Mathematical
Surveys
and
Monographs
Volume 207

Fokker–Planck–Kolmogorov Equations

Vladimir I. Bogachev
Nicolai V. Krylov
Michael Röckner
Stanislav V. Shaposhnikov

American Mathematical Society

Mathematical
Surveys
and
Monographs

Volume 207

Fokker–Planck–Kolmogorov Equations

Vladimir I. Bogachev
Nicolai V. Krylov
Michael Röckner
Stanislav V. Shaposhnikov

American Mathematical Society
Providence, Rhode Island

2010 *Mathematics Subject Classification.* Primary 35-02, 35J15, 35K10, 60J35, 60J60.

For additional information and updates on this book, visit
www.ams.org/bookpages/surv-207

Library of Congress Cataloging-in-Publication Data

Fokker–Planck–Kolmogorov equations /Vladimir I. Bogachev, Nicolai V. Krylov, Michael Röckner, Stanislav V. Shaposhnikov.

pages cm. — (Mathematical surveys and monographs ; volume 207)

Includes bibliographical references and index.

ISBN 978-1-4704-2558-6 (alk. paper)

1. Fokker–Planck equation. 2. Stochastic differential equations. I. Bogachev, V. I. (Vladimir Igorevich), 1961- II. Krylov, N. V. (Nicolai Vladimirovich). III. Röckner, Michael, 1956- IV. Shaposhnikov, Stanislav V.

QA274.23.F65 2015
515′.353—dc23

2015024922

10 9 8 7 6 5 4 3 2 1 20 19 18 17 16 15

Fokker–Planck–Kolmogorov Equations

Preface

This book gives a systematic presentation of the theory of Fokker–Planck–Kolmogorov equations, which are second order elliptic and parabolic equations for measures. This direction goes back to Kolmogorov's works [**527**], [**528**], [**529**] and a number of earlier works in the physics literature by Fokker [**377**], Smoluchowski [**863**], Planck [**781**], and Chapman [**235**]. One of our principal objects is the elliptic operator of the form

$$L_{A,b}f = \operatorname{trace}(AD^2 f) + \langle b, \nabla f\rangle, \quad f \in C_0^\infty(\Omega),$$

where $A = (a^{ij})$ is a mapping on a domain $\Omega \subset \mathbb{R}^d$ with values in the space of nonnegative symmetric linear operators on $\mathbb{R}^d$ and $b = (b^i)$ is a vector field on Ω. In coordinate form, $L_{A,b}$ is given by the expression

$$L_{A,b}f = a^{ij}\partial_{x_i}\partial_{x_j} f + b^i \partial_{x_i} f,$$

where we always assume that the summation is taken over all repeated indices.

With this operator $L_{A,b}$, we associate the weak elliptic equation

$$L_{A,b}^*\mu = 0 \tag{1}$$

for Borel measures on Ω, which is understood in the following sense:

$$\int_\Omega L_{A,b}f\, d\mu = 0 \quad \forall\, f \in C_0^\infty(\Omega), \tag{2}$$

where we assume that b^i, $a^{ij} \in L^1_{\rm loc}(\mu)$. If μ has a density ϱ with respect to Lebesgue measure, then ϱ is sometimes called "an adjoint solution" and the equation is called "an equation in double divergence form". We use the above term "weak elliptic equation for measures". The corresponding equation for the density ϱ is

$$\partial_{x_i}\partial_{x_j}(a^{ij}\varrho) - \partial_{x_i}(b^i\varrho) = 0.$$

If $A = \mathrm{I}$, we obtain the equation $\Delta\varrho - \operatorname{div}(\varrho b) = 0$.

Similarly, one can consider parabolic operators and parabolic Fokker–Planck–Kolmogorov equations for measures on $\Omega\times(0,T)$ of the type

$$\partial_t\mu = L_{A,b}^*\mu.$$

The corresponding equations for densities are

$$\partial_t\varrho(x,t) = \partial_{x_i}\partial_{x_j}\bigl(a^{ij}(x,t)\varrho(x,t)\bigr) - \partial_{x_i}\bigl(b^i(x,t)\varrho(x,t)\bigr), \tag{3}$$

and if we also have an initial distribution μ_0 in a suitable sense, then we arrive at the Cauchy problem for the Fokker–Planck–Kolmogorov equation. However, it is crucial that a priori Fokker–Planck–Kolmogorov equations are equations for measures, not for functions; this becomes relevant when the coefficients are singular or degenerate and, in particular, in the infinite-dimensional case, where no Lebesgue measure exists. It is also important that equation (1) is meaningful under very broad assumptions about A and b: only their local integrability with respect to the regarded solution μ is needed. These coefficients may be quite singular with respect to Lebesgue measure even if the solution admits a smooth density. For example, for an arbitrary infinitely differentiable probability density ϱ on $\mathbb{R}^d$, the measure $\mu = \varrho\,dx$ satisfies the above equation with $A = \mathrm{I}$ and $b = \nabla\varrho/\varrho$, where we set $\nabla\varrho(x)/\varrho(x) = 0$ whenever $\varrho(x) = 0$. This is obvious from the integration by parts formula

$$\int_{\mathbb{R}^d} [\Delta f + \langle \nabla\varrho/\varrho, \nabla f\rangle]\varrho\,dx = \int_{\mathbb{R}^d} \varrho\Delta f + \int_{\mathbb{R}^d} \langle \nabla\varrho, \nabla f\rangle\,dx = 0.$$

Since ϱ may vanish on an arbitrary proper closed subset of $\mathbb{R}^d$, the vector field b can fail to be locally integrable with respect to Lebesgue measure, but it is locally integrable with respect to μ. Also note that in general our solutions need not be more regular than the coefficients (unlike in the case of usual elliptic equations). For example, if $d = 1$ and $b = 0$, then for an arbitrary positive probability density ϱ, the measure $\mu = \varrho\,dx$ satisfies the equation $L^*_{A,0}\mu = 0$ with $A = \varrho^{-1}$.

In this general setting, a study of weak elliptic equations for measures on finite- and infinite-dimensional spaces was initiated in the 1990s in the papers of the first three authors. Actually, the infinite-dimensional case was even a starting point, which was motivated by investigations of infinite-dimensional diffusion processes and other applications in infinite-dimensional stochastic analysis (developed in particular in the works of Albeverio, Høegh-Krohn **[21]** as well as A.I. Kirillov **[511]**–**[516]**). It was realized in the course of these investigations that even infinite-dimensional equations with very nice coefficients often require results on finite-dimensional equations with quite general coefficients. For example, we shall see in Chapter 10 that the finite-dimensional projections μ_n of a measure μ satisfying an elliptic equation on an infinite-dimensional space satisfy elliptic equations whose coefficients are the conditional expectations of the original coefficients with respect to the σ-algebras generated by the corresponding projection operators. As a result, even for smooth infinite-dimensional coefficients, the only information about their conditional expectations is related to their integrability with respect to μ_n, not with respect to Lebesgue measure; in particular, no local boundedness is given.

The theory of elliptic and parabolic equations for measures is now a rapidly growing area with deep and interesting connections to many directions in real analysis, partial differential equations, and stochastic analysis. Let us briefly describe the probabilistic picture behind our analytic framework. Suppose that $\xi = (\xi_t^x)_{t\geqslant 0}$ is a diffusion process in $\mathbb{R}^d$ governed by the stochastic differential equation

$$d\xi_t^x = \sigma(\xi_t^x)dW_t + b(\xi_t^x)dt, \quad \xi_0 = x.$$

The basic concepts related to this equation are recalled in § 1.3. The generator of the transition semigroup $\{T_t\}_{t\geqslant 0}$ has the form $L_{A,b}$, where $A = \sigma\sigma^*/2$. The matrix $A = (a^{ij})$ in the operator $L_{A,b}$ will be called the *diffusion matrix* or *diffusion coefficient*; this differs from the standard form of the diffusion generator by the absence of the factor $1/2$ in front of the second order derivatives, but is more convenient when one deals with equations. The vector field b is called the *drift coefficient* or just the *drift*. The transition probabilities of ξ satisfy the corresponding parabolic equation. Any invariant probability measure μ of ξ (if such exists) satisfies (1), where μ is called invariant for $\{T_t\}_{t\geqslant 0}$ if the following identity holds:

$$\int_{\mathbb{R}^d} T_t f\, d\mu = \int_{\mathbb{R}^d} f\, d\mu \quad \forall f \in C_b(\mathbb{R}^d). \tag{4}$$

Measures satisfying (1) are called *infinitesimally invariant*, because this equation has deep connections with invariance with respect to the corresponding operator semigroups. More precisely, if there is an invariant probability measure μ, then $\{T_t\}_{t\geqslant 0}$ extends to $L^1(\mu)$ and is strongly continuous. Let L be the corresponding generator with domain $D(L)$. Then (4) is equivalent to the equality

$$\int_{\mathbb{R}^d} Lf\, d\mu = 0 \quad \forall f \in D(L).$$

Under reasonable assumptions about A and b, the generator of the semigroup associated with the diffusion governed by the indicated stochastic equation coincides with $L_{A,b}$ on $C_0^\infty(\mathbb{R}^d)$. As we shall see, invariance of the measure in the sense of (4) is not the same as (2). The point is that the class $C_0^\infty(\mathbb{R}^d)$ may be much smaller than $D(L)$. What is important is that the equation is meaningful and can have solutions under assumptions that are much weaker than those needed for the existence of a diffusion, so that this equation can be investigated without any assumptions about the existence of semigroups. On the other hand, there exist very interesting and fruitful relations between equations (2) and (4). For example, if A and b are both Lipschitz and if A is nondegenerate, they are equivalent.

Letting $P(x,t,\,\cdot\,)$ be the corresponding transition probabilities (the distributions of ξ_t^x), the semigroup property reads

$$P(x,t+s,B) = \int_{\mathbb{R}^d} P(u,s,B)\, P(x,t,du), \tag{5}$$

or in the case where there exist densities $p(x,t,y)$,

$$p(x,t+s,y) = \int_{\mathbb{R}^d} p(u,s,y)\, p(x,t,u)\, du.$$

Identity (5) is called the Smoluchowski equation or the Chapman–Kolmogorov equation. In his seminal paper [**527**] Kolmogorov posed the following problems: find conditions for the existence and uniqueness of solutions to the Cauchy problem for (3) and investigate when (5) holds for these solutions. Now, 80 years later, these problems are still not completely solved. However, considerable progress has been achieved; results obtained and some related open problems are discussed in this book.

We shall consider the following problems.

1) Regularity of solutions of equation (2), for example, the existence of densities with respect to Lebesgue measure, the continuity and smoothness of these densities, and certain related estimates (such as L^2-estimates for logarithmic gradients of solutions). In particular, we shall see in Chapter 1 that the measure μ is always absolutely continuous with respect to Lebesgue measure on the set $\{\det A > 0\}$ and has a continuous density from the Sobolev class $W^{p,1}_{loc}$ with $p > d$ provided that the diffusion coefficients a^{ij} are in this class, $|b| \in L^p_{loc}(dx)$ or $|b| \in L^p_{loc}(\mu)$, and the matrix A is positive definite. Global properties of solutions of equations with unbounded coefficients are studied in Chapter 3, where certain global upper and lower estimates for the densities are obtained. We shall also obtain analogous results for parabolic equations in Chapters 6–8.

2) Existence of solutions to elliptic equation (2) and existence of invariant measures in the sense of (4) as well as relations between these two concepts are the subjects of Chapter 2 and Chapter 5. In particular, we shall see in Chapter 5 that under rather general assumptions, for a given probability measure μ satisfying our elliptic equation (2), one can construct a strongly continuous Markov semigroup $\{T_t^\mu\}_{t\geqslant 0}$ on $L^1(\mu)$ such that μ is $\{T_t^\mu\}_{t\geqslant 0}$-invariant and the generator of $\{T_t^\mu\}_{t\geqslant 0}$ coincides with $L_{A,b}$ on $C_0^\infty(\mathbb{R}^d)$. For this, an easy to verify condition is the existence of a Lyapunov function for $L_{A,b}$. In the general case (without any additional assumptions), a bit less is true, namely, μ is only subinvariant for $\{T_t\}_{t\geqslant 0}$. We shall see examples where this really occurs, i.e., where μ is not invariant. Existence of solutions to parabolic equations is addressed in Chapter 6.

3) Various uniqueness problems are considered in Chapters 4 and 5; in particular, uniqueness of invariant measures in the sense of (4) and uniqueness of solutions to (2) in the class of all probability measures. Related interesting problems concern uniqueness of associated semigroups $\{T_t^\mu\}_{t\geqslant 0}$ and the essential self-adjointness of the operator $L_{A,b}$ on $C_0^\infty(\mathbb{R}^d)$ in the case when it is symmetric. Parabolic analogues are considered in Chapter 9.

First, we concentrate on the elliptic case, to which Chapters 1–5 are devoted. In Chapters 6–9 similar problems are studied for parabolic equations; however, parabolic equations appear already in Chapter 5 in relation to semigroups generated by elliptic operators. Chapter 10 is devoted to a brief discussion of infinite-dimensional analogues of the problems listed in 1)–3). The results obtained so far in the infinite-dimensional setting apply to various particular situations, although they cover many concrete examples arising in applications such as stochastic partial differential equations, infinite particle systems, Gibbs measures, and so on. The main purpose of Chapter 10 is to give applications of finite-dimensional results and to demonstrate the universality of certain ideas, methods, and techniques. Finally, in Chapters 2, 6, and 9 we discuss degenerate equations and nonlinear equations for measures; important examples of such equations are Vlasov-type equations. We made some effort to minimize dependencies between the chapters; the proofs of a number of fundamental results are rather difficult and can be omitted without any loss of understanding of the rest.

Every chapter opens with some synopsis mentioning the chief problems and results discussed. The last section of each chapter includes some complementary subsections (the numbers in brackets within these internal contents refer to the corresponding page numbers) and also brief historical and bibliographic comments and exercises. In the Bibliography each item is provided with indication of all pages where it is cited. The Subject Index also includes special notations used.

Acknowledgments. The material in this book has been discussed with many colleagues; we are especially grateful to S. Albeverio, L. Ambrosio, A.-B. Cruzeiro, G. Da Prato, A. Eberle, D. Elworthy, F. Flandoli, M. Fukushima, B. Goldys, I. A. Ibragimov, A. I. Kirillov, A. V. Kolesnikov, V. A. Kondratiev, Yu. G. Kondratiev, V. V. Kozlov, G. Leha, V. A. Liskevich, A. Lunardi, P. Malliavin, I. Marshall, G. Metafune, J. van Neerven, D. Pallara, E. Pardoux, S. I. Pokhozhaev, Yu. V. Prohorov, E. V. Radkevich, A. Rhandi, G. Ritter, T. A. Shaposhnikova, I. Shigekawa, A. Shirikyan, A. N. Shiryaev, Z. Sobol, W. Stannat, N. S. Trudinger, A. Yu. Veretennikov, F. Y. Wang, J. Zabczyk, M. Zakai, T. S. Zhang, and V. V. Zhikov. The Russian Science Foundation grant 14-11-00196 is gratefully acknowledged.

Contents

CHAPTER 1

Stationary Fokker–Planck–Kolmogorov Equations

In this chapter we introduce principal objects related to elliptic equations for measures, an important example of which is the stationary Fokker–Planck–Kolmogorov equation for invariant probabilities of diffusion processes. Although our approach is purely analytic, some concepts related to diffusion processes are explained. Our principal problems are explained and in the rest of this chapter we present the results on existence of densities of solutions to elliptic equations for measures and their local properties such as Sobolev regularity. Thus, it turns out that under broad assumptions our equations for measures are reduced to equations for their densities. However, these equations have a rather special form, which leads to certain properties of solutions that are different from the case of general second order equations.

1.1. Background material

Throughout we shall use the following standard notation. The inner product and norm in $\mathbb{R}^d$ are denoted by $\langle \cdot , \cdot \rangle$ and $| \cdot |$, respectively. The diameter of a set Ω is $\operatorname{diam}\Omega = \sup_{x,y\in\Omega} |x-y|$. The open ball of radius r centered at a is denoted by $U(a,r)$ or $U_r(a)$. The unit matrix is denoted by I. The trace of an operator A is denoted by $\operatorname{tr} A$. The inequality $A \leqslant B$ for operators on $\mathbb{R}^d$ means the estimate $\langle Ah, h\rangle \leqslant \langle Bh, h\rangle$, where $h \in \mathbb{R}^d$, for their quadratic forms. In expressions like $a^{ij}x_iy_i$ and b^ix_i the standard summation rule with respect to repeated indices will be meant. Set $u^+ = \max(u, 0)$, $u^- = -\min(u, 0)$, i.e., $u = u^+ - u^-$.

Throughout "*positive*" means "larger than zero".

The class of all smooth functions with compact support lying in an open set $\Omega \subset \mathbb{R}^d$ is denoted by $C_0^\infty(\Omega)$; the classes of the type $C_b^k(\Omega)$, $C_0^k(\Omega)$ of functions with k continuous derivatives etc. are defined similarly; $C(\Omega)$ and $C_b(\Omega)$ are the classes of continuous and bounded continuous functions. The class of functions whose derivatives up to order k have continuous extensions to the closure of Ω is denoted by the symbol $C^k(\overline{\Omega})$. The support of a function f, i.e., the closure of the set $\{f \neq 0\}$, is denoted by $\operatorname{supp} f$.

A measure μ on a σ-algebra $\mathcal{A}$ in a space Ω is a function $\mu\colon \mathcal{A} \to \mathbb{R}^1$ that is countably additive: $\mu(A) = \sum_{n=1}^{\infty} \mu(A_n)$ whenever $A_n \in \mathcal{A}$ are pairwise disjoint and their union is A. Such a measure is automatically bounded and can be written as $\mu = \mu^+ - \mu^-$, where the measures μ^+ and μ^-, called the positive and negative parts of μ, respectively, are nonnegative and concentrated on disjoint sets $\Omega^+ \in \mathcal{A}$ and $\Omega^+ \in \mathcal{A}$, respectively, such that $\Omega = \Omega^+ \cup \Omega^-$. The measure

$$|\mu| := \mu^+ + \mu^-$$

is called the total variation of the measure μ. The variational norm or the variation of the measure μ is defined by the equality $\|\mu\| := |\mu|(\Omega)$. Let $\mathcal{M}(\Omega)$ be the class of all bounded measures on $(\Omega, \mathcal{A})$ and $\mathcal{P}(\Omega)$ the class of all *probability measures* on $(\Omega, \mathcal{A})$ (i.e., measures $\mu \geqslant 0$ with $\mu(\Omega) = 1$). The simplest probability measure is Dirac's measure δ_a at a point $a \in \Omega$, it equals 1 at the point a and 0 at the complement of a. If $\mu \geqslant 0$ and $\mu(\Omega) \leqslant 1$, then μ is a *subprobability* measure.

It is useful to admit also unbounded measures with values in $[0, +\infty]$ defined similarly. Such a measure is called σ-finite if the space is the union of countably many parts of finite measure. The classical Lebesgue measure on $\mathbb{R}^d$ provides an example. Lebesgue measure of a set Ω will be occasionally denoted by $|\Omega|$. For most of the results discussed below we need only the classical Lebesgue measure and other measures absolutely continuous with respect to it (see below).

We recall that the Borel σ-algebra $\mathcal{B}(E)$ is the smallest σ-algebra containing all open sets of a given space E. The term "a Borel measure μ" will normally mean a finite (possibly signed) countably additive measure on the σ-algebra of Borel sets; cases where infinite measures (say, locally finite measures) are considered will always be specified, except for Lebesgue measure. A Borel measure μ on a subset in $\mathbb{R}^d$ is called locally finite if every point has a neighborhood of finite $|\mu|$-measure.

A finite Borel measure μ on a topological space X is called Radon if, for every Borel set $B \subset X$ and every $\varepsilon > 0$, there is a compact set $K_\varepsilon \subset B$ such that $|\mu|(B \backslash K_\varepsilon) < \varepsilon$. By Ulam's theorem, on all complete separable metric spaces all finite Borel measures are Radon. Throughout we consider only Borel measures.

The integral of a function f with respect to a measure μ over a set A is denoted by the symbols

$$\int_A f(x)\,\mu(dx), \quad \int_A f\,d\mu.$$

For a nonnegative measure μ and $p \in [1, \infty)$, the symbols $L^p(\mu)$ or $L^p(\Omega, \mu)$ denote the space of equivalence classes of μ-measurable functions f such that the function $|f|^p$ is integrable. This space is equipped with the standard norm

$$\|f\|_p := \|f\|_{L^p(\mu)} := \left(\int_\Omega |f|^p\,d\mu\right)^{1/p}.$$

The notation $L^p(\Omega)$ always refers to the classical Lebesgue measure; sometimes we write $L^p(\Omega, dx)$ in order to stress this.

Let $L^\infty(\mu)$ denote the space of equivalence classes of bounded μ-measurable functions equipped with the norm $\|f\|_\infty := \inf_{g \sim f} \sup_x |g(x)|$.

A measure μ is called *separable* if $L^1(\mu)$ is separable (and then so are also all spaces $L^p(\mu)$ for $p < \infty$).

As usual, for $p \in [1, +\infty)$ we set

$$p' := \frac{p}{p-1}.$$

The classical *Hölder inequality* says that

$$\int_\Omega |fg|\,d\mu \leqslant \|f\|_p \|g\|_{p'}, \quad f \in L^p(\mu),\ g \in L^{p'}(\mu).$$

It yields the *generalized Hölder inequality*

$$\int_\Omega |f_1 \cdots f_n|\,d\mu \leqslant \|f_1\|_{p_1} \cdots \|f_n\|_{p_n}, \quad f_i \in L^{p_i}(\mu),\ p_1^{-1} + \cdots + p_n^{-1} = 1.$$

In addition, if $pq \geqslant p+q$, $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then by Hölder's inequality $fg \in L^r(\mu)$ and

(1.1.1) $$\|fg\|_r \leqslant \|f\|_p \|g\|_q \quad \text{if } r = pq/(p+q).$$

The integrability of a function with respect to a signed measure μ is understood as its integrability with respect to the total variation $|\mu|$ of the measure μ; the corresponding classes will be denoted by $L^p(\mu)$ or $L^p(|\mu|)$ and by $L^p(U,\mu)$ or $L^p(U,|\mu|)$ in the case where μ is restricted to a fixed set $U \subset \Omega$.

For a Radon measure μ, the class $L^1_{\text{loc}}(\mu)$ consists of all functions that are integrable with respect to μ on all compact sets.

Let I_A denote the indicator function of the set A, i.e., $I_A(x) = 1$ if $x \in A$, $I_A(x) = 0$ if $x \notin A$.

A measure ν on a σ-algebra $\mathcal{A}$ is called absolutely continuous with respect to a measure μ on the same σ-algebra if the equality $|\mu|(A) = 0$ implies the equality $\nu(A) = 0$; notation: $\nu \ll \mu$. By the Radon–Nikodym theorem this is equivalent to the existence of a function ϱ integrable with respect to $|\mu|$ such that

$$\nu(A) = \int_A \varrho(x)\,\mu(dx), \quad A \in \mathcal{A}.$$

The function ϱ is called the *density* (the *Radon–Nikodym density*) of the measure ν with respect to the measure μ and is denoted by the symbol $d\nu/d\mu$. It is customary to write also

$$\nu = \varrho \cdot \mu \quad \text{or} \quad \nu = \varrho\mu.$$

If $\nu \ll \mu$ and $\mu \ll \nu$, then the measures ν and μ are *equivalent*; notation: $\nu \sim \mu$. This is equivalent to the following property: $\nu \ll \mu$ and $d\nu/d\mu \neq 0$ $|\mu|$-almost everywhere. The term "almost everywhere" is shortened as μ-a.e. (for a signed measure μ, the term "μ-a.e." is understood as "$|\mu|$-a.e.").

A sequence of Borel measures μ_n converges weakly to a Borel measure μ if for every bounded continuous function f one has

$$\lim_{n\to\infty} \int f\,d\mu_n = \int f\,d\mu.$$

A family $\mathcal{M}$ of Radon measures on a topological space X is called uniformly tight if for each $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset X$ such that $|\mu|(X \backslash K_\varepsilon) < \varepsilon$ for all measures $\mu \in \mathcal{M}$. According to the Prohorov theorem, a bounded family of Borel measures on a complete separable metric space is uniformly tight precisely when every infinite sequence in it contains a weakly convergent subsequence (see Bogachev [**125**, Chapter 8]).

Given an open set $\Omega \subset \mathbb{R}^d$ and $p \in [1,+\infty)$, we denote by $W^{p,1}(\Omega)$ or $H^{p,1}(\Omega)$ the Sobolev class of all functions $f \in L^p(\Omega)$ whose generalized partial derivatives $\partial_{x_i} f$ are in $L^p(\Omega)$. A generalized (or Sobolev) derivative is defined by the equality (the integration by parts formula)

$$\int_U \varphi \partial_{x_i} f\,dx = -\int_U f \partial_{x_i} \varphi\,dx, \quad \varphi \in C_0^\infty(\Omega).$$

This space is equipped with the Sobolev norm

$$\|f\|_{p,1} := \|f\|_p + \sum_{i=1}^d \|\partial_{x_i} f\|_p.$$

We also use higher-order Sobolev classes $W^{p,k}(\Omega) = H^{p,k}(\Omega)$ with $k \in \mathbb{N}$, consisting of functions whose Sobolev partial derivatives up to order k are in $L^p(\Omega)$ and equipped with naturally defined norms $\|f\|_{p,k}$, and fractional Sobolev spaces $H^{p,r}(\Omega)$ with noninteger r (the definition is given in § 1.8); the notation with the letter H will normally be used in the case of fractional or parabolic Sobolev classes.

The class $W^{\infty,k}(\Omega)$ consists of functions with bounded Sobolev derivatives up to order k; for example, $W^{\infty,1}(\Omega)$ is the class of bounded Lipschitzian functions. Let $\nabla f = (\partial_{x_1} f, \ldots, \partial_{x_d} f)$.

The class $W_0^{p,k}(\Omega)$ is defined as the closure of $C_0^\infty(\Omega)$ in $W^{p,k}(\Omega)$.

The space $C^{0,\delta}(\Omega)$ consists of Hölder continuous of order $\delta \in (0,1)$ functions f on Ω with finite norm

$$\|f\|_{C^{0,\delta}} := \sup_{x\in\Omega} |f(x)| + \sup_{x,y\in\Omega, x\neq y} |f(x) - f(y)|/|x-y|^\delta.$$

Symbols like $W^{p,1}_{loc}(\mathbb{R}^d)$, $W^{p,1}_{loc}(\Omega)$, $L^p_{loc}(\Omega, \mu)$ denote the classes of functions f such that ζf belongs to the corresponding class without the lower index "loc" for every $\zeta \in C_0^\infty(\mathbb{R}^d)$ or $\zeta \in C_0^\infty(\Omega)$, respectively.

Let $W^{p,-1}(\mathbb{R}^d)$ denote the dual space to $W^{p',1}(\mathbb{R}^d)$ with $p' = p/(p-1)$, $p > 1$.

Let us define weighted Sobolev spaces or classes. Let a nonnegative measure μ on $\mathbb{R}^d$ be given by a locally integrable density ϱ with respect to Lebesgue measure. The class $W^{p,k}(\mu)$ is defined as the completion of $C_0^\infty(\mathbb{R}^d)$ with respect to the Sobolev norm $\|\cdot\|_{p,k,\mu}$ defined similarly to $\|\cdot\|_{p,k}$, but with the measure μ in place of Lebesgue measure. If the density ϱ is continuous and positive, then $W^{p,k}(\mu)$ coincides with the class of functions $f \in W^{p,k}_{loc}(\mathbb{R}^d)$ with $\|f\|_{p,k,\mu} < \infty$. Weighted classes are used below only in a very few places, mostly the classes $W^{p,1}(\mu)$, moreover, in such cases the measure μ has some additional properties, for example, possessing a continuous positive density or a weakly differentiable density, so that the weighted Sobolev classes are well-defined (see, e.g., Bogachev [**126**, § 2.6]).

We shall need the class $W^{d+,1}_{loc}(\Omega)$ consisting of all functions f on an open set Ω such that the restriction of f to each ball U with closure in Ω belongs to $W^{p_U,1}(U)$ for some $p_U > d$, and also the class $L^{d+}_{loc}(\Omega)$ defined similarly.

In the theory of Sobolev spaces and its applications a very important role is played by the following Sobolev embedding theorem (the case $p = 1$ is called the Gagliardo–Nirenberg embedding theorem).

1.1.1. Theorem. (i) *If $p > d$ or $p = d = 1$, then one has the embedding*

$$W^{p,1}(\mathbb{R}^d) \subset C_b(\mathbb{R}^d) = C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d).$$

Moreover, there exists a number $C(p,d) > 0$ such that

$$\|f\|_\infty \leqslant C(p,d)\|f\|_{p,1}, \quad f \in W^{p,1}(\mathbb{R}^d). \tag{1.1.2}$$

(ii) *If $p \in [1,d)$, then $W^{p,1}(\mathbb{R}^d) \subset L^{dp/(d-p)}(\mathbb{R}^d)$, hence $L^q(\mathbb{R}^d) \subset W^{p',-1}(\mathbb{R}^d)$ if $q = dp/(dp+p-d)$, $p > 1$. Moreover, there is a number $C(p,d) > 0$ such that*

$$\|f\|_{dp/(d-p)} \leqslant C(p,d)\|f\|_{p,1}, \quad f \in W^{p,1}(\mathbb{R}^d). \tag{1.1.3}$$

For any bounded domain Ω with Lipschitzian boundary analogous embeddings hold with some number $C(p,d,\Omega)$.

Note that $p' = qd/(d-q)$ in (ii). Actually in place of (1.1.3) the inequality

$$\|f\|_{dp/(d-p)} \leqslant C(p,d)\big\||\nabla f|\big\|_p \quad \forall\, f \in W^{p,1}(\mathbb{R}^d) \tag{1.1.4}$$

holds, which for $p = 1$ is called the Galiardo–Nirenberg inequality; it shows that an integrable function on $\mathbb{R}^d$ with an integrable gradient belongs in fact to the class $L^{d/(d-1)}(\mathbb{R}^d)$, hence also to all $L^p(\mathbb{R}^d)$ with $1 \leqslant p \leqslant d/(d-1)$. For functions with support in the unit ball U we obtain the inequality

$$\|f\|_p \leqslant C(p)\big\||\nabla f|\big\|_p, \quad f \in W_0^{p,1}(U). \tag{1.1.5}$$

Note also the *Poincaré inequality*

$$\|f - f_U\|_p \leqslant C(p)\big\||\nabla f|\big\|_p, \quad f \in W^{p,1}(U),\ f_U = \int_U f\,dx. \tag{1.1.6}$$

A function from the class $W^{d,1}(\mathbb{R}^d)$ need not be even locally bounded, but on every ball U it belongs to all $L^r(U)$.

For higher derivatives the following assertions are valid.

1.1.2. Corollary. *One has the following embeddings.*

(i) *If $kp < d$, then $W^{p,k}(\mathbb{R}^d) \subset L^{dp/(d-kp)}(\mathbb{R}^d)$.*

(ii) *If $kp > d$, then $W^{p,k}(\mathbb{R}^d) \subset C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$.*

(iii) $W^{1,d}(\mathbb{R}^d) \subset C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$.

Hölder norms of Sobolev functions admit the following estimates.

1.1.3. Theorem. *Let $rp > d$, let U be a ball of radius 1 in $\mathbb{R}^d$, and let $f \in W^{p,r}(U)$. Then f has a modification f_0 which satisfies Hölder's condition with exponent $\alpha = \min(1, r - d/p)$, and there exists $C(d,p,r) > 0$ such that for all $x, y \in U$ one has the inequality*

$$|f_0(x) - f_0(y)| \leqslant C(d,p,r)\|f\|_{p,r}|x-y|^\alpha. \tag{1.1.7}$$

If $f \in W_0^{p,r}(U)$, then

$$|f_0(x) - f_0(y)| \leqslant C(d,p,r)\|D^r f\|_{L^p(U)}|x-y|^\alpha, \tag{1.1.8}$$

where $\|D^r f\|_{L^p(U)}$ denotes the $L^p(U)$-norm of the real function

$$x \mapsto \sup_{|v_i| \leqslant 1} |D^r f(x)(v_1, \ldots, v_r)|.$$

A similar assertion is true for domains with sufficiently regular boundaries, but the constants will depend also on the domains.

Unlike the whole space, for a bounded domain $\Omega \subset \mathbb{R}^d$, one has the inclusion $L^p(\Omega) \subset L^r(\Omega)$ whenever $p > r$. This yields a wider spectrum of embedding theorems. We formulate the main results for a ball $U \subset \mathbb{R}^d$. Let us set $W^{q,0} := L^q$.

1.1.4. Theorem. (i) *Let $kp < d$. Then*

$$W^{p,j+k}(U) \subset W^{q,j}(U), \quad q \leqslant \frac{dp}{d-kp},\ j \in \{0,1,\ldots\}.$$

(ii) *Let $kp = d$. Then*

$$W^{p,j+k}(U) \subset W^{q,j}(U), \quad q < \infty,\ j \in \{0,1,\ldots\}.$$

If $p = 1$, then $W^{j+d,1}(U) \subset C_b^j(U)$.

(iii) *Let $kp > d$. Then*

$$W^{p,j+k}(U) \subset C_b^j(U), \quad j \in \{0,1,\ldots\}.$$

In addition, these embeddings are compact operators, with the exception of case (i) *with $q = dp/(d-kp)$.*

Proofs of all these classic results can be found in the book Adams, Fournier [**3**].

For $p > d$ and any function $f \in W^{p,1}(\mathbb{R}^d)$ with support in a ball of radius R one has the estimate

$$\|f\|_{L^\infty} \leqslant C(p,d,R)\big\||\nabla f|\big\|_p.$$

Neither this estimate nor (1.1.4) hold for functions on bounded domains (for example, for constant functions). Also a constant $C(p,d,R)$ cannot be taken independently of R (excepting the case $d = p = 1$), as simple computations with the functions $f_j(x) = \max(1 - |x|/j, 0)$ show.

Under broad assumptions about a set Ω in $\mathbb{R}^d$, the class $W_0^{p,k}(\Omega)$ (defined above as the closure of $C_0^\infty(\Omega)$ in $W^{p,k}(\Omega)$) admits the following description (see Adams, Fournier [**3**, Theorem 5.29 and Theorem 5.37]).

1.1.5. Theorem. *Let Ω be a bounded open set with smooth boundary. Then the class $W_0^{p,k}(\Omega)$ coincides with the set of functions in $W^{p,k}(\Omega)$ whose extensions by zero outside Ω belong to $W^{p,k}(\mathbb{R}^d)$.*

1.1.6. Corollary. *Let Ω be a bounded open set with smooth boundary. Suppose that $f \in W^{p,k}(\Omega)$, where $p > d$. If the continuous version of f vanishes on $\partial\Omega$ along with its derivatives up to order $k-1$, then $f \in W_0^{p,k}(\Omega)$.*

Let U_R be an open ball of radius R. First we want to recall some simple properties of the space $W^{p,-1}(U_R)$, which is the dual of $W_0^{p',1}(U_R)$ for $p \in (1,\infty)$. It is known (see, e.g., Adams, Fournier [**3**, Chapter III, Theorem 3.12]) that every $u \in W^{p,-1}(U_R)$ can be written as

$$u = \partial_{x_i} f^i, \quad f^i \in L^p(U_R),\ i = 1,\ldots,d, \tag{1.1.9}$$

and, for all representations (1.1.9), one has

$$\|u\|_{W^{p,-1}(U_R)} \leqslant \|f\|_{L^p(U_R)}. \tag{1.1.10}$$

By using scaling to control the norms of the embeddings, we arrive at the following well-known lemma (see, e.g., Gilbarg, Trudinger [**409**, Theorem 7.10]).

1.1.7. Lemma. (i) *Let $d' < r < \infty$ and $R > 0$. Then we have the continuous embedding $L^{rd/(r+d)}(U_R) \subset W^{r,-1}(U_R)$. In addition, there exists a number N independent of R such that*

$$\|u\|_{W^{r,-1}(U_R)} \leqslant N\|u\|_{L^{rd/(r+d)}(U_R)} \tag{1.1.11}$$

for all $u \in L^{rd/(r+d)}(U_R)$ and all $R > 0$.

(ii) *Let $1 < r < d'$ and $R > 0$. Then $L^1(U_R) \subset W^{r,-1}(U_R)$ and the embedding operator is bounded. In addition, there exists a number N independent of R such that*

$$\|u\|_{W^{r,-1}(U_R)} \leqslant NR^{1-d/r'}\|u\|_{L^1(U_R)} \tag{1.1.12}$$

for all $u \in L^1(U_R)$ and all $R > 0$.

(iii) *Let $r = d'$, $s > 1$, and $R > 0$. Then $L^s(U_R) \subset W^{r,-1}(U_R)$. In addition, there exists N independent of R such that*

$$\|u\|_{W^{r,-1}(U_R)} \leqslant NR^{2+d/s}\|u\|_{L^s(U_R)} \tag{1.1.13}$$

for all $u \in L^s(U_R)$ and all $R > 0$.