Sborník: Mathematics

PAPER

On uniqueness of probability solutions of the Fokker-Planck-Kolmogorov equation

To cite this article: V. I. Bogachev *et al* 2021 *Sb. Math.* **212** 745

View the article online for updates and enhancements.

You may also like

- Symmetries of stochastic differential equations using Girsanov transformations
Francesco C De Vecchi, Paola Morando and Stefania Ugolini
- Spatially inhomogeneous structures in the solution of Fisher-Kolmogorov equation with delay
S V Aleshin, S D Glyzin and S A Kaschenko
- SOME MATHEMATICAL PROBLEMS OF STATISTICAL HYROMECHANICS
M I Vishik, A I Komech and A V Fursikov

This content was downloaded from IP address 18.118.102.225 on 08/05/2024 at 03:12

Sbornik: Mathematics **212**:6 745–781 *Matematicheskiĭ Sbornik* **212**:6 3–42
DOI: https://doi.org/10.1070/SM9427

On uniqueness of probability solutions of the Fokker-Planck-Kolmogorov equation

V. I. Bogachev, T. I. Krasovitskii and S. V. Shaposhnikov

Abstract. The paper gives a solution to the long-standing problem of uniqueness for probability solutions to the Cauchy problem for the Fokker-Planck-Kolmogorov equation with an unbounded drift coefficient and unit diffusion coefficient. It is proved that in the one-dimensional case uniqueness holds and in all other dimensions it fails. The case of nonconstant diffusion coefficients is also investigated.

Bibliography: 70 titles.

Keywords: Fokker-Planck-Kolmogorov equation, Cauchy problem, uniqueness problem.

§ 1. Introduction

The goal of this work is to give a positive solution to the long-standing problem of uniqueness of a probability solution to the Cauchy problem for the one-dimensional Fokker-Planck-Kolmogorov equation

$$\partial_t\mu = \partial_x^2\mu - \partial_x(b\mu), \qquad \mu_0 = \nu, \tag{1.1}$$

where the initial condition ν is an arbitrary Borel probability measure and the drift coefficient b is a locally bounded Borel function not depending on time t (the problem has remained open even for infinitely differentiable drifts b). By a solution we mean a family of probability measures $\mu = \{\mu_t\}_{t\geqslant 0}$ on the real line, Borel measurable in t and satisfying the integral identity

$$\int \varphi\,d\mu_t - \int \varphi\,d\nu = \int_0^t \int (\varphi'' + b\varphi')\,d\mu_s\,ds \tag{1.2}$$

The research of V. I. Bogachev was carried out with the support of the Russian Foundation for Basic Research (grant no. 20-01-00432-a), the Moscow Center for Fundamental and Applied Mathematics, and the Foundation for the Advancement of Theoretical Physics and Mathematics "Basis" (grant no. 18-1-6-83-1). The research of T. I. Krasovitskii was carried out with the support of the Russian Foundation for Basic Research (grant no. 20-01-00432-a), the Moscow Center for Fundamental and Applied Mathematics, and a stipend from the Foundation for the Advancement of Theoretical Physics and Mathematics "Basis". The research of S. V. Shaposhnikov was carried out with the support of the Russian Foundation for Basic Research (grant no. 20-01-00432-a), the Moscow Center for Fundamental and Applied Mathematics, the Simons-IUM fellowship, and a stipend from the contest "Young Russian Mathematics".

AMS 2020 *Mathematics Subject Classification.* Primary 35Q84.

© 2021 Russian Academy of Sciences (DoM) and London Mathematical Society

for almost all $t \in [0, T]$ for every function $\varphi \in C_0^\infty(\mathbb{R})$, where $T > 0$ is fixed. Under our assumptions the measures μ_t for $t > 0$ are given by probability densities $\varrho(\,\cdot\,, t)$, so equation (1.1) can be written as an equation with respect to the density $\varrho(x, t)$. However, in the setting under consideration it is important that all (or almost all) measures μ_t are probability measures. The significance of this condition will be clear below.

Our main result is this.

Theorem 1.1. *If a probability solution to the Cauchy problem* (1.1) *exists, then it is unique.*

Uniqueness is understood as the equality $\mu_t = \nu_t$ almost everywhere for any two solutions $\{\mu_t\}$ and $\{\nu_t\}$.

We emphasize that there are no global restrictions on the drift coefficient and no assumptions about the behaviour of solutions at infinity or any semigroup properties of solutions. In the two-dimensional case ($x \in \mathbb{R}^2$) the assertion of Theorem 1.1 is false, as Example 4.7 shows (in the three-dimensional case an example of non-uniqueness was given in [11], Ch. 9). Throughout, we cite that book many times, so it is worth mentioning that its Russian version is available on the website of the Russian Foundation for Basic research (https://www.rfbr.ru/rffi/ru/books/o_1896849#1) and the somewhat more complete English version that we cite can be found on the internet.

In addition, we study the Cauchy problem

$$\partial_t \mu_t = \partial_x^2(a\mu_t) - \partial_x(b\mu_t), \qquad \mu_0 = \nu \tag{1.3}$$

with nonconstant diffusion coefficient a and obtain the following result.

Theorem 1.2. *Let a be a positive locally Lipschitz function and let b be a locally bounded Borel function. Suppose that*

$$\int_{-\infty}^{0} \frac{1}{\sqrt{a(x)}}\,dx = \int_{0}^{+\infty} \frac{1}{\sqrt{a(x)}}\,dx = +\infty. \tag{1.4}$$

Then, if a probability solution to the Cauchy problem (1.3) *exists, it is unique. If at least one of these integrals converges, then there exists a locally bounded drift coefficient b* (*which is continuous if a has a continuous derivative and is smooth if a is*) *and an initial distribution given by a locally Lipschitz density* (*smooth if a is*) *for which the simplex of probability solutions to the Cauchy problem is infinite dimensional.*

A similar assertion is true for probability solutions to the one-dimensional stationary Fokker-Planck-Kolmogorov equation (see [11], Proposition 4.1.2 and Example 4.1.1). In the case of the one-dimensional stationary equation

$$\partial_x^2(a\mu) - \partial_x(b\mu) = 0$$

it is easy to write out an explicit formula for a general solution (see [11], § 1.4), which simplifies the investigation considerably. In the parabolic case under consideration no such formula exists and the uniqueness question is difficult. In particular, to construct an example of non-uniqueness in Theorem 1.2 it becomes necessary to employ nontrivial results on the solvability of the initial-boundary value problem for a degenerate parabolic equation.

Note that even for an infinitely differentiable drift coefficient b it can happen that in addition to a unique probability solution with smooth initial distribution there exists another family of nonnegative bounded measures that is a solution with the same initial condition (see Example 4.5).

The Fokker-Planck-Kolmogorov equations for transition probabilities and stationary distributions of diffusion processes in a mathematically rigorous form were first obtained in the fundamental papers [40] and [41] by Kolmogorov. In particular, the problem of investigating the uniqueness of a solution to the Cauchy problem for such equations was formulated in them. In § 15, "A statement of the uniqueness and existence problem for solutions of the second differential equation", in the first paper, one-dimensional equations are considered and the question about the existence of a unique probability solution is posed; multi-dimensional equations are considered in the second paper. Note that the term 'the Fokker-Planck-Kolmogorov equation' is now used for the 'second differential equation' in Kolmogorov's terminology. The book [11] contains references to prior papers in the physics literature, including works by Fokker and Planck. In Kolmogorov's paper [41] the uniqueness of the solution is established for the equation on a compact Riemannian manifold (in the case of coefficients with two continuous derivatives and an initial distribution with continuous density).

The problem of uniqueness of solutions to the Fokker-Planck-Kolmogorov equation in the one-dimensional case was considered in such classics as Feller [28], Yosida [70] and Hille [37], however, in a somewhat different setting connected with semigroups. For example, in [37], the uniqueness in the class of probability solutions was not discussed, but the existence and uniqueness of solutions with certain properties and initial conditions in the domain of definition of the corresponding elliptic operator (see a more precise comment in Remark 4.6 below and also an example showing that such a problem and the problem studied here are not equivalent).

Since Fokker-Planck-Kolmogorov equations in the case of sufficiently regular coefficients are classical parabolic equations, known results on uniqueness of solutions in the classical theory of parabolic equations can, of course, be applied to them, and so we give a brief overview now.

According to Tychonoff's well-known example [65], even for the heat equation $\partial_t u = \partial_x^2 u$ the Cauchy problem can have several solutions. However, as shown by Widder [68], in the class of nonnegative functions the Cauchy problem for the heat equation has a unique solution. In the case of a parabolic equation of general form uniqueness depends not only on the class of functions in which the equation is solved, but also on the coefficients of the equation. Aronson and Besala [2] and [3], Friedman [32] and [33], and Smirnova [61] and [62] obtained various results on uniqueness, namely, uniqueness was investigated in the following classes of functions u: 1) u has a limit as $|x| \to \infty$; 2) for a suitable weight ω, the function $u\omega$ belongs to the space $L^\infty(\mathbb{R}^d \times [0,T])$; 3) for a suitable weight ω, the function $u\omega$ belongs to the space $L^p(\mathbb{R}^d \times [0,T])$. The uniqueness of a solution having a prescribed limit at infinity is established with the aid of the maximum principle and requires that the coefficient c in the equation

$\partial_t u = a^{ij}\partial_{x_i}\partial_{x_j}u + b^i\partial_{x_i}u + cu$ be bounded above, while the uniqueness of solutions integrable with some weight or growing no faster than some function is established under the assumption that the coefficients have at most linear growth. A considerable number of papers are devoted to the study of the uniqueness problem in the class of nonnegative solutions of parabolic equations on $\mathbb{R}^d$ and on smooth Riemannian manifolds (see [38], [48], [50], [52], [53] and [58]). Typical results in this direction establish uniqueness under restrictions on the growth of the coefficients and under the assumption that the parabolic Harnack inequality holds for solutions. In the case of parabolic equations with nonsmooth or degenerate coefficients we mention the method of renormalized solutions (see, for example, [44] by Le Bris and P.-L. Lions), where uniqueness is established in the class of solutions satisfying certain differential inequalities. However, the class of probability solutions of the Fokker-Planck-Kolmogorov equations differs from all the traditional classes listed above. Moreover, there are examples (see [11], Ch. 9, and Example 4.5 below) where the probability solution is unique, but in the class of integrable, bounded or nonnegative functions the Cauchy problem has at least two different solutions. A separate question, though close in spirit, is the uniqueness of a semigroup generated by the corresponding elliptic operator. The study of uniqueness in the classes of Markov and Feller semigroups or semigroups on the space L^p with respect to a fixed measure is the subject of the well-known papers by Feller [29] and [30], Yosida [70], Hille [37], and Wentzell [66] and [67], and some information about recent results can be found in [11], Ch. 5, [1], [4], [24], [25], [45], [49], [64] and [69]. The uniqueness of a semigroup also requires some restrictions on the growth of the coefficients and additional boundary conditions, which are actually necessary restrictions on the domain of definition of the generator of the semigroup. In this paper we do not assume that a probability solution is the result of applying some semigroup to the initial probability measure.

The uniqueness of probability solutions to Fokker-Planck-Kolmogorov equations has also been investigated in many papers; see [7], [8], [10], [15]–[18], [42], [47], [59], [60], and also [11], Ch. 9, where in particular the following results were obtained. Let b be a locally bounded Borel vector field on $\mathbb{R}^d \times [0,T]$. Then, for a probability solution to the Cauchy problem

$$\partial_t\mu_t = \Delta\mu_t - \operatorname{div}(b\mu_t), \qquad \mu_0 = \nu$$

to be unique it suffices that the inequality

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

hold. It also suffices that for at least one probability solution $\mu = \{\mu_t\}$ one of the following conditions be fulfilled:

(i) $(1+|x|)^{-1}|b| \in L^1(\mu_t\,dt, \mathbb{R}^d \times [0,T])$;

(ii) $|b-\beta_\mu| \in L^1(\mu_t\,dt, \mathbb{R}^d \times [0,T])$, where β_μ is the logarithmic derivative of the measure μ, that is, $\beta_\mu(x,t) = \nabla_x\varrho(x,t)/\varrho(x,t)$, where μ is given by the density $\varrho(x,t)$.

In addition, in all dimensions $d \geqslant 3$ examples have been constructed (see [11], Example 9.2.1) in which the Cauchy problem has infinitely many linearly independent probability solutions. The question about uniqueness in dimensions $d=1$ and $d=2$ has remained open. In this paper we give a complete solution of the uniqueness problem for $d \leqslant 2$ in the case when the drift coefficient b depends only on x. In addition, we have obtained auxiliary results for the proof of our main theorems, Theorems 1.1 and 1.2, that are of independent interest:

1) it is shown that the special semigroup $\{T_t^\mu\}_{t\geqslant 0}$ constructed in [11], Ch. 5, for every probability solution μ of the stationary equation defines the minimal nonnegative solution to the Cauchy problem for the first Kolmogorov equation with respect to functions (that is, the usual parabolic equation), and the corresponding semigroup $\{K_t^*\}_{t\geqslant 0}$ on the space of measures defines the minimal nonnegative solution to the Fokker-Planck-Kolmogorov equation;

2) it is shown that if for some probability measure ν on $\mathbb{R}^d$ the family of measures $K_t^*\nu$ generated by the indicated semigroup is a probability solution to the Fokker-Planck-Kolmogorov equation, then for every initial condition the probability solution to this equation is unique.

Since in what follows we discuss not only the one-dimensional, but also the multi-dimensional case, we give the definition of a solution in the general case and mention the results about the regularity of solutions that will be used in this paper. Let $L^p(U)$ denote the usual space L^p of functions on a domain U in $\mathbb{R}^d$ with Lebesgue measure (which is not indicated in this notation), and the symbol $L^p(\mu)$ will denote the space L^p with respect to the measure μ. By $W^{p,1}(U)$ we denote the Sobolev space of functions f in $L^p(U)$ whose generalized partial derivatives $\partial_{x_i} f$ also belong to $L^p(U)$. The norm in $W^{p,1}(U)$ is defined by

$$\|f\|_{W^{p,1}(U)} = \|f\|_{L^p(U)} + \sum_{i\leqslant d} \|\partial_{x_i} f\|_{L^p(U)}.$$

For brevity the second term will be denoted by $\|\partial_x f\|_{L^p(U)}$. Similarly, we introduce the Sobolev space $W^{p,2}(U)$ of functions in $L^p(U)$ with generalized derivatives of the first and second order in $L^p(U)$, with its natural norm $\|f\|_{W^{p,2}(U)}$.

Let a^{ij} and b^i be Borel functions on $\mathbb{R}^d$. We recall that here and throughout the coefficients of the equation depend only on x and are independent of t. Suppose that the matrix $A(x) = (a^{ij}(x))$ is symmetric and nonnegative definite.

Let $T > 0$. A family of Borel probability measures $\{\mu_t\}_{t\in[0,T]}$ on $\mathbb{R}^d$ (that is, $\mu_t \geqslant 0$ and $\mu_t(\mathbb{R}^d) = 1$) is called a probability solution to the Cauchy problem

$$\partial_t \mu_t = \partial_{x_i}\partial_{x_j}(a^{ij}\mu_t) - \partial_{x_i}(b^i\mu), \qquad \mu_0 = \nu, \tag{1.5}$$

where ν is a given probability Borel measure on $\mathbb{R}^d$ (in writing out the equation we omit summation over repeated indices) if for every Borel set E the function $t \mapsto \mu_t(E)$ is Borel measurable (which is equivalent to the integrals with respect to μ_t of smooth functions with compact support being Borel measurable in t), the functions a^{ij} and b^i are integrable on compact sets with respect to the measure $\mu_t\,dt$ on $\mathbb{R}^d \times (0,T)$ and for every function $\varphi \in C_0^\infty(\mathbb{R}^d)$, for almost all $t \in [0,T]$ the equality

$$\int \varphi\,d\mu_t - \int \varphi\,d\nu = \int_0^t \int (a^{ij}\partial_{x_i}\partial_{x_j}\varphi + b^i\partial_{x_i}\varphi)\,d\mu_s\,ds$$

holds. A solution can be also understood as the measure $\mu := \mu_t\,dt$ defined by

$$\mu = \int_0^T \mu_t\,dt.$$

If the measures μ_t are given by their densities $\varrho(\,\cdot\,,t)$ (which is the case under our assumptions), then this measure μ is given by its density $\varrho(x,t)$ in two variables. We write $\mu=\{\mu_t\}$ or $\mu=\mu_t\,dt$ in what follows.

It should be noted that the Borel measurability of the functions $t\mapsto\mu_t(E)$ can be weakened to Lebesgue measurability, since we can take equivalent versions; moreover, it is possible to pick a Borel version for which the above equality (or (1.2) in the one-dimensional case) holds for all t, provided that we allow μ_t to be a probability measure only for almost all t.

In addition to probability solutions, we can introduce solutions of bounded variation $\mu=\mu_t\,dt$ in precisely the same way, including signed solutions, for which $|\mu_t|\,dt$ is assumed to be bounded (such solutions are also called integrable). A solution is called nonnegative if the measures μ_t are nonnegative. However, such a solution need not be a probability solution even if the measure $\mu=\mu_t\,dt$ is probabilistic.

It is known (see [11], Ch. 6) that for a nonnegative solution μ the locally bounded measure $(\det A)^{1/(d+1)}\cdot\mu$ is always absolutely continuous, and if the matrix $A(x)$ is nonsingular for all x, then the measure μ itself is absolutely continuous. If for every ball U there exists a number $C(U)>0$ such that $A(x)\geqslant C(U)I$ for all $x\in U$ and if $a^{ij}\in W^{p,1}(U)$ and $b^i\in L^p(U)$ for some $p>d+2$, then the measure $\mu=\mu_t\,dt$ has a locally Hölder continuous positive density ϱ with respect to Lebesgue measure on $\mathbb{R}^d\times[0,T]$ such that the function $x\mapsto\varrho(x,t)$ belongs to the local Sobolev class $W^{p,1}_{\mathrm{loc}}$. As concerns the properties of solution densities, see also [9], [12], [13], [14] and [19].

If $d=1$, $A=I$ and b is a locally bounded function, then for $t>0$ the function $x\mapsto\varrho(x,t)$ belongs to all local Sobolev classes $W^{p,1}_{\mathrm{loc}}$ with $p\geqslant 1$; in particular, it has a locally absolutely continuous version; moreover, for all $\tau>0$ and $R>0$ the integrals

$$\int_\tau^T\int_{-R}^R|\partial_x\varrho(x,t)|^p\,dx\,dt$$

are finite.

It is worth noting that for the Hölder continuous version of the density ϱ it is only known that almost all densities $\varrho(\,\cdot\,,t)$ are probabilistic, but it is not known whether the continuous version of all these densities must be probabilistic.

Under such broad assumptions about the coefficients, which only have to be locally integrable with respect to the solution μ, but can be locally unbounded, it is not difficult to construct examples of non-uniqueness of probability solutions even in dimension 1 for the stationary equation with unit drift, that is, the equation $\varrho''-(b\varrho)'=0$ with respect to probability densities, as well as for a parabolic equation (see Example 4.8). For this reason we consider locally bounded drift coefficients.

Finally, note that in the case when $A=I$ and the drift coefficient b is infinitely differentiable the solution density ϱ has an infinitely differentiable version on $\mathbb{R}^d\times(0,T)$.

This paper consists of four sections. In § 2 we discuss stationary equations and some properties of special semigroups generated by the corresponding elliptic operators. Section 3 is devoted to the proofs of the main theorems, Theorems 1.1 and 1.2. In § 4 we construct examples of non-uniqueness.

§ 2. Stationary equations and semigroups

In this section we obtain new conditions for the uniqueness of a probability solution to the Cauchy problem in the case when there exists a probability solution of the stationary equation. Of course, a stationary solution does not always exist: for example, if $b=0$, then there are no probability measures satisfying the harmonic equation $\Delta\mu=0$. Thus, the existence of a probability stationary solution can be regarded as an additional restriction on the coefficients of the equation.

Consider the elliptic operator $L_{A,b}$ of the form

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

with Borel coefficients a^{ij} and b^i on $\mathbb{R}^d$ satisfying the following conditions:

(i) the functions a^{ij} are continuous and belong to the Sobolev class $W^{p,1}(U)$ on every ball U in $\mathbb{R}^d$ for some $p>d+2$, and the functions b^i belong to $L^p(U)$;

(ii) the matrix $A(x)=(a^{ij}(x))_{i,j\leqslant d}$ is symmetric and positive definite, and for every ball U there is a number $\lambda(U)>0$ such that $A(x)\geqslant\lambda(U)\cdot\mathrm{I}$ for $x\in U$.

Note that for some assertions used below the condition $p>d$ is sufficient, but for the desired regularity of the semigroup considered below the bound $p>d+2$ is required.

Suppose that there exists a probability measure μ on $\mathbb{R}^d$ satisfying the stationary equation

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

in the sense of the integral identity

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

where it is also assumed that the functions b^i are integrable with respect to μ on all balls (this is automatically fulfilled for locally bounded coefficients, but here we do not assume local boundedness); moreover, here and throughout when integrating over the whole space $\mathbb{R}^d$ we do not indicate the limits of integration. It is known (see [11], Ch. 1) that in this case the measure μ is given by a positive locally Hölder continuous density ϱ with respect to Lebesgue measure such that ϱ belongs to the Sobolev classes $W^{p,1}$ on all balls. Hence the vector field $\nabla\varrho/\varrho$ (the logarithmic gradient of the density) belongs to L^p with respect to Lebesgue measure on balls.

With the aid of the logarithmic gradient we introduce the dual drift

$$\widehat{b} = 2\beta_{A,\mu} - b, \qquad \beta_{A,\mu} = A\nabla\varrho/\varrho + \operatorname{trace}\nabla A = (a^{ij}\partial_{x_j}\varrho/\varrho + \partial_{x_j}a^{ij})_{i=1}^d,$$

which, according to what we have said above, also belongs to $L^p(U)$ on every ball U. The measure μ also satisfies the equation

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

with the dual drift.

According to [11], Theorem 5.2.2, the operators $L_{A,b}$ and $L_{A,\widehat{b}}$ on the domain of definition $C_0^\infty(\mathbb{R}^d)$ extend to the generators $L^\mu_{A,b}$ and $L^\mu_{A,\widehat{b}}$ of strongly continuous sub-Markov semigroups $\{T^\mu_t\}_{t\geqslant0}$ and $\{\widehat{T}^\mu_t\}_{t\geqslant0}$ on the space $L^1(\mu)$ such that the measure μ is subinvariant with respect to them. We recall that the sub-Markov property means that $0 \leqslant T_t^\mu f \leqslant 1$ whenever $0 \leqslant f \leqslant 1$. If, in addition, $T_t^\mu 1 = 1$, then the semigroup is called Markov. The subinvariance of the measure μ means the inequality

$$\int T_t^\mu f\, d\mu \leqslant \int f\, d\mu$$

holds for all nonnegative f. The indicated semigroups are adjoint to each other:

$$\int \xi T_t^\mu \eta\, d\mu = \int \eta \widehat{T}_t^\mu \xi\, d\mu \quad \text{and} \quad \int \xi L_{A,b}\eta\, d\mu = \int \eta L_{A,\widehat{b}}\xi\, d\mu \tag{2.1}$$

for all $\eta, \xi \in C_0^\infty(\mathbb{R}^d)$.

The resolvent $R_\lambda = (\lambda - L_{A,b}^\mu)^{-1}$ of the semigroup $\{T_t^\mu\}_{t\geqslant 0}$ for $\lambda > 0$ is characterized as follows: $(\lambda - L_{A,b}^\mu)^{-1} f$ for $f \in C_0^\infty(\mathbb{R}^d)$ is the limit of solutions u_k to the boundary value problems $\lambda u_k - L_{A,b}u_k = f$, $u_k|_{\partial B_k} = 0$ on the balls B_k of radius $k \in \mathbb{N}$. The resolvent $\{\widehat{T}_t^\mu\}_{t\geqslant 0}$ is described similarly.

However, these semigroups, called canonical, are not always unique strongly continuous semigroups on $L^1(\mu)$ whose generators extend $L_{A,b}$ and $L_{A,\widehat{b}}$. In addition, the measure μ is not always invariant for these semigroups, that is, the identity

$$\int T_t^\mu f\, d\mu = \int f\, d\mu$$

is not always fulfilled for all $f \in C_0^\infty(\mathbb{R}^d)$ (or for all bounded measurable f). Sufficient conditions for the invariance of μ are given in [11], Ch. 5. Note that μ being invariant with respect to one of the two semigroups is equivalent to it being invariant with respect to the other (see [11], Remark 5.2.4), and it is also equivalent to the equality $T_t^\mu 1 = 1$. For example, for $A = I$ a sufficient condition for invariance is the bound $|b(x)| \leqslant C + C|x|$. It is also sufficient that $|b(x)|/(1+|x|)$ be integrable with respect to μ. One more sufficient condition for invariance in terms of μ is this: $|b - \nabla\varrho/\varrho| \in L^1(\mu)$. In the case of a non-constant A the invariance of μ for $\{T_t^\mu\}_{t\geqslant 0}$ is ensured by the inclusions $a^{ij}, |b - \beta_{A,\mu}| \in L^1(\mu)$, as follows from the proofs of Example 5.5.3 and Theorem 5.3.1 in [11]. In particular, if $b = \beta_{A,\mu}$ and $a^{ij} \in L^1(\mu)$, then invariance holds.

Throughout, we assume that the functions a^{ij} are locally Lipschitz (which is stronger than the condition in (i) above).

Lemma 2.1. *For every function $f \in L^1(\mu)$, the family of measures $\nu_t = T_t^\mu f \cdot \mu$ gives a solution to the Cauchy problem for the Fokker-Planck-Kolmogorov equation with the dual drift $\widehat{b}$ and initial condition $f \cdot \mu$.*

In addition, the function $u(x,t) = T_t^\mu f(x)$ is a solution to the Cauchy problem

$$\partial_t u = L_{A,b}u, \qquad u(x,0) = f(x) \tag{2.2}$$

in the following sense: for every $t > 0$ the function $T_t^\mu f$ belongs to the Sobolev class $W^{p,2}(U)$ on every ball U, the function $\|T_t^\mu f\|^p_{W^{p,2}(U)}$ is integrable over every compact interval $[\tau, T_0]$ of $(0,T)$, in $U \times (\tau, T_0)$ the Sobolev derivative $\partial_t u \in L^p(U \times (\tau, T_0))$ exists, equality (2.2) for Sobolev derivatives holds almost everywhere, and the initial condition is also fulfilled in the sense of convergence in $L^1(\mu)$.

If the function f is locally bounded, then $u(x,t)$ is a weak solution in the sense of [2], *which means that for $t>0$ the function $u(\,\cdot\,,t)$ belongs to the Sobolev class $W^{2,1}(U)$ on every ball U, the function $\|u(\,\cdot\,,t)\|_{L^2(U)}$ is bounded on all compact intervals $[0,T_0]\subset[0,T)$, the function $\|\partial_x u(\,\cdot\,,t)\|^2_{L^2(U)}$ is integrable on $[0,T_0]$, and for every function $\psi\in C_0^\infty(\mathbb{R}^d)$ the equality*

$$\begin{aligned}\int u(x,t)\psi(x)\,dx-\int f(x)\psi(x)\,dx\\ =-\int_0^t\int[a^{ij}(x)\partial_{x_j}\psi(x)\partial_{x_i}u(x,s)-b^i\partial_{x_i}u(x,s)\psi(x)\\ +\partial_{x_j}a^{ij}(x)\partial_{x_i}u(x,s)\psi(x)]\,dx\,ds\end{aligned}\tag{2.3}$$

holds.

Proof. Let $\varphi\in C_0^\infty(\mathbb{R}^d)$. Then φ belongs to the domain of definition of the generator of the dual semigroup $\{\widehat{T}_t^\mu\}_{t\geqslant 0}$ and its action on φ coincides with $L_{A,\widehat{b}}\varphi$. In addition,

$$\widehat{T}_t^\mu\varphi(x)-\varphi(x)=\int_0^t L_{A,\widehat{b}}\widehat{T}_s^\mu\varphi(x)\,ds,$$

which yields the equality

$$\begin{aligned}\int\widehat{T}_t^\mu\varphi(x)f(x)\mu(dx)-\int\varphi(x)f(x)\mu(dx)&=\int\int_0^t f(x)L_{A,\widehat{b}}\widehat{T}_s^\mu\varphi(x)\,ds\,\mu(dx)\\ &=\int\int_0^t f(x)\widehat{T}_s^\mu L_{A,\widehat{b}}\varphi(x)\,ds\,\mu(dx).\end{aligned}$$

Therefore, we have

$$\int\varphi(x)T_t^\mu f(x)\mu(dx)-\int\varphi(x)f(x)\mu(dx)=\int_0^t\int T_s^\mu f(x)L_{A,\widehat{b}}\varphi(x)\mu(dx)\,ds,$$

which proves the first assertion.

Let ϱ be a version of the density of the measure μ which is locally Hölder continuous and Sobolev in x. It follows from what we have said about the properties of solution densities that the measure $T_t^\mu f\cdot\mu$ has a Hölder continuous density g on $(0,T)\times\mathbb{R}^d$ such that for almost every t the function $x\mapsto u(x,t)\varrho(x)$ belongs to the Sobolev class $W^{p,1}(U)$ on every ball U and the function $t\mapsto\|\partial_x g(x,t)\|^p_{L^p(U)}$ is integrable on compact intervals in $(0,T)$. Hence the function u possesses the same properties. From the equality obtained above, for $0<\tau<t<T$ we have

$$\int\varphi(x)(u(x,t)-u(x,\tau))\varrho(x)\,dx=\int_\tau^t\int u(x,s)L_{A,\widehat{b}}\varphi(x)\varrho(x)\,dx\,ds.$$

Integrating by parts we transform the right-hand side into the form

$$-\int_\tau^t\int[\partial_{x_i}(a^{ij}(x)\varrho(x)u(x,s))-\widehat{b}^j(x)\varrho(x)u(x,s)]\partial_{x_j}\varphi(x)\,dx\,ds.$$

Now we observe that from (2.1), integrating by parts we obtain the equality

$$\int[-\partial_{x_i}(a^{ij}\xi\varrho)\partial_{x_j}\eta+b^i\partial_{x_i}\eta\xi\varrho]\,dx=\int[-\partial_{x_i}(a^{ij}\eta\varrho)\partial_{x_j}\xi+\widehat{b}^i\partial_{x_i}\xi\eta\varrho]\,dx$$

for all $\eta,\xi\in C_0^\infty(\mathbb{R}^d)$. Passing to the limit, this remains valid for functions η and ξ in the Sobolev class $W^{p,1}(\mathbb{R}^d)$ such that one of them has compact support, in particular, for $\eta(x)=u(x,s)$ with $s>0$ and $\xi=\varphi$. Hence

$$\int\varphi(x)(u(x,t)-u(x,\tau))\varrho(x)\,dx$$
$$=-\int_\tau^t\int[\partial_{x_i}(a^{ij}(x)\varrho(x)\varphi(x))-b^j(x)\varrho(x)\varphi(x)]\partial_{x_j}u(x,s)\,dx\,ds.$$

By a limiting procedure this equality extends to functions $\varphi\in W^{p,1}(\mathbb{R}^d)$ with compact support, in particular, we can take $\varphi=\psi/\varrho$, where $\psi\in C_0^\infty(\mathbb{R}^d)$. This yields identity (2.3) for $[\tau,t]$ in place of $[0,t]$. Thus, in every inner strip u is a weak solution of the direct parabolic equation, which, by known results (see [46]), implies the second assertion of the lemma.

The last assertion follows from Theorem 7.3.11 in [11], where it was assumed that A is globally Lipschitz and both A and A^{-1} are uniformly bounded, but we can see from the proof that for the desired assertion about convergence in L^2 on balls it suffices to have the local Lipschitz property and pointwise invertibility. The theorem just cited also ensures that $u(x,t)$ is bounded on compact sets in $\mathbb{R}^d\times[0,T]$, which implies the integrability of $\|\partial_x u(\,\cdot\,,t)\|^2_{L^2(U)}$ on $[0,T_0]$. Indeed, if we multiply (2.2) by $\varphi(x)^2u(x,t)$, where $\varphi\in C_0^\infty(U)$, integrate over $[\tau,T_0]\times U$ and integrate by parts in the integral of $\varphi^2ua^{ij}\partial_{x_i}\partial_{x_j}u$, then we see that the integral of $\varphi^2a^{ij}\partial_{x_i}u\partial_{x_j}u$ is estimated by the sum of the integrals of $2\varphi\partial_{x_i}\varphi a^{ij}u\partial_{x_j}u$, $\varphi^2\partial_{x_i}a^{ij}u\partial_{x_j}u$, $\varphi^2ub^i\partial_{x_i}u$ and $\varphi^2u\partial_tu$. As u is bounded on $U\times[0,T_0]$, using the inequality $vw\leqslant\varepsilon v^2+\varepsilon^{-1}w^2$ with an appropriate ε enables us to estimate the integral $I(\tau)$ of $\varphi^2|\nabla_xu|^2$ over $U\times[\tau,T_0]$ by $C+I(\tau)^{1/2}$, where C does not depend on τ, which gives the uniform boundedness of $I(\tau)$ up to $\tau-=0$. Therefore, equality (2.3) extends to the whole of the interval $[0,t]$.

Lemma 2.1 is proved.

We observe in passing that so far the boundary behaviour (as $t\to0$) of these solutions has not been sufficiently investigated. In the elliptic case the boundary values of solutions of equations in divergence form on domains were investigated in [34]–[36] under considerably more general assumptions about the diffusion matrix. For general parabolic boundary value problems, see [63].

According to Theorem 5.4.5 in [11], the canonical semigroup $\{T_t^\mu\}_{t\geqslant0}$ is represented by integrable kernels in the form

$$T_t^\mu f(x)=\int f(y)K_t(x,dy),$$

where $K_t(x,dy)$ is a family of subprobability measures on $\mathbb{R}^d$ of the form

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

where the density $p_{A,b}$ is locally Hölder continuous on $(0,T)\times\mathbb{R}^d\times\mathbb{R}^d$. If the measure μ is invariant with respect to the operators T_t^μ, then the measures $K_t(x,dy)$ are probabilistic.

For a bounded measure ν on $\mathbb{R}^d$ we set

$$K_t^*\nu(dy) := \int K_t(x,dy)\nu(dx).$$

Then the family $\{K_t^*\nu\}_{t\geqslant 0}$ is a solution of the Cauchy problem

$$\partial_t\mu = \partial_{x_i}\partial_{x_j}(a^{ij}\mu) - \partial_{x_i}(b^i\mu), \qquad \mu_0 = \nu \tag{2.4}$$

with initial condition ν. If the measure μ is invariant with respect to the operators T_t^μ, then, for every probability measure ν, the measures $K_t^*\nu$ are also probabilistic.

For any bounded Borel function φ we have the equality

$$\int \varphi(x)K_t^*\nu(dx) = \int T_t^\mu\varphi(x)\nu(dx).$$

If the measure ν has a density f with respect to μ, then $K_t^*\nu = \widehat{T}_t^\mu f\cdot\mu$. The semigroup property of T_t^μ implies the semigroup property of K_t^*:

$$K_t^*(K_s^*\nu) = K_{t+s}^*\nu, \qquad t,s\geqslant 0.$$

If a family of measures σ_t on $\mathbb{R}^d$ satisfies equation (2.4), then the measures σ_t have densities $v(\,\cdot\,,t)$ with respect to the measure μ and the function $v(x,t)$ of two arguments possesses a locally Hölder continuous density on $(0,T)\times\mathbb{R}^d$. As in the lemma above, it is straightforward to verify that v satisfies the direct equation

$$\partial_t v = a^{ij}\partial_{x_i}\partial_{x_j}v + \widehat{b}^i\partial_{x_i}v$$

with the dual drift. If the solution $\{\sigma_t\}$ has the initial condition ν, where ν is a bounded measure, then the initial condition for v is the locally bounded measure $\varrho^{-1}\nu$, where this initial condition is understood in the sense of convergence of generalized functions, that is, for every function $\varphi\in C_0^\infty(\mathbb{R}^d)$, as $t\to 0$ the integrals of $\varphi(x)v(x,t)$ converge to the integral of φ against the measure ν/ϱ.

For the next lemma we need the fact that a semigroup with similar properties can be constructed for every ball U in place of the whole space, that is, on $L^1(\mu|_U)$ there is a contractive C_0-semigroup $\{T_t^{U,\mu}\}_{t\geqslant 0}$ of sub-Markov operators for which the measure μ on U is subinvariant and its generator extends the operator $(L_{A,b}, C_0^\infty(U))$. For a bounded function f (or for $f\in L^2(U)$) the function $(x,t)\mapsto T_t^{U,\mu}f(x)$ on $U\times[0,T]$ is defined as the solution to the initial-boundary value problem

$$\partial_t u = L_{A,b}u, \qquad u(x,0) = f, \quad u|_{\partial U\times[0,T]} = 0,$$

which exists, as was shown in [2], p. 634, Theorem 1; moreover, it is unique in the class $L^2([0,T],W_0^{2,1}(U))$ of functions v such that $v(t,\,\cdot\,)$ belongs to the Sobolev class $W_0^{2,1}(U)$ of functions with zero boundary value on ∂U and the function

$\|v(t,\,\cdot\,)\|^2_{W^{2,1}(U)}$ is integrable on $[0,T]$. The semigroup property follows from the uniqueness of a solution, the sub-Markov property follows from the properties of solutions established in [2]. If f is continuous and has a compact support in U, then the solution is continuous on the closure of $U\times[0,T]$. Our conditions on the matrix A are stronger than those assumed in [2], so by virtue of the results in [46] the solution obtained has the property that for each $f\in L^p(U)$ the function $T_t^{U,\mu}f=u(\,\cdot\,,t)$ with $t>0$ belongs to $W_0^{p,1}(U)\cap W^{p,2}(U)$. Then by Lemma 5.2.1 in [11] the integral of $L_{A,b}T_t^{U,\mu}f$ over U is nonpositive if $f\geqslant 0$. Hence for nonnegative functions $f\in C_0^\infty(U)$ we have

$$\int_U T_t^{U,\mu}f\,d\mu-\int_U f\,d\mu=\int_0^t\int_U L_{A,b}T_sf\,d\mu\,ds\leqslant 0,$$

which implies the same estimate for all nonnegative functions $f\in L^1(\mu|_U)$. It follows that the operators $T_t^{U,\mu}$ extend to contractions on all spaces $L^p(\mu|_U)$, $p\geqslant 1$. Note that the role of the measure μ is as follows: the operators $T_t^{U,\mu}$, whose action on bounded functions has no relation to the measure, by virtue of the equation $L^*_{A,b}\mu=0$ turn out to be contractions of the spaces $L^p(\mu|_U)$, but not of the spaces $L^p(U)$ with equivalent norms. The resolvent $w=(L_{A,b}-\lambda)^{-1}f$ for $f\in C_0^\infty(U)$ satisfies the equation $L_{A,b}w-\lambda w=f$ with zero boundary condition, which follows from the equalities

$$(L_{A,b}-\lambda)^{-1}f=\int_0^\infty e^{-\lambda t}T_t^{U,\mu}f\,dt,\qquad L_{A,b}T_t^{U,\mu}f=\frac{d}{dt}T_t^{U,\mu}f.$$

We can also verify that the semigroup $\{T_t^{U,\mu}\}_{t\geqslant 0}$ we have described coincides with the canonical semigroup in Theorem 5.2.2 in [11], constructed for the domain similarly to the case of the whole space. The resolvent $R_\lambda^U=(\lambda-L_{A,b}^{U,\mu})^{-1}$ of the semigroup $\{T_t^{U,\mu}\}_{t\geqslant 0}$ with $\lambda>0$ is characterized as follows: $(\lambda-L_{A,b}^{U,\mu})^{-1}f$ for $f\in C_0^\infty(U)$ is the solution u of the boundary value problem

$$\lambda u-L_{A,b}u=f,\qquad u|_{\partial U}=0.$$

We emphasize that $L_{A,b}^{U,\mu}$ is not the closure of $(L_{A,b},C_0^\infty(U))$. For example, even for the interval $U=(-1,1)$ with Lebesgue measure and $Lu=u''$ the image of $C_0^\infty(U)$ under the operator $L-I$ is not dense in $L^1(U)$, since the integral of $e^x(u''(x)-u(x))$ vanishes for all $u\in C_0^\infty(U)$.

In the case of Hölder continuous coefficients the following fact was establish in the paper [49].

Lemma 2.2. *The canonical semigroup is the limit of the aforementioned semigroups $\{T_t^k\}_{t\geqslant 0}$ corresponding to the operator $L_{A,b}$ with zero boundary conditions on the balls B_k of radius $k\in\mathbb{N}$, which are defined on the spaces $L^1(\mu|_{B_k})$. In particular, if $f\in L^1(\mu)$, $T>0$ and $u_k=T_t^kf$ is the solution to the initial-boundary value problem*

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

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

Proof. It suffices to consider the case $f \geqslant 0$. Suppose in addition that $f \leqslant M$. By the maximum principle $u_{k+1} \geqslant u_k$ and $u_k \leqslant M$ (see [2], p. 634), hence the pointwise limit $u(x,t) = \lim_{k\to\infty} u_k(x,t) \leqslant M$ exists. Therefore, whenever $\operatorname{Re}\lambda > 0$, the pointwise limit

$$\int_0^\infty e^{-\lambda t} u(x,t)\,dt = \lim_{k\to\infty} \int_0^\infty e^{-\lambda t} u_k(x,t)\,dt$$

exists. The construction of the canonical semigroup (see [11], Ch. 5, and the explanations above) yields that the resolvent R_λ of this semigroup satisfies the equality

$$R_\lambda f(x) = \lim_{k\to\infty} R_\lambda^{B_k} f(x) = \lim_{k\to\infty} \int_0^\infty e^{-\lambda t} T_t^k f(x)\,dt,$$

where $R_\lambda^{B_k}$ is the resolvent of the semigroup $\{T_t^k\}_{t\geqslant 0}$, that is, the function $R_\lambda^{B_k} f$ is the solution of the boundary value problem $L_{A,b}u - \lambda u = f$ on B_k. Moreover, for $R_\lambda f$ we have the equality

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} T_t^\mu f(x)\,dt.$$

Therefore,

$$\int_0^\infty e^{-\lambda t} u(x,t)\,dt = \int_0^\infty e^{-\lambda t} T_t^\mu f(x)\,dt.$$

From the estimates $0 \leqslant u \leqslant M$ and $0 \leqslant T_t^\mu f(x) \leqslant M$ we obtain the desired equality $u(x,t) = T_t^\mu f(x)$. Now we drop the assumption that f is bounded. As above, there exists a limit $u(x,t)$ of the increasing sequence of solutions $u_k(x,t)$ of initial-boundary value problems on the balls B_k. For every fixed $M \in \mathbb{N}$, according to what we have proved we have the equality

$$T_t^\mu \min(f,M) = \lim_{k\to\infty} T_t^k \min(f,M).$$

The right-hand side does not exceed $u = \lim_{k\to\infty} T_t^k f$. The left-hand side increases to $T_t^\mu f$ as $M \to \infty$. Hence $T_t^\mu f \leqslant u$. On the other hand $u \leqslant T_t^\mu f$, since for any fixed M and k we have $T_t^k \min(f,M) \leqslant T_t^\mu \min(f,M)$, and so $T_t^k f \leqslant T_t^\mu f$ for all k. The lemma is proved.

Theorem 2.3. *The canonical semigroup $\{T_t^\mu\}_{t\geqslant 0}$ gives the minimal solution of the Cauchy problem* (2.2) *in the following sense*: *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 aforementioned sense from* [2], *then*

$$T_t^\mu f(x,t) \leqslant v(x,t).$$

A similar assertion is true for the dual drift $\widehat{b}$ and the semigroup $\{\widehat{T}_t^\mu\}_{t\geqslant 0}$.

Proof. Let v be an arbitrary nonnegative solution to the Cauchy problem

$$\partial_t v = L_{A,b} v, \qquad v(x,0) = f(x).$$

Then it follows from the maximum principle (see [2], p. 634) that $0 \leqslant u_k \leqslant v$ on B_k, where $u_k(x,t) = T_t^k f(x)$ for $x \in B_k$ is the function from the previous lemma. Hence $T_t^\mu f \leqslant v$. Of course, what we have proved also applies to the dual drift. The theorem is proved.

Below the inequality $\nu_1 \leqslant \nu_2$ for measures means that $\nu_1(B) \leqslant \nu_2(B)$ for all Borel sets. For measures with densities this reduces to an inequality for densities almost everywhere.

Corollary 2.4. *If $\{\sigma_t\}_{t\geqslant 0}$ is some probability solution of the Cauchy problem* (2.4) *with initial condition ν, then $K_t^*\nu \leqslant \sigma_t$.*

Proof. We show that for every nonnegative smooth function φ with compact support the integral with respect to the measure $K_t^*\nu$ is not greater than the integral with respect to the measure σ_t. Fix $t_1 > 0$. It suffices to verify that for every $\varepsilon > 0$ the first integral is not greater than the second plus ε. There is $\tau_1 \in (0, t_1)$ such that

$$\int T_{t_1}^{\mu}\varphi(y) K_\tau^*\nu(dy) \leqslant \int T_{t_1}^{\mu}\varphi(y)\sigma_\tau(dy) + \varepsilon$$

for all $\tau \in [0, \tau_1]$. For $t > 0$ the measures σ_t are given by continuous densities $v(x,t)$ with respect to μ. The function $v(x,t+\tau)$ is a nonnegative solution to the Cauchy problem for the equation $\partial_t v = L_{A,\widehat{b}} v$ with the continuous initial condition $v(x,\tau)$ for $t = 0$. By Theorem 2.3

$$K_t^*\sigma_\tau = \widehat{T}_t^{\mu} v(\cdot,\tau)\cdot\mu \leqslant v(x,t+\tau)\cdot\mu \quad \text{for } t \geqslant 0.$$

Hence for $t \geqslant 0$ and $\tau \in (0,\tau_1]$ we have

$$\int \varphi(y) K_t^*\sigma_\tau(dy) \leqslant \int \varphi(x) v(x,t+\tau)\mu(dx).$$

In addition,

$$\int \varphi(y) K_{t_1}^*\sigma_\tau(dy) = \int T_{t_1}^{\mu}\varphi(y)\sigma_\tau(dy) \geqslant \int T_{t_1}^{\mu}\varphi(y) K_\tau^*\nu(dy) - \varepsilon.$$

Thus,

$$\int T_{t_1}^{\mu}\varphi(y) K_\tau^*\nu(dy) \leqslant \int \varphi(x) v(x,t_1+\tau)\mu(dx) + \varepsilon,$$

that is,

$$\int T_{t_1+\tau}^{\mu}\varphi(y)\nu(dy) \leqslant \int \varphi(x)\sigma_{t_1+\tau}(dx) + \varepsilon,$$

which, as $\tau \to 0$, gives the estimate

$$\int T_{t_1}^{\mu}\varphi(y)\nu(dy) \leqslant \int \varphi(x)\sigma_{t_1}(dx) + \varepsilon,$$

completing the proof.

Corollary 2.5. *If for some probability measure ν all measures $K_t^*\nu$ (or $\widehat{K}_t^*\nu$) are probabilistic, then the Cauchy problem* (2.4) *has a unique probability solution $K_t^*\sigma$ for every initial probability distribution σ.*

Proof. If the measures $K_t^*\nu$ are probabilistic, then the integral of $T_{t/2}^{\mu}1$ with respect to the measure $K_{t/2}^*\nu$ is equal to the integral of $T_t^{\mu}1$ with respect to the measure ν, that is, the integral of 1 with respect to the measure $K_t^*\nu$, is equal to 1. Since the measure $K_t^*\nu$ possesses a positive density and $T^\mu_{t/2}1 \leqslant 1$, this is only possible in the case when $T^\mu_{t/2}1 = 1$. Hence the measure μ is invariant for the canonical semigroup, and for every probability measure σ the family $K_t^*\sigma$ gives a probability solution with initial condition σ. Its minimality implies that there are no other probability solutions. The corollary is proved.

Theorem 2.6. *Let $b = \beta_{A,\mu}$, $a^{ij} \in L^1(\mu)$. Then the probability solution of equation (2.4) is unique.*

If $A = I$ and $b = \nabla V$, where $V \in W^{p,1}_{\mathrm{loc}}$, $p > d$ and $e^V \in L^1(\mathbb{R}^d)$, then the probability solution of equation (2.4) is unique. In particular, this is true for $d = 1$ if $e^B \in L^1(\mathbb{R})$ and $B' = b$.

Proof. The first assertion follows from what has been said above. Indeed, according to the results presented at the beginning of this section, under our conditions the measure μ is invariant for the semigroup $\{T_t^\mu\}_{t\geqslant 0}$, hence $K_t^*\mu = \mu$ is a probability measure for each t. By Corollary 2.5 the Cauchy problem under consideration has a unique probability solution for every initial distribution ν. To prove the second assertion take the measure $\mu = C \exp V\,dx$, where C is picked so that the measure becomes probabilistic. Then $\beta_{A,\mu} = \nabla V$, that is, $b - \beta_{A,\mu} = 0$.

§ 3. Uniqueness in the one-dimensional case

In this section $d = 1$. Set

$$B(x) = \int_0^x b(s)\,ds.$$

The next assertion is inspired by Lemma 9.3 in Feller's paper [29].

Proposition 3.1. *Let b be a locally bounded Borel function and let w be a nonnegative function, absolutely continuous on compact intervals and satisfying the inequality*

$$w'' - bw' \geqslant w$$

in the sense of distributions; moreover, suppose that the limit

$$\lim_{|x|\to\infty} w(x) = q$$

exists and is finite. Then the following assertions are true:

(i) *if $q = 0$, then $w = 0$;*

(ii) *if $q > 0$, then $e^B \in L^1(\mathbb{R})$.*

Analogous assertions are true in a more general case when

$$aw'' - bw' \geqslant w,$$

where $a > 0$ is a locally Lipschitz function, but here in (ii) the integrability of the function $\exp \displaystyle\int_0^x a(s)^{-1} b(s)\,ds$ *is asserted.*

Proof. Since a nonnegative distribution is locally a finite measure, the equality

$$w'' = bw' + w + m$$

holds, where m is a nonnegative Borel measure which is finite on compact intervals. Hence w' as a distribution is given by an ordinary function v of bounded variation on compact intervals.

We show that the function w cannot have positive local maxima. This is well known for twice differentiable solutions (see [54], Ch. 1), but we need the general case. Let z be a point of local maximum. Suppose first (as in the classical result) that at the point z the derivative of the function w exists and is continuous. Then $w'(z) = 0$, hence since b is locally bounded and we have assumed that w' is continuous at z, we have $w'' \geqslant w(z)/2$ in some neighbourhood of the point z in the sense of distributions. Hence $w'(x) > 0$ for all $x > z$ in this neighbourhood, and so $w(x) > w(z)$, contradicting the fact that $w(z)$ is a local maximum. Now we drop the assumption about the existence and continuity of the derivative at z. For the function v of locally bounded variation representing the generalized function w', as noted above, there exist one-sided limits $L = \lim_{x\to z-} v(x)$ and $R = \lim_{x\to z+} v(x)$. Then $L \leqslant R$, since in case $L > R$ the measure w'' must have an atom at the point z with a negative coefficient, which is impossible by the equality $w'' = bw' + w + m$, where $m \geqslant 0$ and the measure with density $bw' + w$ has no atoms. If $L = R$, then we obtain the existence and continuity of w' at the point z and arrive at the case under consideration. If $L < R$, then either $L < 0$ or $R > 0$. Both cases are impossible at a point of local maximum, since in the first case in every interval $(z - \varepsilon, z)$ there exist points with values larger than $w(z)$, and in the second case such points exist in any interval $(z, z + \varepsilon)$.

Let $q = 0$, that is, $w(x) \to 0$ as $|x| \to \infty$. If w assumes a positive value, then w has a positive local maximum, which is impossible. Therefore, $w = 0$.

Now let $q > 0$. Then $w \leqslant q$, since otherwise there exists a point of positive local maximum. There is $z > 0$ such that $w(x) \geqslant q/2$ whenever $|x| \geqslant z$. Such a point can be taken so that $w(z) < q$, since w cannot be constant. Then there exists a point $z_1 > z$ for which $w(x) > w(z)$ for all $x \geqslant z_1$. On $[z_1, +\infty)$ the function w must be increasing, since the existence of points $y > x > z_1$ with $w(y) < w(x)$ yield a local maximum on the interval $[z, y]$ by the inequalities $w(z) < w(y) < w(x)$. Thus, $w' \geqslant 0$ on $(z_1, +\infty)$. Similarly, there exists a point $z_2 < -z$ such that $w'(x) \leqslant 0$ for all $x \leqslant z_2$. Obviously, we can assume that both points z_1 and z_2 are picked such that the function w' of bounded variation is continuous at them and $w'(z_1) > 0$, $w'(z_2) < 0$.

Since w' is of locally bounded variation, on the ray $(z_1, +\infty)$ we obtain the inequality in the sense of distributions:

$$(w'e^{-B})' \geqslant we^{-B}.$$

Therefore, for almost all $x > z_1$ we have

$$w'(x) \geqslant w'(z_1)e^{-B(z_1)}e^{B(x)} + e^{B(x)} \int_{z_1}^{x} w(y)e^{-B(y)}\, dy,$$

and so

$$w'(z_1)e^{-B(z_1)}\int_{z_1}^{+\infty} e^{B(x)}\,dx \leqslant \int_{z_1}^{+\infty} w'(x)\,dx = q - w(z_1) < \infty.$$

Therefore, the function e^B is integrable on $[z_1, +\infty)$. Similarly, we obtain integrability on $(-\infty, z_2]$. Thus, the function e^B is integrable on $\mathbb{R}$.

The second assertion of the proposition is proved similarly, but for B we take

$$B(x) = \int_0^x \frac{b(s)}{a(s)}\,ds + \ln a(x),$$

and the corresponding inequality takes the form $(aw'e^{-B})' \geqslant uw^{-B}$, which leads to the inequality $w'(x) \geqslant u'(z_1)e^{-B(z_1)}e^{B(x)}a(x)^{-1}$ and shows that the function $e^{B(x)}a(x)^{-1}$ is integrable.

Proposition 3.1 is proved.

We now prove our main theorems.

Proof of Theorem 1.1. Suppose that there exist two probability solutions ϱ_1 and ϱ_2. Then by Theorem 2.6

$$e^B \notin L^1(\mathbb{R}).$$

We shall use continuous versions of densities (which exist as we have noted). Set

$$F(x,t) = \int_{-\infty}^x r(y,t)\,dy, \qquad r(y,t) = \varrho_1(y,t) - \varrho_2(y,t), \qquad F(x,0) = 0,$$

that is, F is the difference of the distribution functions of these two solutions. Note that $F(x,t) \to 0$ as $t \to 0$ for all points x, apart from an at most countable set (possible atoms of the common initial distribution). In addition, $-1 \leqslant F(x,t) \leqslant 1$. Finally, for almost all t the function $F(x,t)$ tends to zero as $|x| \to +\infty$. Since we are using the continuous version of the function r, the function F is Borel on $\mathbb{R}\times(0,T)$. With respect to the argument x the function $F(x,t)$ is continuously differentiable.

Let ζ be a smooth probability density with compact support. Set

$$q(t) = \int [\zeta''(x)F(x,t) - b(x)\zeta(x)r(x,t)]\,dx, \qquad q(0) = 0,$$

where here and below we do not indicate the limits of integration when integrating over the whole real line. The function q is Borel and is bounded on $[0,T]$. Hence the function

$$C(t) = \int_0^t q(s)\,ds$$

is Lipschitz on $[0,T]$ and $C(0) = 0$. We show that the bounded Borel function

$$H(x,t) = F(x,t) - \int \zeta(y)F(y,t)\,dy + C(t),$$

on $\mathbb{R}\times[0,T]$, which is continuously differentiable in the variable x, satisfies the equation

$$\partial_t H = \partial_x^2 H - b\partial_x H$$

in the sense of the equality

$$\int \psi(x)H(x,t)\,dx - \int \psi(x)H(x,s)\,dx$$
$$= \int_s^t \int [\psi''(x)H(x,\tau) - \psi(x)b(x)\partial_x H(x,\tau)]\,dx\,d\tau \tag{3.1}$$

for every function $\psi \in C_0^\infty(\mathbb{R})$ and all $t, s \in (0,T)$ with $s \leqslant t$. Note that

$$\partial_x H(x,t) = r(x,t)$$

by the continuity of the version of r under consideration.

By the definition of a solution, for every function $\varphi \in C_0^\infty(\mathbb{R})$ and all $s, t \in (0,T)$ we have the equality

$$\int \varphi(x)r(x,t)\,dx - \int \varphi(x)r(x,s)\,dx = \int_s^t \int [\varphi'' + b\varphi']r\,dx\,d\tau.$$

Substituting $\partial_x F = r$ and integrating by parts, we arrive at

$$\int \varphi'(x)F(x,t)\,dx - \int \varphi'(x)F(x,s)\,dx = \int_s^t \int [\varphi''' F - b\varphi'\partial_x F]\,dx\,d\tau.$$

Let $\psi \in C_0^\infty(\mathbb{R})$. Then there exists a function $\varphi \in C_0^\infty(\mathbb{R})$ such that

$$\varphi'(x) = \psi(x) - \zeta(x)\int \psi(y)\,dy.$$

Substituting into the above equality we obtain

$$\int \psi(x)\Big(F(x,t) - \int \zeta(y)F(y,t)\,dy\Big)\,dx - \int \psi(x)\Big(F(x,s) - \int \zeta(y)F(y,s)\,dy\Big)\,dx$$
$$= \int_s^t \int \Big[\psi''(x)F(x,t) - b(x)\psi(x)\partial_x F(x,t)$$
$$- \psi(x)\Big(\int (\zeta''(y)F(y,\tau) - b(y)\zeta(y)\partial_y F(y,\tau))\,dy\Big)\Big]\,dx\,d\tau$$
$$= \int_s^t \int \Big[\psi''(x)F(x,t) - b(x)\psi(x)\partial_x F(x,t) - \psi(x)q(\tau)\Big]\,dx\,d\tau.$$

Now, to prove (3.1) it suffices to observe that in the integrals containing ψ'' and $\partial_x F$ we can replace $F(x,t)$ by $H(x,t)$, since the difference $H(x,t) - F(x,t)$ does not depend on x.

It follows from equality (3.1) that the function

$$t \mapsto \int \psi(x)H(x,t)\,dx$$

is Lipschitz on $(0,T)$. In addition, on the right-hand side of (3.1) we can integrate the term with ψ'' under the integral twice by parts, which gives us the integral of $\psi(x)\partial_x r(x,t)$. This yields the equality

$$H(x,t) - H(x,s) = \int_s^t [\partial_x r(x,\tau) - b(x)r(x,\tau)]\,d\tau \tag{3.2}$$

for almost all x (for any fixed s and t) and also in the sense of distributions. The left-hand side is continuously differentiable in x.

Consider the bounded Borel function

$$W(x,t) = \frac{H(x,t)^2}{2}$$

on $\mathbb{R} \times [0,T]$, which is also continuously differentiable in the variable x. For $t > 0$ this function satisfies

$$\partial_t W = \partial_x^2 W - b\partial_x W - |\partial_x H|^2,$$

and therefore we have the inequality

$$\partial_x^2 W - b\partial_x W \geqslant \partial_t W.$$

For $t > 0$ the equation and the inequality hold almost everywhere in x and also in the sense of distributions, because the function $x \mapsto \partial_x W(x,t)$ is locally absolutely continuous as the function $x \mapsto r(x,t)$ is (the solution densities are locally Sobolev in x for $t > 0$). In addition,

$$W(x,0) = 0, \qquad W \geqslant 0, \qquad \eta(t) := \lim_{x\to+\infty} W(x,t) = \lim_{x\to-\infty} W(x,t) \geqslant 0.$$

The first equality is fulfilled for all x excepting possibly a countable set, since $F(x,t) \to 0$ as $t \to 0$ for all x in the complement of an at most countable set. The two limits of $W(x,t)$ with respect to x are equal for almost all t because $\lim_{|x|\to\infty} F(x,t) = 0$ for almost all t. Now consider a new function:

$$w(x) = \int_0^T W(x,t)e^{-t}\,dt.$$

This function is nonnegative, bounded, and continuous, and

$$\lim_{|x|\to\infty} w(x) = \int_0^T \eta(t)e^{-t}\,dt.$$

For the sequel we note that

$$w'(x) = \int_0^T H(x,t)r(x,t)e^{-t}\,dt$$

almost everywhere, $w'(x)$ is integrable because H is uniformly bounded, and this expression also defines the generalized derivative, so that the function w is absolutely continuous on compact intervals. To justify this we observe that the function $H(x,t)r(x,t)e^{-t}$ is integrable in both variables on the strip $(-\infty,+\infty) \times [0,T]$, since, for each fixed t, the function $r(x,t)$ is the difference of subprobability densities, hence the integral of $|r(x,t)|$ with respect to the variable x over the real line is no greater than 2. By Fubini's theorem, the function given by the right-hand side is integrable. By integrating by parts against a test function it is readily verified that this expression serves as the generalized derivative of w.

Now we show that w satisfies the inequality

$$w'' - bw' \geqslant w$$

in the sense of distributions. Once this is done, since $e^B \notin L^1(\mathbb{R})$, from Proposition 3.1 we obtain that $w = 0$. Then $W = 0$, hence $H = 0$, that is, $F(x,t)$ does not depend on x, which means that $r(x,t) = 0$.

To justify the desired inequality we observe that for $\tau > 0$ the function

$$w(\tau, x) := \int_\tau^T W(x,t)e^{-t}\,dt$$

satisfies the inequality

$$\begin{aligned}\int_\tau^T \partial_t W(x,t)e^{-t}\,dt &= w(\tau,x) + W(x,T)e^{-T} - W(x,\tau)e^{-\tau}\\ &\geqslant w(\tau,x) - W(x,\tau)e^{-\tau}.\end{aligned}$$

Therefore, the inequality

$$w(\tau,x)'' - b(x)w(\tau,x)' \geqslant w(\tau,x) - W(x,\tau)e^{-\tau}$$

holds in the sense of distributions (with derivatives with respect to x), that is, for every nonnegative smooth function φ with compact support,

$$\begin{aligned}&\int [\varphi''(x)w(\tau,x) - b(x)\,\partial_x w(\tau,x)\varphi(x)]\,dx\\ &\qquad\geqslant \int \varphi(x)w(\tau,x)\,dx - \int \varphi(x)W(x,\tau)e^{-\tau}\,dx,\end{aligned}$$

where the derivative $\partial_x w(\tau,x)$ exists almost everywhere and defines the generalized derivative similarly to the case of the function w. As $\tau \to 0$, the last integral tends to zero, since the function $W(x,t)$ is uniformly bounded and $W(x,\tau) \to 0$ as $\tau \to 0$ for almost all x. The first integral on the right tends to the integral of φw as $\tau \to 0$. The integral of $\varphi''(x)w(\tau,x)$ tends to the integral of $\varphi''(x)w(x)$. The integral of $b(x)\partial_x w(\tau,x)\varphi(x)$ tends to the integral of $b(x)w'(x)\varphi(x)$, because

$$w'(\tau,x) = \int_\tau^T H(x,t)r(x,t)e^{-t}\,dt,$$

and the function $H(x,t)e^{-t}$ is uniformly bounded. Thus, letting $\tau \to 0$ we obtain the desired inequality.

Theorem 1.1 is proved.

Proof of Theorem 1.2. Suppose that condition (1.4) is fulfilled. The function

$$\psi(x) = \int_0^x \frac{1}{\sqrt{a(s)}}\,ds$$

is a diffeomorphism of the real line with locally Lipschitz derivative (this follows from the local Lipschitz property, since a is positive). Let $\varphi = \psi^{-1}$ be the inverse function. The family of measures $\mu_t = \varrho(x,t)\,dx$ is a solution of the Cauchy problem (1.3) precisely when the family of measures $\sigma_t = \sigma(y,t)\,dy$, where $\sigma(y,t) = \varphi'(y)\varrho(\varphi(y),t)$, satisfies the equation

$$\partial_t\sigma_t = \partial_y^2\sigma_t - \partial_y(\beta\sigma_t)$$

with the drift coefficient

$$\beta(y) = b(\varphi(y))\psi'(\varphi(y)) + a(\varphi(y))\psi''(\varphi(y))$$

and initial condition $\widetilde{\nu} = \nu\circ\psi^{-1}$. By Theorem 1.1 the probability solution of the Cauchy problem for σ_t is unique. Therefore, the probability solution of the original Cauchy problem is also unique. Examples of non-uniqueness in the case where condition (1.4) is not fulfilled are constructed in the next section.

§ 4. Examples of non-uniqueness

We start with examples completing the proof of Theorem 1.2.

Let $T = 1$.

Example 4.1. Let a be a locally Lipschitz positive function on the real line such that

$$\int_{-\infty}^{0}\frac{1}{\sqrt{a(s)}}\,ds = y_1 < \infty \quad\text{and}\quad \int_{0}^{+\infty}\frac{1}{\sqrt{a(s)}}\,ds = y_2 < \infty.$$

Set $b = a'/2$. Then there exists a locally Lipschitz probability density ϱ_0 (which is smooth if a is) such that the Cauchy problem

$$\partial_t\varrho = \partial_x^2(a\varrho) - \partial_x(b\varrho), \qquad \varrho(x,0) = \varrho_0(x) \tag{4.1}$$

has infinitely many linearly independent solutions ϱ such that the function ϱ is continuous on $\mathbb{R}\times[0,1]$, continuously differentiable in t, locally Lipschitz in x (if the coefficient a is twice continuously differentiable, then ϱ is also twice continuously differentiable in x), and

$$\varrho(x,t) > 0, \qquad \int\varrho(x,t)\,dx = 1.$$

The function

$$\psi(x) = \int_0^x\frac{1}{\sqrt{a(s)}}\,ds$$

defines a diffeomorphism from the real line onto the interval $J = (-y_1, y_2)$ with locally Lipschitz derivative. We denote the inverse function ψ^{-1} by φ. Changing the variables we arrive at the Cauchy problem

$$\partial_t\sigma = \partial_y^2\sigma, \qquad \sigma(y,0) = \sigma_0(y),$$

where

$$\sigma(y,t) = \varphi'(y)\varrho(\varphi(y),t) \quad\text{and}\quad \sigma_0(y) = \varphi'(y)\varrho_0(\varphi(y)).$$

Below we pick a smooth probability density σ_0 on the interval J determining the required initial density ϱ_0. We observe that if the new Cauchy problem for the function σ has infinitely many linearly independent probability solutions, then the original Cauchy problem has also infinitely many linearly independent probability solutions.

On $[-y_1, y_2] \times [0,1]$ consider the initial-boundary value problem

$$\partial_t \sigma = \partial_y^2 \sigma, \qquad \sigma(y,0) = \sigma_0(y), \qquad \partial_y \sigma(-y_1, t) = \partial_y \sigma(y_2, t) = \theta(t), \tag{4.2}$$

where θ is a continuously differentiable function, $\theta(0) = \theta'(0) = 0$ and σ_0 is a smooth nonnegative function with compact support in (y_1, y_2). It is known (see [43], Theorem 5.3) that there exists a solution σ that is a function on $[-y_1, y_2] \times [0,1]$ which is continuously differentiable in t and twice continuously differentiable in y. We verify that for all $t \in [0,1]$

$$\int_{-y_1}^{y_2} \sigma(y,t)\, dy = \int_{-y_1}^{y_2} \sigma_0(y)\, dy.$$

In fact,

$$\begin{aligned} \frac{d}{dt} \int_{-y_1}^{y_2} \sigma(y,t)\, dy &= \int_{-y_1}^{y_2} \partial_t \sigma(y,t)\, dy = \int_{-y_1}^{y_2} \partial_y^2 \sigma(y,t)\, dy \\ &= \partial_y \sigma(y_2, t) - \partial_y \sigma(-y_1, t) = 0. \end{aligned}$$

Adding a constant to σ (and correspondingly to σ_0) we can assume that $\sigma > 0$. Multiplying σ by a suitable constant, we obtain the solution σ which for every t is a probability density on $[-y_1, y_2]$. If the functions $\theta_1, \dots, \theta_N$ are linearly independent, then the corresponding solutions $\sigma_1, \dots, \sigma_N$ are also linearly independent. To prove this it suffices to observe that $\theta_j(t) = \partial_y \sigma_j(y_2, t)$. Thus, taking linearly independent functions θ, we can construct linearly independent solutions σ. Returning to the original coordinates, we obtain linearly independent probability solutions of the Cauchy problem (4.1) under consideration.

In the following example only one integral from condition (1.4) converges.

Example 4.2. Let a be a locally Lipschitz positive function on the real line such that

$$\int_{-\infty}^{0} \frac{1}{\sqrt{a(s)}}\, ds = y_1 < \infty \quad \text{and} \quad \int_{0}^{+\infty} \frac{1}{\sqrt{a(s)}}\, ds = \infty.$$

Then there exists a locally bounded Borel drift coefficient b (which is continuous if a has a continuous derivative) and an initial condition with a locally Lipschitz probability density ϱ_0 (which is smooth if a is) for which the Cauchy problem

$$\partial_t \varrho = \partial_x^2 (a\varrho) - \partial_x (b\varrho), \qquad \varrho(x,0) = \varrho_0(x) \tag{4.3}$$

has infinitely many linearly independent solutions ϱ with the following properties: the function ϱ is continuous on $\mathbb{R} \times [0,1]$, continuously differentiable in t, locally Lipschitz in x (and twice continuously differentiable in x in the case of a twice continuously differentiable coefficient a), and

$$\varrho(x,t) > 0, \qquad \int \varrho(x,t)\, dx = 1.$$

The function

$$\psi(x) = \int_0^x \frac{1}{\sqrt{a(s)}}\, ds$$

is a diffeomorphism of the real line onto the ray $J = (-y_1, +\infty)$ with locally Lipschitz derivative. Let $\varphi = \psi^{-1}$. In the new coordinates the Cauchy problem (4.3) has the form

$$\partial_t \sigma = \partial_y^2 \sigma - \partial_y(\beta\sigma), \qquad \sigma(y, 0) = \sigma_0(y),$$

where

$$\beta = \psi' b + a\psi'' = \frac{b}{\sqrt{a}} - \frac{1}{2}\frac{a'}{\sqrt{a}},$$

$$\sigma_0(y) = \varphi'(y)\varrho_0(\varphi(y)), \qquad \sigma(y,t) = \varphi'(y)\varrho(\varphi(y), t).$$

Below we pick a smooth coefficient β, from which we find the required coefficient b by the formula $b = \beta\sqrt{a} + a'/2$. It is seen from this formula that if the derivative a' is continuous, then the function b is also continuous and if a is locally Lipschitz, then it is locally bounded. In addition, we pick some initial density σ_0 which will give ϱ_0. We make one more change of coordinates

$$\eta(y) = 1 - (y + y_1 + 1)^{-1},$$

which transforms the ray J into the interval $(0, 1)$. In the new coordinates we arrive at the Cauchy problem

$$\partial_t v = \partial_z^2((1-z)^4 v) - \partial_z(hv), \qquad v(z, 0) = v_0(z).$$

As above, $h = \eta'\beta + \eta''$ and $v(z,t) = \xi'(z)\sigma(\xi(z), t)$, where $\xi = \eta^{-1}$. Now take the third order polynomial

$$h(z) = -4(1-z)^3 - 1,$$

which determines the smooth function $\beta = (h - \eta'')/\eta'$. We have

$$\partial_z^2(1-z)^4 - \partial_z h(z) = 0, \qquad h(0) = -5, \qquad h(1) = -1.$$

Suppose also that v_0 is a smooth function with compact support in the interval $(0, 1)$. Consider the initial-boundary value problem

$$\begin{gathered} \partial_t v = \partial_z^2((1-z)^4 v) - \partial_z(hv), \qquad v(z,0) = v_0(z), \\ v(1,t) = v(0,t) + \partial_z v(0,t) = \theta(t) \end{gathered} \tag{4.4}$$

on $[0,1]\times[0,1]$, where θ is a continuously differentiable function and $\theta(0) = \theta'(0) = 0$. Below we show that this problem has a solution v that is continuously differentiable in t and twice continuously differentiable in z on $[0,1]\times[0,1]$. We verify that

$$\int_0^1 v(z,t)\, dz = \int_0^1 v_0(z)\, dz$$

for all $t \in [0, 1]$. In fact,

$$\begin{aligned} \frac{d}{dt}\int_0^1 v(z,t)\, dz &= \int_0^1 \bigl[\partial_z^2((1-z)^4 v) - \partial_z(hv)\bigr]\, dz \\ &= -v(0,t) - \partial_z v(0,t) + v(1,t) = 0. \end{aligned}$$

The function h is picked in such a way that if v is a solution of the equation, then $v+\mathrm{const}$ is also a solution. Adding a constant we can assume that $v>0$. Next, multiplying by a suitable constant, we can assume that the function $x\mapsto v(x,t)$ is a probability density on $[0,1]$ for every t. Finally, linearly independent solutions correspond to the linearly independent functions θ.

The reasoning we have given here is essentially based on the solvability of the initial-boundary value problem (4.4). The main difficulty is that the equation degenerates at $z=1$. Boundary value problems for degenerate elliptic and parabolic equations are considered in many works, among which we mention [26], [27], [31], [51] and [55]. However, the assertion we need about the solvability of the third boundary value problem is only contained in [26], without proof; furthermore, the proof of this assertion is also omitted from [27], which contains the proofs of the main results from [26]. In the closely related paper [55], the solvability of the third boundary value problem for a degenerate elliptic equation is studied, a particular case of which is, of course, a parabolic equation, but the smoothness of the boundary assumed there is such that the result does not apply to parabolic problems. Moreover, in [55] only a generalized solution is constructed. For these reasons we include the necessary assertion here together with a short justification in the case of one space variable, which is the case we are looking at. A general assertion will be considered in a separate paper.

Let a, b, c and f be infinitely differentiable functions on $\mathbb{R}\times[0,T]$, let h, g and k be infinitely differentiable functions on $[0,T]$, and let u_0 be an infinitely differentiable function with compact support in $(0,1)$. Suppose that $a\geqslant 0$ and consider the following problem:

$$\begin{gathered}\partial_t u=a\partial_x^2u+b\partial_x u+cu+f, \qquad u(x,0)=u_0(x), \qquad u(1,t)=h(t),\\ \partial_x u(0,t)+ku(0,t)=g(t).\end{gathered}\tag{4.5}$$

Proposition 4.3. *Suppose that $a(0,t)>0$ and $a(1,t)=0$, and let*

$$b(1,t)-\partial_x a(1,t)>0.$$

Then there exists a unique solution v of the Cauchy problem (4.5) in the class of functions that are continuously differentiable in t and twice continuously differentiable in x on the rectangle $[0,1]\times[0,T]$.

Proof. Let $u=we^{\lambda t+\gamma x}$. Then

$$\begin{gathered}w_t=a\partial_x^2w+(b+2\gamma a)\partial_x w+(\gamma^2a+\gamma b+c-\lambda)w+fe^{-\lambda t-\gamma x},\\ w(x,0)=u_0(x)e^{-\gamma x}, \qquad w(1,t)=h(t)e^{-\gamma-\lambda t},\\ \partial_x w(0,t)+(k+\gamma)w(0,t)=g(t)e^{-\lambda t}.\end{gathered}$$

Picking constants $\gamma<0$ and $\lambda>0$, we can obtain

$$k+\gamma<0 \quad\text{and}\quad \gamma^2a+\gamma b+c-\lambda<0.$$

Moreover, the condition $b(1,t)-\partial_x a(1,t)$ for our new coefficient $b+2\gamma a$ in place of b is preserved, since $a(1,t)=0$. So in what follows, passing from u to w, we assume that the inequalities $c \leqslant -c_0 < 0$ and $k \leqslant -k_0 < 0$ hold for some numbers c_0 and k_0. Moreover, subtracting a function Q from the solution u which is such that $Q(1,t) = h(t)$ and $\partial_x Q(0,t) + k(t)Q(0,t) = g(t)$, we can assume that $h = g = 0$.

Let $n \in \mathbb{N}$. It is known (see [5], [20], [21], [57] and [39], Ch. 2, § 4, Theorem 4.1) that there exists a solution u_n of the Cauchy problem for the equation with $a+n^{-1}$ in place of a. Let $q_n = u_n^2/2$. Then

$$\partial_t q_n = a\partial_x^2 q_n + b\partial_x q_n + 2cq_n + fu_n - a(\partial_x q_n)^2.$$

Since $fu_n \leqslant c_0 q_n + |f|^2 c_0^{-1}$, we have

$$\partial_t q_n \leqslant a\partial_x^2 q_n + b\partial_x q_n + cq_n + |f|^2 c_0^{-1}.$$

Moreover, $\partial_x q_n(0,t) = -2kq_n(0,t)$, $q_n(x,0) = u_0(x)^2/2$ and $q_n(1,t) = 0$. Let

$$M = 2^{-1}\max|u_0^2| + c_0^{-2}\max|f|^2.$$

Then

$$\partial_t(q_n - M) \leqslant a\partial_x^2(q_n - M) + b\partial_x(q_n - M) + c(q_n - M)$$

and $\partial_x(q_n(0,t) - M) = -2k(q_n(0,t) - M) - kM$, $q_n(x,0) - M \leqslant 0$, $q_n(1,t) - M \leqslant 0$. Since $c < 0$ and $k < 0$, it is clear that the function $q_n - M$ cannot attain a positive maximum. Therefore, $q_n \leqslant M$ and $|u_n| \leqslant \sqrt{2M}$, where M does not depend on n. Passing to a subsequence, we can assume that $\{u_n\}$ converges weakly in $L^2([0,1]\times[0,T])$ to some function u.

By assumption there are positive numbers a_0 and x_0 such that $a(x,t) \geqslant a_0$ on the rectangle $[0,2x_0]\times[0,T]$. According to Theorem 10.1 in [43], passing to a subsequence, we can assume that $\{u_n\}$ converges uniformly on $[0,x_0]\times[0,T]$ to a function u which is continuously differentiable in t and twice continuously differentiable in x on the rectangle $[0,x_0]\times[0,T]$.

According to Theorem 4 in [27], there exists a smooth solution v of the Cauchy problem

$$\partial_t v = a\partial_x^2 v + b\partial_x v + cv + f,$$
$$v(x,0) = u_0(x), \qquad v(1,t) = 0, \qquad v(0,t) = u(0,t).$$

Here we use the condition $b(1,t) - \partial_x a(1,t) > 0$. Consider the difference $r_n = u_n - v$. The function r_n is a solution of the Cauchy problem

$$\partial_t r_n = (a + n^{-1})\partial_x^2 r_n + b\partial_x r_n + cr_n + n^{-1}\partial_x^2 v,$$

$r_n(x,0) = 0$, $r_n(1,t) = 0$ and $r_n(0,t) = u_n(0,t) - u(0,t)$. By the maximum principle (see, for example, [51], Theorem 1.1.2) we have

$$\max_{[0,1]\times[0,T]}|r_n(x,t)| \leqslant \frac{1}{nc_0}\max_{[0,1]\times[0,T]}|\partial_x^2 v(x,t)| + \max_{[0,T]}|u_n(0,t) - u(0,t)|.$$

Hence the sequence $\{r_n\}$ converges uniformly to zero, and u coincides with v. Thus, the function u is continuously differentiable in t and twice continuously differentiable in x on $[0,1]\times[0,T]$, satisfies the initial condition $u = u_0$ for $t = 0$, the boundary condition $\partial_x u + ku = 0$ for $x = 0$ and the boundary condition $u = 0$ for $x = 1$. The uniqueness of the solution we have constructed follows from the maximum principle. Proposition 4.3 is proved.

Remark 4.4. Since in the case where

$$\int_{-\infty}^{0} \frac{1}{\sqrt{a(s)}}\, ds = \int_{0}^{+\infty} \frac{1}{\sqrt{a(s)}}\, ds = \infty$$

the probability solution is unique, the method of constructing examples of non-uniqueness suggested in Examples 4.1 and 4.2 should not work in this case. We show that this is indeed the case. After an appropriate change of coordinates we obtain the Cauchy problem on the interval $(0,1)$ for the equation of the form

$$\partial_t \sigma = \partial_x^2 (A\sigma) - \partial_x (B\sigma),$$

where $A(0) = \partial_x A(0) = A(1) = \partial_x A(1) = 0$. Thus, this equation degenerates at the endpoints of the interval $(0,1)$. To construct a probability solution, we have to ensure that the expression $\displaystyle\int_0^1 \sigma(x,t)\, dx$ is independent of t. Since we are constructing a smooth solution, due to the equation the latter requirement implies the equality

$$0 = \frac{d}{dt} \int_0^1 \sigma(x,t)\, dx = B(0)\sigma(0,t) - B(1)\sigma(1,t).$$

On the other hand, for the problem with boundary conditions at $x = 0$ and $x = 1$ to be solvable, it is necessary (see [51], Ch. 1) that $B(0) > 0$ and $B(1) < 0$. This means that the numbers $\sigma(0,t)$ and $\sigma(1,t)$ must have different signs. In turn, this contradicts the positivity of the solution. It is important to note here that we cannot pick B such that $\partial_x A = B + \text{const}$, because $\partial_x A$ vanishes at the endpoints of the interval $[0,1]$ and B assumes values with different signs at these points. Therefore, it is not possible, by choosing the coefficient B, to ensure that a solution remains a solution after adding a constant and thus guarantee the positivity of the solution.

The idea of constructing an example of non-uniqueness for the Fokker-Planck-Kolmogorov equation with the aid of a change of coordinates and application of the theory of degenerate equations was first suggested in [42] in the case of a stationary equation.

Example 4.5. (i) Consider an example where, for an infinitely differentiable drift coefficient b on the real line, in addition to a unique probability solution with smooth initial distribution there exists another family of nonnegative bounded measures that is a solution with the same initial condition. In [11], Exercise 9.8.47, the following example was proposed (with a hint for the solution):

$$b(x) = -2x(1+x^2)^{-1} - (1+x^2)\arctan x,$$

the initial density is $u_0(x) = (\pi(1+x^2))^{-1}$. We give a solution of this exercise here. One nonnegative integrable solution has a simple explicit form $e^t u_0(x)$. This can be verified directly:

$$u_0'(x) = -2\pi^{-1} x(1+x^2)^{-2},$$
$$u_0''(x) = -2\pi^{-1}(1+x^2)^{-2} + 8\pi^{-1}x^2(1+x^2)^{-3},$$
$$(b(x)u_0(x))' = -2\pi^{-1}(1+x^2)^{-2} + 8\pi^{-1}x^2(1+x^2)^{-3} - \pi^{-1}(1+x^2)^{-1}.$$

However, there is also a unique probability solution. Indeed, since $b(x)x \leqslant 0$, for the function $V(x) = x^2$ we have $V''(x) + b(x)V'(x) = 2 + 2xb(x) \leqslant 2$. It is known (see [11], Theorem 9.4.8, where a Lyapunov function with the estimate $V'' + bV' \leqslant C + CV$ is required) that this guarantees the existence of a unique probability solution of the Cauchy problem (though, the uniqueness also follows from the main theorem of this paper).

(ii) Consider an example where for an infinitely differentiable drift coefficient b on the real line and a smooth initial condition there are no probability solutions, but there exists a unique subprobability solution. We take the same initial condition u_0 as in (i), but the drift is changed as follows:

$$b(x) = -2x(1+x^2)^{-1} + (1+x^2)\arctan x.$$

Now, since $xb(x) \geqslant -2$, for $V(x) = x^2 + 4$ we obtain $V'' + bV' \geqslant -V$, which according to Theorem 9.6.3 in [11] implies the uniqueness of an integrable solution. The calculations in (i) show that $e^{-t}u_0(x)$ is such a solution; moreover, it is a subprobability solution.

In [11], § 9.6, in the multi-dimensional case sufficient conditions on the coefficients are given in terms of Lyapunov functions under which there is at most one integrable solution of the Cauchy problem.

Remark 4.6. Now we explain why the problem we consider is not equivalent to the existence and uniqueness problem studied by Hille [37] for the one-dimensional Fokker-Planck-Kolmogorov equation. For simplicity, we confine ourselves to the case of a unit diffusion coefficient (note that in [37] the opposite notation is used, the drift is denoted by a, but we give the formulations below using our notation). The problem posed in [37], §8, p. 116 (in the case of the equation on the whole real line), is this: find necessary and sufficient conditions in order that for every function $h \in L^1(\mathbb{R})$ with $Lh = h'' - (bh)' \in L^1(\mathbb{R})$ a unique solution $T(x,t,h)$ of the equation $\partial_t u = \partial_x^2 u - \partial_x(ub)$ can be found with initial condition h in the sense of the relation $\|T(\,\cdot\,,t,h) - h\|_{L^1} \to 0$ as $t \to 0$. This setting is called Problem L_0, and in Problem L it is required in addition that the solution with any nonnegative initial condition h be nonnegative and have the same integral over the real line as h (in other words, the solutions with probability initial densities in the domain of definition of the operator L must be probabilistic). The drift coefficient in [37] is assumed to be continuous, but this is a minor technical difference. According to Theorems 8.5 and 8.7 in [37], a necessary and sufficient condition for the solvability of Problem L_0 is that the integral

$$\int_0^x \exp B(y) \int_0^y \exp(-B(u))\,du\,dy, \quad \text{where } B(y) = \int_0^y b(s)\,ds,$$

diverges at $-\infty$ and $+\infty$; for Problem L to be solvable the divergence of the integral

$$\int_0^x \exp(-B(y)) \int_0^y \exp(B(u))\,du\,dy$$

must also diverge at $-\infty$ and $+\infty$. This is the previous condition for the drift $-b$, which makes the conditions for b and $-b$ the same. In both theorems of Hille we have cited the closure of the operator L generates a semigroup on $L^1(\mathbb{R})$. We see in the previous example that, for every initial condition that is a probability measure, there may exist a unique probability solution of the Cauchy problem, but there are also other solutions. Finally, yet another difference between Hille's conditions and our result (in which there are no conditions on the drift at all, apart from local boundedness) is that Hille's solution must exist for every initial condition (although with density in the domain of definition of the operator, while we admit arbitrary initial probability measures), and in our setting solutions can exist for some initial conditions and not for others, but we show that for no initial probability distribution can two different solutions exist. Moreover, uniqueness holds even in the case when the closure of the operator L does not generate any semigroup on $L^1(\mathbb{R})$.

It is also worth noting that Hille's theorem says (for a continuous drift) that, if Hille's condition for Problem L is violated, then either for some initial condition there is no solution or for some other initial condition there are several solutions, while our result shows that the second possibility cannot occur, so that the reason is always the first case. Moreover, it follows from our result that there is no solution for the initial distribution that is the Dirac measure δ_a at some point a. In fact, we can verify that if the solution $\varrho(x,t,a)$ exists for every initial condition of the form δ_a, then due to uniqueness it depends on a Borel measurably, which after averaging over the probability measure ν gives a solution with initial condition ν.

We give an example of the absence of a probability solution for some initial distributions and its existence for others. In the justification of the following example we use the drift $b(x)=-x-6\exp(x^2/2)$, for which the standard Gaussian measure γ on the real line is the stationary solution with initial distribution γ. We verify that here there exist initial probability distributions for which there are no probability solutions; using Hille's terminology this means that Problem L is not solvable for this drift. To this end we show that Hille's condition is violated, namely, the second of the two integrals above converges at $-\infty$. In our case, after changing from x to $-x$, we arrive at the integral over $[0,+\infty)$ of the function $F(x)/F'(x)$, where

$$F(x)=\int_0^x f(y)\,dy, \qquad f(x)=\exp\biggl(-2^{-1}x^2+6\int_0^x \exp\biggl(\frac{y^2}{2}\biggr)\,dy\biggr),$$

the function f is increasing and the estimate $F''(x)/F'(x)\geqslant x^2$ holds. We observe that then also $F'(x)/F(x)\geqslant x^2/8$: indeed, integrating the estimate $F''(x)\geqslant x^2F'(x)$ we obtain

$$\begin{aligned} F'(x)-F'(0) &\geqslant \int_0^x y^2F'(y)\,dy \geqslant \int_{x/2}^x y^2F'(y)\,dy \\ &\geqslant \frac{x^2}{4}\biggl(F(x)-F\biggl(\frac{x}{2}\biggr)\biggr) \geqslant \frac{x^2}{8}F(x), \end{aligned}$$

since $2F(x/2)\leqslant F(x)$. The latter is seen from the equality

$$2F\biggl(\frac{x}{2}\biggr)=2\int_0^{x/2} f(y)\,dy=\int_0^x f\biggl(\frac{y}{2}\biggr)\,dy$$

and the inequality $f(y/2)\leqslant f(y)$. Thus, $F'(x)\geqslant x^2F(x)/8$, which implies that the integral of F/F' converges at $+\infty$. The same conclusion can be obtained by finding an asymptotic relation for the ratio F/F' with the aid of L'Hôpital's rule, which shows that the ratio of $F(x)(-x+6\exp(x^2/2))$ and $f(x)$ tends to 1, that is, $F(x)/F'(x)$ actually decreases even more rapidly. By Hille's criterion there is no probability solution for some initial distribution, but it follows from what we said above that then there is no solution for some Dirac initial measure. Note that by Theorem 6.6.2 in [11] a subprobability solution exists for every initial probability distribution. By the way, this implies the absence of a diffusion process (in the usual sense) generated by the operator L, in spite of the existence of the infinitesimally invariant measure γ for this operator.

We now construct an example of a Fokker-Planck-Kolmogorov equation in dimension $d=2$ for which the Cauchy problem has an infinite-dimensional simplex of probability solutions. We recall that such examples have previously only been constructed for $d\geqslant 3$ (see [11], Ch. 9).

Example 4.7. Let γ be the standard Gaussian measure on the real line given by the density $(2\pi)^{-1/2}\exp(-x^2/2)$. Set

$$b^1(x)=-x-6\exp\biggl(\frac{x^2}{2}\biggr),\qquad b^2(y)=-y.$$

Let $\{T_t\}_{t\geqslant 0}$ be the standard Ornstein-Uhlenbeck semigroup (see, for example, [6]) generated by the operator $L_yu=u''+b^2u'$ on the space $L^1(\gamma)$ and defined by the formula

$$T_tf(x)=\int f\Bigl(e^{-t}x-\sqrt{1-e^{-2t}}\,y\Bigr)\,\gamma(dy).$$

It is straightforward to verify that the measure γ also satisfies the stationary equation with the operator $L_xu=u''+b^1u'$. As indicated in § 2 above, on $L^1(\gamma)$ there exists a sub-Markov semigroup $\{S_t\}_{t\geqslant 0}$ associated with the operator L_x for which γ is a subinvariant measure. However, it is known (and it is important for the sequel) that γ is not an invariant measure for this semigroup (see [11], Exercises 4.5.17 and 5.6.49). The semigroups $\{T_t\}_{t\geqslant 0}$ and $\{S_t\}_{t\geqslant 0}$ also act on every measure ν with density and give nonnegative measures $T_t^*\nu$ and $S_t^*\nu$ by the formulae

$$\int f\,d(T_t^*\nu)=\int T_tf\,d\nu\quad\text{and}\quad\int f\,d(S_t^*\nu)=\int S_tf\,d\nu.$$

In terms of the density g of the measure ν with respect to γ we can write

$$T_t^*\nu=T_tg\cdot\gamma\quad\text{and}\quad S_t^*\nu=S_t^*g\cdot\gamma,$$

where S_t^*g is the action on g of the operator on the space $L^1(\gamma)$ obtained by extending the operator adjoint to S_t from $L^\infty(\gamma)$ (if the density g is bounded, then this is the action of the adjoint operator itself).

For any function u of two variables we set

$$Lu=\partial_x^2u+\partial_y^2u+b^1(x)\,\partial_xu+b^2(y)\,\partial_yu=L_xu+L_yu.$$

Let σ be an arbitrary probability measure with smooth density. Consider the Cauchy problem

$$\partial\mu_t=L^*\mu_t,\qquad\mu_0=\gamma\otimes\sigma.$$

We verify that this problem has infinitely many different probability solutions of the form

$$\mu_t^\alpha = S_t^*\gamma \otimes (T_t^*\sigma - T_t^*\alpha) + \gamma \otimes T_t^*\alpha,$$

where α is an arbitrary probability measure with smooth density; moreover, solutions $\mu_t^{\alpha_j}$ corresponding to linearly independent measures α_j are linearly independent.

First, $\{\mu_t^\alpha\}$ is a solution, since in every variable it satisfies the corresponding equation: $\{S_t^*\gamma\}$ is a solution of the one-dimensional Cauchy problem with the operator L_x and the initial condition γ, the measure γ is the stationary solution for the equation with the operator L_x, $\{T_t^*\sigma\}$ and $\{T_t^*\alpha\}$ are solutions to the Cauchy problems with the operator L_y and the initial conditions σ and α, respectively. Since $\{T_t\}_{t\geqslant 0}$ is obviously a Markov semigroup, $T_t^*\sigma$ and $T_t^*\alpha$ are probability measures for each t. In addition, as noted in § 2, $S_t^*\gamma \leqslant \gamma$ in the sense of the inequality for measures; moreover, $S_t^*\gamma \neq \gamma$ for $t > 0$ due to the lack of invariance (if $S_t^*\gamma = \gamma$ for some $t > 0$, then $S_\tau^*\gamma = \gamma$ for all $\tau \leqslant t$, hence $S_t^*\gamma = \gamma$ for all $t > 0$).

Second, μ_t^α is a nonnegative measure. Indeed,

$$\mu_t^\alpha = S_t^*\gamma \otimes T_t^*\sigma + (\gamma - S_t^*\gamma) \otimes T_t^*\alpha \geqslant 0.$$

In addition, μ_t^α is a probability measure since

$$\mu_t^\alpha(\mathbb{R}^2) = S_t^*\gamma(\mathbb{R}^1) \cdot (T_t^*\sigma - T_t^*\alpha)(\mathbb{R}^1) + \gamma(\mathbb{R}^1) \cdot T_t^*\alpha(\mathbb{R}^1) = S_t^*\gamma(\mathbb{R}^1) \cdot 0 + 1 = 1.$$

We now verify that for linearly independent probability measures α_j the corresponding solutions are linearly independent. Suppose that for some finite collection of numbers c_j the equality

$$\sum_j c_j(S_t^*\gamma \otimes (T_t^*\sigma - T_t^*\alpha_j) + \gamma \otimes T_t^*\alpha_j) = 0$$

holds for all $t \geqslant 0$. This can be written in the form

$$\biggl(\sum_j c_j\biggr) S_t^*\gamma \otimes T_t^*\sigma + (\gamma - S_t^*\gamma) \otimes T_t^*\biggl(\sum_j c_j\alpha_j\biggr) = 0.$$

For $t = 0$ we obtain $\bigl(\sum_j c_j\bigr)\gamma \otimes \sigma = 0$, and therefore $\sum_j c_j = 0$. Hence

$$(\gamma - S_t^*\gamma) \otimes T_t^*\biggl(\sum_j c_j\alpha_j\biggr) = 0.$$

Since $\gamma - S_t^*\gamma$ is a positive measure for all $t > 0$, we have $T_t^*\bigl(\sum_j c_j\alpha_j\bigr) = 0$ for $t > 0$, but this, in turn, implies the equality $\sum_j c_j\alpha_j = 0$. The linear independence of the α_j shows that all the numbers c_j equal zero.

Note that, in this example, non-uniqueness takes place for some smooth, but very special initial condition. Naturally, the question arises about constructing an example where non-uniqueness takes place for some broad class of initial conditions, say, for all Dirac measures. In the case when the variable x is three-dimensional, such examples can be constructed by combining the idea from the example described and the approach based on the theory of degenerate parabolic equations in the spirit of Examples 4.1 and 4.2. A detailed discussion of such examples will be the subject of a separate paper.

Now we give examples showing that if we omit the local boundedness of the drift, the uniqueness of a probability solution can fail for the elliptic equation as well as for the parabolic one. However, the question of uniqueness for drifts that are locally integrable to some power with respect to Lebesgue measure is worth studying (of course, keeping the requirement of the local integrability with respect to the solution).

Example 4.8. Any locally Lipschitz probability density ϱ obviously satisfies the stationary equation $\varrho'' - (b\varrho)' = 0$ with the drift b equal to the logarithmic derivative of ϱ, that is, $b(x) = \varrho'(x)/\varrho(x)$, where we set $b(x) = 0$ if $\varrho(x) = 0$. The function b is locally integrable with weight ϱ. If, say, we take $\varrho(x) = (2\pi)^{-1/2}x^2\exp(-x^2)$, then for the drift $b(x) = 2x^{-1} - 2x$ obtained there are other probability solutions, distinct from the one given by the density ϱ, including the solution with density $\varrho/2$ on $(-\infty, 0)$ and density $3\varrho/2$ on $[0, +\infty)$. In this case the function b is integrable with weight ϱ on the whole real line.

In the parabolic case we consider the function

$$\varrho(x,t) = t^{-2}x^3E(x,t), \qquad E(x,t) = C\exp\biggl(-\frac{x^2}{4t}\biggr), \qquad x \geqslant 0, \quad t > 0,$$

where $C > 0$ is a constant such that the function $\varrho(\,\cdot\,,1)$ is a probability density on $[0,+\infty)$; then all the functions $\varrho(\,\cdot\,,t)$ are too. Set $\varrho(x,t) = 0$ if $x < 0$ and $\varrho_1(x,t) = \varrho(-x,t)$.

Then the functions ϱ and ϱ_1 give distinct probability solutions of the equation

$$\partial_t\varrho = \partial_x^2\varrho - \partial_x(b\varrho), \qquad b(x) = \frac{3}{x},$$

with the initial distribution at $t = 0$ equal to Dirac's measure at the origin.

Proof. We verify that for $t > 0$ and $x \geqslant 0$ the twice differentiable function ϱ satisfies pointwise the equation indicated. Indeed,

$$\partial_t\varrho = -2t^{-3}x^3E(x,t) + \frac{1}{4}t^{-4}x^5E(x,t),$$

$$\partial_x\varrho = 3t^{-2}x^2E(x,t) - \frac{1}{2}t^{-3}x^4E(x,t),$$

$$\begin{aligned}\partial_x^2\varrho &= 6t^{-2}xE(x,t) - \frac{3}{2}t^{-3}x^3E(x,t) - 2t^{-3}x^3E(x,t) + \frac{1}{4}t^{-4}x^5E(x,t)\\ &= 6t^{-2}xE(x,t) - \frac{7}{2}t^{-3}x^3E(x,t) + \frac{1}{4}t^{-4}x^5E(x,t),\end{aligned}$$

$$b\varrho = 3t^{-2}x^2E(x,t),$$

$$\partial_x(b\varrho) = 6t^{-2}xE(x,t) - \frac{3}{2}t^{-3}x^3E(x,t)$$

and

$$\partial_x^2\varrho - \partial_x(b\varrho) = -2t^{-3}x^3E(x,t) + \frac{1}{4}t^{-4}x^5E(x,t) = \partial_t\varrho.$$

At the point $x = 0$ the function $x \mapsto \varrho(x,t)$ with $t > 0$ is twice differentiable and

$$\varrho(0,t) = \partial_x \varrho(0,t) = \partial_x^2 \varrho(0,t) = 0.$$

In addition, for $t > 0$ the function $x \mapsto b(x)\varrho(x,t)$ is continuously differentiable at zero, and its derivative at zero vanishes. Therefore, setting $\varrho(x,t) = 0$ for $x < 0$ we obtain a twice differentiable solution of our equation on the whole real line. Moreover, as $t \to 0$, the measures $\varrho(x,t)\,dx$ converge weakly to the Dirac measure at zero. Finally, we verify that $b\varrho \in L^1(\mathbb{R} \times [0,T])$. Now, this product vanishes if $x < 0$, and for $x > 0$ we have

$$b(x)\varrho(x,t) = Ct^{-2}x^2 \exp\biggl(-\frac{x^2}{4t}\biggr).$$

Then for $t > 0$ we obtain

$$\int_0^{+\infty} b(x)\varrho(x,t)\,dx = Ct^{-1/2}\int_0^{+\infty} u^2 \exp\biggl(-\frac{u^2}{4}\biggr)\,du = C't^{-1/2}.$$

Therefore,

$$\int_0^T \int_0^{+\infty} b(x)\varrho(x,t)\,dx\,dt < \infty.$$

The integrability of $b\varrho$ implies that ϱ is a solution to the Cauchy problem in our sense, that is, in the sense of the integral identity. We now observe that the measures with densities $\varrho_1(x,t) = \varrho(-x,t)$, concentrated on the left half-line, also give a solution. Therefore, our problem has at least two linearly independent solutions.

We should explain the probabilistic nature of this example. It was shown in [22] and [23], Example 1.23, that for the singular drift

$$b(x) = \frac{3}{2}x^{-1}I_{\mathbb{R}\setminus 0}(x)$$

the stochastic differential equation $dX_t = b(X_t)dt + dW_t$ with the initial distribution concentrated at zero has several solutions with different one-point distributions. One of these solutions is the nonnegative Bessel process B_t with the parameter $\alpha = 4$, another solution $X_t = -B_t$ is nonpositive; moreover, $|X_t|$ has the same distribution as B_t. This leads to two different probability solutions $\{\mu_t\}$ and $\{\nu_t\}$ of the corresponding Fokker-Planck-Kolmogorov equation, with respect to which the drift b is integrable on $\mathbb{R} \times [0,1]$. The integrability of the drift is verified explicitly with the aid of the known formula for the density of the distribution of the square B_t^2 of the Bessel process (see [56], Ch. XI, § 1, Corollary 1.4), which also gives the explicit densities of the distributions of B_t and X_t. The coefficients of the Fokker-Planck-Kolmogorov equation for B_t and X_t differ from the ones indicated in our example by some numeric coefficients.

Note that in the last example only the drift coefficient is singular with respect to Lebesgue measure, but the solutions are smooth, and this coefficient is integrable with respect to the solutions. By the way, this shows that in the uniqueness theorems from [11], Ch. 9, with the identity diffusion matrix, the hypotheses cannot be weakened even for $d = 1$ to the inclusion of the drift b in $L^1(\mu)$: in Theorem 9.3.6 in [11], in addition to the condition $|b| \in L^1(\mu)$ the inclusion $|b| \in L^2(\mu, U \times (0,T))$ is also required for every ball U and in Theorem 9.4.3 in [11], in addition to the previous condition, the inclusion $|b| \in L^p(\mu, U \times (0,T))$ with some $p > d+2$ is necessary, that is, $p > 3$ in the one-dimensional case.

Remark 4.9. In this work we have studied equations with coefficients independent of t. In our proofs this has played an essential role, and the uniqueness problem in the case when the coefficients depend on x and t remains open. We only observe that if $a = 1$ and the drift coefficient has the form $b(x,t) = h(t)$, then by the change of the solution $\varrho(x,t)$ for the new function

$$\sigma(x,t) = \varrho(x - H(t), t), \qquad H(t) = \int_0^t h(s)\, ds,$$

the problem reduces to the heat equation. Therefore, in this case the probability solution is unique.

Bibliography

[1] A. Albanese, L. Lorenzi and E. Mangino, "L^p-uniqueness for elliptic operators with unbounded coefficients in $\mathbb{R}^N$", *J. Funct. Anal.* **256**:4 (2009), 1238–1257.

[2] D. G. Aronson, "Non-negative solutions of linear parabolic equations", *Ann. Scuola Norm. Sup. Pisa Fis. Mat.* (3) **22** (1968), 607–694.

[3] D. G. Aronson and P. Besala, "Uniqueness of solutions of the Cauchy problem for parabolic equations", *J. Math. Anal. Appl.* **13**:3 (1966), 516–526.

[4] A. Attalienti and M. Campiti, "Semigroups generated by ordinary differential operators in $L^1(I)$", *Positivity* **8**:1 (2004), 11–30.

[5] E. A. Baderko, "Solvability of boundary-value problems for high-order parabolic equations in regions with curved lateral boundaries", *Differ. Uravn.* **12**:10 (1976), 1781–1792; English transl. in *Differential Equations* **12**:10 (1976), 1253–1261.

[6] V. I. Bogachev, "Ornstein-Uhlenbeck operators and semigroups", *Uspekhi Mat. Nauk* **73**:2(440) (2018), 3–74; English transl. in *Russian Math. Surveys* **73**:2 (2018), 191–260.

[7] V. I. Bogachev, G. Da Prato, M. Röckner and S. V. Shaposhnikov, "On the uniqueness of solutions to continuity equations", *J. Differential Equations* **259**:8 (2015), 3854–3873.

[8] 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. Akad. Nauk* **482**:5 (2018), 489–493; English transl. in *Dokl. Math.* **98**:2 (2018), 475–479.

[9] 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**:11–12 (2001), 2037–2080.

[10] V. I. Bogachev, N. V. Krylov and M. Röckner, "Elliptic and parabolic equations for measures", *Uspekhi Mat. Nauk* **64**:6(390) (2009), 5–116; English transl. in *Russian Math. Surveys* **64**:6 (2009), 973–1078.

[11] V. I. Bogachev, N. V. Krylov, M. Röckner and S. V. Shaposhnikov, *Fokker-Planck-Kolmogorov equations*, Regulyarnaya i Khaoticheskaya Dinamika, Moscow–Izhevsk 2013, 592 pp.; English transl. in Math. Surveys Monogr., vol. 207, Amer. Math. Soc., Providence, RI 2015, xii+479 pp.

[12] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov, "Global regularity and bounds for solutions of parabolic equations for probability measures", *Teor. Veroyatnost. i Primenen.* **50**:4 (2005), 652–674; English transl. in *Theory Probab. Appl.* **50**:4 (2006), 561–581.

[13] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov, "Estimates of densities of stationary distributions and transition probabilities of diffusion processes", *Teor. Veroyatnost. i Primenen.* **52**:2 (2007), 240–270; English transl. in *Theory Probab. Appl.* **52**:2 (2008), 209–236.

[14] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov, "Positive densities of transition probabilities of diffusion processes", *Teor. Veroyatnost. i Primenen.* **53**:2 (2008), 213–239; English transl. in *Theory Probab. Appl.* **53**:2 (2009), 194–215.

[15] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov, "On uniqueness problems related to the Fokker-Planck-Kolmogorov equation for measures", Probl. Mat. Anal., vol. 61, Tamara Rozhkovskaya Publishing House, Novosibirsk 2011, pp. 9–42; English transl. in *J. Math. Sci.* (*N.Y.*) **179**:1 (2011), 7–47.

[16] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov, "On uniqueness of solutions to the Cauchy problem for degenerate Fokker-Planck-Kolmogorov equations", *J. Evol. Equ.* **13**:3 (2013), 577–593.

[17] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov, "Uniqueness problems for degenerate Fokker-Planck-Kolmogorov equations", Probl. Mat. Anal., vol. 78, Tamara Rozhkovskaya Publishing House, Novosibirsk 2015, pp. 31–46; English transl. in *J. Math. Sci.* (*N.Y.*) **207**:2 (2015), 147–165.

[18] 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**:5 (2016), 1262–1300.

[19] V. I. Bogachev and S. V. Shaposhnikov, "Representations of solutions to Fokker-Planck-Kolmogorov equations with coefficients of low regularity", *J. Evol. Equ.* **20**:2 (2020), 355–374.

[20] A. Bove, B. Franchi and E. Obrecht, "Parabolic problems with mixed time dependent lateral conditions", *Comm. Partial Differential Equations* **7**:11 (1982), 1253–1288.

[21] A. Bove, B. Franchi and E. Obrecht, "Boundary value problems with mixed lateral conditions for parabolic operators", *Ann. Mat. Pura Appl.* (4) **131** (1982), 375–413.

[22] A. S. Cherny, "On the strong and weak solutions of stochastic differential equations governing Bessel processes", *Stochastics Stochastic Rep.* **70**:3–4 (2000), 213–219.

[23] A. S. Cherny and H.-J. Engelbert, *Singular stochastic differential equations*, Lecture Notes in Math., vol. 1858, Springer-Verlag, Berlin 2005, viii+128 pp.

[24] A. Eberle, *Uniqueness and non-uniqueness of semigroups generated by singular diffusion operators*, Lecture Notes in Math., vol. 1718, Springer-Verlag, Berlin 1999, viii+262 pp.

[25] A. Eberle, "L^p uniqueness of non-symmetric diffusion operators with singular drift coefficients. I. The finite-dimensional case", *J. Funct. Anal.* **173**:2 (2000), 328–342.

[26] G. M. Fateeva, "Boundary value problems for quasilinear degenerate equations of parabolic type", *Uspekhi Mat. Nauk* **22**:3(135) (1967), 244–245. (Russian)

[27] G. M. Fateeva, "Boundary value problems for degenerate quasilinear parabolic equations", *Mat. Sb.* (*N.S.*) **76(118)**:4 (1968), 537–565; English transl. in *Math. USSR-Sb.* **5**:4 (1968), 509–532.

[28] W. Feller, “Zur Theorie der stochastischen Prozesse. Existenz- und Eindeutigkeitssätze”, *Math. Ann.* **113**:1 (1937), 113–160.

[29] W. Feller, “The parabolic differential equations and the associated semi-groups of transformations”, *Ann. of Math.* (2) **55**:3 (1952), 468–519.

[30] W. Feller, “Generalized second order differential operators and their lateral conditions”, *Illinois J. Math.* **1**:4 (1957), 459–504.

[31] G. Fichera, “On a unified theory of boundary value problems for elliptic parabolic equations of second order”, *Boundary problems in differential equations*, Univ. of Wisconsin Press, Madison, WI 1960, pp. 97–120.

[32] A. Friedman, “On the uniqueness of the Cauchy problem for parabolic equations”, *Amer. J. Math.* **81**:2 (1959), 503–511.

[33] A. Friedman, *Partial differential equations of parabolic type*, Prentice-Hall, Inc., Englewood Cliffs, NJ 1964, xiv+347 pp.

[34] A. K. Gushchin, “The Dirichlet problem for a second-order elliptic equation with an L_p boundary function”, *Mat. Sb.* **203**:1 (2012), 3–30; English transl. in *Sb. Math.* **203**:1 (2012), 1–27.

[35] A. K. Gushchin, “Solvability of the Dirichlet problem for an inhomogeneous second-order elliptic equation”, *Mat. Sb.* **206**:10 (2015), 71–102; English transl. in *Sb. Math.* **206**:10 (2015), 1410–1439.

[36] A. K. Gushchin, “The boundary values of the solutions of an elliptic equation”, *Mat. Sb.* **210**:12 (2019), 67–97; English transl. in *Sb. Math.* **210**:12 (2019), 1724–1752.

[37] E. Hille, “The abstract Cauchy problem and Cauchy’s problem for parabolic differential equations”, *J. Anal. Math.* **3** (1953), 81–196.

[38] K. Ishige and M. Murata, “Uniqueness of nonnegative solutions of the Cauchy problem for parabolic equations on manifolds or domains”, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **30**:1 (2001), 171–223.

[39] A. V. Ivanov, “Quasilinear degenerate and nonuniformly elliptic and parabolic equations of second order”, Tr. Mat. Inst. Steklova, vol. 160, 1982, pp. 3–285; English transl. in *Proc. Steklov Inst. Math.* **160** (1984), 1–288.

[40] A. Kolmogoroff (Kolmogorov), “Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung”, *Math. Ann.* **104**:1 (1931), 415–458.

[41] A. Kolmogoroff (Kolmogorov), “Zur Theorie der stetigen zufälligen Prozesse”, *Math. Ann.* **108**:1 (1933), 149–160.

[42] T. I. Krasovitskii, “Degenerate elliptic equations and nonuniqueness of solutions to the Kolmogorov equation”, *Dokl. Akad. Nauk* **487**:4 (2019), 361–364; English transl. in *Dokl. Math.* **100**:1 (2019), 354–357.

[43] O. A. Ladyženskaja (Ladyzhenskaya), V. A. Solonnikov and N. N. Ural’ceva (Uraltseva), *Linear and quasi-linear equations of parabolic type*, Nauka, Moscow 1967, 736 pp.; English transl. in Transl. Math. Monogr., vol. 23, Amer. Math. Soc., Providence, RI 1968, xi+648 pp.

[44] C. Le Bris and P.-L. Lions, “Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients”, *Comm. Partial Differential Equations* **33**:7–9 (2008), 1272–1317.

[45] L. D. Lemle, “On the L^∞-uniqueness of symmetric diffusion operators on complete non-compact Riemannian manifolds”, *J. Geom. Anal.* **25**:4 (2015), 2375–2385.

[46] G. M. Lieberman, “A mostly elementary proof of Morrey space estimates for elliptic and parabolic equations with VMO coefficients”, *J. Funct. Anal.* **201**:2 (2003), 457–479.

[47] O. A. Manita and S. V. Shaposhnikov, "On the Cauchy problem for Fokker-Planck-Kolmogorov equations with potential terms on arbitrary domains", *J. Dynam. Differential Equations* **28**:2 (2016), 493–518.

[48] P. J. Mendez-Hernandez and M. Murata, "Semismall perturbations, semi-intrinsic ultracontractivity, and integral representations of nonnegative solutions for parabolic equations", *J. Funct. Anal.* **257**:6 (2009), 1799–1827.

[49] G. Metafune, D. Pallara and M. Wacker, "Feller semigroups on $\mathbf{R}^N$", *Semigroup Forum* **65**:2 (2002), 159–205.

[50] M. Murata, "Non-uniqueness of the positive Cauchy problem for parabolic equations", *J. Differential Equations* **123**:2 (1995), 343–387.

[51] O. A. Oleĭnik and E. V. Radkevič (Radkevich), *Second order equations with nonnegative characteristic form*, Moscow Univ., Moscow 2010, 359 pp.; English transl., Springer-Verlag, Boston 1973, vii+259 pp.

[52] Y. Pinchover, "On uniqueness and nonuniqueness of the positive Cauchy problem for parabolic equations with unbounded coefficients", *Math. Z.* **223**:4 (1996), 569–586.

[53] Y. Pinchover, "Topics in the theory of positive solutions of second-order elliptic and parabolic partial differential equations", *Spectral theory and mathematical physics: a Festschrift in honor of Barry Simon's 60th birthday*, Part 1, Proc. Sympos. Pure Math., vol. 76, Part 1, Amer. Math. Soc., Providence, RI 2007, pp. 329–355.

[54] M. H. Protter and H. F. Weinberger, *Maximum principles in differential equations*, Corr. reprint of the 1967 original, Springer-Verlag, New York 1984, x+261 pp.

[55] E. V. Radkevich, "The second boundary value problem for a second order equation with non-negative characteristic form", *Vestnik Moskov. Univ. Ser.* I *Mat. Meh.*, 1967, no. 4, 3–11. (Russian)

[56] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, 3rd ed., Grundlehren Math. Wiss., vol. 293, Springer-Verlag, Berlin 1999, xiv+602 pp.

[57] G. Savaré, "Parabolic problems with mixed variable lateral conditions: an abstract approach", *J. Math. Pures Appl.* (9) **76**:4 (1997), 321–351.

[58] S. Sawyer, "A Fatou theorem for the general one-dimensional parabolic equation", *Indiana Univ. Math. J.* **24**:5 (1974/75), 451–498.

[59] S. V. Shaposhnikov, "On the uniqueness of integrable and probability solutions to the Cauchy problem for the Fokker-Planck-Kolmogorov equations", *Dokl. Akad. Nauk* **439**:3 (2011), 323–328; English transl. in *Dokl. Math.* **84**:1 (2011), 565–570.

[60] S. V. Shaposhnikov, "On uniqueness of a probability solution to the Cauchy problem for the Fokker-Planck-Kolmogorov equation", *Teor. Veroyatnost. i Primenen.* **56**:1 (2011), 77–99; English transl. in *Theory Probab. Appl.* **56**:1 (2012), 96–115.

[61] G. N. Smirnova, "Uniqueness classes of the Cauchy problem for parabolic equations", *Dokl. Akad. Nauk SSSR* **153**:6 (1963), 1269–1272; English transl. in *Soviet Math. Dokl.* **4** (1963), 1828–1831.

[62] G. N. Smirnova, "Cauchy problems for parabolic equations degenerating at infinity", *Mat. Sb.* (*N.S.*) **70(112)**:4 (1966), 591–604; English transl. in Amer. Math. Soc. Transl. Ser. 2, vol. 72, Amer. Math. Soc., Providence, RI 1968, pp. 119–134.

[63] V. A. Solonnikov, "On boundary value problems for linear parabolic systems of differential equations of general form", *Boundary value problems of mathematical physics. Part 3. On boundary value problems for linear parabolic systems of differential equations of general form*, Tr. Mat. Inst. Steklova, vol. 83, 1965, pp. 3–163; English transl. in *Proc. Steklov Inst. Math.* **83** (1965), 1–184.

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

[65] A. Tychonoff, "Théorèmes d'unicité pour l'équation de la chaleur", *Mat. Sb.* **42**:2 (1935), 199–216.

[66] A. D. Venttsel' (Wentzell), "Semi-groups of operators corresponding to a generalized second order differential operator", *Dokl. Akad. Nauk SSSR* **111**:2 (1956), 269–272. (Russian)

[67] A. D. Venttsel' (Wentzell), "On boundary conditions for multidimensional diffusion processes", *Teor. Veroyatnost. i Primenen.* **4**:2 (1959), 172–185; English transl. in *Theory Probab. Appl.* **4**:2 (1959), 164–177.

[68] D. V. Widder, "Positive temperatures on an infinite rod", *Trans. Amer. Math. Soc.* **55** (1944), 85–95.

[69] Liming Wu and Yiping Zhang, "A new topological approach to the L^∞-uniqueness of operators and the L^1-uniqueness of Fokker-Planck equations", *J. Funct. Anal.* **241**:2 (2006), 557–610.

[70] K. Yosida, "Integration of Fokker-Planck's equation in a compact Riemannian space", *Ark. Mat.* **1** (1949), 71–75.

Received 21/APR/20 and 28/NOV/20
Translated by V. BOGACHEV

Vladimir I. Bogachev
Faculty of Mechanics and Mathematics,
Lomonosov Moscow State University,
Moscow, Russia;
National Research University
Higher School of Economics,
Moscow, Russia
E-mail: vibogach@mail.ru

Tikhon I. Krasovitskii
Faculty of Mechanics and Mathematics,
Lomonosov Moscow State University,
Moscow, Russia;
Moscow Center for Fundamental
and Applied Mathematics,
Moscow, Russia
E-mail: tik714@yandex.ru

Stanislav V. Shaposhnikov
Faculty of Mechanics and Mathematics,
Lomonosov Moscow State University,
Moscow, Russia;
National Research University
Higher School of Economics,
Moscow, Russia
E-mail: starticle@mail.ru