ANALYSIS

SECOND EDITION

Elliott H. Lieb
Princeton University

Michael Loss
Georgia Institute of Technology

Graduate Studies
in Mathematics
Volume 14

American Mathematical Society
Providence, Rhode Island

Chapter 6

Distributions

6.1 INTRODUCTION

The notion of a weak derivative is an indispensable tool in dealing with partial differential equations. Its advantage is that it allows one to dispense with subtle questions about differentiation, such as the interchange of partial derivatives. Its main point is that every locally integrable function can be weakly differentiated indefinitely many times, just as though it were a C^∞-function. The weakening of the notion of a derivative makes it easier to find solutions to equations and, once found, these 'weak' solutions can then be analyzed to find out if they are, in fact, truly differentiable in the classical sense. An analogy in elementary algebra might be trying to solve a polynomial equation by rational numbers. It is extremely important, at the beginning of the investigation, to know that solutions always exist in the larger category of real numbers; many techniques are available for this purpose, e.g. Rolle's theorem, that are not available in the category of rationals. Later on one can try to prove that the solutions are, in fact, rational.

A theory developed around the notion that every L^1_{loc}-function is differentiable is the theory of distributions invented by [Schwartz] (see [Hörmander], [Rudin, 1991], [Reed–Simon, Vol. 1]). Although we do not present some of the deeper aspects of this theory we shall state its basic techniques. In the following, for completeness, we define distributions for an arbitrary open set Ω in $\mathbb{R}^n$ but, in fact, we shall mainly need the case $\Omega = \mathbb{R}^n$ in the rest of the book.

6.2 TEST FUNCTIONS (The space $\mathcal{D}(\Omega)$)

Let Ω be an open, nonempty set in $\mathbb{R}^n$; in particular Ω can be $\mathbb{R}^n$ itself. Recall from Sect. 1.1 that $C_c^\infty(\Omega)$ denotes the space of all infinitely differentiable, complex-valued functions whose support is compact and in Ω. Recall also that the support of a continuous function is defined to be the closure of the set on which the function does not vanish, and compactness means that the closed set is also contained in some ball of finite radius. Note that Ω is never compact.

The space of **test functions**, $\mathcal{D}(\Omega)$, consists of all the functions in $C_c^\infty(\Omega)$ supplemented by the following notion of convergence: A *sequence* $\phi^m \in C_c^\infty(\Omega)$ **converges in $\mathcal{D}(\Omega)$ to the function** $\phi \in C_c^\infty(\Omega)$ if and only if there is some *fixed, compact* set $K \subset \Omega$ such that the support of $\phi^m - \phi$ is in K for all m and, for each choice of the nonnegative integers $\alpha_1, \ldots, \alpha_n$,

$$(\partial/\partial x_1)^{\alpha_1} \cdots (\partial/\partial x_n)^{\alpha_n} \phi^m \longrightarrow (\partial/\partial x_1)^{\alpha_1} \cdots (\partial/\partial x_n)^{\alpha_n} \phi \tag{1}$$

as $m \to \infty$, *uniformly* on K. To say that a sequence of continuous functions ψ^m converges to ψ uniformly on K means that

$$\sup_{x\in K} |\psi^m(x) - \psi(x)| \to 0 \quad \text{as} \quad m \to \infty.$$

$\mathcal{D}(\Omega)$ is a linear space, i.e., functions can be added and multiplied by (complex) scalars.

6.3 DEFINITION OF DISTRIBUTIONS AND THEIR CONVERGENCE

A **distribution** T is a continuous linear functional on $\mathcal{D}(\Omega)$, i.e., $T : \mathcal{D}(\Omega) \to \mathbb{C}$ such that for $\phi, \phi_1, \phi_2 \in \mathcal{D}(\Omega)$ and $\lambda \in \mathbb{C}$

$$T(\phi_1 + \phi_2) = T(\phi_1) + T(\phi_2) \quad \text{and} \quad T(\lambda\phi) = \lambda T(\phi), \tag{1}$$

and continuity means that whenever $\phi^n \in \mathcal{D}(\Omega)$ and $\phi^n \to \phi$ in $\mathcal{D}(\Omega)$

$$T(\phi^n) \to T(\phi).$$

Distributions can be added and multiplied by complex scalars. This linear space is denoted by $\mathcal{D}'(\Omega)$, the **dual space of** $\mathcal{D}(\Omega)$.

There is an obvious notion of convergence of distributions: A sequence of distributions $T^j \in \mathcal{D}'(\Omega)$ **converges in** $\mathcal{D}'(\Omega)$ to $T \in \mathcal{D}'(\Omega)$ if, for every $\phi \in \mathcal{D}(\Omega)$, the numbers $T^j(\phi)$ converge to $T(\phi)$.

One might suspect that this kind of convergence is rather weak. Indeed, it is! For example, we shall see in Sect. 6.6, where we develop a notion of the derivative of a distribution, that for *any* converging sequence of distributions their derivatives converge too, i.e., differentiation is a continuous operation in $\mathcal{D}'(\Omega)$. This contrasts with ordinary pointwise convergence because the derivatives of a pointwise converging sequence of functions need not, in general, converge anywhere.

Another instance, as we shall see in Sect. 6.13, is that any distribution can be approximated in $\mathcal{D}'(\Omega)$ by functions in $C^\infty(\Omega)$. To make sense of that statement, we first have to define what it means for a function to be a distribution. This is done in the next section.

6.4 LOCALLY SUMMABLE FUNCTIONS, $L^p_{\rm loc}(\Omega)$

The foremost example of distributions are functions themselves. We begin by defining the space of **locally p^{th}-power summable functions**, $L^p_{\rm loc}(\Omega)$, for $1 \le p \le \infty$. Such functions are Borel measurable functions defined on all of Ω and with the property that

$$\|f\|_{L^p(K)} < \infty \tag{1}$$

for *every* compact set $K \subset \Omega$. Equivalently, it suffices to require (1) to hold when K is any closed ball in Ω.

A sequence of functions $f^1, f^2, \ldots$ in $L^p_{\rm loc}(\Omega)$ is said to **converge** (or **converge strongly**) to f in $L^p_{\rm loc}(\Omega)$ (denoted by $f^j \to f$) if $f^j \to f$ in $L^p(K)$ in the usual sense (see Theorem 2.7) for every compact $K \subset \Omega$. Likewise, f^j **converges weakly** to f if $f^j \rightharpoonup f$ weakly in every $L^p(K)$ (2.9(6)).

Note (for general $p \ge 1$) that $L^p_{\rm loc}(\Omega)$ is a vector space but it does not have a simply defined norm. Furthermore, $f \in L^p_{\rm loc}(\Omega)$ does not imply that $f \in L^p(\Omega)$. Clearly, $L^p_{\rm loc}(\Omega) \supset L^p(\Omega)$ and, if $r > p$, we have the inclusion

$$L^p_{\rm loc}(\Omega) \supset L^r_{\rm loc}(\Omega),$$

by Hölder's inequality (but it is *false*—unless Ω has finite measure—that $L^p(\Omega) \supset L^r(\Omega)$).

As far as distributions are concerned, $L^1_{\rm loc}(\Omega)$ is the most important space. Let f be a function in $L^1_{\rm loc}(\Omega)$. For any ϕ in $\mathcal{D}(\Omega)$ it makes sense to consider

$$T_f(\phi) := \int_\Omega f\phi \, \mathrm{d}x, \tag{2}$$

which obviously defines a linear functional on $\mathcal{D}(\Omega)$. T_f is also continuous since

$$\begin{aligned}|T_f(\phi) - T_f(\phi^m)| &= \left|\int_\Omega (\phi(x) - \phi^m(x))f(x)\,\mathrm{d}x\right| \\ &\le \sup_{x\in K} |\phi(x) - \phi^m(x)| \int_K |f(x)|\,\mathrm{d}x,\end{aligned}$$

which tends to zero by the uniform convergence of the ϕ^m's. Thus T_f is in $\mathcal{D}'(\Omega)$. If a distribution T is given by (2) for some $f \in L^1_{\mathrm{loc}}(\Omega)$, we say that **the distribution T is the function f.** This terminology will be justified in the next section.

An important example of a distribution that is *not* of this form is the so-called **Dirac 'delta-function'**, which is not a function at all:

$$\delta_x(\phi) = \phi(x) \tag{3}$$

with $x \in \Omega$ fixed. It is obvious that $\delta_x \in \mathcal{D}'(\Omega)$. Thus, the delta-measure of Sect. 1.2(6), like any Borel measure, can also be considered to be a distribution. In fact, one can say that it was partly the attempt to understand the true mathematical meaning of the delta function, which had been used so successfully by physicists and engineers, that led to the theory of distributions.

Although $\mathcal{D}(\Omega)$, the space of test functions, is a very restricted class of functions it is large enough to distinguish functions in $\mathcal{D}'(\Omega)$, as we now show.

6.5 THEOREM (Functions are uniquely determined by distributions)

Let $\Omega \subset \mathbb{R}^n$ be open and let f and g be functions in $L^1_{\mathrm{loc}}(\Omega)$. Suppose that the distributions defined by f and g are equal, i.e.,

$$\int_\Omega f\phi = \int_\Omega g\phi \tag{1}$$

for all $\phi \in \mathcal{D}(\Omega)$. Then $f(x) = g(x)$ for almost every x in Ω.

PROOF. For $m = 1, 2, \ldots$ let Ω_m be the set of points $x \in \Omega$ such that $x + y \in \Omega$ whenever $|y| \le \frac{1}{m}$. Ω_m is open. Let j be in $C_c^\infty(\mathbb{R}^n)$ with support in the unit ball and with $\int_{\mathbb{R}^n} j = 1$. Define $j_m(x) = m^n j(mx)$. Fix M. If $m \ge M$, then, by (1) with $\phi(y) = j_m(x - y)$, we have that

$(j_m * f)(x) = (j_m * g)(x)$ for all $x \in \Omega_M$ (see Sect. 2.15 for the definition of the convolution $*$). By Theorem 2.16, $j_m * f \to f$ and $j_m * g \to g$ in $L^1_{\mathrm{loc}}(\Omega_M)$ as $m \to \infty$. Thus $f = g$ in $L^1_{\mathrm{loc}}(\Omega_M)$ and therefore $f(x) = g(x)$ for almost every $x \in \Omega_M$. Finally let M tend to ∞. ■

6.6 DERIVATIVES OF DISTRIBUTIONS

We now define the notion of **distributional** or **weak derivative**. Let T be in $\mathcal{D}'(\Omega)$ and let $\alpha_1, \ldots, \alpha_n$ be nonnegative integers. We define the distribution $(\partial/\partial x_1)^{\alpha_1} \cdots (\partial/\partial x_n)^{\alpha_n} T$, denoted by $D^\alpha T$, by its action on each $\phi \in \mathcal{D}(\Omega)$ as follows:

$$(D^\alpha T)(\phi) = (-1)^{|\alpha|} T(D^\alpha \phi) \tag{1}$$

with the notation

$$|\alpha| = \sum_{i=1}^{n} \alpha_i. \tag{2}$$

The symbol

$$\partial_i T$$

denotes $D^\alpha T$ in the special case $\alpha_i = 1, \alpha_j = 0$ for $j \neq i$.

The symbol ∇T, called the **distributional gradient of** T, denotes the n-tuple $(\partial_1 T, \partial_2 T, \ldots, \partial_n T)$.

If f is a $C^{|\alpha|}(\Omega)$-function (not necessarily of compact support), then

$$(D^\alpha T_f)(\phi) := (-1)^{|\alpha|} \int_\Omega (D^\alpha \phi) f \, \mathrm{d}x = \int_\Omega (D^\alpha f) \phi \, \mathrm{d}x =: T_{D^\alpha f}(\phi),$$

where the middle equality holds by partial integration. Hence the notion of weak derivative extends the classical one and it agrees with the classical one whenever the classical derivative exists and is continuous (see Theorem 6.10 (equivalence of classical and distributional derivatives)). Obviously, in this weak sense, *every distribution is infinitely often differentiable* and this is one of the main virtues of the theory. Note however, that the distributional derivative of a nondifferentiable function (in the classical sense) is not necessarily a function.

Let us show that $D^\alpha T$ actually is a distribution. Obviously it is linear, so we only have to check its continuity on $\mathcal{D}(\Omega)$. Let $\phi^m \to \phi$ in $\mathcal{D}(\Omega)$. Then $D^\alpha \phi^m \to D^\alpha \phi$ in $\mathcal{D}(\Omega)$ since

$$\operatorname{supp}\{D^\alpha \phi^m - D^\alpha \phi\} \subset \operatorname{supp}\{\phi^m - \phi\} \subset K$$

and

$$D^\beta(D^\alpha\phi^m - D^\alpha\phi) = D^{\beta+\alpha}\phi^m - D^{\beta+\alpha}\phi$$

converges to zero uniformly on compact sets. [Here $\beta+\alpha$ simply denotes the multi-index given by $(\beta_1+\alpha_1, \beta_2+\alpha_2, \dots, \beta_n+\alpha_n)$]. Thus, $D^\alpha\phi$ and $D^\alpha\phi^m$ are themselves functions in $\mathcal{D}(\Omega)$ with $D^\alpha\phi^m \to D^\alpha\phi$ as $m\to\infty$. Hence, as $m\to\infty$,

$$(D^\alpha T)(\phi^m) := (-1)^{|\alpha|}T(D^\alpha\phi^m) \longrightarrow (-1)^{|\alpha|}T(D^\alpha\phi) =: (D^\alpha T)(\phi).$$

We end this section by showing that differentiation of distributions is a continuous operation in $\mathcal{D}'(\Omega)$. Indeed, if $T^j(\phi)\to T(\phi)$ for all $\phi\in\mathcal{D}(\Omega)$, then, by the definition of the derivative of a distribution

$$(D^\alpha T^j)(\phi) = (-1)^{|\alpha|}T^j(D^\alpha\phi) \underset{j\to\infty}{\to} (-1)^{|\alpha|}T(D^\alpha\phi) = (D^\alpha T)(\phi)$$

since $D^\alpha\phi\in\mathcal{D}(\Omega)$.

6.7 DEFINITION OF $W^{1,p}_{\mathrm{loc}}(\Omega)$ AND $W^{1,p}(\Omega)$

$L^1_{\mathrm{loc}}(\Omega)$-functions are an important class of distributions, but we can usefully refine that class by studying functions whose distributional first derivatives are also $L^1_{\mathrm{loc}}(\Omega)$-functions. This class is denoted by $W^{1,1}_{\mathrm{loc}}(\Omega)$. Furthermore, just as $L^p_{\mathrm{loc}}(\Omega)$ is related to $L^1_{\mathrm{loc}}(\Omega)$ we can also define the class of functions $W^{1,p}_{\mathrm{loc}}(\Omega)$ for each $1\le p\le\infty$. Thus,

$$\begin{aligned} W^{1,p}_{\mathrm{loc}}(\Omega) = \big\{ f:\Omega\to\mathbb{C} : f\in L^p_{\mathrm{loc}}(\Omega) \text{ and } \partial_i f, \text{ as a distribution} \\ \text{in } \mathcal{D}'(\Omega), \text{ is an } L^p_{\mathrm{loc}}(\Omega)\text{-function for } i=1,\dots,n\big\}. \end{aligned}$$

We urge the reader not to use the symbol ∇f at first, since it is tempting to apply the rules of calculus which we have not established yet. One should just think of ∇f as n functions $\mathbf{g}=(g_1,\dots,g_n)$, each of which is in $L^p_{\mathrm{loc}}(\Omega)$, such that

$$\int_\Omega f\nabla\phi = -\int_\Omega \mathbf{g}\phi \quad \text{for all } \phi\in\mathcal{D}(\Omega).$$

This set of functions, $W^{1,p}_{\mathrm{loc}}(\Omega)$, forms a vector space but not a normed one. We have the inclusion $W^{1,p}_{\mathrm{loc}}(\Omega)\supset W^{1,r}_{\mathrm{loc}}(\Omega)$ if $r>p$.

We can also define $W^{1,p}(\Omega) \subset W^{1,p}_{\text{loc}}(\Omega)$ analogously:

$$W^{1,p}(\Omega) = \{f : \Omega \to \mathbb{C} : f \text{ and } \partial_i f \text{ are in } L^p(\Omega) \text{ for } i = 1, \dots, n\}.$$

We can make $W^{1,p}(\Omega)$ into a normed space, by defining

$$\|f\|_{W^{1,p}(\Omega)} = \left\{ \|f\|^p_{L^p(\Omega)} + \sum_{j=1}^n \|\partial_j f\|^p_{L^p(\Omega)} \right\}^{1/p} \tag{1}$$

and it is complete, i.e., every Cauchy sequence in this norm has a limit in $W^{1,p}(\Omega)$. This follows easily from the completeness of $L^p(\Omega)$ (Theorem 2.7) together with the definition 6.6 of the distributional derivative, i.e., if $f^j \to f$ and $\partial_i f^j \to g_i$, then it follows that $g_i = \partial_i f$ in $\mathcal{D}'(\Omega)$. The proof is a simple adaptation of the one for $W^{1,2}(\Omega) = H^1(\Omega)$ in Theorem 7.3 (see Remark 7.5). We leave the details to the reader.

The spaces $W^{1,p}(\Omega)$ are called **Sobolev spaces**. In this chapter only $W^{1,1}_{\text{loc}}(\Omega)$ will play a role.

The superscript 1 in $W^{1,p}(\Omega)$ denotes the fact that the first derivatives of f are p^{th}-power summable functions.

As with $L^p(\Omega)$ and $L^p_{\text{loc}}(\Omega)$, we can define the notions of **strong** and **weak convergence** in the spaces $W^{1,p}_{\text{loc}}(\Omega)$ or $W^{1,p}(\Omega)$ of a sequence of functions $f^1, f^2, \dots$ to a function f. Strong convergence simply means that the sequence converges strongly to f in $L^p(\Omega)$ and the n sequences $\{\partial_1 f^j\}, \dots, \{\partial_n f^j\}$, formed from the derivatives of f^j, converge in $L^p(\Omega)$ to the n functions $\partial_1 f, \dots, \partial_n f$ in $L^p(\Omega)$. In the case of $W^{1,p}_{\text{loc}}(\Omega)$ we require this convergence only on every compact subset of Ω. Similarly, for weak convergence in $W^{1,p}(\Omega)$ we require that for every $L \in L^p(\Omega)^*$, $L(f^j - f) \to 0$ and, for each i, $L(\partial_i f^j - \partial_i f) \to 0$ as $j \to \infty$. For $W^{1,p}_{\text{loc}}(\Omega)$ we require weak convergence in $W^{1,p}(\mathcal{O})$ for every open set $\mathcal{O}$ with $\mathcal{O} \subset K \subset \Omega$ where K is compact. (Recall Theorem 2.14 for $L^p(\Omega)^*$ when $1 \le p < \infty$.)

Similar definitions apply to $W^{m,p}(\Omega)$ and $W^{m,p}_{\text{loc}}(\Omega)$ with $m > 1$. The first m derivatives of these functions are $L^p(\Omega)$-functions and, similarly to (1),

$$\begin{aligned} \|f\|^p_{W^{m,p}(\Omega)} := \|f\|^p_{L^p(\Omega)} &+ \sum_{j=1}^n \|\partial_j f\|^p_{L^p(\Omega)} \\ &+ \cdots + \sum_{j_1=1}^n \cdots \sum_{j_m=1}^n \|\partial_{j_1} \cdots \partial_{j_m} f\|^p_{L^p(\Omega)}. \end{aligned} \tag{2}$$

● In the following it will be convenient to denote by ϕ_z the function ϕ translated by $z \in \mathbb{R}^n$, i.e.,

$$\phi_z(x) := \phi(x - z). \tag{3}$$

6.8 LEMMA (Interchanging convolutions with distributions)

Let $\Omega \subset \mathbb{R}^n$ be open and let $\phi \in \mathcal{D}(\Omega)$. Let $\mathcal{O}_\phi \subset \mathbb{R}^n$ be the set

$$\mathcal{O}_\phi = \{y \ : \ \mathrm{supp}\{\phi_y\} \subset \Omega\}.$$

It is elementary that $\mathcal{O}_\phi$ is open and not empty. Let $T \in \mathcal{D}'(\Omega)$. Then the function $y \mapsto T(\phi_y)$ is in $C^\infty(\mathcal{O}_\phi)$. In fact, with D_y^α denoting derivatives with respect to y,

$$D_y^\alpha T(\phi_y) = (-1)^{|\alpha|} T((D^\alpha \phi)_y) = (D^\alpha T)(\phi_y). \tag{1}$$

Now let $\psi \in L^1(\mathcal{O}_\phi)$ have compact support. Then

$$\int_{\mathcal{O}_\phi} \psi(y) T(\phi_y)\,\mathrm{d}y = T(\psi * \phi). \tag{2}$$

PROOF. If $y \in \mathcal{O}_\phi$ and if $\varepsilon > 0$ is chosen so that $y + z \in \mathcal{O}_\phi$ for all $|z| < \varepsilon$, we have that for all $x \in \Omega$

$$|\phi_y(x) - \phi_{y+z}(x)| = |\phi(x - y) - \phi(x - y - z)| < C\varepsilon \tag{3}$$

for some number $C < \infty$. This is so because ϕ has continuous derivatives and (since it has compact support) these derivatives are *uniformly* continuous. For the same reason, (3) holds for all derivatives of ϕ (with C depending on the order of the derivative). This means that ϕ_{y+z} converges to ϕ_y as $z \to 0$ in $\mathcal{D}(\Omega)$ (see Sect. 6.2). Therefore, $T(\phi_{y+z}) \to T(\phi_y)$ as $z \to 0$, and thus $y \mapsto T(\phi_y)$ is continuous on $\mathcal{O}_\phi$.

Similarly, we have that

$$\Big| [\phi(x + \delta z) - \phi(x)]/\delta - \nabla\phi(x) \cdot z \Big| \le C'\delta|z|$$

and thus, by a similar argument, $y \mapsto T(\phi_y)$ is differentiable. Continuing in this manner we find that (1) holds.

To prove (2) it suffices to assume that $\psi \in C_c^\infty(\mathcal{O}_\phi)$. To verify this, we use Theorem 2.16 to find, for each $\delta > 0$, $\psi^\delta \in C_c^\infty(\mathcal{O}_\phi)$ so that $\int_{\mathcal{O}_\phi} |\psi^\delta - \psi| < \delta$. In fact, we can assume that $\mathrm{supp}\{\psi^\delta\}$ is contained in some fixed compact subset, K, of $\mathcal{O}_\phi$, independent of δ. Then

$$\left| \int \{\psi(y) - \psi^\delta(y)\} T(\phi_y)\, \mathrm{d}y \right| \le \delta \sup\left\{ |T(\phi_y)| : y \in K \right\}.$$

It is also easy to see that $\psi^\delta * \phi$ converges to $\psi * \phi$ in $\mathcal{D}(\Omega)$ and therefore $T(\psi^\delta * \phi) \to T(\psi * \phi)$.

With ψ now in $C_c^\infty(\mathcal{O}_\phi)$ we note that the integrand in (2) is a product of two C_c^∞-functions. Hence the integral can be taken as a Riemann integral and thus can be approximated by finite sums of the form

$$\Delta_m \sum_{j=1}^m \psi(y_j) T(\phi_{y_j}) \quad \text{with } \Delta_m \to 0 \text{ as } m \to \infty.$$

Likewise, for any multi-index α, $(D^\alpha(\psi * \phi))(x)$ is uniformly approximated by $\Delta_m \sum_{j=1}^m \psi(y_j) D^\alpha \phi(x - y_j)$ as $m \to \infty$ (because $\phi \in C_c^\infty(\Omega)$). Note that for m sufficiently large every member of this sequence has support in a fixed compact set $K \subset \Omega$. Since T is continuous (by definition) and the function $\eta_m(x) = \Delta_m \sum_{j=1}^n \psi(y_j)\phi(x - y_j)$ converges in $\mathcal{D}(\Omega)$ to $(\psi * \phi)(x)$ as $m \to \infty$, we conclude that $T(\eta_m)$ converges to $T(\psi * \phi)$ as $m \to \infty$. ■

6.9 THEOREM (Fundamental theorem of calculus for distributions)

Let $\Omega \subset \mathbb{R}^n$ be open, let $T \in \mathcal{D}'(\Omega)$ be a distribution and let $\phi \in \mathcal{D}(\Omega)$ be a test function. Suppose that for some $y \in \mathbb{R}^n$ the function ϕ_{ty} is also in $\mathcal{D}(\Omega)$ for all $0 \le t \le 1$ (see 6.7(3)). Then

$$T(\phi_y) - T(\phi) = \int_0^1 \sum_{j=1}^n y_j (\partial_j T)(\phi_{ty})\, \mathrm{d}t. \tag{1}$$

As a particular case of (1), suppose that $f \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$. Then, for each y in $\mathbb{R}^n$ and almost every $x \in \mathbb{R}^n$,

$$f(x+y) - f(x) = \int_0^1 y \cdot \nabla f(x + ty)\, \mathrm{d}t. \tag{2}$$

PROOF. Let $\mathcal{O}_\phi = \{z \in \mathbb{R}^n : \phi_z \in \mathcal{D}(\Omega)\}$. It is clearly open and nonempty. Denote the right side of (1) by $F(y)$. Observe that by Lemma 6.8, $z \mapsto (\partial_j T)(\phi_z)$ is a C^∞-function on $\mathcal{O}_\phi$ and $\partial(\partial_j T(\phi_z))/\partial z_i = -\partial_j T(\partial_i \phi_z)$.

With this infinite differentiability in mind we can now interchange derivatives and integrals, and compute

$$\partial_i F(y) = -\sum_{j=1}^{n} \int_0^1 t(\partial_j T)(\partial_i \phi_{ty}) y_j \, \mathrm{d}t + \int_0^1 (\partial_i T)(\phi_{ty}) \, \mathrm{d}t.$$

The first term is, by the definition of the derivative of a distribution,

$$\sum_{j=1}^{n} \int_0^1 tT(\partial_j \partial_i \phi_{ty}) y_j \, \mathrm{d}t = -\int_0^1 \sum_{j=1}^{n} t(\partial_i T)(\partial_j \phi_{ty}) y_j \, \mathrm{d}t,$$

which can be rewritten (for the same reason as before) as

$$\int_0^1 t \frac{\mathrm{d}}{\mathrm{d}t} (\partial_i T)(\phi_{ty}) \, \mathrm{d}t.$$

A simple integration by parts then yields $\partial_i F(y) = (\partial_i T)(\phi_y)$. The function $y \mapsto G(y) = T(\phi_y) - T(\phi)$ is also C^∞ in y (by Lemma 6.8) and also has $(\partial_i T)(\phi_y)$ as its partial derivatives. Since $F(0) = G(0) = 0$, the two C^∞-functions F and G must be the same. This proves (1).

To prove (2), note that since

$$(\partial_j f)(\phi_{ty}) = \int \phi(x)(\partial_j f)(x + ty) \, \mathrm{d}x,$$

(1) implies that

$$\int_{\mathbb{R}^n} \phi(x)[f(x+y) - f(x)] \, \mathrm{d}x = \int_0^1 \sum_{j=1}^{n} y_j \left\{ \int_{\mathbb{R}^n} \phi(x)(\partial_j f)(x + ty) \, \mathrm{d}x \right\} \mathrm{d}t.$$

Since ϕ has compact support, the integrand is (t, x) integrable (even if $\partial_j f \notin L^1(\mathbb{R}^n)$), and hence Fubini's theorem can be used to interchange the t and x integrations. Conclusion (2) then follows from Theorem 6.5. ■

6.10 THEOREM (Equivalence of classical and distributional derivatives)

Let $\Omega \subset \mathbb{R}^n$ be open, let $T \in \mathcal{D}'(\Omega)$ and set $G_i := \partial_i T \in \mathcal{D}'(\Omega)$ for $i = 1, 2, \ldots, n$. The following are equivalent.

(i) *T is a function $f \in C^1(\Omega)$.*

(ii) *G_i is a function $g_i \in C^0(\Omega)$ for each $i = 1, \ldots, n$.*

In each case, g_i is $\partial f / \partial x_i$, the classical derivative of f.

REMARK. The assertion $f \in C^1(\Omega)$ means, of course, that there is a $C^1(\Omega)$-function in the equivalence class of f. A similar remark applies to $g_i \in C^0(\Omega)$.

PROOF. $\boxed{\texttt{(i)} \Rightarrow \texttt{(ii).}}$ $G_i(\phi) = (\partial_i T)(\phi) = -\int_\Omega (\partial_i\phi) f$ by the definition of distributional derivative. On the other hand, the classical integration by parts formula yields

$$\int_\Omega (\partial_i\phi) f = -\int_\Omega \phi(\partial f/\partial x_i)$$

since ϕ has compact support in Ω and $f \in C^1(\Omega)$. Therefore, by the terminology of Sect. 6.4 and Theorem 6.5, G_i is the function $\partial f/\partial x_i$.

$\boxed{\texttt{(ii)} \Rightarrow \texttt{(i).}}$ Fix $R > 0$ and let $\omega = \{x \in \Omega : |x - z| > R \text{ for all } z \notin \Omega\}$. Clearly ω is open and nonempty for R small enough, which we henceforth assume. Take $\phi \in \mathcal{D}(\omega) \subset \mathcal{D}(\Omega)$ and $|y| < R$. Then $\phi_{ty} \in \mathcal{D}(\Omega)$ for $-1 \le t \le 1$. By 6.9(1) and Fubini's theorem

$$\begin{aligned} T(\phi_y) - T(\phi) &= \int_0^1 \sum_{j=1}^n y_j \int_\omega g_j(x)\phi(x - ty)\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_\omega \left\{ \int_0^1 \sum_{j=1}^n g_j(x + ty) y_j \,\mathrm{d}t \right\} \phi(x)\,\mathrm{d}x. \end{aligned} \tag{1}$$

Pick $\psi \in C_c^\infty(\mathbb{R}^n)$ nonnegative with $\operatorname{supp}\{\psi\} \subset B := \{y : |y| < R\}$ and $\int \psi = 1$. The convolution $\int_B \psi(y)\phi(x - y)\,\mathrm{d}y$ with $\phi \in \mathcal{D}(\omega)$ defines a function in $\mathcal{D}(\Omega)$. Integrating (1) against ψ we obtain, using Fubini's theorem,

$$\begin{aligned} &\int_B \psi(y) T(\phi_y)\,\mathrm{d}y - T(\phi) \\ &\quad = \int_\omega \left\{ \sum_{j=1}^n \int_B \psi(y) \int_0^1 y_j g_j(x + ty)\,\mathrm{d}t\,\mathrm{d}y \right\} \phi(x)\,\mathrm{d}x. \end{aligned} \tag{2}$$

The first term on the left is $\int_\omega \phi(x) T(\psi_x)\,\mathrm{d}x$, which follows from Lemma 6.8 by noting that ψ_x for $x \in \omega$ is an element of $\mathcal{D}(\Omega)$. Hence

$$T(\phi) = \int_\omega \left\{ T(\psi_x) - \sum_{j=1}^n \int_B \psi(y) \int_0^1 y_j g_j(x + ty)\,\mathrm{d}t\,\mathrm{d}y \right\} \phi(x)\,\mathrm{d}x,$$

which displays T explicitly as a function, which we denote by f.

Finally, by Theorem 6.9(2)

$$f(x+y) - f(x) = \int_0^1 \sum_{j=1}^{n} g_j(x+ty)y_j \, \mathrm{d}t \tag{3}$$

for $x \in \omega$ and $|y| < R$. The right side is

$$\sum_{j=1}^{n} g_j(x)y_j + o(|y|)$$

and this proves that $f \in C^1(\omega)$ with derivatives g_i. This suffices, since x can be arbitrarily chosen in Ω by choosing R to be small enough. ■

The following is a special case of Theorem 6.10, which we state separately for emphasis.

6.11 THEOREM (Distributions with zero derivatives are constants)

Let $\Omega \subset \mathbb{R}^n$ be a connected, open set and let $T \in \mathcal{D}'(\Omega)$. Suppose that $\partial_i T = 0$ for each $i = 1, \ldots, n$. Then there is a constant C such that

$$T(\phi) = C \int_\Omega \phi$$

for all $\phi \in \mathcal{D}(\Omega)$. (See Exercise 1.23 for 'connected' and Exercise 6.12 for a generalization.)

PROOF. By Theorem 6.10, T is a $C^1(\Omega)$-function, f, and $\partial f/\partial x^i = 0$. Application of 6.10(3) to f shows that f is constant. ■

6.12 MULTIPLICATION AND CONVOLUTION OF DISTRIBUTIONS BY C^∞-FUNCTIONS

A useful fact is that distributions can be multiplied by C^∞-functions. Consider T in $\mathcal{D}'(\Omega)$ and ψ in $C^\infty(\Omega)$. Define the product ψT by its action on $\phi \in \mathcal{D}(\Omega)$ as

$$(\psi T)(\phi) := T(\psi\phi) \tag{1}$$

for all $\phi \in \mathcal{D}(\Omega)$. That ψT is a distribution follows from the fact that the product $\psi\phi \in C_c^\infty(\Omega)$ if $\phi \in C_c^\infty(\Omega)$. Moreover, if $\phi^n \to \phi$ in $\mathcal{D}(\Omega)$, then

$\psi\phi^n \to \psi\phi$ in $\mathcal{D}(\Omega)$. To differentiate ψT we simply apply the product rule, namely

$$\partial_i(\psi T)(\phi) = \psi(\partial_i T)(\phi) + (\partial_i\psi)T(\phi), \tag{2}$$

which is easily verified from the basic definition 6.6(1) and Leibniz's differentiation formula $\partial_i(\psi\phi) = \phi\partial_i\psi + \psi\partial_i\phi$ for C^∞-functions.

Observe that when $T = T_f$ for some f in $L^1_{\text{loc}}(\Omega)$, then $\psi T = T_{\psi f}$. If, moreover, $f \in W^{1,p}_{\text{loc}}(\Omega)$, then $\psi f \in W^{1,p}_{\text{loc}}(\Omega)$ and (2) reads

$$\partial_i(f\psi)(x) = f(x)\partial_i\psi(x) + \psi(x)(\partial_i f)(x) \tag{3}$$

for almost every x. The same holds for $W^{1,p}(\Omega)$ and it also clearly extends to $W^{k,p}_{\text{loc}}(\Omega)$ and $W^{k,p}(\Omega)$.

The convolution of a distribution T with a $C^\infty_c(\mathbb{R}^n)$-function j is defined by

$$(j * T)(\phi) := T(j_R * \phi) = T\left(\int_{\mathbb{R}^n} j(y)\phi_{-y}\,\mathrm{d}y\right) \tag{4}$$

for all $\phi \in \mathcal{D}(\mathbb{R}^n)$, where $j_R(x) := j(-x)$. Since $j_R * \phi \in C^\infty_c(\mathbb{R}^n)$, $j * T$ makes sense and is in $\mathcal{D}'(\mathbb{R}^n)$. The reader can check that when T is a function, i.e., $T = T_f$, then, with this definition, $(j * T_f)(\phi) = T_{j*f}(\phi)$ where $(j * f)(x) = \int_{\mathbb{R}^n} j(x-y)f(y)\,\mathrm{d}y$ is the usual convolution.

► Note the requirement that j must have compact support.

6.13 THEOREM (Approximation of distributions by C^∞-functions)

Let $T \in \mathcal{D}'(\mathbb{R}^n)$ and let $j \in C^\infty_c(\mathbb{R}^n)$. Then there exists a function $t \in C^\infty(\mathbb{R}^n)$ (depending only on T and j) such that

$$(j * T)(\phi) = \int_{\mathbb{R}^n} t(y)\phi(y)\,\mathrm{d}y \tag{1}$$

for every $\phi \in \mathcal{D}(\mathbb{R}^n)$. If we further assume that $\int_{\mathbb{R}^n} j = 1$, and if we set $j_\varepsilon(x) = \varepsilon^{-n}j(x/\varepsilon)$ for $\varepsilon > 0$, then $j_\varepsilon * T$ converges to T in $\mathcal{D}'(\mathbb{R}^n)$ as $\varepsilon \to 0$.

PROOF. By definition we have that

$$(j * T)(\phi) := T(j_R * \phi) = T\left(\int_{\mathbb{R}^n} j(y - \cdot)\phi(y)\,\mathrm{d}y\right)$$

which, by Lemma 6.8, equals $\int_{\mathbb{R}^n} T(j(y-\cdot))\phi(y)\,\mathrm{d}y$. If we now define $t(y) := T(j(y-\cdot))$, then, by 6.8(1), $t \in C^\infty(\mathbb{R}^n)$. This proves (1). To verify the convergence of $j_\varepsilon * T$ to T, simply observe that

$$\begin{aligned}(j_\varepsilon * T)(\phi) &:= T\left(\int_{\mathbb{R}^n} j_\varepsilon(y)\phi_{-y}\,\mathrm{d}y\right)\\ &= \int_{\mathbb{R}^n} j_\varepsilon(y)T(\phi_{-y})\,\mathrm{d}y = \int_{\mathbb{R}^n} j(y)T(\phi_{-\varepsilon y})\,\mathrm{d}y\end{aligned} \tag{2}$$

by changing variables. It is clear that the last term in (2) tends to $T(\phi)$ since $T(\phi_{-y})$ is C^∞ as a function of y, and j has compact support. ■

● The **kernel** or **null-space** of a distribution $T \in \mathcal{D}'(\Omega)$ is defined by $\mathcal{N}_T = \{\phi \in \mathcal{D}(\Omega) : T(\phi) = 0\}$. It forms a closed linear subspace of $\mathcal{D}(\Omega)$. The following theorem about the intersection of kernels is useful in connection with Lagrange multipliers in the calculus of variations. (See Sect. 11.6.)

6.14 THEOREM (Linear dependence of distributions)

Let $S_1, \dots, S_N \in \mathcal{D}'(\Omega)$ be distributions. Suppose that $T \in \mathcal{D}'(\Omega)$ has the property that $T(\phi) = 0$ for all $\phi \in \bigcap_{i=1}^N \mathcal{N}_{S_i}$. Then there exist complex numbers $c_1, \dots, c_N$ such that

$$T = \sum_{i=1}^N c_i S_i. \tag{1}$$

PROOF. Without loss of generality it can be assumed that the S_i's are linearly independent. First, we show that there exist N fixed functions $u_1, \dots, u_N \in \mathcal{D}(\Omega)$ such that every $\phi \in \mathcal{D}(\Omega)$ can be written as

$$\phi = v + \sum_{i=1}^N \lambda_i(\phi)u_i \tag{2}$$

for some $\lambda_i(\phi) \in \mathbb{C}$, $i = 1, \dots, N$, and $v \in \bigcap_{i=1}^N \mathcal{N}_{S_i}$. To see this consider the set of vectors

$$V = \{\underline{S}(\phi) : \phi \in \mathcal{D}(\Omega)\}, \tag{3}$$

where $\underline{S}(\phi) = (S_1(\phi), \dots, S_N(\phi))$. It is obvious that V is a vector space of dimension N since the S_i's are linearly independent. Hence there exist functions $u_1, \dots, u_N \in \mathcal{D}(\Omega)$ such that $\underline{S}(u_1), \dots, \underline{S}(u_N)$ span V. Thus, the $N \times N$ matrix given by $M_{ij} = S_i(u_j)$ is invertible. With

$$\lambda_i(\phi) = \sum_{j=1}^{N} (M^{-1})_{ij} S_j(\phi), \tag{4}$$

it is easily seen that (2) holds.

Applying T to formula (2) yields (using $T(v) = 0$)

$$T(\phi) = \sum_{i,j=1}^{N} (M^{-1})_{ij} T(u_i) S_j(\phi),$$

which gives (1) with $c_i = \sum_{j=1}^{N} (M^{-1})_{ji} T(u_j)$. ∎

6.15 THEOREM ($C^\infty(\Omega)$ is 'dense' in $W^{1,p}_{\mathrm{loc}}(\Omega)$)

Let f be in $W^{1,p}_{\mathrm{loc}}(\Omega)$. For any open set $\mathcal{O}$ with the property that there exists a compact set $K \subset \Omega$ such that $\mathcal{O} \subset K \subset \Omega$, we can find a sequence $f^1, f^2, f^3, \dots \in C^\infty(\mathcal{O})$ such that

$$\|f - f^k\|_{L^p(\mathcal{O})} + \sum_i \|\partial_i f - \partial_i f^k\|_{L^p(\mathcal{O})} \to 0 \quad \text{as} \quad k \to \infty. \tag{1}$$

PROOF. For $\varepsilon > 0$ consider the function $j_\varepsilon * f$, where $j_\varepsilon(x) = \varepsilon^{-n} j(x/\varepsilon)$ and j is a C^∞-function with support in the unit ball centered at the origin with $\int_{\mathbb{R}^n} j(x)\,\mathrm{d}x = 1$. For any open set $\mathcal{O}$ with the properties stated above we have that $j_\varepsilon * f \in C^\infty(\mathcal{O})$ if ε is sufficiently small since, on $\mathcal{O}$,

$$D^\alpha (j_\varepsilon * f)(x) = \int_{\mathbb{R}^n} (D^\alpha j_\varepsilon)(x - y) f(y)\,\mathrm{d}y$$

for derivatives of any order α. Further, since $\mathcal{O} \subset K \subset \Omega$ with K compact, we can assume, by choosing ε small enough, that

$$\mathcal{O} + \mathrm{supp}\{j_\varepsilon\} := \{x + z : x \in \mathcal{O}, z \in \mathrm{supp}\{j_\varepsilon\}\} \subset K.$$

Thus, since

$$\partial_i \int_K j_\varepsilon(x - y) f(y)\,\mathrm{d}y = \int_K j_\varepsilon(x - y)(\partial_i f)(y)\,\mathrm{d}y,$$

and since f and $\partial_i f$ are in $L^p(K)$ for $i = 1, \dots, n$, (1) follows from Theorem 2.16 by choosing $\varepsilon = 1/k$ with k large enough. ∎

● The reader is invited to jump ahead for the moment and compare Theorem 6.15 for $p = 2$ with the much deeper Meyers–Serrin Theorem 7.6 (density of $C^\infty(\Omega)$ in $H^1(\Omega)$). The latter easily generalizes to $p \neq 2$, i.e., to $W^{1,p}(\Omega)$ and, in each case, implies 6.15. The important point is that if $f \in H^1(\Omega)$, then $\nabla f(x)$ can go to infinity as x goes to the boundary of Ω. Thus, convergence of the smooth functions f^k to f in the $H^1(\Omega)$-norm, as in 7.6, is not easy to achieve. Theorem 6.15 only requires convergence arbitrarily close to, but not up to, the boundary of Ω. The sequence f^k in 6.15 is allowed to depend on the open subset $\mathcal{O} \subset \Omega$. In contrast, in 7.6 the fixed sequence f^k must yield convergence in $H^1(\Omega)$. On the other hand, a function in $W^{1,2}_{\mathrm{loc}}(\Omega)$ need not be in $H^1(\Omega)$; it need not even be in $L^1(\Omega)$.

6.16 THEOREM (Chain rule)

Let $G : \mathbb{R}^N \to \mathbb{C}$ be a differentiable function with bounded and continuous derivatives. We denote it explicitly by $G(s_1, \dots, s_N)$. If

$$u(x) = (u_1(x), \dots, u_N(x))$$

denotes N functions in $W^{1,p}_{\mathrm{loc}}(\Omega)$, then the function $K : \Omega \to \mathbb{C}$ given by

$$K(x) = (G \circ u)(x) = G(u(x))$$

is in $W^{1,p}_{\mathrm{loc}}(\Omega)$ and

$$\frac{\partial}{\partial x_i} K = \sum_{k=1}^{N} \frac{\partial G}{\partial s_k}(u) \cdot \frac{\partial u_k}{\partial x_i} \tag{1}$$

in $\mathcal{D}'(\Omega)$.

If $u_1, \dots, u_N$ are in $W^{1,p}(\Omega)$, then K is also in $W^{1,p}(\Omega)$ and (1) *holds—provided we make the additional assumption, in case $|\Omega| = \infty$, that $G(0) = 0$.*

PROOF. It suffices to prove that $K \in W^{1,p}(\mathcal{O})$ and verify formula (1) for any open set $\mathcal{O}$ with the property that $\mathcal{O} \subset C \subset \Omega$, with C compact.

By Theorem 6.15 we can find a sequence of functions $\phi^m = (\phi_1^m, \dots, \phi_N^m)$ in $(C^\infty(\mathcal{O}))^N$ such that (with an obvious abuse of notation)

$$\|\phi^m - u\|_{W^{1,p}(\mathcal{O})} \to 0 \tag{2}$$

as $m \to \infty$. By passing to a subsequence we may assume that $\phi^m \to u$ pointwise a.e. and $\frac{\partial}{\partial x_i}\phi^m \to \frac{\partial}{\partial x_i}u$ pointwise a.e. for all $i = 1, \dots, n$. Set $K^m(x) = G(\phi^m(x))$. Since

$$\max_i \left| \frac{\partial G}{\partial s_i} \right| \leq M,$$

a simple application of the fundamental theorem of calculus and Hölder's inequality in $\mathbb{R}^N$ shows that for $s,t \in \mathbb{R}^N$

$$|G(s) - G(t)| \le MN^{1/p'} \left(\sum_{i=1}^{N} |s_i - t_i|^p \right)^{1/p} . \tag{3}$$

Here $1/p + 1/p' = 1$. Since $\mathcal{O} \subset C$ and G is bounded, K is in $L^p_{\mathrm{loc}}(\mathcal{O})$.

Next, for $\psi \in \mathcal{D}(\Omega)$

$$\begin{aligned} \int_\Omega \frac{\partial \psi}{\partial x_k} K^m \,\mathrm{d}x &= -\int_\Omega \psi \frac{\partial}{\partial x_k} K^m \,\mathrm{d}x \\ &= -\sum_{l=1}^{N} \int_\Omega \psi \frac{\partial G}{\partial s_l}(\phi^m) \frac{\partial}{\partial x_k} \phi_l^m \,\mathrm{d}x. \end{aligned} \tag{4}$$

In (4) the ordinary chain rule for C^1-functions has been used. Using (3) we find that

$$|K(x) - K^m(x)| \le MN^{1/p'} \left(\sum_{i=1}^{N} |u_i(x) - \phi_i^m(x)|^p \right)^{1/p} ,$$

which implies that $K^m \to K$ in $L^p(\mathcal{O})$, and therefore the left side of (4) tends to $\int_\Omega \frac{\partial \psi(x)}{\partial x_k} K(x)\,\mathrm{d}x$. Each term on the right side can be written as

$$\int_{\mathcal{O}} \psi \frac{\partial G}{\partial s_l}(\phi^m) \frac{\partial}{\partial x_k} u_l \,\mathrm{d}x + \int_{\mathcal{O}} \psi \frac{\partial G}{\partial s_l}(\phi^m) \left(\frac{\partial}{\partial x_k} \phi_l^m - \frac{\partial}{\partial x_k} u_l \right) \mathrm{d}x. \tag{5}$$

The first term tends to

$$\int_{\mathcal{O}} \psi \frac{\partial G}{\partial s_l}(u) \frac{\partial u_l}{\partial x_k} \,\mathrm{d}x$$

by dominated convergence and the second tends to zero since $\frac{\partial G}{\partial s_l}$ is uniformly bounded and $\frac{\partial \phi_l^m}{\partial x_k} - \frac{\partial u_l}{\partial x_k} \to 0$ in $L^p(\mathcal{O})$. Clearly $\frac{\partial G}{\partial s_l}(u) \frac{\partial u_l}{\partial x_k}$, which is a bounded function times an $L^p(\mathcal{O})$-function, is itself in $L^p(\mathcal{O})$.

To verify the second statement about $W^{1,p}(\Omega)$, note that $\partial G/\partial s_k$ is bounded for all $k = 1, 2, \ldots, N$ and, since $\nabla u_k \in L^p(\Omega)$, it follows from (1) that $\nabla K \in L^p(\Omega)$ also. The only thing to check is that K itself is in $L^p(\Omega)$. It follows from (3) that

$$|K(x)|^p \le A + B \sum_{k=1}^{N} |u_k(x)|^p, \tag{6}$$

where A and B are some constants. If $|\Omega| < \infty$, (6) implies that $K \in L^p(\Omega)$. If $|\Omega| = \infty$ we have to use the assumption $G(0) = 0$, which implies that we can take $A = 0$ in (6). Again, $K \in L^p(\Omega)$. ■

6.17 THEOREM (Derivative of the absolute value)

Let f be in $W^{1,p}(\Omega)$. Then the absolute value of f, denoted by $|f|$ and defined by $|f|(x) = |f(x)|$, is in $W^{1,p}(\Omega)$ with $\nabla|f|$ being the function

$$(\nabla|f|)(x) = \begin{cases} \frac{1}{|f|(x)}(R(x)\nabla R(x) + I(x)\nabla I(x)) & \text{if } f(x) \neq 0, \\ 0 & \text{if } f(x) = 0; \end{cases} \tag{1}$$

here $R(x)$ and $I(x)$ denote the real and imaginary parts of f. In particular, if f is real-valued,

$$(\nabla|f|)(x) = \begin{cases} \nabla f(x) & \text{if } f(x) > 0, \\ -\nabla f(x) & \text{if } f(x) < 0, \\ 0 & \text{if } f(x) = 0. \end{cases} \tag{2}$$

Thus $\left|\nabla|f|\right| \leq |\nabla f|$ a.e. if f is complex-valued and $\left|\nabla|f|\right| = |\nabla f|$ a.e. if f is real-valued.

PROOF. We follow [Gilbarg–Trudinger]. That $|f|$ is in $L^p(\Omega)$ follows from the definition of $\|f\|_p$. Further, since

$$\left|\frac{1}{|f|}(R\nabla R + I\nabla I)\right|^2 \leq (\nabla R)^2 + (\nabla I)^2 \tag{3}$$

pointwise, $\nabla|f|$ is also in $L^p(\Omega)$ once the claimed equality (1) is proved. Consider the function

$$G_\varepsilon(s_1, s_2) = \sqrt{\varepsilon^2 + s_1^2 + s_2^2} - \varepsilon. \tag{4}$$

Obviously $G_\varepsilon(0,0) = 0$ and

$$\left|\frac{\partial G_\varepsilon}{\partial s_i}\right| = \left|\frac{s_i}{\sqrt{\varepsilon^2 + s_1^2 + s_2^2}}\right| \leq 1. \tag{5}$$

Hence, by 6.16, the function $K_\varepsilon(x) = G_\varepsilon(R(x), I(x))$ is in $W^{1,p}(\Omega)$ and for all ϕ in $\mathcal{D}(\Omega)$

$$\begin{aligned} \int_\Omega \nabla\phi(x) K_\varepsilon(x)\, \mathrm{d}x &= -\int_\Omega \phi(x)\nabla K_\varepsilon(x)\, \mathrm{d}x \\ &= -\int_\Omega \phi(x) \frac{R(x)\nabla R(x) + I(x)\nabla I(x)}{\sqrt{\varepsilon^2 + |f(x)|^2}}\, \mathrm{d}x. \end{aligned} \tag{6}$$

Since $K_\varepsilon(x) \le |f(x)|$ and

$$\left|\frac{R(x)\nabla R(x) + I(x)\nabla I(x)}{\sqrt{\varepsilon^2 + |f(x)|^2}}\right| \le |\nabla f(x)|^2,$$

and since the two functions (4) and (5) converge pointwise to the claimed expressions as $\varepsilon \to 0$, the result follows by dominated convergence. ■

6.18 COROLLARY (Min and Max of $W^{1,p}$-functions are in $W^{1,p}$)

Let f and g be two real-valued functions in $W^{1,p}(\Omega)$. Then the minimum of $(f(x), g(x))$ and the maximum of $(f(x), g(x))$ are functions in $W^{1,p}(\Omega)$ and the gradients are given by

$$\nabla \max(f(x), g(x)) = \begin{cases} \nabla f(x) & \textit{when } f(x) > g(x), \\ \nabla g(x) & \textit{when } f(x) < g(x), \\ \nabla f(x) = \nabla g(x) & \textit{when } f(x) = g(x), \end{cases} \tag{1}$$

$$\nabla \min(f(x), g(x)) = \begin{cases} \nabla g(x) & \textit{when } f(x) > g(x), \\ \nabla f(x) & \textit{when } f(x) < g(x), \\ \nabla f(x) = \nabla g(x) & \textit{when } f(x) = g(x). \end{cases} \tag{2}$$

PROOF. That these two functions are in $W^{1,p}(\Omega)$ follows from the formulas

$$\min(f(x), g(x)) = \frac{1}{2}\big[(f(x) + g(x)) - |f(x) - g(x)|\big]$$

and

$$\max(f(x), g(x)) = \frac{1}{2}\big[(f(x) + g(x)) + |f(x) - g(x)|\big].$$

The formulas (1) and (2) follow immediately from Theorem 6.17 in the cases where $f(x) > g(x)$ or $f(x) < g(x)$. To understand the case $f(x) = g(x)$ consider

$$h(x) = (f(x) - g(x))_+ = \frac{1}{2}\big\{|f(x) - g(x)| + (f(x) - g(x))\big\}.$$

Obviously $|h|(x) = h(x)$, and hence by 6.17

$$\nabla h(x) = \nabla |h|(x) = 0 \quad \text{when } f(x) \le g(x).$$

But again by 6.17 $\nabla h(x) = \frac{1}{2}(\nabla(f - g))(x)$, when $f(x) = g(x)$ and hence

$$(\nabla f)(x) = (\nabla g)(x) \quad \text{when } f(x) = g(x),$$

which yields (1) and (2) in the case $f(x) = g(x)$. ■

It is an easy exercise to extend the above result to truncations of $W^{1,p}(\Omega)$-functions defined by

$$f_{<\alpha}(x) = \min(f(x), \alpha).$$

The gradient is then given by

$$(\nabla f_{<\alpha})(x) = \begin{cases} \nabla f(x) & \text{if } f(x) < \alpha, \\ 0 & \text{otherwise.} \end{cases}$$

Analogously, define

$$f_{>\alpha}(x) = \max(f(x), \alpha).$$

Then

$$(\nabla f_{>\alpha})(x) = \begin{cases} \nabla f(x) & \text{if } f(x) > \alpha, \\ 0 & \text{otherwise.} \end{cases}$$

Note that when Ω is unbounded $f_{<\alpha} \in W^{1,p}(\Omega)$ only if $\alpha \geq 0$, and $f_{>\alpha} \in W^{1,p}(\Omega)$ only if $\alpha \leq 0$.

The foregoing implies that if $u \in W^{1,1}_{\mathrm{loc}}(\Omega)$, if $\alpha \in \mathbb{R}$ and if $u(x) = \alpha$ on a set of positive measure in $\mathbb{R}^n$, then $(\nabla u)(x) = 0$ for almost every x in this set. This can be derived easily from 6.18. The following theorem, to be found in [Almgren–Lieb], generalizes this fact by replacing the single point $\alpha \in \mathbb{R}$ by a Borel set A of zero measure. Such sets need not be 'small', e.g., A could be all the rational numbers, and hence A could be dense in $\mathbb{R}$. Note that if f is a Borel measurable function, then $f^{-1}(A) := \{x \in \mathbb{R}^n : f(x) \in A\}$ is a Borel set, and hence is measurable. This follows from the statement in Sect. 1.5 and Exercise 1.3 that $x \mapsto \chi_A(f(x))$ is measurable.

6.19 THEOREM (Gradients vanish on the inverse of small sets)

Let $A \subset \mathbb{R}$ be a Borel set with zero Lebesgue measure and let $f : \Omega \to \mathbb{R}$ be in $W^{1,1}_{\mathrm{loc}}(\Omega)$. Let

$$B = f^{-1}(A) := \{x \in \Omega : f(x) \in A\} \subset \Omega.$$

Then $\nabla f(x) = 0$ for almost every $x \in B$.

PROOF. Our goal will be to establish the formula

$$\int_\Omega \phi(x)\chi_{\mathcal{O}}(f(x))\nabla f(x)\,\mathrm{d}x = -\int_\Omega \nabla\phi(x) G_{\mathcal{O}}(f(x))\,\mathrm{d}x \tag{1}$$

for each open set $\mathcal{O} \subset \mathbb{R}$. Here $\chi_\mathcal{O}$ is the characteristic function of $\mathcal{O}$ and $G_\mathcal{O}(t) = \int_0^t \chi_\mathcal{O}(s)\,\mathrm{d}s$. Equation (1) is just like the chain rule except that $G_\mathcal{O}$ is not in $C^1(\mathbb{R})$. Assuming (1) for the moment, we can conclude the proof of our theorem as follows. By the outer regularity of Lebesgue measure we can find a decreasing sequence $\mathcal{O}^1 \supset \mathcal{O}^2 \supset \mathcal{O}^3 \supset \cdots$ of open sets such that $A \subset \mathcal{O}^j$ for each j and $\mathcal{L}^1(\mathcal{O}^j) \to 0$ as $j \to \infty$. Thus $A \subset C := \bigcap_{j=1}^\infty \mathcal{O}^j$ (but it could happen that A is strictly smaller than C) and $\mathcal{L}^1(C) = 0$. By definition, $G_j(t) := G_{\mathcal{O}_j}(t)$ satisfies $|G_j(t)| \le \mathcal{L}^1(\mathcal{O}^j)$, and thus $G_j(t)$ goes uniformly to zero as $j \to \infty$. The right side of (1) (with $\mathcal{O}$ replaced by $\mathcal{O}^j$) therefore tends to zero as $j \to \infty$. On the other hand, $\chi_j := \chi_{\mathcal{O}_j}$ is bounded by 1, and $\chi_j(f(x)) \to \chi_{f^{-1}(C)}(x)$ for every $x \in \mathbb{R}^n$. By dominated convergence, the left side of (1) converges to $\int_\Omega \phi \chi_{f^{-1}(C)} \nabla f$, and this equals zero for every $\phi \in \mathcal{D}(\Omega)$. By the uniqueness of distributions, the function $\chi_{f^{-1}(C)}(x) \nabla f(x) = 0$ for almost every x, which is what we wished to prove.

It remains to prove (1). Observe that every open set $\mathcal{O} \subset \mathbb{R}$ is the union of countably many disjoint open intervals. (Why?) Thus $\mathcal{O} = \bigcup_{j=1}^\infty U_j$ with $U_j = (a_j, b_j)$. Since f is a function, $f^{-1}(U_j)$ is disjoint from $f^{-1}(U_k)$ when $j \ne k$. By the countable additivity of measure, therefore, it suffices to prove (1) when $\mathcal{O}$ is just one interval (a, b). We can easily find a sequence $\chi^1, \chi^2, \chi^3, \ldots$ of continuous functions such that $\chi^j(t) \to \chi_\mathcal{O}(t)$ for *every* $t \in \mathbb{R}$ and $0 \le \chi^j(t) \le 1$ for every $t \in \mathbb{R}$. The everywhere (not just almost everywhere) convergence is crucial and we leave the simple construction of $\{\chi^j\}$ to the reader. Then, with $G^j = \int_0^t \chi^j$, equation (1) is obtained by taking the limit $j \to \infty$ on both sides and using dominated convergence. The easy verification is again left to the reader. ∎

● An amusing—and useful—exercise in the computation of distributional derivatives is the computation of Green's functions. Let $y \in \mathbb{R}^n$, $n \ge 1$, and let $G_y : \mathbb{R}^n \to \mathbb{R}$ be defined by

$$\begin{aligned} G_y(x) &= -|\mathbb{S}^1|^{-1} \ln(|x-y|), && n = 2, \\ G_y(x) &= [(n-2)|\mathbb{S}^{n-1}|]^{-1} |x-y|^{2-n}, && n \ne 2, \end{aligned} \tag{2}$$

where $|\mathbb{S}^{n-1}|$ is the area of the unit sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$.

$$|\mathbb{S}^0| = 2, \quad |\mathbb{S}^1| = 2\pi, \quad |\mathbb{S}^2| = 4\pi, \quad |\mathbb{S}^{n-1}| = 2\pi^{n/2}/\Gamma(n/2).$$

These are the **Green's functions for Poisson's equation in $\mathbb{R}^n$**. Recall that the **Laplacian**, Δ, is defined by $\Delta := \sum_{i=1}^n \partial^2/\partial x_i^2$. The notation notwithstanding, $G_y(x)$ is actually *symmetric*, i.e., $G_y(x) = G_x(y)$.

6.20 THEOREM (Distributional Laplacian of Green's functions)

In the sense of distributions,

$$-\Delta G_y = \delta_y, \tag{1}$$

where δ_y is Dirac's delta measure at y (often written as $\delta(x-y)$).

PROOF. To prove (1) we can take $y=0$. We require

$$I := \int_{\mathbb{R}^n} (\Delta\phi) G_0 = -\phi(0)$$

when $\phi \in C_c^\infty(\mathbb{R}^n)$. Since $G_0 \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, it suffices to show that

$$-\phi(0) = \lim_{r\to 0} I(r),$$

where

$$I(r) := \int_{|x|>r} \Delta\phi(x) G_0(x)\,\mathrm{d}x.$$

We can also restrict the integration to $|x| < R$ for some R since ϕ has compact support. However, when $|x| > 0$, G_0 is infinitely differentiable and $\Delta G_0 = 0$. We can evaluate $I(r)$ by partial integration, and note that boundary integrals at $|x| = R$ vanish. Thus, denoting the set $\{x : r \le |x| \le R\}$ by A,

$$\begin{aligned} I(r) &= \int_A (\Delta\phi) G_0 = -\int_A \nabla\phi\cdot\nabla G_0 + \int_{|x|=r} G_0 \nabla\phi\cdot\nu \\ &= -\int_{|x|=r} \phi\nabla G_0\cdot\nu + \int_{|x|=r} G_0\nabla\phi\cdot\nu, \end{aligned} \tag{2}$$

where ν is the unit outward normal to A. On the sphere $|x| = r$, we have $\nabla G_0 \cdot \nu = |\mathbb{S}^{n-1}|^{-1} r^{-n+1}$, and therefore the penultimate integral in (2) is

$$-\int_{|x|=r} \phi\nabla G_0\cdot\nu = -|\mathbb{S}^{n-1}|^{-1}\int_{\mathbb{S}^{n-1}} \phi(r\omega)\,\mathrm{d}\omega,$$

which converges to $-\phi(0)$ as $r \to 0$, since ϕ is continuous. The last integral in (2) converges to zero as $r \to 0$ since $\nabla\phi\cdot\nu$ is bounded by some constant, while $\big||x|^{n-1}G_0(x)\big| < |x|^{1/2}$ for small $|x|$. Thus, (1) has been verified. ■

6.21 THEOREM (Solution of Poisson's equation)

Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, $n \geq 1$. Assume that for almost every x the function $y \mapsto G_y(x)f(y)$ is summable (here, G_y is Green's function given before 6.20) *and define the function $u : \mathbb{R}^n \to \mathbb{C}$ by*

$$u(x) = \int_{\mathbb{R}^n} G_y(x)f(y)\,\mathrm{d}y. \tag{1}$$

Then u satisfies:

$$u \in L^1_{\mathrm{loc}}(\mathbb{R}^n), \tag{2}$$

$$-\Delta u = f \quad \textit{in } \mathcal{D}'(\mathbb{R}^n). \tag{3}$$

Moreover, the function u has a distributional derivative that is a function; it is given, for almost every x, by

$$\partial_i u(x) = \int_{\mathbb{R}^n} (\partial G_y/\partial x_i)(x)f(y)\,\mathrm{d}y. \tag{4}$$

When $n = 3$, for example, the partial derivative is

$$(\partial G_y/\partial x_i)(x) = -\frac{1}{4\pi}|x-y|^{-3}(x_i - y_i). \tag{5}$$

REMARKS. (1) A trivial consequence of the theorem is that $\mathbb{R}^n$ can be replaced by any open set $\Omega \subset \mathbb{R}^n$. Suppose $f \in L^1_{\mathrm{loc}}(\Omega)$ and $y \mapsto G_y(x)f(y)$ is summable over Ω for almost every $x \in \Omega$. Then (see Exercises)

$$u(x) := \int_{\Omega} G_y(x)f(y)\,\mathrm{d}y \tag{6}$$

is in $L^1_{\mathrm{loc}}(\Omega)$ and satisfies

$$-\Delta u = f \quad \text{in } \mathcal{D}'(\Omega). \tag{7}$$

(2) The summability condition in Theorem 6.21 is equivalent to the condition that the function $w_n(y)f(y)$ is summable. Here

$$w_n(y) = \begin{cases} (1+|y|)^{2-n}, & n \geq 3, \\ \ln(1+|y|), & n = 2, \\ |y|, & n = 1. \end{cases} \tag{8}$$

The easy proof of this equivalence is left to the reader as an exercise. (It proceeds by decomposing the integral in (1) into a ball containing x, and its complement in $\mathbb{R}^n$. The contribution from the ball is easily shown to be finite for almost every x in the ball, by Fubini's theorem.)

(3) It is also obvious that any solution to equation (7) has the form $u+h$, where u is defined by (6) and where $\Delta h = 0$. Hence h is a harmonic function on Ω (see Sect. 9.3). Since harmonic functions are infinitely differentiable (Theorem 9.4), it follows that *every* solution to (7) is in $C^k(\Omega)$ if and only if $u \in C^k(\Omega)$.

PROOF. To prove (2) it suffices to prove that $I_B := \int_B |u| < \infty$ for each ball $B \subset \mathbb{R}^n$. Since $|u(x)| \le \int_{\mathbb{R}^n} |G_y(x) f(y)|\, dy$, we can use Fubini's theorem to conclude that

$$I_B \le \int_{\mathbb{R}^n} H_B(y)|f(y)|\, dy \quad \text{with} \quad H_B(y) = \int_B |G_y(x)|\, dx.$$

It is easy to verify (by using Newton's Theorem 9.7, for example) that if B has center x_0 and radius R, then $H_B(y) = |B||G_y(x_0)|$ for $|y - x_0| \ge R$ for $n \ne 2$ and $H_B(y) = |B||G_y(x_0)|$ when $|y - x_0| \ge R + 1$ when $n = 2$ (in order to keep the logarithm positive). Moreover, $H_B(y)$ is bounded when $|y - x_0| < R$. From this observation it follows easily that $I_B < \infty$. (Note: Fubini's theorem allows us to conclude both that u is a measurable function and that this function is in $L^1_{\mathrm{loc}}(\mathbb{R}^n)$.)

To verify (3) we have to show that

$$-\int u \Delta \phi = \int f \phi \tag{9}$$

for each $\phi \in C_c^\infty(\mathbb{R}^n)$. We can insert (1) into the left side of (9) and use Fubini's theorem to evaluate the double integral. But Theorem 6.20 states that $-\int_{\mathbb{R}^n} \Delta\phi(x) G_y(x)\, dx = \phi(y)$, and this proves (9).

To prove (4) we begin by verifying that the integral in (4) (call it $V_i(x)$) is well defined for almost every $x \in \mathbb{R}^n$. To see this note that $|(\partial G_y/\partial x_i)(x)|$ is bounded above by $c|x-y|^{1-n}$, which is in $L^1_{\mathrm{loc}}(\mathbb{R}^n)$. The finiteness of $V_i(x)$ follows as in Remark (2) above. Next, we have to show that

$$\int_{\mathbb{R}^n} \partial_i \phi(x) u(x)\, dx = -\int_{\mathbb{R}^n} \phi(x) V_i(x)\, dx \tag{10}$$

for all $\phi \in C_c^\infty(\mathbb{R}^n)$. Since the function $(x, y) \to (\partial_i \phi)(x) G_y(x) f(y)$ is $\mathbb{R}^n \times \mathbb{R}^n$ summable, we can use Fubini's theorem to equate the left side of (10) to

$$\int_{\mathbb{R}^n} \left\{ \int_{\mathbb{R}^n} (\partial_i \phi)(x) G_y(x)\, dx \right\} f(y)\, dy. \tag{11}$$

A limiting argument, combined with integration by parts, as in 6.20(2), shows that the inner integral in (11) is

$$-\int_{\mathbb{R}^n} \phi(x) \partial G_y/\partial x_i)(x)\, dx$$

for every $y \in \mathbb{R}^n$. Applying Fubini's theorem again, we arrive at (4). ■

● The next theorem may seem rather specialized, but it is useful in connection with the potential theory in Chapter 9. Its proof (which does not use Lebesgue measure) is an important exercise in measure theory. We shall leave a few small holes in our proof that we ask the reader to fill in as further exercises. Among other things, this theorem yields a construction of Lebesgue measure (Exercise 5).

6.22 THEOREM (Positive distributions are measures)

Let $\Omega \subset \mathbb{R}^n$ be open and let $T \in \mathcal{D}'(\Omega)$ be a positive distribution (meaning that $T(\phi) \geq 0$ for every $\phi \in \mathcal{D}(\Omega)$ such that $\phi(x) \geq 0$ for all x). We denote this fact by $T \geq 0$.

Our assertion is that there is then a unique, positive, regular Borel measure μ on Ω such that $\mu(K) < \infty$ for all compact $K \subset \Omega$ and such that for all $\phi \in \mathcal{D}(\Omega)$

$$T(\phi) = \int_\Omega \phi(x)\mu(\,\mathrm{d}x). \tag{1}$$

Conversely, any positive Borel measure with $\mu(K) < \infty$ for all compact $K \subset \Omega$ defines a positive distribution via (1).

REMARK. The representation (1) shows that a positive distribution can be extended from $C_c^\infty(\Omega)$-functions to a much larger class, namely the Borel measurable functions with compact support in Ω. This class is even larger than the continuous functions of compact support, $C_c(\Omega)$.

The theorem amounts to an extension, from $C_c(\Omega)$-functions to $C_c^\infty(\Omega)$-functions, of what is known as the **Riesz–Markov representation theorem**. See [Rudin, 1987].

PROOF. In the following, *all* sets are understood to be subsets of Ω. For a given open set $\mathcal{O}$ denote by $\mathcal{C}(\mathcal{O})$ the set of all functions $\phi \in C_c^\infty(\Omega)$ with $0 \leq \phi(x) \leq 1$ and $\operatorname{supp}\phi \subset \mathcal{O}$. Clearly, this set is not empty. (Why?) Next we define for any open set $\mathcal{O}$

$$\mu(\mathcal{O}) = \sup\{T(\phi) : \phi \in \mathcal{C}(\mathcal{O})\}. \tag{2}$$

For the empty set $\varnothing$ we set $\mu(\varnothing) = 0$. The nonnegative set function μ has the following properties:

(i) $\mu(\mathcal{O}_1) \leq \mu(\mathcal{O}_2)$ if $\mathcal{O}_1 \subset \mathcal{O}_2$,

(ii) $\mu(\mathcal{O}_1 \cup \mathcal{O}_2) \leq \mu(\mathcal{O}_1) + \mu(\mathcal{O}_2)$,

(iii) $\mu\left(\bigcup_{i=1}^\infty \mathcal{O}_i\right) \leq \sum_{i=1}^\infty \mu(\mathcal{O}_i)$ for every countable family of *open* sets $\mathcal{O}_i$.

Property (i) is evident. The second property follows from the following fact (F) whose proof we leave as an exercise for the reader:

(F) For any compact set K and open sets $\mathcal{O}_1, \mathcal{O}_2$ such that $K \subset \mathcal{O}_1 \cup \mathcal{O}_2$ there exist functions ϕ_1 and ϕ_2, both C^∞ in a neighborhood $\mathcal{O}$ of K, such that $\phi_1(x) + \phi_2(x) = 1$ for $x \in K$ and $\phi \cdot \phi_1 \in C_c^\infty(\mathcal{O}_1), \phi \cdot \phi_2 \in C_c^\infty(\mathcal{O}_2)$ for any function $\phi \in C_c^\infty(\mathcal{O})$.

Thus, any $\phi \in \mathcal{C}(\mathcal{O}_1 \cup \mathcal{O}_2)$ can be written as $\phi_1 + \phi_2$ with $\phi_1 \in \mathcal{C}(\mathcal{O}_1)$ and $\phi_2 \in \mathcal{C}(\mathcal{O}_2)$. Hence $T(\phi) = T(\phi_1) + T(\phi_2) \le \mu(\mathcal{O}_1) + \mu(\mathcal{O}_2)$ and property (ii) follows. By induction we find that

$$\mu\left(\bigcup_{i=1}^{m} \mathcal{O}_i\right) \le \sum_{i=1}^{m} \mu(\mathcal{O}_i).$$

To see property (iii) pick $\phi \in \mathcal{C}\left(\bigcup_{i=1}^\infty \mathcal{O}_i\right)$. Since ϕ has compact support, we have that $\phi \in \mathcal{C}\left(\bigcup_{i\in I} \mathcal{O}_i\right)$ where I is a finite subset of the natural numbers. Hence, by the above,

$$T(\phi) \le \mu\left(\bigcup_{i\in I} \mathcal{O}_i\right) \le \sum_{i\in I} \mu(\mathcal{O}_i) \le \sum_{i=1}^{\infty} \mu(\mathcal{O}_i),$$

which yields property (iii).

For *every* set A define

$$\mu(A) = \inf\{\mu(\mathcal{O}) : \mathcal{O} \text{ open}, A \subset \mathcal{O}\}. \tag{3}$$

The reader should not be confused by this definition. We have defined a set function, μ, that measures *all* subsets of Ω, but only for a special subcollection will this function be a measure, i.e., be countably additive. This set function μ will now be shown to have the properties of an **outer measure**, as defined in Theorem 1.15 (constructing a measure from an outer measure), i.e.,

(a) $\mu(\varnothing) = 0$,

(b) $\mu(A) \le \mu(B)$ if $A \subset B$,

(c) $\mu\left(\bigcup_{i=1}^\infty A_i\right) \le \sum_{i=1}^\infty \mu(A_i)$ for *every* countable collection of sets $A_1, A_2, \dots$.

The first two properties are evident. To prove (c) pick open sets $\mathcal{O}_1, \mathcal{O}_2, \dots$ with $A_i \subset \mathcal{O}_i$ and $\mu(\mathcal{O}_i) \le \mu(A_i) + 2^{-i}\varepsilon$ for $i = 1, 2, \dots$. Now

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \le \mu\left(\bigcup_{i=1}^{\infty} \mathcal{O}_i\right) \le \sum_{i=1}^{\infty} \mu(\mathcal{O}_i)$$

by (b) and (iii), and hence

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i) + \varepsilon,$$

which yields (c) since ε is arbitrary. By Theorem 1.15 the sets A such that $\mu(E) = \mu(E \cap A) + \mu(E \cap A^c)$ for every set E form a sigma-algebra, Σ, on which μ is countably additive.

Next we have to show that all open sets are measurable, i.e., we have to show that for any set E and any open set $\mathcal{O}$

$$\mu(E) \geq \mu(E \cap \mathcal{O}) + \mu(E \cap \mathcal{O}^c). \tag{4}$$

The reverse inequality is obvious. First we prove (4) in the case where E is itself open; call it V.

Pick any function $\phi \in \mathcal{C}(V \cap \mathcal{O})$ such that $T(\phi) \geq \mu(V \cap \mathcal{O}) - \varepsilon/2$. Since $K := \operatorname{supp} \phi$ is compact, its complement, U, is open and contains $\mathcal{O}^c$. Pick $\psi \in \mathcal{C}(U \cap V)$ such that $T(\psi) \geq \mu(U \cap V) - \varepsilon/2$. Certainly

$$\begin{aligned}\mu(V) &\geq T(\phi) + T(\psi) \geq \mu(V \cap \mathcal{O}) + \mu(V \cap U) - \varepsilon \\ &\geq \mu(V \cap \mathcal{O}) + \mu(V \cap \mathcal{O}^c) - \varepsilon\end{aligned}$$

and since ε is arbitrary this proves (4) in the case where E is an open set. If E is arbitrary we have for any open set V with $E \subset V$ that $E \cap \mathcal{O} \subset V \cap \mathcal{O}$, $E \cap \mathcal{O}^c \subset V \cap \mathcal{O}^c$, and hence $\mu(V) \geq \mu(E \cap \mathcal{O}) + \mu(E \cap \mathcal{O}^c)$. This proves (4). Thus we have shown that the sigma-algebra Σ contains all open sets and hence contains the Borel sigma-algebra. Hence the measure μ is a Borel measure.

By construction, this measure is outer regular (see (3) above). We show next that it is inner regular, i.e., for any measurable set A

$$\mu(A) = \sup\{\mu(K) : K \subset A,\ K \text{ compact}\}. \tag{5}$$

First we have to establish that compact sets have finite measure. We claim that for K compact

$$\mu(K) = \inf\{T(\psi) : \psi \in C_c^\infty(\Omega),\ \psi(x) = 1 \text{ for } x \in K, \psi \geq 0\}. \tag{6}$$

The set on the right side is not empty. Indeed for K compact and $K \subset \mathcal{O}$ open there exists a C_c^∞-function ψ such that $\operatorname{supp} \psi \subset \mathcal{O}$ and $\psi := 1$ on K. (Such a ψ was constructed in Exercise 1.15 *without the aid of Lebesgue measure.*)

Now (6) follows from the following fact which we ask the reader to prove as an exercise: $\mu(K) \le T(\psi)$ for any $\psi \in C_c^\infty(\Omega)$ with $\psi \equiv 1$ on K and $\psi \ge 0$. Given this fact, choose $\varepsilon > 0$ and choose $\mathcal{O}$ open such $\mu(K) \ge \mu(\mathcal{O}) - \varepsilon$. Also pick $\psi \in C_c^\infty(\Omega)$ with $\operatorname{supp}\psi \subset \mathcal{O}$ and $\psi \equiv 1$ on K. Then $\mu(K) \le T(\psi) \le \mu(\mathcal{O}) \le \mu(K) + \varepsilon$. This proves (6).

It is easy to see that for $\varepsilon > 0$ and every measurable set A with $\mu(A) < \infty$ there exists an open set $\mathcal{O}$ with $A \subset \mathcal{O}$ and $\mu(\mathcal{O} \sim A) < \varepsilon$. Using the fact that Ω is a countable union of closed balls, the above holds for any measurable set, i.e., even if A does not have finite measure. We ask the reader to prove this.

For $\varepsilon > 0$ and a measurable set A we can find $\mathcal{O}$ with $A^c \subset \mathcal{O}$ such that $\mu(\mathcal{O} \sim (A^c)) < \varepsilon$. But

$$\mathcal{O} \sim (A^c) = \mathcal{O} \cap A = A \sim (\mathcal{O}^c)$$

and $\mathcal{O}^c$ is closed. Thus for any measurable set A and $\varepsilon > 0$ one can find a closed set $\mathcal{C}$ such that $\mathcal{C} \subset A$ and $\mu(A \sim \mathcal{C}) < \varepsilon$. Since any closed set in $\mathbb{R}^n$ is a countable union of compact sets, the inner regularity is proven.

Next we prove the representation theorem. The integral $\int_\Omega \phi(x)\mu(\,dx)$ defines a distribution R on $\mathcal{D}(\Omega)$. Our aim is to show that $T(\phi) = R(\phi)$ for all $\phi \in C_c^\infty(\Omega)$. Because $\phi = \phi_1 - \phi_2$ with $\phi_{1,2} \ge 0$ and $\phi_{1,2} \in C_c^\infty(\Omega)$ (as Exercise 1.15 shows), it suffices to prove this with the additional restriction that $\phi \ge 0$. As usual, if $\phi \ge 0$,

$$R(\phi) = \int_0^\infty m(a)\,\mathrm{d}a = \lim_{n\to\infty} \frac{1}{n} \sum_{j\ge 1} m(j/n) \tag{7}$$

where $m(a) = \mu(\{x : \phi(x) > a\})$. The integral in (7) is a Riemann integral; it always makes sense for nonnegative monotone functions (like m) and it always equals the rightmost expression in (7). For each n, the sum in (7) has only finitely many terms, since ϕ is bounded.

For n fixed we define compact sets K_j, $j = 0, 1, 2, \ldots$, by setting $K_0 = \operatorname{supp}\phi$ and $K_j = \{x : \phi(x) \ge j/n\}$ for $j \ge 1$. Similarly, denote by O^j the open sets $\{x : \phi(x) > j/n\}$ for $j = 1, 2, \ldots$. Let χ_j and χ^j denote the characteristic functions of K_j and O^j. Then, as is easily seen,

$$\frac{1}{n}\sum_{j\ge 1} \chi^j < \phi < \frac{1}{n}\sum_{j\ge 0} \chi_j .$$

Since ϕ has compact support, all the sets have finite measure by (6).

For $\varepsilon > 0$ and $j = 0, 1, \ldots$ pick U_j open such that $K_j \subset U_j$ and $\mu(U_j) \le \mu(K_j) + \varepsilon$. Next pick $\psi_j \in C_c^\infty(\mathbb{R}^n)$ such that $\psi_j \equiv 1$ on K_j and $\operatorname{supp}\psi_j \subset$

U_j. We have shown above that such a function exists. Obviously $\phi \leq \frac{1}{n}\sum_{j\geq 0}\psi_j$ and hence

$$T(\phi) \leq \frac{1}{n}\sum_{j\geq 0} T(\psi_j) \leq \frac{1}{n}\sum_{j\geq 0}\mu(U_j) \leq \frac{1}{n}\sum_{j\geq 0}\mu(K_j) + \varepsilon.$$

By the inner regularity we can find, for every open set $\mathcal{O}^j$ of finite measure, a compact set $C^j \subset \mathcal{O}_j$ such that $\mu(C^j) \geq \mu(\mathcal{O}^j) - \varepsilon$ and, in the same fashion as above, conclude that $T(\phi) \geq \frac{1}{n}\sum_{j\geq 1}\mu(\mathcal{O}^j) - \varepsilon$. Since $\varepsilon > 0$ is arbitrary,

$$\frac{1}{n}\sum_{j\geq 1}\mu(\mathcal{O}^j) \leq T(\phi) \leq \frac{1}{n}\sum_{j\geq 0}\mu(K_j).$$

By noting that $K_j \subset \mathcal{O}^{j-1}$ for $j \geq 1$, we have

$$\frac{1}{n}\sum_{j\geq 1} m(j/n) \leq T(\phi) \leq \frac{1}{n}\sum_{j\geq 1} m(j/n) + \frac{2}{n}\mu(K_0),$$

which proves the representation theorem. The uniqueness part is left to the reader. ■

● In Sects. 6.19–6.21 the Green's function G_y for $-\Delta$ was exhibited. As a further important exercise in distribution theory, which will be needed in Sect. 12.4, we next discuss the Green's function for $-\Delta + \mu^2$ with $\mu > 0$. It satisfies (*cf.* 6.20(1))

$$(-\Delta + \mu^2)G_y^\mu = \delta_y. \tag{8}$$

This function is called the **Yukawa potential**, at least for $n = 3$, and played an important role in the theory of elementary particles (mesons), for which H. Yukawa won a Nobel prize. As in the case of G_y, the function G_y^μ is really a function of $x - y$ (in fact, a function only of $|x - y|$) which we call $G^\mu(x - y)$. In the following, G_0 is G_y with $y = 0$.

6.23 THEOREM (Yukawa potential)

For each $n \geq 1$ and $\mu > 0$ there is a function G_y^μ that satisfies 6.22(8) *in $\mathcal{D}'(\mathbb{R}^n)$ and is given by*

$$G_y^\mu(x) = G^\mu(x - y), \tag{1}$$

$$G^\mu(x) = \int_0^\infty (4\pi t)^{-n/2}\exp\left\{-\frac{|x|^2}{4t} - \mu^2 t\right\} \mathrm{d}t. \tag{2}$$

The function G^μ, which (2) shows is symmetric decreasing, satisfies

(i) $G^\mu(x) > 0$ *for all* x.

(ii) $\int_{\mathbb{R}^n} G^\mu(x)\mathrm{d}x = \mu^{-2}$.

(iii) *As* $x \to 0$,

$$G^\mu(x) \to 1/2\mu \quad \textit{for } n = 1, \tag{3}$$

$$\frac{G^\mu(x)}{G_0(x)} \to 1 \quad \textit{for } n > 1. \tag{4}$$

(iv) $-\left[\log G^\mu(x)\right]/(\mu|x|) \to 1$ *as* $|x| \to \infty$.

From (3), (4) *we see that* G^μ *is in* $L^q(\mathbb{R}^n)$ *if* $1 \le q \le \infty$ $(n = 1)$, $1 \le q < \infty$ $(n = 2)$, *and* $1 \le q < n/(n-2)$ $(n \ge 3)$. *Also,* $G^\mu \in L_w^{n/(n-2)}(\mathbb{R}^n)$ $(n \ge 3)$. (*See Sect.* 4.3 *for* L_w^q.)

(v) *If* $f \in L^p(\mathbb{R}^n)$, *for some* $1 \le p \le \infty$, *then*

$$u(x) = \int_{\mathbb{R}^n} G_y^\mu(x) f(y)\mathrm{d}y \tag{5}$$

is in $L^r(\mathbb{R}^n)$ *and satisfies*

$$(-\Delta + \mu^2)u = f \quad \textit{in } \mathcal{D}'(\mathbb{R}^n), \tag{6}$$

with $p \le r \le \infty$ $(n = 1)$; $p \le r \le \infty$ *when* $p > 1$ *and* $1 \le r < \infty$ *when* $p = 1$ $(n = 2)$; *and* $p \le r \le np/(n-2p)$ *when* $1 < p < n/2$, $p \le r \le \infty$ *when* $p \ge n/2$, *and* $1 \le r < n/(n-2)$ *when* $p = 1$ $(n \ge 3)$. *Moreover,* (5) *is the unique solution to* (6) *with the property that it is in* $L^r(\mathbb{R}^n)$ *for some* $r \ge 1$.

(vi) *The Fourier transform of* G^μ *is*

$$\widehat{G^\mu}(p) = \left([2\pi p]^2 + \mu^2\right)^{-1}. \tag{7}$$

REMARKS. (1) The function $(4\pi t)^{-n/2} \exp\left\{-|x|^2/4t\right\}$ is the 'heat kernel', which is discussed further in Sect. 7.9.

(2) The following are examples in one and three dimensions, respectively.

$$G^\mu(x) = \frac{1}{2\mu} \exp\{-\mu|x|\}, \qquad n = 1,$$

$$G^\mu(x) = \frac{1}{4\pi|x|} \exp\{-\mu|x|\}, \qquad n = 3. \tag{8}$$

PROOF. It is extremely easy to verify that the integral in (2) is finite for all $x \neq 0$ and that (i) and (ii) are true. To prove 6.22(8) we have to show that

$$\int_{\mathbb{R}^n} G^\mu(x)(-\Delta + \mu^2)\phi(x)\mathrm{d}x = \phi(0) \tag{9}$$

for all $\phi \in C_c^\infty(\mathbb{R}^n)$. We substitute (2) in (9), do the x-integration before the t-integration, and then integrate by parts in x. For $t > 0$,

$$(-\Delta+\mu^2)(4\pi t)^{-n/2}\exp\left\{-\frac{|x|^2}{4t} - \mu^2 t\right\} = -\frac{\partial}{\partial t}(4\pi t)^{-n/2}\exp\left\{-\frac{|x|^2}{4t} - \mu^2 t\right\}.$$

Thus, the left side of (9) is

$$\begin{aligned}
&- \lim_{\varepsilon\to 0}\int_\varepsilon^\infty \left[\int_{\mathbb{R}^n} \phi(x)\frac{\partial}{\partial t}(4\pi t)^{-n/2}\exp\left\{-\frac{|x|^2}{4t} - \mu^2 t\right\}\mathrm{d}x\right]\mathrm{d}t \\
&\quad = - \lim_{\varepsilon\to 0}\int_\varepsilon^\infty \frac{\partial}{\partial t}\left[\int_{\mathbb{R}^n} \phi(x)(4\pi t)^{-n/2}\exp\left\{-\frac{|x|^2}{4t} - \mu^2 t\right\}\mathrm{d}x\right]\mathrm{d}t \\
&\quad = + \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n} \phi(x)(4\pi\varepsilon)^{-n/2}\exp\left\{-\frac{|x|^2}{4\varepsilon}\right\}\mathrm{d}x \\
&\quad = \quad \phi(0)
\end{aligned}$$

since $(4\pi\varepsilon)^{-n/2}\exp\{-|x|^2/4\varepsilon\}$ converges in $\mathcal{D}'$ to δ_0 as $\varepsilon \to 0$ (check these steps!). Thus (9) is proved, and hence 6.22(8).

The proof of (6) is even easier than the proof of Theorem 6.21(1–3). Again, Fubini's theorem plus integration by parts does the job. The r-summability of u follows from Young's (or the Hardy-Littlewood-Sobolev) inequality and the fact that $G^\mu \in L^1(\mathbb{R}^n)$. Since $u \in L^p(\mathbb{R}^n)$, and hence vanishes at infinity, the uniqueness assertion after (6) is equivalent to the assertion that the only solution to $(-\Delta + \mu^2)u = 0$ in some $L^r(\mathbb{R}^n)$ is $u \equiv 0$. This will be proved in Sect. 9.11.

We leave items (iii) and (iv) as exercises. They are evidently true for $n = 1$ and 3.

Item (vi) can be proved either by direct computation from (2) or else by multiplying 6.22(8) by $\exp\{-2\pi i(p, x)\}$ and integrating. ■

● In Sect. 6.7 we defined the weak convergence of a sequence of functions $f^1, f^2, \ldots$ in $W^{1,p}(\Omega)$ with $1 \leq p \leq \infty$ by the statement that f^j converges to f if and only if f^j and each of its n partial derivatives $\partial_i f^j$ converges in the usual sense of weak $L^p(\Omega)$ convergence. While such a notion of convergence makes sense, the reader may wonder what the dual space of $W^{1,p}(\Omega)$ actually is and whether the notion of convergence, as defined in Sect. 6.7, agrees with the fundamental definition in 2.9(6). The answer is 'yes', as the next theorem shows.

The question can be restated as follows. Let $g_0, g_1, \dots, g_n$ be $n+1$ functions in $L^{p'}(\Omega)$ and, for all $f \in W^{1,p}(\Omega)$, set

$$L(f) = \int_\Omega g_0 f + \sum_{i=1}^n \int_\Omega g_i \partial_i f, \tag{9}$$

which, obviously, defines a continuous linear functional on $W^{1,p}(\Omega)$. If *every* continuous linear functional has this form, then we have identified the dual of $W^{1,p}(\Omega)$ and the Sect. 6.7 definition agrees with the standard one.

Two things are worth noting. One is that, with L given, the right side of (9) may not be unique because f and ∇f are not independent. For example, if the g_i are C_c^∞ functions, then the $n+1$-tuple $g_0, g_1, \dots, g_n$ gives the same L as $g_0 - \sum_i \partial_i g_i, 0, \dots, 0$. Another thing to note is that (9) really defines a continuous linear functional on the vector space consisting of $n+1$ copies of $L^p(\Omega)$ (which can be written as $\times^{(n+1)} L^p(\Omega)$ or as $L^p(\Omega; \mathbb{C}^{(n+1)})$). In this bigger space a continuous linear functional defines the g_i uniquely. In other words, $W^{1,p}(\Omega)$ can be viewed as a closed subspace of $\times^{(n+1)} L^p(\Omega)$ and our question is whether every continuous linear functional on $W^{1,p}(\Omega)$ can be extended to a continuous linear functional on the bigger space. The Hahn-Banach theorem guarantees this, but we give a proof below for $1 \le p < \infty$ that imitates our proof in Sect. 2.14.

6.24 THEOREM (The dual of $W^{1,p}(\Omega)$)

Every continuous linear functional L on $W^{1,p}(\Omega)$ $(1 \le p < \infty)$ can be written in the form 6.23(9) *above for some choice of $g_0, g_1, \dots, g_n$ in $L^{p'}(\Omega)$.*

PROOF. Let $\mathcal{H} = \times^{(n+1)} L^p(\Omega)$, i.e., an element h of $\mathcal{H}$ is a collection of $n+1$ functions $h = (h_0, \dots, h_n)$, each in $L^p(\Omega)$. Likewise, we can consider the space $\Xi = \Omega \times \{0, 1, \dots, n\}$, i.e., a point in Ξ is a pair $y = (x, j)$ with $x \in \Omega$ and $j \in \{0, 1, 2, \dots, n\}$. We equip Ξ with the obvious product sigma-algebra, an element of which can be viewed as a collection of $n+1$ elements of the Borel sigma-algebra on Ω, i.e., $A = (A_0, \dots, A_n)$ with $A_j \subset \Omega$. Finally, we put the obvious measure on A, namely $\mu(A) = \sum_j \mathcal{L}^n(A_j)$. Thus, $\mathcal{H} = L^p(\Xi, \mathrm{d}\mu)$ and $\|h\|_p^p = \sum_{j=0}^n \|h_j\|_p^p$.

Think of $W^{1,p}(\Omega)$ as a subset of $\mathcal{H} = L^p(\Xi, \mathrm{d}\mu)$, i.e., $f \in W^{1,p}(\Omega)$ is mapped into $\widetilde{f} = (f, \partial_1 f, \dots, \partial_n f)$. With this correspondence, we have that $\widetilde{W}$, the imbedding of $W^{1,p}(\Omega)$ in $\mathcal{H}$, is a closed subset and it is also a subspace (i.e., it is a linear space). Likewise, the kernel of L, namely $K = \{f \in W^{1,p}(\Omega) : L(f) = 0\} \subset W^{1,p}(\Omega)$, defines a closed (why?) subspace of $\mathcal{H}$ (which we call $\widetilde{K}$). L corresponds to a linear functional $\widetilde{L}$ on $\widetilde{W}$ whose kernel is $\widetilde{K}$.

Consider, first, $1 < p < \infty$. Lemma 2.8 (Projection on convex sets) is valid and (assuming that $L \neq 0$) we can find an $\widetilde{f} \in \widetilde{W}$ so that $\widetilde{L}(\widetilde{f}) \neq 0$, i.e., $\widetilde{f} \notin \widetilde{K}$. Then, by 2.8(2), there is a function $\widetilde{Y} \in L^{p'}(\Xi, \mathrm{d}\mu)$ such that $\mathrm{Re} \int_\Xi (\widetilde{g} - \widetilde{h})\widetilde{Y} \leq 0$ for some $\widetilde{h} \in \widetilde{K}$ and for all $\widetilde{g} \in \widetilde{K}$. Since $\widetilde{K}$ is a linear space (over the complex numbers) this implies that $\int_\Xi (\widetilde{g} - \widetilde{h})\widetilde{Y} = 0$ for all $\widetilde{g} \in \widetilde{K}$ (why?), which, in turn, implies that $\int_\Xi \widetilde{f}\widetilde{Y} = 0$ for all $\widetilde{f} \in \widetilde{K}$ (why?).

The proof is now finished in the manner of Theorem 2.14. For $p = 1$ the second part of Theorem 2.14 also extends to the present case. ■

Exercises for Chapter 6

1. Fill in the details in the last paragraph of the proof of Theorem 6.19, i.e., (a) Construct the sequence χ^j that converges everywhere to $\chi_{(\text{interval})}$; (b) Complete the dominated convergence argument.

2. Verify the summability condition in Remark (2), equation (8) of Theorem 6.21.

3. Prove fact (F) in Theorem 6.22.

4. Prove that for K compact, $\mu(K)$ (defined in 6.22(3)) satisfies $\mu(K) \leq T(\psi)$ for $\psi \in C_c^\infty(\Omega)$ and $\psi \equiv 1$ on K.

5. Notice that the proof of Theorem 6.22 (and its antecedents) used only the Riemann integral and not the Lebesgue integral. Use the conclusion of Theorem 6.22 to prove the existence of Lebesgue measure. See Sect. 1.2.

6. Prove that the distributional derivative of a monotone nondecreasing function on $\mathbb{R}$ is a Borel measure.

7. Let $\mathcal{N}_T$ be the null-space of a distribution, T. Show that there is a function $\phi_0 \in \mathcal{D}$ so that every element $\phi \in \mathcal{D}$ can be written as $\phi = \lambda\phi_0 + \psi$ with $\psi \in \mathcal{N}_T$ and $\lambda \in \mathbb{C}$. One says that the null-space $\mathcal{N}_T$ has 'codimension one'.

8. Show that a function f is in $W^{1,\infty}(\Omega)$ if and only if $f = g$ a.e. where g is a function that is bounded and **Lipschitz continuous** on Ω, i.e., there exists a constant C such that

$$|g(x) - g(y)| \leq C|x - y| \quad \text{for all } x, y \in \Omega.$$

9. Verify Remark (1) in Theorem 6.21 that in this theorem $\mathbb{R}^n$ can be replaced by any open subset of $\mathbb{R}^n$.

10. Consider the function $f(x) = |x|^{-n}$ on $\mathbb{R}^n$. Although this function is not in $L^1_{loc}(\mathbb{R}^n)$, it is defined as a distribution for test functions on $\mathbb{R}^n$ that vanish at the origin, by

$$T_f(\phi) = \int_{\mathbb{R}^n} |x|^{-n}\phi(x)dx.$$

   a) Show that there is a distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ that agrees with T_f for functions that vanish at the origin. Give an explicit formula for one such T.
   b) Characterize all such T's. Theorem 6.14 may be helpful here.

11. Functions in $W^{1,p}(\mathbb{R}^n)$ can be very rough for $n \geq 2$ and $p \leq n$.
   a) Construct a spherically symmetric function in $W^{1,p}(\mathbb{R}^n)$ that diverges to infinity as $x \to 0$.
   b) Use this to construct a function in $W^{1,p}(\mathbb{R}^n)$ that diverges to infinity at every rational point in the unit cube.

   ► *Hint.* Write the function in b) as a sum over the rationals. How do you prove that the sum converges to a $W^{1,p}(\mathbb{R}^n)$ function?

12. *Generalization of* 6.11. Show that if $\Omega \subset \mathbb{R}^n$ is connected and if $T \in \mathcal{D}'(\Omega)$ has the property that $D^\alpha T = 0$ for all $|\alpha| = m + 1$, then T is a multinomial of degree at most m, i.e., $T = \sum_{|\alpha| \leq m} C_\alpha x^\alpha$.

13. Prove 6.23(4) in the case $n > 2$.

14. Prove 6.23(4) in the case $n = 2$.

15. Prove 6.23, item (iv).

16. Carry out the explicit calculation of the Fourier transform of the Yukawa potential from 6.23(2), as indicated in the last line of the proof of Theorem 6.23. Likewise, justify the alternative derivation, i.e., by multiplying 6.22(8) by $\exp\{-2\pi i(p,x)\}$ and integrating. The point is that $\exp\{-2\pi i(p,x)\}$ does not have compact support and so is not in $\mathcal{D}(\mathbb{R}^n)$.

17. Verify formulas 6.23(8) for the Yukawa potential.

18. The proof of Theorem 6.24 is a bit subtle. Write up a clear proof of the "why's" that appear there.

19. Using the definition of weak convergence for $W^{1,p}(\Omega)$ (see Sect. 6.7) formulate and prove the analog of Theorem 2.18 (bounded sequences have weak limits) for $W^{1,p}(\Omega)$.

20. **Hanner's inequality for** $W^{m,p}$. Show that Theorem 2.5 holds for $W^{m,p}(\Omega)$ in place of $L^p(\Omega)$.

21. For $n \geq 2$ and $p \leq n$ construct a nonzero function f in $W^{1,p}(\mathbb{R}^n)$ with the property that, for every rational point y, $\lim_{x\to y} f(x)$ exists and equals zero. (Can an $f \in C^0(\mathbb{R}^n)$ have this property?)