THEORY PROBAB. APPL.
Vol. 68, No. 3, pp. 342–369

Translated from Russian Journal

KOLMOGOROV PROBLEMS ON EQUATIONS FOR STATIONARY AND TRANSITION PROBABILITIES OF DIFFUSION PROCESSES*

V. I. BOGACHEV[†], M. RÖCKNER[‡], AND S. V. SHAPOSHNIKOV[†]

(*Translated by V. I. Bogachev*)

Abstract. The paper gives a survey of several directions of research connected with the works of A. N. Kolmogorov on parabolic and elliptic Fokker–Planck–Kolmogorov equations for transition and stationary probabilities of diffusion processes. We present the fundamental results on existence of solutions, their uniqueness, and the properties of solution densities. Open questions in this area are mentioned.

Key words. Fokker–Planck–Kolmogorov equation, transition probability, invariant measure, Cauchy problem

DOI. 10.1137/S0040585X97T991507

1. Introduction.

In the papers of A. N. Kolmogorov [45], [46], [47], and [48] published in the early 1930s, diffusion processes in $\mathbf{R}^d$ and finite-dimensional Riemannian manifolds were considered, and for stationary and transition probabilities of diffusions the equations were derived, which are now called Fokker–Planck–Kolmogorov equations. These equations had appeared earlier in the works of Fokker [39] and Planck [58] in physics, about which Kolmogorov learned soon after publication of his first paper [45] (in [48] the term "Fokker–Planck equation" was already in use, and the corresponding remark was made in the Russian translation of [45] published in 1938). Similar equations were also considered by Smoluchowski [63] and Chapman [32] prior to Kolmogorov. However, it is unlikely that a timely acquaintance with the works of predecessors would influence his first paper on this subject, since completely different problems were posed and solved in this paper. This is also seen from Kolmogorov's subsequent papers [46], [47], where Fokker, Planck, and Smoluchowski are cited.

The stationary or elliptic Fokker–Planck–Kolmogorov equation with respect to a measure μ on $\mathbf{R}^d$ has the form

$$\sum_{i,j}\partial_{x_i}\partial_{x_j}(a^{ij}\mu)-\sum_{i}\partial_{x_i}(b^i\mu)=0 \tag{1.1}$$

with some functions a^{ij} and b^i. Below, we give a precise definition of a solution, but, in the first papers of classical authors, the coefficients and solutions were assumed to be sufficiently smooth, so that the equation was understood in the usual sense.

The parabolic equation with the same coefficients has the form

$$\partial_t\mu_t=\sum_{i,j}\partial_{x_i}\partial_{x_j}(a^{ij}\mu_t)-\sum_{i}\partial_{x_i}(b^i\mu_t), \tag{1.2}$$

---

[*] Received by the editors January 28, 2023; published electronically November 7, 2023. This work was supported by the Russian Science Foundation grant 22-11-00015 (fulfilled at the Lomonosov Moscow State University). Originally published in the Russian journal *Teoriya Veroyatnostei i ee Primeneniya*, 68 (2023), pp. 420–455.
https://doi.org/10.1137/S0040585X97T991507

[†] Moscow Lomonosov State University, Moscow, Russia, and National Research University "Higher School of Economics," Moscow, Russia (vibogach@mail.ru, starticle@mail.ru).

[‡] Bielefeld University, 33615 Bielefeld, Germany (roeckner@math.uni-bielefeld.de).

and the Cauchy problem for this equation is complemented by initial distribution μ_0. For nice coefficients and solutions, this can be also understood in the usual sense, but the general definition is given below.

In the theory of equations with partial derivatives, elliptic equations of the indicated form are often called "double divergence form equations" in order to distinguish them from divergence form equations

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

which, in turn, differ from direct equations

$$\sum_{i,j} a^{ij}\partial_{x_i}\partial_{x_j}\mu + \sum_i b^i\partial_{x_i}\mu = 0.$$

Parabolic equations are fulfilled for transition probabilities of diffusion processes, and elliptic equations are fulfilled for stationary distributions.

In section 15 "Setting of the problem about uniqueness and existence of solutions to the second differential equation" of [45], one-dimensional equations are discussed, and the question about existence of a unique probability solution is posed, while multidimensional equations are considered in the second paper [46]. Note that the term "Fokker–Planck–Kolmogorov equation" is now used for the "second differential equation" according to Kolmogorov's terminology. Kolmogorov was mostly speaking of parabolic equations, but, in section 18, stationary equations were briefly discussed also. Note also that Kolmogorov denoted the drift coefficient by A, and the diffusion coefficient was denoted by B^2. The original problem (which was the Kolmogorov–Chapman equation rather than the parabolic equation) was posed in section 11 for densities of transition probabilities in the following way: a nonnegative function $f(t_1, x, t_2, y)$ is measurable in the sense of Borel with respect to the arguments x, y and satisfies the equations

$$\int_{-\infty}^{+\infty} f(t_1, x, t_2, y)\, dy = 1, \tag{85}$$

$$f(t_1, x, t_3, z) = \int_{-\infty}^{+\infty} f(t_1, x, t_2, y) f(t_2, y, t_3, z)\, dy \tag{86}$$

(for convenience of references, we keep the numeration of formulas from [45]). Further, in section 15, Kolmogorov writes the following:

> The main question concerning uniqueness of solutions is the following: under what conditions can one assert that, for given s and x, only one nonnegative function $f(s, x, t, y)$ of variables t, y can exist, defined for all values y and $t > s$ and satisfying equation (133) together with conditions (142), (143)? For some important special cases this question can be given a positive answer; this applies for example to all the cases considered in the next two sections.
>
> Suppose now that the functions $A(t, y)$ and $B^2(t, y)$ are given in advance; one can pose the question whether there exists a nonnegative function $f(s, x, t, y)$, which, on the one hand, satisfies equations (85) and (86) (as was shown in section 11, these requirements are necessary in order that $f(s, x, t, y)$ define a stochastic system), and, on the other hand, after passing to the limits by formulae (122) and (124) would yield these given functions $A(t, y)$ and $B^2(t, y)$.

> For solving such a problem, one can, for example, first determine some nonnegative solution of our second differential equation (133), satisfying conditions (142), (143), and next investigate whether this is indeed a solution to our problem. Consequently, the following two general questions arise:
>
> (1) Under what conditions does there exist such a solution to equation (133)?
>
> (2) Under what conditions can one assert that this solution satisfies additionally equations (85) and (86)?
>
> There are good grounds to believe that these conditions have a sufficiently general character.

Equation (133) is exactly the parabolic equation (1.2) we are interested in, which in Kolmogorov's notation had the form

$$\frac{\partial}{\partial t}f(s,x,t,y) = -\frac{\partial}{\partial y}[A(t,y,s,x)f(s,x,t,y)] + \frac{\partial^2}{\partial y^2}[B^2(t,y,s,x)f(s,x,t,y)],$$

condition (142) coincides with (85), and condition (143) is the relationship

$$\int_{-\infty}^{+\infty}(y-x)^2 f(s,x,t,y)\,dy \to 0, \qquad t \to s, \tag{143}$$

which expresses some convergence of the transition probability to the Dirac delta-measure at the point x. Finally, conditions (122) and (124) relate the coefficients A and B with the transition densities: the first one is the equality

$$B^2(s,x) = \lim_{\Delta\to 0}\frac{1}{2\Delta}\int_{-\infty}^{+\infty}(y-x)^2 f(s,x,s+\Delta,y)\,dy, \tag{1.3}$$

and the second one is the equality

$$A(s,x) = \lim_{\Delta\to 0}\frac{1}{\Delta}\int_{-\infty}^{+\infty}(y-x) f(s,x,s+\Delta,y)\,dy. \tag{1.4}$$

Actually, this is a limit relationship for the variances and means of the transition probabilities. These two conditions are meaningful only under the additional assumption of existence of the second moment of the transition densities. It is shown in section 5 that, under broad conditions, these relationships are automatically fulfilled when the main equation holds. So, below, we do not consider them (they become important in the case of the Kolmogorov–Chapman equations rather than the Fokker–Planck–Kolmogorov equations).

Theorems on existence and uniqueness of a probability solution to the Cauchy problem for the parabolic Fokker–Planck–Kolmogorov equation and also of a probability solution to the stationary equation (1.1) were proved by Kolmogorov in his paper [47] in the case of a compact Riemannian manifold (Kolmogorov called such manifolds "closed") under the assumption of the existence of continuous first and second derivatives of the coefficients. As far as the properties of solutions are concerned, Kolmogorov established positivity of their densities. These results led to questions about analogous theorems for the space $\mathbf{R}^d$ and noncompact manifolds. Naturally, questions arose about the properties of solutions in more general situations. Below, we discuss the principal achievements in these directions, complementing the content of surveys [15], [14], and [6].

In section 2, we present the fundamental results on existence of solutions to stationary equations and the properties of these solutions. The uniqueness problems in the stationary case are considered in section 3. In section 4, we begin a transition to evolution equations: we discuss the operator semigroups that are connected to solutions to stationary equations and, in some sense, are generated by elliptic operators. The existence of solutions to the Cauchy problem for Fokker–Planck–Kolmogorov equations and the properties of their densities is the subject of section 5. Uniqueness problems in the parabolic case are considered in section 6. Finally, in section 7, some remarks are made about estimates of distances between solutions to linear equations and their applications to nonlinear Fokker–Planck–Kolmogorov equations, an actively developing modern direction of study. In all sections, open problems are mentioned.

2. Stationary equations: Existence of solutions and their properties. Let us proceed with precise formulations of the Kolmogorov problems, their modern settings, and a survey of achievements in studying problems and some remaining open questions. We deal below with equations on the whole space $\mathbf{R}^d$, and just briefly comment on the case of domains or manifolds. Thus, suppose that on $\mathbf{R}^d$ we are given an operator (matrix) mapping $x \mapsto A(x) = (a^{ij}(x))_{i,j\leqslant d}$, the matrices $A(x)$ are nonnegatively definite, and their elements $a^{ij}(x)$ are Borel measurable. Then $A(x)$ is called the diffusion coefficient. In addition, suppose that we are given a Borel vector field $b = (b^i)_{i\leqslant d}$; it is called the drift coefficient. These mappings generate a second-order elliptic differential operator

$$\begin{aligned} L_{A,b}f(x) &= \sum_{i,j} a^{ij}(x)\partial_{x_i}\partial_{x_j}f(x) + \sum_i b^i(x)\partial_{x_i}f(x) \\ &= \operatorname{trace}(A(x)D^2f(x)) + \langle b(x), \nabla f(x)\rangle, \end{aligned}$$

also written in shorthand as

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

with the standard rule of summation over repeated indices. So far, this expression has a merely formal character, and for f we can take, say, functions with two (usual or Sobolev) derivatives.

For a Borel measure μ (possibly, signed) that is bounded on $\mathbf{R}^d$ or on compact sets, the stationary Fokker–Planck–Kolmogorov equation has the form

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

and is understood in the following sense: the functions a^{ij} and b^i must be integrable on compact sets with respect to the measure μ (in the case of a signed measure with respect to its total variation $|\mu|$), and the identity

$$\int L_{A,b}f(x)\,\mu(dx) = 0 \quad \forall\, f \in C_0^\infty(\mathbf{R}^d)$$

must hold, where $C_0^\infty(\mathbf{R}^d)$ denotes the class of all infinitely differentiable functions with compact support. Thus, the equation is considered only for measures with respect to which the coefficients are locally integrable. Of course, if the coefficients are locally bounded, then the equation is meaningful for all locally bounded measures. As for any differential equation, questions arise about existence of solutions, their uniqueness, and various properties. Kolmogorov discussed these questions in the class of probability solutions, but other settings are possible; in particular, we also consider solutions in the class of measures of bounded variation.

It is important to note at once that a solution to the stationary equation is not always an invariant measure of the diffusion semigroup connected with the operator $L_{A,b}$ (however, under the conditions assumed by Kolmogorov these are equivalent properties). Semigroups and their invariant measures are discussed in section 4, and for now we consider only solutions to the elliptic equation.

Not every reasonable equation has nonzero solutions. For example, for the Laplace operator $L_{A,b} = \Delta$ (i.e., $A = \mathrm{I}$, $b = 0$), a solution to our equation is a measure with a density harmonic in the sense of distributions, but then also in the usual sense. Hence in the class of measures of bounded variation, there are no nonzero solutions (the situation changes in the case of manifolds). The question about uniqueness also requires some precision, because the set of all solutions admits multiplication by constants. In this section, we state two main results about existence of solutions and their properties.

The following existence theorem (see [18] and [15]), which develops and generalizes results of Khasminskii [43], [44], uses the concept of a Lyapunov function. Recall that a real function V on a topological space is called compact if all sets $\{V \leqslant c\}$ are compact. A function V is called quasi-compact if the space can be represented as the union of increasing sets $\{V \leqslant c_k\}$ with some numbers c_k. A continuous function V on $\mathbf{R}^d$ is compact precisely when $V(x) \to +\infty$ as $|x| \to +\infty$. By a Lyapunov function for the operator $L_{A,b}$, one usually means a compact or quasi-compact function V with some estimates for $L_{A,b}V$.

Let $p \geqslant 1$, $k \in \mathbb{N}$. Let $W^{p,k}(\Omega)$ denote the Sobolev space of functions on a domain Ω belonging to $L^p(\Omega)$ along with their Sobolev derivatives up to order k. By $\|f\|_{p,k}$ we denote the Sobolev norm, which is the sum of the L^p-norms of the function and its derivatives up to order k. The class of functions on $\mathbf{R}^d$ the restrictions of which to every ball Ω belong to $W^{p,k}(\Omega)$ is denoted by $W^{p,k}_{\mathrm{loc}}(\mathbf{R}^d)$. The class $W^{d+,k}_{\mathrm{loc}}(\mathbf{R}^d)$ consists of all functions f such that, for every ball Ω, there exists a number $p = p(\Omega) > d$ for which $f \in W^{p,k}(\Omega)$. The symbols $L^p_{\mathrm{loc}}(\mathbf{R}^d)$, $L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$, $L^p_{\mathrm{loc}}(\mu)$, and $L^{d+}_{\mathrm{loc}}(\mu)$ are defined similarly.

A function $f \in L^1_{\mathrm{loc}}(\mathbf{R}^d)$ belongs to the class VMO if there is a modulus of continuity ω (i.e., ω is an increasing continuous function on $[0, +\infty)$ and $\omega(0) = 0$) such that

$$\sup_{z\in\mathbf{R}^d,\, 0<r<t} r^{-2d} \int_{|x-z|\leqslant r,\, |y-z|\leqslant r} |f(x) - f(y)|\, dx\, dy \leqslant \omega(t).$$

A function belongs to VMO on balls if from every ball it extends to a function of class VMO. A continuous function belongs to VMO on balls.

We say that a function f satisfies the Dini condition if, for every ball $B \subset \mathbf{R}^d$, there is a modulus of continuity ω_B such that

$$|f(x) - f(y)| \leqslant \omega_B(|x-y|) \quad \forall\, x, y \in B, \qquad \int_0^1 \frac{\omega_B(t)}{t}\, dt < \infty.$$

A weaker condition is the Dini mean oscillation condition employed in [34], [33]: for every ball B, there exists a modulus of continuity ω_B such that the function $\omega_B(t)/t$ is integrable on $[0, 1]$, and, for all $r \in (0, 1]$, the inequality

$$\sup_{x\in B} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f_{B(x,r)}|\, dy \leqslant \omega_B(r)$$

holds, where f_B denotes the normalized mean of the function f over the ball B, i.e., the integral of $f/|B|$ over B, and $|B|$ is the volume of B.

Theorem 2.1. *Suppose that, for the operator $L_{A,b}$, one can find a quasi-compact function $V \in W^{d,2}_{\mathrm{loc}}(\mathbf{R}^d)$ and numbers $C, R > 0$ such that*

$$L_{A,b}V(x) \leqslant -C \quad \text{if } |x| > R. \tag{2.1}$$

Either of the following conditions is sufficient for the existence of a Borel probability measure μ on $\mathbf{R}^d$ satisfying the equation $L^*_{A,b}\mu = 0$:

(i) *The coefficients A and b are continuous, and the second derivatives of V are locally bounded.*

(ii) *The coefficients A and b are locally bounded, the operators $A(x)$ are invertible, and the function $1/\det A$ is locally bounded.*

(iii) *The matrices A and A^{-1} are locally bounded, A belongs to* VMO *on balls (e.g., is continuous), and $b^i \in L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$.*

Assertion of the theorem in cases (i) and (ii) can be found in [15, Chap. 2]. In papers [24] and [25], the assertion in case (iii) is proved for a compact Lyapunov function, but it remains valid for a quasi-compact function, because the proof gives a positive locally finite measure, which, as it follows from the results in [15, section 2.3], is finite in the case of a quasi-compact Lyapunov function.

The listed conditions are not necessary, and a probability solution can exist even if they all fail. Some other sufficient conditions can be found in the book [15], but the conditions with Lyapunov functions have proved the most useful for applications. Moreover, as the next result shows (see [15, Proposition 5.3.9]), under additional constraints on the coefficients, the existence of a Lyapunov function follows from the existence of a probability solution.

Theorem 2.2. *If A^{-1} is locally bounded, $a^{ij} \in W^{d+,1}_{\mathrm{loc}}(\mathbf{R}^d)$, $b^i \in L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$, and there exists a probability solution μ to the equation $L^*_{A,b}\mu = 0$ such that*

$$\frac{\operatorname{trace} A(x)}{1+|x|^2}, \frac{|b(x)|}{1+|x|} \in L^1(\mu),$$

then there exists a function $V \in W^{d+,2}_{\mathrm{loc}}(\mathbf{R}^d)$ such that $V(x) \to +\infty$ and $L_{A,b}V(x) \to -\infty$ as $|x| \to +\infty$.

The hypotheses of the theorem are fulfilled if, for example, $|a^{ij}(x)| \leqslant C + C|x|^2$ and $|b^i(x)| \leqslant C + C|x|$.

The following question remains open.

Question 1. Suppose that there exists a nonzero bounded signed measure μ satisfying $L^*_{A,b}\mu = 0$. Does there exist a probability solution to this equation?

For $d > 1$, this question is open even in the case where $A = \mathrm{I}$ and b is infinitely differentiable.

The properties of solutions in the cases listed in Theorem 2.1 can differ substantially. In many problems, the following properties are of interest: existence of densities with respect to Lebesgue measure, their local boundedness and separateness from zero, their continuity and differentiability, and upper and lower bounds. In the case of smooth coefficients and a nondegenerate matrix A, the classical elliptic regularity results give existence of a smooth density. However, the regularity of the solution cannot be better in the general case than that of the diffusion matrix, which differs from the case of direct or divergence form equations. For example, the equation $(A\mu)'' = 0$ on the real line has a solution with density $1/A$, and hence every probability measure with positive density serves as a solution to an equation with a positive diffusion coefficient. It was shown in [61] that, for a Hölder continuous nondegenerate diffusion matrix and a locally bounded drift, the solution densities are also Hölder continuous. Let us state some principal results with the simplest formulations. Unlike the existence theorem, we are now also speaking of signed solutions.

THEOREM 2.3. *Suppose that* $\det A > 0$ *everywhere and* μ *is a bounded measure satisfying the equation* $L^*_{A,b}\mu = 0$.

(i) *If* $\mu \geqslant 0$, *then* μ *has a density.*

(ii) *If* A *and* b *are infinitely differentiable, then* μ *has an infinitely differentiable density.*

(iii) *If* A *is continuous,* $a^{ij} \in W^{d+,1}_{\mathrm{loc}}(\mathbf{R}^d)$, $b^i \in L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$, *or* $b^i \in L^{d+}_{\mathrm{loc}}(\mu)$, *then* μ *has a locally Hölder continuous density of class* $W^{d+,1}_{\mathrm{loc}}(\mathbf{R}^d)$. *If, in addition,* $\mu \geqslant 0$ *is a nonzero measure and* $b^i \in L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$, *then this density has no zeros.*

(iv) *If* A *is locally Hölder continuous of order* α *and* $b^i \in L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$, *then the density of the measure* μ *is locally Hölder continuous of the same order.*

(v) *If* A *satisfies the mean oscillation Dini condition* (*e.g., the usual Dini condition*), *and* $b^i \in L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$, *then the density of the measure* μ *is continuous. If, in addition, the measure* μ *is nonnegative and nonzero, then the continuous version of the density has no zeros.*

(vi) *If* A *and* A^{-1} *are locally bounded, the functions* a^{ij} *belong to the class* VMO *on balls, and* $b^i \in L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$, *then the density of the measure* μ *is locally integrable to any power.*

For proofs of (i)–(iv) and (vi), see [12], [15, Chap. 1], [29], and [25]; the continuity in (v) is proved in [34], [33]; and the assertion about positivity of densities is established in [25] with the aid of a new version of the Harnack inequality (for the zero drift shown also in [34] and [33]).

By using examples constructed by Bauman [4] (see [15, section 1.6]), one can give examples of probability solutions to the stationary equation on the plane with continuous coefficients and a nondegenerate diffusion matrix such that their densities have no versions bounded on a ball. Thus, the condition for the continuity of a density is located between the continuity of A and the Dini mean oscillation condition. The latter condition so far is also the most general one for the Harnack inequality for positive solutions. The following question remains open.

Question 2. Is assertion (i) of the previous theorem valid for signed solutions?

In addition, it is not clear whether the assertion in [29] about the local exponential integrability of densities in the case of locally bounded coefficients is true: justification of Corollary 2.3 in [29] contains a gap. A result in this direction is obtained in [25], where a condition on the modulus of continuity of A is imposed.

In the previous theorem, local properties of solutions are considered, but there are also results on global properties, such as inclusions in Sobolev classes on the whole space and global estimates. Let us give the principal results (see [15], [13], and [56]). We start with estimates on the logarithmic gradient and a condition for existence of the Fisher information. Such estimates were first obtained in paper [17].

For $a^{ij} \in L^1_{\mathrm{loc}}(\mathbf{R}^d)$, we set

$$a := (a^1, \dots, a^d), \qquad a^j := \sum_{i=1}^d \partial_i a^{ij}.$$

Theorem 2.4. *Suppose that the mapping A is uniformly bounded and uniformly Lipschitz, and there exists $\alpha > 0$ such that $A(x) \geqslant \alpha \cdot \mathrm{I}$. Suppose also that a Borel probability measure μ on $\mathbf{R}^d$ satisfies $L^*_{A,b}\mu = 0$, where $|b| \in L^2(\mu)$. Then*

(i) *$\mu = \varrho\, dx$, where $\varrho = \varphi^2$ and $\varphi \in W^{2,1}(\mathbf{R}^d)$, which gives the inclusion $\varrho \in W^{1,1}(\mathbf{R}^d)$ and also $\varrho \in L^{d/(d-2)}(\mathbf{R}^d)$ if $d > 2$;*

(ii) $\displaystyle\int_{\mathbf{R}^d} \left|\frac{\nabla\varrho}{\varrho}\right|^2 \varrho\, dx = 4\int_{\mathbf{R}^d} |\nabla\varphi|^2\, dx \leqslant \frac{1}{\alpha^2}\int_{\mathbf{R}^d} |b-a|^2\, d\mu;$

(iii) *the mapping $\nabla\varrho/\varrho$ coincides μ-a.e. with the orthogonal projection of the vector field $A^{-1}(b-a)$ onto the closure of the set of mappings $\{\nabla u |\, u \in C_0^\infty(\mathbf{R}^d)\}$ in the space $L^2(\mu, \mathbf{R}^d)$ with the inner product*

$$\langle F, G\rangle_2 := \int_{\mathbf{R}^d} \langle AF, G\rangle\, d\mu,$$

and thus

$$\int_{\mathbf{R}^d} \left|\frac{\sqrt{A}\,\nabla\varrho}{\varrho}\right|^2 \varrho\, dx \leqslant \int_{\mathbf{R}^d} |A^{-1/2}(b-a)|^2\, d\mu.$$

If $A = \mathrm{I}$, then assertions (ii) *and* (iii) *are true with $\alpha = 1$ and $a = 0$.*

Let us give a condition for the global Sobolev differentiability.

Theorem 2.5. *Suppose that a Borel probability measure μ on $\mathbf{R}^d$ satisfies the equation $L^*_{A,b}\mu = 0$, where A and A^{-1} are uniformly bounded, A is Lipschitz and $|b| \in L^p(\mu)$ with some $p > d$. Then the continuous version ϱ of the density of the measure μ is uniformly bounded and $\varrho \in W^{p,1}(\mathbf{R}^d)$.*

It had been open for a long time whether $\varrho \in W^{1,1}(\mathbf{R}^d)$ when $|b| \in L^1(\mu)$ until it was shown in the paper [16] that, for $d > 1$, there is a probability solution μ to the equation $L^*_{\mathrm{I},b}\mu = 0$ for which $|b| \in L^1(\mu)$, but the density ϱ does not belong to $W^{1,1}(\mathbf{R}^d)$, i.e., $|\nabla\varrho|$ does not belong to $L^1(\mathbf{R}^d)$. This can be done even with a smooth drift b, but in the general case there is an example in which the function $|\nabla\varrho|$ is integrable on no ball. On the other hand, it is proved in the same paper that ϱ belongs to the fractional Sobolev class $H^{r,\alpha}(\mathbf{R}^d)$ whenever $1 < r < d/(d-1)$, $\alpha < 1 - d(r-1)/r$.

Finally, let us mention upper and lower bounds on densities. For simplification of formulations we consider the case of the unit diffusion matrix.

Theorem 2.6. *Suppose that a Borel probability measure $\mu = \varrho\, dx$ on $\mathbf{R}^d$ satisfies the equation $L^*_{\mathrm{I},b}\mu = 0$, where $|b| \in L^p(\mu)$ with some $p > d$. Suppose also that Φ is a positive function of class $W^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$ such that $\Phi \in L^1(\mu)$ and $|\nabla\Phi| \in L^p(\mu)$. Then there exists a number $C > 0$ such that*

$$\varrho(x) \leqslant \frac{C}{\Phi(x)}.$$

For example, if $|b| \in L^p(\mu)$ with $p > d$ and μ has all moments, then, for every $k > 0$, there is a number $C_k > 0$ such that $\varrho(x) \leqslant C_k(1+|x|)^{-k}$.

If it is known that $\exp\{\alpha|x|^\beta\} \in L^1(\mu)$ with some $\alpha, \beta > 0$ and $|b(x)| \leqslant C\exp\{\delta|x|^\beta\}$ with some $\delta < \alpha/d$, then, for every $r < \beta/d$, there is a number $C_r > 0$ such that $\varrho(x) \leqslant C_r \exp\{-r|x|^\beta\}$.

It is surprising that such estimates, fulfilled under broad assumptions, are often rather sharp, which is seen from the following lower bounds.

THEOREM 2.7. *Let ϱ be a continuous density of a probability solution to the equation $L^*_{\mathrm{I},b}\mu = 0$ such that $|b(x)| \leqslant V(|x|/\theta)$, where $\theta > 1$ and $V > 0$ is a continuous increasing function on $[0, +\infty)$. Then there exists a number $K > 0$ such that*

$$\varrho(x) \geqslant \varrho(0) \exp\{-K(1 + V(|x|)|x|)\}.$$

For example, if $|b(x)| \leqslant c_1|x|^\beta + c_2$, then

$$\varrho(x) \geqslant \varrho(0) \exp\{-K(1 + |x|^{\beta+1})\}.$$

If also $\limsup_{|x|\to\infty} |x|^{-\beta-1}\langle b(x), x\rangle < 0$, then one has the two-sided estimate

$$\exp\{-K_1(1 + |x|^{\beta+1})\} \leqslant \varrho(x) \leqslant \exp\{-K_2(1 + |x|^{\beta+1})\}.$$

Papers [5], [28], and [9] contain information about equations with the unit diffusion matrix and drifts of the following form: $b(x) = -x + v(x)$. If $v(x) = 0$, then we obtain the classical Ornstein–Uhlenbeck operator, for which a unique invariant probability measure is the standard Gaussian measure γ on $\mathbf{R}^d$. For a nonzero field v, usually there are no explicit solutions, but it is useful to describe the properties of solutions through their densities with respect to the measure γ rather than with respect to Lebesgue measure. In particular, it is shown in [28] that if $|v| \in L^1(\mu)$, then the density $f = d\mu/d\gamma$ for any $\alpha < 1/4$ satisfies the estimate

$$\int_{\mathbf{R}^d} f[\ln(f + 1)]^\alpha \, d\gamma \leqslant C(\alpha)\bigl[1 + \|v\|_{L^1(\mu)}\bigl(\ln(1 + \|v\|_{L^1(\mu)})\bigr)^\alpha\bigr]$$

with constant $C(\alpha)$ independent of d. It is proved in [9] that if $|v| \in L^p(\mu)$ with some $p > 2$, then $f[\ln(1 + f)]^\alpha \in L^1(\gamma)$ whenever $\alpha < \min(2, (p + 2)/4)$. If v is bounded, then $\exp\{\varepsilon|\ln\max(f, 1)|^2\} \in L^1(\gamma)$ for all $\varepsilon < (2\pi\| \, |v| \, \|_\infty)^{-2}$. In addition, $|\nabla f| \in L^p(\gamma)$ for all $p > 1$. Moreover, the latter is true under a weaker condition than the boundedness of v; namely, inclusion of $|v|$ to some Orlicz class is sufficient.

Solutions of unbounded variation to stationary equations on the whole space are considered in [15], [53], [54], and [55].

Finally, we note that the results about local properties of solutions remain also valid for equations on domains and manifolds, but the situation is different for global properties. For example, there exist connected manifolds with nonconstant positive integrable harmonic functions, which gives probability solutions to the equation $\Delta\mu = 0$ with the zero drift. Some global properties are transferred to manifolds under additional geometric assumptions like restrictions on curvature; see [27] and [15] on this topic.

3. Stationary equations: Uniqueness of solutions.

In the one-dimensional case, the stationary equation has the form $(A\mu)'' = (b\mu)'$, hence $(A\mu)' = b\mu + c$, where c is a constant. If $A = 1$, then $\mu' = b\mu + c$. In the case where the coefficient b is locally Lebesgue integrable, this equation is solved explicitly, and the density of the measure μ has the form

$$\varrho(x) = c_1 \exp\{B(x)\} + c\exp\{B(x)\} \int_0^x \exp\{-B(y)\} \, dy,$$

$$B(x) = \int_0^x b(y) \, dy.$$

A simple analysis shows (see [15, section 4.1]) that, in this case, at most one probability solution can exist. However, in the case of a locally Lebesgue nonintegrable coefficient b many linearly independent probability solutions can exist. For example, the measures with densities $\varrho_1(x) = 2(2\pi)^{-1/2}x^2\exp\{-x^2/2\}I_{(-\infty,0]}(x)$ and $\varrho_2(x) = \varrho_1(-x)$ satisfy the equation with $A = 1$ and $b(x) = -x + 2/x$. In this example, b even belongs to $L^2(\mu)$ for all solutions.

It was shown in [26] (see also [15, Example 4.2.1]) that, for $d > 1$, even the infinite differentiability of b does not guarantee the uniqueness of a probability solution to the equation $L^*_{\mathrm{I},b}\mu = 0$. An example for $\mathbf{R}^2$ is as follows:

$$b^1(x, y) = -x - 2ye^{(x^2-y^2)/2}, \qquad b^2(x, y) = -y - 2ye^{(y^2-x^2)/2}.$$

Here, one solution is the standard Gaussian measure, and another is given with respect to it by the smooth bounded density

$$c\int_{-\infty}^{x} e^{-s^2/2}\,ds + c\int_{-\infty}^{y} e^{-s^2/2}\,ds.$$

Moreover, effectively verified conditions are obtained in [60] on the coefficients to ensure the existence of infinitely many linearly independent probability solutions. In particular, this holds in the indicated explicit example. In [10], conditions are obtained for existence of two solutions to imply the existence of infinitely many linearly independent probability solutions. In [50], a method of constructing examples of nonuniqueness was suggested based on change of coordinates and passing to a degenerate boundary value problem.

Question 3. Let $A = \mathrm{I}$ and let b be infinitely differentiable. Suppose that a probability solution to the equation $L^*_{A,b} = \mu$ is not unique. Can the simplex of all probability solutions be finite-dimensional?

There are various sufficient conditions for uniqueness of probability solutions. The next result is proved in the paper [31].

THEOREM 3.1. *Suppose that A satisfies the Dini condition, A^{-1} is locally bounded, and $b^i \in L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$. Then a probability solution μ to the equation $L^*_{A,b}\mu = 0$ is unique if either of the following conditions is fulfilled*:

(i) $(1 + |x|)^{-2}|a^{ij}(x)|, (1 + |x|)^{-1}|b^i(x)| \in L^1(\mu)$;

(ii) *there exists a function $V \in C^2(\mathbf{R}^d)$ with $\lim_{|x|\to\infty} V(x) = +\infty$ and $L_{A,b}V \leqslant C_1 + C_2V$.*

If $a^{ij} \in W^{d+,1}_{\mathrm{loc}}(\mathbf{R}^d)$, then a sufficient condition for the uniqueness of a probability solution $\mu = \varrho\,dx$ is the μ-integrability of the functions a^{ij}, $b^i - \sum_j(\partial_{x_j}a^{ij} + a^{ij}\partial_{x_j}\varrho/\varrho)$.

Question 4. Let A and A^{-1} be bounded on $\mathbf{R}^d$ and $b(x) = -x$. Can several probability solutions exist?

A unique probability solution to the equation $L^*_{A,b}\mu = 0$ does not exclude the existence of nonzero signed solutions even on the real line (see [15, Example 4.1.3]). A sufficient condition for the absence of such solutions under the hypotheses of assertion (iii) in Theorem 2.3 is the existence of a function $V \in C^2(\mathbf{R}^d)$ for which $V(x) \to +\infty$ as $|x| \to +\infty$ and $L_{A,b}V(x) \geqslant -C$, $|\sqrt{A(x)}\nabla V(x)| \leqslant CV(x)$ with some number $C > 0$; see [15, Theorem 4.1.9]. Here, the inequality for $L_{A,b}V$ is opposite to the one which guarantees the existence of a probability solution, so it is assumed that there is a probability solution. However, if b is locally bounded, then [15, Corollary 4.3.7] gives the absence of nonzero signed solutions without this assumption.

4. Generated semigroups. Here, we discuss semigroups connected with solutions to stationary Fokker–Planck–Kolmogorov equations and also invariant measures of these semigroups.

Recall that a family of continuous linear operators T_t, $t \geqslant 0$, acting on a Banach space E, is called an operator semigroup if $T_0 = I$ and $T_{t+s} = T_t T_s$ for all $t, s \geqslant 0$. Such a semigroup is called strongly continuous (or a C_0-semigroup) if $\lim_{t\to 0} T_t x = x$ for every $x \in E$. It can be easily derived from the Banach–Steinhaus theorem that the mapping $t \mapsto T_t x$ is continuous on the whole half-line $[0, +\infty)$. It is known that, for a strongly continuous semigroup $\{T_t\}_{t\geqslant 0}$, the linear subspace $D(L)$ of vectors $h \in E$ such that there exists a limit $Lh = \lim_{t\to 0} t^{-1}(T_t h - h)$ with respect to the norm in E is everywhere dense. The operator L on the domain of definition $D(L)$ is called the generator of the semigroup. If $\varphi \in D(L)$, then one has the equality

$$T_t\varphi = \varphi + \int_0^t T_s L\varphi\, ds.$$

Typical generators of semigroups are elliptic operators, but usually the whole subspace $D(L)$ is not known in advance even for an operator generating a semigroup. The operators $L_{A,b}$ we consider are initially defined on the set $C_0^\infty(\mathbf{R}^d)$, so that there is no Banach space. If there is a probability measure μ on $\mathbf{R}^d$ satisfying $L_{A,b}^*\mu = 0$, then such natural Banach spaces appear: one can take $E = L^1(\mu)$ or $E = L^2(\mu)$. Why do we need a measure, and why not take for E the space of bounded continuous functions or the space of bounded Borel functions with the sup-norm? The point is that, in the case of unbounded coefficients A, b (or just b), the operator semigroups, which should be regarded as generated by our operators $L_{A,b}$, are usually not strongly continuous on such spaces. For example, this happens for the Ornstein–Uhlenbeck semigroup given by the explicit formula

$$T_t f(x) = \int_{\mathbf{R}^d} f\bigl(e^{-t}x - \sqrt{1-e^{-2t}}\, y\bigr)\, \gamma_d(dy),$$

where γ_d is the standard Gaussian measure (it is invariant with respect to this semigroup). Formal calculations show that the generator of this semigroup must be the Ornstein–Uhlenbeck operator

$$Lf(x) = \Delta f(x) - \langle x, \nabla f(x)\rangle,$$

and this is true indeed if for E we take $L^1(\gamma_d)$ or $L^2(\gamma_d)$ rather than the spaces of bounded functions indicated above.

It turns out that, in the general case, one can also associate with the operator $L_{A,b}$ some canonical strongly continuous semigroup on $L^1(\mu)$.

Recall that a bounded linear operator T on the space $L^p(\mu)$, where $1 \leqslant p \leqslant \infty$, is called sub-Markov if $0 \leqslant Tf \leqslant 1$ whenever $0 \leqslant f \leqslant 1$, $f \in L^\infty(\mu)$. If in addition $T1 = 1$, then T is called Markov.

The measure μ is called invariant with respect to an operator T on $L^\infty(\mu)$ (or on the space of bounded Borel functions in the case of a Borel measure on a topological space) if

$$\int Tf\, d\mu = \int f\, d\mu \quad \forall\, f \in L^\infty(\mu).$$

For a Borel probability measure on $\mathbf{R}^d$ and an operator T continuous with respect to the norm of $L^1(\mu)$, it suffices to have this equality for all functions f in $C_0^\infty(\mathbf{R}^d)$.

If, in place of this identity, the inequality

$$\int Tf\,d\mu \leqslant \int f\,d\mu$$

holds for all nonnegative functions f in $L^\infty(\mu)$, then μ is called a subinvariant measure for T.

Suppose now that a Borel probability measure μ on $\mathbf{R}^d$ satisfies the stationary equation $L^*_{A,b}\mu = 0$, where A is continuous, $\det A > 0$, $a^{ij} \in W^{d+,1}_{\mathrm{loc}}(\mathbf{R}^d)$, and $b^i \in L^{d+}_{\mathrm{loc}}(\mathbf{R}^d)$. As we know, the measure μ has a continuous positive density $\varrho \in W^{d+,1}_{\mathrm{loc}}(\mathbf{R}^d)$, and hence we can define the mappings

$$\beta_\mu = \frac{\nabla\varrho}{\varrho}, \quad \beta_{\mu,A} = (\beta^i_{\mu,A})^d_{i=1}, \quad \beta^i_{\mu,A} = \partial_{x_j} a^{ij} + a^{ij}\,\frac{\partial_{x_j}\varrho}{\varrho}.$$

The vector field

$$\widehat{b} = 2\beta_{\mu,A} - b$$

is called the dual drift. It is straightforward to verify the identity

$$\int \psi L_{A,b}\varphi\,d\mu = \int \varphi L_{A,\widehat{b}}\psi\,d\mu \quad \forall\,\varphi,\psi \in C^\infty_0(\mathbf{R}^d).$$

In addition,

$$L^*_{A,\widehat{b}}\mu = 0.$$

Let B_n be the ball of radius n centered at the origin in $\mathbf{R}^d$.

THEOREM 4.1. (i) *Under the stated assumptions, there exist closed extensions* $\big(L^\mu_{A,b}, D(L^\mu_{A,b})\big)$ *and* $\big(L^\mu_{A,\widehat{b}}, D(L^\mu_{A,\widehat{b}})\big)$ *of the operators* $\big(L_{A,b}, C^\infty_0(\mathbf{R}^d)\big)$ *and* $\big(L_{A,\widehat{b}}, C^\infty_0(\mathbf{R}^d)\big)$, *respectively, that are the generators of sub-Markov contracting* C_0*-semigroups* $\{T^\mu_t\}_{t\geqslant 0}$ *and* $\{\widehat{T}^\mu_t\}_{t\geqslant 0}$ *on* $L^1(\mu)$ *with the following properties*:

(a) *For every bounded measurable function* f *on* $\mathbf{R}^d$ *with compact support, the function* $(I - L^\mu_{A,b})^{-1}f$ *is the limit in* $L^1(\mu)$ *of the functions* u_n *that are solutions to the Dirichlet problems* $(I - L_{A,b})u_n = f$ *on* B_n *with zero boundary conditions, and the analogous assertion is true for the operator* $(I - L^\mu_{A,\widehat{b}})^{-1}$. *In addition, the measure* μ *is subinvariant for both semigroups* $\{T^\mu_t\}_{t\geqslant 0}$ *and* $\{\widehat{T}^\mu_t\}_{t\geqslant 0}$;

(b) *the indicated semigroups are adjoint, i.e.*,

$$\int_\Omega g\,T^\mu_t f\,d\mu = \int_\Omega f\,\widehat{T}^\mu_t g\,d\mu, \qquad f,g \in L^\infty(\mu). \tag{4.1}$$

The same equality is also true for the corresponding families of resolvents $\{R^\mu_\alpha\}_{\alpha>0}$ *and* $\{\widehat{R}^\mu_\alpha\}_{\alpha>0}$, *where* $R^\mu_\alpha = (\alpha\cdot I - L_{A,b})^{-1}$.

(ii) *The semigroup* $\{T^\mu_t\}_{t\geqslant 0}$ *has the following property*: *for every function* $\psi \in C^\infty_0(\mathbf{R}^d)$ *and every* $t \geqslant 0$, *the function* $T^\mu_t\psi$ *possesses a continuous modification such that as* $t \to 0$ *these modifications converge to* ψ *uniformly on compact sets.*

The described semigroup $\{T^\mu_t\}_{t\geqslant 0}$ is called canonical.

If the matrix A is locally Lipschitz, then, as shown in [11, Lemma 2.2], the canonical semigroup is the limit of the semigroups $\{T^k_t\}_{t\geqslant 0}$ corresponding to the operator $L_{A,b}$ on the balls B_k of radius $k \in \mathbb{N}$ with zero boundary conditions and defined on the spaces $L^1(\mu|_{B_k})$. In particular, if $f\in L^1(\mu)$, $T>0$ and $u_k=T_t^k f$ is the solution of the boundary value problem

$$\partial_t u_k = L_{A,b}u_k, \qquad u_k|_{\partial B_k\times[0,T]}=0, \quad u_k(x,0)=f(x) \text{ if } x\in B_k,$$

then $T_t^\mu f(x)=\lim_{k\to\infty}u_k(x,t)$ in $L^1(\mu)$ as $t\in[0,T]$.

In addition, under the same additional assumption of local Lipschitzness of A, it is shown in [11, Theorem 2.3] that the canonical semigroup $\{T_t^\mu\}_{t\geqslant 0}$ defines a minimal solution to the Cauchy problem

$$\partial_t u = L_{A,b}u, \quad u(x,0)=f(x) \tag{4.2}$$

in the following sense: for every $t>0$ the function $T_t^\mu f$ belongs to the Sobolev class $W^{p,2}_{\mathrm{loc}}(\mathbf{R}^d)$, for every ball U the function $\|T_t^\mu f\|^p_{W^{p,2}(U)}$ is integrable over each compact interval $[\tau,T_0]$ in $(0,T)$, in $U\times(\tau,T_0)$ there exists the Sobolev derivative $\partial_t u\in L^p(U\times(\tau,T_0))$, equality (4.2) for Sobolev derivatives is true almost everywhere, and the initial condition is fulfilled also in the sense of convergence in $L^1(\mu)$. The minimality is understood as follows: if f is a μ-integrable nonnegative continuous function and $v(x,t)$ is some nonnegative solution to this Cauchy problem with initial condition f in the weak sense, i.e., for all $t>0$, the function $v(\,\cdot\,,t)$ belongs to the Sobolev class $W^{2,1}(U)$ on every ball U, the function $\|v(\,\cdot\,,t)\|_{L^2(U)}$ is bounded on all intervals $[0,T_0]\subset[0,T)$, the function $\|\partial_x v(\,\cdot\,,t)\|^2_{L^2(U)}$ is integrable on $[0,T_0]$, and for every function $\psi\in C_0^\infty(\mathbf{R}^d)$ the equality

$$\begin{aligned}\int v(x,t)\psi(x)\,dx &- \int f(x)\psi(x)\,dx \\ = -\int_0^t \int \big[a^{ij}(x)&\partial_{x_j}\psi(x)\partial_{x_i}v(x,s) - b^i(x)\partial_{x_i}v(x,s)\psi(x) \\ &+ \partial_{x_j}a^{ij}(x)\partial_{x_i}v(x,s)\psi(x)\big]\,dx\,ds\end{aligned}$$

is true, then $T_t^\mu f(x,t)\leqslant v(x,t)$. The analogous assertion is true for the dual drift $\widehat{b}$ and the corresponding semigroup $\{\widehat{T}_t^\mu\}_{t\geqslant 0}$.

It is important to note that the generator of the canonical semigroup is an extension of the closure of the operator $L_{A,b}$ on $C_0^\infty(\mathbf{R}^d)$ in $L^1(\mu)$, but it can be a strict extension. The following fact is true (see [15, Proposition 5.2.5 and Theorem 5.3.1]).

THEOREM 4.2. *Under the assumptions about A and b stated before Theorem* 4.1 *the following conditions are equivalent.*

(i) *The indicated closure of $L_{A,b}$ is the generator of a strongly continuous operator semigroup on $L^1(\mu)$.*

(ii) *The set $(L_{A,b}-I)(C_0^\infty(\mathbf{R}^d))$ is everywhere dense in $L^1(\mu)$.*

(iii) *There exists a unique strongly continuous operator semigroup on $L^1(\mu)$ whose generator is an extension of $L_{A,b}$ on $C_0^\infty(\mathbf{R}^d)$.*

(iv) *The measure μ is invariant with respect to $\{T_t^\mu\}_{t\geqslant 0}$.*

(v) *The equality $T_t^\mu 1=1$ holds, i.e., $\{T_t^\mu\}_{t\geqslant 0}$ is a Markov semigroup.*

Under either of these conditions the canonical semigroups $\{T_t^\mu\}_{t\geqslant 0}$ and $\{\widehat{T}_t^\mu\}_{t\geqslant 0}$ are Markov, and the measure μ is invariant for both. In addition, the measure μ is a unique probability solution to the equation $L^*_{A,b}\mu=0$.

Canonical semigroups are not always unique strongly continuous semigroups on $L^1(\mu)$ whose generators extend $L_{A,b}$ and $L_{A,\widehat{b}}$: see examples in [15]. In addition, the measure μ is not always invariant for these semigroups.

Question 5. Can it happen that there is a strongly continuous sub-Markov semigroup on $L^1(\mu)$ that differs from the canonical one and whose generator extends $L_{A,b}$ if $A = \mathrm{I}$ and b is smooth?

Sufficient conditions for invariance of μ are given in [15, Chap. 5]. Note that invariance of μ for one of the two semigroups is equivalent to invariance with respect to the other (see [15, Remark 5.2.4]).

Theorem 4.3. *Conditions* (i)–(v) *are fulfilled if there is a compact function* $V \in C^2(\mathbf{R}^d)$ *and numbers* $\alpha > 0$ *and* $R > 0$ *for which*

$$L_{A,b}V(x) \leqslant \alpha V(x) \quad \textit{for a.e. } x \textit{ with } |x| \geqslant R.$$

For example, it suffices to have the following estimate outside a ball:

$$-\frac{2}{1+|x|^2}\,\langle A(x)x, x\rangle + \operatorname{trace} A(x) + \langle b(x), x\rangle \leqslant C|x|^2 \ln|x|.$$

For $A = \mathrm{I}$, a sufficient condition is the estimate $|b(x)| \leqslant C + C|x|\ln|x|$. However, here $\ln|x|$ cannot be replaced by $|\ln|x||^r$ with $r > 1$.

There are sufficient conditions without Lyapunov functions. For example, in the case $A = \mathrm{I}$ the integrability of $|b(x)|/(1+|x|)$ with respect to μ is sufficient. Yet another sufficient condition for invariance of the measure μ in terms of μ itself is this: $|b - \nabla\varrho/\varrho| \in L^1(\mu)$. In the case of a nonconstant A, invariance of μ for $\{T_t^\mu\}_{t\geqslant 0}$ is ensured by the inclusions $a^{ij}, |b - \beta_{A,\mu}| \in L^1(\mu)$, which follows from justification of Example 5.5.3 and Theorem 5.3.1 in [15]. In particular, if $b = \beta_{A,\mu}$ and $a^{ij} \in L^1(\mu)$, then invariance holds.

Let us observe that if we divide the coefficients of the operator $L_{A,b}$ by the function $\theta = |L_{A,b}V| + 1$, taking some smooth compact function V, then the new operator $L_{A/\theta,b/\theta} = \theta^{-1}L_{A,b}$ will satisfy the estimate from Theorem 4.3. However, the original measure μ need not satisfy the equation with this operator. The equation $(\theta^{-1}L_{A,b})^*\nu = 0$ is obviously satisfied by the measure $\nu = \theta\cdot\mu$. Suppose that $L_{A,b}V \in L^1(\mu)$. Then we can assume that ν is a probability measure. Nevertheless, we still cannot apply Theorem 4.3, because it contains some local conditions on the coefficients. Major problems are connected with the new diffusion coefficient A/θ, since $L_{A,b}V$ includes the term $\langle b, \nabla V\rangle$. If the functions b^i belong to $W^{d+,1}_{\mathrm{loc}}(\mathbf{R}^d)$, then the required local conditions are fulfilled. In particular, all hypotheses of Theorem 4.3 are fulfilled for the new operator if b and V are smooth. But what conclusion does this theorem enable us to derive? It says that the measure $\theta\cdot\mu$ is a unique probability solution to the equation with $\theta^{-1}L_{A,b}$ and is a unique invariant probability measure for the corresponding canonical semigroup. However, for the original equation it only guarantees uniqueness among probability solutions with respect to which the function $L_{A,b}V$ is integrable.

With the aid of the canonical semigroup one can construct a solution to the Cauchy problem for the parabolic Fokker–Planck–Kolmogorov equation discussed in the subsequent sections. Next, we mention some properties of this semigroup important for this procedure, while the assertion about parabolic equations is postponed until section 5.

Theorem 4.4. *Let* μ *be a probability measure on* $\mathbf{R}^d$ *and let* $L^*_{A,b}\mu = 0$, *where* $a^{ij} \in C(\mathbf{R}^d) \cap W^{p,1}_{\mathrm{loc}}(\mathbf{R}^d)$, $\det A > 0$, $b^i \in L^p_{\mathrm{loc}}(\mathbf{R}^d)$, *and* $p > d+2$. *Then there exists a locally Hölder continuous positive function* $p_{A,b}(t,x,y)$ *on* $(0,+\infty) \times \mathbf{R}^d \times \mathbf{R}^d$ *such that the measures*

$$K_t(x,dy) = p_{A,b}(t,x,y)\,dy$$

are subprobabilities, and, for every function $f \in L^1(\mu)$, *the function*

$$x \mapsto K_t f(x) := \int_\Omega f(y)\,p_{A,b}(t,x,y)\,dy$$

serves as a μ*-version of* $T^\mu_t f$ *such that the function* $(t,x) \mapsto K_t f(x)$ *is continuous on the product* $(0,+\infty) \times \mathbf{R}^d$.

In addition, if there is a bounded Borel measure ν *invariant for* $\{K_t\}_{t\geqslant 0}$, *i.e.*,

$$\nu = K^*_t \nu(dy) := \int_{\mathbf{R}^d} K_t(x,dy)\,\nu(dx) \quad \forall\, t \geqslant 0,$$

then $\nu = c\mu$ *for some constant* c. *In particular, if* $\nu \neq 0$, *then the measure* μ *itself is also invariant. Hence* $\{K_t\}_{t\geqslant 0}$ *cannot have invariant probability measures different from* μ.

Question 6. It is not known whether this theorem is true under our usual assumption $p > d$ in place of $p > d+2$.

Uniqueness of a probability invariant measure for the diffusion semigroup generated by the operator $L_{\mathrm{I},b}$ with a smooth drift was proved by Varadhan [66, section 31], who also raised a question about generalization of this result to more general coefficients. Such generalizations were obtained in [1], [12], and [15]; the previous theorem gives a typical result.

Question 7. What are optimal conditions on A and b for uniqueness of invariant probability measures for semigroups whose generators extend $L_{A,b}$?

Remark 4.1. It is asserted in [43, Lemma 5.4] that the existence of an invariant probability measure for the diffusion process (with three times differentiable coefficients) is equivalent to the existence of a positive solution $L_{A,b}u = -1$ on the complement to a compact set. Let us show that this is true under our local assumptions about the coefficients if, in addition, we have the following estimates (as in the paper [42]):

$$|a^{ij}(x)| \leqslant C + C|x|^2, \quad |b^i(x)| \leqslant C + C|x|.$$

Indeed, in this case the existence of a probability solution μ to the equation $L^*_{A,b}\mu = 0$ by Theorem 2.2 yields a nonnegative Lyapunov function $V \in W^{d+,2}_{\mathrm{loc}}(\mathbf{R}^d)$ such that $V(x) \to +\infty$ and $L_{A,b}V(x) \to -\infty$ as $|x| \to +\infty$. Let us take a closed ball U centered at the origin outside of which $L_{A,b}V(x) \leqslant -1$. For any ball U_n of radius n large enough we take the solution u_n to the Dirichlet problem $L_{A,b}u_n = -1$ on the ring $U_n \setminus K$ with zero boundary condition. These solutions are nonnegative (positive on the interiors of the rings). By the maximum principle $u_n \leqslant u_{n+1}$ on $U_n \setminus K$, since $L_{A,b}(u_{n+1} - u_n) = 0$ on $U_n \setminus U$, $u_{n+1} - u_n = 0$ on ∂U and $u_{n+1} - u_n \geqslant 0$ on ∂U_n because $u_n = 0$ on ∂U_n and $u_{n+1} \geqslant 0$. In addition, $u_n \leqslant V$ on $U_n \setminus U$, since $L_{A,b}(V - u_n) \leqslant 0$ on $U_n \setminus U$ and $V - u_n \geqslant 0$ on the boundary of $U_n \setminus U$. Therefore, there exists a finite positive limit $u(x) = \lim_{n\to\infty} u_n(x)$ outside U. It is readily verified that, under our conditions on A and b, the restrictions of the functions u_n to every ball Ω outside U are bounded with respect to the Sobolev norm in $W^{p,2}(\Omega)$ with some $p = p(\Omega) > d$. Hence the function u belongs to $W^{d+,2}_{\mathrm{loc}}(\mathbf{R}^d)$ and satisfies $L_{A,b}u = -1$ outside U.

Conversely, if a positive function $u \in W^{d+,2}_{\mathrm{loc}}(\mathbf{R}^d)$ satisfies $L_{A,b}u = -1$ outside a ball U, then the function $V(x) = u(x) + \varepsilon \ln(|x|^2 + 1)$ for small $\varepsilon > 0$ satisfies the inequality $L_{A,b}V(x) \leqslant -1/2$ outside U. It belongs to $W^{d+,2}_{\mathrm{loc}}(\mathbf{R}^d)$, and we have $V(x) \to +\infty$ as $|x| \to +\infty$.

However, without the indicated restriction on the growth of A and b, no invariant probability measures for the canonical semigroup $\{T^{\mu}_t\}_{t\geqslant 0}$ can exist even in the case of a unique probability solution to the equation $L^*_{A,b}\mu = 0$. Let us consider such an example on the real line. Let

$$b(x) = -2x + 6e^{x^2}, \quad A(x) = 1, \quad Lu = u'' + bu'.$$

The unique probability solution μ to the stationary equation has density $\pi^{-1/2}e^{-x^2}$. The function

$$w(x) = \int_{-\infty}^{x} e^{-s^2}\, ds$$

satisfies the condition $L_{1,b}w(x) = -4e^{-x^2} + 6 \geqslant w(x)$. Hence there are no invariant probability measures for the canonical semigroup (see Exercise 5.6.49 in [15], according to which the measure μ is not invariant, but by [15, Theorem 5.4.5] there are no other invariant probability measures). However, there is a positive solution to the equation $L_{1,b}u = -1$ on the whole real line. Indeed, we set

$$B(x) = \int_0^x b(s)\, ds = -x^2 + 6\int_0^x e^{s^2}\, ds.$$

Solving the equation

$$u'' + bu' = -1,$$

we obtain

$$u(x) = C_2 - \int_0^x \biggl(C_1 + \int_0^t e^{B(s)}\, ds\biggr) e^{-B(t)}\, dt.$$

Since $B(s) \leqslant -s^2$ if $s \leqslant 0$, the integral of $e^{B(s)}$ over $(-\infty, 0]$ is less than $\sqrt{\pi} < 2$. Let $C_1 = 2$. Then, for all t,

$$2 + \int_0^t e^{B(s)}\, ds \geqslant 0,$$

and, for $x \leqslant 0$, we have

$$-\int_0^x \biggl(2 + \int_0^t e^{B(s)}\, ds\biggr) e^{-B(t)}\, dt \geqslant 0.$$

Observe that

$$\lim_{t\to+\infty} e^{t^2 - B(t)} \int_0^t e^{B(s)}\, ds = \lim_{t\to+\infty} \frac{e^{B(t)}}{(-2t + e^{t^2})e^{B(t)-t^2}} = 1.$$

Therefore, letting $t \to +\infty$, we obtain

$$e^{-B(t)} \int_0^t e^{B(s)}\, ds \sim e^{-t^2}, \qquad \int_0^{+\infty} \biggl(2 + \int_0^t e^{B(s)}\, ds\biggr) e^{-B(t)}\, dt < \infty.$$

Taking a sufficiently large constant C_2, we conclude that $u > 0$. Thus, there exists a positive solution to the equation $Lu = -1$ on the whole real line, but there is no invariant measure for the minimal semigroup. It is now easy to construct an example of a smooth field b on the plane for which there is a positive solution to the equation $L_{\mathrm{I},b}u = -1$, but there are no probability solutions to the stationary equation, and hence no invariant measures for the semigroup. To this end, we set $b^1(x,y) = b(x)$, $b^2(x,y) = 0$, where b is given as above. If a probability solution to the stationary equation exists, then its projection on the axis of ordinates is a probability solution with zero drift, which is impossible. Apparently, the other implication in Lemma 5.4 of [43] is also false without additional restrictions on the growth of coefficients.

Note that in this survey we discuss only analytic aspects of the theory of Fokker–Planck–Kolmogorov equations but do not touch upon questions about existence and properties of diffusion processes connected with these equations. On such questions, in the case of irregular coefficients of the equation, see [51], [52], [53], [54], [55], and [64].

5. Evolution equations: Existence of solutions and their properties. Let us proceed to parabolic equations. Suppose we are given Borel functions a^{ij}, b^i on $\mathbf{R}^d_T := \mathbf{R}^d \times (0,T)$, $T > 0$, and the matrix $A(x,t) = (a^{ij}(x,t))_{i,j\leqslant d}$ is nonnegative definite. The parabolic Fokker–Planck–Kolmogorov equation

$$\partial_t \mu = \partial_{x_i}\partial_{x_j}(a^{ij}\mu) - \partial_{x_i}(b^i\mu) \tag{5.1}$$

for a Borel measure μ on $\mathbf{R}^d_T$ (possibly, signed) is understood similarly to the elliptic case: the coefficients a^{ij}, b^i are integrable with respect to $|\mu|$ on compact sets in $\mathbf{R}^d_T$, and, for every function $\varphi \in C_0^\infty(\mathbf{R}^d_T)$,

$$\int_{\mathbf{R}^d_T} [\partial_t\varphi + L_{A,b}\varphi]\, d\mu = 0.$$

Let us introduce the Cauchy problem for (5.1) with an initial condition. We do this in the special case where the measure μ is represented as

$$\mu(dx\,dt) = \mu_t(dx)\,dt$$

by means of a family of locally finite measures μ_t on $\mathbf{R}^d$ such that, for every Borel set B with compact closure, the function $t \mapsto \mu_t(B)$ is Lebesgue measurable and the function $t \mapsto |\mu_t|(B)$ is integrable on compact intervals in $(0,T)$. Then, for every bounded Borel function f with compact support, its integral against μ_t is Lebesgue measurable in t and the previous equality by definition means that

$$\int_{\mathbf{R}^d_T} f\,d\mu = \int_0^T \int_{\mathbf{R}^d} f(x,t)\,\mu_t(dx)\,dt.$$

Such measures μ_t exist under broad assumptions, in particular, if μ is absolutely continuous. We write $\mu = (\mu_t)_{t\in(0,T)}$ or $\mu = \mu_t\,dt$.

We call a locally finite Borel measure ν on $\mathbf{R}^d$ the initial condition for $\mu = (\mu_t)_{t\in(0,T)}$ and write $\mu|_{t=0} = \nu$ or $\mu_0 = \nu$ if, for every function $f \in C_0^\infty(\mathbf{R}^d)$, there exists a full measure set $J_f \subset (0,T)$ such that

$$\int_{\mathbf{R}^d} f(x)\,\nu(dx) = \lim_{t\to 0,\, t\in J_f} \int_{\mathbf{R}^d} f(x)\,\mu_t(dx). \tag{5.2}$$

This condition is weaker than weak convergence of the measures μ_t to the measure ν as $t \to 0$. If the integral of f against μ_t is continuous on $(0,T)$, then $J_f = (0,T)$. This will be the case under the conditions on the coefficients imposed below.

The Cauchy problem in this sense will be written in the form

$$\partial_t \mu = L^*_{A,b}\mu, \quad \mu|_{t=0} = \nu. \tag{5.3}$$

If A and b are bounded on all sets of the form $U \times [0,T]$, where U is a ball in $\mathbf{R}^d$, then (5.3) is equivalent to the identity

$$\int \varphi\, d\mu_t - \int \varphi\, d\nu = \int_0^t \int L_{A,b}\varphi\, d\mu_s\, ds \tag{5.4}$$

for every function $\varphi \in C_0^\infty(\mathbf{R}^d)$ and almost all t (with the corresponding measure zero set depending on φ).

For a given probability measure ν, by $\mathcal{M}_\nu$ we denote the set of all nonnegative solutions $\mu = \mu_t\, dt$ to problem (5.3) for which $\mu_t(\mathbf{R}^d) \leqslant 1$ for almost all points t, i.e., almost all measures μ_t are subprobabilities.

For parabolic equations, there are also a priori estimates with Lyapunov functions (see [15, section 7.1]), and we give a typical result with a simple formulation.

THEOREM 5.1. *Let $\mu = (\mu_t)_{0<t<T}$ be a solution to the Cauchy problem with initial condition ν that is a subprobability measure on $\mathbf{R}^d$ such that all μ_t are also subprobability measures. Suppose that there is a positive function $W \in C^2(\mathbf{R}^d)$ such that $\lim_{|x|\to+\infty} W(x) = +\infty$ and, for some number $C > 0$ and all $(x,t) \in \mathbf{R}^d_T$,*

$$L_{A,b}W(x,t) \leqslant C + CW(x).$$

Then, for almost all $t \in (0,T)$,

$$\int_{\mathbf{R}^d} W(x)\,\mu_t(dx) \leqslant \exp\{Ct\} + \exp\{Ct\}\int_{\mathbf{R}^d} W(x)\,\nu(dx).$$

For example, if we have the estimates

$$|a^{ij}(x,t)| \leqslant C(1+|x|^2), \qquad |b^i(x,t)| \leqslant C(1+|x|), \tag{5.5}$$

and the measure ν has finite moment of order r, then

$$\int_{\mathbf{R}^d} |x|^r\,\mu_t(dx) \leqslant e^{c_1 t} - 1 + e^{c_2 t}\int_{\mathbf{R}^d} |x|^r\,\nu(dx).$$

Conditions for existence of solutions to the parabolic equation are considerably broader than for the elliptic one.

THEOREM 5.2. *Let ν be a probability measure on $\mathbf{R}^d$.*

(i) *The set $\mathcal{M}_\nu$ is nonempty if A, A^{-1}, and b are bounded on sets of the form $U \times [0,T]$, where U is a ball.*

(ii) *The local boundedness of b in* (i) *can be replaced by the condition $|b| \in L^p(U \times [0,T])$ with some $p > d+2$ if $a^{ij}(\,\cdot\,,t) \in W^{p,1}_{\mathrm{loc}}(\mathbf{R}^d)$ and $\sup_{t\in(0,T)} \|a^{ij}(\,\cdot\,,t)\|_{W^{p,1}(U)} < \infty$ for every ball U.*

(iii) *Finally, if there exist a function $V \in C^2(\mathbf{R}^d)$ and a number $C \geqslant 0$ such that*

$$\lim_{|x|\to+\infty} V(x) = +\infty, \quad L_{A,b}V(x,t) \leqslant C + CV(x),$$

then, for every solution, almost all measures μ_t are probabilities.

Note that in case (ii) in this theorem one can find a version $\varrho(x,t)$ of the solution density continuous on $\mathbf{R}^d\times(0,T)$ (see below) and for this version all measures μ_t will be probabilities if $V\in L^1(\nu)$. Indeed, by Theorem 5.1 there exists a number M such that, for almost all $t\in(0,T)$, the integral of $V(x)\varrho(x,t)$ over $\mathbf{R}^d$ does not exceed M. Due to the continuity of $\varrho(x,t)$ in t this is true for all t. Since $V(x)\to+\infty$ as $|x|\to+\infty$, the measures $\varrho(\,\cdot\,,t)\,dx$ are uniformly tight, whence it follows that they are all probabilities.

In the general case, this is false even for $d=1$, $A(x)=1$, and a smooth drift. Let us give an example of a smooth solution $\varrho(x,t)$ to the equation

$$\partial_t\varrho=\partial_x^2\varrho-\partial_x(b\varrho),$$

where $b\in C^\infty(\mathbf{R}^2)$, such that the function $x\mapsto\varrho(t,x)$ for all $t\neq1$ is a probability density, but for $t=1$ it is not. Let σ be a smooth positive probability density on the real line, for example, the standard Gaussian density. We set

$$\varrho(x,t)=\frac{1}{2}[(1-t)^2\sigma((1-t)^2x)+\sigma(x)],$$
$$b(x,t)=\frac{1}{\varrho(x,t)}\int_0^x[\partial_y^2\varrho(y,t)-\partial_t\varrho(y,t)]\,dy.$$

Both functions are smooth, $\varrho>0$, and for $t\neq1$ the function $x\mapsto\varrho(x,t)$ is the half-sum of two probability densities, so it is also a probability density, but for $t=1$ we have $\varrho(x,1)=\sigma(x)/2$. Here, the equality $\partial_x(b\varrho)=\partial_x^2\varrho-\partial_t\varrho$ holds.

Let us consider the special case where A and b do not depend on t and there exists a probability solution μ to the stationary equation $L_{A,b}^*\mu=0$. In this case, a solution to the Cauchy problem (5.3) can be obtained explicitly with the aid of the canonical semigroup in the situation of Theorem 4.4. If a bounded measure ν is given on $\mathbf{R}^d$ and

$$K_t^*\nu(dy):=\int_{\mathbf{R}^d}K_t(x,dy)\,\nu(dx)=\int_{\mathbf{R}^d}p_{A,b}(t,x,y)\,\nu(dx)\,dy,$$

then the measure $\sigma=K_t^*\nu(dy)\,dt$ satisfies $\partial_t\sigma=L_{A,b}^*\sigma$ for all $T>0$. In addition, it gives a solution to the Cauchy problem (5.3). Note that, for absolutely continuous measures ν, these properties follow at once from the properties of canonical semigroups. Indeed, for $\nu=g\cdot\mu$ we can take $\nu_t=\widehat{T}_t^\mu g\cdot\mu$. Then, for every function $\varphi\in C_0^\infty(\mathbf{R}^d)$,

$$\int_{\mathbf{R}^d}T_t^\mu\varphi g\,d\mu=\int_{\mathbf{R}^d}\varphi g\,d\mu+\int_0^t\int_{\mathbf{R}^d}T_s^\mu L_{A,b}\varphi g\,d\mu\,ds,$$

which can be written as

$$\int_{\mathbf{R}^d}\varphi\,d\nu_t=\int_{\mathbf{R}^d}\varphi g\,d\nu+\int_0^t\int_{\mathbf{R}^d}L_{A,b}\varphi g\,d\nu_s\,ds.$$

The following is known about densities of solutions to parabolic equations.

We first recall that a function f on $\mathbf{R}^d\times(0,+\infty)$ belongs to the class VMO_x with respect to the variable x if there exists a modulus of continuity ω_0 such that

$$\sup_{\substack{(x_0,t)\in\mathbf{R}^d\times(0,+\infty)\\0<r\leqslant R}} r^{-2d-2}\int_t^{t+r^2}\int_{\substack{|x-x_0|<r\\|y-x_0|<r}}|f(x,s)-f(y,s)|\,dx\,dy\,ds\leqslant\omega_0(R).$$

Theorem 5.3. *Suppose that a measure* μ *on* $\mathbf{R}^d \times (0,T)$ *satisfies* (5.1).

(i) *If* $\mu \geqslant 0$ *and* $\det A > 0$, *then* μ *has a density* ϱ.

(ii) *If* A *on compact sets is Hölder continuous in* x *uniformly in* t *and* $\det A > 0$, *then* μ *has a density* ϱ *also for signed solutions.*

(iii) *If* A *and* A^{-1} *are locally bounded,* $a^{ij} \in \mathrm{VMO}_x$, $b^i \in L^q_{\mathrm{loc}}(\mathbf{R}^d \times (0,+\infty))$, *where* $q > d+2$, *then* μ *has a density in all* $L^p_{\mathrm{loc}}(\mathbf{R}^d \times (0,+\infty))$.

(iv) *If in* (iii) *the matrix* A *satisfies the Dini condition, then there exists a continuous version of the density of the measure* μ.

(v) *If in* (iv) A *and* A^{-1} *are globally bounded and* $|b| \in L^p(|\mu|)$, *where* $p > d+2$, *then the solution density is bounded on every set* $\mathbf{R}^d \times [r_1, r_2]$, $r_1, r_2 > 0$.

(vi) *If* A *and* A^{-1} *are locally bounded, for some* $p > d+2$ *for every ball* U *we have* $\sup_t \|a^{ij}(\,\cdot\,,t)\|_{W^{p,1}(U)} < \infty$, *and* $b^i \in L^p_{\mathrm{loc}}(\mathbf{R}^d \times (0,+\infty))$, *then* μ *has a continuous density* ϱ, *and for every ball* U *and every* $r_1, r_2 > 0$,

$$\int_{r_1}^{r_2} \|\varrho(\,\cdot\,,t)\|^p_{W^{p,1}(U)}\,dt < \infty.$$

The proofs of these assertions can be found in [15] and [30].

As in the elliptic case, it remains open whether (i) holds for signed solutions.

In the parabolic case, there are also results about global properties of densities, such as membership in Sobolev classes, and upper and lower bounds, which are analogous to the results presented above in the elliptic case (see [57] and [15, Chaps. 7 and 8]).

Let us return to the Kolmogorov conditions (1.3) and (1.4) and show that they are fulfilled for continuous coefficients with estimates (5.5). In our notation, in the multidimensional case, these conditions are the relationships

$$a^{ij}(s,x) = \lim_{\delta\to 0} \frac{1}{2\delta} \int_{\mathbf{R}^d} (y_i - x_i)(y_j - x_j)\,\mu_{s+\delta}(dy), \tag{5.6}$$

$$b^i(s,x) = \lim_{\delta\to 0} \frac{1}{\delta} \int_{\mathbf{R}^d} (y_i - x_i)\,\mu_{s+\delta}(dy), \tag{5.7}$$

where (μ_t) is a solution to the Cauchy problem with initial condition $\nu = \delta_x$ at time s.

Theorem 5.4. *Let the coefficients* A *and* b *be continuous and satisfy estimates* (5.5). *Then equalities* (5.6) *and* (5.7) *hold.*

Proof. Suppose for notational simplicity that the coefficients do not depend on time and $s = 0$, $x = 0$. Under our assumptions the measures μ_t have all moments, and the moment of order r is uniformly bounded in t from every compact interval (see [15, section 7.1]). Hence the defining identity (5.4) is fulfilled also for the functions $\varphi(y) = y_i y_j$ rather than just for functions with compact support. For the former functions, the integral form of the Cauchy problem with the initial condition at $t = 0$ equal to the Dirac measure at zero is written as

$$\int_{\mathbf{R}^d} y_i y_j\,\mu_\delta(dy) = \int_0^\delta \int_{\mathbf{R}^d} [2a^{ij}(y) + y_i b^j(y) + y_j b^i(y)]\,\mu_t(dy)\,dt.$$

We have to verify that the integral of $\delta^{-1}[2a^{ij}(y) + y_i b^j(y) + y_j b^i(y)]$ against the measure $\mu_t(dy)\,dt$ over $\mathbf{R}^d \times [0,\delta]$ tends to $2a^{ij}(0)$ as $\delta \to 0$. To this end, it suffices to establish the following: if a function g on $\mathbf{R}^d$ is continuous, $g(0) = 0$, and $|g(y)| \leqslant C + C|y|^2$, then the integral of $\delta^{-1}g(y)$ against the measure $\mu_t(dy)\,dt$

over $\mathbf{R}^d \times [0,\delta]$ tends to zero as $\delta \to 0$. If the function g vanishes outside a ball centered at the origin, then this is true due to the aforementioned uniform boundedness of moments of the measures μ_t. Hence our assertion reduces to the case of a function g with support in a ball U. Moreover, with the aid of uniform approximations, we can assume that this function belongs to $C_0^\infty(\mathbf{R}^d)$. In this case,

$$\delta^{-1}\int_{\mathbf{R}^d} g\,d\mu_\delta = \delta^{-1}\int_0^\delta\int_{\mathbf{R}^d} L_{A,b}g\,d\mu_s\,ds,$$

since the integral of g against the Dirac measure at zero is zero. The absolute value of the right-hand side is estimated by $M := \sup_U |L_{A,b}g|$. Hence the absolute value of the integral of the function $\delta^{-1}g$ against the measure $\mu_t\,dt$ over $\mathbf{R}^d \times [0,\delta]$ does not exceed $M\delta$.

If we take $\varphi(y) = y_i$, then similarly we find that the integral of y_i against μ_t equals the integral of b^i against the measure $\mu_t\,dt$ over $\mathbf{R}^d \times [0,\delta]$, which after dividing by δ tends to $b^i(0)$ as $\delta \to 0$ by the same reasoning. The same justification works in the case of coefficients depending on time. Theorem 5.4 is proved.

Let us also mention the so-called Ambrosio–Figalli–Trevisan superposition principle (see [2], [38], and [65]) connecting solutions to the Cauchy problem for the Fokker–Planck–Kolmogorov equation with solutions to martingale problems. According to this principle, under suitable conditions on A and b, for every probability solution (μ_t) to the Cauchy problem (5.3), such that the mapping $t \mapsto \mu_t$ is continuous in the weak topology, there exists a probability measure P_ν on the space of continuous trajectories $\Omega = C([0,T],\mathbf{R}^d)$ such that ν is the distribution of $\omega(0)$, μ_t with $t > 0$ is the distribution of $\omega(t)$, and for every function $f \in C_0^\infty(\mathbb{R}^d)$ the process

$$\xi(\omega,t) = f(\omega(t)) - f(\omega(0)) - \int_0^t L_{A,b}f(\omega(s),s)\,ds$$

is a martingale with respect to the measure P_ν and the filtration $\mathcal{F}_t = \sigma(\omega(s)\colon s \leqslant t)$, $t \geqslant 0$. The most general sufficient condition on A and b known so far to ensure such a representation is obtained in paper [23]:

$$\int_0^T\int_{\mathbf{R}^d} \frac{\|A(x,t)\| + |\langle b(x,t),x\rangle|}{1+|x|^2}\,\mu_t(dx)\,dt < \infty.$$

In terms of the coefficients without reference to the solution, the following estimate is sufficient:

$$\|A(x,t)\| + |\langle b(x,t),x\rangle| \leqslant C(1+|x|^2).$$

The following question remains open.

Question 8. Does the superposition principle follow from the existence of a Lyapunov function $V \in C^2(\mathbf{R}^d)$ such that $V(x) \to +\infty$ as $|x| \to +\infty$ and $L_{A,b}V(x,t) \leqslant CV(x)$?

6. Evolution equations: Uniqueness of solutions. Some sufficient conditions for uniqueness of probability solutions to the Cauchy problem for the Fokker–Planck–Kolmogorov equation follow from general results of the older papers [40], [3], and [62]. For example, for smooth coefficients and a nondegenerate diffusion matrix, these results are applicable in the case of the estimates $\|A(x,t)\| \leqslant C(1+|x|^2)$, $|b(x,t)| \leqslant C(1+|x|)$. A close problem of uniqueness of semigroups was studied by such classical authors as Feller [35], [36], [37], Yosida [67], and Hille [41]. In the case of coefficients independent of time, a complete answer to the question about uniqueness of the solution to the Cauchy problem for the Fokker–Planck–Kolmogorov equations is given in [11], where it was shown that, in any dimension greater than 1, there is no uniqueness even for the unit diffusion matrix and a smooth drift, while in dimension 1 uniqueness holds; more precisely, the following theorem is true.

Theorem 6.1. *Let* (μ_t) *be a solution to the Cauchy problem* $\partial_t\mu_t = \partial_x^2\mu_t - \partial_x(b\mu_t)$, $\mu_0 = \nu$, *where all measures* μ_t *are probabilities. If the coefficient* b *is locally bounded and depends only on* x, *then such a solution is unique.*

Question 9. Is Theorem 6.1 true if b depends on both variables?

In Example 4.5 in [11], besides a unique probability solution, there is another positive solution with finite measures μ_t, so it is important that we deal with probability measures μ_t. For a nonconstant diffusion coefficient, the following result is obtained in the cited paper.

Theorem 6.2. *Let* a *be a positive locally Lipschitz function on* $\mathbf{R}$, *and let* b *be a locally bounded Borel function on* $\mathbf{R}$. *Suppose that*

$$\int_{-\infty}^{0} \frac{1}{\sqrt{a(x)}}\,dx = \int_{0}^{+\infty} \frac{1}{\sqrt{a(x)}}\,dx = +\infty.$$

If a probability solution to the Cauchy problem

$$\partial_t\mu_t = \partial_x^2(a\mu_t) - \partial_x(b\mu_t), \quad \mu_0 = \nu,$$

exists, then it is unique. If at least one of these integrals converges, then there exists a locally bounded coefficient drift b (*continuous if* a *has a continuous derivative and smooth if so is* a) *and an initial distribution given by a locally Lipschitz density* (*smooth if so is* a) *for which the simplex of probability solutions to the Cauchy problem is infinite-dimensional.*

In the multidimensional case, there are the following sufficient conditions for uniqueness. Suppose that, for every ball U in $\mathbf{R}^d$, the operators $A(x,t)$ are Lipschitz uniformly in $t \in (0,T)$ in $x \in U$, and $A^{-1}(x,t)$ is bounded in $x \in U$, and also $b^i \in L^p_{\text{loc}}(\mathbf{R}^d \times (0,T))$ for some $p > d$.

Theorem 6.3. *Let* μ *be a probability solution to problem* (5.3), *and let either of the following conditions be fulfilled*:

(i) $a^{ij}/(1+|x|^2)$, $b^i/(1+|x|) \in L^1(\mu)$;

(ii) *there exists a positive function* $V \in C^2(\mathbf{R}^d)$ *along with a number* C *such that* $V(x) \to +\infty$ *as* $|x| \to +\infty$ *and*

$$L_{A,b}V(x,t) \leqslant C + CV(x).$$

Then μ *is a unique probability solution to problem* (5.3).

It is now appropriate to comment on (86) in section 11 of Kolmogorov's paper, which is usually called the Kolmogorov–Chapman equation and has the form (in our notation)

$$\mu_{s,x,u} = \int \mu_{t,y,u}\,\mu_{s,x,t}(dy)$$

for the solution $\mu_{s,x,t}$ to the Cauchy problem with the coefficients $A(x)$ and $b(x)$, independent of t, and Dirac's measure at the point x as the initial distribution at time s with $s \leqslant t \leqslant u$. We consider locally bounded coefficients A and b. Suppose that a probability solution to the Cauchy problem exists and is unique for every initial probability measure. Then $\mu_{s,x,t}$ depends on the difference $t-s$, so it suffices to consider the measures $\mu(x,t) = \mu_{0,x,t}$. For these measures, the Kolmogorov–Chapman equation is written in the form

$$\mu(s+t,x) = \int \mu(s,y)\,\mu(t,x)(dy). \tag{6.1}$$

Under our assumption about uniqueness, this equation is a corollary of the Fokker–Planck–Kolmogorov equation. Indeed, the measure $\mu(s+t,x)\,dt$ for any fixed x is a solution to the Cauchy problem with initial condition $\mu(s,x)$ at $t=0$. The right-hand side of (6.1) at $t=0$ is the same measure $\mu(s,x)$, since $\mu(0,x)=\delta_x$. In addition, the right-hand side multiplied by the measure dt also satisfies the Fokker–Planck–Kolmogorov equation. Indeed, it equals

$$\eta(t,s,x) = \int \mu(t,y)\,\mu(s,x)(dy).$$

Hence, for every function $\varphi \in C_0^\infty(\mathbf{R}^d \times (0,\infty))$, the integral of $\partial_t\varphi + L_{A,b}\varphi$ against the measure $\eta(t,s,x)\,dt$ is zero, since the integral of $\partial_t\varphi + L_{A,b}\varphi$ against the measure $\mu(t,y)$ also is zero.

Under our assumption of uniqueness the solution also possesses the semigroup property in the following sense: denoting by $\mu(t,\nu)$ the value of the solution with initial condition ν at time t, we obtain

$$\mu(t+s,\nu) = \mu(t,\mu(s,\nu)). \tag{6.2}$$

According to [59], in the case of bounded continuous coefficients and without assumptions about uniqueness of solutions to the Cauchy problem, one can select a family of probability solutions $\mu(t,\nu)$ for all probability measures ν in such a way that the semigroup identity (6.2) is fulfilled.

Question 10. What are the broadest conditions under which a selection of a solution with the property (6.2) is possible?

Note that if probability solutions are given only for Dirac initial conditions, and the solution $\mu(t,x)$ is Borel measurable in (t,x), then solutions for all initial distributions ν can be defined by the formula

$$\mu(t,\nu) = \int \mu(t,x)\,\nu(dx).$$

This is verified directly by definition, taking into account the local boundedness of coefficients.

7. Distances between solutions and nonlinear equations.

In [8], [19], [20], [21], and [22], some estimates are obtained for distances between solutions to stationary and evolution Fokker–Planck–Kolmogorov equations in terms of some distances between the coefficients. These estimates are applied to the study of nonlinear equations, where coefficients can depend on solutions. Let us give the main result of paper [8]. Let $\|\cdot\|_{\mathrm{TV}}$ denote the total variation norm, and let $W_1(\mu,\sigma)$ denote the

Kantorovich distance between probability measures μ and σ, with finite first moments defined by the formula

$$W_1(\mu,\sigma)=\sup\biggl\{\int f\,d(\mu-\sigma)\colon f\in \mathrm{Lip}_1\biggr\},$$

where Lip_1 is the class of all Lipschitz functions, with the Lipschitz constant equal to 1.

Suppose that probability measures $\mu=\varrho_\mu\,dx$ and $\sigma=\varrho_\sigma\,dx$ on $\mathbf{R}^d$ are solutions to the stationary equation with the coefficients A_μ, b_μ and A_σ, b_σ, respectively, and there are numbers $\Lambda>0$ and $\alpha>0$ such that

$$|a^{ij}_\mu(x)-a^{ij}_\mu(y)|\leqslant\Lambda|x-y|,\quad |a^{ij}_\sigma(x)-a^{ij}_\sigma(y)|\leqslant\Lambda|x-y|\qquad\forall\,x,y\in\mathbf{R}^d,$$
$$A_\mu\geqslant\alpha\cdot\mathrm{I},\quad A_\sigma\geqslant\alpha\cdot\mathrm{I},$$

and also $b^i_\mu, b^i_\sigma\in L^p_{\mathrm{loc}}(\mathbf{R}^d)$ with some $p>d$.

We set

$$h^i_\mu=b^i_\mu-\partial_{x_j}a^{ij}_\mu,\quad h^i_\sigma=b^i_\sigma-\partial_{x_j}a^{ij}_\sigma,$$
$$\Phi=\frac{(A_\mu-A_\sigma)\nabla\varrho_\sigma}{\varrho_\sigma}-(h_\mu-h_\sigma).$$

If $A_\mu=A_\sigma$, then $\Phi=b_\sigma-b_\mu$.

THEOREM 7.1. *Under the stated assumptions, let* $b_\mu\in L^1(\mu+\sigma)$, $\Phi\in L^1(\sigma)$, *and* $|x|\in L^1(\sigma)$, *and let there exist a number* $\kappa>d^2\Lambda^2/(4\alpha)$ *such that, for all* $x,y\in\mathbf{R}^d$,

$$\langle b_\mu(x)-b_\mu(y),x-y\rangle\leqslant-\kappa|x-y|^2.$$

Then μ *has finite first moment and*

$$W_1(\mu,\sigma)\leqslant\frac{1}{m}\int_{\mathbf{R}^d}|\Phi|\,d\sigma,\qquad m=\kappa-\frac{d^2\Lambda^2}{4\alpha}.$$

In addition, there exists a number $C>0$, *depending only on* d, α, Λ, *and* κ, *such that*

$$\|\mu-\sigma\|_{\mathrm{TV}}\leqslant C\int_{\mathbf{R}^d}|\Phi|\,d\sigma.$$

COROLLARY 7.1. *If, under the assumptions of the theorem,* $A_\mu=A_\sigma$, *then, for the solutions* μ *and* σ, *the obtained estimates have the following form*:

$$W_1(\mu,\sigma)\leqslant\frac{1}{m}\int_{\mathbf{R}^d}|b_\mu-b_\sigma|\,d\sigma,\quad \|\mu-\sigma\|_{\mathrm{TV}}\leqslant C\int_{\mathbf{R}^d}|b_\mu-b_\sigma|\,d\sigma.$$

Estimates of this kind can be useful for diverse versions of the Kantorovich problem of optimal transportation of measures (see the recent survey [7]).

The nonlinear stationary Fokker–Planck–Kolmogorov equation is also determined by the diffusion matrix A and the drift coefficient b, which can now depend on the solution: A and b are defined on $\mathbf{R}^d\times\Pi$, where Π is a subset of the space of measures on $\mathbf{R}^d$. The equation has the form

$$L^*_{A(\mu),b(\mu)}\mu=0.$$

Thus, the solution μ satisfies the usual equation with the coefficients $a^{ij}(x,\mu)$ and $b^i(x,\mu)$. Substantial differences with linear equations already arise in the case $A = \mathrm{I}$. A typical example of a drift is

$$b(x,\mu) = \int b(x,y)\,\mu(dy).$$

If $b(x,y) = b_0(x-y)$, then a nonlinear Vlasov-type equation arises (see [49] and [15]). Nonlinear parabolic Fokker–Planck–Kolmogorov equations are introduced similarly.

Acknowledgment. S. V. Shaposhnikov, who is a winner of "The Young Mathematics of Russia" contest, is grateful to its jury and sponsors.

REFERENCES

[1] S. Albeverio, V. Bogachev, and M. Röckner, *On uniqueness of invariant measures for finite- and infinite-dimensional diffusions*, Comm. Pure Appl. Math., 52 (1999), pp. 325–362, https://doi.org/10.1002/(SICI)1097-0312(199903)52:3%3C325::AID-CPA2%3E3.0.CO;2-V

[2] L. Ambrosio, *Transport equation and Cauchy problem for non-smooth vector fields*, in Calculus of Variations and Nonlinear Partial Differential Equations, Lecture Notes in Math. 1927, Springer, Berlin, 2008, pp. 1–41, https://doi.org/10.1007/978-3-540-75914-0_1.

[3] D. G. Aronson and P. Besala, *Uniqueness of solutions of the Cauchy problem for parabolic equations*, J. Math. Anal. Appl., 13 (1966), pp. 516–526, https://doi.org/10.1016/0022-247X(66)90046-1.

[4] P. Bauman, *Equivalence of the Green's functions for diffusion operators in* $\mathbf{R}^n$*: A counterexample*, Proc. Amer. Math. Soc., 91 (1984), pp. 64–68, https://doi.org/10.2307/2045270.

[5] V. I. Bogachev, *Ornstein–Uhlenbeck operators and semigroups*, Russian Math. Surveys, 73 (2018), pp. 191–260, https://doi.org/10.1070/RM9812.

[6] V. I. Bogachev, *Stationary Fokker–Planck–Kolmogorov equations*, in Stochastic Partial Differential Equations and Related Fields, Springer Proc. Math. Stat. 229, Springer, Cham, 2018, pp. 3–24, https://doi.org/10.1007/978-3-319-74929-7_1.

[7] V. I. Bogachev, *Kantorovich problem of optimal transportation of measures: New directions of research*, Russian Math. Surveys, 77 (2022), pp. 769–817, https://doi.org/10.4213/rm10074e.

[8] V. I. Bogachev, A. I. Kirillov, and S. V. Shaposhnikov, *Distances between stationary distributions of diffusions and solvability of nonlinear Fokker–Planck–Kolmogorov equations*, Theory Probab. Appl., 62 (2018), pp. 12–34, https://doi.org/10.1137/S0040585X97T988460.

[9] V. I. Bogachev, E. D. Kosov, and A. V. Shaposhnikov, *Regularity of solutions to Kolmogorov equations with perturbed drifts*, Potential Anal., 58 (2023), pp. 681–702, https://doi.org/10.1007/s11118-021-09954-9.

[10] V. I. Bogachev, T. I. Krasovitskii, and S. V. Shaposhnikov, *On non-uniqueness of probability solutions to the two-dimensional stationary Fokker–Planck–Kolmogorov equation*, Dokl. Math., 98 (2018), pp. 475–479, https://doi.org/10.1134/S1064562418060212.

[11] V. I. Bogachev, T. I. Krasovitskii, and S. V. Shaposhnikov, *On uniqueness of probability solutions of the Fokker–Planck–Kolmogorov equation*, Sb. Math., 212 (2021), pp. 745–781, https://doi.org/10.1070/SM9427.

[12] V. I. Bogachev, N. V. Krylov, and M. Röckner, *On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions*, Comm. Partial Differential Equations, 26 (2001), pp. 2037–2080, https://doi.org/10.1081/PDE-100107815.

[13] V. I. Bogachev, N. V. Krylov, and M. Röckner, *Elliptic equations for measures: Regularity and global bounds of densities*, J. Math. Pures Appl. (9), 85 (2006), pp. 743–757, https://doi.org/10.1016/j.matpur.2005.11.006.

[14] V. I. Bogachev, N. V. Krylov, and M. Röckner, *Elliptic and parabolic equations for measures*, Russian Math. Surveys, 64 (2009), pp. 973–1078, https://doi.org/10.1070/RM2009v064n06ABEH004652.

[15] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov, *Fokker–Planck–Kolmogorov Equations*, Math. Surveys Monogr. 207, Amer. Math. Soc., Providence, RI, 2015, https://doi.org/10.1090/surv/207.

[16] V. I. Bogachev, S. N. Popova, and S. V. Shaposhnikov, *On Sobolev regularity of solutions to Fokker–Planck–Kolmogorov equations with drifts in* L^1, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl., 30 (2019), pp. 205–221, https://doi.org/10.4171/RLM/843.

[17] V. I. Bogachev and M. Röckner, *Regularity of invariant measures on finite and infinite dimensional spaces and applications*, J. Funct. Anal., 133 (1995), pp. 168–223, https://doi.org/10.1006/jfan.1995.1123.

[18] V. I. Bogachev and M. Röckner, *A generalization of Khasminskii's theorem on the existence of invariant measures for locally integrable drifts*, Theory Probab. Appl., 45 (2001), pp. 363–378, https://doi.org/10.1137/S0040585X97978348.

[19] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov, *Distances between transition probabilities of diffusions and applications to nonlinear Fokker–Planck–Kolmogorov equations*, J. Funct. Anal., 271 (2016), pp. 1262–1300, https://doi.org/10.1016/j.jfa.2016.05.016.

[20] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov, *The Poisson equation and estimates for distances between stationary distributions of diffusions*, J. Math. Sci. (N.Y.), 232 (2018), pp. 254–282, https://doi.org/10.1007/s10958-018-3872-3.

[21] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov, *Convergence in variation of solutions of nonlinear Fokker–Planck–Kolmogorov equations to stationary measures*, J. Funct. Anal., 276 (2019), pp. 3681–3713, https://doi.org/10.1016/j.jfa.2019.03.014.

[22] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov, *On convergence to stationary distributions for solutions of nonlinear Fokker–Planck–Kolmogorov equations*, J. Math. Sci. (N.Y.), 242 (2019), pp. 69–84, https://doi.org/10.1007/s10958-019-04467-8.

[23] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov, *On the Ambrosio–Figalli–Trevisan superposition principle for probability solutions to Fokker–Planck–Kolmogorov equations*, J. Dynam. Differential Equations, 33 (2021), pp. 715–739, https://doi.org/10.1007/s10884-020-09828-5.

[24] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov, *Applications of Zvonkin's transform to stationary Kolmogorov equations*, Dokl. Math., 106 (2022), pp. 318–321, https://doi.org/10.1134/S1064562422050064.

[25] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov, *Zvonkin's transform and the regularity of solutions to double divergence form elliptic equations*, Comm. Partial Differential Equations, 48 (2023), pp. 119–149, https://doi.org/10.1080/03605302.2022.2139724.

[26] V. I. Bogachev, M. Röckner, and W. Stannat, *Uniqueness of solutions of elliptic equations and uniqueness of invariant measures of diffusions*, Sb. Math., 193 (2002), pp. 945–976, https://doi.org/10.1070/SM2002v193n07ABEH000665.

[27] V. I. Bogachev, M. Röckner, and F.-Y. Wang, *Elliptic equations for invariant measures on finite and infinite dimensional manifolds*, J. Math. Pures Appl. (9), 80 (2001), pp. 177–221, https://doi.org/10.1016/S0021-7824(00)01187-9.

[28] V. I. Bogachev, A. V. Shaposhnikov, and S. V. Shaposhnikov, *Log-Sobolev-type inequalities for solutions to stationary Fokker–Planck–Kolmogorov equations*, Calc. Var. Partial Differential Equations, 58 (2019), 176, https://doi.org/10.1007/s00526-019-1625-x.

[29] V. I. Bogachev and S. V. Shaposhnikov, *Integrability and continuity of solutions to double divergence form equations*, Ann. Mat. Pura Appl. (4), 196 (2017), pp. 1609–1635, https://doi.org/10.1007/s10231-016-0631-2.

[30] V. I. Bogachev and S. V. Shaposhnikov, *Representations of solutions to Fokker–Planck–Kolmogorov equations with coefficients of low regularity*, J. Evol. Equ., 20 (2020), pp. 355–374, https://doi.org/10.1007/s00028-019-00532-6.

[31] V. I. Bogachev and S. V. Shaposhnikov, *Uniqueness of a probability solution to the Kolmogorov equation with a diffusion matrix satisfying Dini's condition*, Dokl. Math., 104 (2021), pp. 322–325, https://doi.org/10.1134/S1064562421060041.

[32] S. Chapman, *On the Brownian displacements and thermal diffusion of grains suspended in a non-uniform fluid*, Proc. R. Soc. Lond. Ser. A, 119 (1928), pp. 34–54, https://doi.org/10.1098/rspa.1928.0082.

[33] H. Dong, L. Escauriaza, and S. Kim, *On* C^1*,* C^2*, and weak type-*$(1,1)$ *estimates for linear elliptic operators: Part* II, Math. Ann., 370 (2018), pp. 447–489, https://doi.org/10.1007/s00208-017-1603-6.

[34] H. Dong and S. Kim, *On* C^1*,* C^2*, and weak type-*$(1,1)$ *estimates for linear elliptic operators*, Comm. Partial Differential Equations, 42 (2017), pp. 417–435, https://doi.org/10.1080/03605302.2017.1278773.

[35] W. Feller, *Zur Theorie der stochastischen Prozesse: Existenz- und Eindeutigkeitssätze*, Math. Ann., 113 (1937), pp. 113–160, https://doi.org/10.1007/BF01571626.

[36] W. Feller, *The parabolic differential equations and the associated semi-groups of transformations*, Ann. of Math. (2), 55 (1952), pp. 468–519, https://doi.org/10.2307/1969644.

[37] W. Feller, *Generalized second order differential operators and their lateral conditions*, Illinois J. Math., 1 (1957), pp. 459–504, https://doi.org/10.1215/ijm/1255380673.

[38] A. Figalli, *Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients*, J. Funct. Anal., 254 (2008), pp. 109–153, https://doi.org/10.1016/j.jfa.2007.09.020.

[39] A. D. Fokker, *Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld*, Ann. Phys., 348 (1914), pp. 810–820, https://doi.org/10.1002/andp.19143480507.

[40] A. Friedman, *On the uniqueness of the Cauchy problem for parabolic equations*, Amer. J. Math., 81 (1959), pp. 503–511, https://doi.org/10.2307/2372754.

[41] E. Hille, *The abstract Cauchy problem and Cauchy's problem for parabolic differential equations*, J. Analyse Math., 3 (1954), pp. 81–196, https://doi.org/10.1007/BF02803587.

[42] A. M. Il'in and R. Z. Has'minskii, *Asymptotic behavior of solutions of parabolic equations and an ergodic property of nonhomogeneous diffusion processes*, in Ten Papers on Functional Analysis and Measure Theory, Amer. Math. Soc. Transl. Ser. 2 49, Amer. Math. Soc., Providence, RI, 1966, pp. 241–268, https://doi.org/10.1090/trans2/049/10.

[43] R. Z. Khas'minskii, *Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations*, Theory Probab. Appl., 5 (1960), pp. 179–196, https://doi.org/10.1137/1105016.

[44] R. Z. Khasminskii, *Stochastic Stability of Differential Equations*, 2nd ed., Stoch. Model. Appl. Probab. 66, Springer, Heidelberg, 2012, https://doi.org/10.1007/978-3-642-23280-0.

[45] A. N. Kolmogorov, *On analytical methods in probability theory*, in Selected Works of A. N. Kolmogorov, Vol. II, Math. Appl. (Soviet Ser.) 26, Kluwer Acad. Publ., Dordrecht, 1992, pp. 62–108.

[46] A. N. Kolmogorov, *On the limit theorems of probability theory*, in Selected Works of A. N. Kolmogorov, Vol. II, Math. Appl. (Soviet Ser.) 26, Kluwer Acad. Publ., Dordrecht, 1992, pp. 147–155.

[47] A. N. Kolmogorov, *On the theory of continuous random processes*, in Selected Works of A. N. Kolmogorov, Vol. II, Math. Appl. (Soviet Ser.) 26, Kluwer Acad. Publ., Dordrecht, 1992, pp. 156–168.

[48] A. N. Kolmogorov, *On the reversibility of the statistical laws of nature*, in Selected Works of A. N. Kolmogorov, Vol. II, Math. Appl. (Soviet Ser.) 26, Kluwer Acad. Publ., Dordrecht, 1992, pp. 209–215.

[49] V. V. Kozlov, *The generalized Vlasov kinetic equation*, Russian Math. Surveys, 63 (2008), pp. 691–726, https://doi.org/10.1070/RM2008v063n04ABEH004549.

[50] T. I. Krasovitskii, *Degenerate elliptic equations and nonuniqueness of solutions to the Kolmogorov equation*, Dokl. Math., 100 (2019), pp. 354–357, https://doi.org/10.1134/S1064562419040112.

[51] N. V. Krylov, *On diffusion processes with drift in L_d*, Probab. Theory Related Fields, 179 (2021), pp. 165–199, https://doi.org/10.1007/s00440-020-01007-3.

[52] N. V. Krylov, *On diffusion processes with drift in L_{d+1}*, Potential Anal., 59 (2023), pp. 1013–1037, https://doi.org/10.1007/s11118-022-09988-7.

[53] H. Lee, W. Stannat, and G. Trutnau, *Analytic Theory of Itô-Stochastic Differential Equations with Non-Smooth Coefficients*, Springer Briefs Probab. Math. Stat., Springer, Singapore, 2022, https://doi.org/10.1007/978-981-19-3831-3.

[54] H. Lee and G. Trutnau, *Existence and regularity of infinitesimally invariant measures, transition functions and time-homogeneous Itô-SDEs*, J. Evol. Equ., 21 (2021), pp. 601–623, https://doi.org/10.1007/s00028-020-00593-y.

[55] H. Lee and G. Trutnau, *Existence and uniqueness of (infinitesimally) invariant measures for second order partial differential operators on Euclidean space*, J. Math. Anal. Appl., 507 (2022), 125778, https://doi.org/10.1016/j.jmaa.2021.125778.

[56] G. Metafune, D. Pallara, and A. Rhandi, *Global properties of invariant measures*, J. Funct. Anal., 223 (2005), pp. 396–424, https://doi.org/10.1016/j.jfa.2005.02.001.

[57] G. Metafune, D. Pallara, and A. Rhandi, *Global properties of transition probabilities of singular diffusions*, Theory Probab. Appl., 54 (2010), pp. 68–96, https://doi.org/10.1137/S0040585X97984012.

[58] M. Planck, *Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie*, Sitzungsber. Preuß. Akad. Wiss., 1917, pp. 324–341.

[59] M. Rehmeier, *Existence of flows for linear Fokker–Planck–Kolmogorov equations and its connection to well-posedness*, J. Evol. Equ., 21 (2021), pp. 17–31, https://doi.org/10.1007/s00028-020-00569-y.

[60] S. V. Shaposhnikov, *On nonuniqueness of solutions to elliptic equations for probability measures*, J. Funct. Anal., 254 (2008), pp. 2690–2705, https://doi.org/10.1016/j.jfa.2008.02.003.

[61] P. Sjögren, *Harmonic spaces associated with adjoints of linear elliptic operators*, Ann. Inst. Fourier (Grenoble), 25 (1975), pp. 509–518, https://doi.org/10.5802/aif.595.

[62] G. N. Smirnova, *Cauchy problems for parabolic equations degenerating at infinity*, in Fifteen Papers on Analysis, Amer. Math. Soc. Transl. Ser. 2 72, Amer. Math. Soc., Providence, RI, 1968, pp. 119–134, https://doi.org/10.1090/trans2/072/09.

[63] M. von Smoluchowski, *Über Brownsche Molekularbewegung unter Einwirkung äußerer Kräfte und den Zusammenhang mit der verallgemeinerten Diffusionsgleichung*, Ann. Phys., 353 (1916), pp. 1103–1112, https://doi.org/10.1002/andp.19163532408.

[64] W. Stannat, *(Nonsymmetric) Dirichlet operators on L^1: Existence, uniqueness and associated Markov processes*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 28 (1999), pp. 99–140.

[65] D. Trevisan, *Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients*, Electron. J. Probab., 21 (2016), 22, https://doi.org/10.1214/16-EJP4453.

[66] S. R. S. Varadhan, *Lectures on Diffusion Problems and Partial Differential Equations*, Tata Inst. Fund. Res. Stud. Lect. Math. Phys. 64, Tata Inst. Fund. Res., Bombay, 1980.

[67] K. Yosida, *Integration of Fokker–Planck's equation in a compact Riemannian space*, Ark. Mat., 1 (1949), pp. 71–75, https://doi.org/10.1007/BF02590470.