period 2π. If a distribution has period T, then one can always make a scaling transformation of the argument to change the period to 2π. If $f(x + T) = f(x)$, then the distribution $f(\frac{Tx}{2\pi})$ is 2π periodic.

THEOREM **22.3**. *Let f be a periodic distribution with period 2π. Then*

$$f(x) = \sum_n a_n e^{inx}\,, \quad a_n = \frac{1}{2\pi}\Big(f(x), \eta(x)e^{-inx}\Big)\,, \tag{22.7}$$

where η is any test function with the characteristic property

$$\sum_n \eta(x + 2\pi n) = 1\,. \tag{22.8}$$

*Conversely, if f is given by a convergent Fourier series (**22.2**), then $(f(x), \eta(x)e^{-inx}) = 2\pi a_n$ for any test function η satisfying (**22.8**). So, any periodic distribution is uniquely defined by its Fourier coefficients.*

Let us first show that a test function η with the stated characteristic property exists. Recall that if χ is a characteristic function of an interval (a, b), then the regularization $\omega_\varepsilon * \chi$, where ω_ε is a hat function, is a test function that takes values between 0 and 1 so that

$\omega_\varepsilon * \chi(x) = 1$ if $a+\varepsilon < x < b-\varepsilon$. In the limit $a \to -\infty$ and $b \to \infty$, $\omega_\varepsilon * \chi(x) \to 1$ for any x. Let $\chi_m(x)$ be a characteristic function for an interval $(c - 2\pi m, c + 2\pi(m+1))$ where c is any number and m is a non-negative integer. Then $\omega_\varepsilon * \chi_m(x) \to 1$ as $m \to \infty$ for any x. On the other hand,

$$\chi_m(x) = \sum_{|n|\le m} \chi_0(x + 2\pi n) \quad a.e.$$

Therefore

$$1 = \lim_{m\to\infty} \omega_\varepsilon * \chi_m(x) = \sum_n \omega_\varepsilon * \chi_0(x + 2\pi n)\,.$$

So, one can take $\eta = \omega_\varepsilon * \chi_0$. It is clear that $\eta = \phi * \chi_0$ would also satisfy (**22.8**) for any normalized test function ϕ, meaning that $\phi * 1 = \int \phi(x)dx = 1$.

If f is locally integrable and 2π periodic. Then

$$\begin{aligned} a_n &= \frac{1}{2\pi}\int f(x)\eta(x)e^{-inx}dx = \frac{1}{2\pi}\sum_k \int_{2\pi k}^{2\pi(k+1)} f(x)\eta(x)e^{-inx}dx \\ &= \frac{1}{2\pi}\sum_k \int_0^{2\pi} f(x)\eta(x+2\pi k)e^{-inx}dx = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}dx \end{aligned}$$

by periodicity of f and (**22.8**). The order of integration and summation can be interchanged because the series (**22.8**) converges uniformly on $[0, 2\pi]$. So, the theory of classical trigonometric Fourier series is included into the result (**22.7**).

Let $f \in \mathcal{D}'$ be 2π periodic. From the analysis in the previous section, it also follows that for any test function φ,

$$\varphi(x) \sum_{|n|\le m} \eta(x + 2\pi n) \;\to\; \varphi(x) \quad \text{in } \mathcal{D}$$

as $m \to \infty$. Therefore by continuity, linearity, and periodicity of the functional f one infers that

$$\begin{aligned} (f, \varphi) &= \lim_{m\to\infty} \sum_{|n|\le m} \Big(f(x), \eta(x+2\pi n)\varphi(x)\Big) \\ &= \lim_{m\to\infty} \sum_{|n|\le m} \Big(f(x+2\pi n), \eta(x+2\pi n)\varphi(x)\Big) \\ &= \lim_{m\to\infty} \sum_{|n|\le m} \Big(f(x), \eta(x)\varphi(x-2\pi n)\Big) = (f, \eta\phi)\,, \end{aligned}$$

where $\phi(x) = \sum_n \varphi(x-2\pi n)$ is a periodic C^∞ function with period 2π. The last equality follows from continuity of f and the analysis from the previous section (see the derivation of (**22**.**6**)). The function ϕ can be expanded into a Fourier series

$$\phi(x) = \lim_{m\to\infty} \phi_m(x)\,, \quad \phi_m(x) = \sum_{|n|\le m} b_n e^{-inx}\,,$$

where the Fourier coefficient b_n can written in the form

$$\begin{aligned} b_n &= \frac{1}{2\pi}\int_0^{2\pi} \phi(x)\, e^{inx}\, dx = \frac{1}{2\pi}\sum_k \int_0^{2\pi} \varphi(x-2\pi k)\, e^{inx}\, dx \\ &= \frac{1}{2\pi}\sum_k \int_{2\pi k}^{2\pi(k+1)} \varphi(x) e^{inx} dx = \frac{1}{2\pi}\Big(e^{inx}, \varphi(x)\Big)\,. \end{aligned}$$

The order of integration and summation can be interchanged thanks to uniform convergence of the series on $[0, 2\pi]$. The partial sums ϕ_m and its derivatives $D^\beta\phi_m$ converge uniformly to ϕ and $D^\beta\phi$, respectively, on any compact for any β. Therefore $\eta\phi_m \to \eta\phi$ in $\mathcal{D}$ as $m\to\infty$ by the same reasoning as in the previous section. The conclusion of the theorem follows from the linearity and continuity of f:

$$\begin{aligned} (f,\varphi) = (f,\eta\phi) &= \lim_{m\to\infty}\Big(f,\eta\phi_m\Big) = \lim_{m\to\infty}\sum_{|n|\le m} b_n\Big(f(x), \eta(x)e^{-inx}\Big) \\ &= \lim_{m\to\infty}\sum_{|n|\le m} a_n\Big(e^{inx}, \varphi(x)\Big) \end{aligned}$$

for any test function φ. This means that any 2π periodic distribution f can be expanded into the Fourier series (**22**.**2**).

Now suppose that f is defined by a convergent series (**22**.**2**) for some choice of $\{a_n\}$. Then

$$\begin{aligned} \Big(f(x), \eta(x)e^{-inx}\Big) &= \lim_{m\to\infty}\sum_{|k|\le m} a_k\Big(e^{ikx}, \eta(x)e^{-inx}\Big) \\ &= \lim_{m\to\infty}\sum_{|k|\le m} a_k \int_0^{2\pi} e^{i(k-n)x} dx = 2\pi a_n \end{aligned}$$

because e^{ikx} is a regular periodic distribution. The proof is complete.

For example, let $f(x) = \sum_n \delta(x-2\pi n)$. Then choosing η so that $\eta(0) = 1$, one gets $a_n = \frac{1}{2\pi}$, and the Poisson summation formula is recovered.

There is a simple consequence of Theorems **22**.**2** and **22**.**3** which is a particular case of the structure theorem in Sec.**21**.**7**.

COROLLARY **22.3**. *Any periodic distribution is a repetitive derivative of a continuous function.*

Indeed, any periodic distribution with period T can be reduced to a distribution with period 2π by the aforementioned scaling transformation. Then the 2π periodic distribution can be expanded into the Fourier series by Theorem **22.3**. By Theorem **22.2**, the Fourier coefficients of any periodic distribution $a_n = O(|n|^k)$ for some k as $|n| \to \infty$ because the Fourier series converges in $\mathcal{D}'$. Therefore a continuous function (**22.3**) can be constructed and (**22.3**) holds for any periodic distribution.

22.4.1. Convolution of periodic distributions. Let f and g be periodic distributions with period 2π. They are uniquely defined by their Fourier coefficients

$$g(x) = \sum_n a_n e^{inx}\,, \quad f(x) = \sum_n b_n e^{inx}\,.$$

A *convolution* of periodic distributions f and g is a distribution defined by the Foureir series

$$f * g(x) = 2\pi \sum_n a_n b_n e^{inx}\,.$$

Note that the Fourier series converges in $\mathcal{D}'$ because $a_n = O(|n|^k)$ and $b_n = O(|n|^m)$ so that $a_n b_n = O(|n|^{k+m})$. So, the convolution is a 2π periodic distribution. If g and f are defined by functions integrable on $(0, 2\pi)$ such that

$$\int_0^{2\pi}\int_0^{2\pi} |f(x-y)g(y)|\, dxdy < \infty\,.$$

Then by Fubini's theorem the convolution

$$f * g(x) = \int_0^{2\pi} f(x-y)g(y)\, dy\,.$$

exists and is integrable on $(0, 2\pi)$. It is also 2π periodic by periodicity of f and g. Therefore the Fourier coefficients of $f * g$ are

$$c_n = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{2\pi} f(x-y)g(y)e^{-inx}\, dy\, dx = b_n \int_0^{2\pi} e^{-iny} g(y)\, dy = 2\pi a_n b_n\,,$$

as required.

The convolution is commutative and associative, like the product of classical functions:

$$f * g = g * f\,, \quad (f * g) * h = f * (g * h)\,.$$

The derivative of the convolution has the following property

$$D(f*g) = Df*g = f*Dg\,.$$

Put $\sum_n \delta(x+2\pi n) = \delta_{2\pi}(x)$. Then by the Poisson summation formula $\delta_{2\pi} * g = g$ for any periodic $g \in \mathcal{D}'$. The distribution $\delta_{2\pi}$ is sometimes called a *Dirac comb* with period 2π. It plays the role of the unit periodic distribution relative the multiplication defined by the convolution for 2π periodic distributions.

By analogy with algebra of ordinary functions, one can try to define *division* of a distribution f by a distribution g as the product $g^{-1} * f$ where the *reciprocal* distribution g^{-1} is a solution the equation

$$g^{-1} * g = \delta_{2\pi}\,.$$

If g^{-1} exists as a periodic distribution, then the *convolution equation* has a unique solution:

$$g*h = f \quad \Rightarrow \quad h = g^{-1} * f\,.$$

Unfortunately, not every periodic distribution has the reciprocal. If g^{-1} exists, then its Fourier coefficients must be reciprocals $1/a_n$ so that $a_n \neq 0$ necessarily. For example, take $g(x) = \sum_n e^{-n^2} e^{inx}$. Then g^{-1} does not exist because its Fourier coefficients do not satisfy the hypotheses of Theorem **22.2**. More generally, if g is defined by a C^∞ periodic function. Then $|n|^k a_n \to 0$ as $n \to \infty$ for any k. Therefore the reciprocals $1/a_n$ grow faster than any power $|n|^k$ and the Fourier series with coefficients $1/a_n$ does not converge in $\mathcal{D}'$.

22.5. Exercises.

1. Let $f(x) = 1 - |x|$ if $|x| < 1$ and $f(x+2) = f(x)$. Show that

$$f''(x) = 2\sum_n e^{i\pi(2n+1)x}$$

in the sense of distributions.

2. Poisson summation formula for smooth functions. Let φ be a test function and

$$\mathcal{F}[\varphi](k) = \int_{-\infty}^{\infty} \varphi(x)\, e^{ixk}\, dx$$

be its Fourier transform. Show that

$$\sum_n \varphi(2\pi n) = \frac{1}{2\pi}\sum_n \mathcal{F}[\varphi](n)$$

3. Show that the Fourier series

$$f(x) = \sum_{n=1}^{\infty} n^2 \cos(nx)$$

converges in the sense of distributions and express its sum in terms of shifted delta-functions and its derivatives. In particular, find an explicit expression for (f, φ) where $\varphi \in \mathcal{D}$ in terms of point values of φ and, possibly, point values of the derivatives of φ.

23. Fourier method for differential equations

Let us find all 2π periodic distributions u that satisfy a differential equation

$$L(D)u(x) = f(x)$$

where $L(D)$ is a linear differential operator of order p with constant coefficients, and f is a 2π periodic distribution. If a solution exists, then it is uniquely defined by its Fourier series. Put

$$f(x) = \sum_n a_n e^{inx}\,, \quad u(x) = \sum_n b_n e^{inx}\,.$$

Derivatives of u are obtained by term-by-term differentiation of their Fourier series. Therefore

$$L(in)b_n = a_n\,.$$

If $L(in) \neq 0$ for any integer n, then the solution exists and is unique. It can be written as the convolution of periodic distributions

$$u(x) = G * f(x) = \sum_n \frac{a_n}{L(in)}\, e^{inx}\,, \tag{23.1}$$

$$G(x) = \frac{1}{2\pi}\sum_n \frac{e^{inx}}{L(in)}\,. \tag{23.2}$$

For differential operators of order $p \geq 2$, the series (**23**.**2**) converges uniformly and, hence, $G(x)$ is continuous and 2π periodic. For $p = 1$, G is piecewise continuous (shown below). So, if f is integrable on $(0, 2\pi)$, then by Fubini's theorem the solution can be written as the convolution integral

$$u(x) = G * f(x) = \int_0^{2\pi} G(x-y)f(y)\,dy\,,$$

that defines a regular periodic distribution.

23.1. Differential equations as a convolution equation. The differential equation studied above can also be cast as the convolution equation. By the properties of the convolution

$$L(D)u = L(D)(\delta_{2\pi} * u) = L(D)\delta_{2\pi} * u = f\,,$$

where $\delta_{2\pi}$ is the Dirac comb. The reciprocal of $L(D)\delta_{2\pi}$ exists if $L(in) \neq 0$ for any integer n and, in this case, the reciprocal is equal to the Green's function:

$$u = [L(D)\delta_{2\pi}]^{-1} * f = G * f\,.$$

23.2. Regular periodic boundary value problem. Equation (**23**.1) solves the posed problem. However a practical use of it is somewhat complicated because of summation of the Fourier series. If the inhomogeneity f is a regular distribution, then the integral representation of the convolution looks technically simpler to deal with than summation of the Fourier series. Is it possible to find the sum (**23**.**2**)? Let us try to do so.

By Theorem **22**.**2** and the Poisson summation formula, the distribution G is a periodic solution to the equation

$$L(D)G(x) = \sum_n \delta(x - 2\pi n)\,. \tag{23.3}$$

Note that the right-hand side is equal to $\delta(x)$ in $(-\pi, \pi)$. As noted above, the distribution G is defined by a (piecewise) continuous 2π periodic function. Let $G_0(x) = G(x)$ in $(-\pi, \pi)$ and $G_0(x) = 0$ a.e. otherwise. Then

$$G_0(x) = \frac{1}{2\pi}\sum_n \frac{e^{inx}}{L(in)}\,, \quad x \in (-\pi, \pi)\,, \tag{23.4}$$

$$G(x) = \sum_n G_0(x + 2\pi n)\,. \tag{23.5}$$

The distribution G_0 satisfies the equation

$$L(D)G_0(x) = \delta(x)\,, \quad x \in (-\pi, \pi)\,. \tag{23.6}$$

A solution to this equation is not unique as one can always add a solution to the associate homogeneous equation $G_0 \to G_0 + h$ where $L(D)h = 0$. Recall that a general solution h is a linear combination of p linearly independent solutions. So, h has exactly p real parameters. All linearly independent solutions can be written in the form $x^k e^{\gamma x}$ where γ is a root (real or complex) of $L(\gamma) = 0$ of multiplicity m and in this case $k = 0, 1, ..., m - 1$. So, h is a C^∞ function in any open interval.

Thus, not every solution to (**23**.**6**) would serve the purpose in (**23**.**5**) as G is shown to be unique. Note that near $x = \pi$, the Dirac comb vanishes, $\delta_{2\pi}(x) = 0$ for $|x - \pi| < a$ for small enough $a > 0$. Therefore $G(x)$ must be a solution to the homogeneous equation near $x = \pi$. On the other had, by (**23**.**5**) $G(x) = G_0(x) + G_0(x - 2\pi)$ near $x = \pi$. In the intervals $(-\pi, 0)$ and $(0, \pi)$ the function G_0 satisfies the homogeneous equation and, hence, it is smooth. Therefore $G(x)$ is defined by a *piecewise smooth* function near $x = \pi$ because the sum $G_0(x) + G_0(x - 2\pi)$ and any of its derivatives have jump discontinuities in general at $x = \pi$. Owing to the rule of distributional differentiation of piecewise smooth functions

$$G^{(k)}(x) = \{G^{(k)}\} + \sum_{m=0}^{k-1} c_m \delta^{(m)}(x-\pi)\,, \quad |x-\pi| < a\,,$$

where the coefficients $c_m = G_0^{(k-1-m)}(-\pi) - G_0^{(k-1-m)}(\pi)$ define the magnitude of the jump discontinuities of classical derivatives $\{G^{(k-1-m)}\}$ at $x = \pi$. This shows that G cannot satisfy (**23.3**) near $x = \pi$ unless the distribution G_0 satisfies the periodic boundary conditions

$$G_0^{(k)}(-\pi) = G_0^{(k)}(\pi)\,, \quad k = 0, 1, ..., p-1\,. \tag{23.7}$$

In this case, $c_m = 0$ for any $m = 0, 1, ..., p-1$, and $L(D)G = \{L(D)G\} = 0$ near $x = \pi$ as well as near $x = \pi + 2\pi n$ for any integer n, as required.

Moreover, if such G_0 exists, then it is unique. Any two solutions differs by a solution to the homogeneous equation but now any such solution is also required to obey periodic boundary conditions (**23.7**). Such a non-trivial solution exists only if $L(\gamma) = 0$ has a root $\gamma = in$ for some integer n because $e^{-\gamma\pi} = e^{\gamma\pi}$ only if $\gamma = in$. However, G exists and is unique precisely under the condition that $L(in) \neq 0$ for any integer n.

The operator $L(D)$ for which $L(in) \neq 0$ is called a *regular differential operator* on a circle. Periodic functions can be viewed as functions on a unit circle owing to the boundary conditions (**23.7**). So, the periodic distribution G satisfying (**23.3**) exists and is unique if and only if the distribution G_0 solving the *regular boundary value problem* (**23.6**) and (**23.7**) exists and is unique. It remains to solve the boundary value problem.

23.2.1. Solving regular boundary value problems on a circle. Let

$$G_0(x) = G_{0+}(x)\,, \quad x \in (0, \pi)\,, \quad G_0(x) = G_{0-}(x)\,, \quad x \in (-\pi, 0)\,.$$

As noted, the functions $G_{0\pm}$ are smooth and satisfy the homogeneous equation $L(D)G_{0\pm} = 0$. Each solution has p parameters. Consequently, G_0 is piecewise smooth and has $2p$ parameters. The boundary conditions impose p conditions on these parameters:

$$G_{0+}^{(k)}(\pi) = G_{0-}^{(k)}(-\pi)\,, \quad k = 0, 1, ..., p-1\,.$$

Then according to the distributional differentiation of a piecewise smooth function, (**23.6**) is reduced to

$$L(D)G_0 = \{L(D)G_0\} + \sum_{k=0}^{p-1} c_k \delta^{(k)} = \delta\,.$$

where c_k are defined by magnitudes of jump discontinuities of classical derivatives of G_0 at $x = 0$. By construction $\{L(D)G_0\} = 0$. Therefore (**23**.**6**) is satisfied only if $c_0 = 1$ and $c_k = 0$ for $k \neq 0$, assuming that the coefficient at the p^{th} derivative in $L(D)$ is equal to 1. To get rid off all terms containing derivatives of the delta function, one must demand that G_0 and its derivatives up order $p - 2$ are continuous at $x = 0$. Thus, another p conditions on G_0 are

$$G_{0-}^{(k)}(0) = G_{0+}^{(k)}(0)\,, \quad k = 0, 1, ..., p-2\,, \quad G_{0+}^{(p-1)}(0) - G_{0-}^{(p-1)}(0) = 1\,.$$

It is interesting to note that Eq. (**23**.**4**) offers a summation formula for trigonometric Fourier series with coefficients being reciprocals of polynomials. The function G defined by (**23**.**7**) is called a *Green's function of the operator $L(D)$ on a circle*.

23.2.2. Examples. Let $L(D) = D + \gamma$ for some real γ. Then

$$G_{0\pm}(x) = A_\pm e^{-\gamma x}$$

for some constants $A_\pm$. The periodicity boundary condition requires that $A_+ = A_- e^{2\pi\gamma}$. Therefore

$$\begin{aligned} G_0'(x) + \gamma G_0(x) &= \{G_0'(x)\} + \gamma G_0(x) + \operatorname*{disc}_{x=0}[G_0]\delta(x) \\ &= (A_+ - A_-)\delta(x)\,. \end{aligned}$$

The equation is satisfied if $A_+ - A_- = 1$ and

$$G_0(x) = \frac{1}{e^{2\pi\gamma} - 1} \begin{cases} e^{-\gamma x} & , x \in (-\pi, 0) \\ e^{\gamma(2\pi - x)} & , x \in (0, \pi) \end{cases}$$

It is straightforward to verify that

$$\int_{-\pi}^{\pi} G_0(x) e^{-inx} dx = \frac{1}{in + \gamma}\,,$$

and therefore

$$G(x) = \sum_n G_0(x + 2\pi n) = \frac{1}{2\pi} \sum_n \frac{e^{inx}}{in + \gamma}\,.$$

The Fourier series converges to $G_0(x)$ everywhere in $(-\pi, \pi)$ but $x = 0$ (see (**22**.**1**)).

Let $L(D) = -D^2 + \gamma^2$ where $\gamma > 0$. To find the Green's function, one has to solve the following problem:

$$-G_0'' + \gamma^2 G_0 = \delta\,, \quad G_0(-\pi) = G_0(\pi)\,,\; G_0'(-\pi) = G_0'(\pi)\,.$$

The homogeneous equation has two linearly independent solutions, $e^{\pm\gamma x}$. Therefore

$$G_{0\pm}(x) = A_\pm e^{\gamma x} + B_\pm e^{-\gamma x}\,,$$

for some constants $A_\pm$ and $B_\pm$. The boundary conditions and the continuity conditions are satisfied only if $A_\pm$ and $B_\pm$ are such that

$$G_0(x) = \frac{1}{2\gamma \sinh(\pi\gamma)} \begin{cases} \cosh[\gamma(x+\pi)] \ , \ x \in (-\pi, 0) \\ \cosh[\gamma(x-\pi)] \ , \ x \in (0, \pi) \end{cases}$$

Technical details are left to the reader as an exercise. It is not difficult to see that the periodicity conditions are fulfilled:

$$G_0(\pm\pi) = \frac{1}{2\gamma \sinh(\pi\gamma)} \,, \quad G_0'(\pm\pi) = 0 \,.$$

The function G_0 is continuous at $x = 0$ and its classical derivative has a required jump discontinuity:

$$G_0(0^\pm) = \frac{\cosh(\pi\gamma)}{2\gamma \sinh(\pi\gamma)} \,, \quad \{G_0'(0^\pm)\} = \mp\frac{1}{2} \,,$$

where the argument $0^\pm$ denotes the left and right limits of the function at $x = 0$. Therefore $-G_0'' = -\{G_0'\}' = -\{G_0''\} + \delta$. It is also not difficult to show that

$$\int_{-\pi}^{\pi} G_0(x) e^{-inx} dx = \frac{1}{n^2 + \gamma^2} \,,$$

so that the periodic extension of G_0 to the whole $\mathbb{R}$ reads

$$G(x) = \sum_n G_0(x + 2\pi n) = \frac{1}{2\pi} \sum_n \frac{e^{inx}}{n^2 + \gamma^2} \,.$$

It satisfies (**23**.**3**). The sum of the Fourier series in the above equation is equal to $G_0(x)$ for every $x \in [-\pi, \pi]$.

23.3. Classical boundary value problem on a circle. Green's function found in the previous section can be used to construct an integral representation of classical solutions to regular boundary value problems on a circle.

PROPOSITION **23**.**5**. *Let f be a continuous function on $[-\pi, \pi]$. Put*

$$u(x) = \int_{-\pi}^{\pi} G(x - y) f(y)\, dy \,, \quad |x| < \pi \,,$$

*where G is given by (**23**.**5**) and G_0 is a solution to the boundary value problem (**23**.**6**) and (**23**.**7**). Then $u \in C^p(-\pi, \pi) \cap C^{p-1}[-\pi, \pi]$ and is a unique solution to the boundary value problem:*

$$L(D)u = f \,, \quad u^{(k)}(-\pi) = u^{(k)}(\pi) \,, \quad k = 0, 1, ..., p - 1 \,.$$

To verify smoothness of u, the boundary conditions, and the equation, it is convenient to write u explicitly via the smooth functions $G_{0\pm}$:

$$u(x) = \Big(\int_{-\pi}^{x-\pi} G_{0+}(x-y) + \int_{x-\pi}^{x} G_{0-}(x-y-2\pi) + \int_{x}^{\pi} G_{0-}(x-y) \Big) f(y)\, dy\,, \quad x \in [0,\pi]\,,$$
$$u(x) = \Big(\int_{-\pi}^{x} G_{0+}(x-y) + \int_{x}^{x+\pi} G_{0-}(x-y) + \int_{x+\pi}^{\pi} G_{0-}(x-y+2\pi) \Big) f(y)\, dy\,, \quad x \in [-\pi,0]\,.$$

It follows from this representation that $u(\pi) = u(-\pi)$. Since f is continuous, the (classical) derivative u' can be computed by the fundamental theorem of calculus. It has same integral representation as u where $G_{0\pm}$ are replaced by the derivatives $G'_{0\pm}$. The boundary terms arising from differentiation vanish thanks to boundary and continuity conditions for $G_{0\pm}$. The same holds for u'' and other derivatives. The equation is verified by taking a combination of derivatives to make $L(D)u$ and using that $G_{0\pm}$ are annihilated by $L(D)$. The technical details are left to the reader as an exercise.

The convolution with the Green's function is an operator that is inverse to the differential operator $L(D)$. It is noteworthy to make an analogy with the linear algebra problem $Ax = b$ where A is a square matrix and b is a given vector. The problem has a unique solution if A is invertible and, in this case, $x = A^{-1}b$. A matrix is invertible if and only if the homogeneous equation $Ax = 0$ has only the trivial solution $x = 0$. Here the differential operator L acts in a special class of functions that satisfy boundary conditions. If $L(D)h = 0$ has only the trivial solution subject to the boundary conditions, then L is invertible and the solution to $Lu = f$ is unique and given by $u = L^{-1}f = G * f$.

23.3.1. Well-posedness of the problem. A problem is well posed if its solution exists, is unique, and depends continuously on parameters. In other words, small variations of parameters produce small variations of the solution. Let u and $\tilde{u}$ be solutions to the above problem for inhomogeneities f and $\tilde{f}$. Then

$$|u(x) - \tilde{u}(x)| \le \int_{-\pi}^{\pi} |G(x-y)||f(y) - \tilde{f}(y)|\, dy \le 2\pi M \sup |f(x) - \tilde{f}(x)|\,, \quad M = \sup |G(x)|\,,$$

where $M < \infty$ because the Green's function is bounded. This inequality holds for all x and, hence, for any $\varepsilon > 0$

$$\sup |f(x) - \tilde{f}(x)| \le \varepsilon \quad \Rightarrow \quad \sup |u(x) - \tilde{u}(x)| \le 2\pi M\varepsilon \,.$$

This shows that the classical boundary value problem for a regular differential operator on a circle is well posed.

23.4. Green's functions for a singular differential operator on a circle. Suppose that $L(in) = 0$ for some integers $n \in Z_L$. By analogy with the linear algebra problem, the operator L will be called *singular*. In this case, the problem has no solution if $a_n \neq 0$ at least for one $n \in Z_L$. If $a_n = 0$ for all $n \in Z_L$, then the problem has a solution but it is not unique because b_n, $n \in Z_L$, are arbitrary so that

$$u(x) = \sum_{n\in Z_L} b_n e^{inx} + \sum_{n\notin Z_L} \frac{a_n}{L(in)}\, e^{inx} = \sum_{n\in Z_L} b_n e^{inx} + G_s * f(x) \,.$$

The distribution

$$G_s(x) = \frac{1}{2\pi} \sum_{n\notin Z_L} \frac{e^{inx}}{L(in)}$$

will be called a *Green's function for a singular differential operator L on a circle*. The objective is to find an explicit form of G_s as a periodic function of x and thereby to obtain an integral representation for the solution when f is a regular distribution:

$$u(x) = \sum_{n\in Z_L} b_n e^{inx} + \int_{-\pi}^{\pi} G_s(x-y) f(y)\, dy \,,$$

subject to the conditions that the Fourier coefficients of f vanish for all $n \in Z_L$. Note also that, if L is real, then $-n \in Z_L$ if $n \in Z_L$.

To accomplish this task, the same strategy will be employed as in the case of a regular L. But it requires modifications. First, the function G_s does not satisfy (**23.3**). The new equation for the Green's function reads

$$L(D)G_s = \frac{1}{2\pi} \sum_{n\notin Z_L} e^{inx} = \sum_{n} \delta(x - 2\pi n) - \frac{1}{2\pi} \sum_{n\in Z_L} e^{inx} \,.$$

Therefore G_s can be written in the form (**23.5**) where G_0 satisfies the equation

$$L(D)G_0(x) = \delta(x) - \frac{1}{2\pi} \sum_{n\in Z_L} e^{inx} \,, \tag{23.8}$$

and the boundary conditions (**23.7**). A solution can be found in the exactly same way as in the regular case. The smooth functions $G_{0\pm}$

are general solutions to the above equation in the intervals $(0, \pm\pi)$, that is, when $\delta(x)$ is omitted in the right-hand side. So, G_0 is again a piecewise smooth with a jump discontinuity at $x = 0$. This implies that $G_{0\pm}$ satisfy the same boundary conditions and the same continuity conditions at $x = 0$ as in the regular case. The difference is that $G_{0\pm}$ are general solutions to the *non-homogeneous* problem. A solution G_0 constructed in this way is not unique because G_0 can be changed by adding a linear combination of e^{inx} with $n \in Z_L$.

To eliminate this ambiguity, note that G_s is, by construction, a periodic extension of an integrable function G_0 by the Fourier series from $(-\pi, \pi)$ to $\mathbb{R}$. Therefore, the Fourier coefficients of G_0 must vanish for all $n \in Z_L$:

$$\int_{-\pi}^{\pi} G_0(x) e^{-inx}\, dx = 0\,, \quad n \in Z_L\,.$$

The convolution $G_s * f$ is not affected by these *orthogonality* conditions because f must satisfy the same conditions. The orthogonality conditions make G_0 a unique solution to (**23.8**) that satisfies (**23.7**) such that

$$G_s(x) = \sum_n G_0(x - 2\pi n) = \sum_{n \notin Z_L} \frac{1}{L(in)}\, e^{inx}\,.$$

By construction, the Fourier series converges to $G_0(x)$ for all $x \in [-\pi, \pi]$.

23.4.1. Example. Let $L(D) = D^2 + 1$. Then $Z_L = \{\pm 1\}$ so that

$$G_s'' + G_s = \delta(x) - \frac{1}{\pi}\cos(x)\,, \quad G_s(\pi) = G_s(-\pi)\,, \quad G'(-\pi) = G_s'(\pi)\,.$$

A general solution for $x < 0$ and $x > 0$ reads

$$G_{s\pm}(x) = A_\pm \cos(x) + B_\pm \sin(x) - \frac{x}{2\pi}\, \sin(x)\,.$$

The boundary conditions yield $A_+ = A_-$ and $B_+ = B_1 + 1$. The continuity conditions, $\operatorname{disc} G_s = 0$ and $\operatorname{disc}\{G_s'\} = 1$ at $x = 0$, do not impose any further restrictions on the real parameters $A_\pm$ and $B_\pm$. Therefore a general solution reads

$$G_s(x) = A_- \cos(x) + B_- \sin(x) - \frac{x}{2\pi}\, \sin(x) + \theta(x)\sin(x)\,.$$

Finally, the orthogonality conditions, which are convenient to write in the real form,

$$\int_{-\pi}^{\pi} G_s(x)\sin(x)\, dx = 0\,, \quad \int_{-\pi}^{\pi} G_s(x)\cos(x)\, dx = 0$$

yield $B_- = -\frac{1}{2}$ and $A_- = \frac{1}{4\pi}$ so that

$$G_s(x) = \frac{1}{4\pi}\cos(x) - \frac{x}{2\pi}\sin(x) + \frac{1}{2}\varepsilon(x)\sin(x) = 1 - 2\sum_{n=2}^{\infty}\frac{\cos(nx)}{n^2-1},$$

where $\varepsilon(x)$ is the sign function, for any $x \in [-\pi, \pi]$ because $L(in) = 1 - n^2$.

23.5. Regularization of distributional solutions. Let $f_\varepsilon = \phi_\varepsilon * f$ be a C^∞ regularization of a periodic distribution f. By periodicity of f, its regularization is also periodic. Let u be a solution to $L(D)u = f$. Then by (**21**.**7**)

$$f_\varepsilon = \phi_\varepsilon * f = \phi_\varepsilon * L(D)u = L(D)(\phi_\varepsilon * u).$$

This shows that a smooth periodic function $u_\varepsilon = \phi_\varepsilon * u$ converges to the distributional solution u as $\varepsilon \to 0^+$ and, hence,

$$u_\varepsilon(x) = \int_0^{2\pi} G(x-y) f_\varepsilon(y)\, dy \;\to\; u(x) \quad \text{in } \mathcal{D}',$$

as $\varepsilon \to 0^+$. If L is singular, then G is to be replaced by G_s to obtain a particular solution.

Thus, for every distributional solution there exists a smooth solution that is arbitrary close to the distributional one in the sense that for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|(f,\varphi) - (f_a,\varphi)| < \varepsilon, \quad \varphi \in \mathcal{D}$$

for all $a < \delta$. So, from the perspective that measured physical quantities are distributions rather than classical functions, the above relation shows that distributional solutions are just as good as smooth ones but the former are much easier to find via formal Fourier series as one does not dot need to worry about uniform convergence at all when studying smoothness of the solution. Later it will be shown these ideas can also be extended to linear partial differential equation and other orthogonal bases in $\mathcal{L}_2(\Omega)$.

23.6. Exercises.

1. Dumped harmonic oscillator on a circle. (i) Find the Green's function for the operator $L(D) = D^2 - 2\gamma D + \omega^2$ on a circle.
(ii) Solve the periodic boundary value problem

$$L(D)u(x) = x, \quad x \in (-\pi, \pi), \quad u(-\pi) = u(\pi),\ u'(-\pi) = u'(\pi).$$

24. Taylor approximations of distributions

In applications, smooth functions are often approximated by Taylor polynomials near a particular point or by some asymptotic expansions, e.g., for large values of arguments. Can these concepts be extended to distributions? The answer is affirmative but with some new peculiarities characteristic to distributions. These techniques are used to study asymptotic behavior and other approximations of distributional solutions of partial differential equations, like Green's functions for differential operators. Here basic mathematical foundations are formulated to achieve the goal and illustrated by Taylor asymptotic expansions of distributions. Applications to Green's functions will be considered later.

24.1. On the alternative definition of distributional derivatives. Equation (**13**.**4**) gives an alternative definition of the distributional derivatives. Let us show that it is equivalent to the distributional derivative defined as the adjoint of differentiation on the space functions. In Sec.**13.2.5** it was argued that it is sufficient to show that

$$\frac{\varphi(x-a)-\varphi(x)}{a} = \psi_a(x) \to -\varphi'(x) \quad \text{in } \mathcal{D} \quad \text{as } a \to 0\,.$$

Indeed, the support of ψ_a lies in $|x| \le R+1$ for all $|a| \le 1$ (the support is independent of a) if the support of φ is in $[-R, R]$. By the Taylor theorem, there exists a point x_a between $x-a$ and x such that

$$\varphi(x-a) = \varphi(x) - a\varphi'(x) + \frac{1}{2}\, a^2\varphi''(x_a)\,.$$

Since all derivatives of a test function are bounded,

$$\Big|\psi_a(x) + \varphi'(x)\Big| \le \frac{|a|}{2}\, \sup|\varphi''|\,,$$

which holds for any x. Therefore one can take the supremum in the left-hand side:

$$\sup|\psi_a + \varphi'| \le \frac{|a|}{2}\, \sup|\varphi''|\,.$$

Hence, ψ_a converges to $-\varphi'$ uniformly:

$$\lim_{a\to 0} \sup|\psi_a + \varphi'| = 0\,.$$

Similarly

$$\Big|D^\beta\psi_a(x) + D^\beta\varphi'(x)\Big| \le \frac{|a|}{2}\, \sup|D^\beta\varphi''|$$

from which the uniform convergence of all derivatives of ψ_a to the corresponding derivatives of $-\varphi'$ follows. Thus, ψ_a converges to $-\varphi'$

in $\mathcal{D}$ and hence the distributional derivative can also be computed via the distributional limit (**13**.**4**).

It is also clear that the same line of arguments can be applied to the distribution f' to find the second distributional derivative f'' and so on. This shows that *the classical definition of the derivative via the limit can be extended to all distributions and it agrees with the rule (***21**.**1***)*.

Recall that a differentiable function has a good linear approximation. Equation (**13**.**4**) allows us to extend this concept to distributions. Any distribution f admits the following asymptotic representation

$$f(x+a) = f(x) + af'(x) + O'(a^2)\,,$$

where $O'(a^2)$ denotes a distribution with the characteristic property that $(O'(a^2), \varphi) = O(a^2)$ as $a \to 0$ for any test function φ. Since any distribution can be differentiated any number of times, one can find asymptotic expansions of distributions that are similar to Taylor polynomial approximations of smooth functions.

24.2. Asymptotic power series for distributions. Let $f(x;a)$ be a distribution from $\mathcal{D}'(\mathbb{R}^N)$ for every $a \in \mathbb{R}^M$. *A distribution $f(x;a)$ is said to be from class $C^p(\Omega)$ in parameters a if for every test function $\varphi(x)$,*

$$u(a) = \Big(f(x;a), \varphi(x)\Big) \in C^p(\Omega)\,.$$

For the sake of simplicity, let $a \in \mathbb{R}$. A generalization to any Euclidean space is straightforward. Since the derivative $u'(a)$ exist for any a, by the completeness theorem, there exists a distribution $f_1(x;a)$ such that

$$u'(a) = \lim_{h\to 0}\Big(\frac{f(x;a+h) - f(x;a)}{h}, \varphi(x)\Big) = (f_1(x;a), \varphi(x))\,,$$

Owing to the definition of partial derivatives of classical functions, one writes for brevity

$$f_1(x;a) = \lim_{h\to 0}\frac{f(x;a+h) - f(x;a)}{h} \stackrel{\text{def}}{=} \frac{\partial}{\partial a}f(x;a)\,.$$

The distribution f_1 will be called the derivative of f with respect to a parameter a. If the distribution is from class C^p in parameter a, then

$$f_k(x;a) = \frac{\partial}{\partial a}f_{k-1}(x;a)\,, \quad k \le p\,.$$

Near a particular point, say, $a = 0$, the function u can be approximated by a Taylor polynomial

$$u(a) = \sum_{k=0}^{n-1} \frac{u^{(k)}(0)}{k!}\, a^k + O(a^n)\,, \quad n \le p\,,$$

where

$$u^{(k)}(0) = (f_k(x;0), \varphi(x))\,.$$

A Taylor polynomial approximation of $f(x;a)$ in parameter $a \to 0$ is defined by

$$f(x;a) = P_n(x;a) + O'(a^n)\,, \quad P_n(x;a) = \sum_{k=0}^{n-1} \frac{f_k(x)}{k!}\, a^k\,,$$

where the symbol $O'(a^n)$ stand for a distribution with the characteristic property $(O'(a^n), \varphi) = O(a^n)$ as $a \to 0$.

Suppose that $u \in C^\infty$. Then in the *formal* limit $n \to \infty$ one has :

$$f(x;a) \sim \sum_{k=0}^{\infty} \frac{f_k(x;0)}{k!}\, a^k \stackrel{\text{def}}{=} \lim_{n\to\infty} P_n(x;a)\,.$$

This is a *formal* series because nothing is known about its convergence in $\mathcal{D}'$ for $a \neq 0$. In fact, as is shown in the next section, asymptotic series often have no limit in $\mathcal{D}'$. However the series does have the *characteristic property* of the power series representation:

$$\lim_{a\to 0} \frac{1}{a^n}\Big(f(x;a) - P_n(x;a)\Big) = \frac{f_n(x;0)}{n!}\,,$$

where the limit is understood in the distributional sense. For this reason is called an *asymptotic* power series expansion of the distribution f in parameter a when $a \to 0$.

24.2.1. Distributions smooth in a particular variable. The same construction can be developed for particular variables in a distribution. Consider a distribution $f(x,y)$ of two variables $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^M$. For every such distribution one can define a distribution

$$g(y) \in \Big(f(x,y), \varphi(x)\Big) \in \mathcal{D}'(\mathbb{R}^M)\,, \quad \varphi \in \mathcal{D}(\mathbb{R}^N)\,,$$

by the rule

$$\Big(g(y), \psi(y)\Big) \stackrel{\text{def}}{=} \Big(f(x,y), \varphi(x)\psi(y)\Big)\,, \quad \psi \in \mathcal{D}(\mathbb{R}^M)\,.$$

Note that the product $\varphi(x)\psi(y)$ is a test function of two variables and if $\psi_n \to 0$ in $\mathcal{D}(\mathbb{R}^M)$, then $\varphi\psi_n \to 0$ in $\mathcal{D}(\mathbb{R}^{N+M})$ so that g is a linear and continuous functional on $\mathcal{D}(\mathbb{R}^M)$. Let us find derivatives $D^\beta g$. One has

$$(D^\beta g, \psi) = (-1)^\beta \Big(g, D^\beta \psi) = (-1)\Big(f(x,y), \varphi(x) D^\beta \psi(y)\Big) \\ = \Big(D^\beta_y f(x,y), \varphi(x)\psi(y)\Big)$$

This shows that

$$D^\beta g(y) = D^\beta \Big(f(x,y), \varphi(x)\Big) = \Big(D^\beta_y f(x,y), \varphi(x)\Big)\,.$$

If g is defined by a smooth function, then its distributional and classical derivatives coincide. *A distribution $f(x,y)$ is said to be from class $C^p(\Omega)$ in the variable y if $g \in C^p(\Omega)$.*

Let $M = 1$ for simplicity. The distribution g and its derivative have *pointwise values*. This implies that if $y = 0$ lies in Ω, then $D^\beta_y f(x,0) \in \mathcal{D}'(\mathbb{R}^N)$. Then a Taylor polynomial approximation to $f(x,y)$ in the variable y about $y = 0$ can be defined by

$$f(x,y) = \sum_{k=0}^{n-1} \frac{D^k_y f(x,0)}{k!}\, y^k + O'(y^n)\,,$$

where the same convention is used that $O'(y^n)$ denotes a distribution in the variable x with the characteristic property that $(O'(y^n), \varphi) = O(y^n)$ as $y \to 0$.

If $f(x,y)$ is from class C^∞ in the variable y, then one can take the formal limit $n \to \infty$ and obtain an asymptotic power series expansion of f in the variable y. There is no guarantee that the power series converges to f in $\mathcal{D}'(\mathbb{R}^{N+M})$ and, hence, cannot be used in calculations in the place of f, unless the convergence is established. A generalization to the case $M > 1$ is obtained just by replacing single variable Taylor polynomials by the corresponding multi-variable Taylor polynomials in the above Taylor polynomial approximations.

24.3. Asymptotic Taylor expansion of distributions. Let $f \in \mathcal{D}'(\mathbb{R})$. Consider a shifted distribution $f(x;a) = f(x+a)$. Then the function

$$u(a) = \Big(f(x+a), \varphi(x)\Big) = \Big(f(x), \varphi(x-a)\Big) = (\varphi^- * f)(a)\,.$$

is from class C^∞ by Proposition **19.3** and

$$u^{(n)}(0) = \Big(f(x), D^n_a \varphi(x-a)\Big)\Big|_{a=0} = (-1)^n (f, \varphi^{(n)}) = (f^{(n)}, \varphi)\,.$$

Therefore an asymptotic power series expansion in a reads

$$(\mathbf{24.1})\qquad f(x+a)\sim\sum_{n=0}^{\infty}\frac{f^{(n)}(x)}{n!}\,a^n \overset{\text{def}}{=} \lim_{n\to\infty}P_n(x;a)\,,$$
$$P_n(x;a)=f(x)+f'(x)a+\cdots+\frac{f^{(n)}(x)}{(n-1)!}\,a^{n-1}\,.$$

It is similar to the Taylor expansion of smooth functions and will be called an *asymptotic Taylor expansion* of a distribution f. In general, the series does not converge in $\mathcal{D}'$ for $a\neq 0$ but it has the characteristic property of the Taylor series

$$(\mathbf{24.2})\qquad \frac{1}{a^n}\Big(f(x+a)-P_n(x;a)\Big)\to\frac{1}{n!}\,f^{(n)}(x)\quad\text{in }\mathcal{D}'$$

as $a\to 0$.

It is not difficult to see that there are distributions for which the equality in (**24.1**) is not possible for any $a\neq 0$. If support of f is not the whole $\mathbb{R}$, say, $\operatorname{supp} f=K$ is a closed and bounded interval. Then the support of $f(x+a)$ is an interval K_a that is obtained by shifting K by a distance a. Therefore K_a has a point x_a that is at a non-zero distance from K, and there exists a test function η_a supported in a neighborhood of x_a and vanishing in a neighborhood of K. For example, one can take a bump function for the point set $x=x_a$ whose support does not intersect K. Then $(P_n,\eta_a)=0$ for all n, while $(f(x+a),\eta_a)\neq 0$. Recall also that if f has a bounded support, then $u=\varphi^-*f$ is a test function by Proposition **19.3** and, hence, cannot be a real analytic function. There are points near which u cannot be given by a power series.

For instance, consider the asymptotic Taylor expansion of the shifted delta function:

$$\delta(x+a)\sim\delta(x)+\sum_{n=1}^{\infty}\frac{a^n}{n!}\,\delta^{(n)}(x)\,.$$

The left and right-hand sides have non-intersecting supports if $a\neq 0$ and, hence, cannot be equal. Furthermore, if the series $\sum_n c_n\delta^{(n)}$ converges to a distribution f in $\mathcal{D}'$, then the support of f consists of the single point $x=0$. By the structure theorem for distributions with a point support, this is possible only if finitely many c_n are not equal to zero. Therefore the asymptotic Taylor expansion for a shifted delta-function does *not* converge in $\mathcal{D}'$ for any $a\neq 0$.

It is therefore interesting to find a class of distributions for which the equality holds in (**24.1**). In order for the asymptotic Taylor expansion to converge in $\mathcal{D}'$ it is sufficient that $u = \varphi^- * f$ is a real analytic function for any test function φ, that is, u is defined by a power series near any point. Let us describe distributions for which the convolution with any test function is a real analytic function.

Suppose that a distribution is defined by a real analytic function $f(x)$ so that its values near any point are given by a convergent Taylor series about that point:

$$f(x+a) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}\, a^n\,, \quad |a| < r\,,$$

where the radius of convergence $r > 0$ can depend on x. Then

$$u(a) \sim \sum_{n=0}^{\infty} \frac{(f^{(n)}(x), \varphi(x))}{n!}\, a^n\,.$$

If the series converges for some $a \neq 0$, then its radius of convergence depends on φ. A pointwise convergence of the series for $f(x+a)$ does not yet guarantee its convergence in $\mathcal{D}'$. One should make sure that there is no test function for which the radius of convergence vanishes. Let us further assume that a real analytic function f can be extended analytically into a strip $|\mathrm{Im}\, z| < r$ in the complex plane so that

$$f(y) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}\, (y-x)^n\,, \quad |x-y| < r\,,$$

for any real x and y. By setting $y = x + a$, one can see that for any $|a| < r$ the Taylor series in (**24.1**) converges pointwise for all x.

Let us investigate the convergence of the series in $\mathcal{D}'$. A real analytic function f is known to have the following characteristic property[5]. For any $0 < b < r$ and any bounded and closed interval $[-R, R]$, there exists a constant M such that

$$|f^{(n)}(x)| \leq M b^{-n} n!\,, \quad |x| \leq R\,,$$

and all n. This shows that the series converges uniformly on any compact, which is sufficient for the convergence of the series in the distributional sense (recall Theorem **22.1**). Indeed, fix $|a| < r$. Then there exists $|a| < b < r$ for which the above inequality holds. Therefore for any test function φ supported in some interval $[-R, R]$, there exists a constant M such that

$$|(f^{(n)}, \varphi)| \leq \int_{-R}^{R} |f^{(n)}(x)\varphi(x)|\, dx \leq M b^{-n} n! \int |\varphi(x)|\, dx\,.$$

---

[5]S.G. Krantz and H.R. Parks, A Primer Of Real Analytic Functions (2nd Edition)

This inequality implies that the power series for $u(a)$ converges absolutely for any test function φ by the comparison test:

$$\sum_{n\geq 0}\frac{|a|^n}{n!}|(f^{(n)},\varphi)|\leq M\int|\varphi(x)|\,dx\sum_{n\geq 0}\Big(\frac{|a|}{b}\Big)^n<\infty$$

because $|a|<b$. Thus, for a regular distribution defined by a real analytical function that has a holomorphic extension into a strip in the complex plane the asymptotic Taylor expansion is a convergent series in $\mathcal{D}'$. It turns out that the converse is also true, and the assertion can also be extended to distributions of several variables[6].

THEOREM **24.1**. *Let $f\in\mathcal{D}'$ and $y\in\mathbb{R}^N$ be a vector with non-zero components. The asymptotic Taylor expansion*

$$f(x+ay)\sim\sum_{n\geq 0}\frac{a^n}{n!}\,(y,D_x)^nf(x)$$

along the line ay, $a\in\mathbb{R}$, is a convergent Taylor series in $\mathcal{D}'$ if and only if there exist $r_j>0$, $j=1,2,....,N$ such that f is a real analytic function on $\mathbb{R}^N$ which has a holomorphic extension $f(z)$, $z\in\mathbb{C}^N$, into a strip $|\mathrm{Im}\,z_j|<r_j$.

So, the class of distributions for which the asymptotic Taylor expansion define a convergent series in $\mathcal{D}'$ consists of regular distributions defined by rather smooth (holomorphic) functions. Nonetheless they are useful in applications for distributional solutions of partial differential equations. As an example, the multipole asymptotic expansion of a potential of the electric or gravitational field created by a collection of point particles is studied below in Sec. **24.5**. Asymptotic expansions of distributions are not limited to Taylor expansions[7]. An example is given in the next section.

24.4. Pizzetti's formula. Let φ be a smooth function on $\mathbb{R}^3$. Consider its integral average over the sphere $|x|=a$ in the limit $a\to 0^+$:

$$\frac{1}{4\pi a^2}\oint_{|x|=a}\varphi(x)\,dS=\frac{1}{4\pi}\oint_{|y|=1}\varphi(ay)\,dS\sim\sum_{n=0}^{\infty}\frac{\Delta^n\varphi(0)}{(2n+1)!}\,a^{2n}\,,$$

where the asymptotic expansion is obtained by a formal expansion of $\varphi(ay)$ into a Taylor series with a subsequent evaluation of the integral of products $y_{j_1}\cdots y_{j_{2n}}$ in each term of the series, which can be done in

[6] B. Stanković, J. Math. Anal. Appl. 203 (1996) 31-37

[7] For additional reading see, e.g., S. Pipilović, B. Stanković, and J. Vindas, Asymptotic behavior of generalized functions, WSCP, 2012

spherical coordinates (as shown below). This result was obtained by Pizzetti in 1909. It has a generalization to $\mathbb{R}^N$:

$$\frac{1}{\sigma_N}\oint_{|y|=1}\varphi(ay)\,dS\sim\sum_{n=0}^{\infty}\frac{\Delta^n\varphi(0)}{2^n n!N(N+2)\cdots(N+2n-2)}\,a^{2n}\,.$$

Recall a spherical delta function δ_{S_a} introduced in Sec.**15.5**. If φ is a test function, then the left-hand side of Pizzetti's equation is given by (δ_{S_a},φ) divided by the surface area of the sphere. Since $\Delta^n\varphi(0)=(\Delta^n\delta,\varphi)$, Pizzetti's equation is nothing but an asymptotic expansion of the spherical delta function in the limit of zero radius:

$$\frac{1}{\sigma_N a^{N-1}}\delta_{S_a}(x)\sim\sum_{n=0}^{\infty}\frac{\Delta^n\delta(x)}{2^n n!N(N+2)\cdots(N+2n-2)}\,a^{2n}\,.$$

Note that the series cannot converge to the spherical delta function in $\mathcal{D}'$ because each term in the series is supported at the origin $x=0$ whereas δ_{S_a} is supported on the sphere $|x|=a>0$ so that all terms of the series vanish on any test function whose support contains the sphere $|x|=a$ but does not contain $x=0$. In fact, the series does not converge in $\mathcal{D}'$ at all by the structure theorem for distributions supported at a single point.

So, the above relation is *not* an equality in $\mathcal{D}'$ and must be viewed only in the asymptotic sense. For example, the first and second terms in the expansion are understood as the distributional limits:

$$\lim_{a\to0^+}\frac{1}{\sigma_N a^{N-1}}\delta_{S_a}(x)=\delta(x)\,,$$
$$\lim_{a\to0^+}\frac{1}{a^2}\Big(\frac{1}{\sigma_N a^{N-1}}\,\delta_{S_a}(x)-\delta(x)\Big)=\frac{1}{2N}\,\Delta\delta(x)\,.$$

One can also write

$$\frac{1}{\sigma_N a^{N-1}}\delta_{S_a}(x)=\delta(x)+O'(a^2)\,,$$
$$\frac{1}{\sigma_N a^{N-1}}\,\delta_{S_a}(x)=\delta(x)-\frac{a^2}{2N}\,\Delta\delta(x)+O'(a^4)\,,$$

where the symbol $O'(a^n)$ is defined in the previous section. The terms of higher order in a are interpreted similarly.

The first limit follows from the integral mean value theorem. The second one means that the Laplacian of a smooth function φ at a point can be found by using the mean values of φ on a sphere centered at the point and the value of φ at the center of the sphere:

$$\lim_{a\to 0^+} \frac{1}{a^2}\Big(\frac{1}{\sigma_{_N} a^{N-1}}\int_{|x|=a} \varphi(x)\,dS - \varphi(0)\Big) = \frac{1}{2N}\Delta\varphi(0)\,.$$

If the limit is omitted, then, according to the asymptotic expansion, the relation can be used to approximate $\Delta\varphi$ with accuracy of order $O(a^2)$ as $a \to 0^+$. Let us prove this relation. The other terms in the asymptotic expansion can be calculated in a similar way.

Using Einstein's summation rule for repeated indices and changing variables $x = ay$ in the left-hand side, the latter is reduced to

$$\begin{aligned}&\frac{1}{\sigma_{_N} a^2}\int_{|y|=1}\Big(\varphi(ay)-\varphi(0)\Big)\,dS\\ &= \frac{1}{\sigma_{_N} a^2}\int_{|y|=1}\Big(a\partial_j\varphi(0)\hat{y}_j + \frac{a^2}{2}\partial_j\partial_k\varphi(0)\hat{y}_j\hat{y}_k + O(a^3)\Big)\,dS\end{aligned}$$

and the result follows because

$$\int_{|y|=1}\hat{y}_j\,dS = 0\,,$$

$$\int_{|y|=1}\hat{y}_j\hat{y}_k\,dS = \delta_{jk}\int_{|y|=1}\hat{y}_j^2\,dS = \frac{\sigma_{_N}}{N}\,\delta_{jk}\,. \tag{24.3}$$

The integrals should be invariant under rotations about the origin. Therefore the first integral vanishes whereas the second one must be proportional to δ_{jk} because $U_{j'j}U_{k'k}\delta_{jk} = (UU^T)_{j'k'} = \delta_{j'k'}$ for any orthogonal matrix U, and no other matrix $d_{jk} \neq \delta_{jk}$ with this property exists. The last integral does not depend on j by the rotational symmetry. It is therefore convenient to take the component that is the projection of $\hat{y}$ onto the axis of a spherical coordinate system (see Sec.**8.3.1**). This component is equal to $\cos(\phi)$. Then the integral over the sphere is reduced to an iterated integral over the $N-2$ dimensional sphere that lies in the $N-1$ hyperplane perpendicular to the selected axis and the integral with respect to ϕ:

$$\begin{aligned}\int_{|y|=1}\hat{y}_j^2\,dS &= \sigma_{_{N-1}}\int_0^\pi \cos^2(\phi)\sin^{N-2}(\phi)\,d\phi\\ &= \sigma_{_{N-1}}\int_0^1 s^{\frac{1}{2}}(1-s)^{\frac{N-3}{2}}\,ds = \sigma_{_{N-1}}B\Big(\frac{N-1}{2},\frac{3}{2}\Big) = \frac{\sigma_{_N}}{N}\end{aligned}$$

where $s = \cos^2(\phi)$, $B(p,q)$ is the Euler beta function

$$B(p,q) = \int_0^1 (1-s)^{p-1}s^{q-1}ds = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

and the explicit form of σ_N was used.

24.5. Multipole expansion in electrostatics. In Sec.**10.7**, it was shown that an electrostatic potential created by charges compactly distributed with a sufficiently smooth density ρ is given by the convolution integral

$$u(x) = -\frac{1}{4\pi}\int \frac{\rho(y)\,d^3y}{|x-y|} = -\frac{1}{4\pi}\,\frac{1}{|x|} * \rho(x)\,.$$

If support of ρ lies in a ball $|y| \leq R$, let us investigate u for $R \ll |x|$. Using the Taylor expansion

$$\frac{1}{|x-y|} = \frac{1}{|x|} + \frac{(x,y)}{|x|^3} + \frac{3(x,y)^2 - |y|^2|x|^2}{2|x|^5} + O\Big(\frac{1}{|x|^4}\Big)\,,$$

in the integral representation of the solution, one infers that

$$u(x) = -\frac{Q}{4\pi|x|} - \frac{(x,P)}{4\pi|x|^3} - \frac{(x,Mx)}{|x|^5} + O\Big(\frac{R^3}{|x|^4}\Big)\,, \qquad (\mathbf{24}.4)$$

where the scalar Q, vector P, and matrix M are called the total charge, dipole and quadrupole moments, respectively, and they are given by

$$Q = \int \rho(y)\,d^3y\,, \quad P_j = \int y_j\rho(y)\,d^3y\,,$$
$$M_{ij} = \frac{1}{2}\int (3y_iy_j - |y|^2\delta_{ij})\rho(y)\,d^3y\,.$$

The corresponding terms in the asymptotic expansion are called *Coulomb, dipole, and quadrupole* potentials. The expansion can be continued to obtain the so-called *multipole asymptotic expansion* of the solution to the Poisson equation.

Consider a finite collection point-like electric charges q_p positioned at x_p. Then the charge density can be modeled by

$$\rho(x) = \sum_p q_p\delta(x - x_p)\,.$$

Define a diameter of the system by $R = \max_{p,p'}|x_p - x_{p'}|$. It follows from the asymptotic Taylor expansion of a shifted delta function that in the limit $R \to 0^+$, the density has the following asymptotic expansion

$$\rho(x) = Q\delta(x) - (P,\nabla)\delta(x) + (\nabla,\tilde{M}\nabla)\delta(x) + O'(R^3)\,,$$

where $O'(R^3)$ is a distribution such that $(O'(R^3),\varphi) = O(R^3)$ for any test function φ, and

$$Q = \sum_p q_p\,, \quad P_j = \sum_p q_p x_{pj}\,, \quad \tilde{M}_{ij} = \frac{1}{2}\sum_p q_p x_{pi} x_{pj}\,.$$

The terms in the asymptotic expansion of ρ are called the *Coulomb, dipole, and quadrupole charge densities*, respectively, for a system of point charges. Let us find the corresponding expansion for a distributional solution to the Poisson equation $\Delta u = \rho$ in the limit $R \to 0^+$.

For any distribution ρ with a bounded support, its regularization $\rho_a = \phi_a * f$ is a test function. Then it follows from (**21.7**) that

$$\rho_a = \phi_a * \rho = \phi_a * \Delta u = \Delta(\phi_a * u) = \Delta u_a\,,$$

where u_a is a regularization of the distributional solution u. Therefore

$$u_a(x) = -\frac{1}{4\pi}\,\frac{1}{|x|} * \rho_a(x) \;\to\; u(x) \quad \text{in } \mathcal{D}'$$

as $a \to \infty$. So, to find u, one has to calculate the limit.

If δ_a is a regularization of δ, then by (**21.7**) ρ_a is a linear combination of δ_a and its derivatives. On the other hand, the convolution $\frac{1}{|x|} * \delta_a = \delta_a * \frac{1}{|x|}$ is a regularization of the distribution $\frac{1}{|x|}$. Therefore by (**21.7**)

$$\frac{1}{|x|} * D^\beta \delta_a = D^\beta(\delta_a * \frac{1}{|x|}) = \delta_a * D^\beta \frac{1}{|x|} \;\to\; D^\beta \frac{1}{|x|} \quad \text{in } \mathcal{D}'$$

as $a \to 0^+$, and, hence,

$$u(x) = -\frac{1}{4\pi}\Big(\frac{Q}{|x|} - (P, \nabla)\,\frac{1}{|x|} + (\nabla, \tilde{M}\nabla)\,\frac{1}{|x|}\Big) + O\Big(\frac{R^3}{|x|^4}\Big)\,.$$

The needed distributional derivatives are given in (**21.11**) and (**21.12**). They coincide with the corresponding classical derivatives if $x \neq 0$. By substituting the derivatives, one can see that the distributional potential coincides with (**24.4**) for $x \neq 0$ where

$$M_{ij} = 3\tilde{M}_{ij} - \mathrm{Tr}(\tilde{M})\,\delta_{ij} = \frac{1}{2}\sum_p q_p\Big(3x_{pi}x_{pj} - |x_p|^2\delta_{ij}\Big)$$

is the quadrupole moment of a system of point-like electric charges. High-order terms in the multipole asymptotic expansion can be found by using the corresponding higher-order terms in the asymptotic Taylor expansion of a shifted delta function.

In the asymptotic region $|x| \to \infty$ of the potential, the leading contribution is given by the Coulomb potential if the total charge of the system is not zero. If the total charge is equal to zero, then the leading term is given by the dipole potential. It takes into account a separation of between the mean positions of all positive and negative charges. Let $q = \sum_{q_p>0} q_p$ be the total positive charge. Then the total negative charge is equal to $-q$. The mean positions of positive and negative charges are defined by

$$x_+ = \frac{1}{q}\sum_{q_p>0} q_p x_p\,, \quad x_- = \frac{1}{q}\sum_{q_p<0} |q_p| x_p\,,$$

so that the dipole moment is $P = q(x_+ - x_-)$. The vector P is directed from negative to positive charges. The dipole potential is the difference of two Coulomb potentials created by point opposite charges located at x_+ and x_- in the limit when $|x_+ - x_-| \to 0$:

$$\rho(x) = q\delta(x - x_+) - q\delta(x - x_-) = -(P, \nabla)\delta(x) + O'(|x_+ - x_-|^2)\,.$$

24.6. Exercises.

1. Prove the asymptotic relations in $\mathcal{D}'(\mathbb{R})$

$$\text{(i)} \quad \delta(x + a) - \delta(x - a) = 2a\delta'(x) + O'(a^3)\,,$$
$$\text{(ii)} \quad \theta(x + a) = \theta(x) + a\delta(x) + O'(a^2)$$

as $a \to 0$.

2. Find a distribution $g(x; a) \in \mathcal{D}'(\mathbb{R})$ such that

$$\text{(i)} \quad \frac{1}{x + a + i0} = \frac{1}{x + i0} + g(x; a) + O'(a^3)\,,$$
$$\text{(ii)} \quad \mathcal{P}\frac{1}{x + a} = \mathcal{P}\frac{1}{x} + g(x; a) + O'(a^n)\,, \quad n \geq 1\,.$$

as $a \to 0$.

3. Find a distribution $g \in \mathcal{D}'(\mathbb{R}^N)$, $N > 2$, such that

$$\frac{1}{|x|^2 - m^2 + i0} = \frac{1}{|x|^2} + m^2 g(x) + O'(m^4)\,.$$

25. Product of distributions

The product of classical functions is defined as the product of their pointwise values: $(fg)(x) = f(x)g(x)$. This product is commutative and associative. Is it possible to define a commutative and associative product in the space of all distributions? The answer is not straightforward because distributions do not have pointwise values.

Suppose that such a product exists. Then if this products agrees with the product of classical functions, at least for regular distributions, then there is an obvious problem because the product of two locally integrable functions is not generally locally integrable. For example, the product of $f(x) = g(x) = |x|^{-1/2}$ is $|x|^{-1}$ which is not locally integrable on $\mathbb{R}$. Thus, the product of regular distributions cannot be the product of the corresponding classical functions. If one wants to define a product of distributions as close as possible to the product of classical functions, then one can demand at least that this product agrees with (**18.1**) that defines the product of distributions one of which is a smooth function.

25.1. No-go theorem for a product of distributions. Let the dot be used to denote a product in $\mathcal{D}'$ (whether it agrees with (**18.1**) or not), then the following properties are required to hold for any distributions f, g, and h:

$$\begin{aligned} \text{existence}: &\quad f\cdot g \in \mathcal{D}' \\ \text{commutativity}: &\quad f\cdot g = g\cdot f \\ \text{associativity}: &\quad f\cdot(g\cdot h) = (f\cdot g)\cdot h \end{aligned}$$

The following theorem due to L. Schwartz holds.

Theorem **25.1**. (L. Schwartz)
There exist no product in $\mathcal{D}'$ that is commutative and associative.

The assertion can easily be understood in the case of one-variable distributions if, in addition, it is assumed that the rule of multiplication of a distribution by a smooth function agrees with the distributional multiplication:

$$a(x)\cdot f(x) = a(x)f(x)\,, \quad f\in\mathcal{D}'\,, \quad a\in C^\infty\,.$$

Then the following chain of equalities must hold:

$$\begin{aligned} \delta(x) = 1\cdot\delta(x) &= \Big(x\cdot\mathcal{P}\frac{1}{x}\Big)\cdot\delta(x) = \Big(\mathcal{P}\frac{1}{x}\cdot x\Big)\cdot\delta(x) \\ &= \mathcal{P}\frac{1}{x}\cdot\Big(x\cdot\delta(x)\Big) = \mathcal{P}\frac{1}{x}\cdot 0 = 0\,. \end{aligned}$$

Here the second equality follows from $x\mathcal{P}\frac{1}{x} = 1$ in $\mathcal{D}'$, the third equality is a consequence of the commutativity, the forth equality is deduced from the associativity, and finally the distributional relation $x\delta(x) = 0$, commutativity of the product, and multiplication by the zero function were used. The delta-function is not equal to the zero distribution and, hence, no product with the said properties exists in the space of distributions.

This no-go theorem does not preclude us from constructing a product of distributions on some subsets of $\mathcal{D}'$. The very definition of the product depends on features (or properties) of distributions that are required to be preserved for the product. Here the approach based on the *localization theorem* for distribution will be discussed. This product is closest to the product of classical functions. Another approach based on the *Fourier transform of distributions* will be discussed later. It is used in quantum field theories. It should also be noted that there exists an associative and commutative product in a subset of $\mathcal{D}'$ known as a *convolution*, which is important in applications to partial differential equations. It also will be studied later in detail.

25.2. Partitions of unity. A *partition of unity* for a set Ω is a sum (or series) of test functions that is equal to 1 in Ω. Of course, one should show that a partition of unity exists. Let $\Omega = \mathbb{R}$. Then Ω can be viewed as the union of open bounded intervals $\Omega_n = (n, n+2)$, where n ranges over all integers. Each Ω_n is the proper subset of Ω. Similarly, $\Omega = (0,\infty)$ can aslo be represented as a countable union of open bounded intervals $\Omega_n = (\frac{1}{n}, \frac{2}{n})$ for positive integers $n > 0$, and $\Omega_n = (1-n, 3-n)$ if $n \leq 0$. Each open interval Ω_n is a proper subset of Ω. Open sets that are unions of open bounded sets that are proper subsets in the union can be constructed in $\mathbb{R}^N$. Then the following assertion holds.

Let an open set $\Omega \subseteq \mathbb{R}^N$ be the union of open bounded sets Ω_n, $n = 1, 2, \ldots$, where every Ω_n is a proper subset of Ω,

$$\Omega = \bigcup_{n\geq 1} \Omega_n\,, \qquad \overline{\Omega}_n \subset \Omega\,,$$

such that any compact intersects only finitely many of Ω_n. Then there exist test functions $\varphi_n \in \mathcal{D}(\Omega_n)$ such that

$$\sum_n \varphi_n(x) = 1\,, \quad x \in \Omega\,.$$

To prove the assertion, consider the set K_1 that consists of all points of Ω_1 that are not in any Ω_n, $n \geq 2$. All boundary points of Ω_1 must be in at least one of the other sets Ω_n, $n \geq 2$, because Ω_1 is a proper subset of Ω (the boundary $\partial\Omega_1 \subset \Omega$) and, furthermore, there are only finitely many of Ω_n that intersect Ω_1. Therefore K_1 is a closed and bounded subset of Ω_1 and, hence, there exists an open set Ω_1' that contains K_1 and is a proper subset of Ω_1 (see the proof of Corollary **14.1**). If Ω_1 is replaced by its open subset Ω_1' in the collection $\{\Omega_n\}$, the new collection has the same property:

$$\Omega = \Omega_1' \bigcup \Big(\bigcup_{n\geq 2} \Omega_n \Big) .$$

In other words, Ω_1 can be reduced (but not too much because $K_1 \subset \Omega_1' \subset \Omega_1$) so that the union is still equal to Ω.

In the new collection, the same reduction operation is applied to obtain an open proper subset $\Omega_2' \subset \Omega_2$. The union of the new collection where Ω_2 is replaced by Ω_2' is equal to Ω. Continuing this reduction procedure, a new collection $\{\Omega_n'\}$ is obtained such that

$$\Omega = \bigcup_{n\geq 1} \Omega_n' ,$$

where Ω_n' is a proper open subset of Ω_n. By Corollary **14.1**, there exists a test function $\eta_n \in \mathcal{D}(\Omega_n)$ such that

$$0 \leq \eta_n(x) \leq 1 , \quad \eta_n(x) = 1 , \; x \in \Omega_n' \subset \Omega_n .$$

Then the required decomposition of the unit function is given by

$$\varphi_n(x) = \frac{\eta_n(x)}{\eta(x)} , \quad \eta(x) = \sum_{k\geq 1} \eta_k(x) \geq 1 .$$

The series converges because any $x \in \Omega$ belongs to finitely many Ω_n'.

Corollary **25.4**. *Let K be a compact in $\mathbb{R}^N$ that is covered by finitely many open bounded sets Ω_n*

$$K \subset \bigcup_{n=1}^{m} \Omega_n .$$

Then there exists a family of test functions $\psi_n \in \mathcal{D}(\Omega_n)$ taking their values in $[0, 1]$ such that

$$\sum_{n=1}^{m} \varphi_n(x) = 1 , \quad x \in K .$$

Let $\{\Omega_n'\}_1^m$ be a new cover of K such that Ω_n' is a proper open subset of Ω_n for every n. The assertion would follow from the partition of unity constructed above for an open set $\Omega = \cup_{n=1}^m \Omega_n' \supset K$. Let us show that such a cover exists. In every Ω_n, take a subset of points whose distance to the boundary $\partial\Omega_n$ is greater than $\delta > 0$:

$$\Omega_{n\delta} = \{x \in \Omega_n \,|\, d(x, \partial\Omega_n) > \delta\} .$$

Then open sets $\Omega_{n\delta}$, $n = 1, 2, ..., m$ and all $\delta > 0$, also form an open cover for K. By the Heine-Borel theorem, any open cover of a compact has a finite subcover. For every n, this finite subcover either contains $\Omega_{n\delta_n}$ for some $\delta_n > 0$ or does not contain Ω_n at all. So, put $\Omega'_n = \Omega_{n\delta_n}$ in the former case and Ω'_n is any proper open subset of Ω_n in the latter case. Then $\Omega = \cup_{n=1}^m \Omega'_n \supset K$ as required.

25.3. The localization theorem for distributions. An ordinary function is zero in a region if and only if it vanishes at every point of the region. In contrast to a function, it only makes sense to say that a distribution can vanish in a neighborhood of a point. Clearly, if a distribution vanishes in an open set Ω, then it vanishes in a neighborhood of any point of Ω. It turns out that the converse is also true.

PROPOSITION **25.6**. *In order for a distribution to vanish in an open set it is necessary and sufficient that the distribution vanishes in a neighborhood of every point of the set:*

$$f(x) = 0\,, \quad x \in \Omega \quad \Leftrightarrow \quad f(x) = 0\,, \quad x \in B_a(x_0)\,, \quad x_0 \in \Omega\,.$$

To prove this assertion, take a test function $\varphi \in \mathcal{D}(\Omega)$. Its support $K = \operatorname{supp}\varphi$ is compact in Ω. Then the union of neighborhoods $B_\delta(x_0)$ over all $x_0 \in K$ covers K (δ generally depends on x_0). By the Heine-Borel theorem, this cover contains a finite subcover of K, say, $\Omega_k = B_{\delta_k}(x_k)$, $k = 1, 2, ..., n$. By Corollary **25.4**, there exist test functions $\varphi_k \in \mathcal{D}(\Omega_k)$ such that

$$\varphi_1(x) + \varphi_2(x) + \cdots + \varphi_n(x) = 1\,, \quad x \in K$$

Then $\phi_k = \varphi_k\varphi$ is a test function with support in Ω_k so that by the hypothesis, $(f, \phi_k) = 0$. Therefore

$$(f, \varphi) = \sum_{k=1}^{n} (f, \varphi_k\varphi) = \sum_{k=1}^{n} (f, \phi_k) = 0\,.$$

which means that $f(x) = 0$ in Ω.

THEOREM **25.2**. (Localization theorem for distributions)
Let $\{\Omega_\alpha\}$ be an arbitrary collection of open sets whose union contains an open set $\Omega \subseteq \mathbb{R}^N$, and for every α there exists a distribution from $\mathcal{D}'(\Omega_\alpha)$ such that $f_\alpha = f_\beta$ on $\Omega_\alpha \cap \Omega_\beta$ whenever the intersection is not

empty. Then there exists a unique distribution $f \in \mathcal{D}'(\Omega)$ *such that* $f = f_\alpha$ *on* Ω_α *for all* α.

First, it should be noted that if f exists, then it is unique. Indeed, if f_1 and f_2 are are two such distributions, then its difference $f = f_1 - f_2$ vanishes in every Ω_α and, hence, vanishes in the whole Ω by Proposition **25.6**. So, it is sufficient to construct one distribution f with the said properties.

Recall that the support of a test function from $\mathcal{D}(\Omega)$ is a proper subset of $\Omega_R = \Omega \cap B_R$ for some ball B_R. Then one can always find a compact set $K \subset \Omega \cap B_R$ that contain the support of φ. For example, one can remove a neighborhood of the boundary $\partial\Omega_R$ of a sufficiently small radius from Ω_R. Since $\operatorname{supp}\varphi \subseteq K \subset \Omega$, by the Heine-Borel theorem there exists a finite subcollection of open sets $\{\Omega_k\}_{k=1}^n$ whose union contains K (here $\alpha_k = k$ for brevity). By Corollary **25.4**, there exist test functions $\varphi_k \in \mathcal{D}(\Omega_k)$, whose sum is equal to 1 in K and, hence, in $\operatorname{supp}\varphi$. By the hypothesis, $f = f_k$ in Ω_k and, hence, $(f, \phi_k) = (f_k, \phi_k)$ where $\phi_k = \varphi_k\varphi \in \mathcal{D}(\Omega_k)$ for any $\varphi \in \mathcal{D}(\Omega)$. Therefore the value of f on φ can be calculated via the values of f_k:

$$(f,\varphi) = \sum_{k=1}^{n}(f,\varphi_k\varphi) = \sum_{k=1}^{n}(f,\phi_k) = \sum_{k=1}^{n}(f_k,\phi_k)\,. \tag{25.1}$$

The number (f, φ) seems to depend on the choice of K, its finite cover $\{\Omega_k\}$, and a partition of unity associated with the cover. But (f, φ) must be unique for every φ in order to define a functional on $\mathcal{D}(\Omega)$. Despite its appearance, the rule (**25.1**) does define a linear continuous functional. First, let us show that for any choice of K, the number (f, φ) is independent of the choice of $\{\Omega_k\}$ and $\{\varphi_k\}$. Let $\{\tilde{\Omega}_k\}_{k=1}^m$ be a different finite cover of K and $\{\tilde{\varphi}_k\}_{k=1}^m$ be the associated partition of unity. For any test function $\varphi \in \mathcal{D}(\Omega)$, put $\phi_k = \varphi_k\varphi$ and $\tilde{\phi}_j = \tilde{\varphi}_j\varphi$. Then the product $\tilde{\varphi}_j\phi_k$ is a test function from $\mathcal{D}(\tilde{\Omega}_j \cap \Omega_k)$ and vanishes if $\tilde{\Omega}_j$ and Ω_k do not intersect. It follows from the hypothesis that

$$f_k(x) = f_j(x)\,,\quad x \in \tilde{\Omega}_j \cap \Omega_k \;\Rightarrow\; (f_k, \tilde{\varphi}_j\phi_k) = (f_j, \tilde{\varphi}_j\phi_k) = (f_j, \varphi_k\tilde{\phi}_j)\,.$$

Since the sum of $\tilde{\varphi}_j$ is equal to 1 on K and, hence, on $\operatorname{supp}\phi_k \subseteq K$ and similarly, the sum of φ_k is equal to 1 on $\operatorname{supp}\tilde{\phi}_j \subseteq K$, one infers that

$$\sum_{k=1}^{n}(f_k,\phi_k) = \sum_{j=1}^{m}\sum_{k=1}^{n}(f_k,\tilde{\varphi}_j\phi_k) = \sum_{j=1}^{m}\sum_{k=1}^{n}(f_j,\varphi_k\tilde{\phi}_j) = \sum_{j=1}^{m}(f_j,\tilde{\phi}_j)$$

This shows that the number (f, φ) is independent of the choice of the subcover of K and the associated partition of unity.

Let K and $\tilde{K}$ be any two closed bounded subsets of Ω that contain $\operatorname{supp}\varphi$. Then their union $K \cup \tilde{K}$ also contain $\operatorname{supp}\varphi$. If $\{\Omega_k\}$ and $\{\tilde{\Omega}_j\}$ are finite covers of K and $\tilde{K}$, respectively, their union is also a finite cover of K and of $\tilde{K}$. Let $(f, \varphi)_1$, $(f, \varphi)_2$, and $(f, \varphi)_{12}$ be the values of (f, φ) calculated for the cover of K, the cover of $\tilde{K}$, and the union of two covers. Since (f, φ) does not depend on the choice of a cover of K, $(f, \varphi)_1 = (f, \varphi)_{12}$, and similarly $(f, \varphi)_2 = (f, \varphi)_{12}$, which implies that $(f, \varphi)_1 = (f, \varphi)_2$. Thus, the rule (**25.1**) defines a functional on $\mathcal{D}(\Omega)$. Its continuity and linearity follows from continuity and linearity of f_α.

It remains to verify that $f = f_\alpha$ in Ω_α for all α. Since $\mathcal{D}(\Omega_\alpha) \subset \mathcal{D}(\Omega)$, for any test function $\varphi_\alpha \in \mathcal{D}(\Omega_\alpha)$, there exists a compact $K \subset \Omega$ such that $\operatorname{supp}\varphi_\alpha \subseteq K$. Then for any partition of unity $\{\varphi_k\}$ for K, the product $\varphi_k\varphi_\alpha$ is a test function from $\mathcal{D}(\Omega_k \cap \Omega_\alpha)$, and

$$(f_k, \varphi_k\varphi_\alpha) = (f_\alpha, \varphi_k\varphi_\alpha)$$

because $f_k = f_\alpha$ in the intersection $\Omega_k \cap \Omega_\alpha$. By the rule (**25.1**),

$$(f, \varphi_\alpha) = \sum_{k=1}^{n}(f_k, \varphi_k\varphi_\alpha) = \sum_{k=1}^{n}(f_\alpha, \varphi_k\varphi_\alpha) = (f_\alpha, \varphi_\alpha)\,.$$

as required.

Theorem **25.2** also shows that any distribution is uniquely defined by its local values (by values on test functions with with supports in a neighborhood of each point). The rule (**25.1**) is often referred to as a *piecewise gluing rule* for distributions. If a distribution is known on a collection of open sets that cover Ω, then a unique distribution on Ω can be obtained by (**25.1**) by gluing its values in the open sets.

25.3.1. Example. *Let $\{x_n\}$ be a sequence in $\mathbb{R}$ that has no limit points, that is, any bounded interval contains finitely many points of the sequence. Suppose $f \in \mathcal{D}'$ is a distribution that is equal to $\delta(x - x_n)$ in any open interval that contains x_n and no other points of the sequence, and $f(x) = 0$ in any open interval that has no points of the sequence. Then*

$$f(x) = \sum_n \delta(x - x_n)\,.$$

Let $x_n < x_m$ be two neighboring points of the sequence (there are no points of the sequence in (x_n, x_m)). Let I_n and I_m be open non-intersecting intervals containing x_n and x_m, respectively. Let $I_n^0 \subset (x_n, x_m)$ be an open subinterval that intersects with I_n and I_m. If the sequence is bounded from below, then $x_k < x_n$ for all $n \neq k$ and some k.

In this case, $I_k = (-\infty, x_k + \delta)$ for some small enough $\delta > 0$. Similarly, if the sequence is bounded from above, then $x_n < x_m$ for all $n \neq m$ and some m. In this case, $I_k = (x_m - \delta, \infty)$ for some small enough $\delta > 0$. The union of all I_n and all I_n^0 is equal to $\mathbb{R}$. By construction

$$f(x) = \delta(x - x_n)\,, \quad x \in I_n\,, \quad f(x) = 0\,, \quad x \in I_n^0\,.$$

The support of a test function φ lies in $|x| \leq R$ which contains finitely points of the sequence, $|x_n| \leq R$. Therefore the interval $|x| \leq R$ is covered by finitely many I_n and I_n^0. Let φ_k and φ_n^0 be a partition of unit for $|x| \leq R$ where $\varphi_k \in \mathcal{D}(I_k)$ and $\varphi_n^0 \in \mathcal{D}(I_n^0)$:

$$\sum_{|x_k| \leq R} \varphi_k(x) + \sum_{|x_n| \leq R} \varphi_n^0(x) = 1\,, \quad |x| \leq R\,.$$

For each k, there exists a neighborhood of x_k that lies in I_k and does not intersect I_n^0. Therefore $\varphi_k(x_k) = 1$. By the rule (**25.1**)

$$(f, \varphi) = \sum_{|x_k| < R} \varphi_k(x_k)\varphi(x_k) = \sum_{|x_k| < R} \varphi(x_k)$$

which proves the assertion.

Proposition **25.7**. *Let f be a distribution from $\mathcal{D}'(\mathbb{R}^N)$ such that $f(x) = f_n(x)$ in Ω_n, $n = 1, 2, \ldots$, where open sets Ω_n are not intersecting with each other, and $f(x) = 0$ in any open set that does not intersect any Ω_n. Then*

$$f(x) = \sum_n f_n(x)\,.$$

A proof of this assertion is left to the reader as an exercise. It should also be noted that the above proposition applies to the case of a change of variables in distributions discussed in Sec.**20.3.5** and extends the result to any distribution (not just a delta-function).

25.4. Product of distributions by the localization method. Let f and g be distributions from $\mathcal{D}'(\Omega)$. Suppose that $\Omega = \Omega_f \cup \Omega_g$, where Ω_f and Ω_g are open sets and f is such that its reduction on $\mathcal{D}(\Omega_f)$ is a regular distribution defined by a C^∞ function a_f, that is,

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

and g is such that its reduction on $\mathcal{D}(\Omega_g)$ is a regular distribution defined by a C^∞ function a_g. Then the product $f(x)g(x)$ can be defined on Ω_f and Ω_g by

$$f(x)g(x) = a_f(x)g(x)\,, \quad x \in \Omega_f\,,$$
$$f(x)g(x) = a_g(x)f(x)\,, \quad x \in \Omega_g\,.$$

These products agrees on the intersection $\Omega_g \cap \Omega_f$. For any test function $\varphi \in \mathcal{D}(\Omega_g \cap \Omega_f)$, one infers that

$$(fg, \varphi) = (a_f g, \varphi) = (g, a_f\varphi) = (a_g, a_f\varphi) = (a_f a_g, \varphi)$$

so that $fg = a_f a_g$ in $\Omega_g \cap \Omega_f$. Thus, fg is a linear continuous functional on $\mathcal{D}(\Omega_f)$ and on $\mathcal{D}(\Omega_g)$. By the localization theorem, it is a linear continuous functional on $\mathcal{D}(\Omega)$ and, hence, $fg \in \mathcal{D}'(\Omega)$ defined by the rule (**25.1**)

$$(fg, \varphi) = (f, a_g\psi_g\varphi) + (g, a_f\psi_f\varphi)\,, \quad \varphi \in \mathcal{D}(\Omega)$$

where $\psi_f \in \mathcal{D}(\Omega_f)$ and $\psi_g \in \mathcal{D}(\Omega_g)$ such that

$$\psi_f(x) + \psi_g(x) = 1\,, \quad x \in \operatorname{supp}\varphi\,.$$

25.4.1. Example. Let $f(x) = \mathcal{P}\frac{1}{x}$ and $g(x) = \delta(x - x_0)$ where $x_0 > 0$. Then take $\Omega_f = (-\infty, x_0 - \delta)$ and $\Omega_g = (\delta, \infty)$ where $0 < \delta < \frac{x_0}{2}$. Then $\Omega = \Omega_f \cup \Omega_g = \mathbb{R}$. Then

$$\mathcal{P}\frac{1}{x} = \frac{1}{x}\,, \quad x > \delta\,; \quad \delta(x - x_0) = 0\,, \quad x < x_0 - \delta$$

Therefore the product fg exists in $\mathcal{D}'$ because

$$\mathcal{P}\frac{1}{x}\delta(x - x_0) = \frac{1}{x}\,\delta(x - x_0)\,, \quad x > \delta\,,$$
$$\mathcal{P}\frac{1}{x}\delta(x - x_0) = 0\,, \quad x < x_0 - \delta\,,$$

and for any $\varphi \in \mathcal{D}$,

$$\Big(\mathcal{P}\frac{1}{x}\delta(x - x_0), \varphi(x)\Big) = \Big(\delta(x - x_0), \frac{1}{x}\psi_g\varphi\Big) = \frac{1}{x_0}\psi_g(x_0)\varphi(x_0)$$

where $\psi_g(x_0) = 1$ because $x_0 \notin \Omega_f$ and $\psi_f \in \mathcal{D}(\Omega_f)$ so that $\psi_f(x_0) = 0$. Therefore

$$\mathcal{P}\frac{1}{x}\delta(x - x_0) = \frac{1}{x_0}\delta(x - x_0)\,.$$

25.4.2. Singular support of a distribution. The above idea can be extended to define the product fg of distributions f and g if f and g behave near every point in the same way as in the example considered above, that is, one of them is defined by a smooth function. Recall that the localization theorem guarantees that any distribution from $\mathcal{D}'(\Omega)$ is uniquely defined by its values in a neighborhood of each point of Ω by the rule (**25.1**). However, not any pair of distributions would have such a property. For example, the distributions $f(x) = \mathcal{P}\frac{1}{x}$ and $g(x) = \delta(x)$ cannot be represented by a smooth function in any neighborhood of $x = 0$. Therefore it would not be possible to define their product.

DEFINITION **25.1**. (Regular points and singular support of $f \in \mathcal{D}'$) *A point x_0 is called a regular point of a distribution f if there exists a neighborhood $U(x_0)$ such that f is equal to a smooth function $a_f \in C^\infty$ in it:*

$$(f, \varphi) = \int a_f(x)\varphi(x)\, d^N x\,, \quad \varphi \in \mathcal{D}(U(x_0))\,.$$

The complement of all regular points of f is called the singular support of f and denoted by $\operatorname{Ssupp} f$.

It follows from this definition that *the singular support of f is a closed subset in the support of f.*

For example, if $f(x) = \mathcal{P}\frac{1}{x}$, then its singular support consists of the single point $x = 0$. The singular support of spherical delta function coincides with its support being a sphere. If $f(x)$ is locally integrable, then it can have no regular point. For example, if $f(x)$ is nowhere differentiable, then its singular support coincides with the whole support.

25.4.3. Product of distributions with non-intersecting singular supports. Suppose that the singular supports of distributions f and g do not have common points:

$$\operatorname{Ssupp} f \cap \operatorname{Ssupp} g = \emptyset\,.$$

Then near any point x_0, one of them is equal to a smooth function and, hence, their product can be defined near x_0. Let a_f and a_g be C^∞ functions such that

$$f(x) = a_f(x)\,, \quad x \in U(x_0)\,, \quad x_0 \in \operatorname{Ssupp} g\,,$$
$$g(x) = a_g(x)\,, \quad x \in U(x_0)\,, \quad x_0 \in \operatorname{Ssupp} f\,.$$

Since the singular supports of f and g are closed sets in $\mathbb{R}^N$ and, by hypothesis, have no common points, the functions a_f and a_f exist for any x_0. Then in a neighborhood of x_0, the product is defined by

$$(fg)(x) = a_f(x)g(x)\,, \quad x \in U(x_0)\,, \quad x_0 \in \operatorname{Ssupp} g$$
$$(fg)(x) = a_g(x)f(x)\,, \quad x \in U(x_0)\,, \quad x_0 \in \operatorname{Ssupp} f$$

By the localization theorem, there exists a unique distribution $fg \in \mathcal{D}'$ whose reduction to a neighborhood of any point x_0 is given by the above distribution on $\mathcal{D}'(U(x_0))$. The rule (**25.1**) allows us to find (fg, φ) for any $\varphi \in \mathcal{D}$.

It should be noted that the condition that the singular supports of distributions have no common points is rather restrictive. For example, the product cannot be defined by the localization method if $f(x)$ and $g(x)$ are regular distributions that are not smooth near one common point even though the product $f(x)g(x)$ is a locally integrable function. For example, if $f(x) = g(x) = |x|$, $x \in \mathbb{R}^N$, then the singular supports of f and g consist of the single point $x = 0$. So, the localization method does not apply. However, $f(x)g(x) = |x|^2$ which is a smooth function, hence, defines a regular distribution. The product of some distributions with intersecting singular supports can be defined via their Fourier transforms as shown later.

25.5. Product of distributions via a regularization. Any distribution can be viewed as a distributional limit of a sequence of smooth functions obtained by a regularization of the distribution. Let f and g be distributions. Let $f_n \to f$ in $\mathcal{D}'$ where $\{f_n\} \subset C^\infty$. Then the product $f_n(x)g(x)$ is well defined in $\mathcal{D}'$. By the completeness theorem, if the sequence $(f_n g, \varphi)$ converges for any test function φ, then there exists a distribution h such that $(h, \varphi) = \lim_{n\to\infty}(f_n g, \varphi)$. So, it is tempting to define a product of any distributions by this limiting procedure:

$$(fg, \varphi) \overset{\text{def}}{=} \lim_{n\to\infty} (f_n g, \varphi) = \lim_{n\to\infty} (g, f_n\varphi)$$

provided the limit exists for any test function φ.

If the singular supports of f and g do not intersect, then this definition of the product agrees with the product defined by the localization method. Indeed, the convergence in $\mathcal{D}'$ implies the convergence in $\mathcal{D}'(U(x_0))$. Therefore $f_n \to f = a_f$ in $\mathcal{D}'(U(x_0))$ for any x_0 from the singular support of g. Furthermore $f_n g = f_n a_g$ in $U(x_0)$ if x_0 lies in the singular support of f and, hence, $f_n g \to f a_g$ in $\mathcal{D}'(U(x_0))$. The conclusion does not depend of the choice of a regularizing sequence f_n. It also does not matter which of the two distributions in the product fg is to be regularized in this case.

Suppose that the singular supports of f and g have common points. Then, first, the product defined via a regularization is not commutative, meaning that it can depend on which of the two distributions in the product is regularized. If one agrees that the distribution to the left in the product is to be regularized to find the product, then $fg \neq gf$ in general because the distributional limits of $f_n g$ and $g_n f$ may not be the same even if they exists. Second, the limit of $f_n g$ can depend on the choice of the sequence $f_n \to f$.

This is illustrated by the following example. Let $f(x) = \theta(x)$ and $g(x) = \delta(x)$. Let g be regularized by the hat function $\omega_a \to \delta$ as $a \to 0^+$. Then for any test function φ,

$$(\omega_a\theta, \varphi) = \int_0^\infty \omega_a(x)\varphi(x)\,dx = \int_0^\infty \omega_a(x)\Big(\varphi(x) - \varphi(0)\Big)\,dx + \frac{1}{2}\varphi(0)\,.$$

The integral in this identity vanishes in the limit $a \to 0^+$ because the support of ω_a is the interval $[-a, a]$ so that the integration interval can be reduced to $[0, a]$ and in this interval $|\varphi(x) - \varphi(0)| \leq |x| \sup|\varphi'| \leq a \sup|\varphi'|$ so that

$$\Big|\int_0^\infty \omega_a(x)\Big(\varphi(x) - \varphi(0)\Big)\,dx\Big| \leq \frac{1}{2}\sup|\varphi'|a \to 0$$

Thus, the product $\delta\theta$ is the limit

$$\omega_a(x)\theta(x) \to \frac{1}{2}\delta(x)$$

On the other hand, let η_a be a bump function for $[0, \infty)$ such that $\eta_a(x) = 1$ if $x > -a$ and $\eta_a(x) = 0$ if $x < -3a$. This function is the convolution of ω_a and the shifted step function $\theta(x+2a)$. Then $\eta_a \to \theta$ in $\mathcal{D}'$. Indeed for any test function

$$(\eta_a, \varphi) = \int_{-3a}^0 \eta_a(x)\varphi(x)\,dx + \int_0^\infty \varphi(x)\,dx \to (\theta, \varphi)$$

because $0 \leq \eta_a(x) \leq 1$ and the first integral vanishes in the limit $a \to 0^+$

$$\Big|\int_{-3a}^0 \eta_a(x)\varphi(x)\,dx\Big| \leq 3a \sup|\varphi| \to 0\,.$$

The product in the reversed order, $\theta\delta$, is the limit

$$\eta_a(x)\delta(x) = \eta_a(0)\delta(x) = \delta(x) \to \delta(x)$$

as $a \to 0^+$, so that

$$\theta(x)\delta(x) \neq \delta(x)\theta(x)\,.$$

Let us take a different regularization of θ obtained by the convolution $\omega_a * \theta$. Then

$$(\omega_a * \theta)(x)\delta(x) = (\omega_a * \theta)(0)\delta(x) = \int_{-\infty}^{0} \omega_a(y)\, dy\, \delta(x) = \frac{1}{2}\delta(x)$$

by the normalization property of the hat function so that the distributional limits of $(\omega_a * \theta)\delta$ and $\eta_a\delta$ are different whereas the distributional limits of $\omega_a * \theta$ and η_a are equal.

25.5.1. Remark. Products of distributions appear in treatments of nonlinear equations with distributional sources (e.g. a motion of a point particle in Einstein's general relativity) based on perturbation theory. Products of Green's functions for some differential operators occur in quantum (field) theory. Without a properly defined product of distributions, a mathematically consistent treatment of such problems is not possible.

25.6. Exercises.

1. Let $x_0 > 0$. Find the product of distributions

(i) $$\frac{1}{x \pm i0}\, \delta'(x - x_0)\,,$$

(ii) $$\mathcal{P}\frac{1}{|x|}\, \delta(x^2 - x_0^2)\,,$$

(iii) $$\frac{1}{x - x_0 + i0}\mathcal{P}\frac{1}{x^2}\,,$$

using the localization method. Give an explicit rule for the value of the product on any test function from $\mathcal{D}$. Express the answers for (i) and (ii) via a linear combination of shifted delta functions, and the answer for (iii) via a linear combination of Sokhotsky and principal value distributions.

2. Suppose that $a(x)$ is from C^∞ such that $a(0) \neq 0$, all zeros of a are simple, and the set of all zeros $\{x_n\}$ has no limit point. Show that there exists a regularization of $1/a(x)$ such that the product defined by the localization method

$$f(x) = \operatorname{Reg} \frac{1}{a(x)}\mathcal{P}\frac{1}{x}$$

is a solution to the equation $a(x)f(x) = \mathcal{P}\frac{1}{x}$.

3. Let $P(x)$ and $Q(x)$ be polynomials with no common zeros. Find distributional regularizations of the reciprocals of $P(x)$, $Q(x)$, and the product $P(x)Q(x)$ such that

$$\operatorname{Reg}\frac{1}{P(x)}\operatorname{Reg}\frac{1}{Q(x)} = \operatorname{Reg}\frac{1}{P(x)Q(x)}$$

where the product of distributions is defined by the localization method.

26. Algebraic distributional equations

Consider the following equation

$$a(x)f(x) = g(x)\,, \quad x \in \Omega \subseteq \mathbb{R}^N\,,$$

where a is a smooth function and g is a given distribution. The objective is to find the most general distribution $f \in \mathcal{D}'(\Omega)$ that satisfies this equation. The equation is linear and, hence, its general solution, if it exists, must have the form

$$f(x) = f_p(x) + h(x)\,, \quad a(x)h(x) = 0\,,$$

where $f_p(x)$ is a particular solution and h is a general solution to the associated homogeneous equation.

26.1. Division problem in distributional algebraic equations. For example, one has to find a distribution f that satisfies the equation

$$xf(x) = \mathcal{P}\frac{1}{x}\,.$$

The equation makes perfect sense because $xf(x)$ is a distribution for any distribution f. If one formally divides the equation by x, one would get a meaningless expression $f(x) = \frac{1}{x}\mathcal{P}\frac{1}{x}$. The function x^{-1} is not smooth and f is not a distribution. Alternatively, if the equation can be formally multiplied by $\mathcal{P}\frac{1}{x}$, then assuming associativity and commutativity of the product $\mathcal{P}\frac{1}{x}(xf) = (x\mathcal{P}\frac{1}{x})f = f$ so that $f = (\mathcal{P}\frac{1}{x})^2$. Apart from that no associative and commutative product exists in $\mathcal{D}'$ due to the Schwartz theorem and, hence, the formal algebraic manipulations in the left-hand side of the equation are not valid, the resulting expression for f is the squared principal value distribution which cannot be defined by the localization method. This simple example shows that *conventional algebraic rules for solving algebraic equations fail or make no sense in the case of distributions because there is no associative and commutative product in $\mathcal{D}'$ and it is not always possible to divide a distribution by a smooth function*. The latter is known as the *division problem for distributions*.

26.2. Localization method. Suppose that a is analytic. By the localization theorem, f can be recovered from its values in a neighborhood of any point. Let $a(x_0) \neq 0$. Then there exists a neighborhood $U(x_0)$ in which $a(x) \neq 0$ by continuity of a so that the reciprocal $1/a(x)$ is from C^∞ near x_0 and the solution exists and is unique:

$$f(x) = \frac{1}{a(x)}\,g(x)\,, \quad x \in U(x_0)\,.$$

Indeed, for any test function $\varphi \in \mathcal{D}(U(x_0))$, $\phi(x) = \varphi(x)/a(x)$ is also a test function from $\mathcal{D}(U(x_0))$ so that if f is a solution, then

$$(f, \varphi) = (f, a\phi) = (af, \phi) = (g, \phi) = \left(\frac{1}{a}\, g, \varphi\right),$$

as required. Suppose that $a(x_0) = 0$. Then $a(x) = (x - x_0)^n b(x)$ for some integer $n > 0$, where b is from class C^∞ and $b(x_0) \neq 0$. Therefore there exists a neighborhood $U(x_0)$ in which $b(x) \neq 0$. By the above analysis, the original equation is equivalent to the equation

$$(x - x_0)^n f(x) = \frac{1}{b(x)} g(x), \quad x \in U(x_0).$$

Thus, the problem is reduced to finding a general solution near points where analytic a has zeros. Let us first solve the associated homogeneous equation.

26.3. General distributional solution to $x^n f(x) = 0$. *Let n be a positive integer. Consider the equation*

$$x^n f(x) = 0, \quad f \in \mathcal{D}'(\mathbb{R})$$

Any solution to this equation must have the form

$$f(x) = \sum_{k=1}^{n-1} c_k\, \delta^{(k)}(x),$$

where c_k are constants.

To prove the assertion, one has to show that if f is a solution, then there exist constants c_k such that

$$(f, \varphi) = \sum_{k=1}^{n-1} c_k (\delta^{(k)}, \varphi) = \sum_{k=1}^{n-1} (-1)^k c_k \varphi^{(k)}(0)$$

for any test function φ. To do so, let us show first that for any test function $\varphi(x)$, there exists a test function $\psi(x)$ such that

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

where $p_{n-1}(x)$ is the Taylor polynomial for φ about $x = 0$ of order $n - 1$, and $\eta(x)$ is a test function with support in $|x| < a$, for some $a > 0$, and $\eta(x) = 1$ in a neighborhood of $x = 0$. If the support of φ is an interval $[0, R]$. Then $\varphi^{(m)}(0) = 0$ for any $m \geq 0$. By l'Hospital's rule it follows that

$$\lim_{x\to 0^+} \frac{\varphi(x)}{x^m} = \lim_{x\to 0^+} \frac{\varphi^{(m)}(x)}{m!} = 0$$

Therefore in this case

$$\psi(x) = \frac{\varphi(x)}{x^n} \in \mathcal{D}$$

is a test function with support being $[0, R]$. Suppose that a neighborhood of $x = 0$ lies in the support of φ. Then, if $x \neq 0$,

$$\psi(x) = \frac{1}{x^n}\left(\varphi(x) - \eta(x)\sum_{k=0}^{n-1}\frac{\varphi^{(k)}(0)}{k!}\,x^k\right)$$

The support of ψ is bounded because the supports of φ and η are bounded. The function ψ has continuous derivatives of any order for all $x \neq 0$. So, it is sufficient to check if $\psi(x)$ has a smooth extension to $x = 0$. For a sufficiently small a and any $m > n$, by Taylor's theorem

$$\varphi(x) = \sum_{k=0}^{m}\frac{\varphi^{(k)}(0)}{k!}\,x^k + O(x^{m+1}), \quad |x| < a$$

It follows that

$$\psi(x) = \sum_{k=n}^{m}\frac{\varphi^{(k)}(0)}{k!}\,x^{k-n} + O(x^{m-n+1}), \quad |x| < a$$

because $\eta(x) = 1$ if $|x| < a$. Since m is arbitrary, $\psi(x)$ is smooth near x. So, $\psi \in C^\infty$. In particular,

$$\psi^{(k)}(0) = \varphi^{(k+n)}(0)$$

for any $k \geq 0$.

Let $f(x)$ be a solution to the stated equation. Then

$$\begin{aligned}(f,\varphi) &= (f,\eta p_{n-1}) + (f, x^n\psi) = (f,\eta p_{n-1}) + (x^n f,\psi) = (f,\eta p_{n-1})\\ &= \sum_{k=0}^{n-1}\frac{(f,x^k\eta)}{k!}\,\varphi^{(k)}(0) = \sum_{k=0}^{n-1}\frac{(f,x^k\eta)}{k!}\,(-1)^k(\delta^{(k)},\varphi)\end{aligned}$$

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 $x = 0$.

An alternative proof. It follows from the equation $x^n f(x) = 0$ that f must be supported at a single point $x = 0$. By the structure theorem for such distributions, f is a linear combination of $\delta^{(k)}$, $k = 0, 1, ..., m$. Substituting a general linear combination of $\delta^{(k)}$ into the equation and using the results of Sec.**18.1**, one can show that only $\delta^{(k)}$ with $m \leq n-1$ contribute to f.

26.4. General solution to the homogeneous equation. Suppose that zeros of a form a sequence $\{x_n\}$. This sequence cannot have limit points by analyticity of a. Indeed, suppose that $\{x_n\}$ (or any its subsequence) converges to x_0. Then by continuity $a(x_0) = 0$ and by analyticity $a(x) = (x - x_0)^k b(x)$ for some positive integer k, where $b(x_0) \neq 0$. By continuity, $b(x) \neq 0$ in a neighborhood of x_0 and, hence, x_0 cannot be a limit point of $\{x_n\}$. By Proposition **25.7**, a general solution to the homogeneous equation reads

$$f(x) = \sum_n f_n(x)\,, \quad f_n(x) = \sum_{k=1}^{k_n} c_{nk}\delta^{(k)}(x - x_n)\,.$$

where c_{nk} are arbitrary constants and $k_n > 0$ is the order of zero x_n. The series series converges in $\mathcal{D}'$ because

$$(f, \varphi) = \sum_{|x_n|<R} (f_n, \varphi)\,, \quad \operatorname{supp}\varphi \subseteq [-R, R]\,,$$

and any bounded interval has only finitely many zeros of a since $\{x_n\}$ has no limit points.

For example, a general solution to the equation

$$\sin^2(x) f(x) = 0\,, \quad f \in \mathcal{D}'\,,$$

has the form

$$f(x) = \sum_{n=-\infty}^{\infty} \Big(A_n\delta(x - \pi n) + B_n\delta'(x - \pi n)\Big)$$

for any constants A_n and B_n.

26.5. Particular solution for smooth inhomogeneity. A particular solution can be found using the localization theorem in combination with the concept of a distributional regularization of singular functions if the inhomogeneity is a smooth function. The idea is first illustrated by a simple example.

26.5.1. Example. Let us find a general distributional solution to the equation

$$(x-a)(x-b)f(x)=1\,,\quad a\neq b\,.$$

A general solution to the associated homogeneous equation

$$(x-a)(x-b)h(x)=0$$

is obtained by the localization method,

$$h(x)=A\delta(x-a)+B\delta(x-b)\,,$$

for any constants A and B.

To find a particular solution, note that a formal particular solution $[(x-a)(x-b)]^{-1}$ is not a locally integrable function so it does not define a distribution. One has to find a distributional extension of this function to singular points with the properties

$$(x-a)(x-b)\,\mathrm{Reg}\,\frac{1}{(x-a)(x-b)}=1\,,$$
$$\mathrm{Reg}\,\frac{1}{(x-a)(x-b)}=\frac{1}{(x-a)(x-b)}\,,\quad x\in\Omega$$

in any open interval Ω that does not contain a and b. For example, the distributions

$$\mathcal{P}\frac{1}{x}\,,\qquad \frac{1}{x-i0^+}\,,\qquad \frac{1}{x+i0^+}$$

satisfy the equation $xg(x)=1$ and can be viewed as different distributional extensions of a formal solution $\frac{1}{x}$ to the singular point $x=0$.

Consider a partial fraction decomposition of the formal singular solution

$$\frac{1}{(x-a)(x-b)}=\frac{1}{a-b}\left(\frac{1}{x-a}-\frac{1}{x-b}\right)$$

Then the right side can be turned into a distribution by

$$g(x)=\mathrm{Reg}\,\frac{1}{(x-a)(x-b)}=\frac{1}{a-b}\left(\mathcal{P}\frac{1}{x-a}-\mathcal{P}\frac{1}{x-b}\right)$$

It is not difficult to verify that g is a required particular solution:

$$\begin{aligned}(x-a)(x-b)g(x) &= \frac{1}{a-b}\Bigg((x-b)(x-a)\mathcal{P}\frac{1}{x-a}\\ &\qquad -(x-a)(x-b)\mathcal{P}\frac{1}{x-b}\Bigg)\\ &= \frac{1}{a-b}\Big((x-b)-(x-a)\Big)=1\end{aligned}$$

A regularization or distributional extension is not unique. For example one can also use Sokhotsky distributions to regularize the formal solution

$$\mathrm{Reg}\,\frac{1}{(x-a)(x-b)} = \frac{1}{a-b}\left(\frac{1}{x-a\pm i0^+} - \frac{1}{x-b\pm i0^+}\right).$$

This method can be used to find a particular solution to the equation

$$P(x)f(x) = Q(x) \quad \Rightarrow \quad f(x) = \mathrm{Reg}\,\frac{Q(x)}{P(x)} \;\in\; \mathcal{D}',$$

where P and Q are polynomials. The rational function $Q(x)/P(x)$ can be expanded into a sum of partial fractions and each singular term can then be regularized by using the principal value or Sokhotsky distributions to obtain a distributional solution. The details are left to the reader as an Exercise. This method is used to construct Fourier transforms of Green's functions of differential operators.

26.5.2. Use of generalized principal value distributions. It is easy to see that

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

Indeed, the Taylor polynomial of order $n-2$ about $x=0$ is zero for the function $x^n\varphi(x)$ for any test function φ. Therefore

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

Second, if $\varphi\in\mathcal{D}(a,b)$ and (a,b) does not contain $x=0$, then any Taylor polynomial of φ about $x=0$ is zero so that

$$\Big(\mathcal{P}\frac{1}{x^n},\varphi\Big) = \int_a^b\frac{\varphi(x)}{x^n}\,dx\,,\quad \varphi\in\mathcal{D}(a,b)$$

where $0<a<b$ or $a<b<0$. This means that the generalized principal value distribution is equal to the corresponding reciprocal power function on any open interval that does not contain $x=0$:

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

Let use this idea to find a solution to the equation

$$a(x)f(x) = 1$$

where a is analytic. Let $\{x_n\}$ be the set of zeros of an analytic function $a(x)$ and m_n be a multiplicity of x_n. It is assumed that $x_{n-1}<x_n$ (the

set is ordered). Then a solution f can be obtained by the localization method. In a neighborhood of x_n, it is given by

$$f(x) = f_n(x) = \frac{(x-x_n)^{m_n}}{a(x)}\,\mathcal{P}\frac{1}{(x-x_n)^{m_n}}\,, \quad x \in U(x_n)$$

and for any $x' \neq x_n$

$$f(x) = \frac{1}{a(x)}\,, \quad x \in U(x')\,.$$

Let δ_n be the midpoint of the interval (x_n, x_{n+1}). Note that $\delta_n \geq \delta > 0$ for some δ because the sequence $\{x_n\}$ has no limits points. Define a distribution $\mathcal{P}\frac{1}{a(x)}$ by the rule

$$\textbf{(26.1)} \qquad \Big(\mathcal{P}\frac{1}{a(x)}, \varphi\Big) = \sum_n p.v. \int_{\delta_{n-1}}^{\delta_n} \frac{\varphi(x) - p_{m_n-2}(x)}{a(x)}$$

where $p_k(x)$ is the Taylor polynomial of φ about $x = x_n$ of order k. Then

$$\mathcal{P}\frac{1}{a(x)} = f_n(x)\,, \quad x \in U(x_n)\,, \quad \mathcal{P}\frac{1}{a(x)} = \frac{1}{a(x)}\,, \quad x \in U(x')\,.$$

By the localization theorem, any distribution is uniquely defined by its values in a neighborhood of any point. Therefore the linear functional defined by the rule (**26.1**) is a distribution for any analytic function a. It will be called a *principal value regularization* of the reciprocal of an analytic function.

If $g \in C^\infty$ and a is analytic, then a particular solution to $af = g$ is given by the generalized principal value distribution multiplied by g:

$$f(x) = g(x)\,\mathcal{P}\frac{1}{a(x)}\,.$$

26.5.3. Example. Let us find a general solution to the distributional equation

$$\sin^2(x) f(x) = \cos(x)$$

A general solution to the associated homogeneous equation was found at the end of Sec.**26.4**. A particular solution is obtained by the method from the previous section

$$f(x) = \cos(x)\mathcal{P}\frac{1}{\sin^2(x)}$$

The value of this distribution on any test function φ is computed by the rule

$$(f,\varphi) = \sum_n p.v. \int_{\pi n - \frac{\pi}{2}}^{\pi n + \frac{\pi}{2}} \frac{\varphi(x) - \varphi(\pi n)}{\sin^2(x)} \cos(x)\, dx$$

The series converges owing to the boundedness of support of φ (the summation is restricted to all n such that $\pi|n| < R$ if $\operatorname{supp}\varphi \subset (-R,R)$).

26.6. Particular solution for distributional inhomogeneity. There are two basic methods for solving an algebraic distributional equation if the inhomogeneity is not a smooth function.

26.6.1. Particular solution via the product of distributions. Consider the equation $a(x)f(x) = g(x)$ where a is analytic and $g \in \mathcal{D}'$. If the set of zeros of analytic function a does not intersect the singular support of the inhomogeneity g, then a particular solution can be obtained using the product of distributions defined by the localization method:

$$f(x) = \mathcal{P}\,\frac{1}{a(x)}\, g(x)\,.$$

Owing to the above identity, $a(x)\mathcal{P}\,\frac{1}{a(x)} = 1$, the distribution f satisfies the said equation. Here the generalized principal value can also be replaced by any regularization $\operatorname{Reg}\frac{1}{a(x)}$ that satisfies the equation $a(x)f(x) = 1$

For example, let us find a general solution to the equation

$$(x - x_0)f(x) = \frac{1}{x + i0}$$

where $x_0 \neq 0$. So, the product

$$f(x) = \mathcal{P}\frac{1}{x - x_0}\frac{1}{x - i0}$$

exists in $\mathcal{D}'$ and solves the equation. In the localization method for computing the product, one can take two open sets, $|x| < 2\delta$ and $|x| > \delta$, where $0 < 2\delta < x_0$, as a cover of $\mathbb{R}$. Then for any test function $\varphi \in \mathcal{D}$,

$$\Big(\mathcal{P}\frac{1}{x - x_0}\frac{1}{x - i0}, \varphi\Big) = \lim_{a\to 0^+} \int\limits_{|x|<\delta} \frac{\varphi(x)dx}{(x - x_0)(x + ia)} + p.v. \int\limits_{|x|>\delta} \frac{\varphi(x)\, dx}{x(x - x_0)}\,.$$

Using the partial fraction decomposition in the integrand

$$\frac{1}{x(x - x_0)} = \frac{1}{x_0}\Big(\frac{1}{x - x_0} - \frac{1}{x}\Big)$$

the product of distributions can be reduced to

$$\mathcal{P}\frac{1}{x-x_0}\frac{1}{x-i0} = \frac{1}{x_0}\left(\frac{1}{x+i0} - \mathcal{P}\frac{1}{x-x_0}\right).$$

26.6.2. Overlapping singularities. If the set of zeros of analytic function a and the singular support of the inhomogeneity g have common points, then a particular solution near these points cannot be obtained by the product of distributions based on the localization method. As mentioned earlier, there exists a generalization of the product of distributions based on the Fourier method. It can be used to obtain a solution if it exists. However, there is a simpler method that works in special cases. It is illustrated by examples.

26.6.3. Generalized principal value as inhomogeneity. Consider the equation

$$a(x)f(x) = g(x)\mathrm{Reg}\frac{1}{b(x)},$$

where a and b are analytic, g is from class C^∞, and the distributional regularization of $1/b(x)$ satisfies the equation $b(x)\mathrm{Reg}\frac{1}{b(x)} = 1$. Then any solution f to this equation is also a solution to the equation

$$b(x)a(x)f(x) = g(x).$$

Note that the converse is *not* true. If f is a solution to the latter equation, then $f + h$ is also a solution if $bh = 0$. The set of solutions to the former equation does not have this freedom. Since the product of two analytic function is an analytic function, a general solution is given by

$$f(x) = \mathcal{P}\frac{1}{b(x)a(x)} + \sum_n \sum_{k=1}^{m_n} C_{nk}\delta^{(k)}(x - x_n)$$

where m_n is the multiplicity of a root x_n of the product $a(x)b(x)$. As noted, this distribution may not satisfy the original equation for any choice of constants C_{nk}. So, the solution is to be substituted in the original equation to find all C_{nk} for which the equation is satisfied. It is clear, that C_{nk} associated with the roots of $a(x)$ remain arbitrary. But the coefficients associated with roots of b that are different from the roots of a cannot be arbitrary. Furthermore, if x_n is a common root of a and b then all coefficients C_{nk} for which k exceeds the multiplicity of x_n in $a(x) = 0$ cannot be arbitrary either.

26.6.4. Example. Let us find a general solution to the equation

$$xf(x) = \frac{1}{x + i0}\,.$$

By multiplying the equation by x, one concludes that any solution is also a solution to the equation

$$x^2 f(x) = 1\,.$$

A general solution to this equation reads

$$f(x) = \mathcal{P}\frac{1}{x^2} + c_1\delta(x) + c_2\delta'(x)\,.$$

Let us substitute this distribution into the original equation

$$xf(x) = \mathcal{P}\frac{1}{x} + c_2\delta'(x) = \mathcal{P}\frac{1}{x} - c_2\delta(x) = \frac{1}{x + i0}$$

It follows from the Sokhotsky equation that the last equality is true only if $c_2 = i\pi$. Thus, a general solution can be written in the form

$$f(x) = \mathcal{P}\frac{1}{x^2} + i\pi\delta'(x) + c_1\delta(x)\,,$$

where c_1 is a constant.

26.6.5. Inhomogeneity with a point support. Consider the equation

$$a(x)f(x) = \delta^{(k)}(x)\,,$$

where a is analytic. Let $b \in C^\infty$. Put $b^{(n)}(0) = b_n$. Then using the technique from Sec.**18.1**, it is not difficult to obtain that

$$(\mathbf{26.2})\qquad b(x)\delta^{(k)}(x) = \sum_{n=0}^{k} B_n\delta^{(n)}(x)\,,\quad B_n = (-1)^{k-n}\binom{k}{n} b_{k-n}\,.$$

Suppose that $a(0) \neq 0$, then a particular solution is given by

$$f(x) = \frac{1}{a(x)}\,\delta^{(k)}(x) = \sum_{n=0}^{k} B_n\delta^{(n)}(x)\,,$$

where B_n are given by (**26.2**) for $b(x) = \frac{1}{a(x)}$.

Suppose that $a(x) = x^m b(x)$, where b is analytic and $b(0) \neq 0$. Then using the identity

$$x^{k+1}\delta^{(k)}(x) = 0\,,$$

any solution to the said equation is also a solution to

$$x^{m+k+1}b(x)f(x) = 0\,.$$

Since $b(0) \neq 0$, a general solution to this equation is the sum

$$f(x) = \sum_{k=n+1}^{m+n+1} c_k \delta^{(k)}(x) + h(x)\,, \quad a(x)h(x) = 0\,,$$

where h is a general solution to the associated homogeneous equation. Therefore a particular solution is obtained by choosing coefficients c_k so that the original equation is satisfied:

$$\sum_{n=k+1}^{m+k+1} c_n a(x)\delta^{(n)}(x) = \delta^{(k)}(x)\,.$$

The left-hand side can be transformed in to a linear combination of the delta function and its derivatives using (**26**.**2**) for $b(x) = a(x)$ so that the coefficients c_k are found by matching the coefficients at the corresponding derivatives of the delta functions in the left and right hand sides of the equation.

26.6.6. Example. Let us find a general solution to the equation

$$xf(x) = \delta'(x)$$

Any solution to this equation is also a solution to the equation

$$x^3 f(x) = x^2\delta'(x) = 0$$

Then all solutions to the original equation are linear combinations

$$f(x) = c\delta(x) + c_1\delta'(x) + c_2\delta''(x)$$

The first term is a general solution to the homogeneous equation $xf(x) = 0$. So a particular solution must be a linear combination of δ' and δ'' The substitution of this distribution into the original equation yields

$$\delta'(x) = xf(x) = c_1 x\delta'(x) + c_2 x\delta''(x) = -c_1\delta(x) + 2c_2 x\delta'(x)$$

Therefore $c_1 = 0$ and $c_2 = \frac{1}{2}$ so that a general solution reads

$$f(x) = c\delta(x) + \frac{1}{2}\delta''(x)\,.$$

26.6.7. Superposition principle. If the distributions f_1 and f_2 satisfy the equations $af_1 = g_1$ and $af_2 = g_2$, then the sum $f = f_1 + f_2$ is a solution to $af = g_1 + g_2$ by linearity of the equation. This is called the *superposition principle* for linear equations. For example a general solution to the equation

$$xf(x) = 2\delta'(x) - \frac{3}{x + i0}$$

is obtained by the superposition principle using the particular solutions from the two above examples:

$$f(x) = c\delta(x) + \delta''(x) - 3i\pi\delta'(x) - \mathcal{P}\frac{3}{x^2}\,.$$

26.7. Higher dimensional generalizations. Let a smooth function $a(x)$ have isolated zeros $x_n \in \mathbb{R}^N$ of finite orders. Then any distribution f that satisfies the equation $a(x)f(x) = 0$ is the sum

$$f(x) = \sum_n f_n(x)\,,$$

where $f_n(x)$ is a distribution with support being the single point x_n. This follows from the localization theorem. By the structure theorem for distribution with point supports, each f_n is a linear combination of $D^\beta\delta(x - x_n)$. By substituting such combinations into the equation, the most general form of f_n can be found. In practice, finding independent parameters in linear combinations of derivatives of multi-variable delta-functions is often a tedious combinatorial problem. It is solved best via the Fourier transform of distributions and, for this reason, will be discussed later.

If $a(x) = 0$ defines a smooth M surface in $\mathbb{R}^N$, then a solution can be constructed from delta functions on the surface and possibly its derivatives. However, a general solution is often difficult to obtain. The Fourier method is helpful for this purpose if the surface is defined by zeros of a multi-variable polynomial.

26.8. Exercises.

1. Find a general distributional solution $f \in \mathcal{D}'(\mathbb{R})$ to the equations

(i) $x^2 f(x) = 1\,,$

(ii) $(x-a)^2 f(x) = x\,,$

(iii) $x^2(x-1)f(x) = x^2 + 1\,,$

(iv) $(x-a)f(x) = \delta'(x)\,,\ a \neq 0$

2. Find a general distributional solution to

$$(x-a)^n(x-b)^m f(x) = h(x)$$

where n and m are positive integers, and $h(x) > 0$ is a C^∞ function.

3. Find a general distributional solution to the equation

$$P(x)f(x) = Q(x)\,,$$

where P and Q are polynomials.

4. Find a general distributional solution to the equation

$$xf(x) = 2\mathcal{P}\frac{1}{x}\,.$$

5. Find a general distributional solution to the equation

$$\begin{aligned} &\text{(i)} \quad \sin(x)f(x) = x\,, \\ &\text{(ii)} \quad \sin(x)f(x) = \delta(x)\,, \\ &\text{(iii)} \quad \sin(x)f(x) = \mathcal{P}\frac{1}{x}\,. \end{aligned}$$

Give an explicit rule for computing (f, φ) where $\varphi \in \mathcal{D}$.

8. Suppose that $a(x)$ is a C^∞ function that has only one zero $a(0) = 0$ and the root $x = 0$ has infinite multiplicity, that is, $a^{(n)}(0) = 0$ for all $n \geq 0$ (so a is not analytic at $x = 0$). For example, $a(x) = \exp(-\frac{1}{|x|})$ if $x \neq 0$ and $a(0) = 0$. Show that the distribution

$$f_m(x) = \sum_{n=0}^{m} c_n \delta^{(n)}(x) \;\in\; \mathcal{D}'(\mathbb{R})\,,$$

is a solution to the equation $a(x)f(x) = 0$ for any m and any choice of constants c_n. However the formal series

$$f(x) = \lim_{m\to\infty} f_m(x) = \sum_{n=0}^{\infty} c_n \delta^{(n)}(x) \;\notin\; \mathcal{D}'(\mathbb{R})$$

is not a distribution for a general choice of constants c_n, that is, the distributional sequence f_m does not converge in $\mathcal{D}'$.

27. Integration of distributions

27.1. Antiderivative of a distribution. Consider the following distributional equation

$$f'(x) = \delta(x)$$

It is satisfied by the step-function up to an additive constant:

$$f(x) = \theta(x) + c$$

One can say that $\theta(x) + c$ is an antiderivative of the delta function. In this regard, two questions can be posed:

(i) *does any distribution f have a distributional antiderivative F in the sense that*

$$F'(x) = f(x)\,?$$

(ii) *what is the most general antiderivative of a distribution?*

The following theorem answers these questions.

THEOREM **27.1**. (Antiderivative of a distribution)
For any distribution f there exists a distribution F such that $F'(x) = f(x)$, and F is unique up to an additive constant.

Let us first investigate uniqueness. Suppose that a distribution F exists. Then for any test function ψ,

$$F' = f \quad \Rightarrow \quad (F, \psi') = -(f, \psi)$$

Therefore the distribution F is defined on a subset of $\mathcal{D}$ that consists of derivatives of all test functions. So, F has to be extended to the whole $\mathcal{D}$ in order to be an element of $\mathcal{D}'$. The following relation between a set of derivatives of all test functions and $\mathcal{D}$ can be established.

For any test function φ there exists a test function ψ such that

$$\varphi(x) = \psi'(x) + (1, \varphi)\eta(x)$$

where η is a particular test function with the property

$$(1, \eta) = \int_{-\infty}^{\infty} \eta(x)\, dx = 1\,.$$

For example, one can take $\eta(x) = \omega_a(x)$ (a hat function). A proof of this assertion is done by verifying that

$$\psi(x) = \int_{-\infty}^{x} \Big(\varphi(y) - (1, \varphi)\eta(y)\Big)\, dy\,, \qquad (\mathbf{27.1})$$

is a test function. Clearly, ψ is from C^∞. Let the supports of φ and η be in $[-R, R]$ and $[-a, a]$, respectively. Then the support of ψ lies in the interval:

$$\operatorname{supp}\psi \subset [-b, b]\,, \qquad b = \max\{R, a\}\,.$$

Indeed, if $x < b$, then integrand is identically zero and so is $\psi(x)$. If $x > b$, then the first integral is equal to $(1, \varphi)$ (the integral of φ over its whole support), while the second integral is equal to $(1, \varphi)(1, \eta) = (1, \varphi)$ and, hence, cancels the first one so that $\psi(x) = 0$.

Therefore for any $\varphi \in \mathcal{D}$

$$(F, \varphi) = (F, \psi') + (1, \varphi)(F, \eta) = -(f, \psi) + (F, \eta)(1, \varphi)\,.$$

The constant $C = (F, \eta)$ is the value of the distribution F at a particular test function η (independent of φ). So, the above relation can be restated as

$$(F, \varphi) = -(f, \psi) + (C, \varphi)\,. \tag{\textbf{27.2}}$$

If F exists, then its value at any test function is uniquely defined up to an additive constant distribution.

It remains to show that F is a distribution, that is, the functional (**27.2**) is linear and continuous. To do so, let us use the standard trick. Define a transformation of $\mathcal{D}$ into itself by

$$T: \quad \varphi \to T(\varphi) = -\psi\,,$$

where ψ is defined by (**27.1**). If T is linear and continuous on $\mathcal{D}$, then its adjoint maps $\mathcal{D}'$ into itself $\mathcal{D}'$ by the rule $(T^*(f), \varphi) = (f, T(\varphi)) = -(f, \psi)$. Therefore $F = T^*(f) + C$ by (**27.2**). A constant function defines a distribution. So, if $T^*(f)$ is a distribution, then F is also a distribution. Linearity of T is obvious. Let $\varphi_n \to 0$ in $\mathcal{D}$. One has to show that $T(\varphi_n)$ is also a null sequence in $\mathcal{D}$. One has

$$T(\varphi_n) = -\int_{-\infty}^{x} \varphi_n(y)\,dy + (1, \varphi_n)\int_{-\infty}^{x} \eta(y)\,dy\,.$$

The support of all φ_n lies in $[-R, R]$ (by definition of topology in $\mathcal{D}$) and, hence, the support of all $T(\varphi_n)$ is in $[-b, b]$ where b is defined above. Then

$$\begin{aligned} |T(\varphi_n)(x)| &\le 2R\sup|\varphi_n| + |(1, \varphi_n)| \cdot 2a\sup|\eta| \\ &\le 2R\Big(1 + 2a\sup|\eta|\Big)\sup|\varphi_n| \end{aligned}$$

because the supports of φ_n and η in $[-R, R]$ and $[-a, a]$, respectively. The estimate holds for all x and, hence,

$$\sup|T(\varphi_n)| \le 2R\Big(1 + 2a\sup|\eta|\Big)\sup|\varphi_n|\,.$$

It follows from this inequality, that if $\sup|\varphi_n| \to 0$, then $\sup|T(\varphi_n)| \to 0$. Since

$$D^\beta T(\varphi_n) = D^{\beta-1}\varphi_n - (1,\varphi_n)D^{\beta-1}\eta\,,\quad \beta > 0\,,$$

similar estimates hold for derivatives of $T(\varphi_n)$:

$$\sup|D^\beta T(\varphi_n)| \le \sup|D^{\beta-1}\varphi_n| + 2R\sup|\varphi_n|\sup|D^{\beta-1}\eta|\,.$$

This shows that the convergence $\sup|D^\beta T(\varphi_n)| \to 0$ follows from the convergence $\sup|D^\alpha\varphi_n| \to 0$ for any $\alpha \ge 0$. Thus, T is continuous on $\mathcal{D}$ and, hence, $T^*(f)$ is a distribution and so is $F = T^*(f) + C$.

Finally, let us verify that $F' = f$ in $\mathcal{D}'$. For any $\varphi \in \mathcal{D}$,

$$(F',\varphi) = -(F,\varphi') = -(T^*(f),\varphi') - (C,\varphi') = -(f, T(\varphi')) = (f,\varphi)$$

because $(C,\varphi') = 0$ for any constant C and

$$T(\varphi')(x) = -\int_{-\infty}^x \varphi'(y)dy - (1,\varphi')\int_{-\infty}^x \eta(y)\,dy = -\varphi(x)\,.$$

27.2. Indefinite integral of a distribution. Let us define a particular antiderivative of a distribution f denoted by $D^{-1}f$ as

$$D^{-1}f = T^*(f)\,.$$

By construction, the antiderivative D^{-1} is a linear and continuous transformation of $\mathcal{D}'$ into itself. If $f_n \to f$ in $\mathcal{D}'$, then $D^{-1}f_n \to D^{-1}f$ in $\mathcal{D}'$ by continuity of the adjoint transformation. The *indefinite integral of a distribution* is defined as the most general antiderivative:

$$\int f(x)\,dx = D^{-1}f(x) + C\,.$$

Clearly, the operation D^{-1} can be repeated any number of times. So, an antiderivative of a distribution f of order β, $D^{-\beta}f$, is a distribution.

27.3. Classical and distributional antiderivatives. Let $f(x)$ be continuous. Put

$$F(x) = \int_{x_0}^x f(y)\,dy$$

for some x_0. Then F is a classical antiderivative of f because $dF(x) = f(x)dx$ by the fundamental theorem of calculus. Let us calculate the distributional antiderivative $D^{-1}f$. By integration by parts one infers that

$$(D^{-1}f,\varphi) = (f, T(\varphi)) = \int_{-\infty}^{\infty} f(x)T(\varphi)(x)\,dx = \int_{-\infty}^{\infty} T(\varphi)dF$$
$$= \int_{-\infty}^{\infty} F(x)\,dT(\varphi) = -(F,\psi') = (F,\varphi) - (1,\varphi)(F,\eta)\,.$$

Thus, the classical and distributional antiderivatives are equal for continuous functions up to an additive constant

$$D^{-1}f(x) = \{D^{-1}f\} + C, \quad C = (\{D^{-1}f\},\eta)\,.$$

This implies that the classical and distributional indefinite integrals are equal for all continuous functions. Note that different choices of η in the definition of $D^{-1} = T^*$ simply leads to different constants C so that the classical and distributional indefinite integrals coincide whenever the former exists.

In particular,

$$D^{-1}x^n = \frac{x^{n+1}}{n+1} + C\,, \quad C = \frac{(x^{n+1},\eta)}{n+1}\,, \quad n = 0, 1, ...$$

It follows from this relation that a distributional antiderivative of order β is unique up to an additive polynomial of order $\beta - 1$. For example, the double indefinite integral of a distribution f reads

$$\int \left(\int f(x)\,dx\right) dx = \int \left(D^{-1}f(x) + C_1\right) dx$$
$$= D^{-2}f(x) + C_1 x + C_2$$

27.4. Examples. Let us find $D^{-1}\delta(x)$ using the definition of the action of D^{-1} on $\mathcal{D}'$. One has

$$(D^{-1}\delta,\varphi) = -(\delta,\psi) = -\psi(0) = -\int_{-\infty}^{0} \varphi(x)\,dx + (1,\varphi)\int_{-\infty}^{0} \eta(x)\,dx$$

The integrals are the values of the regular distribution $-\theta(-x)$, where θ is the step function, at a test function. Therefore

$$D^{-1}\delta(x) = -\theta(-x) + C\,, \qquad C = \int_{-\infty}^{0} \eta(x)\,dx\,.$$

The constant C depends on the choice of η. If one takes $\eta(-x) = \eta(x)$, then the condition $(1,\eta) = 1$ implies that $C = \frac{1}{2}$ and

$$D^{-1}\delta(x) = -\theta(-x) + \frac{1}{2} = \frac{1}{2}\varepsilon(x)$$

If η is such that

$$\int_{-\infty}^{0} \eta(x)\, dx = 1$$

then

$$D^{-1}\delta(x) = -\theta(-x) + 1 = \theta(x)$$

So,

$$\int \delta(x)\, dx = \theta(x) + C$$

Finding antiderivatives of distributions does not always require the use of the rule (**27.2**) and (**27.1**), just like in the case of smooth functions derivatives of basic distributions allows us to find antiderivatives of other distributions. For example,

$$D\mathcal{P}\frac{1}{x} = -\mathcal{P}\frac{1}{x^2} \quad \Rightarrow \quad D^{-1}\mathcal{P}\frac{1}{x^2} = -\mathcal{P}\frac{1}{x} + C$$
$$D\ln(|x|) = \mathcal{P}\frac{1}{x} \quad \Rightarrow \quad D^{-1}\mathcal{P}\frac{1}{x} = \ln(|x|) + C$$

27.4.1. Integration by parts. The Leibniz rule is also helpful for finding antiderivatives (just like the integration by parts for classical functions). If $a(x)$ is a smooth function and f is a distribution, then integrating the Leibniz identity

$$a(x)f'(x) = \Big(a(x)f(x)\Big)' - a'(x)\, f(x)$$

one infers the *integration by parts* for distributions

$$\int a(x)f'(x)\, dx = a(x)f(x) - \int a'(x)f(x)\, dx\,.$$

For example,

$$\int \Big(x\delta'(x)\Big)\, dx = x\delta(x) - \int \delta(x)\, dx = -\theta(x) + C$$

because $x\delta(x) = 0$.

27.5. Distributional initial value problem. The classical initial value problem is to find a function F whose derivative and value at a particular point are known:

$$F'(x) = f(x)\,, \quad F(x_0) = F_0$$

It has a unique solution for a continuous f:

$$F(x) = F_0 + \int_{x_0}^{x} f(y)\, dy$$

If f is locally integrable, then F is absolutely continuous and the above equality holds almost everywhere. So, the problem has a unique solution in the space of regular distributions.

If f is a general distribution, then it does not have point values. So, the initial condition fixes the value of the functional f on a particular test function η:

$$F' = f\,, \qquad (F,\eta) = F_0\,.$$

The problem has a unique solution

$$F(x) = F_0 + D^{-1}f(x)$$

where the test function η used in the definition of $D^{-1}f$ is the test function at which the "initial" data are set. In this case

$$(D^{-1}f,\eta) = 0$$

because the test function ψ vanishes if $\varphi = \eta$ in (**27.1**). This is similar to the above classical solution where the antiderivative is defined so that $D^{-1}f(x_0) = 0$.

27.6. Higher dimensional generalizations. Let $x \in \mathbb{R}$ and $y \in \mathbb{R}^N$. The objective is to find a general distributional solution to the equation

$$D_x F(x,y) = f(x,y)\,, \tag{27.3}$$

where $f \in \mathcal{D}'(\mathbb{R}^{N+1})$. If f is a regular distribution defined by a continuous function, then

$$F(x,y) = \int_{x_0}^{x} f(s,y)\,ds + g(y) = D_x^{-1}f(x,y) + g(y)$$

for some function g independent of x. Let us show that this solution can be extended to distributions. In other words, *a general solution is an antiderivative of the distribution f with respect to x up to an additive distribution that is independent of x*. Here the concept of an antiderivative with respect to a particular variable needs to be defined.

Consider a transformation $T : \mathcal{D}(\mathbb{R}^{N+1}) \to \mathcal{D}(\mathbb{R}^{N+1})$ defined by the rule

$$T(\varphi)(x,y) = -\int_{-\infty}^{x} \varphi(s,y)\,ds + \phi(y)\int_{-\infty}^{x} \eta(s)\,ds$$
$$\phi(y) = \int_{-\infty}^{\infty} \varphi(x,y)\,dx = \Big(1(x), \varphi(x,y)\Big) \in \mathcal{D}(\mathbb{R}^N)\,,$$

where η is the same as in (**27.1**). If the variable y is omitted, then this transformation is identical to (**27.1**) up the overall sign. So, a proof that $T(\varphi)$ is a test function is identical to that given for (**27.1**) if one notes that ϕ is a test function obtained by the reduction of variables. Since the latter is a linear continuous transformation (see Sec.**17.3**), the transformation T is also linear and continuous. Then its adjoint defines a transformation of distributions $T^* : \mathcal{D}'(\mathbb{R}^{N+1}) \to \mathcal{D}'(\mathbb{R}^{N+1})$ by the usual rule $(T^*(f), \varphi) = (f, T(\varphi))$. Let us calculate the distributional derivative $D_x T^*(f)$. One infers that

$$(D_x T^*(f), \varphi) = -(T^*(f), D_x\varphi) = -(f, T^*(D_x\varphi)) = (f, \varphi)\,.$$

The latter equality follows from

$$\Big(1(x), D_x\varphi(x,y)\Big) = -\Big(D_x 1(x), \varphi(x,y)\Big) = 0$$

justified by the analysis in Sec.**19.4** so that

$$T(D_x\varphi)(x,y) = -\int_{-\infty}^{x} D_s\varphi(s,y)\,ds = -\varphi(x,y)$$

This shows that the distribution $T^*(f)$ is an antiderivative of f with respect to x,

$$D_x T^*(f)(x,y) = f(x,y)\,.$$

So put

$$D_x^{-1} f(x,y) = T^*(f)(x,y)\,.$$

This distribution is a particular solution to (**27.3**).

It follows from the definition of $T(\varphi)$ that

$$\varphi(x,y) = -D_x T(\varphi)(x,y) + \phi(y)\eta(x)\,.$$

Let F be any solution to (**27.3**). Define a distribution $g(y) \in \mathcal{D}'(\mathbb{R}^N)$ by the rule

$$(g,\phi) = \Big(F(x,y), \eta(x)\phi(y)\Big)\,, \quad \eta \in \mathcal{D}(\mathbb{R})\,, \quad \phi \in \mathcal{D}(\mathbb{R}^N)\,.$$

The linearity and continuity of the functional g follows from the linearity and continuity of the functional F. Then for any $\varphi \in \mathcal{D}(\mathbb{R}^{N+1})$,

$$\begin{aligned}(F,\varphi) &= -(F, D_x\psi) + (F, \phi\eta) = (f,\psi) + (g,\phi)\\ &= (T^*(f),\varphi) + \Big(1(x)g(y), \varphi(x,y)\Big)\,,\end{aligned}$$

where $1(x)g(y)$ stands for a distribution of $N+1$ variables that is independent of variable x (see Sec.**17.3**). Thus, a general solution to (**27.3**) reads

$$F(x,y) = \int f(x,y)\,dx = D_x^{-1} f(x,y) + g(y)\,,$$

where g is any distribution that is independent of x. So, an antiderivative and indefinite integral with respect to any variable can be defined for distributions of several variables. The properties of such distributional antiderivatives and indefinite integrals are identical to their classical analogs.

28. Differential equations for distributions

Let $a_k(x)$, $k = 0, 1, ..., n$, be smooth functions of a real variable x. Consider a differential equation

$$L(D)f = \sum_{k=0}^{n} a_k(x) D^k f(x) = g(x)\,,$$

where g is a given distribution. The problem is to find all distributions $f \in \mathcal{D}'(\mathbb{R})$ that satisfy the equation or show that none exists. If $x \in \mathbb{R}^N$, then the distribution $f \in \mathcal{D}'(\mathbb{R}^N)$ is a solution to a partial differential equation.

Let $\varphi_{1,2}$ be test functions. The *formal adjoint* L^* of L is defined by the relation

$$(L\varphi_1, \varphi_2) = \int L\varphi_1 \varphi_2\, dx = \int \varphi_1 L^* \varphi_2\, dx = (\varphi_1, L^*\varphi_2)\,,$$

where the second equality is obtained by integration by parts. Therefore

$$L^*(D)\varphi = \sum_{k=0}^{n} (-1)^k D^k (a_k \varphi)\,.$$

If $f \in \mathcal{D}'$ is a solution to a differential equation, then for any test function φ,

$$(L(D)f, \varphi) = (f, L^*(D)\varphi) = (g, \varphi)\,. \tag{28.1}$$

The differential operator L is linear. Therefore any two solutions differ by a solution to the associated homogeneous equation. So, a general solution has the form

$$f(x) = f_p(x) + h(x)\,, \quad L(D)h = 0\,,$$

where f_p is a particular solution. So, the first problem is to find a general solution to the homogeneous equation. A general theory for the latter problem is already difficult enough for classical functions. For example, second-order ordinary differential equations require the Frobenius theory when $a_2(x)$ has zeros to construct classical solutions near the zeros. Classical solutions are in general singular functions at zeros of a_2 and may not have distributional extensions to the zeros. The stated problem for ordinary differential equations is generally more of a theoretical significance than of importance for applications. So, the discussion will be limited to first-order equations and to some simple examples of equations of other types to illustrate peculiarities of distributional differential equations that are not present in the classical theory of ordinary differential equations. A general theory for equations with constant coefficients will be developed later and extended to partial differential equations.

It should be noted that if the coefficients are not smooth everywhere, then the problem of multiplication of distributions arises in this equation. It must be resolved somehow so that the equation makes sense in the space of distributions, e.g., by restricting the class of distributions in which a solution is sought. For example, the problem can still make sense for distributions in an interval, $f \in \mathcal{D}'(\Omega)$, $\Omega \subset \mathbb{R}$, if the coefficients are smooth on Ω, $a_k \in C^\infty(\Omega)$ for all k.

28.1. Simple examples. Let us find a general distributional solution to the equation

$$x^2 f'(x) = 1\,.$$

Put $g(x) = f'(x)$. Then it was shown earlier that

$$g(x) = \mathcal{P}\frac{1}{x^2} + C_1\delta(x) + C_2\delta'(x)$$

where C_1 and C_2 are constants. Therefore

$$f(x) = \int g(x)\,dx = -\mathcal{P}\frac{1}{x} + C_1\theta(x) + C_2\delta(x) + C_3$$

where C_3 is an integration constant. It is noteworthy that in contract to the classical case, a general distributional solution has three arbitrary constants for a first-order equation. But just like in the classical case, arbitrary constants are parameters of a general solution to the associated homogeneous equation.

28.1.1. A second-order equation. Let us find a general solution to

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

By multiplying the equation by a smooth function x and using the identity, $x\delta(x) = 0$, it is concluded that any solution of the equation is among solutions of the homogeneous equation:

$$x^2 f''(x) = 0 \quad \Rightarrow \quad f''(x) = C_1\delta(x) + C_2\delta'(x)$$

The substitution of f'' into the original equation yields

$$C_2 x\delta'(x) = \delta(x) \quad \Rightarrow \quad C_2 = -1\,.$$

Therefore

$$\begin{aligned} f(x) &= \iint \Big(-\delta'(x) + C_1\delta(x)\Big)\,dx\,dx = \int \Big(-\delta(x) + C_1\theta(x) + C_2\Big)\,dx \\ &= -\theta(x) + C_1 x\theta(x) + C_2 x + C_3\,, \end{aligned}$$

where C_1, C_2, and C_3 are arbitrary constants. Note that $-\theta(x)$ is a particular solution, whereas the other terms comprise a general solution to the associated homogeneous equation.

28.1.2. General solution to a homogeneous 2D wave equation. Let us find a general distributional solution to the 2D wave equation discussed in Sec.**21.6.3**. The change of variables $x_{pm} = x \pm ct$ is legitimate for distributions of two variables and the chain rule holds. Therefore, denoting the partial derivatives with respect to $x_\pm$ by $D_\pm$, the said equation is equivalent to

$$\begin{aligned}(D_t^2 - c^2 D_x^2)u &= [c^2(D_+ - D_-)^2 - c^2(D_+ + D_-)^2]u \\ &= -4c^2 D_+ D_- u = 0\,.\end{aligned}$$

The latter equation is of type (**27.3**) and can be integrated using distributional antiderivatives with respect to $x_\pm$:

$$\begin{aligned}D_- u = h(x_-) \quad \Rightarrow \quad u(x,t) &= D_-^{-1} h(x_-) + f(x_+) \\ &= g(x-ct) + f(x+ct)\,,\end{aligned}$$

where $h' = g$ and f and g are arbitrary distributions of one real variable.

For example, $u(x,t) = \delta(x-ct)$ describes a propagation of a delta-like pulse in the direction of increasing x with speed c. It can be viewed as an idealization of a propagating of a narrow pulse represented by a classical (smooth) solution obtained from a distributional one by, e.g., a convolution regularization $u_a(x,y) = \omega_a(x-ct)$ where $\omega_a(x_-)$ is a hat function.

28.2. Localization method. Example. Recall that any distribution is uniquely defined by its values in neighborhoods of all points. Then a differential equation can be solved near each point and the solution in the whole region is obtained by the localization theorem using (**25.1**). To illustrate the concept, let us find a general solution to the following equation

$$a(x)f'(x) = \delta^{(k)}(x)\,, \quad k \ge 0\,,$$

where $a(x)$ is a smooth function. If f_k is a particular solution to this equation, then by linearity of the equation, the distribution

$$f(x) = \sum_{k=0}^{M} c_k f_k(x)$$

is a particular solution to the equation $af' = g$ with inhomogeneity being a general distribution with point support

$$g(x) = \sum_{k=0}^{M} c_k \delta^{(k)}(x)\,.$$

The problem is solved using the result from Sec.**26.6.5**. If $a(x)$ does not vanish anywhere, then the reciprocal of a is from C^∞. Then

$$(\mathbf{28.2})\qquad f(x) = \int \frac{1}{a(x)}\delta^{(k)}(x)\,dx = C + B_0\theta(x) + \sum_{n=0}^{k-1} B_{n+1}\delta^{(n)}(x)\,,$$

where the coefficients B_n are given by (**26.2**) for $b(x) = 1/a(x)$ and C is an integration constant. Note that a constant C is a general solution to the associated homogeneous problem.

If $a(x)$ has zeros, then due to the division problem for distributions this method is not applicable. It is further assumed that each zero has a finite multiplicity and the sequence of zeros of a has no limit points. Under the latter assumption, one can construct a countable cover of $\mathbb{R}$ by open intervals such that each interval contains just one root of a. Let $\{x_n\}$ be a sequence where $x_0 = 0$ and x_n, $n \neq 0$, are roots of a (x_0 is not necessarily a root). The terms of the sequence are ordered so that $x_n < x_{n+1}$. Let I_n be an open interval containing x_n such that $\mathbb{R} = \cup_n I_n$. Put $f(x) = f_n(x)$ if $x \in I_n$ where f_n is a general solution near x_n. These local solutions must match in the intersections $I_n \cap I_{n+1}$ so that

$$(\mathbf{28.3})\qquad f_n(x) = f_{n+1}(x)\,,\quad x \in I_n \cap I_{n+1} \subset (x_n, x_{n+1})\,.$$

Once all f_n have been found, the distribution f is obtained by (**25.1**). For the sake of avoiding unessential technical complications that might be fogging the concept, let us make a simplifying assumption that all roots of a are simple. The case of roots of higher multiplicity can be treated in the same way by the method developed in Sec.**26.4** but technicalities are more involved. Specific examples are better to work out individually following the concept.

If x_n is a simple root, then $a(x) = (x - x_n)b(x)$ where $b(x_n) \neq 0$ and b is from C^∞ near x_n. Therefore or $n \neq 0$,

$$a(x)f_n'(x) = 0 \quad\Rightarrow\quad f_n'(x) = C\delta(x - x_n)\,,$$

for some constant C. It is now easy to integrate the equation. A general solution is convenient to write in the form

$$f_n(x) = C_n^- + (C_n^+ - C_n^-)\theta(x - x_n)\,.$$

Note that $f_n'(x) = 0$ for $x \neq x_n$ so that it is defined by constant functions C_- and C_+ for $x < x_n$ and $x > x_n$ in I_n, respectively. The solutions f_n and f_{n+1} must coincide in the interval (x_n, x_{n+1}). So the matching condition (**28.3**) for two neighboring non-zero roots is

$$C_n^+ = C_{n+1}^- .$$

This shows that a general solution to the homogeneous equation is given by a general piecewise constant function with jump discontinuities at the roots of a.

Our next objective is to find a general solution in I_0,

$$a(x)f_0'(x) = \delta^{(k)}(x)\,, \quad x \in I_0\,,$$

and match it with solutions in $I_{\pm 1}$. Suppose $a_0 = a(0) \neq 0$. In this case, there is no division problem in I_0 and a general solution coincides with (**28.2**). The matching conditions in $(x_{-1}, 0)$ and $(0, x_1)$ require that

$$C = C_{-1}^+\,, \quad C_1^- = B_0 + C_{-1}^+ .$$

So, a general solution is

$$f(x) = h(x) + \sum_{n=0}^{k-1} B_{n+1}\delta^{(n)}(x)\,, \quad \operatorname*{disc}_{x=0} h = B_0\,,$$

where $k > 0$, h is a general piecewise constant functions with jump discontinuities at roots of a and $x = 0$ (the discontinuity at $x = 0$ must have a fixed value B_0), and $f = h$ if $k = 0$.

Suppose a has a (simple) root $x = 0$. In this case, the equation is equivalent to

$$xf_0'(x) = \sum_{n=0} B_n\delta^{(n)}(x)\,, \quad x \in I_0\,,$$

where B_n are given by (**26.2**) for $b(x) = x/a(x)$ which is from C^∞ near $x = 0$ and $b(0) \neq 0$. If g_n is a solution to $xg_n = \delta^{(n)}$, then by linearity

$$f_0'(x) = \sum_{n=0}^{k} B_n g_n(x)\,, \quad xg_n = \delta^{(n)} .$$

The distribution g_n must have a point support because it vanishes for $x \neq 0$. Therefore it is a linear combination of the delta function and its derivatives. The coefficients in the linear combination are easy to find:

$$x g_n(x) = \sum_{m=0}^{M} c_m x \delta^{(m)}(x) = -\sum_{m=0}^{M} c_m m \delta^{(m-1)}(x) = \delta^{(n)}(x)\,,$$
$$g_n(x) = c_0\delta(x) - \frac{1}{n+1}\,\delta^{(n+1)}(x)\,,$$

where c_0 is an arbitrary constant. Note that the first term is nothing but a general solution to the associated homogeneous equation. Substituting g_n and taking an antiderivative of f_0', one infers that

$$f_0(x) = C_{-1}^{+} + (C_1^{-} - C_{-1}^{+})\theta(x) - \sum_{n=0}^{k} \frac{B_n}{n+1}\,\delta^{(n)}(x)\,,$$

where c_0 and the integration constant are chosen so that $f_0(x) = C_{-1}^{+}$ in $(x_{-1}, 0)$ and $f_0(x) = C_1^{-}$ in $(0, x_1)$. Thus, a general solution is given by

$$f(x) = h(x) - \sum_{n=0}^{k} \frac{B_n}{n+1}\,\delta^{(n)}(x)\,,$$

where h is a general piecewise constant function with discontinuities at roots of a.

28.3. Linear first-order equations. The most general linear first-order equation reads

$$b(x)f'(x) + a(x)f(x) = h(x)\,, \quad h \in \mathcal{D}'\,, \quad a, b \in C^{\infty}\,.$$

If $b(x)$ has no zeros, then the problem is fully analogous to the corresponding classical problem as shown below. If b has zeros, then the problem can have no solution in $\mathcal{D}'$ at all.

For example, consider the following homogeneous equation

$$x^2 f'(x) + \nu f(x) = 0\,, \quad f \in \mathcal{D}'\,,$$

where ν is a constant. For $x \neq 0$, a distributional solution, if it exists, must coincide with the classical solution:

$$f(x) = C_+ e^{\frac{\nu}{x}}\,, \quad x > 0\,, \quad f(x) = C_- e^{\frac{\nu}{x}}\,, \quad x < 0\,,$$

where $C_\pm$ are arbitrary constants. Let $f(x) = \{f(x)\}$ for $x \neq 0$ where $\{f\}$ is the above classical solution. The function $\{f\}$ is singular and does not define any distribution on $\mathbb{R}$. It must be extended to $x = 0$. However, in Sec.**16.3.3** it was shown that such $\{f\}$ is too singular and has no distributional extension to $x = 0$. So the problem does not have any non-trivial solution in $\mathcal{D}'(\mathbb{R})$. A non-trivial solution exists in $\mathcal{D}'(\Omega)$

for any open interval Ω that does not contain the singular point $x = 0$ and, in this case, the distributional and classical solutions coincide.

28.3.1. Case $b \neq 0$. The equation can be divided by b to obtain an equivalent equation. Therefore without loss of generality, put $b(x) = 1$ and consider a substitution

$$f(x) = \exp\Big(-\int_0^x a(y)\,dy\Big)g(x)\,,$$

where g is a unknown distribution. Since the exponential is smooth function that vanishes nowhere, the substitution makes sense in $\mathcal{D}'$ and is invertible. For any $f \in \mathcal{D}'$ there exists a unique $g \in \mathcal{D}'$ and vice versa. By the Leibniz rule, g is shown to satisfy the equation

$$\exp\Big(-\int_0^x a(y)\,dy\Big)g'(x) = h(x)\,.$$

Therefore, a general solution reads

$$f(x) = e^{-\int_0^x a(y)\,dy}\int e^{\int_0^x a(y)\,dy}h(x)\,dx\,. \qquad (\mathbf{28}.\mathbf{4})$$

The indefinite (distributional) integral contains an additive integration constant C. The term $Ce^{-\int_0^x a(y)\,dy}$ is a general solution to the associated homogeneous equation. In particular, if $h(x) = \delta(x)$, one gets

$$f(x) = \Big(\theta(x) + C\Big)e^{-\int_0^x a(y)\,dy}\,,$$

where C is a constant.

28.4. General homogeneous equation. Let us analyze a general solution to the associated homogeneous equation

$$b(x)f'(x) + a(x)f(x) = 0\,, \quad f \in \mathcal{D}'\,,$$

where b has zeros. As noted, the existence of a solution depends on the behavior of b and a near zeros of b. The approach is based on the localization theorem. It is assumed that b and a are smooth and zeros of b have finite multiplicity. The sequence of zeros has no limit point so that the existence of f can be studied in an open interval containing a single zero of b. If solutions near each zero exist, then a general solution in $\mathcal{D}'$ is obtained by (**25**.**1**) (or by fulfilling matching conditions) as illustrated in Sec.**28.2**.

Let $b(0) = 0$. Then near $x = 0$

$$\frac{a(x)}{b(x)} = q(x) + \frac{\nu}{x} + p_n(x)\,, \quad p_n(x) = \frac{\nu_2}{x^2} + \cdots + \frac{\nu_n}{x^n}\,,$$

for some integer $n > 1$ and an analytic function $q(x)$. Then a distributional solution must coincide with the classical solution near $x = 0$

$$f(x) = \{f(x)\} = C_\pm e^{-\int (a/b)\,dx} = \frac{C_\pm}{|x|^\nu} e^{-P_n(x)} e^{-Q(x)}\,, \quad x \neq 0\,,$$

where $C_\pm$ are constants for $x > 0$ and $x < 0$ as in the example above, and P_n and Q are antiderivatives of p_n and q, respectively. Due to the factor e^{-P_n}, the function $\{f\}$ is too singular at $x = 0$ to have a distributional regularization in a neighborhood of $x = 0$ as in the example considered in Sec.**16.3.3**. So, no distributional solution exists if $p_n \neq 0$.

If a/b is bounded near a particular point, then this point is called a *regular point* of the equation, otherwise it is called a *singular point*. Near every regular point a distributional solution exists and coincide with a classical solution. Suppose that

$$\lim_{x\to x_0} \frac{(x - x_0)a(x)}{b(x)} = \nu\,.$$

In this case, x_0 is called a *regular singular point* of the equation. The parameters ν and multiplicity m of the root x_0 of b are called *indices* of the regular singular point. Near every such point, a distributional solution also exists. Indeed, a classical solution

$$\{f(x)\} \sim \frac{e^{-Q(x)}}{|x - x_0|^\nu}\,,$$

is locally integrable if $\nu < 1$ and, hence, defines a distribution in a neighborhood of $x = 0$. If $\nu \geq 1$, $\{f\}$ is not locally integrable, but any reciprocal of a power function admits a distributional extension to the singular point $x = x_0$. For example, one can always take the principal value regularization defined in (**16**.**2**). Thus, if all singular points of the equation are regular singular points, then a distributional solution exists near every point. Between any two regular singular points, the distributional solution is given by the classical solution which is a smooth function. It remains to show that local distributional solutions have enough parameters to fulfill the latter matching conditions. Then by the localization theorem, the equation has a solution in $\mathcal{D}'(\mathbb{R})$.

28.4.1. The structure of a general solution. Let $\{x_s\}$ be a sequence of regular singular points with indices ν_s and m_s ordered so that $x_s < x_{s+1}$ It is assumed that the sequence has no limit point (or any bounded interval contains at most finitely many points of $\{x_s\}$. Let I_s be non-overlapping open intervals obtained from $\mathbb{R}$ by removing singular points

x_s. Then the distributional solution coincides with the classical one in any of these intervals

$$\{f_s(x)\} = C_s \exp\Big(-\int \frac{a(x)}{b(x)}\,dx\Big)\,, \quad x \in I_s$$

where C_s is an arbitrary constant. Between any two singular points x_s and x_{s+1} there are two equivalent representations of a/b:

$$\frac{a(x)}{b(x)} = \frac{\nu_s}{x - x_s} + q_s(x) = \frac{\nu_{s+1}}{x - x_{s+1}} + q_{s+1}(x)\,,$$

where q_s and q_{s+1} are smooth functions in (x_s, x_{s+1}). Therefore there are two representations of the classical solution

$$\begin{aligned}\{f_s(x)\} &= \frac{C_s}{|x - x_s|^{\nu_s}} \exp\Big(-\int_{x_s}^{x} q_s(y)\,dy\Big)\,,\\ &= \frac{C_s'}{|x - x_{s+1}|^{\nu_{s+1}}} \exp\Big(-\int_{x_{s+1}}^{x} q_{s+1}(y)\,dy\Big)\,,\end{aligned}$$

in the interval $x_s < x < x_{s+1}$. The constants C_s and C_s' are proportional to one another, and the proportionality coefficient can be defined by matching the solutions at any particular point in (x_s, x_{s+1}). For definitiveness, the first representation is used in any bounded interval I_s. If $\{x_s\}$ is bounded from above, then in the interval $(\max\{x_s\}, \infty)$ the first representation is used for $x_s = \max\{x_s\}$. If the sequence $\{x_s\}$ is bounded from below, then the classical solution in the interval $(-\infty, \min\{x_s\})$ is defined by the second representation $x_s = \min\{x_s\}$.

Each function $\{f_s\}$ is extended to $\mathbb{R}$ by zero outside of I_s so that

$$\{f(x)\} = \sum_s \{f_s(x)\}\,,$$

is a general classical solution wherever it exists. By construction

$$\lim_{x \to x_s^{\pm}} |x - x_s|^{\nu_s} \{f(x)\} = C_s^{\pm}\,,$$

where $C_s^+ = C_s$ are independent parameters of $\{f\}$, whereas $C_s^- = C_{s-1}'$ is proportional to C_{s-1}^+. If the sequence $\{x_s\}$ is bounded from below, then $C_s^- = C_s$ is also an independent parameters of $\{f\}$ in the interval $(-\infty, \min\{x_s\})$. This shows that if any of $\nu_s > 1$, $\{f(x)\}$ is not locally integrable and does not define a distribution. However it has a distributional regularization by means of the principal value regularization defined in Sec.**16.3.2**. Recall that a distributional regularization is not unique. It is defined up to an additive distribution supported in the set to which $\{f\}$ is extended. Therefore by the localization theorem any distributional solution, if it exists, must have the form

$$f(x) = \mathcal{P}_r\{f(x)\} + F(x)\,, \quad \mathrm{supp}F = \{x_s\}\,.$$

Since the support of F consists of isolated points,

$$F(x) = \sum_s F_s(x)\,, \quad F_s(x) = \sum_{n=0}^{M_s} A_{sn}\delta^{(n)}(x - x_s)$$

for some choice of constants A_{sn}. By construction f satisfies the equation in any I_s. The objective is to show that there always exists a choice of parameters C_s and A_{sn} for which the equation is satisfied in a neighborhood of each x_s.

The formal adjoint of the differential operator $L(D) = bD + a$ is defined by

$$L^*(D)\varphi = -D(b\varphi) + a\varphi$$

for any test function φ. It follows from (**28**.**1**) that the distribution F satisfies the equation

$$(L(D)F, \varphi) = -\Big(\mathcal{P}_r\{f\}, L^*(D)\varphi\Big) \equiv G\,.$$

The right-hand side of this equation G can be expressed as the sum over s of linear combinations of derivatives $\varphi^{(n)}(x_s) = (-1)^n(\delta^{(n)}(x - x_s), \varphi(x))$ because F is a distribution supported at $\{x_s\}$. To do so, define neighborhoods of singular points:

$$I'_s = (x_s - c, x_s + c)\,,$$

where c is small enough so that I'_s contains only one singular point x_s. Let I^c be the complement of the union of all I'_s. Then by Sec.**16.3.2** and continuity of the integral

$$G = \sum_s \lim_{\varepsilon\to 0^+} \int_{I'_{\varepsilon s}} \{f(x\}\Big(L^*\varphi(x) - P_s(x)\Big)\,dx + \int_{I^c} \{f(x)\}L^*\varphi(x)\,dx\,,$$

where $I'_{\varepsilon s}$ is defined by $\varepsilon < |x - x_s| < c$, and $P_s(x)$ is a Taylor polynomial for the test function $L^*\varphi$ about x_s of the minimal order defined by the condition that the function $\{f\}(L^*\varphi - P_s)$ is integrable on I'_s. If $\{f\}L^*\varphi$ is integrable on I_s, then $P_s = 0$. The classical solution $\{f\}$ is smooth in every $I'_{\varepsilon s}$ and in I^c. Integrating by parts in every integral containing the derivative $D(b\varphi)$ and using that $L(D)\{f(x)\} = 0$ in every $I'_{\varepsilon s}$ and in I^c, the last equation is reduced to

$$G = \sum_s \lim_{\varepsilon\to 0^+} \left(\int_{I'_{\varepsilon s}} \{f(x\}P_s(x)\,dx - \{f(x)\}b(x)\varphi(x)\Big|_{x_s-\varepsilon}^{x_s+\varepsilon}\right)\,.$$

It remains to find the polynomial P_s and calculate the integral and the limit. It should be stressed that the limit always exists for every s because $\mathcal{P}_r\{f\}$ is a distribution by construction. If x_s is not in the support of φ, then any Taylor polynomial of φ about x_s vanishes. This implies that the sum contains only finitely many terms because support of any test function contains at most finitely many points from $\{x_s\}$.

To simplify notations, put $y = x - x_p$. Near x_s, the coefficients b and a have the form

$$b(x) = y^{m_s}(b_0 + b_1 y + \cdots + b_k y^k + O(y^{k+1}))\,,$$
$$a(x) = y^{m_s - 1}(a_0 + a_1 y + \cdots + a_k + O(y^{k+1}))\,,$$

where, by the hypothesis, $m_s \geq 1$, $b_0 \neq 0$, and $\nu_s = a_0/b_0$. Put

$$\varphi(x) = \varphi_0 + \varphi_1 y + \cdots + \varphi_k y^k + O(y^{k+1})\,.$$

Then

$$\begin{aligned} L^*\varphi(x) &= y^{m_s-1}(\gamma_0 + \gamma_1 y + \cdots + \gamma_k y^k) + O(x^{k+m_p}) \\ &= P_s(x) + O(y^{k+m_p})\,, \\ \gamma_k &= \sum_{n=0}^{k} (a_n - (m_s - k)b_n)\varphi_{k-n}\,. \end{aligned}$$

Therefore $P_s = 0$ if $\nu_s - m_s \leq 0$. Indeed, one has $\gamma_0 = (\nu_s - m_s)b_0\varphi_0$. If $\gamma_0 \neq 0$, then $\{f\}L^*\varphi \sim |y|^{m_s - \nu_s - 1}$ is integrable if $\nu_s - m_s < 0$. If $\gamma_0 = 0$ or $\nu_s - m_s = 0$ because $b_0 \neq 0$, then $\{f\}L^*\varphi \sim |\gamma_1|$ is bounded near $y = 0$. If $\nu_s - m_s > 0$, then the minimal integer $k = k_s \geq 0$ for which the function

$$\{f(x)\}(L^*\varphi(x) - P_s(x)) = O(|y|^{k+m_s-\nu_s})$$

is integrable in a neighborhood of $y = 0$ is uniquely defined by the condition

$$\nu_s - m_s - 1 < k_s \leq \nu_s - m_s\,.$$

Thus, for each s, the limit in G is a linear combination of $\varphi_n = \varphi^{(n)}(x_s)/n!$, $0 \leq n \leq k_s$. As noted, this implies that the distribution F_s satisfies the equation

$$L(D)F_s(x) = \sum_{n=0}^{k_s} B_{sn}\delta(x - x_s)\,, \tag{28.5}$$

where the coefficients B_{sn} are uniquely determined by arbitrary parameters $C_s^+ = Q_s(x_s)$, $C_s^- = Q_{s-1}(x_s)$, and functions $a(x)$ and $b(x)$. It is always possible to choose Q_s so that $C_s^+ = C_s$ (a parameter of the classical solution in I_s) so that C_s^- is proportional to C_{s-1}. The distribution $L(D)F_s$ can always be written as a linear combinations of $\delta(x-x_s)$ and its derivatives. The coefficients are linear combinations of parameters A_{sn}. Therefore A_{sn} satisfy a linear system of equation. It is proved in the next section that *there always exists a choice of parameters C_s for which the linear system has a solution.*

28.4.2. The existence of a solution. [8]

Let us show that (**28.5**) always has a solution for some choice of parameters C_s. It is convenient to make a substitution

$$F_s(x) = e^{-Q_s(x)} G_s(x-x_s)\,, \quad Q_s'(x) = q_s(x)\,,$$

and rewrite the equation in terms of the variable $y = x - x_s$:

$$y^{m_s} G_s'(y) + \nu_s y^{m_s-1} G_s(y) = e^{Q_s(y+x_s)} \sum_{n=0}^{k_s} B_{sn}\delta^{(n)}(y) = \sum_{n=0}^{k_s} B_{sn}^Q \delta^{(n)}(y)\,,$$

where

$$G_s(y) = e^{Q_s(y+x_s)} \sum_{n=0}^{M_s} A_{sn}\delta^{(n)}(y) = \sum_{n=0}^{M_s} A_{sn}^Q \delta^{(n)}(y)\,.$$

The coefficients B_{sn}^Q and A_{sn}^Q are uniquely determined by B_{sn} and A_{sn}, respectively, and vice versa because e^{Q_s} does not vanish anywhere (see Sec.**18.1**). Using the results of Sec.**18.1**, the equation is further reduced to

$$\sum_{n=m_s-1}^{M_s} \alpha_{sn} A_{sn}^Q \delta^{(n-m_s+1)} = \sum_{n=0}^{k_s} B_{sn}^Q \delta^{(n)}\,,$$

$$\alpha_{sn} = -\frac{(-1)^{m_s} n!}{(n-m_s+1)!}\,(\nu_s - n - 1)\,.$$

Therefore A_{sn}^Q remains arbitrary for $0 \le n \le m_s - 2$ (no such parameters are present if $m_s = 1$). By shifting the summation index, one infers that

$$(-1)^{m_s} \frac{(n+m_s-1)!}{n!} (\nu_s - m_s - n) A_{s(n+m_s-1)}^Q = B_{sn}^Q$$

for $0 \le n \le k_s$. If ν_s is not an integer, then $M_s = k_s - m_s - 1$, and A_{sn}^Q are uniquely determined for $m_s - 1 \le n \le M_s$. Thus, the distribution F_s always exists if ν_s is not an integer and has It is concluded that a solution F_s exists and the distributional solution $f(x)$ near x_s has m_s free parameters, one parameter C_s and the others are A_{sn}^Q, $0 \le n \le m_s - 2$.

[8] This section is NOT complete yet and to be revised

Suppose that ν_s is an integer. The case $P_s = 0$ (and, hence, $B_{sn} = 0$) is possible only if $\nu_s = m_s \geq 1$. In this case, the equation for G_s is reduced to

$$(y^{m_s} G_s)' = 0 \quad \Rightarrow \quad G_s(y) = \sum_{n=0}^{m_s-1} A^Q_{sn} \delta^{(n)}(y)\,.$$

where A_{sn} are arbitrary. If $\nu_s - m_s > 0$ (or $P_s \neq 0$) and $\nu_s \geq 2$ is an integer, then $k_s = \nu_s - m_s > 0$. The last equation in the system for A_{sn} yields the condition on $C_s^{\pm}$:

$$B_{sn} = 0\,, \quad n = k_s = \nu_s - m_s\,,$$

whereas the coefficient A_{sn} remains arbitrary for $n = \nu_s - 1$.

In particular, if $\nu_s - m_s \leq 0$, then $P_s = 0$ and, hence, $B_{sn} = 0$ so that

$$F_s(y) = \sum_{n=0}^{m_s-2} A_{sn} \delta^{(n)}(y)\,, \quad m_s \geq 2\,, \quad F_s(y) = 0\,, \quad m_s = 1\,.$$

Recall that a distributional regularization is unique up to an additive distribution supported in the set to which the extension is carried out. Therefore any distributional solution near a root $x = 0$ must have the form

$$f(x) = \mathcal{P}_r\{f(x)\} + f_0(x)\,, \quad f_0(x) = \sum_{n=0}^{M} C_n \delta^{(n)}(x)\,,$$

where f_0 is a general distribution supported at $x = 0$. Note that $C_{\pm}$ and C_n are to be determined by substituting this distribution into the equation. The parameters $C_{\pm}$ are to be used to fulfill the matching conditions in open intervals between singular regular points of the equation. The objective is to investigate if they are independent parameters of the solution near each root of b.

Near $x = 0$, the coefficients b and a have the form

$$b(x) = x^m(b_0 + b_1 x + O(x^2))\,, \quad a(x) = x^{m-1}(a_0 + a_1 x + O(x^2))\,,$$

where, by the hypothesis, $b_0 \neq 0$, while a_0 can be zero, and $\nu = a_0/b_0$. It is convenient to make a substitution

$$f(x) = e^{-Q(x)} g(x)\,,$$

where Q is defined above, and reduce the equation an equivalent one

$$L(D)g(x) = x^m Dg(x) + \nu x^{m-1} g(x) = 0$$

near $x = 0$. As noted, any solution to this equation must have the form

$$g(x) = \mathcal{P}_r\{g(x)\} + g_0(x)\,, \quad g_0(x) = \sum_{n=0}^{M} C_n \delta^{(n)}(x)\,, \quad \{g(x)\} = \frac{C_\pm}{|x|^\nu}\,,$$

for some choice of parameters $C_\pm$ and C_n because $e^{-Q(x)}$ is smooth and vanish nowhere. The parameters are such that Eq. (**28**.**1**) is satisfied:

To calculate $\mathcal{P}_r\{g\}$, one has to find a Taylor polynomial p^L of a minimal order for the test function $L^*(D)\varphi$ about $x = 0$ such that $\{g\}(L^*\varphi - p^L)$ is locally integrable. Then

$$(\mathcal{P}_r\{g\}, L^*(D)\varphi) = \int_{|x|<1} \{g(x)\}\Big(L^*(D)\varphi(x) - p^L(x)\Big)\,dx + \int_{|x|>1} \{g(x)\} L^*(D)\varphi(x)\,dx\,.$$

In the first integral, the integration interval is reduced to $\varepsilon < |x| < 1$ with the subsequent limit $\varepsilon \to 0^+$ (by continuity of the integral). Then the integration by parts is carried out in all terms containing $D(x^m\varphi)$. Owing to that $L(D)\{g\} = 0$ for $|x| > \varepsilon$ and cancellation of the boundary terms at $x = 1$ in both the integrals, the equation is reduced to

$$(\mathcal{P}_r\{g\}, L^*(D)\varphi) = -\lim_{\varepsilon\to 0^+}\Big(\int_{\varepsilon<|x|<1} \{g(x)\} p^L(x)\,dx - \varepsilon^m\{g(\varepsilon)\}\varphi(\varepsilon) - (-\varepsilon)^m\{g(-\varepsilon)\}\varphi(-\varepsilon)\Big)\,.$$

It remains to find the polynomial p^L, calculate the integral and the limit.

Put $\varphi^{(n)}(0) = \varphi_n$. Then

$$L^*(D)\varphi(x) = x^{m-1}\sum_{n=0}^{k} (\nu - n - 1)\varphi_n x^n + O(x^{k+m})\,.$$

If $\nu - m + 1 < 1$ or $\nu < m$, the product $\{g\}L^*\varphi$ is integrable on the interval $|x| < 1$. Therefore $p^L = 0$. Otherwise,

$$p^L(x) = x^{m-1}\sum_{n=0}^{k} (\nu - n - 1)\varphi_n x^n\,, \tag{28.6}$$

where the order $k \geq 0$ is determined by $\nu - m \leq k < \nu - m + 1$.

28.4.3. The case $\nu < m$. [9] Since $p^L(x) = 0$,

$$(\mathcal{P}_r\{g\}, L^*(D)\varphi) = O(\varepsilon^{m-\nu}) \to 0$$

as $\varepsilon \to 0^+$ because $m-\nu > 0$. By (**28.5**) the distribution g_0 is anihillated by $L(D)$:

$$x^m g_0'(x) + x^{m-1} g_0(x) = 0\,.$$

Let $m = 1$ (a simple zero of b). In this case, one has to find a general distribution supported at $x = 0$ that satifies the equation

$$(\mathbf{28.7})\qquad x g_0'(x) + \nu g_0(x) = \sum_{n=0}^{M} (\nu - n - 1) C_n \delta^{(n)}(x) = 0$$

Therefore $C_n = 0$ because $\nu - n - 1 \neq 0$ for $\nu < m = 1$ and $n \geq 0$. It follows from $f = e^{-Q} g$ that the distributional and classical solutions coincide:

$$f(x) = \{f(x)\}\,,\quad \nu < m = 1\,,$$

as already noted earlier.

If $m \geq 2$, then the parameters C_n satisfy the condition that is obtained from (**28.7**) by multiplying the latter by x^{m-1}. Using the result of Sec.**18.1**, it is concluded that

$$(\nu - n - 1) C_n = 0\,,\quad n \geq m - 1\,,$$

whereas no conditions on C_n for $n = 0, 1, ..., m-2$ arise because $x^{m-1}\delta^{(n)}(x) = 0$ for $n < m-1$. Since $\nu < m$, $C_n = 0$ for $n \geq m-1$. The distribution $f_0 = e^{-Q} g_0$ also has exactly $m-1$ arbitrary parameters because $e^{-Q(0)} \neq 0$ and, hence, a general solution has the form

$$(\mathbf{28.8})\qquad f(x) = \mathcal{P}_r\{f(x)\} + \sum_{n=0}^{m-2} C_n \delta^{(n)}(x)\,,\quad \nu < m\,,\quad m \geq 2\,,$$

and depends on $m+1$ independent parameters, $C_\pm$ and C_n.

28.4.4. The case $\nu \geq m$. [10] Using (**28.6**) to evaluate the integral and a Taylor approximation of $\varphi(\pm\varepsilon)$, the limit is found to be

$$(\mathcal{P}_r\{g\}, L^*\varphi) = -\sum_{n=0}^{k} \left((-1)^{m+n} C_- - C_+\right) \varphi_n\,.$$

[9] This section is NOT complete yet and to be revised

[10] This section is NOT complete yet and to be revised

Since $\varphi_n = (-1)^n(\delta^{(n)}, \varphi)/n!$, it follows from (**28.5**) that the distribution g_0 satisfies the equation

$$x^m g_0'(x) + x^{m-1}g_0(x) = \sum_{n=0}^{k} A_n\delta^{(n)}(x)\,, \quad A_n = \frac{(-1)^m C_- - (-1)^n C_+}{n!}\,.$$

Using the results of Sec.**18.1**, the equation is reduced to

$$\sum_{n=m-1}^{M} \alpha_{mn}C_n\delta^{(n-m+1)} = \sum_{n=0}^{k} A_n\delta^{(n)}\,,$$
$$\alpha_{mn} = -\frac{(-1)^m n!}{(n-m+1)!}\,(\nu - n - 1)\,.$$

If ν is not an integer, then α_{mn} cannot be zero. In this case, the equation implies that C_n remain arbitrary for $n = 0, 1, ..., m-2$ (if $m = 1$, then no such C_n are present in the solution),

$$C_n = \frac{A_{n-m+1}}{\alpha_{mn}}\,, \quad m - 1 \le n \le k + m - 1\,.$$

and $C_n = 0$ for $n \ge k + m$. So, a general solution is given by the distribution (**28.8**) plus a linear combination of $\delta^{(n)}$, $n \ge m-1$, with coefficients uniquely determined by parameters $C_\pm$.

If $\nu \ge m$ is an integer, then $\nu = m + k$ by definition of the integer k. Therefore $\alpha_{mn} = 0$ if $n = \nu - 1$. This implies that $A_k = 0$ or

$$C_- = (-1)^\nu C_+$$

and the coefficient $C_{\nu-1}$ is also arbitrary in addition to C_n, $n \le m-2$. Under the above condition on $C_\pm$

$$\mathcal{P}_r\{g(x)\} = C_+\mathcal{P}\frac{1}{x^\nu}\,.$$

It also follows from the explicit form of A_n that $A_{k-2l} = 0$ for any non-negative integer $l \le k/2$ so that $C_n = C_+\beta_n$ for $m-1 \le n \le \nu-2$ and $\beta_{\nu-2l-1} = 0$ for $1 \le l \le k/2$. A general solution has again $m+1$ parameters:

$$g(x) = C_+\Big(\mathcal{P}\frac{1}{x^\nu} + \sum_{n=m-1}^{\nu-2} \beta_n\delta^{(n)}(x)\Big) + \sum_{n=0}^{m-2} C_n\delta^{(n)}(x) + C_{\nu-1}\delta^{(\nu-1)}(x)\,.$$

The distribution $f(x) = e^{-Q(x)}g(x)$ contains a linear combination of $\delta^{(n)}$ for $0 \le n \le \nu-1$ but among ν coefficients only $m+1$ are independent.

28.4.5. Matching conditions. [11] The above analysis shows that a distributional solution always exists near a regular singular point of the equation. Let $\{x_n\}$ be a sequence of all such points (and no other singular points exist), ν_n and m_n be parameters of a singular point $x = x_n$, and $C_n^{\pm}$ be parameters of the classical solution near $x = x_n$. If $\nu_n \geq m_n$ is an integer, then $C_n^- = (-1)^{\nu} C_n^+$. In the interval (x_n, x_{n+1}), the distributional solution coincides with the classical one

$$f(x) = \{f_n(x)\} = C_n^0 \exp\Big(- \int \frac{a(x)}{b(x)}\, dx \Big)\,, \quad x \in (x_n, x_{n+1})\,,$$

where C_n^0 is an arbitrary constant. There are two equivalent representations on the integrand in this interval

$$\frac{a(x)}{b(x)} = \frac{\nu_n}{x - x_n} + q_n(x) = \frac{\nu_{n+1}}{x - x_{n+1}} + q_{n+1}(x)\,.$$

If the first representation is used in the integral, the local solution near $x = x_n$ is obtained in (x_n, x_{n+1}). Therefore $C_n^+ = C_n^0$. If the second representation is used, the local solution near $x = x_{n+1}$ is obtained in (x_n, x_{n+1}) and, hence, $C_{n+1}^- = C_n^0$. Thus, the local distributional solutions obtained near every singular point define a distribution on $\mathbb{R}$ if

$$C_n^+ = C_{n+1}^-\,.$$

This condition can always be fulfilled. This is obvious if none of ν_n is an integer. In this case, $C_n^+ = C_n^0$ are arbitrary. Suppose that all $\nu_n \geq m_n$ are integers. Then the only independent parameters to be used for matching are C_n^+. In this case, the matching condition reads $C_{n+1}^+ = (-1)^{\nu_{n+1}} C_n^+$ or $C_{n+1}^0 = (-1)^{\nu_{n+1}} C_n^0$. Therefore the integration constants C_n^0 in every interval (x_n, x_{n+1}) must be equal up to a sign.

28.5. Exercises.

1. Find general first and second antiderivatives of Sokhotsky distributions

$$D^{-1}\, \frac{1}{x \pm i0^+}\,, \qquad D^{-2}\, \frac{1}{x \pm i0^+}$$

[11] This section is NOT complete yet and to be revised

2. Find a general distributional solution to each of the following equations

$$
\begin{aligned}
&\text{(i)} \qquad (x-a)(x-b)f'(x) = 1 \\
&\text{(ii)} \qquad (x-a)(x-b)f''(x) = \delta(x) \\
&\text{(iii)} \qquad f'(x) + a(x)f(x) = \delta'(x)\,, \quad a \in C^\infty \\
&\text{(iv)} \qquad xf'(x) + xa(x)f(x) = \delta(x)\,, \quad a \in C^\infty
\end{aligned}
$$

Hint: Multiply Equations (ii) and (iv) by x.

3. Let $f(x,t)$ be a distribution of two real variables. Find a general solution to an inhomogeneous distributional wave equation:

$$(D_t^2 - c^2 D_x)u(x,t) = f(x,t)$$

Express the answer in terms of antiderivatives of f with respect to x and t.