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

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

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

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

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*

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

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

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

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

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

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

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

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

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 $R > 0$.

1.2. Elliptic equations

For convenience of later references we collect here a number of known results about second order elliptic equations. Throughout $\Delta = \partial_{x_1}^2 + \cdots + \partial_{x_d}^2$ is the Laplace operator. An elliptic operator (or a "nondivergence form elliptic operator") is an expression

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

where a^{ij}, b^i and c are functions on $\mathbb{R}^d$, $A = (a^{ij})_{i,j\leqslant d}$, $b = (b^i)_{i=1}^d$ and the summation over repeated upper and lower indices is meant, moreover, $A = A^* \geqslant 0$. Such operators should be distinguished from "divergence form" operators

$$\mathcal{L}_{A,b,c}u = \partial_{x_i}(a^{ij}\partial_{x_j}u + b^i u) + cu,$$

to which it is customary to ascribe also more general operators

$$\mathcal{L}_{A,b,\beta,c}u = \partial_{x_i}(a^{ij}\partial_{x_j}u + b^i u) + \beta^i\partial_{x_i}u + cu, \quad \beta = (\beta^i).$$

As we shall see below, different forms of operators lead not only to different properties of solutions to the equations of the form $L_{A,b,c}u = f$ (a "direct" or "nondivergence form" equation) or $\mathcal{L}_{A,b,c}v = f$ (a "divergence form" equation), but even to different settings of problems. We note at once that our principal object — a stationary Fokker–Planck–Kolmogorov equation — is in general something third.

Nondivergence and divergence form equations are most often solved in Hölder classes (functions with Hölder continuous derivatives up to the second order) and in Sobolev classes. Let us mention the basic facts about Dirichlet problems on domains and about equations on the whole space.

A function u in the class $W_0^{p,1}(\Omega)$ on an open set Ω in $\mathbb{R}^d$ is called a solution of the equation

$$\mathcal{L}_{A,b,\beta,c}u = \nu, \quad \text{where } \nu \in W_0^{p,-1}(\Omega),\ p > 1,$$

if a^{ij}, b^i, β^i, c are measurable, $a^{ij}|\nabla u|, b^i u, \beta^i|\nabla u|, cu \in L^p(\Omega)$, and

$$\int_\Omega \bigl[-\langle A\nabla u - ub, \nabla\varphi\rangle + \varphi\langle\beta, \nabla u\rangle + cu\varphi\bigr]\,dx = \nu(\varphi)$$

for all functions $\varphi \in W_0^{p/(p-1),1}(\Omega)$ or, equivalently, for all $\varphi \in C_0^\infty(\Omega)$, where $\nu(\varphi)$ is the value of the functional ν at φ. In the case of bounded coefficients the required integrability conditions are automatically fulfilled.

Let Ω be a bounded domain in $\mathbb{R}^d$ with smooth boundary, let functions a^{ij} be Hölder continuous on the closure of Ω, and let the matrix $A(x)$ be symmetric and positive definite on $\overline{\Omega}$. It is known (see, for example, Gilbarg, Trudinger [**409**, Theorem 6.14], Krylov [**552**, Theorem 6.5.3]) that for every function $f \in C_0^\infty(\Omega)$ there is a function $u \in C^2(\overline{\Omega})$ such that $u = 0$ on $\partial\Omega$ and

$$a^{ij}\partial_{x_i}\partial_{x_j}u = f \quad \text{on } \Omega.$$

It is known (see [**409**, Lemma 9.17] or Krylov [**556**, Theorem 2, p. 242]) that for every $r > 1$ there is a number C_r independent of f such that

$$\|u\|_{W^{r,2}(\Omega)} \leqslant C_r\|f\|_{L^r(\Omega)}. \tag{1.2.1}$$

If A is merely continuous on $\overline{\Omega}$, then for any $f \in L^r(\Omega)$ the equation $L_{A,b,c}u = f$ with lower order terms $b^i, c \in L^\infty(\Omega)$ has a solution in the space $W^{r,2}(\Omega)\cap W_0^{r,1}(\Omega)$ if $c \leqslant 0$ (say, if $c = 0$). In this case also the indicated estimate holds.

The most general known conditions on the second order coefficients ensuring the solvability in Sobolev classes are formulated in terms of the class VMO consisting of locally integrable functions a on $\mathbb{R}^d$ for each of which there is a positive continuous function ω on $[0,+\infty)$ with $\omega(0)=0$ such that

$$\sup_{z\in\mathbb{R}^d, r<R} r^{-2d}\int_{U_r(z)\times U_r(z)} |a(x)-a(y)|\,dx\,dy \leqslant \omega(R) \quad \forall\, R>0.$$

This class contains all uniformly continuous functions, but includes also some locally unbounded functions. Note that the inclusion $W^{d,1}(\mathbb{R}^d)\subset VMO$ holds. About these conditions, see Dong [**308**], [**309**], Krylov [**557**] and the references given there. The following result is proved in Krylov [**555**].

1.2.1. Theorem. *Suppose that*

$$a^{ij}\in VMO,\ A(x)\geqslant \varepsilon\cdot \mathrm{I},\ |a^{ij}|+|b^i|+|\beta^i|+|c|\leqslant K.$$

Then, for every $p>1$, there are numbers λ_0 and M depending only on p,d,K,ε and a common for all a^{ij} function ω from the condition of the membership in VMO such that for all $\lambda\geqslant\lambda_0$ and $f,g^1,\dots,g^d\in L^p(\mathbb{R}^d)$ the equations

$$L_{A,b,c}u-\lambda u=f \quad \textit{and} \quad \mathcal{L}_{A,b,\beta,c}v-\lambda v=f+\operatorname{div} g,\ \ g=(g^1,\dots,g^d),$$

have unique solutions $u\in W^{p,2}(\mathbb{R}^d)$ and $v\in W^{p,1}(\mathbb{R}^d)$ and

$$\lambda\|u\|_p+\|u\|_{p,2}\leqslant M\|(L_{A,b,c}-\lambda)u\|_p, \tag{1.2.2}$$

$$\|v\|_{p,1}\leqslant M(\|f\|_p+\|g\|_p). \tag{1.2.3}$$

Thus, the operator $L_{A,b,c}-\lambda$ is an isomorphism between the spaces $W^{p,2}(\mathbb{R}^d)$ and $L^p(\mathbb{R}^d)$, $\mathcal{L}_{A,b,\beta,c}-\lambda$ is an isomorphism between $W^{p,1}(\mathbb{R}^d)$ and $W^{p,-1}(\mathbb{R}^d)$.

Estimate (1.2.3) means that $\|v\|_{p,1}\leqslant M\|\mathcal{L}_{A,b,\beta,c}v-\lambda v\|_{p,-1}$.

1.2.2. Corollary. *Under the conditions indicated in the theorem, for every ball U, whenever $\lambda\geqslant\lambda_0$, we have the estimate*

$$\|u\|_{W_0^{p,1}(U)}\leqslant M\|\mathcal{L}_{A,b,\beta,c}u-\lambda u\|_{W^{p,-1}(U)}, \quad u\in W_0^{p,1}(U). \tag{1.2.4}$$

PROOF. Note that by defining u by zero outside U we obtain a function in $W^{p,1}(\mathbb{R}^d)$ with the same norm, but the norm of $\mathcal{L}_{A,b,\beta,c}u$ will change, so for justifying (1.2.4) we use a different reasoning. We take a sequence of functions $u_n\in C_0^\infty(U)$ converging to u in $W^{p,1}(U)$. Set $\mathcal{L}=\mathcal{L}_{A,b,\beta,c}$. Then

$$\|u_n\|_{W^{p,1}(U)}=\|u_n\|_{W^{p,1}(\mathbb{R}^d)}\leqslant M\|\mathcal{L}u_n-\lambda u_n\|_{W^{p,-1}(\mathbb{R}^d)}=M\|\mathcal{L}u_n-\lambda u_n\|_{W^{p,-1}(U)},$$

where the left-hand side converges to $\|u_n\|_{W^{p,1}(U)}$ and the right-hand side converges to $M\|\mathcal{L}u-\lambda u\|_{W^{p,-1}(U)}$, which follows by the estimate

$$\|\mathcal{L}v\|_{W^{p,-1}(U)}\leqslant \big\||A\nabla v|+|\beta|\,|\nabla v|+|bv|+|cv|\big\|_{L^p(U)}$$

and the boundedness of the coefficients. □

For equations without lower order terms (or under some other additional conditions) one can take $\lambda=0$. For the proof of the following result under more general conditions (in particular, with a bounded domain with C^1-boundary in place of a ball), see Auscher, Qafsaoui [**78**], Byun [**214**]. Let us derive it from the previous corollary.

1.2.3. Proposition. *Under the conditions on A from the theorem, for the operator $\mathcal{L}_A = \mathcal{L}_{A,0,0,0}$ and every ball U one can find $M > 0$ such that*

$$\|u\|_{W_0^{p,1}(U)} \leqslant M \|\mathcal{L}_A u\|_{W^{p,-1}(U)}, \quad u \in W_0^{p,1}(U). \tag{1.2.5}$$

Moreover, for any $f, g^1, \ldots, g^d \in L^p(U)$ the equation

$$\mathcal{L}_A u = f + \operatorname{div} g, \quad g = (g^1, \ldots, g^d)$$

has a unique solution in $W_0^{p,1}(U)$.

Proof. First we observe that our estimate yields the existence of a solution taking into account that the number M in (1.2.4) according to Theorem 1.2.1 depends on A only through p, d, K, ε and the function ω. Indeed, we can approximate A in $L^p(U)$ by a sequence of smooth mappings A_k with common parameters indicated above. The sequence of solutions $u_k \in W_0^{p,1}(U)$ of the equations $\mathcal{L}_{A_k} u_k = f + \operatorname{div} g$ turns out to be bounded in $W_0^{p,1}(U)$, hence a subsequence $\{u_{k_n}\}$ converges weakly in $W_0^{p,1}(U)$ to some function u, which obviously will be a solution to $\mathcal{L}_A u = f + \operatorname{div} g$.

We now establish estimate (1.2.5). Suppose that it fails. Then there exist functions $u_n \in W_0^{p,1}(U)$ such that

$$\|u_n\|_{W_0^{p,1}(U)} = 1, \quad \|\mathcal{L}_A u_n)\|_{W^{p,-1}(U)} \leqslant 1/n.$$

We observe that $\{u_n\}$ converges in $W_0^{p,1}(U)$, since otherwise there is a subsequence $\{v_n\}$ with $\|v_n - v_k\|_{W_0^{p,1}(U)} \geqslant c > 0$, whence we obtain

$$\|(\mathcal{L}_A - \lambda_0)(v_n - v_k)\|_{W^{p,-1}(U)} \geqslant c/M,$$

hence $\|v_n - v_k\|_{W^{p,-1}(U)} \geqslant c/(2M\lambda_0)$ for sufficiently large n. This contradicts the compactness of the embedding $W_0^{p,1}(U) \subset W^{p,-1}(U)$. Thus, there is $u = \lim\limits_{n\to\infty} u_n$ in $W_0^{p,1}(U)$. Then $\|u\|_{W_0^{p,1}(U)} = 1$, but $\mathcal{L}_A u = 0$, i.e., the integral of $\langle A\nabla u, \nabla\varphi\rangle$ vanishes for all functions $\varphi \in C_0^\infty(U)$, then also for all $\varphi \in W_0^{p',1}(U)$, whence it follows that $u = 0$. Indeed, if $p \geqslant 2$, then the integral of $\langle A\nabla u, \nabla u\rangle$ over U vanishes, which is only possible if $u = 0$, since $u \in W_0^{p,1}(U)$. The estimate proven for $p \geqslant 2$ yields also the existence of a solution, as observed above.

We can now complete our proof of (1.2.5) in the case $1 < p < 2$. It remains to show that $u = 0$ if $u \in W_0^{p,1}(U)$ and $\mathcal{L}_A u = 0$. As shown above, we can solve the equation $\mathcal{L}_A w = \operatorname{sign} u$ in $W_0^{p',1}(U)$. Then the integral of $|u|$ equals the integral of $-\langle A\nabla w, \nabla u\rangle$, which equals the vanishing integral of $-\langle \nabla w, A\nabla u\rangle$. □

1.2.4. Corollary. *Let the conditions on A indicated in Theorem* 1.2.1 *hold and $u \in W_0^{q,1}(U)$ for some $q > 1$. If*

$$\mathcal{L}_A u = f + \operatorname{div} g, \quad g = (g^1, \ldots, g^d),$$

where $f, g^i \in L^p(U)$ and $p > q$, then $u \in W_0^{p,1}(U)$.

Proof. Let $w \in W_0^{p,1}(U)$ be a solution of the equation $\mathcal{L}_A u = f + \operatorname{div} g$, which exists by Proposition 1.2.3. Then the difference $v = u - w \in W_0^{q,1}(U)$ satisfies the homogeneous equation $\mathcal{L}_A v = 0$, but this equation has only zero solution in the class $W_0^{q,1}(U)$. Therefore, $u = w$ almost everywhere. □

Below we need the following technical assertion which follows from the previous proposition and embedding theorems.

1.2.5. Lemma. *Let p and q be two numbers satisfying the estimates $p \geqslant d$, $q \geqslant p'$, but not such that $p = d = q'$. Let $R_1 > 0$. Assume that the functions $a^{ij} \in W^{p,1}(U_{R_1})$ are continuous and $A \geqslant \lambda \cdot \mathrm{I}$ for some $\lambda > 0$. Then, there exist $N > 0$ and $R_0 > 0$ depending only on p, q, d, λ, R_1, the modulus of continuity of A, $\|a^{ij}\|_{W^{p,1}(U_{R_1})}$, and the rate of decreasing to zero of $\|\nabla a^{ij}\|_{L^d(U_R)}$ as $R \to 0$, such that for all $R < R_0$ and $\varphi \in W_0^{q,1}(U_R)$, one has*

$$f := a^{ij}\partial_{x_i}\partial_{x_j}\varphi \in W^{q,-1}(U_R) \quad \text{and} \quad \|\nabla\varphi\|_{L^q(U_R)} \leqslant N\|f\|_{W^{q,-1}(U_R)}. \tag{1.2.6}$$

Proof. We may assume that $R_1 = 1$. Note $f \in W^{q,-1}(U_R)$. This follows from the fact that, for every bounded function $\zeta \in W^{p,1}(U_R)$, the operator $\psi \mapsto \zeta\psi$ is continuous on $W_0^{q',1}(U_R)$ by the estimate

$$\|\nabla(\zeta\psi)\|_{L^{q'}(U_R)} \leqslant C\|\nabla\psi\|_{L^{q'}(U_R)}.$$

Indeed, if $q' < d$, we have $|\nabla\zeta| \in L^p(U_R)$ and $\psi \in L^{q'd/(d-q')}(U_R)$. Hence by Hölder's inequality $|\psi\nabla\zeta| \in L^s(U_R)$, where

$$s = \frac{pq'd/(d-q')}{p + q'd/(d-q')} = \frac{pq'd}{pd - pq' + q'd} = q'\frac{pd}{pd - pq' + q'd} \geqslant q'.$$

In addition,

$$\begin{aligned}\|\psi\nabla\zeta\|_{L^s(U_R)} &\leqslant \|\psi\|_{L^{q'd/(d-q')}(U_R)}\|\nabla\zeta\|_{L^p(U_R)} \\ &\leqslant C\|\nabla\psi\|_{L^{q'}(U_R)}\|\nabla\zeta\|_{L^p(U_R)}.\end{aligned}$$

The case $q' > d$, where ψ is bounded and $q' \leqslant p$ (since $q \geqslant p'$), and the case $q' = d$, where $q' < p$ and ψ is in all $L^r(U_R)$, $r < \infty$, are similar. We note that

$$\partial_{x_i}(a^{ij}\partial_{x_j}\varphi) = f + \partial_{x_i}a^{ij}\partial_{x_j}\varphi =: f + g.$$

By the previous proposition

$$\|\nabla\varphi\|_{L^q(U_R)} \leqslant N_1\Bigl(\|g\|_{W^{q,-1}(U_R)} + \|f\|_{W^{q,-1}(U_R)}\Bigr), \tag{1.2.7}$$

where N_1 is independent of $R \in (0,1]$ and φ. Let $a = (a^j)$, $a^j = \partial_{x_i}a^{ij}$.

Now we consider three cases.

Case $q > d'$. By Lemma 1.1.7(i) and (1.1.1) we have

$$\|g\|_{W^{q,-1}(U_R)} \leqslant N_2\|a\|_{L^d(U_R)}\|\nabla\varphi\|_{L^q(U_R)},$$

which along with (1.2.7) yields

$$\|\nabla\varphi\|_{L^q(U_R)} \leqslant N_1N_2\|a\|_{L^d(U_R)}\|\nabla\varphi\|_{L^q(U_R)} + N_1\|f\|_{W^{q,-1}(U_R)}. \tag{1.2.8}$$

We emphasize that N_1 and N_2 are independent of R and f and note that, since $|\nabla a^{ij}| \in L^p(U_R)$ and $p \geqslant d$, we can choose R so small that

$$N_1N_2\|a\|_{L^d(U_R)} \leqslant 1/2.$$

For such an R, inequality (1.2.8) implies (1.2.6).

Case $p' < q \leqslant d'$. In that case it follows from $q > p'$ that, for r defined by $rd/(r+d) = pq/(p+q)$, we have $r > d' \geqslant q$. Therefore, for $R \in (0,1)$, we obtain

$$\|g\|_{W^{q,-1}(U_R)} \leqslant N_3R^{d(r-q)/(rq)}\|g\|_{W^{r,-1}(U_R)} \leqslant N\|a\|_{L^p(U_R)}\|\nabla\varphi\|_{L^q(U_R)},$$

and we can finish the proof as above.

Case $q = p' < d'$. As is easy to see, this is the only remaining case. By Lemma 1.1.7 and (1.1.1), for $R \in (0,1)$, we have

$$\|g\|_{W^{q,-1}(U_R)} \leqslant N_4 R^{1-d/p} \|g\|_{L^1(U_R)} \leqslant N_4 \|a\|_{L^p(U_R)} \|\nabla\varphi\|_{L^q(U_R)},$$

and the argument from the first case applies again to complete the proof. □

In § 1.7 we return to divergence form equations.

An important role in the theory of partial differential equations is played by various maximum principles. These principles can be of the following types:

1) the weak maximum principle asserts that, under certain conditions, if $Lu \leqslant 0$ in a domain Ω, $u|_{\partial\Omega} \geqslant 0$, then $u \geqslant 0$ in Ω; this maximum principle is discussed in Chapter 2; note that if $Lu = 0$ and $L1 = 0$, then the maximum and minimum of u are attained at the boundary;

2) the strong maximum principle asserts that, under certain conditions (see Gilbarg, Trudinger [**409**, § 3.2]), if $Lu = 0$ and u attains its minimum or maximum in the interior of Ω, then u is a constant. Let us give a precise formulation.

1.2.6. Theorem. *Suppose that* $L_{A,b}u \geqslant 0$ *in a connected open set* Ω*, where* $u \in C^2(\Omega)$*,* $c_1 \cdot \mathrm{I} \leqslant A(x) \leqslant c_2 \cdot \mathrm{I}$ *with constant* $c_1, c_2 > 0$*, and* b *is bounded. If* u *attains its maximum in the interior of* Ω*, then* u *is constant in* Ω*.*

1.3. Diffusion processes

Fokker–Planck–Kolmogorov equations arise naturally in the study of diffusion processes. Here we recall basic concepts and consider some examples. These concepts are not used in the main part of the book, but some acquaintance with them is useful for understanding the origins of the central problems of the book and the character of the most important applications of the presented analytical results.

First we define the concept of a Markov transition function on a measurable space $(X, \mathcal{X})$ in which all singletons belongs to $\mathcal{X}$. Suppose we are given a nonempty set $T \subset \mathbb{R}$. A function $(s,x,t,B) \mapsto P(s,x,t,B)$ defined for all $s,t \in T$ with $s \leqslant t$, $x \in X$ and $B \in \mathcal{X}$ is called a Markov transition function if

1) for all fixed s,t,x, the function $B \mapsto P(s,x,t,B)$ is a probability measure on $\mathcal{X}$ and for $s = t$ it is Dirac's measure at the point x;

2) for all fixed s,t,B, the function $x \mapsto P(s,x,t,B)$ is measurable with respect to $\mathcal{X}$;

3) whenever $s,t,u \in T$ and $s \leqslant t \leqslant u$, for all $x \in X$ and $B \in \mathcal{X}$ we have the equality

$$P(s,x,u,B) = \int_X P(t,y,u,B)\, P(s,x,t,dy), \tag{1.3.1}$$

called the *Chapman–Kolmogorov equation*.

A random process $\{\xi_t\}_{t \in T}$ with values in X is called a Markov process with the given transition function $P(s,x,t,B)$ if, for all $t,u \in T$ with $t \leqslant u$ and $B \in \mathcal{X}$, the function $P(t,\xi_t,u,B)$ serves as a conditional probability $P(\xi_u \in B|\mathcal{F}_{\leqslant t})$ with respect to the σ-algebra $\mathcal{F}_{\leqslant t}$ generated by the random elements ξ_s with $s \leqslant t$.

The quantity $P(s,x,t,B)$ can be interpreted as the probability of hitting the set B by the process at the time t under the condition that it is at the point x at the time $s \leqslant t$. So the measures $P(s,x,t,\,\cdot\,)$ are also called the transition probabilities of the process. It is also possible to consider Markov families $\{\xi_{s,x,t}\}$ for which $s \leqslant t$ and $\xi_{s,x,s} = x$. Certainly, in the general case there is no requirement that the process must be at a fixed point at the initial time. If $P(s,x,t,\,\cdot\,)$ depends on s,t through $t-s$, then the process is called *homogeneous*; in this case

$$P(s,x,t,\,\cdot\,) = P(0,x,t-s,\,\cdot\,) =: P(x,t-s,\,\cdot\,).$$

The one-dimensional distributions P_t of the process $\{\xi_t\}_{t\in T}$ are defined by the equality $P_t(B) := P(\xi_t \in B)$. A necessary and sufficient condition that a process be Markov with the given transition function is the equality

$$(1.3.2)\quad P\bigl((\xi_{t_1},\dots,\xi_{t_n})\in C\bigr) = \int_X\cdots\int_X I_C(x_1,\dots,x_n) \times P(t_{n-1},x_{n-1},t_n,dx_n)\cdots P(t_1,x_1,t_2,dx_2)P_{t_1}(dx_1)$$

for all $C\in\mathcal{X}^n$ and $t_i\in T$ with $t_1<\cdots<t_n$.

A somewhat more general concept is obtained if in place of the family of the σ-algebras $\mathcal{F}_{\leqslant t}$ we take an increasing family of σ-algebras $\mathcal{F}_t$ with the property that ξ_t is $\mathcal{F}_t$-measurable.

Let $U(x,\varepsilon)=\{y\colon |x-y|<\varepsilon\}$, $V(x,\varepsilon)=\{y\colon |x-y|>\varepsilon\}$. A Markov process with values in $\mathbb{R}^d$ with transition probabilities $P(s,x,t,B)$ is called a *diffusion process* or a *diffusion* (see Wentzell [**937**] or Gikhman, Skorokhod [**408**]) if there is a mapping $b\colon \mathbb{R}^d\times[0,+\infty)\to\mathbb{R}^d$, called the *drift coefficient*, and a mapping $(x,t)\mapsto A(x,t)$ with values in the space of symmetric operators on $\mathbb{R}^d$, called the *diffusion coefficient* or *diffusion matrix*, such that

(i) for all $\varepsilon>0$, $t\geqslant 0$ and $x\in\mathbb{R}^d$ we have

$$\lim_{h\to 0} h^{-1}P\bigl(t,x,t+h,V(x,\varepsilon)\bigr)=0,$$

(ii) for some $\varepsilon>0$ and all $t\geqslant 0$, $x\in\mathbb{R}^d$ we have

$$\lim_{h\to 0} h^{-1}\int_{U(x,\varepsilon)}(y-x)\,P(t,x,t+h,dy)=b(x,t),$$

(iii) for some $\varepsilon>0$ and all $t\geqslant 0$, $x,z\in\mathbb{R}^d$ we have

$$\lim_{h\to 0} h^{-1}\int_{U(x,\varepsilon)}\langle y-x,z\rangle^2\,P(t,x,t+h,dy)=2\langle A(x,t)z,z\rangle.$$

If A and b do not depend on t, then the diffusion is homogeneous.

1.3.1. Proposition. *Suppose that relations* (i)–(iii) *hold locally uniformly in* x *and the functions* a^{ij}, b^i *are locally bounded. Then the transition probabilities satisfy the parabolic Fokker–Planck–Kolmogorov equation*

$$\partial_t\mu = \partial_{x_i}\partial_{x_j}(a^{ij}\mu)-\partial_{x_i}(b^i\mu)$$

in the sense of generalized functions (*see Chapter* 6). *If* ν *is a finite Borel measure on* $\mathbb{R}^d$ *and*

$$\mu_t(dx)=\int_{\mathbb{R}^d}P(0,y,t,dx)\,\nu(dy),$$

then the measure $\mu=\mu_t(dx)\,dt$ *gives a solution to the Cauchy problem with the initial condition* $\mu\big|_{t=0}=\nu$.

PROOF. We give a brief justification, see details in Wentzell [**937**, § 11.2] or Gikhman, Skorokhod [**408**, Chapter 1, § 1]. Let $f \in C_0^\infty(\mathbb{R}^d)$. Then

$$\frac{d}{dt}\int_{\mathbb{R}^d} f(y)\,P(s,x,t,dy) = \lim_{h\to 0} h^{-1}\Bigl(\int_{\mathbb{R}^d} f(y)\,P(s,x,t+h,dy) - \int_{\mathbb{R}^d} f(z)\,P(s,x,t,dz)\Bigr).$$

By the Chapman–Kolmogorov equation the right-hand side equals

$$\lim_{h\to 0}\int_{\mathbb{R}^d} h^{-1}\int_{\mathbb{R}^d}\bigl(f(y)-f(z)\bigr)\,P(t,z,t+h,dy)\,P(s,x,t,dz).$$

By using conditions (i)–(iii) and Taylor's expansion for f, we obtain

$$\lim_{h\to 0} h^{-1}\int_{\mathbb{R}^d}\bigl(f(y)-f(z)\bigr)\,P(t,z,t+h,dy) = a^{ij}(z,t)\partial_{z_i}\partial_{z_j}f(z) + b^i(z,t)\partial_{z_i}f(z).$$

Since convergence as $h\to 0$ is uniform in z, we have

$$\frac{d}{dt}\int_{\mathbb{R}^d} f(y)\,P(s,x,t,dy) = \lim_{h\to 0}\int_{\mathbb{R}^d}\bigl(a^{ij}(z,t)\partial_{z_i}\partial_{z_j}f(z) + b^i(z,t)\partial_{z_i}f(z)\bigr)\,P(s,x,t,dz).$$

Thus, we have proved that the transition probabilities satisfy the indicated equation. In addition, for each function $\zeta \in C_0^\infty(\mathbb{R}^d)$, condition (i) gives the equality

$$\lim_{h\to 0}\int_{\mathbb{R}^d}\zeta(y)\,P(s,x,s+h,dy) = \zeta(x),$$

i.e., $P(s,x,t,dy)$ satisfies the condition $P\big|_{t=s} = \delta_x$. This proves the last assertion in the case where $\nu = \delta_x$. The general case follows by integration with respect to ν. □

In the case where the transition probabilities $P(s,x,t,dy)$ are given by densities $\varrho(s,x,t,y)$ with respect to Lebesgue measure, in the variables (y,t) they satisfy the above Fokker–Planck–Kolmogorov equation (also called the forward Kolmogorov equation)

$$\partial_t\varrho(s,x,t,y) = \partial_{y_i}\partial_{y_j}\bigl(a^{ij}(y,t)\varrho(s,x,t,y)\bigr) - \partial_{y_i}\bigl(b^i(y,t)\varrho(s,x,t,y)\bigr),$$

and in the variables (x,s) they satisfy the backward Kolmogorov equation

$$-\partial_s\varrho(s,x,t,y) = a^{ij}(x,s)\partial_{x_i}\partial_{x_j}\varrho(s,x,t,y) + b^i(x,s)\partial_{x_i}\varrho(s,x,t,y).$$

If A and b do not depend on t, then

$$P(s,x,t,dy) = P(0,x,t-s,dy)$$

under broad assumptions, i.e., the transition probabilities are determined by the probabilities

$$P(x,t,dy) = P(0,x,t,dy).$$

In case the latter have densities $\varrho(x,t,y)$, the backward Kolmogorov equation takes the form

$$\partial_t\varrho(x,t,y) = a^{ij}(x)\partial_{x_i}\partial_{x_j}\varrho(x,t,y) + b^i(x)\partial_{x_i}\varrho(x,t,y).$$

The differential operator

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

is called the *generator* of the given process. This terminology is connected with the fact that under suitable conditions the operators

$$T_t f(x) = \int_{\mathbb{R}^d} f(y)\, P(x,t,dy)$$

form a semigroup in a suitable functional space (semigroups will be discussed in Chapter 5). Then the forward Kolmogorov equation (which is the Fokker–Planck–Kolmogorov equation) reads $\partial_t T_t f = T_t L f$ and the backward equation becomes $\partial_t T_t f = L T_t f$. Say, if $A = \mathrm{I}$, $b^i \in C_b^\infty(\mathbb{R}^d)$, then the given formal relations have the usual meaning for $f \in C_b^\infty(\mathbb{R}^d)$.

In the case of a homogeneous process an important concept of a *stationary distribution* or *invariant measure* of the process (or of its transition semigroup) arises. This is a probability measure μ such that

$$\mu(B) = \int_{\mathbb{R}^d} P(y,t,B)\,\mu(dy) \quad \forall\, t \geqslant 0,\ B \in \mathcal{B}(\mathbb{R}^d).$$

It follows from what has been said that any stationary distribution satisfies the *stationary Fokker–Planck–Kolmogorov equation*

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

which will be the main object of study in Chapters 1–5.

Diffusion processes can be considered also in a broader sense, for example, one can consider almost surely continuous Markov processes in $\mathbb{R}^d$ such that their transition probabilities $P(s,x,t,dy)$ satisfy the Fokker–Planck–Kolmogorov equation with the initial condition $P\big|_{t=s} = \delta_x$. Such processes are called *quasi-diffusions*.

Since the distribution of a Markov process (or the family of its finite-dimensional distributions) is completely determined by its initial distribution and its transition probabilities, uniqueness of a probability solution to the Cauchy problem for the Fokker–Planck–Kolmogorov equation yields the weak uniqueness of the diffusion process whose transition probabilities satisfy this equation.

We recall that the conditional expectation of an integrable function ξ on a probability space $(\Omega, \mathcal{A}, \mu)$ with respect to a σ-algebra $\mathcal{B} \subset \mathcal{A}$ is a $\mathcal{B}$-measurable integrable function $\mathbb{E}_\mu[\xi|\mathcal{B}]$ such that

$$\int_\Omega \eta\xi\, d\mu = \int_\Omega \eta\mathbb{E}_\mu[\xi|\mathcal{B}]\, d\mu$$

for every bounded $\mathcal{B}$-measurable function η.

Jensen's inequality for the conditional expectation says that if V is a convex function and $V(\xi) \in L^1(\mu)$, then a.e.

$$V(\mathbb{E}_\mu[\xi|\mathcal{B}]) \leqslant \mathbb{E}_\mu[V(\xi)|\mathcal{B}].$$

A real or vector random process $\{\xi_t\}_{t\in T}$ with a directed index set T is called a *martingale* with respect to a family of σ-algebras $\mathcal{F}_t$ that is increasing in the sense that $\mathcal{F}_s \subset \mathcal{F}_t$ whenever $s \leqslant t$, provided that the element ξ_s is measurable with respect to $\mathcal{F}_s$, integrable and almost surely $\xi_s = \mathbb{E}[\xi_t|\mathcal{F}_s]$ for $t \geqslant s$, where $\mathbb{E}[\xi_t|\mathcal{F}_s]$ is the conditional expectation (the existence of the conditional expectation follows by the integrability).

One of the most important examples of a Markov process which is also a martingale (and one of the most important for applications processes) is the Wiener process.

1.3.2. Example. A Wiener process (or a Brownian motion) $\{w_t\}_{t\geqslant 0}$ is a real random process with the following properties:

(i) the trajectory $t\mapsto w_t(\omega)$ is continuous for every ω and $w_0=0$,

(ii) the random variables $w_{t_1}, w_{t_2}-w_{t_1},\ldots,w_{t_n}-w_{t_{n-1}}$ are independent, whenever $0\leqslant t_1<t_2<\cdots<t_n$,

(iii) for each t the random variable w_t is Gaussian with zero mean and variance t, i.e., $\mathbb{E}w_t^2=t$.

A Wiener process in $\mathbb{R}^d$ is just a collection $(w_t^1,\ldots,w_t^d)$ of independent Wiener processes. The existence of Wiener processes is not straightforward and is proved in many textbooks (see, e.g., Wentzell [**937**]).

Let $\mathcal{F}_t$ be the σ-algebra generated by the variables w_s with $s\leqslant t$. Then with respect to this family the Wiener process is a martingale, since for $s\leqslant t$ the conditional expectation of w_t-w_s with respect to $\mathcal{F}_s$ vanishes by condition (ii). In addition, the Wiener process is Markov with respect to the indicated family with the transition function $P(s,x,t,\,\cdot\,)$ defined as follows: if $s=t$, then this is Dirac's measure at the point x, if $s<t$, then this is the Gaussian measure with mean x and variance $t-s$, i.e., the measure with density $y\mapsto(2\pi)^{-1/2}\exp\bigl[-(y-x)^2/(2t-2s)\bigr]$. The Chapman–Kolmogorov equation is verified directly. The Markov property is verified by means of (1.3.2).

A diffusion process with a nonzero drift is not a martingale, which can be seen from the Itô equation (see below).

The most important way of constructing diffusion processes is solving stochastic differential equations. First we introduce the Itô integral.

Suppose we are given a Wiener process $\{w_t\}_{t\geqslant 0}$ and a process $\{\xi_t\}_{t\geqslant 0}$ measurable in (ω,t) such that the variable ξ_t is measurable with respect to the σ-algebra $\mathcal{F}_t$ generated by the variables w_s with $s\leqslant t$ (such a process is called adapted). Let also $T>0$. Suppose that

$$\int_0^T \mathbb{E}|\xi_t|^2\,dt<\infty.$$

The stochastic Itô integral

$$\int_0^T \xi_t\,dw_t$$

is defined as follows. If there are points $0=t_1<\cdots<t_n=T$ such that $\xi_t=\xi_{t_i}$ whenever $t_i\leqslant t<t_{i+1}$, then this integral is naturally defined as the sum $\sum_{i=1}^n \xi_{t_i}(\omega)\bigl(w_{t_{i+1}}(\omega)-w_{t_i}(\omega)\bigr)$. In the general case the function $\xi_t(\omega)$ is approximated in $L^2(P\otimes dt)$ by a sequence of functions of the indicated form and it is proved that the stochastic integrals of the approximations converge in $L^2(P)$; the limit is taken for the stochastic integral of the original process.

Similarly one defines the stochastic integral with respect to a Wiener process in $\mathbb{R}^d$ of a real-valued or operator-valued process.

Suppose now we are given Borel functions σ and b on $\mathbb{R}$. If

$$\xi_t=\xi_{t_0}+\int_{t_0}^t \sigma(\xi_s)\,dw_s+\int_{t_0}^t b(\xi_s)\,ds,\ \ t_0\leqslant t\leqslant T,$$

then the adapted process $\{\xi_t\}_{t\in[t_0,T]}$ is called a *strong solution* of the stochastic differential equation

$$d\xi_t = \sigma(\xi_t)dw_t + b(\xi_t)dt \tag{1.3.3}$$

on $[t_0, T]$ with the initial distribution ξ_{t_0}. This equation is a symbolic expression for the previous integral equation. Similarly stochastic equations are introduced for processes in $\mathbb{R}^d$, when b is a vector field, σ is an operator-valued mapping.

It is known (see Wentzell [**937**] or Krylov [**551**]) that in the case of Lipschitzian coefficients σ and b for any $\mathcal{F}_{t_0}$-measurable square-integrable random variable ξ_{t_0} this stochastic equation has a unique solution and this solution is a diffusion process with the drift b and the diffusion coefficient $\sigma\sigma^*/2$. There are also more subtle results (see Ikeda, Watanabe [**473**]). The following theorem on existence of a strong solution is proved in Gyöngy, Krylov [**432**] (we give its version for the whole space).

1.3.3. Theorem. *Suppose that $\sigma = (\sigma^{ij})_{i,j\leqslant d}$ and b are Borel mappings from $\mathbb{R}^d\times[0,+\infty)$ to the space of matrices on $\mathbb{R}^d$ and to $\mathbb{R}^d$, respectively, such that for every $k \in \mathbb{N}$ there are a positive function M_k integrable on $[0, k]$ and a number $\varepsilon_k > 0$ such that*

$$|\sigma^{ij}(x,t) - \sigma^{ij}(y,t)|^2 \leqslant M_k(t)|x-y|^2,\ \ |b(x,t)| + |\sigma^{ij}(x,t)| \leqslant M_k(t),$$

$$A(x,t) := 2^{-1}\sigma(x,t)\sigma^*(x,t) \geqslant \varepsilon_k\cdot \mathrm{I} \quad \textit{if } |x|\leqslant k,\quad t\in[0,k],\ i,j\leqslant d.$$

Let $V \geqslant 0$ be a function on $\mathbb{R}^d\times[0,+\infty)$ with continuous first and second derivatives in x and a continuous derivative in t such that for some increasing sequence of bounded domains D_k covering $\mathbb{R}^d$ for each $T > 0$ we have

$$L_{A,b}V(x,t) \leqslant M(t)V(t,x), \qquad \inf_{x\in\partial D_k,\, t\leqslant T} V(x,t) \to +\infty,$$

where M is a locally integrable function on $[0,+\infty)$. Then (1.3.3) has a unique strong solution on $[0,+\infty)$.

In Chapter 9 also the concept of weak solution will be mentioned and the related concept of martingale problem (which in turn is strongly related to Fokker–Planck–Kolmogorov equations).

Apart from the Wiener process, the *Ornstein–Uhlenbeck process* is very useful in applications. This process is given by the linear stochastic equation (scalar or vector)

$$d\xi_t = dw_t - 2^{-1}\xi_t\, dt.$$

It can be expressed via the Wiener process by the formula

$$\xi_t = e^{-t/2}\xi_0 + e^{-t/2}\int_0^t e^{s/2}\, dw_s.$$

For $\xi_0 = x_0$ the process $\widehat{\xi_t} = e^{-t/2}x_0 + e^{-t/2}w_{e^t-1}$ has the same finite-dimensional distributions as ξ_t (but does not satisfy the above stochastic equation). The generator of the Ornstein–Uhlenbeck process has the form $L/2$, where the operator

$$L\varphi(x) = \Delta\varphi(x) - \langle x, \nabla\varphi(x)\rangle$$

is called the *Ornstein–Uhlenbeck operator*.

We observe that the Wiener process has no stationary probability measures, but the standard Gaussian measure is invariant for the Ornstein–Uhlenbeck process and the Ornstein–Uhlenbeck semigroup (see also Examples 1.4.7, 5.1.1 and Exercise 5.6.57).

Let us also mention the *Itô formula*. If the process $\xi_t = (\xi_t^1, \ldots, \xi_t^d)$ in $\mathbb{R}^d$ satisfies the equation $d\xi_t = \sigma(\xi_t)dw_t + b(\xi_t)dt$, then for any smooth function f the scalar process $f(\xi_t)$ satisfies the equation

$$\begin{aligned} df(\xi_t) &= \sum_{i=1}^d \partial_{x_i} f(\xi_t) d\xi_t^i + \frac{1}{2}\sum_{i,j=1}^d \partial_{x_i}\partial_{x_j} f(\xi_t) d\xi_t^i d\xi_t^j \\ &= \sum_{i,k=1}^d \partial_{x_i} f(\xi_t)\sigma^{ik}(\xi_t) dw_t^k + \sum_{i,k=1}^d \partial_{x_i} f(\xi_t)\sigma^{ik}(\xi_t) b^i(\xi_t) dt \\ &+ \frac{1}{2}\sum_{i,j,k=1}^d \partial_{x_i}\partial_{x_j} f(\xi_t)\sigma^{ik}(\xi_t)\sigma^{jk}(\xi_t) dt. \end{aligned}$$

The Itô formula will not be used in this book, but some acquaintance with it is useful for better understanding the methods of obtaining certain estimates. For example, the integral of the function f with respect to the solution of the parabolic Fokker–Planck–Kolmogorov equation at the moment t is usually the expectation of $f(\xi_t)$. For example, if $\sigma = \mathrm{I}$, then by the Itô formula this gives the expectation

$$\mathbb{E}f(\xi_0) + \mathbb{E}\int_0^t \Bigl[\langle \nabla f(\xi_t), b(\xi_t)\rangle + \frac{1}{2}\Delta f(\xi_t)\Bigr]\, dt.$$

Under the integral we have $Lf(\xi_t)$, $L = \Delta/2 + b\cdot\nabla$. If $Lf \leqslant C + Cf$, then the right-hand side is estimated by the integral of $C + C\mathbb{E}f(\xi_t)$ over $[0,t]$, which enables us to estimate $\mathbb{E}f(\xi_t)$ by means of the known Gronwall inequality (see Exercise 7.5.3).

1.4. Basic problems

Suppose we are given a locally finite Borel measure μ (possibly signed) on an open set $\Omega \subset \mathbb{R}^d$, a Borel function c on Ω, a Borel vector field $b = (b^i)$ on Ω, and a matrix-valued mapping $A = (a^{ij})_{i,j\leqslant d}$ on Ω such that the functions a^{ij} are Borel measurable. For $\varphi \in C^\infty(\Omega)$ let us set

$$L_{A,b}\varphi := \sum_{i,j\leqslant d} a^{ij}\partial_{x_i}\partial_{x_j}\varphi + \sum_{i\leqslant d} b^i\partial_{x_i}\varphi, \quad L_{A,b,c}\varphi = L_{A,b}\varphi + c\varphi.$$

We shall also consider the divergence form operators

$$\mathcal{L}_{A,b}\varphi := \sum_{i,j\leqslant d} \partial_{x_i}(a^{ij}\partial_{x_j}\varphi) + \sum_{i\leqslant d} b^i\partial_{x_i}\varphi$$

and the correspondingly defined operators $\mathcal{L}_{A,b,c}$.

1.4.1. Definition. We say that μ satisfies the equation

$$L_{A,b,c}^*\mu = 0 \tag{1.4.1}$$

in Ω if $a^{ij}, b^i, c \in L^1_{\mathrm{loc}}(|\mu|)$ and one has

$$\int_\Omega L_{A,b,c}\varphi(x)\,\mu(dx) = 0 \quad \forall \varphi \in C_0^\infty(\Omega). \tag{1.4.2}$$

For a given measure ν on Ω the equation

$$L_{A,b,c}^*\mu = \nu \tag{1.4.3}$$

is defined similarly as the identity $\displaystyle\int_\Omega L_{A,b,c}\varphi\, d\mu = \int_\Omega \varphi\, d\nu$.

For $c = 0$ we arrive at the equations $L^*_{A,b}\mu = 0$ and $L^*_{A,b}\mu = \nu$, the first of which is called the *stationary Fokker–Planck–Kolmogorov equation*; when $c \neq 0$ it is called the *Fokker–Planck–Kolmogorov equation with a potential.*

Equation (1.4.1) can be written as the equality

$$\partial_{x_i}\partial_{x_j}(a^{ij}\mu) - \partial_{x_i}(b^i\mu) + c\mu = 0$$

in the sense of generalized functions. If it is known in advance that the measure μ is given by a density ϱ of class $W^{1,1}_{\rm loc}$ (which is true under broad assumptions, as we shall see below) and the functions a^{ij} are locally bounded and belong to $W^{1,1}_{\rm loc}$ and the functions $\partial_{x_j}a^{ij}\varrho$ are integrable, then the integration by parts in (1.4.2) yields the equation

$$\partial_{x_i}(a^{ij}\partial_{x_j}\varrho) - \partial_{x_i}\bigl((b^i - \partial_{x_j}a^{ij})\varrho\bigr) + c\varrho = 0,$$

also understood in the sense of generalized functions.

If the coefficients are smooth and it is known in advance that the measure μ is given by a smooth density ϱ (which is true if the matrix $A(x)$ is not degenerate), then the double integration by parts in (1.4.2) yields the usual equation

$$a^{ij}\partial_{x_i}\partial_{x_j}\varrho + 2\partial_{x_i}a^{ij}\partial_{x_j}\varrho - b^i\partial_{x_i}\varrho - \partial_{x_i}b^i\varrho + c\varrho = 0.$$

Unlike the direct elliptic equations of the form $L_{A,b,c}u = 0$, the density of a solution of equation (1.4.1) even with Lipcshitzian coefficients and a nondegenerate matrix A may fail to belong to the second class $W^{1,2}_{\rm loc}$ (simple examples are given below).

The equation

$$\mathcal{L}^*_{A,b,c}\mu = 0$$

is defined similarly, but it requires additional assumptions about either a^{ij} or μ (which will be made in appropriate places), because it is necessary to give meaning to the integral of $\mathcal{L}_{A,b,c}\varphi$ with respect to μ. For example, if $a^{ij} \in C^1(\Omega)$, then we write $\partial_{x_i}(a^{ij}\partial_{x_j}\varphi)$ as $\partial_{x_i}a^{ij}\partial_{x_j}\varphi + a^{ij}\partial_{x_i}\partial_{x_j}\varphi$ and use the previous definition.

Let us give a precise definition of a solution of the elliptic equation

$$\mathcal{L}^*_{A,b,c}\mu = 0 \tag{1.4.4}$$

for Borel measures μ on Ω, where $\mathcal{L}$ is an elliptic second order operator of divergence form

$$\mathcal{L}\varphi(x) := \partial_{x_i}\bigl(a^{ij}(x)\partial_{x_j}\varphi(x)\bigr) + b^i(x)\partial_{x_i}\varphi(x).$$

The interpretation of this equation is as usual: the functions a^{ij} and b^i must be integrable on every compact set in Ω with respect to the measure μ and, for every function $\varphi \in C_0^\infty(\Omega)$, we must have the equality

$$\int_\Omega \mathcal{L}_{A,b,c}\varphi\, d\mu = 0.$$

However, the latter can be understood in one of the following two ways.

(I) One has $a^{ij} \in W^{1,1}_{\rm loc}(\Omega)$, the functions a^{ij}, $\partial_{x_i}a^{ij}$, and b^i are Borel measurable and locally integrable with respect to $|\mu|$, and

$$\int_\Omega [a^{ij}\partial_{x_i}\partial_{x_j}\varphi + \partial_{x_i}a^{ij}\partial_{x_j}\varphi + b^j\partial_{x_j}\varphi]\, d\mu = 0. \tag{1.4.5}$$

(II) The measure μ possesses a density ϱ in the class $W^{1,1}_{\rm loc}(\Omega)$ such that the functions $a^{ij}\partial_{x_i}\varrho$ and $b^i\varrho$ are locally Lebesgue integrable and

$$\int_\Omega [-a^{ij}\partial_{x_i}\varrho\partial_{x_j}\varphi + b^i\partial_{x_i}\varphi\varrho]\, dx = 0. \tag{1.4.6}$$

Clearly, if the coefficients a^{ij} are locally Sobolev and the functions $\partial_{x_i} a^{ij}\varrho$ are locally integrable, then (1.4.6) can be written as (1.4.5).

Throughout we deal with the case where the matrix A is symmetric and nonnegative, but this is not needed for the definition (unlike for most of the results).

Probability solutions are those that are probability measures. *Integrable solutions* are those given by integrable densities (possibly signed).

In general, equation (1.4.1) can fail to have nonzero solutions in the class of bounded measures (take $\Omega = \mathbb{R}^1$, $A = 1$, $b = 0$, then the equation $\mu'' = 0$ means that the density of μ is linear), it can have many solutions even in the class of probability measures, and its solutions can be quite singular (e.g., if $A = 0$ and also $b = 0$, then any measure is a solution). However, even in the generality under consideration some positive information is available.

The one-dimensional case is much simpler than the multidimensional case.

1.4.2. Proposition. *Let $d = 1$ and let Ω be an interval (x_0, x_1). Suppose that $A > 0$ on Ω. Then, any measure μ satisfying the equation $L^*_{A,b,c}\mu = \nu$ is absolutely continuous with respect to Lebesgue measure and has a density ϱ of the form $\varrho = \varrho_0/A$, where ϱ_0 is absolutely continuous on every compact subinterval in Ω.*

If $c = 0$ and b/A is locally Lebesgue integrable, $x_2 \in (x_0, x_1)$ is fixed, then

$$\varrho(x) = A(x)^{-1}E(x)\biggl(C_1 + \int_{x_0}^{x} \frac{C_2 + F(t)}{E(t)}\,dt\biggr), \tag{1.4.7}$$

$$E(x) := \exp\int_{x_2}^{x} \frac{b(t)}{A(t)}\,dt, \quad F(x) := \nu\bigl((x_0, x)\bigr).$$

If $A = 1$, $c = 0$, $\nu = 0$, $\Omega = (-1, 1)$, and b is locally Lebesgue integrable on the interval $(-1, 1)$, then

$$\varrho(x) = \biggl(k_1 + k_2\int_0^x \exp\biggl(-\int_0^s b(t)\,dt\biggr)\,ds\biggr)\exp\int_0^x b(t)\,dt, \tag{1.4.8}$$

where k_1 and k_2 are constants.

PROOF. We have the identity

$$\int_\Omega (A\varphi'' + b\varphi' + c\varphi)\,d\mu = \int_\Omega \varphi\,d\nu \quad \forall \varphi \in C_0^\infty(\Omega),$$

which can be written as the equality

$$(A\mu)'' - (b\mu)' + c\mu = \nu$$

in the sense of distributions. Hence the distributional derivative of $(A\mu)' - b\mu$ is a locally bounded measure, i.e., $(A\mu)' - b\mu$ is a function of locally bounded variation. This shows that the distributional derivative of $A\mu$ is a locally bounded measure as well. Hence $A\mu$ is absolutely continuous and has a density ϱ_0. Therefore, μ is absolutely continuous. Now it is seen from our reasoning that the distributional derivative of $A\mu$ is a locally integrable function, so that ϱ_0 admits a locally absolutely continuous version. In the case $A = 1$, $c = 0$, $\nu = 0$, we arrive at the equation $\mu'' - (b\mu)' = 0$, whence $\mu' - b\mu = k_2$ for some constant k_2. If b is locally Lebesgue integrable, this equation can be explicitly solved. The general case reduces to this one by passing to the measure $A\mu$. □

Even in this simplest one-dimensional case we observe that a solution μ can fail to have a continuous density if A is positive but not continuous. We actually see that in the case of nondegenerate A (i.e., $\det A \neq 0$) the regularity of solutions is essentially the regularity of A. We shall see below that in higher dimensions the picture is similar, although the proofs involve much deeper techniques. Another simple observation is that without any assumptions of nondegeneracy on A we obtain that the measure $A\cdot\mu$ is absolutely continuous. A highly nontrivial analogue of this is true also in the multidimensional case.

Sometimes it is useful to construct an equation for which a given function is a solution.

1.4.3. Example. In the one-dimensional case for any two smooth functions f and g with everywhere nonzero Wronskian $W = f'g - fg'$ it is easy to write the equation $L^*_{1,b,c}\mu = 0$ with smooth coefficients for which they form a basis in the space of solutions. To this end we equate the determinant of the matrix with the rows (u, u', u''), (f, f', f''), (g, g', g'') to zero, which gives a second order equation $-Wu''+Au'+Bu = 0$ satisfied by f and g. Dividing by $-W$ we obtain the equation $u'' - (A/W)u' - (B/W)u = 0$, which can be written as $L^*_{1,b,c}u = 0$, $b = A/W$, $c = b' - B/W$. For nonzero c it is not always possible to find an explicit solution, but a new degree of freedom appears, which leads to some effects impossible in the case where $c = 0$.

Let us consider one more instructive example.

1.4.4. Example. Let $\varrho \in W^{1,1}_{loc}(\mathbb{R}^d)$ and let $\mu = \varrho\, dx$. Then μ satisfies the equation $L^*_{\mathrm{I},b}\mu = 0$ with

$$b := \frac{\nabla\varrho}{\varrho}, \quad \text{where } b(x) := 0 \text{ whenever } \varrho(x) = 0.$$

Indeed, $|b|$ is locally $|\mu|$-integrable. For any $\varphi \in C_0^\infty(\mathbb{R}^d)$, by the integration by parts formula we have

$$\int [\Delta\varphi + \langle b, \nabla\varphi\rangle]\, \varrho\, dx = \int [-\langle\nabla\varphi, \nabla\varrho\rangle + \langle b, \nabla\varphi\rangle\varrho]\, dx = 0$$

since $b\varrho = \nabla\varrho$ almost everywhere due to the fact that $\nabla\varrho$ vanishes almost everywhere on the set $\{\varrho = 0\}$ (see Exercise 1.8.19).

1.4.5. Definition. Let $\varrho \in W^{1,1}_{loc}(\mathbb{R}^d)$. The mapping $\nabla\varrho/\varrho$, where we set by definition $\nabla\varrho(x)/\varrho(x) := 0$ if $\varrho(x) = 0$, is called the *logarithmic gradient* of the measure μ or of the density ϱ.

In this example, we can even choose ϱ to be infinitely differentiable, but b can be quite singular with respect to Lebesgue measure. For instance, given a proper closed subset $Z \subset \mathbb{R}^d$, we can find a probability density $\varrho \in C^\infty(\mathbb{R}^d)$ with $Z = \{\varrho = 0\}$; in this way one can even obtain b that is not Lebesgue locally integrable on a closed set of positive Lebesgue measure. The simplest example of a singularity is this:

$$\varrho(x) = x^2 \exp(-x^2/2)/\sqrt{2\pi}, \quad b(x) = x + 2x^{-1}. \tag{1.4.9}$$

In the case of smooth coefficients and nondegenerate A all solutions are smooth. This is a corollary of the following classical result (see Taylor [**894**, Chapter III], Trèves [**898**, Chapter I]), which is often referred to as Weyl's regularity theorem for the equation

$$L_{A,b,c}\mu = \nu.$$

1.4.6. Theorem. *Suppose that* a^{ij}, b^i, $c \in C^\infty(\Omega)$ *and* $\det A > 0$. *If* μ *is a distribution on* Ω *such that* $L_{A,b,c}\mu \in C^\infty(\Omega)$, *then* $\mu \in C^\infty(\Omega)$.

Therefore, if a measure ν *on* Ω *has an infinitely differentiable density, then any measure* μ *on* Ω *satisfying the equation* $L^*_{A,b,c}\mu = \nu$ *possesses an infinitely differentiable density.*

The second assertion follows from the first one, since in the case of smooth coefficients the equation $L^*_{A,b,c}\mu = \nu$ can be written as the equality

$$\partial_{x_i}\partial_{x_j}(a^{ij}\mu) - \partial_{x_i}(b^i\mu) + c\mu = \nu$$

in the sense of distributions, which can be rewritten as

$$a^{ij}\partial_{x_i}\partial_{x_j}\mu + (\partial_{x_i}\partial_{x_j}a^{ij})\mu + 2\partial_{x_i}a^{ij}\partial_{x_j}\mu - (\partial_{x_i}b^i)\mu - b^i\partial_{x_i}\mu + c\mu = \nu,$$

i.e., as $L_{A,b_0,c_0}\mu = \nu$ with some smooth coefficients b_0 and c_0.

Explicitly solvable equations are rather rare, although there are important cases when they appear.

1.4.7. Example. Let μ be the standard Gaussian measure on $\mathbb{R}^d$, i.e., a probability measure with the standard Gaussian density

$$\varrho(x) = (2\pi)^{-d/2}\exp(-|x|^2/2).$$

Its logarithmic gradient has an especially simple form:

$$\frac{\nabla\varrho(x)}{\varrho(x)} = -x.$$

According to the previous example, μ satisfies the equation $L^*\mu = 0$, where L is the Ornstein–Uhlenbeck operator

$$L\varphi(x) = \Delta\varphi(x) - \langle x, \nabla\varphi(x)\rangle,$$

already encountered above. This operator plays an important role in analysis, probability theory, and the most diverse applications. We shall see below that any bounded measure σ on $\mathbb{R}^d$ satisfying the equation $L^*\sigma = 0$ has the form $\sigma = k\mu$, where k is a constant. It is worth noting that the operator L has an eigenbasis in $L^2(\gamma)$. For $d = 1$ an eigenbasis is formed by the Hermite–Chebyshev polynomials

$$H_0 = 1, \quad H_n(t) = \frac{(-1)^n}{\sqrt{n!}}e^{t^2/2}\frac{d^n}{dt^n}\bigl(e^{-t^2/2}\bigr), \quad n > 0.$$

Here $LH_n = -nH_n$. For $\mathbb{R}^d$ an eigenbasis is formed by the polynomials

$$H_{k_1,\ldots,k_d}(x_1,\ldots,x_d) = H_{k_1}(x_1)\cdots H_{k_d}(x_d), \quad k_i \geqslant 0.$$

Here

$$LH_{k_1,\ldots,k_d} = -(k_1 + \cdots + k_d)H_{k_1,\ldots,k_d}.$$

In this case the operator L is obviously symmetric in $L^2(\mu)$ on the domain of definition $C_0^\infty(\mathbb{R}^d)$, but this is not always true, as one can see from the following result.

1.4.8. Proposition. *Suppose that a nonnegative locally finite measure μ with a density $\varrho \in W^{p,1}_{loc}(\mathbb{R}^d)$ satisfies the equation $L^*_{A,b}\mu = 0$, where $a^{ij} \in W^{p,1}_{loc}(\mathbb{R}^d)$, $b^i \in L^p_{loc}(\mathbb{R}^d)$, $p \geqslant 2$. Then the symmetry of the operator $L_{A,b}$ on domain $C_0^\infty(\mathbb{R}^d)$ in $L^2(\mu)$, i.e., the identity*

$$\int_{\mathbb{R}^d} \varphi L_{A,b}\psi \, d\mu = \int_{\mathbb{R}^d} \psi L_{A,b}\varphi \, d\mu, \quad \varphi, \psi \in C_0^\infty(\mathbb{R}^d),$$

is equivalent to the almost everywhere equality

$$A\nabla\varrho = \varrho b_0, \quad b_0^i := b^i - \sum_{j=1}^d \partial_{x_j} a^{ij}.$$

For $A = \mathrm{I}$ the symmetry is equivalent to the equality $\nabla\varrho = \varrho b$.

Proof. Indeed, by the integration by parts formula the indicated identity is equivalent to the identity

$$\int_{\mathbb{R}^d} \langle \varphi\nabla\psi - \psi\nabla\varphi, A\nabla\varrho - \varrho b_0\rangle \varrho \, dx = 0,$$

which by the identity

$$\int_{\mathbb{R}^d} \langle \nabla(\varphi\psi), A\nabla\varrho - \varrho b_0\rangle \, dx = 0$$

that follows from the equation turns out to be equivalent to the relation

$$\int_{\mathbb{R}^d} \varphi\langle \nabla\psi, A\nabla\varrho - \varrho b_0\rangle \, dx = 0, \quad \varphi, \psi \in C_0^\infty(\mathbb{R}^d).$$

The latter is equivalent to the equality

$$\langle \nabla\psi, A\nabla\varrho - \varrho b_0\rangle = 0 \quad \text{a.e. for every function } \psi \in C_0^\infty(\mathbb{R}^d),$$

i.e., is the announced equality, since for ψ we can take a function that coincides with the coordinate function x_i on a given cube, which yields that $(A\nabla\varrho)^i = \varrho b_0^i$. $\square$

Note that the symmetry of the operator $L_{A,b}$ on domain $C_0^\infty(\mathbb{R}^d)$ in $L^2(\mu)$ implies that $L^*_{A,b}\mu = 0$ under the much weaker assumption that $a^{ij}, b^i \in L^2_{loc}(\mu)$, since for $\varphi \in C_0^\infty(\mathbb{R}^d)$ we can take $\psi \in C_0^\infty(\mathbb{R}^d)$ in such a way that $\psi = 1$ on the support of φ.

We shall see below that under broad assumptions any solution of the equation $L^*_{A,b,c}\mu = 0$ has the same smoothness as the diffusion coefficient A. However, even in the one-dimensional case it is easy to find an example where the smoothness of the solution does not exceed that of A.

1.4.9. Example. Let us take a probability measure μ with a smooth density that satisfies the equation $L^*_{\mathrm{I},b_0}\mu = 0$, e.g., let μ be the standard Gaussian measure and $b_0(x) = -x$. If now g is any Borel function with $1 \leqslant g \leqslant 2$, then the measure $g \cdot \mu$ satisfies the equation $L^*_{A,b}\mu = 0$ with $A = g^{-1}\mathrm{I}$ and $b = g^{-1}b_0$. In particular, in this way we can obtain an example, where A and b are Hölder continuous and A is uniformly nondegenerate, but the density of μ is not weakly differentiable and its Hölder order is not greater than that of A.

1.5. Existence of densities

We now turn to conditions for the existence of densities of solutions. Suppose that $A = (a^{ij})_{i,j=1}^d$ is a Borel measurable mapping on an open set $\Omega \subset \mathbb{R}^d$ with values in the space of nonnegative symmetric operators on $\mathbb{R}^d$. The main results of this section assert that under broad assumptions any solution of the equation $L^*_{A,b,c}\mu = 0$ has a density on the set where $\det A > 0$, and it is possible to estimate certain L^p-norms of the density.

For the proof we need the classical result following from the Riesz theorem and asserting that every linear functional Λ defined on a linear subspace E in $L^p(\lambda)$, where λ is a nonnegative σ-finite measure and $p \in [1, \infty)$, and satisfying the estimate

$$\Lambda(f) \leqslant C\|f\|_{L^p(\lambda)}, \quad f \in E,$$

is given by means of some function $g \in L^{p/(p-1)}(\lambda)$ in the form

$$\Lambda(f) = \int fg\,d\lambda, \quad f \in E.$$

In addition, we also need the following corollary of a very deep maximum principle due to A.D. Aleksandrov.

1.5.1. Theorem. *For every smooth positive function f on a uniformly convex smooth domain Ω (the principal curvatures of $\partial\Omega$ are separated from zero, e.g., Ω is a ball) there is a convex function $z \in C^2(\Omega) \cap C(\overline{\Omega})$ such that $z|_{\partial\Omega} = 0$ and*

$$\alpha^{ij}\partial_{x_i}\partial_{x_j}z \geqslant d|\det(\alpha^{ij})|^{1/d} f \quad \text{on } \Omega$$

for every nonnegative symmetric matrix (α^{ij}) and

$$\sup_{x\in\Omega} |z(x)| \leqslant C(d, \Omega)\|f\|_{L^d(\Omega)}.$$

Proof. It is known (see Gilbarg, Trudinger [**409**, Theorem 17.23]) that for every smooth positive function f on a uniformly convex smooth domain Ω there exists a convex solution $z \in C^2(\Omega) \cap C(\overline{\Omega})$ of the Dirichlet problem

$$\det(D^2 z) = f^d, \quad z|_\Omega = 0. \tag{1.5.1}$$

Then by A.D. Aleksandrov's maximum principle (see [**409**, Theorem 9.1]) we have

$$\sup_{x\in\Omega} |z(x)| \leqslant C(d, \Omega)\|f\|_{L^d(\Omega)}.$$

If now $\alpha = (\alpha^{ij})$ is a nonnegative symmetric $d\times d$-matrix, we have

$$\alpha^{ij}\partial_{x_i}\partial_{x_j}z = \operatorname{tr}(\alpha D^2 z) \geqslant d|\det\alpha \det(D^2 z)|^{1/d} = d|\det\alpha|^{1/d} f,$$

since $\operatorname{tr}(AB) \geqslant d|\det(AB)|^{1/d}$ for any nonnegative symmetric $d\times d$-matrices, because $\operatorname{tr}(AB) = \operatorname{tr}(\sqrt{B}A\sqrt{B})$ and $\sqrt{B}A\sqrt{B}$ is a nonnegative matrix. □

Here is one of the main results in this section.

1.5.2. Theorem. *Suppose that the matrix $A(x)$ is symmetric and nonnegative-definite for every x. Let μ be a locally finite Borel measure on Ω (possibly signed) such that $a^{ij} \in L^1_{\mathrm{loc}}(\Omega, \mu)$, and for some $C > 0$ one has*

$$\int_\Omega a^{ij}\partial_{x_i}\partial_{x_j}\varphi\,d\mu \leqslant C(\sup_\Omega |\varphi| + \sup_\Omega |\nabla\varphi|) \tag{1.5.2}$$

for all nonnegative $\varphi \in C_0^\infty(\Omega)$. Then the following assertions are true.

(i) *If* μ *is nonnegative, then* $(\det A)^{1/d}\mu$ *has a density in* $L^{d'}_{\rm loc}(\Omega, dx)$.

(ii) *If* A *is locally Hölder continuous and* $\det A > 0$, *then* μ *has a density which belongs to* $L^r_{\rm loc}(\Omega, dx)$ *for every* $r \in [1, d')$.

Proof. We shall start with case (ii) which is simpler. Let U_0 be a ball with compact closure in Ω and let $\zeta \in C_0^\infty(\Omega)$ be such that $0 \leqslant \zeta \leqslant 1$, $\zeta = 1$ on U_0 and the support of ζ belongs to a ball $U \subset \Omega$. Let us consider the measure $\nu = \zeta\,\mu$. By substituting $\zeta\psi$ in place of φ in (1.5.2), for every nonnegative smooth function ψ on Ω, we obtain

$$\int_U a^{ij}\partial_{x_i}\partial_{x_j}\psi\, d\nu \leqslant C_1(\sup_U |\psi| + \sup_U |\nabla\psi|), \tag{1.5.3}$$

where

$$C_1 = C + \bigl(C + 2d^2 \sup_{i,j} \|a^{ij}\|_{L^1(U,\mu)}\bigr) \sup_U |\nabla\zeta| + \|a^{ij}\|_{L^1(U,\mu)} \sup_U |\partial_{x_i}\partial_{x_j}\zeta|$$

is independent of ψ. It is easily seen that (1.5.3) remains true for every nonnegative $\psi \in C^2(\overline{U})$. By considering the function $\psi + \sup|\psi|$, we arrive at the estimate

$$\left|\int_U a^{ij}\partial_{x_i}\partial_{x_j}\psi\, d\nu\right| \leqslant C_1(\sup_U |\psi| + \sup_U |\nabla\psi|) \quad \forall\, \psi \in C^2(\overline{U}). \tag{1.5.4}$$

Now let $r > d$. As we noted in § 1.2, for every $f \in C_0^\infty(U)$ there exists a function $u \in C^2(\overline{U})$ such that

$$a^{ij}\partial_{x_i}\partial_{x_j}u = f$$

on U and $u = 0$ on ∂U. Moreover, there exists a constant C_2 independent of f such that

$$\|u\|_{W^{r,2}(U)} \leqslant C_2\|f\|_{L^r(U)}.$$

By the Sobolev embedding theorem, we obtain

$$\sup_U |\nabla u| + \sup_U |u| \leqslant C_3\|f\|_{L^r(U)}.$$

Together with (1.5.4) this yields

$$\int_U f\, d\nu \leqslant C_1C_3\|f\|_{L^r(U)} \quad \forall\, f \in C_0^\infty(U). \tag{1.5.5}$$

Hence ν is absolutely continuous with $\nu = g\,dx$, $g \in L^{r'}(U)$.

Let us now consider case (i). The above reasoning does not work in this case even for bounded uniformly nondegenerate A, since the equation $a^{ij}\partial_{x_i}\partial_{x_j}u = f$ need not be solvable; for continuous A, the solution u of this equation is only in $W^{r,2}$ and not in C^2, hence one cannot pass from C_0^∞-functions to u in (1.5.4). In order to overcome this difficulty, we need the assumption that μ is nonnegative. As above, by considering a suitable function ζ, we arrive at estimate (1.5.4) for the measure $\nu = \zeta\mu$ on the open ball $U_{R_0}(x_0)$. Note that the support of the measure ν is contained in a ball $U_R(x_0)$ of radius $R = R_0 - 2r$, where $r > 0$. In that case, instead of solving the elliptic equation, we shall employ Theorem 1.5.1, according to which, for every nonnegative continuous function f on $\mathbb{R}^d$ vanishing outside the closed ball $\overline{U_{R_0}}(x_0)$, there exists a nonnegative continuous concave function z (the convex function from the theorem with the minus sign) on $\overline{U_{R_0}}(x_0)$ with the following property:

$$-\alpha^{ij}\partial_{x_i}\partial_{x_j}z \geqslant |\det(\alpha^{ij})|^{1/d} f$$

in $U_{R_0}(x_0)$ for every nonnegative matrix (α^{ij}) and

$$\sup_{U_{R_0}(x_0)} z \leqslant N\|f\|_{L^d(U_{R_0}(x_0))},$$

where N is independent of f and (α^{ij}). Let g be a fixed smooth probability density on $\mathbb{R}^d$ whose support is contained in the unit ball centered at the origin. For any locally integrable function v, we set

$$v_\varepsilon = v * g_\varepsilon, \quad g_\varepsilon(x) = \varepsilon^{-d} g(\varepsilon^{-1}x).$$

Then, for every nonnegative matrix (α^{ij}) and every $\varepsilon \in (0, r)$, one has the estimates

$$-\alpha^{ij}\partial_{x_i}\partial_{x_j} z_\varepsilon(x) \geqslant |\det(\alpha^{ij})|^{1/d} f_\varepsilon(x), \tag{1.5.6}$$

$$\sup_{U_R(x_0)} |z_\varepsilon| \leqslant N\|f_\varepsilon\|_{L^d(U_{R_0}(x_0))} \leqslant N\|f\|_{L^d(U_{R_0}(x_0))} \tag{1.5.7}$$

on $U_{R+r}(x_0)$, where N is independent of f, α^{ij}, and ε. Clearly, the functions z_ε are smooth, nonnegative and concave on $U_{R+r}(x_0)$ if $\varepsilon < r$. We observe that, for every nonnegative continuously differentiable concave function w on $U_{R+r}(x_0)$, one has

$$|\nabla w(x)| \leqslant r^{-1} \sup_{y \in U_{R+r}(x_0)} w(y) \quad \forall\, x \in U_R(x_0).$$

This estimate follows by considering the one-dimensional case. Together with (1.5.4), (1.5.6) applied to $\alpha^{ij} = a^{ij}(x)$ and (1.5.7), this yields the estimate

$$\begin{aligned}\int |\det(a^{ij})|^{1/d} f_\varepsilon\, d\nu &\leqslant \left|\int a^{ij}\partial_{x_i}\partial_{x_j} z_\varepsilon\, d\nu\right| \leqslant C_1 \sup_{U_R(x_0)} \Bigl(|\nabla z_\varepsilon| + |z_\varepsilon|\Bigr) \\ &\leqslant C_1 N(1 + r^{-1})\|f\|_{L^d(U_{R_0}(x_0))}.\end{aligned}$$

As in case (ii), we complete the proof. □

Notice that in assertion (ii) one cannot expect that the density of μ is continuous even for infinitely differentiable a^{ij}, which is seen if one takes $d = 1$, $\Omega = (-1, 1)$, $A = 1$ and $\mu(dx) = I_{(0,+\infty)}\, dx$.

We do not know whether assertion (i) remains true for signed measures.

1.5.3. Corollary. *Let μ be a locally finite (possibly signed) Borel measure on Ω and let a^{ij}, b^i, $c \in L^1_{\mathrm{loc}}(\Omega, \mu)$. Assume that*

$$\int_\Omega (L_{A,b}\varphi + c\varphi)\, d\mu \leqslant 0 \quad \textit{for all nonnegative} \ \ \varphi \in C_0^\infty(\Omega). \tag{1.5.8}$$

Then the following assertions are true.

(i) *If μ is nonnegative, then the locally finite measure* $(\det A)^{1/d}\mu$ *has a density in $L^{d'}_{\mathrm{loc}}(\Omega, dx)$.*

(ii) *If A is locally Hölder continuous and* $\det A > 0$, *then μ has a density which belongs to $L^r_{\mathrm{loc}}(\Omega, dx)$ for every $r \in [1, d')$.*

In particular, the above statements are true if (1.4.1) *holds.*

Proof. It suffices to note that, for every bounded open $\Omega_0 \subset \overline{\Omega}_0 \subset \Omega$, one has

$$\left|\int_{\Omega_0} (b^i\partial_{x_i}\varphi + c\varphi)\, d\mu\right| \leqslant \sup_{\Omega_0} |\nabla\varphi| \int_{\Omega_0} |b|\, d|\mu| + \sup_{\Omega_0} |\varphi| \int_{\Omega_0} |c|\, d|\mu|$$

for every smooth function φ with support in Ω_0. □

In assertion (ii) of this corollary one cannot expect the density of μ to be Hölder continuous, since for $d=1$ and $A=1$ one can take the measure μ with density

$$\exp\int_0^x b(t)\,dt$$

with a suitable function b (see Exercise 1.8.11).

The previous corollary has the following important generalization with the same proof concerned with the nonhomogeneous equation $L^*_{A,b,c}\mu=\nu$ with a measure on the right.

1.5.4. Corollary. *Let μ and ν be two locally finite (possibly signed) Borel measures on Ω and let a^{ij}, b^i, $c\in L^1_{\rm loc}(\Omega,\mu)$. Assume that*

$$\int_\Omega \Bigl[L_{A,b}\varphi + c\varphi\Bigr]\,d\mu = \int_\Omega \varphi\,d\nu \quad \textit{for all nonnegative } \ \varphi\in C_0^\infty(\Omega). \tag{1.5.9}$$

Then the following assertions are true.

(i) *If μ is nonnegative, then the locally finite measure* $(\det A)^{1/d}\mu$ *has a density in $L^{d'}_{\rm loc}(\Omega,dx)$.*

(ii) *If A is locally Hölder continuous and* $\det A>0$*, then μ has a density which belongs to $L^r_{\rm loc}(\Omega,dx)$ for every $r\in[1,d')$.*

1.5.5. Remark. (i) Assertions (i) of Theorem 1.5.2, Corollary 1.5.3, and Corollary 1.5.4 for nonnegative measures extend to the case when μ is a σ-finite nonnegative Borel measure on Ω (not necessarily locally bounded). Indeed, (1.5.2), (1.5.8), and (1.5.9) make sense also for σ-finite μ provided that $a^{ij},b^i,c\in L^1_{\rm loc}(\Omega,\mu)$. One can find a probability measure μ_0 such that $\mu=f\,\mu_0$, where f is a positive Borel function. Let

$$a_0^{ij}:=fa^{ij},\ b_0^i:=fb^i,\ c_0:=fc,\ A_0=(a_0^{ij})_{i,j\leqslant d},\ b_0=(b_0^i)_{i\leqslant d}.$$

Clearly, $a_0^{ij},b_0^i,c_0\in L^1_{\rm loc}(\mu_0)$ and μ_0 satisfies the hypotheses of the above mentioned assertions with A_0, b_0, and c_0 in place of A, b, and c. Hence the measure $(\det A_0)^{1/d}\mu_0$ has a density $\varrho\in L^{d'}_{\rm loc}(\Omega,dx)$. Since we have the equality $(\det A_0)^{1/d}=f(\det A)^{1/d}$, this means that $(\det A)^{1/d}\mu$ has the same density.

(ii) Assume that the hypotheses of Corollary 1.5.3(i) are fulfilled. Suppose that the ball $U_{R_1}(x_0)$ of radius $R_1>0$ centered at a point x_0 is contained in Ω. Then, for every $R<R_1$ and $r<d'$, there exists a number N depending only on R_1,R,r,d such that the density ϱ_A of $(\det A)^{1/d}\mu$ satisfies the estimate

$$\|\varrho_A\|_{L^{d'}(U_R)}\leqslant N\bigl\|1+|b|+|c|\bigr\|_{L^1(U_{R_1},\mu)}.$$

In addition, for fixed d, the number N can be chosen as a locally bounded function of R_1,R,r. This follows from the proof of Theorem 1.5.2.

(iii) Assume that the hypotheses of Corollary 1.5.3(ii) are fulfilled. Let $U_{R_1}(x_0)$ belong to Ω. Then, for every $R<R_1$ and $r<d'$, there exists a number N depending only on R_1,R,r,d, $\inf_{U_{R_1}}\det A$, $\sup_{i,j}\sup_{U_{R_1}}|a^{ij}|$, and the Hölder norm of A on U_{R_1} such that the density ϱ of μ satisfies the estimate

$$\|\varrho\|_{L^r(U_R)}\leqslant N\bigl\|1+|b|+|c|\bigr\|_{L^1(U_{R_1},\mu)}.$$

In addition, for fixed d, the number N can be chosen as a locally bounded function of the indicated quantities. This also follows from the proof of Theorem 1.5.2.

Let us consider an elliptic operator

$$Lu = a^{ij}\partial_{x_i}\partial_{x_j}u + b^i\partial_{x_i}u + cu,$$

where the coefficients a^{ij}, b^i and c are bounded Borel measurable functions on a domain $\Omega \subset \mathbb{R}^d$, the matrix $A(x) = \bigl(a^{ij}(x)\bigr)_{1\leqslant i,j\leqslant d}$ is symmetric and for some positive constants $\lambda > 0$ and $\gamma > 0$ one has

$$\lambda \mathrm{I} \leqslant A(x) \leqslant \lambda^{-1}\mathrm{I} \quad \forall\, x \in \Omega.$$

Moreover, we assume that for every ball $U(x_0, r) \subset \Omega$ we have

$$\sup_{x\in U(x_0,r)} \Bigl[r|b(x)| + r^2|c(x)|\Bigr] \leqslant \lambda^{-1}.$$

The following theorem was obtained in Bauman [**94**]. We say that a nonnegative Borel measure μ satisfies the inequality $L^*\mu \leqslant 0$ in Ω if

$$\int_\Omega L\varphi\, d\mu \leqslant 0 \quad \text{whenever } \varphi \in C_0^\infty(\Omega), \varphi \geqslant 0.$$

1.5.6. Theorem. *Suppose that $U(x_0, r) \subset \Omega$ and $0 < \sigma < \gamma < 1$. There exists a constant $C > 0$ depending only on γ, σ, λ, and d such that if a Borel measure μ is a nonnegative solution of the inequality $L^*\mu \leqslant 0$, then*

$$\mu\bigl(U(x_0, \gamma r)\bigr) \leqslant C\mu\bigl(U(x_0, \sigma r)\bigr).$$

Proof. The theorem follows if we prove that there exists a number $\theta \in (0, 1)$ such that for all $\gamma \in (\theta, 1)$ there holds the estimate

$$\mu\bigl(U(x_0, \gamma r)\bigr) \leqslant C\mu\bigl(U(x_0, \theta r)\bigr),$$

where C depends only on γ, θ, λ, and d. Indeed, iterations of the above estimate will imply that

$$\mu\bigl(U(x_0, \gamma r)\bigr) \leqslant C^k\mu\bigl(U(x_0, \theta^k r/\gamma^{k-1})\bigr).$$

By choosing k so that $\theta^k < \sigma\gamma^{k-1}$ we obtain the assertion of the theorem. Changing variables we may assume that $r = 1$ and $x_0 = 0$.

Set $\varphi(u) = \exp(-u^{-1})$ if $u > 0$ and $\varphi(u) = 0$ if $u \leqslant 0$. We have

$$\begin{aligned} L\varphi(1 - |x|^2) &= -\varphi'(1 - |x|^2)\bigl(2\mathrm{tr}A(x) + 2\langle b(x), x\rangle\bigr) \\ &\quad - 4\varphi''(1 - |x|^2)\langle A(x)x, x\rangle + c(x)\varphi(1 - |x|^2) \\ &\geqslant \varphi(1 - |x|^2)(1 - |x|^2)^{-4}\Bigl(\langle A(x)x, x\rangle\bigl(4 - 2(1 - |x|^2)\bigr) \\ &\qquad - (1 - |x|^2)^2\bigl(2\mathrm{tr}A(x) + 2\langle b(x), x\rangle\bigr) + c(x)(1 - |x|^2)^4\Bigr). \end{aligned}$$

We can choose $\theta \in (0, 1)$ such that

$$\begin{aligned} \langle A(x)x, x\rangle\bigl(4 - 2(1 - |x|^2)\bigr) - (1 - |x|^2)^2\bigl(2\mathrm{tr}A(x) + 2\langle b(x), x\rangle\bigr) \\ + c(x)(1 - |x|^2)^4 \geqslant C_0 > 0 \end{aligned}$$

for every x with $\theta < |x| < 1$. Here C_0 depends only on λ, θ, and d. We have $L\varphi \geqslant 0$ on $U(0, 1) \setminus U(0, \theta)$ and for every $\gamma \in (\theta, 1)$ we have $L\varphi \geqslant C_1$ on $U(0, \gamma) \setminus U(0, \theta)$, where C_1 depends only on λ, θ, γ, and d. Note also that there exists a constant

$C_2 > 0$ depending only on λ and d such that $|L\varphi| \leqslant C_2$ on $U(0,1)$. Since $\mu \geqslant 0$ and $L^*\mu \leqslant 0$, we obtain

$$C_1\mu\bigl(U(0,\gamma)\setminus U(0,\theta)\bigr) \leqslant \int_{U(0,\gamma)\setminus U(0,\theta)} L\varphi\, d\mu \leqslant \int_{U(0,1)\setminus U(0,\theta)} L\varphi\, d\mu$$
$$\leqslant -\int_{U(0,\theta)} L\varphi\, d\mu \leqslant C_2\mu\bigl(U(0,\theta)\bigr).$$

Hence $\mu\bigl(U(0,\gamma)\bigr) \leqslant (C_1 + C_2)\mu\bigl(U(0,\theta)\bigr)$. □

The following theorem is a type of reverse Hölder's inequality. It follows immediately from Theorem 1.5.6 and Theorem 1.5.2.

1.5.7. Theorem. *Suppose that μ is a nonnegative solution of the inequality $L^*\mu \leqslant 0$ on Ω. Let $\gamma > 1$. Then μ has a density ϱ with respect to Lebesgue measure and there exists a constant $C > 0$ depending only on λ, γ, and d such that for every ball $U(x_0,r)$ with $U(x_0,\gamma r) \subset \Omega$, one has*

$$\biggl(\int_{U(x_0,r)} \varrho^{d/(d-1)}\, dx\biggr)^{(d-1)/d} \leqslant C|U(x_0,r)|^{-1/d} \int_{U(x_0,r)} \varrho\, dx.$$

The next interesting fact was discovered in Gehring [**403**].

1.5.8. Lemma. *Let Q be an arbitrary cube in $\mathbb{R}^d$ and let g be a nonnegative function in $L^q_{\rm loc}(\mathbb{R}^d)$ with $q > 1$. Suppose that for almost all $x \in Q$ the inequality*

$$|U|^{-1}\int_U g^q\, dx \leqslant C\biggl(|U|^{-1}\int_U g\, dx\biggr)^q$$

holds for every ball U centered at x. Then, there is a number $\varepsilon > 0$, depending only on q, C and d, such that $g \in L^p(Q)$ for every $p \in [q, q+\varepsilon)$ and

$$|Q|^{-1}\int_Q g^p\, dx \leqslant \frac{\varepsilon}{q+\varepsilon-p}\biggl(|Q|^{-1}\int_Q g^q\, dx\biggr)^{p/q}.$$

1.5.9. Corollary. *Let μ be a nonnegative solution of the inequality $L^*\mu \leqslant 0$ on Ω and let U be a ball with closure in Ω. Then μ has a density ϱ with respect to Lebesgue measure such that there is a number $\varepsilon > 0$, depending only on λ and U, with the property that $\varrho \in L^p(U)$ for every p in the interval $\bigl[1, d/(d-1)+\varepsilon\bigr)$.*

Recall that a Borel measure μ belongs to the Muckenhoupt class A_∞ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\mu(E) \leqslant \varepsilon\mu(U)$ whenever $U \subset \Omega$ is a ball and $E \subset U$ is a Borel set with $|E| \leqslant \delta|U|$.

1.5.10. Corollary. *Let μ be a nonnegative solution of the inequality $L^*\mu \leqslant 0$ on Ω. Let $\gamma > 1$. Then there exists a constant $C > 0$, depending only on λ, γ and d, such that for every ball $U(x_0,r)$ with $U(x_0,\gamma r) \subset \Omega$ and for every Borel set $E \subset U(x_0,r)$ we have*

$$\frac{\mu(E)}{\mu\bigl(U(x_0,r)\bigr)} \leqslant C\Bigl(\frac{|E|}{|U(x_0,r)|}\Bigr)^{1/d},$$

in particular, μ belongs to A_∞ on every subdomain Ω' with $\overline{\Omega'} \subset \Omega$.

PROOF. Let ϱ be a density of the measure μ. Applying Hölder's inequality and Theorem 1.5.7 we obtain

$$\mu(E) = \int_{U(x_0,r)} I_E \varrho\, dx \leqslant |E|^{1/d} \biggl(\int_{U(x_0,r)} \varrho^{d/(d-1)}\, dx\biggr)^{(d-1)/d}$$
$$\leqslant C|E|^{1/d} |U(x_0,r)|^{-1/d} \mu\bigl(U(x_0,r)\bigr),$$

as required. □

1.5.11. Remark. (i) According to Coifman, Fefferman [**254**, Theorem V] and Muckenhoupt [**728**, Theorem 1], the last corollary implies that ϱ is in the Muckenhoupt class A_p for some $p > 1$ on every subset Ω' with $\overline{\Omega'} \subset \Omega$, i.e., there exists a constant $C_1 > 0$ such that

$$\biggl(\frac{1}{|U|}\int_U \varrho\, dx\biggr)\biggl(\frac{1}{|U|}\int_U \varrho^{-1/(p-1)}\, dx\biggr)^{p-1} \leqslant C_1,$$

for every ball $U \subset \Omega'$, where C_1 depends only on λ, d, and Ω'.

(ii) Moreover, for every $\varepsilon > 0$ there exists a constant β such that, for every ball $U \subset \Omega'$, Lebesgue measure of the set

$$\biggl\{ y \in U\colon\ \beta |U|^{-1} \int_U \varrho\, dx \leqslant \varrho(y) \leqslant \beta^{-1} |U|^{-1} \int_U \varrho\, dx \biggr\}$$

is not less than $(1-\varepsilon)|U|$. This assertion may be interpreted as a generalized Harnack principle.

(iii) There exist constants $C_2 > 0$ and $k > 1$ depending only on λ, d, and Ω' such that

$$\biggl(\int_{\Omega'} |\varphi|^{kp} \varrho\, dx\biggr)^{1/(kp)} \leqslant C_2 \biggl(\int_{\Omega'} |\nabla\varphi|^p \varrho\, dx\biggr)^{1/p}$$

for every $\varphi \in C_0^\infty(\Omega')$. See Fabes, Kenig, Serapioni [**347**] for a proof.

1.6. Local properties of densities

We now proceed to the regularity results. Throughout this section we assume that $A(x)$ is symmetric and positive and $A(x)$ *is continuous in* x. By the Sobolev embedding theorem, the continuity assumption is automatically satisfied for some version of A if $a^{ij} \in W^{p,1}_{\rm loc}$, where $p > d$. In Theorem 1.4.6 we have already considered the case of smooth coefficients.

Let us consider the case where the coefficients are only Hölder continuous. The following result was proved in Sjögren [**861**].

1.6.1. Theorem. *Suppose that the coefficients* a^{ij}, b^i, c *are locally Hölder continuous in* Ω *and* $\det A > 0$. *Then any solution* μ *of the equation* $L^*_{A,b,c}\mu = 0$ *has a locally Hölder continuous density.*

Note that the solutions in [**861**] were a priori locally integrable functions, but by the above results the theorem remains true for measures. It would be interesting to study the case where only the coefficients a^{ij} are Hölder continuous. The continuity of all coefficients does not guarantee the Hölder continuity of a solution even if $d = 1$ and $A > 0$. However, it is not clear whether densities of solutions are continuous in the case where the coefficients are just continuous and A is uniformly elliptic. Without the requirement of uniform ellipticity, when A is just nondegenerate and continuous, one can construct a discontinuous probability solution on $\mathbb{R}^d$

with $d > 1$, using an example from Bauman [**95**] (which in turn employs a construction from Modica, Mortola [**720**]). In this example on a disc U in the plane a uniformly elliptic operator L_A with continuous A is such that there is a locally unbounded integrable function $\varrho \geqslant 0$ on $\mathbb{R}^2$ with $L_A^*(\varrho\, dx) = 0$. Taking a diffeomorphism $G\colon \mathbb{R}^2 \to U$, $G = (g^1, g^2)$ with a positive Jacobian, we obtain that the measure μ with the locally unbounded density $\varrho \circ G \det DG$ satisfies the equation $L_{Q,b}^*\mu = 0$ with continuous coefficients, where $Q = (q^{mk})$, $q^{mk} = a^{ij}\partial_{x_j} g^k \partial_{x_i} g^m$, and $b^k = a^{ij}\partial_{x_i}\partial_{x_j} g^k$.

We now proceed to the most difficult case where the diffusion coefficient is somewhat better than Hölder continuous, but is not smooth, and we want to have some Sobolev regularity of densities of solutions. One of the reasons why this is important is that, having established the Sobolev regularity of our solution, we can rewrite the equation $L_{A,b,c}^*\mu = 0$ for μ as a classical equation for its density ϱ in the sense of weak solutions: indeed, integrating by parts, we find that

$$\int_\Omega [a^{ij}\partial_{x_i}\varrho\partial_{x_j}\varphi + \partial_{x_i}a^{ij}\partial_{x_j}\varphi\varrho + b^i\partial_{x_i}\varphi\varrho + c\varrho]\, dx = 0 \quad \forall\, \varphi \in C_0^\infty(\Omega).$$

The difference between the main idea of the proofs in this section and that of the previous one is that now we verify that the solution determines a functional not on L^p, but on a negative Sobolev class, which gives the membership of the measure in a positive Sobolev class.

1.6.2. Theorem. *Let $d \geqslant 2$, $p \geqslant d$, $1 < q < \infty$, and $R_1 > 0$. Suppose that $a^{ij} \in W^{p,1}(U_{R_1})$ and $A \geqslant \lambda\mathrm{I}$, where $\lambda > 0$. Then there exist numbers $R_0 > 0$ and $N_0 > 0$ with the following properties. Let $R < R_0$ and let μ be a measure of finite total variation on U_R such that for any $\varphi \in C_0^2(U_R) := C^2(\overline{U_R}) \cap \{u\colon\ u|_{\partial U_R} = 0\}$ we have the bound*

$$\left|\int_{U_R} a^{ij}\partial_{x_i}\partial_{x_j}\varphi\, d\mu\right| \leqslant N\|\nabla\varphi\|_{L^q(U_R)} \tag{1.6.1}$$

with a number N independent of φ. Furthermore, assume one of the following:

a) *$p > d$ or*

b) *$p = d > q'$ and $\mu \in \bigcup_{r>1} L^r(U_R)$, where we identify μ with its density.*

Then $\mu \in W_0^{q'\wedge p,1}(U_R)$ and

$$\|\mu\|_{W_0^{q'\wedge p,1}(U_R)} \leqslant N_0. \tag{1.6.2}$$

In addition, the radius R_0 can be taken such that it depends only on p, q, d, λ, R_1, $\|a^{ij}\|_{W^{p,1}(U_{R_1})}$, and the rate of decrease of $\|\nabla a^{ij}\|_{L^d(U_R)}$ as $R \to 0$, and N_0 depends on the same quantities and N.

PROOF. We break the proof into three cases.

Case $q \geqslant p'$ and $q \neq d'$. Take $f = (f^1, \ldots, f^d) \in C^2(\overline{U_R})$ and solve the equation

$$a^{ij}\partial_{x_i}\partial_{x_j}\varphi = \partial_{x_i} f^i$$

in U_R with zero boundary conditions. If a) holds, then $p > d$ and A is Hölder continuous in U_R and, by Hölder space theory (see § 1.2), there exists a unique solution $\varphi \in C_0^2(U_R)$ of our problem, which we can substitute into estimate (1.6.1). If b) holds, then, since A is continuous, $\partial_{x_i}\partial_{x_j}\varphi$ are summable to any power by

L^p-theory (see § 1.2), and, owing to $\mu \in \bigcup_{r>1} L^r(U_R)$, we again can substitute φ into inequality (1.6.1). By Lemma 1.2.5 and (1.6.1) we have

$$\left|\int_{U_R} \partial_{x_i} f^i \, \mu \, dx\right| \leqslant N \|\partial_{x_i} f^i\|_{W^{q,-1}(U_R)},$$

which implies our claim. We emphasize that we have established the inclusion of μ to $W_0^{q',1}(U_R)$, not just to $W^{q',1}(U_R)$, since the dual to $W^{q,-1}(U_R)$ is the former smaller space.

Case $q = d' \geqslant p'$. In this case by our assumptions we have $p > d$, so that Lemma 1.2.5 is still applicable.

Case $1 < q < p' < d'$. As is easy to see, this is the only remaining case. Observe that, of course, (1.6.1) is satisfied with $r = (p' + d')/2$ in place of q and, by the first case, we have $\mu \in W_0^{r',1}(U_R)$ if R is sufficiently small. Since $r' > d$, by the Sobolev embedding theorem, μ is bounded in U_R. Furthermore, we note that (1.6.1) means that

$$\varphi \mapsto \int_{U_R} a^{ij} \partial_{x_i} \partial_{x_j} \varphi \mu \, dx$$

is a linear functional defined on a dense subspace $C_0^2(U_R)$ of $W_0^{q,1}(U_R)$ and bounded in the $W_0^{q,1}(U_R)$-norm. By the duality between $W_0^{q,1}(U_R)$ and $W^{q',-1}(U_R)$, we have that

$$\int_{U_R} a^{ij} \partial_{x_i} \partial_{x_j} \varphi \mu \, dx = \int_{U_R} f \varphi \, dx,$$

where $f \in W^{q',-1}(U_R) \subset W^{p,-1}(U_R)$. Thus, μ is a generalized solution of the equation

$$\partial_{x_j}(a^{ij} \partial_{x_i} \mu) = f - \partial_{x_j}(\partial_{x_i} a^{ij} \mu) =: g.$$

Here $\partial_{x_i} a^{ij} \mu \in L^p(U_R)$, since μ is bounded, so that $g \in W^{p,-1}(U_R)$. Since μ belongs to $W_0^{r',1}(U_R)$ and $r' > d \geqslant 2$, we conclude that $\mu \in W_0^{p,1}(U_R)$ by Corollary 1.2.4, which is applicable, since $p > d$. □

1.6.3. Remark. The proof of this theorem actually shows that if μ has compact support in U_{R_1} and (1.6.1) holds for all $\varphi \in C_0^\infty(U_{R_1})$, then $\mu \in W_0^{q' \wedge p,1}(U_R)$ for some $R < R_1$. Moreover, even without the assumption of compactness of support, one can show that $\mu \in W_{\mathrm{loc}}^{q' \wedge p,1}(U_R)$, but this requires some extra work (Exercise 1.8.16).

This theorem yields at once a certain low regularity of solutions to our elliptic equations.

1.6.4. Corollary. *Suppose that* $p > d \geqslant 2$, $a^{ij} \in W_{\mathrm{loc}}^{p,1}(\Omega)$, $\det A > 0$, *and* μ *satisfies the equation* $L_{A,b}^* \mu = 0$, *where* $b \in L_{\mathrm{loc}}^r(\mu)$ *for some* $r > 1$. *Then* μ *has a density in the class* $W_{\mathrm{loc}}^{\alpha,1}(\Omega)$ *for each* $\alpha < dr/(dr - r + 1)$.

Proof. Let us take $\eta \in C_0^\infty(\Omega)$ with support in a ball $U \subset \Omega$ and $0 \leqslant \eta \leqslant 1$. Consider the measure $\mu_0 = \eta \cdot \mu$. We know that μ has a density in $L_{\mathrm{loc}}^s(\Omega)$ with any $s < d/(d-1)$, which will be denoted also by μ. For every $\varphi \in C_0^\infty(U)$ we have

$$\eta a^{ij} \partial_{x_i} \partial_{x_j} \varphi = L_{A,b,c}(\eta\varphi) - \varphi a^{ij} \partial_{x_i} \partial_{x_j} \eta - 2a^{ij} \partial_{x_i} \varphi \partial_{x_j} \eta - \varphi b^i \partial_{x_i} \eta - \eta b^i \partial_{x_i} \varphi - c\eta\varphi.$$

Let $q = \alpha'$. By Hölder's inequality, the integral of $\eta b^i \partial_{x_i}\varphi$ with respect to μ is estimated by

$$\|\nabla\varphi\|_{L^q(U)}\|\eta b\mu\|_{L^\alpha(U)} \leqslant \|\nabla\varphi\|_{L^q(U)}\|\eta b\|_{L^r(\mu)}\|\mu\|_{L^s(U)}^{(r-\alpha)/(r\alpha)},$$

where $s = (r\alpha-\alpha)/(r-\alpha) < d(d-1)$, since $\alpha < dr/(dr-r+1)$, so $\|\mu\|_{L^s(U)}$ is finite. The integrals of the remaining terms are estimated similarly. In particular, the integral of $\eta c\varphi$ is estimated by $\|\varphi\|_\infty\|\eta c\|_{L^1(\mu)}$ and the norm $\|\varphi\|_\infty$ is estimated by $C\|\nabla\varphi\|_{L^q(U)}$, since $q > d$ due to the inequality $dr/(dr-r+1) < d/(d-1)$, which is readily verified. Therefore,

$$\int_U a^{ij}\partial_{x_i}\partial_{x_j}\varphi\, d\mu_0 \leqslant C(\eta, A, b, c)\|\nabla\varphi\|_q.$$

Hence $\eta\mu \in W_0^{\alpha,1}(U)$, which yields our assertion. □

More can be obtained if b is better integrable.

1.6.5. Theorem. *Let $\mu = \varrho\, dx$, $\varrho \in L^r_{\rm loc}(\Omega, dx)$, $a^{ij} \in W^{p,1}_{\rm loc}(\Omega)$, where $p > d$, $r \in (p', \infty)$. Suppose that the mapping A^{-1} is locally bounded and we are given functions*

$$\beta \in L^p_{\rm loc}(\Omega, dx) + L^p_{\rm loc}(\Omega, \mu) \quad \text{and} \quad \gamma \in L^{pd/(p+d)}_{\rm loc}(\Omega, dx) + L^{pd/(p+d)}_{\rm loc}(\Omega, \mu)$$

such that for every $\varphi \in C_0^\infty(\Omega)$ we have

$$\left|\int_\Omega a^{ij}(x)\partial_{x_i}\partial_{x_j}\varphi(x)\,\mu(dx)\right| \leqslant \int_\Omega \bigl(|\varphi(x)|\,|\gamma(x)| + |\nabla\varphi(x)|\,|\beta(x)|\bigr)\,|\mu|(dx).$$

Then $\varrho \in W^{p,1}_{\rm loc}(\Omega)$.

PROOF. Let $\gamma = \gamma_1 + \gamma_2$, $\gamma_1 \in L^{pd/(p+d)}_{\rm loc}(\Omega, dx)$ and $\gamma_2 \in L^{pd/(p+d)}_{\rm loc}(\Omega, \mu)$. Let also $\beta = \beta_1 + \beta_2$, where $\beta_1 \in L^p_{\rm loc}(\Omega, dx)$ and $\beta_2 \in L^p_{\rm loc}(\Omega, \mu)$.

Note that we can assume that $r \neq p'd/(d-p')$, since otherwise we could just slightly decrease the number r. Since $r > p'$, we have $pr > p + r$. Then

$$q := \frac{pr}{pr-p-r} > 1, \quad q' = \frac{pr}{p+r} > 1.$$

According to (1.1.1), $\beta_1\varrho \in L^{q'}_{\rm loc}(\Omega)$ provided $\beta_1 \in L^p_{\rm loc}(\Omega)$. The same is true for $\beta_2 \in L^p_{\rm loc}(\Omega, \mu)$, since

$$|\beta_2|^{q'}|\varrho|^{q'} = |\beta_2|^{pr/(p+r)}|\varrho|^{r/(p+r)}|\varrho|^{(pr-r)/(p+r)},$$

where $|\beta_2|^{pr/(p+r)}|\varrho|^{r/(p+r)} \in L^s_{\rm loc}(\Omega)$ and $|\varrho|^{(pr-r)/(p+r)} \in L^{s'}_{\rm loc}(\Omega)$ with the number $s = (p+r)/r$.

Since $r \neq p'd/(d-p')$, we have $q \neq d$. If $q > d$, then $r > p'd/(d-p')$ and for every ball U_R with $\overline{U_R} \subset \Omega$ we have by Hölder's inequality

$$\|\gamma\varrho\|_{L^1(U_R)} \leqslant \|\gamma_1\|_{L^{pd/(p+d)}(U_R)}\|\varrho\|_{L^{p'd/(d-p')}(U_R)} + \|\gamma_2\|_{L^1(U_R,|\mu|)}.$$

Let $q < d$. Define k by $kdq'/(d+q') = r$, which gives

$$\frac{k'dq'}{d+q'} = \frac{pd}{p+d},$$

because $(d+q')/(dq') = 1/r + 1/p + 1/d$ due to $1/q' = 1/r + 1/p$. Hence by Hölder's inequality with the exponents k' and k we obtain

$$\|\gamma\varrho\|_{L^{dq'/(d+q')}(U_R)} \leqslant \|\gamma_1\|_{L^{pd/(p+d)}(U_R)}\|\varrho\|_{L^r(U_R)} + \|\gamma_2\|_{L^{pd/(p+d)}(U_R,|\mu|)}\|\varrho\|_{L^s(U_R)},$$

where $s = r(1-(d+q')/dq') < r$.

Observe that for every number $R > 0$ such that $U_R := U_R(x_0) \subset \Omega$, whenever $\eta \in C_0^\infty(U_R)$ and $\varphi \in C_0^2(U_R)$, one has

$$\begin{aligned}
&\Bigl|\int_{U_R} a^{ij}\partial_{x_i}\partial_{x_j}\varphi(\eta\varrho)\,dx\Bigr| \leqslant \Bigl|\int_{U_R} a^{ij}\partial_{x_i}\partial_{x_j}(\varphi\eta)\varrho\,dx\Bigr| \\
&\quad + \Bigl|\int_{U_R} a^{ij}\partial_{x_i}\partial_{x_j}\eta(\varphi\varrho)\,dx\Bigr| + 2\int_{U_R}\|A\|\,|\nabla\eta|\,|\nabla\varphi|\,|\varrho|\,dx \\
&\leqslant N_1\int_{U_R}(|\varphi|+|\nabla\varphi|)|\beta\varrho|\,dx \leqslant N_2\|\beta\varrho\|_{L^{q'}(U_R)}\|\nabla\varphi\|_{L^q(U_R)} = N_3\|\nabla\varphi\|_{L^q(U_R)},
\end{aligned} \tag{1.6.3}$$

where the constants N_1, N_2, and N_3 are independent of φ. The last inequality above is due to the estimate $\|\varphi\|_{L^{dq/(d-q)}(U_R)} \leqslant N\|\nabla\varphi\|_{L^q(U_R)}$ if $q < d$ and the estimate $\|\varphi\|_{L^\infty(U_R)} \leqslant N\|\nabla\varphi\|_{L^q(U_R)}$ if $q > d$ with some constant N. It follows by Theorem 1.6.2 that $\eta\varrho \in W_0^{q'\wedge p,1}(U_R)$ if R is small enough, and, since we can take any point as x_0 and $q' < p$, we have

$$\varrho \in W_{\rm loc}^{q',1}(\Omega). \tag{1.6.4}$$

Moreover, if $q' < d$, then by the Sobolev embedding $\varrho \in L_{\rm loc}^{r_1}(\Omega)$ with

$$r_1 = q'd/(d-q') = \frac{prd}{(p+r)d - pr}.$$

The inequality $q' < d$ is equivalent to $r < pd/(p-d)$. Thus, on the interval

$$\bigl(p/(p-1), pd/(p-d)\bigr)$$

we obtain a mapping $T\colon r \mapsto r_1$ with the property that if $\varrho \in L_{\rm loc}^r(\Omega)$, then we have $\varrho \in L_{\rm loc}^{r_1}(\Omega)$. It is easy to see that

$$\frac{r_1}{r} = \frac{pd}{pd - r(p-d)} \geq \frac{pd}{pd - p'(p-d)} = \frac{d'}{p'} > 1,$$

where the first inequality is due to $pr > p + r$ and the second one is due to $p > d$. Hence after finitely many applications of T to the given number r we will come to

$$s \in \bigl(p/(p-1), pd/(p-d)\bigr)$$

such that $t = T(s) \geq pd/(p-d)$ and $\mu \in L^t(U_R)$. Actually, without loss of generality we may assume that $t > pd/(p-d)$, since otherwise we could just slightly decrease the initial point r (and increase the number of iterations of T). This shows that we could assume from the very beginning that $r > pd/(p-d)$ that is $q' > d$. In that case (1.6.4) implies that the function μ is locally bounded, which shows that (1.6.3) is true with $q' = p$. Now it only remains to apply again Theorem 1.6.2. □

1.6.6. Remark. The condition on the density of μ in Theorem 1.6.5 can be replaced by the condition that $\beta, \gamma \in L_{\rm loc}^1(\Omega, \mu)$. This follows by Theorem 1.5.2.

1.6.7. Corollary. *Let μ be a locally finite Borel measure on Ω satisfying the equation $L_{A,b,c}^*\mu = 0$. Let A^{-1} be locally bounded in Ω with $a^{ij} \in W_{\rm loc}^{p,1}(\Omega)$, where $p > d$, and let either*

(i) $b^i \in L_{\rm loc}^p(\Omega, dx)$, $c \in L_{\rm loc}^{pd/(p+d)}(\Omega, dx)$

or

(ii) $b^i \in L_{\rm loc}^p(\Omega, \mu)$, $c \in L_{\rm loc}^{pd/(p+d)}(\Omega, \mu)$. *Then μ has a density in $W_{\rm loc}^{p,1}(\Omega)$ that is locally Hölder continuous.*

PROOF. It suffices to take $\beta = |b|$, $\gamma = |c|$ and apply Theorem 1.6.5. □

1.6.8. Corollary. *Let μ be a locally finite Borel measure on U_R. Suppose that the mapping A^{-1} is locally bounded on U_R with $a^{ij} \in W^{p,1}_{\rm loc}(U_R)$, where $p > d$, $\partial_{x_i} a^{ij} \in L^p_{\rm loc}(\mu)$, and b^i, $c \in L^p_{\rm loc}(\mu)$. Suppose that*

$$\int_{U_R} \Bigl[a^{ij}\partial_{x_i}\partial_{x_j}\varphi + \partial_{x_i}a^{ij}\partial_{x_j}\varphi + b^i\partial_{x_i}\varphi + c\varphi\Bigr]\,d\mu = 0 \quad \forall\,\varphi \in C_0^\infty(U_R).$$

Then μ has a density in $W^{p,1}_{\rm loc}(U_R)$ that is locally Hölder continuous.

Corollary 1.6.7 can be generalized as follows.

1.6.9. Corollary. *Let $p > d$, let $a^{ij} \in W^{p,1}_{\rm loc}(\Omega)$, b^i, f^i, $c \in L^p_{\rm loc}(\Omega)$, and let A^{-1} be locally bounded in Ω. Assume that μ is a locally finite Borel measure on Ω such that b^i, $c \in L^1_{\rm loc}(\Omega, \mu)$ and, for every function $\varphi \in C_0^\infty(\Omega)$, one has*

$$\int_\Omega \Bigl[a^{ij}\partial_{x_i}\partial_{x_j}\varphi + b^i\partial_{x_i}\varphi + c\varphi\Bigr]\,d\mu = \int_\Omega f^i\partial_{x_i}\varphi\,dx.$$

Then μ has a density in $W^{p,1}_{\rm loc}(\Omega)$.

It is easily seen that in Corollary 1.6.7 one cannot omit the hypotheses that A^{-1} is locally bounded and $a^{ij} \in W^{p,1}_{\rm loc}$. Indeed, if A and b vanish at a point x_0, then Dirac's measure at x_0 satisfies our elliptic equation. In particular, if it is not given in advance that μ is absolutely continuous, then one cannot take an arbitrary Lebesgue version of A. We have already seen in Example 1.4.9 that a solution may fail to be more regular than A. Also, the condition $p > d$ is essential for the membership of μ in a Sobolev class even if $A = {\rm I}$ (see the example below). However, if μ is a probability measure on $\mathbb{R}^d$, then the condition $|b| \in L^2(\mu)$ implies that $\mu = \varrho\,dx$ with $\varrho \in W^{1,1}(\mathbb{R}^d)$ and $|\nabla\varrho|^2/\varrho \in L^1(\mathbb{R}^d)$ (see § 3.1).

1.6.10. Example. Let $d > 3$ and

$$L^*F = \Delta F - {\rm div}\,(Fb) - F = \Delta F + \alpha\partial_{x_i}(x_i|x|^{-2}F) - F,$$

where $\alpha = d - 3$ and

$$b(x) = -\alpha x|x|^{-2} = \nabla(|x|^{-\alpha})/|x|^{-\alpha}.$$

Then the function $F(x) = (e^r - e^{-r})r^{-(d-2)}$, $r = |x|$, is locally Lebesgue integrable and $L^*F = 0$ in the sense of distributions, but F does not belong to $W^{2,1}_{\rm loc}(\mathbb{R}^d)$. Here $|b| \in L^{d-\varepsilon}_{\rm loc}(\mathbb{R}^d)$ for all $\varepsilon > 0$. In a similar way, if the term $-F$ is omitted in the equation above, then the function $F(x) = r^{-(d-3)}$ has the same properties.

PROOF. Observe that $\partial_{x_i}F$, $\partial_{x_i}\partial_{x_j}F$ are locally Lebesgue integrable. Hence the equation $L^*F = 0$ follows easily from the equation

$$f'' + \frac{(d-1+\alpha)}{r}f' + \alpha\frac{d-2}{r^2}f - f = 0$$

on $(0,\infty)$, which is satisfied for the function $f(r) = (e^r - e^{-r})r^{-(d-2)}$. It remains to note that F, ∇F, and ΔF are locally Lebesgue integrable, since $f(r)r^{d-1}$, $f'(r)r^{d-1}$, and $f''(r)r^{d-1}$ are locally bounded, but ∇F is not Lebesgue square-integrable at the origin. If $d \geqslant 6$, then F is also not Lebesgue square-integrable at the origin. In the case without the term $-F$ in the equation similar calculations show that $F(x) = r^{-(d-3)}$ has the same properties. □

1.7. Regularity of solutions to divergence type equations

Here we present several important results related to divergence form equations. First we consider the equation

$$\partial_{x_i}(a^{ij}\partial_{x_j}u) = 0. \tag{1.7.1}$$

Suppose that $\lambda_1 \cdot \mathrm{I} \leqslant A(x) \leqslant \lambda_2 \cdot \mathrm{I}$. What we call its solution can be defined in different ways depending on the properties of the coefficients a^{ij} (see § 1.4). If the functions a^{ij} are merely measurable and locally bounded, then we require that the solution must satisfy the inclusion $u \in W^{1,1}_{\mathrm{loc}}$ and the identity

$$\int_{\mathbb{R}^d} a^{ij}\partial_{x_i}\varphi\partial_{x_j}u\,dx = 0 \quad \forall\,\varphi \in C_0^\infty(\mathbb{R}^d).$$

In this case E. De Giorgi (see De Giorgi [**286**]) showed that any solution in $W^{2,1}_{\mathrm{loc}}$ is locally Hölder continuous. J. Serrin (see Serrin [**842**]) constructed an example showing that the membership in $W^{p,1}_{\mathrm{loc}}$ with $p < 2$ does not ensure the local boundedness of a solution, and his conjecture that in the case of Hölder continuous coefficients any solution in $W^{1,1}_{\mathrm{loc}}$ belongs automatically to $W^{2,1}_{\mathrm{loc}}$ was proved by H. Brezis (see Brezis [**207**], Ancona [**48**]), even under somewhat weaker assumptions: the solution must belong to the class of functions of bounded variation and the coefficients must be Dini continuous. In addition, in the case of merely continuous a^{ij}, Brezis proved that any solution in $W^{p,1}_{\mathrm{loc}}$ with some $p > 1$ belongs to all $W^{q,1}_{\mathrm{loc}}$ with $q < \infty$. However, in Jin, Maz'ya, Van Schaftingen [**492**] an example was constructed showing that for $p = 1$ this is not true.

Let now all functions a^{ij} be locally Lipschitzian. Then equation (1.7.1) can be written as $L^*_{A,b}u = 0$ with $b^i = \sum_{j=1}^d \partial_{x_j}a^{ij}$, hence a priori solutions from L^1_{loc} are admissible. In this situation, in Zhang, Bao [**955**] the conjecture of Brezis was proved that all solutions belong to all classes $W^{q,2}_{\mathrm{loc}}$ with $q < \infty$; the problem was to prove the inclusion in $W^{2,1}_{\mathrm{loc}}$, then the classical results increase the regularity.

Let us consider a general divergence form equation. Set

$$\mathcal{L}u = \partial_{x_i}\big(a^{ij}\partial_{x_j}u - b^iu\big) + \beta^i\partial_{x_i}u + cu, \tag{1.7.2}$$

where functions a^{ij}, b^i, β^i and c are measurable on a bounded open set $\Omega \subset \mathbb{R}^d$.

Suppose that

$$\lambda_1 \cdot \mathrm{I} \leqslant A(x) \leqslant \lambda_2 \cdot \mathrm{I}, \quad \lambda_1, \lambda_2 > 0,$$

$g = |c| + \sum_{i=1}^d \big[|a^i|^2 + |b^i|^2\big] \in L^s(\Omega)$, $f \in L^s(\Omega)$, $g^i \in L^{2s}(\Omega)$, $s > d/2$. We shall say that a function $u \in W^{2,1}_{\mathrm{loc}}(\Omega)$ satisfies the equation

$$\mathcal{L}u = f + \mathrm{div} g$$

if we have the equality

$$\int_\Omega \Big[\langle A\nabla u, \nabla\varphi\rangle\,dx + u\langle b, \nabla\varphi\rangle + \langle\beta, \nabla u\rangle\varphi + cu\varphi\Big]\,dx = \int_\Omega \big[\langle g, \nabla\varphi\rangle - f\varphi\big]\,dx$$

for every function $\varphi \in C_0^\infty(\Omega)$. For such solutions the following important result of Trudinger [**904**] holds (its first assertion was proved already in the first edition of the book Ladyzhenskaya, Ural'tseva [**577**, Theorem 14.1]).

1.7.1. Theorem. *Any solution u has a locally Hölder continuous version and for every ball $U_R(z)$ in Ω and $r < R$ one has the inequality*

$$\sup_{x,y\in U_r(z)} |u(x)-u(y)| \leqslant C_1 r^\alpha \Bigl(\sup_{U_R(z)}|u| + \|f\|_{L^s(\Omega)} + \max_i \|g^i\|_{L^{2s}(\Omega)}\Bigr),$$

where the numbers C_1 and $\alpha>0$ depend only on $d, s, R, \lambda_1, \lambda_2, \|g\|_{L^s(\Omega)}$. If $f=0$ and $g=0$ and $u\geqslant 0$, then u satisfies Harnack's inequality

$$\sup_{x\in U_r(z)} u(x) \leqslant C_2 \inf_{x\in U_r(z)} u(x),$$

where the number C_2 depends only on the same quantities as C_1.

For solutions of the equation $L^*_{A,b,c}\mu = 0$ we obtain the following.

1.7.2. Corollary. *Let μ be a nonnegative locally finite Borel measure on a domain Ω in $\mathbb{R}^d$ satisfying the equation $L^*_{A,b,c}\mu = 0$. Let A^{-1} be locally bounded in Ω with $a^{ij}\in W^{p,1}_{\rm loc}(\Omega)$, where $p>d$, and let $b^i\in L^p_{\rm loc}(\Omega,dx)$, $c\in L^{p/2}_{\rm loc}(\Omega,dx)$. Then the continuous density ϱ of μ has the following property: for every compact set K contained in a connected open set U with compact closure in Ω, one has*

$$\sup_K \varrho \leqslant C \inf_K \varrho,$$

where the number C depends only on the quantities $\|a^{ij}\|_{W^{p,1}(U)}$, $\|b\|_{L^p(U)}$, $\|c\|_{L^p(U)}$, $\inf_U \det A$, and K. In particular, ϱ does not vanish in U if it is not identically zero in U.

The dependence of C on the indicated quantities will be studied in Chapter 3. The assumption that $b^i\in L^p_{\rm loc}(\Omega,dx)$ in Theorem 1.7.2 cannot be replaced by the alternative assumption from Corollary 1.6.7 that $b^i\in L^p_{\rm loc}(\Omega,\mu)$. Indeed, it suffices to take $b=\nabla\varrho/\varrho$ such that ϱ is a probability density which has zeros, but $|b|\in L^p(\mu)$ (for example, we can take ϱ which behaves like $\exp(-x^{-2})$ in a neighborhood of the origin).

1.7.3. Proposition. *Suppose that the hypotheses of the previous corollary are fulfilled and Ω is connected. Let μ be some positive measure on Ω satisfying the equation $L^*_{A,b,c}\mu=0$. Then, any other solution μ_0 can be written as $\mu_0 = f\cdot\mu$, where $f\in W^{p,2}_{\rm loc}(\Omega)$.*

Proof. Suppose first that $d>1$. Then $p>2$. We know that μ and μ_0 have continuous densities ϱ and ϱ_0, respectively, in the class $W^{p,1}_{\rm loc}(\Omega)$ and that ϱ has no zeros in Ω. Set $f=\varrho_0/\varrho$. Then $\mu_0 = f\cdot\mu$ and $f\in W^{p,1}_{\rm loc}(\Omega)$. We have $a^i := \sum_{j=1}^d \partial_{x_j} a^{ij} \in L^p_{\rm loc}(\Omega)$. Set $a:=(a^i)$. Let us verify that f satisfies the elliptic equation

$$a^{ij}\varrho\partial_{x_i}\partial_{x_j} f + \langle \nabla f, 2\varrho a + 2A\nabla\varrho - \varrho b\rangle = 0 \tag{1.7.3}$$

in the sense of weak solutions in the class $W^{p,1}_{\rm loc}(\Omega)$, i.e., in the sense of the identity

$$\int_\Omega \bigl[-\varphi\partial_{x_i}(a^{ij}\varrho)\partial_{x_j} f - \langle \varrho A\nabla f,\nabla\varphi\rangle + \langle\nabla f, 2\varrho a + 2A\nabla\varrho - \varrho b\rangle\varphi\bigr]\,dx = 0$$

for all $\varphi\in C_0^\infty(\Omega)$. This will yield the desired inclusion $f\in W^{p,2}_{\rm loc}(\Omega)$, since we have $\varrho a^{ij}\in W^{p,1}_{\rm loc}(\Omega)$, ϱA is nondegenerate, $a\varrho$, $b\varrho$, $c\varrho\in L^p_{\rm loc}(\Omega)$. In order to establish equality (1.7.3) we observe that the equality $L^*_{A,b,c}\mu_0 = L^*_{A,b,c}\mu = 0$ and the integration by parts formula give the identities

$$\int [-\partial_{x_i}(a^{ij}\varrho f)\partial_{x_j}\varphi + \langle f\varrho b, \nabla\varphi\rangle + c\varrho f\varphi]\, dx = 0, \tag{1.7.4}$$

$$\int [-\partial_{x_i}(a^{ij}\varrho)\partial_{x_j}\varphi + \langle \varrho b, \nabla\varphi\rangle + c\varrho\varphi]\, dx = 0 \tag{1.7.5}$$

for all $\varphi \in C_0^\infty(\Omega)$. Since $a^{ij}, \varrho, f \in W^{p,1}_{\rm loc}(\Omega)$ and $p > 2$, it follows that equality (1.7.5) remains true for all functions φ of the form $\varphi = f\psi$ with $\psi \in C_0^\infty(\Omega)$. This yields the identity

$$\int [-\partial_{x_i}(a^{ij}\varrho) f\partial_{x_j}\varphi - \partial_{x_i}(a^{ij}\varrho)\varphi\partial_{x_j} f + \langle \varrho b, f\nabla\varphi\rangle + \langle \varrho b, \varphi\nabla f\rangle + c\varrho f\varphi]\, dx = 0$$

for all $\varphi \in C_0^\infty(\Omega)$. Subtracting this identity from (1.7.4) and differentiating the products by the Leibniz formula we arrive at (1.7.3). In the case $d = 1$ this reasoning does not apply if $p < 2$, but in this case a simple direct proof works: we have $(Af\varrho)' = f\varrho b + \psi$ and $(A\varrho)' = b\varrho + k$, where ψ is the indefinite integral of $cf\varrho$ and k is constant. Then $f' = (\psi - kf)(A\varrho)^{-1}$. □

In the rest of this section we discuss a priori estimates of solutions on a bounded domain $\Omega \subset \mathbb{R}^d$, which will play an important role in the proofs of the theorems on existence and regularity of solutions. The simplest a priori estimate is obtained by substituting in the identity defining the equation $\mathcal{L}u = f + {\rm div} g$ the function $u\varphi$, where the function $\varphi \in C_0^\infty(\Omega)$ equals 1 on Ω' and $\overline{\Omega'} \subset \Omega$. Suppose for simplicity that $b = \beta = c = 0$ and $f, g^i \in L^2(\Omega)$. Then we immediately obtain

$$\|u\|_{W^{2,1}(\Omega')} \leqslant C\big(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}\big).$$

The next result generalizes this estimate to the case of the operator $\mathcal{L}$ of the form (1.7.2) with all coefficients and the space $W^{p,1}$.

For functions on Ω we shall write that $a^{ij} \in VMO$ if a^{ij} extends to all of $\mathbb{R}^d$ as a function in VMO. We recall that the membership of a function a^{ij} in the class VMO is expressed in terms of a certain function ω denoted below by the symbol ω_A (see § 1.2). This condition is weaker than the uniform continuity. In the case of uniformly continuous coefficients we can take for ω_A the modulus of continuity. For Hölder continuous coefficients, ω_A can be easily expressed in terms of their Hölder norms.

Suppose that the coefficients a^{ij}, b, β and c satisfy the following conditions with some numbers $\lambda_1, \lambda_2 > 0$:

$$\begin{aligned} &a^{ij} \in VMO,\ a^{ij} = a^{ji},\ \lambda_1 \cdot {\rm I} \leqslant A(x) \leqslant \lambda_2 \cdot {\rm I},\\ &b^i, \beta^i \in L^d(\Omega),\ c \in L^{d/2}(\Omega) \text{ if } 2 \leqslant q < d,\\ &b^i, \beta^i \in L^s(\Omega),\ c \in L^{s/2}(\Omega) \text{ with some } s > d \text{ if } 2 \leqslant q = d,\\ &b^i, \beta^i \in L^q(\Omega),\ c \in L^{dq/(d+q)}(\Omega) \text{ if } q > d. \end{aligned} \tag{1.7.6}$$

As above, let $U(x, r)$ denote the ball of radius r centered at x and let $|U(x, r)|$ be its volume.

For $\gamma \geqslant 1, r > 0$ and $\eta \in L^\gamma(\Omega)$ we set

$$\Theta_{\gamma,\eta}(r) = \sup_{x\in\Omega} \biggl(\int_{U(x,r)\cap\Omega} |\eta(y)|^\gamma\, dy\biggr)^{1/\gamma}.$$

We recall Young's inequality. If $x > 0$, $y > 0$, $\varepsilon > 0$, $\delta > 0$, $\gamma > 0$, and $\delta^{-1} + \gamma^{-1} = 1$, then

$$xy \leqslant \varepsilon \frac{x^\delta}{\delta} + \varepsilon^{1-\gamma} \frac{y^\gamma}{\gamma}. \tag{1.7.7}$$

1.7.4. Theorem. *Suppose that the coefficients of the operator $\mathcal{L}$ satisfy conditions (1.7.6) and $q \geqslant 2$. Let $g \in L^q(\Omega)$, $f \in L^p(\Omega)$, where $p = dq/(d+q)$ if $q \neq d$ and $p > d/2$ if $q = d$. If a function $u \in W^{q,1}(\Omega)$ is a solution to the equation $\mathcal{L}u = f + \operatorname{div} g$ on Ω, then for every open set Ω' with compact closure in Ω we have the estimate*

$$\|u\|_{W^{q,1}(\Omega')} \leqslant C\big(\|u\|_{L^1(\Omega)} + \|g\|_{L^q(\Omega)} + \|f\|_{L^p(\Omega)}\big),$$

where the constant C depends only on $\Omega, \Omega', d, \lambda_1, \lambda_2, q$, ω_A and on the rate of convergence of the quantities $\Theta_{d,b}(r)$, $\Theta_{d,\beta}(r)$, $\Theta_{d/2,c}(r)$ to zero as $r \to 0$ in case $2 \leqslant q < d$ and on the number s and the norms $\|b^i\|_{L^s}$, $\|\beta^i\|_{L^s}$, $\|c\|_{L^{s/2}}$ in case $q = d$, and, finally, on the norms $\|b^i\|_{L^q}$, $\|\beta^i\|_{L_q}$, $\|c\|_{L^{dq/(d+q)}}$ in case $q > d$.

Proof. Let $U = U(a,r)$ be a ball with closure in Ω and $0 < r < 1$ and let a function ζ in $C_0^\infty(U)$ be such that $0 \leqslant \zeta \leqslant 1$, $\zeta(x) > 0$ in U and $\zeta(x) = 1$ on the twice smaller ball $U(a, r/2)$, and also

$$J(s) := \sup_x |\nabla\zeta(x)|\zeta(x)^{-s} + \sup_{i,j,x} |\partial_{x_i}\partial_{x_j}\zeta(x)|\zeta(x)^{-s} < \infty, \quad 0 < s < 1.$$

These conditions can be easily ensured by taking $\zeta(x) = \psi(|x|/r)$, $\psi \in C_0^\infty(\mathbb{R})$, $0 \leqslant \psi \leqslant 1$, $\psi(y) = 0$ if $|y| \geqslant 1$, $\psi(y) > 0$ if $|y| < 1$, $\psi(y) = 1$ if $|y| \leqslant 1/2$ and $\psi(y) = \exp\big[(y^2-1)^{-1}\big]$ near the points -1 and 1. We shall use as a cut-off function only this ζ.

1. We first prove the theorem for $q = 2$. To this end, we shall estimate the norm $\|u\zeta\|_{W^{2,1}(U)}$. Substituting in the identity defining the equation $\mathcal{L}u = f + \operatorname{div} g$ the function $\varphi = u\zeta^2$ and integrating, we have

$$\int_\Omega \big[a^{ij}u_{x_j}u_{x_i}\zeta^2\,dx + 2a^{ij}u_{x_j}u\zeta\zeta_{x_i} + b^i u_{x_i}u\zeta^2 + 2b^i u^2\zeta_{x_i}\zeta + g^i u_{x_i}\zeta^2 + 2g^i u\zeta\zeta_{x_i} + c^i u_{x_i}u\zeta^2 + hu^2\zeta^2 + fu\zeta^2\big]\,dx = 0.$$

Using inequality (1.7.7) with $\alpha = \beta = 2$ and a sufficiently small $\varepsilon > 0$ and the first condition in (1.7.6), we find that

$$\|\zeta\nabla u\|_{L^2}^2 \leqslant C_1\big(\|u\nabla\zeta\|_{L^2}^2 + \|bu\zeta\|_{L^2}^2 + \|\beta u\zeta\|_{L^2}^2 + \|\sqrt{|c|}u\zeta\|_{L^2}^2 + \|g\zeta\|_{L^2}^2 + \|fu\zeta^2\|_{L^1}\big), \tag{1.7.8}$$

where $C_1 = C_1(d, m, M)$. Let us estimate the summands in the right-hand side of this inequality separately. Set $t = d$ if $d > 2$ and $t = (s+2)/2$ if $d = 2$. It is clear that $2 < t < s$ if $n = d$. In order to estimate $\|\eta u\zeta\|_{L^2}$, where η equals b, c or $\sqrt{|h|}$, we apply Hölder's inequality with the exponent $t/2$ and the Sobolev embedding $W^{2,1} \subset L^{2t/(t-2)}$. We have

$$\|\eta u\zeta\|_2 \leqslant \|\eta\|_t \|u\zeta\|_{2t/(t-2)} \leqslant C(d,t)\|\eta\|_t\|u\zeta\|_{2,1}, \tag{1.7.9}$$

where $\|\eta\|_t = \Theta_{d,\eta}(r)$ if $t = d$, $\|\eta\|_t \leqslant |U(a,r)|^{1-t/s}\|\eta\|_s$ if $2 < t < s$. Applying Hölder's inequality with the exponents p, $p' = p(p-1)^{-1}$, where $p = 2d(d+2)^{-1}$

if $d > 2$ and $p > 1$ if $d = 2$, inequality (1.7.7) with $\gamma = \delta = 2$ and the Sobolev embedding $W^{2,1} \subset L^{p'}$, we obtain

$$\begin{aligned}\|fu\zeta^2\|_1 &\leqslant \|f\zeta\|_p\|u\zeta\|_{p'} \leqslant C(d,p)\|f\zeta\|_p\|u\zeta\|_{2,1}\\ &\leqslant \varepsilon^2\|u\zeta\|_{2,1}^2 + \varepsilon^{-2}C(d,p)^2\|f\zeta\|_p^2.\end{aligned}$$

Let us estimate $\|u\nabla\zeta\|_{L^2}$. Let $l = 4(d+2)^{-1}$ if $d > 2$ and $0 < l < 1$ if $d = 2$. By the Sobolev embedding theorem $W^{2,1} \subset L^{(2-l)/(1-l)}$. Applying Hölder's inequality, we arrive at the inequalities

$$\begin{aligned}\|u\nabla\zeta\|_2 = \|u\zeta^{(2-l)/2}\zeta^{-(2-l)/2}\nabla\zeta\|_2 &\leqslant J(1-l/2)\biggl(\int_U |u\zeta|^{2-l}|u|^l\,dx\biggr)^{1/2}\\ \leqslant J(1-l/2)\|u\|_1^{l/2}\|u\zeta\|_{(2-l)/(1-l)}^{(2-l)/2} &\leqslant J(1-l/2)C(n,l)\|u\|_1^{l/2}\|u\zeta\|_{2,1}^{(2-l)/2}.\end{aligned}$$

Young's inequality (1.7.7) with $\alpha = 2/l$ and $\beta = 2/(2-l)$ gives

$$\|u\nabla\zeta\|_2 \leqslant \varepsilon\|u\zeta\|_{2,1} + \varepsilon^{-l/(2-l)}C(d,l)\|u\|_{L^1}.$$

According to estimate (1.7.9) with $\eta = 1$, we have

$$\|u\zeta\|_{L^2(U(a,r))} \leqslant C(d,t)|U(a,r)|^{1/t}\|u\zeta\|_{2,1}.$$

Substituting the obtained estimates in (1.7.8) and taking into account that by the Leibniz formula

$$\|u\zeta\|_{2,1} \leqslant \|u\nabla\zeta\|_2 + \|\zeta\nabla u\|_2 + \|u\zeta\|_2,$$

we find that

$$\begin{aligned}\|u\zeta\|_{W^{2,1}(U)} &\leqslant C_2\varepsilon^{-1}\bigl(\|u\|_{L^1(U)} + \|g\|_{L^2(U)} + \|f\|_{L^p(U)}\bigr)\\ &\quad + C_2\bigl(\|b\|_{L^t(U)} + \|\beta\|_{L^t(U)} + \|c\|_{L^{t/2}(U)}^{1/2} + |U(a,r)|^{1/t} + \varepsilon\bigr)\|u\zeta\|_{W^{2,1}(U)},\end{aligned}$$

where $C_2 = C_2(d,p,t,m,M)$. Choosing $r > 0$ and $\varepsilon > 0$ such that

$$C_2\Bigl(\|b\|_{L^t(U(a,r))} + \|c\|_{L^t(U(a,r))} + \|h\|_{L^{t/2}(U(a,r))}^{1/2} + |U(a,r)|^{1/t} + \varepsilon\Bigr) < \frac{1}{2},$$

we arrive at the estimate $\|u\zeta\|_{2,1} \leqslant 2C_2\varepsilon^{-1}\bigl(\|u\|_1 + \|g\|_2 + \|f\|_p\bigr)$. By using a smooth partition of unity associated with a finite covering of the domain Ω' by balls of radius r with closure in Ω, we obtain the required estimate.

2. Let us consider the case $2 < q < d$. As above, we start with an estimate of the quantity $\|u\zeta\|_{W^{q,1}(U)}$. We may assume that A is extended to the whole space $\mathbb{R}^d$ with preservation of all conditions. By Theorem 1.2.1, for a sufficiently large number $\lambda = \lambda(d,q,m,M,\omega_A) > 0$ there is a number $N = N(d,q,m,M,\omega_A) > 1$ such that for every function $w \in W^{q,1}(\mathbb{R}^d)$ with compact support the generalized function $\mathcal{A}w = \partial_{x_i}(a^{ij}\partial_{x_j}w) - \lambda w$ satisfies the inequality

$$\|w\|_{W^{q,1}(\mathbb{R}^d)} \leqslant N\|\mathcal{A}w\|_{W^{q,-1}(\mathbb{R}^d)}. \tag{1.7.10}$$

The function $w = \zeta u$ satisfies the equation

$$\begin{aligned}\mathcal{A}w = -\lambda\zeta u + a^{ij}\partial_{x_i}\zeta\partial_{x_j}u + \partial_{x_i}(a^{ij}u\partial_{x_j}\zeta) - \zeta\partial_{x_i}(b^iu)&\\ + \beta^i\partial_{x_i}u\zeta + cu\zeta - \zeta\partial_{x_i}g^i + f\zeta.&\end{aligned} \tag{1.7.11}$$

According to (1.7.10) it suffices to estimate the norm of the right-hand side in $W^{q,-1}(U)$ through the parameters indicated in the theorem. Let us rewrite the right-hand side of the last equality in the following form convenient for our later use:

$$\mathcal{A}w = \partial_{x_i}(a^{ij}u\partial_{x_j}\zeta - b^i u\zeta - g^i\zeta) - \lambda\zeta u + a^{ij}\partial_{x_i}\zeta\partial_{x_j}u + b^i u\partial_{x_i}\zeta + \beta^i\partial_{x_i}(u\zeta) - \beta^i u\partial_{x_i}\zeta + cu\zeta + g^i\partial_{x_i}\zeta + f\zeta. \tag{1.7.12}$$

We recall that $\|\partial_{x_i}\eta^i\|_{q,-1} \leqslant \|\eta\|_q$ for every vector function $\eta \in L^q$ and by the embedding theorem we have $L^p \subset W^{q,-1}$ if $p = dq(d+q)^{-1}$. Therefore, the norm of the right-hand side is estimated by

$$\|a^{ij}u\partial_{x_j}\zeta\|_q + \|b^i u\zeta\|_q + \|g^i\zeta\|_q + C(d,p)\bigl[\lambda\|\zeta u\|_p + \|a^{ij}\partial_{x_i}\zeta\partial_{x_j}u\|_p + \|b^i u\partial_{x_i}\zeta\|_p + \|\beta^i\partial_{x_i}(u\zeta)\|_p + \|\beta^i u\partial_{x_i}\zeta\|_p + \|hu\zeta\|_p + \|g^i\partial_{x_i}\zeta\|_p + \|f\zeta\|_p\bigr], \tag{1.7.13}$$

where $C = C(d,p)$ is the constant from the embedding theorem. We observe that

$$\|a^{ij}u\partial_{x_j}\zeta\|_q \leqslant M\|u\nabla\zeta\|_q, \quad \|g^i\partial_{x_i}\zeta\|_p \leqslant \|\nabla\zeta\|_d\|g\|_q.$$

Applying Hölder's inequality with the exponents d/q, $d/(d-q)$ and the exponents $(d+q)/q$, $(d+q)/d$, we obtain

$$\|b^i u\zeta\|_q + C(d,p)\bigl[\|b^i u\partial_{x_i}\zeta\|_p + \|c^i\partial_{x_i}(u\zeta)\|_p + \|\beta^i u\partial_{x_i}\zeta\|_p\bigr] \leqslant C(d,q,p)\bigl[(\|b\|_d + \|c\|_d)\|u\zeta\|_{q,1} + (\|b\|_d + \|\beta\|_d)\|u\nabla\zeta\|_q\bigr].$$

Hölder's inequality with the exponents $(d+q)/2q$, $(d+q)/(d-q)$ and the Sobolev embedding theorem give

$$\lambda\|\zeta u\|_p + \|cu\zeta\|_p \leqslant C(d,q)\Bigl(\lambda|U(a,r)|^{2q/(d+q)} + \|c\|_{d/2}\Bigr)\|u\zeta\|_{q,1}.$$

It remains to estimate $\|u\nabla\zeta\|_{L^q}$. Let $l = q^2/(dq+q-d)$. We observe that $0<l<1$ and $(q-l)/(1-l) = dq/(d-q)$. By the Sobolev embedding theorem and Hölder's inequality we have

$$\|u\nabla\zeta\|_q = \|\zeta^{1-l/q}u\zeta^{-1+l/q}\nabla\zeta\|_q \leqslant J\bigl(1-l/q\bigr)\biggl(\int_U |u\zeta|^{q-l}|u|^l\,dx\biggr)^{1/q} \leqslant J\bigl(1-l/q\bigr)\|u\|_1^{l/q}\|u\zeta\|_{(q-l)/(1-l)}^{(q-l)/q} \leqslant J\bigl(1-l/q\bigr)C(d,l,q)\|u\|_1^{l/q}\|\zeta u\|_{q,1}^{(q-l)/q}.$$

According to Young's inequality (1.7.7) with $\alpha = q/l, \beta = q/(q-l)$, we obtain

$$\|u\nabla\zeta\|_q \leqslant \varepsilon\|u\zeta\|_{q,1} + C(d,l,q,\varepsilon)\|u\|_1.$$

Similarly to the case $q=2$ we pick $\varepsilon>0$ and $r>0$ sufficiently small. We obtain

$$\|u\zeta\|_{q,1} \leqslant C_3\bigl(\|\nabla u\nabla\zeta\|_p + \|u\|_1 + \|g\|_q + \|f\|_p\bigr).$$

By using a partition of unity we arrive at the estimate

$$\|u\|_{W^{q,1}(\Omega')} \leqslant C_4\bigl(\|u\|_{W^{p,1}(\Omega'')} + \|u\|_{L^1(\Omega)} + \|g\|_{L^q(\Omega)} + \|f\|_{L^p(\Omega)}\bigr)$$

for any domains $\Omega' \subset \Omega'' \subset \Omega$ with $\overline{\Omega'} \subset \Omega''$, $\overline{\Omega''} \subset \Omega$. The constant C_4 depends only on $\Omega', \Omega'', \Omega$ and the quantities indicated in the hypotheses of the theorem. Let us set

$$p_k = dp_{k-1}(d+p_{k-1}),\ p_0 = q,\ p_1 = p.$$

Since $p_k/p_{k-1} \leqslant d/(d+1)$, we can find a natural number K such that $p_K \leqslant 2$. Let $\{\Omega_k\}_{0\leqslant k\leqslant K}$ be a family of open sets such that $\Omega' \subset \Omega_k \subset \Omega$ and $\overline{\Omega_k} \subset \Omega_{k+1}$. Then, as shown above, whenever $0 \leqslant k \leqslant K$, we have the estimate

$$(1.7.14)\quad \|u\|_{W^{p_{k-1},1}(\Omega_{k-1})} \leqslant C(k)\big(\|u\|_{W^{p_k,1}(\Omega_k)} + \|u\|_{L^1(\Omega)} + \|g\|_{L^q(\Omega)} + \|f\|_{L^p(\Omega)}\big).$$

Since $p_k \leqslant 2$ for $k = K$, we estimate $\|u\|_{W^{p_K,1}}$ via $\|u\|_{W^{2,1}}$ and use the already established estimate for the case $q = 2$. Repeatedly applying inequality (1.7.14) for each k, in finitely many steps we obtain the required estimate for $\|u\|_{W^{p_0,1}(\Omega')}$, which completes the proof in the case $2 < q < d$.

Let us consider the case $q > d$. As in the previous case, the norm of the right-hand side in equality (1.7.11) is estimated by (1.7.13). Applying condition (1.7.6) and Hölder's inequality with the exponents $(d+q)/q$ and $(d+q)/d$, we obtain the following inequalities:

$$\begin{aligned}\|a^{ij}u\partial_{x_j}\zeta\|_q + C(d,p)\big[\|b^i u\partial_{x_i}\zeta\|_p + \|\beta^i u\partial_{x_i}\zeta\|_p\big] &\leqslant C(d,p)(M + \|b\|_d + \|\beta\|_d)\|u\nabla\zeta\|_q \\ &\leqslant C(d,p)\big(M + |U(a,r)|^{(q-d)/d}\|b\|_q + |U(a,r)|^{(q-d)/d}\|\beta\|_q\big)\|u\nabla\zeta\|_q.\end{aligned}$$

Let us estimate $\|u\nabla\zeta\|_q$. Since $q > d$, by the Sobolev embedding theorem

$$\|u\zeta\|_{L^\infty(U(a,r))} \leqslant C(n,p)r^{(q-d)/q}\|u\zeta\|_{W^{q,1}(U(a,r))}.$$

Then

$$\begin{aligned}\|u\nabla\zeta\|_q &= \bigg(\int_U |u\zeta|^{q-1}|u|\,|\nabla\zeta|^q\zeta^{1-q}\,dx\bigg)^{1/q} \\ &\leqslant C(d,q)J\big((q-1)/q\big)\|u\|_{q,1}^{(q-1)/q}\|u\|_1^{1/q} \leqslant \varepsilon\|u\|_{q,1} + \varepsilon^{-1}\widetilde{C}(d,q)\|u\|_1.\end{aligned}$$

Applying again the Sobolev embedding theorem, we obtain

$$\|b^i u\zeta\|_q + C(d,p)\big[\lambda\|u\zeta\|_p + \|cu\zeta\|_p\big] \leqslant C(d,q,p)r^{(q-d)/q}(1 + \|b\|_q + \|c\|_p)\|u\zeta\|_{q,1}.$$

It remains to observe that

$$\|\beta^i\partial_{x_i}(u\zeta)\|_p \leqslant |U(a,r)|^{(q-d)/q}\|\beta\|_q\|u\|_{q,1},\ \ \|g^i\partial_{x_i}\zeta\|_p \leqslant \|g\|_q\|\nabla\zeta\|_d.$$

Similarly to the previous cases we pick $\varepsilon > 0$ and $r > 0$ sufficiently small and obtain that

$$\|u\zeta\|_{q,1} \leqslant C_5\big(\|\nabla u\nabla\zeta\|_p + \|u\|_1 + \|g\|_q + \|f\|_p\big).$$

By using a partition of unity, we arrive at the estimate

$$\|u\|_{W^{q,1}(\Omega')} \leqslant C_6\big(\|u\|_{W^{p,1}(\Omega'')} + \|u\|_{L^1(\Omega)} + \|g\|_{L^q(\Omega)} + \|f\|_{L^p(\Omega)}\big)$$

for any domains $\Omega' \subset \Omega'' \subset \Omega$ with $\overline{\Omega'} \subset \Omega''$, $\overline{\Omega''} \subset \Omega$. Note that $p < d$ and we can use the previous step for estimating the norm $\|u\|_{p,1}$. The proof in the case $q = d$ is similar. □

It is clear that the norm $\|u\|_{L^1(\Omega)}$ can be replaced (by Hölder's inequality) with the norm $\|u\|_{L^q(\Omega)}$.

The next assertion describes the dependence of the constant from the theorem on the diameter of the domain in the simplest case where $b = \beta = 0$ and $c = f = 0$.

1.7.5. Corollary. *Let $\Omega = U(z,R)$, where $R<1$, and let $u \in W^{q,1}_{\rm loc}\bigl(U(z,R)\bigr)$ satisfy the equation $\partial_{x_i}(a^{ij}\partial_{x_i}u - g^i) = 0$, where $q>d$ and* (1.7.6) *is fulfilled. Then, whenever $\lambda<1$, we have*

$$\|u\|_{W^{q,1}(U(z,\lambda R))} \leqslant R^{-1}C\bigl(\|u\|_{L^q(U(z,R))} + R\|g\|_{L^q(U(z,R))}\bigr),$$

where $C = C(d,q,\alpha,\omega_A,\lambda)$.

PROOF. Let us change the coordinates $x = z+Ry$. Then the obtained function $v(y) = u(z+Ry)$ satisfies in $U(0,1)$ the equation

$$\partial_{y_j}\bigl(a^{ij}(z+Ry)\partial_{y_i}v(y) + Rg^i(z+Ry)\bigr) = 0.$$

By assumption, $a^{ij} \in VMO$ with the function ω_A, so we may assume that the function ω_A does not change, since $R<1$. Then for the function v we have

$$\|v\|_{W^{q,1}(U(0,\lambda))} \leqslant C(d,q,\alpha,\|A\|_{C^{0,\delta}},\lambda)\bigl(\|v\|_{L^q(U(0,1))} + R\|g\|_{L^q(U(0,1))}\bigr).$$

Making the inverse change of coordinates and taking into account that $R<1$ we obtain

$$\begin{aligned}\|v\|_{W^{q,1}(U(0,\lambda))} &= R^{-d/q}\|u\|_{L^q(U(z,\lambda R))} + R^{1-d/q}\|\nabla u\|_{L^q(U(z,\lambda R))}\\ &\geqslant R^{1-d/q}\|u\|_{W^{q,1}(U(z,\lambda R))}.\end{aligned}$$

Similarly, we have

$$\begin{aligned}\|v\|_{L^q(U(0,1))}+R\|h\|_{L^q(U(0,1))} &= R^{-d/q}\|u\|_{L^q(U(z,R))}+R^{1-d/q}\|g\|_{L^q(U(z,R))}\\ &= R^{-d/q}\bigl(\|u\|_{L^q(U(z,R))} + R\|g\|_{L^q(U(z,R))}\bigr).\end{aligned}$$

On the basis of these estimates we obtain

$$\|u\|_{W^{q,1}(U(z,\lambda R))} \leqslant C(d,q,\alpha,\|A\|_{C^{0,\delta}},\lambda)R^{-1}\bigl(\|u\|_{L^q(U(z,R))} + R\|g\|_{L^q(U(z,R))}\bigr),$$

as required. □

From Theorem 1.7.4 we deduce an estimate on the whole domain Ω for the function $u \in W^{q,1}_0(\Omega)$ in the case where the boundary $\partial\Omega$ is sufficiently regular.

Recall that the boundary $\partial\Omega$ is of class C^1 if it can be locally made flat by a diffeomorphism of class C^1.

1.7.6. Corollary. *Suppose that the coefficients of the operator $\mathcal{L}$ satisfy conditions* (1.7.6) *and $q \geqslant 2$. Let also $g \in L^q(\Omega)$, $f \in L^p(\Omega)$, where $p = dq/(d+q)$ if $q \neq d$ and $p > d/2$ if $q = d$. If Ω is a bounded domain with boundary of class C^1 and a function $u \in W^{q,1}_0(\Omega)$ is a solution of the equation $\mathcal{L}u = f + {\rm div} g$ in Ω, then*

$$\|u\|_{W^{q,1}_0(\Omega)} \leqslant C\bigl(\|u\|_{L^1(\Omega)} + \|g\|_{L^q(\Omega)} + \|f\|_{L^p(\Omega)}\bigr),$$

where the constant C depends only on the same quantities as in Theorem 1.7.4.

PROOF. It suffices to consider the following situation: $u \in W^{q,1}(K)$, where K is a cube one of the faces of which belongs to the hyperplane $\{x_d = 0\}$ and the cube itself belongs to the open half-space $\{x_d > 0\}$. Suppose that $u\psi \in W^{q,1}_0(\Omega)$ for every function $\psi \in C^\infty(\overline{K})$ vanishing in a neighborhood of every face excepting $x_d = 0$. This actually means that $u = 0$ when $x_d = 0$. Set $y = (x_1,\dots,x_{d-1})$. For any $x_d < 0$ we set $u(y,x_d) = -u(y,-x_d)$ and

$$A(y,x_d) = A(y,-x_d),\ b(y,x_d) = -b(y,-x_d),\ \beta(y,x_d) = -\beta(y,-x_d),$$

$$c(y,x_d) = c(y,-x_d),\ f(y,x_d) = f(y,-x_d),\ g^i(y,x_d) = -g^i(y,-x_d).$$

Let K' be the cube obtained by reflecting K with respect to $\{x_d = 0\}$. The function u belongs to $W_0^{1,q}(K \cup K')$ and satisfies the equation $\mathcal{L}u = f + \mathrm{div} g$ in $K \cup K'$. Let K_1 be a cube inside of K such that all faces of K_1 are strictly inside of K, excepting one face belonging to the hyperplane $\{x_d = 0\}$. By Theorem 1.7.4 we have

$$\|u\|_{W^{q,1}(K_1)} \leqslant C\bigl(\|u\|_{L^1(K\cup K')} + \|g\|_{L^q(K\cup K')} + \|f\|_{L^p(K\cup K')}\bigr).$$

Thus, we obtain the required estimate up to the boundary $\{x_d = 0\}$. □

Existence of solutions to divergence form equations is discussed in Chapter 2, the increasing of the Sobolev regularity is considered in § 1.8(ii).

1.8. Complements, comments, and exercises

(i) Fractional Sobolev classes (43). (ii) Increasing Sobolev regularity of solutions (47). (iii) Renormalized solutions (48). (iv) Generalizations of the maximum principle of A.D. Aleksandrov and k-Hessians (49). Comments (50). Exercises (53).

1.8(i). Fractional Sobolev classes

In the case where the diffusion matrix A is infinitely differentiable a somewhat more special result holds in terms of the scale of fractional Sobolev classes. Given $p \in (1, +\infty)$ and $s \in \mathbb{R}$, we set

$$H^{p,s}(\mathbb{R}^d) := (I - \Delta)^{-s/2}\bigl(L^p(\mathbb{R}^d)\bigr), \quad \|f\|_{p,s} = \|(I - \Delta)^{s/2} f\|_p,$$

where the operator $(I - \Delta)^{-s/2}$ is applied in the sense of the space of tempered distributions $\mathcal{S}'(\mathbb{R}^d)$; it can be defined via the Fourier transform by using the operator of multiplication by the function $(1 + |x|^2)^{-s/2}$. If $s \geqslant 0$, then the space $H^{p,s}(\mathbb{R}^d)$ is continuously embedded into $L^p(\mathbb{R}^d)$. For $s \in \mathbb{N}$ the class $H^{p,s}(\mathbb{R}^d)$ coincides with $W^{p,s}(\mathbb{R}^d)$ and the respective norms are equivalent. The class $H^{p,s}_{\mathrm{loc}}(\mathbb{R}^d)$, where $s \in \mathbb{R}$, $p > 1$, consists of all functions f such that $\varphi f \in H^{p,s}(\mathbb{R}^d)$ for all $\varphi \in C_0^\infty(\mathbb{R}^d)$.

In the proofs below we use the following well-known lemma. Let Ω be a domain in $\mathbb{R}^d$ and A a mapping on Ω with values in the space of positive symmetric operators on $\mathbb{R}^d$.

1.8.1. Lemma. *Suppose that $a^{ij} \in C^\infty(\Omega)$ and $\det A > 0$.*

(i) *Let $r \in (-\infty, \infty)$ and $p > 1$. If u is a distribution such that $L_A u \in H^{p,r}_{\mathrm{loc}}(\Omega)$, then $u \in H^{p,r+2}_{\mathrm{loc}}(\Omega)$; also if $u \in H^{p,r}_{\mathrm{loc}}(\Omega)$, then $\partial_{x_i} u \in H^{p,r-1}_{\mathrm{loc}}(\Omega)$, $1 \leqslant i \leqslant d$.*

(ii) *We have $H^{p,1}_{\mathrm{loc}}(\Omega) \subset L^{dp/(d-p)}_{\mathrm{loc}}(\Omega)$ and $L^p_{\mathrm{loc}}(\Omega) \subset H^{dp/(d-p),-1}_{\mathrm{loc}}(\Omega)$ whenever $1 < p < d$, and $H^{p,1}_{\mathrm{loc}}(\Omega) \subset C^{1-d/p}_{\mathrm{loc}}(\Omega)$ if $p > d$, so that in the latter case all functions in $H^{p,1}_{\mathrm{loc}}(\Omega)$ are locally bounded. In addition, whenever $q > p > 1$, we have the inclusion $L^p_{\mathrm{loc}}(\Omega) \subset H^{q,d/q-d/p}_{\mathrm{loc}}(\Omega)$.*

(iii) *If μ is a locally bounded Radon measure on Ω, then $\mu \in H^{p,-m}_{\mathrm{loc}}(\Omega)$, whenever $p > 1$ and $m > d(1 - 1/p)$.*

Proof. Assertion (i) is well-known: its first statement is a well-known elliptic regularity result (see Taylor [**894**, Chapter III]) and the second statement follows from the boundedness of Riesz's transforms. Assertion (ii) is just the Sobolev embedding theorem (mentioned in § 1.1 for $H^{p,1}$). Assertion (iii) follows from this embedding theorem, because for any regular sub-domain U of Ω one has the embedding $H^{q,m}(U) \subset C(\overline{U})$ if $qm > d$, whence by duality we obtain that the space $H^{q/(q-1),-m}(U) = [H_0^{q,m}(U)]^*$ contains all finite measures on U. □

We formulate the following result for $d > 1$ just because the case $d = 1$ is elementary and has already been discussed. In addition, we include in the formulation some assertions which follow also from the already mentioned results (but the proof we give is direct and does not use the results above).

1.8.2. Theorem. *Under the same assumptions about A as in the lemma, let $d \geqslant 2$ and let μ and ν be Radon measures on Ω (possibly signed). Let a mapping $b = (b^i)\colon \Omega \to \mathbb{R}^d$ and a function $c\colon \Omega \to \mathbb{R}$ be such that $|b|,\ c \in L^1_{\rm loc}(\Omega, \mu)$. Suppose that $L^*_{A,b,c}\mu = \nu$. Then the following assertions are true.*

(i) *One has $\mu \in H^{p,1-d(p-1)/p-\varepsilon}_{\rm loc}(\Omega)$ for any $p \geqslant 1$ and any $\varepsilon > 0$. Here $1 - d(p-1)/p > 0$ if $p \in \bigl[1, d/(d-1)\bigr)$ and, in particular, μ admits a density $F \in L^p_{\rm loc}(\Omega)$ for any $p \in \bigl[1, d/(d-1)\bigr)$.*

(ii) *If $|b| \in L^\gamma_{\rm loc}(\Omega, \mu)$, $c \in L^{\gamma/2}_{\rm loc}(\Omega, \mu)$ and $\nu \in L^{d/(d-\gamma+2)}_{\rm loc}(\Omega)$ where $d \geqslant \gamma > 1$, then $F := d\mu/dx \in H^{p,1}_{\rm loc}(\Omega)$ for any $p \in [1, d/(d-\gamma+1))$. In particular, $F \in L^p_{\rm loc}(\Omega)$ for any $p \in \bigl[1, d/(d-\gamma)\bigr)$, where we set $d/(d-\gamma) := \infty$ if $\gamma = d$.*

(iii) *If $\gamma > d$ and either*

(a) *$|b| \in L^\gamma_{\rm loc}(\Omega)$ and $c, \nu \in L^{\gamma d/(d+\gamma)}_{\rm loc}(\Omega)$,*

or

(b) *$|b| \in L^\gamma_{\rm loc}(\Omega, \mu)$, $c \in L^{\gamma d/(d+\gamma)}_{\rm loc}(\Omega, \mu)$, and $\nu \in L^{\gamma d/(d+\gamma)}_{\rm loc}(\Omega)$,*

then μ admits a density $F \in H^{\gamma,1}_{\rm loc}(\Omega)$. In particular, $F \in C^{1-d/\gamma}_{\rm loc}(\Omega)$.

PROOF. (i) We have in the sense of distributions

$$L_A\mu = \partial_{x_i}\bigl((b^i - \partial_{x_j}a^{ij})\mu\bigr) - \partial_{x_i}a^{ij}\partial_{x_j}\mu - c\mu + \nu \tag{1.8.1}$$

on Ω. Here by Lemma 1.8.1(iii) the right-hand side belongs to $H^{p,-m-1}_{\rm loc}(\Omega)$ if $m > d(1-1/p)$. By Lemma 1.8.1(i) we conclude $\mu \in H^{p,-m+1}_{\rm loc}(\Omega)$, which leads to the result after substituting $m = d(1-1/p) + \varepsilon$.

Before we prove (ii) and (iii), we need some preparations. Fix $p_1 > 1$ and assume that $F = d\mu/dx \in L^{p_1}_{\rm loc}(\Omega)$. Such a number p_1 exists by (i). Set

$$r := r(p_1) := \frac{\gamma p_1}{\gamma - 1 + p_1} \tag{1.8.2}$$

and observe that owing to the inequalities $1 < \gamma$ and $p_1 > 1$, we have $1 < r < \gamma$. Next, starting with the formula

$$|bF|^r = (|b||F|^{1/\gamma})^r |F|^{r-r/\gamma}$$

and using Hölder's inequality (with $s = \gamma/r > 1$ and $t := s/(s-1) = \gamma/(\gamma-r)$ and the relations $|b|\,|F|^{1/\gamma} \in L^\gamma_{\rm loc}(\Omega)$ and $F \in L^{p_1}_{\rm loc}(\Omega)$, we obtain that $b^iF \in L^r_{\rm loc}(\Omega)$. By Lemma 1.8.1(i) one has

$$b^iF \in H^{r,0}_{\rm loc}(\Omega), \quad (b^iF)_{x_i} \in H^{r,-1}_{\rm loc}(\Omega). \tag{1.8.3}$$

(ii) Set

$$q := q(p_1) := \frac{\gamma p_1}{\gamma - 2 + 2p_1} \vee 1, \tag{1.8.4}$$

and note that $q > 1 \Leftrightarrow \gamma > 2 \Leftrightarrow q < \gamma/2$, in particular, $q < \gamma$ in any case. Hence repeating the above argument with the triple c, $\gamma/2$, q in place of $|b|$, γ, r, we obtain that

$$cF \in L^q_{\rm loc}(\Omega). \tag{1.8.5}$$

Fix $p_1 > 1$ such that $F := \frac{d\mu}{dx} \in L^{p_1}_{\rm loc}(\Omega)$ and let r, q be as in (1.8.2), (1.8.4), respectively. Since $\gamma \leqslant d$, we have $q < d$, which by (1.8.5) and assertions (ii) and (iii) of Lemma 1.8.1 implies that $cF \in H^{dq/(d-q),-1}_{\rm loc}(\Omega)$ if $q > 1$ and that $cF \in H^{s,-1}_{\rm loc}(\Omega)$ for any $s \in (1, d/(d-1))$ if $q = 1$.

It turns out that if $p_1 < d/(d-\gamma)$, then

$$cF \in H^{r,-1}_{\rm loc}(\Omega). \tag{1.8.6}$$

Indeed, if $q > 1$, then (1.8.6) follows from the fact that if $p_1 \in \big(1, d/(d-\gamma)\big)$, then the inequality $r \leqslant dq/(d-q)$ holds. If $q = 1$, then $\gamma \leqslant 2$ and (1.8.6) follows from the fact that $r < d/(d-\gamma+1) \leqslant d/(d-1)$ for $p_1 < d/(d-\gamma)$.

Finally, by Lemma 1.8.1 (ii) we obtain that $\nu \in H^{d/(d-\gamma+1),-1}_{\rm loc}(\Omega)$ if $\gamma > 2$ and $\nu \in H^{s,-1}_{\rm loc}(\Omega)$ for any $s \in \big(1, d/(d-1)\big)$ if $\gamma \leqslant 2$. In the same way as above, $\nu \in H^{r,-1}_{\rm loc}(\Omega)$ whenever $1 < p_1 < d/(d-\gamma)$. This along with (1.8.3) and (1.8.6) shows that the right-hand side of (1.8.1) is now in $H^{r,-1}_{\rm loc}(\Omega)$. By Lemma 1.8.1(i) we have

$$\mu \in H^{r,1}_{\rm loc}(\Omega) \tag{1.8.7}$$

and by Lemma 1.8.1(ii) we have $F \in L^{p_2}_{\rm loc}(\Omega)$, where

$$p_2 := \frac{dr}{d-r} = \frac{d\gamma p_1}{d\gamma - d + (d-\gamma)p_1} =: f(p_1).$$

Thus, we obtain

$$p_1 \in \left(1, \frac{d}{d-\gamma}\right) \text{ and } F \in L^{p_1}_{\rm loc}(\Omega) \Longrightarrow F \in L^{f(p_1)}_{\rm loc}(\Omega).$$

One can easily check that $p_2 = f(p_1) > p_1$ if $p_1 < d/(d-\gamma)$, and that the only positive solution of the equation $q = f(q)$ is $q = d/(d-\gamma)$. Therefore, by taking p_1 in $\big(1, d/(d-1)\big)$, which is possible by (i), and by defining $p_{k+1} = f(p_k)$ we obtain an increasing sequence of numbers $p_k \uparrow d/(d-\gamma)$, which implies that $F \in L^p_{\rm loc}(\Omega)$ for any $p < d/(d-\gamma)$.

But as $p_k \nearrow d/(d-\gamma)$, the sequence of numbers $r(p_k)$ defined according to equality (1.8.2) increases to the limit

$$\frac{\gamma d/(d-\gamma)}{\gamma - 1 + d/(d-\gamma)} = \frac{d}{d-\gamma+1}\ .$$

By (1.8.7) this proves (ii).

(iii) First we consider case (b). By the last assertion in (ii) we have $F \in L^{p_1}_{\rm loc}(\Omega)$ for any finite $p_1 > 1$. Let $r := r(p_1)$ be defined as in (1.8.2). Then $1 < r < \gamma$ and inclusions (1.8.3) hold. Set

$$q := q(p_1) := \frac{\frac{d\gamma}{d+\gamma}\, p_1}{\frac{d\gamma}{d+\gamma} - 1 + p_1}\ . \tag{1.8.8}$$

If $2 \leqslant d < \gamma$, then $d\gamma/(d+\gamma) > 1$. Therefore, since $p_1 > 1$, it follows that $1 < q < d\gamma/(d+\gamma)$. Hence repeating the arguments that led to (1.8.3) with the triple

c, $\frac{d\gamma}{d+\gamma}$, q in place of $|b|$, γ, r, we obtain $cF \in L^q_{\mathrm{loc}}(\Omega)$, thus, $cF \in H^{dq/(d-q),-1}_{\mathrm{loc}}(\Omega)$ by assertion (ii) of the lemma. Observe that, as $p_1 \to \infty$, we have

$$r \uparrow \gamma, \quad q \uparrow \frac{d\gamma}{d+\gamma}, \quad \frac{dq}{d-q} \uparrow \gamma.$$

Therefore, combining this with our assumption that ν is contained in the class $L^{d\gamma/(d+\gamma)}_{\mathrm{loc}}(\Omega)$, which by assertion (ii) of the lemma is contained in $H^{\gamma,-1}_{\mathrm{loc}}(\Omega)$, and by taking p_1 large enough, we see that the right-hand side of (1.8.1) is in $H^{\gamma-\varepsilon,-1}_{\mathrm{loc}}(\Omega)$ for any $\varepsilon \in (0, \gamma-1)$. By Lemma 1.8.1(ii) we conclude that $F \in H^{\gamma-\varepsilon,1}_{\mathrm{loc}}(\Omega)$. Since $\gamma > d$, the function F is locally bounded. Now we see that above we can take $p_1 = \infty$ and therefore the right-hand side of (1.8.1) is in $H^{\gamma,-1}_{\mathrm{loc}}(\Omega)$, which by assertion (i) of the lemma gives us the desired result.

In the remaining case (a) we take $p_1 > \gamma/(\gamma-1)$ and assume that $F \in L^{p_1}_{\mathrm{loc}}(\Omega)$. Then instead of (1.8.2) and (1.8.8) we define

$$r := r(p_1) := \frac{\gamma p_1}{\gamma + p_1}, \quad q := q(p_1) := \frac{\frac{d\gamma}{d+\gamma} p_1}{\frac{d\gamma}{d+\gamma} + p_1} \vee 1 \tag{1.8.9}$$

and observe that, since $p_1 > \gamma/(\gamma-1)$, we have $r > 1$, which (because of the relation $p_1^{-1} + \gamma^{-1} = r^{-1}$) allows us to apply Hölder's inequality starting with $|bF|^r = |b|^r|F|^r$ to conclude that (1.8.3) holds. Since $c \in L^1_{\mathrm{loc}}(\Omega, \mu)$,

$$\frac{d\gamma}{d+\gamma} > 1 \quad \text{and} \quad \left(\frac{d\gamma}{d+\gamma}\right)^{-1} + p_1^{-1} = q^{-1},$$

we also have that $cF \in L^q_{\mathrm{loc}}(\Omega)$. Obviously, $q < d$. As in part (ii) this yields that $cF \in H^{dq/(d-q),-1}_{\mathrm{loc}}(\Omega)$ if $q > 1$ and $cF \in H^{s,-1}_{\mathrm{loc}}(\Omega)$ for any $s \in (1, d/(d-1))$ if $q = 1$. We assert that (1.8.6) holds (with $r = r(p_1)$ as in (1.8.9)) for all $p_1 > \gamma/(\gamma-1)$, $p_1 \neq d\gamma/(d\gamma - d - \gamma)$.

Indeed, if $q > 1$, then $dq/(d-q) = r$. If $q = 1$, then $p_1 \leqslant d\gamma/(d\gamma - d - \gamma)$. But since $p_1 \neq d\gamma/(d\gamma - d - \gamma)$, we have $p_1 < d\gamma/(d\gamma - d - \gamma)$, which is equivalent to the inequality $r < d/(d-1)$.

Thus, since $\nu \in L^{d\gamma/(d+\gamma)}_{\mathrm{loc}}(\Omega) \subset H^{\gamma,-1}_{\mathrm{loc}}(\Omega) \subset H^{r,-1}_{\mathrm{loc}}(\Omega)$, because $r < \gamma$, assertion (i) in the lemma yields the following:

$$\Bigl(p_1 > \frac{\gamma}{\gamma-1},\ p_1 \neq \frac{d\gamma}{d\gamma - d - \gamma},\ F \in L^{p_1}_{\mathrm{loc}}(\Omega)\Bigr) \Longrightarrow F \in H^{r,1}_{\mathrm{loc}}(\Omega). \tag{1.8.10}$$

If $r < d$, then the latter in turn implies by assertion (ii) in Lemma 1.8.1 that $F \in L^{p_2}_{\mathrm{loc}}(\Omega)$. Let us summarize what has been shown:

$$\begin{gathered}\Bigl(p_1 > \tfrac{\gamma}{\gamma-1},\ p_1 \neq \tfrac{d\gamma}{d\gamma-d-\gamma},\ r := \tfrac{\gamma p_1}{\gamma+p_1} < d,\ F \in L^{p_1}_{\mathrm{loc}}(\Omega)\Bigr) \\ \Longrightarrow F \in L^{p_2}_{\mathrm{loc}}(\Omega),\end{gathered} \tag{1.8.11}$$

where

$$p_2 := \frac{dr}{d-r} = \frac{d\gamma p_1}{d\gamma - (\gamma-d)p_1} > \frac{d\gamma}{d\gamma - (\gamma - d)} p_1.$$

Also, note that $\gamma/(\gamma-1) < d/(d-1) < \frac{d\gamma}{\gamma d - d - \gamma}$, so that by (i) we can find a number p_1 to start with. Then starting with p_1 close enough to $d/(d-1)$, by iterating (1.8.11) we always increase p by some factor greater than $d\gamma/(d\gamma - (\gamma - d)) > 1$. While doing this, we can obviously choose the first p so that the iterated numbers

p will never be equal to $d\gamma/(d\gamma - d - \gamma)$ and the corresponding numbers r will not coincide with d. After several steps we shall come to the situation where $r > d$, and then we conclude from (1.8.10) that F is locally bounded (one cannot keep iterating (1.8.11) infinitely because of the restriction $r < d$). As in case (b), we can now easily complete the proof. □

Example 1.6.10 shows that assertion (iii) of this theorem may fail if $\gamma > d$ is replaced by $\gamma = d - \varepsilon$. Then F does not even need to be in $H^{2,1}_{\rm loc}(\Omega)$.

1.8(ii). Increasing the Sobolev regularity of solutions

The following result of Morrey is known (see Morrey [**723**, Theorem 5.5.3, p. 154], where A is continuous) about raising the integrability of a solution of the full equation with the operator $\mathcal{L}_{A,b,\beta,c}$ on a domain $\Omega \subset \mathbb{R}^d$, $d > 1$, where $a^{ij} \in VMO$ and $A(x) \geqslant \delta \cdot \mathrm{I}$ with some $\delta > 0$.

1.8.3. Theorem. *Let $d/(d-1) \leqslant q < r$ and let $u \in W^{q,1}_{\rm loc}(\Omega)$ satisfy the equation*

$$\mathcal{L}_{A,b,\beta,c} u = f + \operatorname{div} g, \quad g = (g^1, \ldots, g^d),$$

where $f \in L^{dr/(d+r)}_{\rm loc}(\Omega)$, $g^i \in L^r_{\rm loc}(\Omega)$, $b^i, \beta^i \in L^s_{\rm loc}(\Omega)$, $c \in L^t_{\rm loc}(\Omega)$, $s = d$ and $t = d/2$ if $r < d$, $s > d$ and $t > d/2$ if $r = d$, $s = r$ and $t = dr/(d+r)$ if $r > d$. Then $u \in W^{r,1}_{\rm loc}(\Omega)$.

PROOF. 1. Under the assumptions of the theorem for every ball $U(x_0,\varepsilon)$ in Ω of a sufficiently small radius ε there is a function $v \in W^{r,1}_0\big(U(x_0,\varepsilon)\big)$ satisfying the equation

$$\partial_{x_i}(a^{ij}\partial_{x_j} v - b^i v) + \beta^i \partial_{x_i} v = f + \partial_{x_i} g^i. \tag{1.8.12}$$

It suffices to establish the existence in the case of smooth coefficients, since approximating the coefficients by smooth functions and using the estimate

$$\|u\|_{W^{r,1}_0(U(x_0,\varepsilon))} \leqslant N(\|g\|_{L^r(U(x_0,\varepsilon))} + \|f\|_{L^{dr/(d+r)}(U(x_0,\varepsilon))}),$$

one can construct a sequence of smooth functions converging to the solution. However, it is important that the radius ε of the ball $U(x_0,\varepsilon)$ could depend only on ω_A and the quantities δ, $\|\beta\|_{L^s}$, and $\|b\|_{L^s}$ and be independent of the smoothness of the coefficients.

For constructing a solution in the case of smooth coefficients it suffices to show that for a sufficiently small ball $U(x_0,\varepsilon)$ the solution of the homogeneous equation ($f = g^i = 0$) in the class $W^{2,1}_0\big(U(x_0,\varepsilon)\big)$ must be zero. Then, by the Fredholm alternative (see Proposition 2.1.4), there exists a solution of the nonhomogeneous equation on this ball. Thus, let $v \in W^{2,1}_0\big(U(x_0,\varepsilon)\big)$ and

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

Multiplying by v and integrating by parts we obtain

$$\varepsilon \int_{U(x_0,\varepsilon)} |\nabla v|^2\,dx \leqslant \int_{U(x_0,\varepsilon)} |b+\beta|\,|v|\,|\nabla v|\,dx.$$

We recall that by the Sobolev inequality $\|v\|_{L^{2d/(d-2)}(U(x_0,\varepsilon))} \leqslant c(d)\|\nabla v\|_{L^2(U(x_0,\varepsilon))}$, where $c(d)$ does not depend on ε. Applying Hölder's inequality, we obtain

$$\int_{U(x_0,\varepsilon)} |b+\beta|\,|v|\,|\nabla v|\,dx \leqslant c(d)\|b+\beta\|_{L^d(U(x_0,\varepsilon))}\|\nabla v\|^2_{L^2(U(x_0,\varepsilon))}.$$

Choosing ε so small that $c(d)\|b+\beta\|_{L^d(U(x_0,\varepsilon))} \leqslant \delta/2$, we obtain the inequality

$$\int_{U(x_0,\varepsilon)} |\nabla v|^2\, dx \leqslant 0.$$

Therefore, $v = 0$ on $U(x_0, \varepsilon)$.

Thus, we have shown uniqueness of a solution in $W_0^{2,1}\bigl(U(x_0,\varepsilon)\bigr)$, hence also in $W_0^{r,1}\bigl(U(x_0,\varepsilon)\bigr)$ with $r \geqslant 2$. We observe once again that for constructing a solution in the smooth case this uniqueness is enough. If $r \leqslant 2$, then some additional reasoning is required.

We show that the solution v is unique. It suffices to show that the solution v of the homogeneous equation must be zero. To this end we solve the adjoint equation with the right-hand side $f = \operatorname{sign} v$ and $g = 0$, i.e.,

$$\partial_{x_i}(a^{ij}\partial_{x_i} w - \beta^i w) + b^i w = \operatorname{sign} v.$$

Since the right-hand side is bounded, there is a solution $w \in W_0^{m,1}\bigl(U(x_0,\varepsilon)\bigr)$ with m as close to d as we wish (so $w \in L^p\bigl(U(x_0,\varepsilon)\bigr)$ for any p). Multiplying this equation by v and integrating by parts we conclude that $v = 0$. Thus, on a sufficiently small ball $U(x_0,\varepsilon)$ there is a unique solution $v \in W_0^{r,1}\bigl(U(x_0,\varepsilon)\bigr)$ of equation (1.8.12).

2. Let $r < d$ and $\zeta \in C_0^\infty(\Omega)$, where the support of ζ belongs to a ball U of a sufficiently small radius indicated above. Then

(1.8.13)
$$\partial_{x_i}\bigl(a^{ij}\partial_{x_j}(u\zeta) - b^i(u\zeta)\bigr) + \beta^i\partial_{x_i}(u\zeta) = \partial_{x_i}(a^{ij}u\partial_{x_j}\zeta - g^i\zeta) + a^{ij}\partial_{x_i}\zeta\partial_{x_j}u + b^i u\partial_{x_i}\zeta - \beta^i u\partial_{x_i}\zeta + cu\zeta + g^i\partial_{x_i}\zeta + f\zeta.$$

Set

$$\widetilde{f} = a^{ij}\partial_{x_i}\zeta\partial_{x_j}u + b^i u\partial_{x_i}\zeta - \beta^i u\partial_{x_i}\zeta + cu\zeta + g^i\partial_{x_i}\zeta + f\zeta,\ \widetilde{g}^i = a^{ij}u\partial_{x_j}\zeta - g^i\zeta.$$

Then

$$\partial_{x_i}\bigl(a^{ij}\partial_{x_j}(u\zeta) - b^i(u\zeta)\bigr) + \beta^i\partial_{x_i}(u\zeta) = \widetilde{f} + \partial_{x_i}\widetilde{g}^i.$$

We observe that $\widetilde{f} \in L^q(U)$ and $\widetilde{g}^i \in L^{dq/(d-q)}(U)$. Then by the previous step we have $u\zeta \in W_0^{dq/(d-q),1}(U)$. Note that $dq/(d-q) \geqslant q(d-1)/(d-2)$, since $q \geqslant d/(d-1)$. Repeating this reasoning on a smaller ball we again increase the smoothness of our solution until we obtain the inclusion to the class $W^{r,1}$. The cases $r = d$ and $r > d$ are similar. □

1.8(iii). Renormalized solutions

Let $\Omega \subset \mathbb{R}^d$. Suppose we are given a nonnegative nonzero function $W \in L^1_{\mathrm{loc}}(\Omega)$ satisfying the equation

$$\partial_{x_i}\partial_{x_j}(a^{ij}W) = 0, \tag{1.8.14}$$

where $A = A^*$ is infinitely differentiable, $\lambda^{-1}\cdot \mathrm{I} \leqslant A(x) \leqslant \lambda\cdot\mathrm{I}$, $x\in\Omega$. A renormalized with respect to W solution of equation (1.8.14) is a function w such that we have $wW \in L^1_{\mathrm{loc}}(\Omega)$ and wW in place of W satisfies (1.8.14). It turns out that the renormalized solutions possess many nice properties. For example, the maximum principle holds for them. Indeed, the function w satisfies the equation

$$a^{ij}\partial_{x_i}\partial_{x_j}w + 2W^{-1}\partial_{x_j}(a^{ij}W)\partial_{x_i}w = 0,$$

for the solutions of which the classical strong maximum principle holds. In the next theorem we have collected some typical results from the papers Bauman [**94**], Escauriaza [**341**], [**342**].

1.8.4. Theorem. *Let w be a renormalized with respect to W solution of equation* (1.8.14). *Then the following assertions are true.*

(i) *Harnack's inequality. Let $w \geqslant 0$. For every ball $U(x_0, r)$ with $U(x_0, 5r) \subset \Omega$ one has the inequality*

$$\sup_{U(x_0,r)} w(x) \leqslant C(\lambda, d) \inf_{U(x_0,r)} w(x).$$

(ii) *Hölder's continuity. There is a number $\alpha = \alpha(\lambda, d) \in (0, 1]$ such that for every ball $U(x_0, r) \subset \Omega$ one has the estimate*

$$|w(x) - w(y)| \leqslant C(\lambda, d)\Bigl(\frac{|x - y|}{r}\Bigr)^{\alpha} \sup_{U(x_0,r)} |w(z)|.$$

(iii) *Boundedness of solutions. For every ball $U(x_0, r) \subset \Omega$ one has the estimate*

$$\sup_{U(x_0,r/2)} |w(x)| \leqslant C(\lambda, d)\biggl(\int_{U(x_0,r)} |w|W\, dx\biggr)\biggl(\int_{U(x_0,r)} W\, dx\biggr)^{-1}.$$

The constants $C(\lambda, d)$ and $\alpha(\lambda, d)$ depend only on λ and d and are independent of the smoothness of A.

Assertion (i) is obtained in [**94**, Theorem 4.4]. Assertion (ii) is a corollary of (i) and the fact that any constant is a renormalized solution. Finally, (iii) is proved in [**341**, Theorem 2.3]. In [**342**], these results are used for obtaining some estimates on the Green's function.

1.8(iv). Generalizations of the maximum principle of A.D. Aleksandrov and k-Hessians

For the proof of the existence of densities in § 1.5 we have used the maximum principle of A.D. Aleksandrov (see Theorem 1.5.1). Here we consider some of its generalizations. Let Ω be a convex bounded domain in $\mathbb{R}^d$ with boundary of class C^1, e.g., a ball. It is known (see Gilbarg, Trudinger [**409**, Theorem 8.15]) that if $u \in C^2(\Omega) \bigcap C_0(\overline{\Omega})$ satisfies the equation $\Delta u = f$, then

$$\sup_{\Omega} |u| \leqslant C(d, \Omega, p)\|f\|_{L^p(\Omega)}$$

for any $p > d/2$. According to Aleksandrov's maximum principle, if a convex function $u \in C^2(\Omega) \bigcap C_0(\overline{\Omega})$ satisfies the equation $\det D^2 u = f$, then

$$\sup_{\Omega} |u| \leqslant C(d, \Omega)\|f\|^{1/d}_{L^1(\Omega)}.$$

It turns out that some intermediate estimates hold.

Let $S_k(u)$ be the sum of the principal k-minors in the matrix D^2u. In this case $S_1(u) = \Delta u = \mathrm{tr} D^2 u$ and $S_d(u) = \det D^2 u$. The expression $S_k(u)$ is called the k-Hessian. A function $u \in C^2(\Omega) \bigcap C_0(\overline{\Omega})$ is called k-admissible if $S_j(u) \geqslant 0$ for all indices $j \leqslant k$. The set of all k-admissible functions will be denoted by $\Phi^k_0(\Omega)$. Let

$$\|u\|_{\Phi^k_0(\Omega)} = \biggl(\int_{\Omega} -uS_k(u)\, dx\biggr)^{1/(k+1)}.$$

The next result was proved in Wang [**934**, Theorem 5.1].

1.8.5. Theorem. *Let* $u \in \Phi_0^k(\Omega)$.

(i) *If* $1 \leqslant k < d/2$, *then* $\|u\|_{L^{p+1}(\Omega)} \leqslant C\|u\|_{\Phi_0^k(\Omega)}$ *for each* $p+1 \in [1, k^*]$, *where we set* $k^* = d(k+1)/(d-2k)$.

(ii) *If* $k = d/2$, *then* $\|u\|_{L^p(\Omega)} \leqslant C\|u\|_{\Phi_0^k(\Omega)}$ *for all* $p \geqslant 1$.

(iii) *If* $d/2 < k \leqslant d$, *then* $\|u\|_{L^\infty(\Omega)} \leqslant C\|u\|_{\Phi_0^k(\Omega)}$.

The number C *depends only on* d, k, *and* $\operatorname{diam}\Omega$.

Note that in the case $k = d/2$ in Tian, Wang [**895**] the following was proved: there are numbers $\alpha(d) > 0$ and $C(d,\Omega) > 0$ such that

$$\int_\Omega \exp\biggl(\alpha\Bigl(\frac{|u|}{\|u\|_{\Phi_0^k(\Omega)}}\Bigr)^\beta\biggr)\,dx \leqslant C(d,\Omega), \quad \text{where } 1 \leqslant \beta \leqslant (d+2)/d.$$

By using Moser's iterations, one can derive from this result analogs of Aleksandrov's maximum principle for k-Hessians. The next corollary is proved in [**934**, Theorem 5.5].

1.8.6. Corollary. *Let* $u \in \Phi_0^k(\Omega)$ *and* $S_k(u) = f$. *Then*

$$\sup_\Omega |u| \leqslant C\|f\|_{L^p(\Omega)}^{1/k},$$

where $p > d/(2k)$ *if* $k \leqslant d/2$ *and* $p = 1$ *if* $k > d/2$.

Let $\sigma_k(\lambda) = \sum_{1\leqslant i_1<i_2<\ldots<i_k\leqslant d} \lambda_{i_1}\cdots\lambda_{i_k}$ be the kth elementary symmetric polynomial. Set

$$\Gamma_k = \{\lambda \in \mathbb{R}^d\colon \sigma_j(\lambda) \geqslant 0, j = 1, 2, \ldots, k\}, \quad \Gamma_k^* = \{\lambda \in \mathbb{R}^d\colon \langle\lambda, \mu\rangle \geqslant 0\,\forall \mu \in \Gamma_k\}.$$

Let A be a symmetric matrix and let $\lambda(A)$ be the vector of eigenvalues of A. Suppose that $\lambda(A) \in \Gamma_k^*$. Set

$$\varrho_k^*(A) = \inf\bigl\{\langle\lambda(A), \mu\rangle\colon \mu \in \Gamma_k, \sigma_k(\mu) \geqslant 1\bigr\}$$

For example, $\lambda(A) \in \Gamma_1^*$ precisely when $A = g\cdot \mathrm{I}$, where g is a nonnegative number. In this case $\varrho_k^*(A) = g$.

The next related result is proved in Kuo, Trudinger [**565**].

1.8.7. Corollary. *Let* $u \in C^2(\Omega)\bigcap C_0(\overline{\Omega})$ *and* $\operatorname{tr}(AD^2u) \geqslant f$. *Suppose that* $\lambda(A) \in \Gamma_k^*$ *and* $\varrho_k^*(A) > 0$. *Then*

$$\sup_\Omega u \leqslant C\|f/\varrho_k^*(A)\|_{L^q(\Omega)},$$

where $q = k$ *if* $k > d/2$ *and* $q > d/2$ *if* $k \leqslant d/2$. *The number* C *depends only on* d, k, q, *and* Ω.

Comments

Bibliographic materials related to the outstanding Russian mathematician Andrey Nikolaevich Kolmogorov and the great German physicist, a Nobel prize winner Max Planck, whose names are in the title of this book, can be found in Kolmogorov [**530**] and Klyauc, Frankfurt [**517**], where some additional references are given. Adriaan Daniël Fokker (17.VIII.1887 – 24.IX.1972) is a Dutch physicist, a member of the Royal Dutch Academy. He was born in the island of Java, a Dutch colony at the time, in 1904 – 1905 was a student at the Polytechnic school in Delft, in 1906 – 1913 studied physics at the University of Leiden with H. Lorentz, on

October 24, 1913, defended his thesis “Over Brown’sche bewegingen in het stralingsveld, en waarschijnlijkheids-beschouwingen in de stralingstheorie”, the main results of which were published in his paper [**377**] (where the stationary equation was considered). Later he continued his studies with Albert Einstein. After his military service in World War I he worked in 1917 – 1918 as an assistant of H. Lorentz and P. Ehrenfest, in 1923 – 1927 he was Professor at the High Technical School in Delft, in 1928 – 1955 he was curator of the physical cabinet of the Teylers Museum in Haarlem and was Professor at the University of Leiden. Fokker’s main scientific works are devoted to radiation, X-rays, physics of electron, general relativity, gravitation, the theory of fluctuations, and the theory of gyroscope. The equation bearing his name was needed for establishing the distribution law of the average energy of a rotating electric dipole, for this purpose Fokker developed a method first used by Einstein [**329**] for describing the Brownian motion. In addition, Fokker was an expert in music theory, he was enthusiast of the 31 equal temperament, his 31-tone equal-tempered organ, which was installed in Teyler’s Museum in Haarlem in 1951, is called the Fokker organ. It is worth mentioning that Kolmogorov and Planck also were amateurs and connoisseurs of music, moreover, Planck was playing at a professional level.

There is an extensive literature on the theory of Sobolev spaces, see Adams [**2**], Adams, Fournier [**3**], Besov, Il’in, Nikolskiĭ [**112**], Bogachev [**126**], Brezis [**208**], Burenkov [**212**], Evans, Gariepy [**344**], Gol’dshteĭn, Reshetnyak [**412**], Haroske, Triebel [**440**], Krylov [**556**], Kufner, Sändig [**558**], Leoni [**602**], Maz’ja [**682**], [**683**], Runst, Sickel [**828**], Stein [**877**], Triebel [**899**], [**900**], and Ziemer [**965**], where numerous additional references can be found. In [**126**], [**558**] and Zhikov [**961**] also weighted Sobolev classes are considered.

Many books are devoted to the general theory of elliptic second order linear partial differential equations, see Agmon [**5**], Agmon, Douglis, Nirenberg [**6**], Bers, John, Schechter [**109**], Borsuk [**198**], Borsuk, Kondratiev [**199**], Chen, Wu [**237**], Demengel, Demengel [**293**], Egorov, Kondratiev [**325**], Gilbarg, Trudinger [**409**], Han, Lin [**438**], Hörmander [**461**], Helffer [**447**], Kenig [**501**], Kondrat’ev, Landis [**533**], Koshelev [**537**], Kozlov, Maz’ya, Rossmann [**541**], Kresin, Maz’ya [**542**], Krylov [**552**], [**556**], Ladyzhenskaya, Ural’tseva [**577**], Landis [**581**], Lions, Magenes [**618**], Maugeri, Palagachev, Softova [**681**], Maz’ya, Rossmann [**685**], Miranda [**713**], Nazarov, Plamenevsky [**737**], Oleĭnik, Radkevič [**757**], Radkevich [**798**], Rempel, Schulze [**803**], Sauvigny [**834**], Shimakura [**857**], Shishmarev [**859**], Stampacchia [**870**], Stroock [**882**], Troianiello [**901**], Volpert [**924**], Wu, Yin, Wang [**944**], and also Garroni, Menaldi [**400**].

However, Fokker–Planck–Kolmogorov equations have significant specific features and so far have not become the subject of a separate exposition, although some of their aspects are discussed in depth in books with probabilistic motives, see Gihman, Skorokhod [**407**], Krylov [**549**], Kushner [**570**], Soize [**865**] (where also explicit solutions are considered), Stroock [**882**], Stroock, Varadhan [**884**]. These specific features are connected, on the one hand, with the fact that such equations by their nature are equations with respect to measures (sometimes they are called “double divergence form” equations and their solutions are called “adjoint solutions”, in the case of irregular coefficients they cannot be written as divergence form equations or as direct equations), and, on the other hand, with unusual for the classical theory classes of solutions (say, integrable on the whole space, but without any restrictions on growth or smoothness class). A study of such equations goes back to Kolmorogov's works [**527**], [**528**], [**529**] and a series of earlier works in physics by Fokker [**377**], Smoluchowski [**863**], Planck [**781**], and Chapman [**235**], where equations for probability densities were considered (see also Hostinský [**466**]). The informative survey Fuller [**394**] lists also Lord Rayleigh and L. Bachelier among predecessors.

Connections between elliptic operators and diffusions are discussed in the books Bass [**91**] and Pinsky [**780**].

Traditionally, second order elliptic equations are solved in Hölder classes or in Sobolev classes depending on the properties of the coefficients. It is more convenient to consider the major problems of this book in Sobolev classes even in the case of smooth coefficients, which is due to the significant role of various a priori estimates in terms of quantities like Sobolev norms.

The existence of densities of solutions under minimal assumptions is based on the A.D. Aleksandrov estimates (see [**34**]–[**37**]). Unlike direct elliptic equations and divergence form equations, Fokker–Planck–Kolmogorov equations can have solutions whose regularity is not higher than the regularity of the diffusion coefficients. There are many works devoted to generalizations of Aleksandrov's estimates, see Cabré [**215**], Kuo, Trudinger [**565**]. Elliptic inequalities of the type $L^*u \geqslant 0$ were considered long ago, see, for example, Littman [**628**], [**629**].

Theorem 1.7.4 makes precise the statement of a result formulated by Ch. Morrey in his book [**723**, p. 156] not quite correctly (with $\Omega' = \Omega$). The assertion given there with $\Omega' = \Omega$ would be false, for example, for the Laplace equation on a ball. A proof of Morrey's estimate with an investigation of the dependence of the constant on the coefficients was given in Shaposhnikov [**843**] with the same inaccuracy as in [**723**]. Actually, the reasoning in Shaposhnikov [**843**] yields exactly the estimate we give, as explained in Shaposhnikov [**846**], and an estimate with $\Omega = \Omega'$ is possible only for solutions with zero boundary values on a domain with a sufficiently regular boundary (see Corollary 1.7.6). We observe that in the existing applications of Morrey's theorem actually only the presented correct statement was used, although in some papers it was formulated with the indicated inaccuracy (see, for example, Bogachev, Röckner [**160**] and Bogachev, Krylov, Röckner [**152**]). The proof of the corrected statement was given in Bogachev, Röckner, Shaposhnikov [**165**] and Shaposhnikov [**846**], where even a more general fact is proved.

Various results which can be regarded as results about properties of densities of solutions to elliptic Fokker–Planck–Kolmogorov equations with coefficients of a rather general form were obtained in the books cited above and also in the papers Krylov [**547**], Sjögren [**861**], Bauman [**94**], [**95**], Escauriaza [**341**], [**342**], Escauriaza, Kenig [**343**], Fabes, Stroock [**348**], Gushchin [**430**], Maz'ya, McOwen [**684**]; note also a more abstract approach of Herve [**452**]. A systematic study of the whole complex of these problems was initiated in Bogachev, Röckner [**157**], [**158**], [**160**], Bogachev, Krylov, Röckner [**149**], [**152**] and continued by many authors. Stationary Fokker–Planck–Kolmogorov equations with various special restrictions on the coefficients are considered in Arapostathis, Borkar [**55**], Bensoussan [**105**], Noarov [**745**]–[**749**] and in the works cited on concrete occasions in the subsequent chapters. Fokker–Planck–Kolmogorov equations can be also considered on more singular manifolds such as fractals and metric measure spaces.

Exercises

1.8.8. Suppose that probability measures μ and ν on the real line satisfy the equations $L^*_{1,b_1,c_1}\mu = 0$ and $L^*_{1,b_2,c_2}\nu = 0$. Show that the measure $\sigma = \mu\otimes\nu$ satisfies the equation $L^*_{\mathrm{I},b,c}\sigma = 0$ with $b(x,y) = \bigl(b_1(x), b_2(y)\bigr)$, $c(x,y) = c_1(x) + c_2(y)$.

1.8.9. Let a measure μ on $\mathbb{R}^d$ satisfy the equation $L^*_{A,b}\mu = 0$. Write down the equation to which the measure $\nu = \mu\circ F^{-1}$ satisfies, where $F = (F^i)$ is a diffeomorphism of $\mathbb{R}^d$ with inverse $G = (G^i)$.

HINT: Make the change of variable $x = G(y)$ in the integral identity, substituting $\varphi = \psi\circ F$, obtain the equation $L^*_{Q,h}\nu = 0$ with the matrix $Q = (q^{mk})$, where $q^{mk}(y) = a^{ij}\bigl(G(y)\bigr)\partial_{x_i}F^k\bigl(G(y)\bigr)\partial_{x_j}F^m\bigl(G(y)\bigr)$, and the drift $h = (h^k)$, where $h^k(y) = a^{ij}\bigl(G(y)\bigr)\partial_{x_i}\partial_{x_j}F^k\bigl(G(y)\bigr) + b^i\bigl(G(y)\bigr)\partial_{x_i}F^k\bigl(G(y)\bigr)$.

1.8.10. Suppose that a probability measure μ on the real line satisfies the equation $L^*_{1,b}\mu = 0$ on $(-\infty, 0)$ and on $(0, +\infty)$ with some continuous function b on the real line. Is it true that μ satisfies this equation on the whole real line?

1.8.11. Give an example of an absolutely continuous function on $[0,1]$ that does not satisfy the Hölder condition of any order.

1.8.12. (Bogachev, Röckner [**160**]) Let $A_k = (a^{ij}_k)$ be a sequence of continuous mappings on $\mathbb{R}^d$ with values in the space of symmetric matrices and let $b_k = (b^i_k)$ be a sequence of Borel vector fields on $\mathbb{R}^d$. Suppose that for every ball $U_r \subset \mathbb{R}^d$ of radius r there exist numbers $c_r > 0$, $\alpha_r > 0$, and $p = p_r > d$ such that

$$A_k \geqslant c_r\mathrm{I}, \quad \|a^{ij}_k\|_{W^{p,1}(U_r)} + \|b^i_k\|_{L^p(U_r)} \leqslant \alpha_r \quad \text{for all } i,j,k.$$

Assume that there are probability measures μ_k on $\mathbb{R}^d$ such that $L^*_{A_k,b_k}\mu_k = 0$. Then the measures μ_k have continuous strictly positive densities that are uniformly Hölder continuous on every ball. If the sequence $\{\mu_k\}$ is uniformly tight, then it has compact closure in the variation norm, and every measure μ in its closure has a continuous strictly positive density of class $W^{p,1}(U_r)$ for every $r > 0$.

HINT: It follows from our hypotheses and Theorem 1.6.5 that the measures μ_k have continuous densities f_k. Since the functions f_k are probability densities, we obtain by (1.6.2) that, for every $r > 0$, the sequence $\{f_k\}$ is bounded in $W^{p,1}(U_r)$. By the Sobolev embedding theorem $\{f_k\}$ is uniformly Hölder continuous on U_r, in particular, has compact closure with respect to the sup-norm. If $\{\mu_k\}$ is uniformly tight, then some subsequence $\{\mu_{k_i}\}$ converges weakly to some probability measure μ. Passing to a subsequence once again we may assume that the functions f_{k_i} converge uniformly on compact sets and are uniformly bounded in $W^{p,1}(U_r)$ for each $r > 0$. Hence μ has a density $f \in W^{p,1}(U_r)$. Then we obtain a continuous and strictly positive version of f. Therefore, the probability measures μ_{k_i} converge to μ in the variation norm. This reasoning applies to any subsequence in $\{\mu_k\}$, whence we obtain the desired conclusion.

1.8.13. (Bogachev, Röckner [**160**]) The assertion of the previous exercise can be generalized as follows. Let Ω be an open set in $\mathbb{R}^d$ that is the union of increasing open sets Ω_k such that the closure of Ω_k is compact and contained in Ω_{k+1}. Let μ_k be probability measures on Ω_k satisfying the equations $L^*_{A_k,b_k}\mu_k = 0$ on Ω_k, where each A_k is a continuous mapping on Ω_k with values in the set of nonnegative symmetric matrices, the mappings A_k are uniformly bounded on compact sets in the

$W^{p,1}$-norm with some $p > 1$, the mappings A_k^{-1} are uniformly bounded on compact sets, and Borel vector fields b_k on the sets Ω_k are uniformly bounded in the $L^p(\mathbb{R}^d)$-norm on compact sets. Then the analogue of the assertion of the previous exercise is true. The same is true for Riemannian manifolds of dimension d.

1.8.14. Let $(X, \mathcal{B}, \mu)$ be a probability space, where the measure μ is separable, and let $(S, \mathcal{S})$ be a measurable space. Suppose that for each $s \in S$ we are given a μ-integrable function ξ_s such that for every set $B \in \mathcal{B}$ the integral of ξ_s over B is an $\mathcal{S}$-measurable function of s. Prove that for every $s \in S$ one can choose a version of ξ_s such that the function $(x, s) \mapsto \xi_s(x)$ is $\mathcal{B}\otimes\mathcal{S}$-measurable.

Hint: If μ is Lebesgue measure on $[0,1]$ and $f_s \in L^2(\mu)$, then we can take the standard trigonometric basis $\{e_n\}$ in $L^2(\mu)$ and set $\xi_s(x) = \sum_{n=1}^{\infty}(\xi_s, e_n)e_n(x)$ at the points of convergence, by using the Carleson theorem on convergence of this series almost everywhere and making ξ_s zero outside; the set of convergence belongs to $\mathcal{B}\otimes\mathcal{S}$, which ensures the $\mathcal{B}\otimes\mathcal{S}$-measurability of the obtained version. In the general case we can use the existence of a Schauder basis $\{\varphi_n\}$ in $L^1(\mu)$ consisting of $\mathcal{B}$-measurable functions with the property that for each $f \in L^1(\mu)$ the series $f = \sum_{n=1}^{\infty} c_n(f)\varphi_n$ converges μ-a.e.; the coefficients $c_n(f)$ are continuous linear functionals on $L^1(\mu)$, hence they are represented as the integrals of $f\psi_n$ with certain $\psi_n \in L^\infty(\mu)$, which ensures the measurability of $s \mapsto c_n(f_s)$. Finally, a Schauder basis with the indicated property can be constructed as follows: by the separability of the measure there is a countable collection of sets $B_n \in \mathcal{B}$ the linear span of the indicators of which is dense in $L^1(\mu)$; the σ-algebra $\mathcal{B}_n$ generated by $B_1, \ldots, B_n$ can be also generated by a partition of X into finitely many disjoint parts $B_{n,1}, \ldots, B_{n,k_n}$, the conditional expectations of f with respect to $\mathcal{B}_n$ converge to f in norm and almost everywhere by the martingale convergence theorem, and these conditional expectations can be represented as partial sums of a series in Haar-type functions constructed by means of the indicated decreasing partitions. Another construction can be found in the hint to Exercise 6.10.71 in Bogachev [**125**].

1.8.15. Let $\xi\colon\ [0,1] \to L^1(\mu)$ be a continuous mapping, where μ is a probability measure on a measurable space $(\Omega, \mathcal{B})$. Prove that there is a function $\eta\colon\ \Omega\times\mathbb{R} \to \mathbb{R}$, called a measurable modification of ξ, such that it is measurable with respect to $\mathcal{B}\otimes\mathcal{B}(\mathbb{R})$ and for each $t \in \mathbb{R}$ the equality $\eta(\omega, t) = \xi(t)(\omega)$ holds for μ-a.e. ω.

Hint: Use the previous exercise considering μ on the σ-algebra generated by $\xi(t), t \in \mathbb{Q}$; see also Neveu [**742**, § III.4].

1.8.16. Justify Remark 1.6.3.

1.8.17. Prove that if on the closed ball in $\mathbb{R}^d$ a uniformly Hölder continuous sequence of functions converges in measure, then it converges uniformly.

1.8.18. Let $u \in W^{p,2}(U) \cap W_0^{p,1}(U)$, where the set $U \subset \mathbb{R}^d$ is bounded and open and $p > d$. By using Corollary 1.1.6 show that for all $c > 0$ and $\alpha > 1$ the function $w := (u+c)^\alpha - c^\alpha$ belongs to $W_0^{p,2}(U)$.

1.8.19. Let $f \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$. Show that $\nabla f(x) = 0$ a.e. on $f^{-1}(0)$.

1.8.20. Prove that the diffusion process on the real line given by the equation $d\xi_t = dw_t + f(\xi_t)dt$, where $f'(x) + f(x)^2 = ax^2 + bx + c$, has an invariant probability measure only under the condition that $f(x) = \alpha x + \beta$, $\alpha < 0$.

Hint: See Zeitouni [**953**].