Cambridge University Press
978-0-521-88651-2 - Partial Differential Equations for Probabilists
Daniel W. Stroock
Excerpt
More information

CHAPTER 1

Kolmogorov's Forward, Basic Results

The primary purpose of this chapter is to present some basic existence and uniqueness results for solutions to second order, parabolic, partial differential equations. Because this book is addressed to probabilists, the treatment of these results will follow, in so far as possible, a line of reasoning which is suggested by thinking about these equations in a probabilistic context. For this reason, I begin by giving an explanation of why, under suitable conditions, Kolmogorov's forward equations for the transition probability function of a continuous path Markov process is second order and parabolic. Once I have done so, I will use this connection with Markov processes to see how solutions to these equations can be constructed using probabilistically natural ideas.

1.1 Kolmogorov's Forward Equation

Recall that a *transition probability function* on $\mathbb{R}^N$ is a measurable map $(t,x) \in [0,\infty) \times \mathbb{R}^N \longmapsto P(t,x) \in \mathbf{M}_1(\mathbb{R}^N)$, where $\mathbf{M}_1(\mathbb{R}^N)$ is the space of Borel probability measures on $\mathbb{R}^N$ with the topology of *weak convergence*,[1] which, for each $x \in \mathbb{R}^N$, satisfies $P(0,x,\{x\}) = 1$ and the *Chapman–Kolmogorov equation*[2]

$$P(s+t,x,\Gamma) = \int P(t,y,\Gamma)\, P(s,x,dy) \quad \text{for all } s,t \in [0,\infty) \text{ and } \Gamma \in \mathcal{B}_{\mathbb{R}^N}. \tag{1.1.1}$$

Kolmogorov's forward equation is the equation which describes, for a fixed $x \in \mathbb{R}^N$, the evolution of $t \in [0,\infty) \longmapsto P(t,x) \in \mathbf{M}_1(\mathbb{R}^N)$.

1.1.1. Derivation of Kolmogorov's Forward Equation: In order to derive Kolmogorov's forward equation, we will make the assumption that

$$L\varphi(x) \equiv \lim_{h \searrow 0} \frac{1}{h} \int \big(\varphi(y) - \varphi(x)\big)\, P(h,x,dy) \tag{1.1.2}$$

[1] That is, the smallest topology for which the map $\mu \in \mathbf{M}_1(\mathbb{R}^N) \longmapsto \int \varphi\, d\mu \in \mathbb{R}$ is continuous whenever $\varphi \in C_{\mathrm{b}}(\mathbb{R}^N;\mathbb{R})$. See Chapter III of [53] for more information.
[2] $\mathcal{B}_{\mathbb{R}^N}$ denotes the Borel σ-algebra over $\mathbb{R}^N$.

Cambridge University Press
978-0-521-88651-2 - Partial Differential Equations for Probabilists
Daniel W. Stroock
Excerpt
More information

exists for each $x \in \mathbb{R}^N$ and $\varphi \in C_c^\infty(\mathbb{R}^N;\mathbb{R})$, the space of infinitely differentiable, real-valued functions with compact support. Under mild additional conditions, one can combine (1.1.2) with (1.1.1) to conclude that

$$\begin{aligned}\frac{d}{dt}\int \varphi(y)\,P(t,x,dy) &= \int L\varphi(y)\,P(t,x,dy) \quad \text{or, equivalently,}\\ \int \varphi(y)\,P(t,x,dy) &= \varphi(x) + \int_0^t \left(\int L\varphi(y)\,P(\tau,x,dy)\right)d\tau\end{aligned} \tag{1.1.3}$$

for $\varphi \in C_c^2(\mathbb{R}^N;\mathbb{R})$. This equation is called Kolmogorov's forward equation because it describes the evolution of $P(t,x,dy)$ in terms of its *forward variable* y, the variable giving the distribution of the process at time t, as opposed the the *backward variable* x which gives the initial position.

Thinking of $\mathbf{M}_1(\mathbb{R}^N)$ as a subset of $C_c^\infty(\mathbb{R}^N;\mathbb{R})^*$, the dual of $C_c^\infty(\mathbb{R}^N;\mathbb{R})$, one can rewrite (1.1.3) as

$$\frac{d}{dt}P(t,x) = L^\top P(t,x), \tag{1.1.4}$$

where $L^\top$ is the adjoint of L. Kolmogorov's idea was to recover $P(t,x)$ from (1.1.4) together with the initial condition $P(0,x) = \delta_x$, the unit point mass at x. Of course, in order for his idea to be of any value, one must know what sort of operator L can be. A general answer is given in § 2.1.1 of [55]. However, because this book is devoted to differential equations, we will not deal here with the general case but only with the case when L is a differential operator. For this reason, we add the assumption that L is *local*[3] in the sense that $L\varphi(x) = 0$ whenever φ vanishes in a neighborhood of x. Equivalently, in terms of $P(t,x)$, locality is the condition

$$\lim_{h\searrow 0}\frac{1}{h}P\big(h,x,B(x,r)\complement\big) = 0, \quad x \in \mathbb{R}^N \text{ and } r > 0. \tag{1.1.5}$$

LEMMA[4] 1.1.6. *Let $\{\mu_h : h \in (0,1)\} \subseteq \mathbf{M}_1(\mathbb{R}^N)$, and assume that*

$$A\varphi \equiv \lim_{h\searrow 0}\frac{1}{h}\int \big(\varphi(y) - \varphi(0)\big)\,\mu_h(dy)$$

exists for each $\varphi \in C_c^\infty(\mathbb{R}^N;\mathbb{R})$. Then A is a linear functional on $C_c^\infty(\mathbb{R}^N;\mathbb{R}) \oplus \mathbb{R}$ which satisfies the minimum principle

$$\varphi(0) = \min_{x\in\mathbb{R}^N}\varphi(x) \implies A\varphi \geq 0. \tag{1.1.7}$$

[3] Locality of L corresponds to path continuity of the associated Markov process.
[4] Readers who are familiar with Petrie's characterization of local operators may be surprised how simple it is to prove what, at first sight, might appear to be a more difficult result. Of course, the simplicity comes from the minimum principle, which allows one to control everything in terms of the action of A on quadratic functions.

Cambridge University Press
978-0-521-88651-2 - Partial Differential Equations for Probabilists
Daniel W. Stroock
Excerpt
More information

Moreover, if

$$\lim_{h\searrow 0}\frac{1}{h}\mu_h\big(B(0,r)\complement\big)=0\quad\textit{for all } r>0,$$

then A is local. Finally, if A is a linear functional on $C^\infty_c(\mathbb{R}^N;\mathbb{R})\oplus\mathbb{R}$, then A is local and satisfies the minimum principle if and only if there exists a non-negative, symmetric matrix[5] $a=((a_{ij}))_{1\le i,j\le N}\in\mathrm{Hom}(\mathbb{R}^N;\mathbb{R}^N)$ *and a vector* $b=(b_i)_{1\le i\le N}\in\mathbb{R}^N$ *such that*

$$A\varphi=\frac{1}{2}\sum_{i,j=1}^N a_{ij}\partial_{x_i}\partial_{x_j}\varphi(0)+\sum_{i=1}^N b_i\partial_{x_i}\varphi(0)\quad\textit{for all } \varphi\in C^\infty_c(\mathbb{R}^N;\mathbb{R}).$$

Proof: The first assertion requires no comment. To prove the "if" part of the second assertion, suppose A is given in terms of a and b with the prescribed properties. Obviously, A is then local. In addition, if φ achieves its minimum value at 0, then the first derivatives of φ vanish at 0 and its Hessian is non-negative definite there. Thus, after writing $\sum_{i,j=1}^N a_{ij}\partial_{x_i}\partial_{x_j}\varphi(0)$ as the trace of $a(0)$ times the Hessian of φ at 0, the non-negativity of Af comes down to the fact that the product of two non-negative definite, symmetric matrices has a non-negative trace, a fact that can be seen by first writing one of them as the square of a symmetric matrix and then using the commutation invariance properties of the trace.

Finally, suppose that A is local and satisfies the minimum principle. To produce the required a and b, we begin by showing that $A\varphi=0$ if φ vanishes to second order at 0. For this purpose, choose $\eta\in C^\infty_c\big(\mathbb{R}^N;[0,1]\big)$ so that $\eta=1$ on $B(0,1)$ and $\eta=0$ off $B(0,2)$, and set $\varphi_R(x)=\eta(R^{-1}x)\varphi(x)$ for $R>0$. Then, by locality, $A\varphi=A\varphi_R$ for all $R>0$. In addition, by Taylor's Theorem, there exists a $C<\infty$ such that $|\varphi_R|\le CR\psi$ for $R\in(0,1]$, where $\psi(x)\equiv\eta(x)|x|^2$. Hence, by the minimum principle applied to $CR\psi\mp\varphi_R$, $|A\varphi|=|A\varphi_R|\le CRA\psi$ for arbitrarily small R's.

To complete the proof from here, set $\psi_i(x)=\eta(x)x_i$, $\psi_{ij}=\psi_i\psi_j$, $b_i=A\psi_i$, and $a_{ij}=A\psi_{ij}$. Given φ, consider

$$\tilde\varphi=\varphi(0)+\frac{1}{2}\sum_{i,j=1}^N\partial_{x_i}\partial_{x_j}\varphi(0)\psi_{ij}+\sum_{i=1}^N\partial_{x_i}\varphi(0)\psi_i.$$

By Taylor's Theorem, $\varphi-\tilde\varphi$ vanishes to second order at 0, and therefore $A\varphi=A\tilde\varphi$. At the same time, by the minimum principle applied to the constant functions $\pm\varphi(0)$, A kills the first term on the right. Hence,

$$A\varphi=\frac{1}{2}\sum_{i,j=1}^N\partial_{x_i}\partial_{x_j}\varphi(0)A\psi_{ij}+\sum_{i=1}^N\partial_{x_i}\varphi(0)A\psi_i,$$

[5] We will use $\mathrm{Hom}(\mathbb{R}^M;\mathbb{R}^N)$ to denote the vector space of linear transformation from $\mathbb{R}^M$ to $\mathbb{R}^N$.

and so it remains to check that a is non-negative definite. But, if $\xi \in \mathbb{R}^N$ and $\psi_\xi(x) \equiv \eta(x)^2(\xi, x)^2_{\mathbb{R}^N} = \sum_{i,j=1}^N \xi_i\xi_j\psi_{ij}(x)$, then, by the minimum principle, $0 \le 2A\psi_\xi = (\xi, a\xi)_{\mathbb{R}^N}$. □

Since the origin can be replaced in Lemma 1.1.6 by any point $x \in \mathbb{R}^N$, we now know that, when (1.1.5) holds, the operator L which appears in Kolmogorov's forward equation has the form

$$L\varphi(x) = \frac{1}{2}\sum_{i,j=1}^N a_{ij}(x)\partial_{x_i}\partial_{x_j}\varphi(x) + \sum_{i=1}^N b_i(x)\partial_{x_i}\varphi(x), \tag{1.1.8}$$

where $a(x) = \big((a_{ij}(x))\big)_{1\le i,j\le N}$ is a non-negative definite, symmetric matrix for each $x \in \mathbb{R}^N$. In the probability literature, a is called the diffusion coefficient and b is called the drift coefficient.

1.1.2. Solving Kolmogorov's Forward Equation: In this section we will prove the following general existence result for solutions to Kolmogorov's forward equation. Throughout we will use the notation $\langle\varphi,\mu\rangle$ to denote the integral $\int\varphi\,d\mu$ of the function φ with respect to the measure μ.

THEOREM 1.1.9. *Let $a : \mathbb{R}^N \longrightarrow \mathrm{Hom}(\mathbb{R}^N;\mathbb{R}^N)$ and $b : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be continuous functions with the properties that $a(x) = \big((a_{ij}(x))\big)_{1\le i,j\le N}$ is symmetric and non-negative definite for each $x \in \mathbb{R}^N$ and*

$$\Lambda \equiv \sup_{x\in\mathbb{R}^N} \frac{\mathrm{Trace}\big(a(x)\big) + 2\big(x, b(x)\big)^+_{\mathbb{R}^N}}{1+|x|^2} < \infty. \tag{1.1.10}$$

Then, for each $\nu \in \mathbf{M}_1(\mathbb{R}^N)$, there is a continuous $t \in [0,\infty) \longmapsto \mu(t) \in \mathbf{M}_1(\mathbb{R}^N)$ which satisfies

$$\langle\varphi,\mu(t)\rangle - \langle\varphi,\nu\rangle = \int_0^t \langle L\varphi,\mu(\tau)\rangle\,d\tau, \tag{1.1.11}$$

for all $\varphi \in C^2_c(\mathbb{R}^N;\mathbb{C})$, where L is the operator in (1.1.8). Moreover,

$$\int (1+|y|^2)\,\mu(t,dy) \le e^{\Lambda t}\int(1+|x|^2)\,\nu(dx), \quad t\ge 0. \tag{1.1.12}$$

Before giving the proof, it may be helpful to review the analogous result for ordinary differential equations. Indeed, when applied to the case when $a = 0$, our proof is exactly the same as the usual one there. Namely, in that case, except for the initial condition, there should be no randomness, and so, when we remove the randomness from the initial condition by taking $\nu = \delta_x$, we expect that $\mu_t = \delta_{X(t)}$, where $t \in [0,\infty) \longmapsto X(t) \in \mathbb{R}^N$ satisfies

$$\varphi\big(X(t)\big) - \varphi(x) = \int_0^t \big(b(X(\tau)), \nabla\varphi(X(\tau)\big)_{\mathbb{R}^N}\,d\tau.$$

Equivalently, $t \rightsquigarrow X(t)$ is an integral curve of the vector field b starting at x. That is,

$$X(t) = x + \int_0^t b\big(X(\tau)\big)\,d\tau.$$

To show that such an integral curve exists, one can use the following Euler approximation scheme. For each $n \geq 0$, define $t \rightsquigarrow X_n(t)$ so that $X_n(0) = x$ and

$$X_n(t) = X_n(m2^{-n}) + (t - m2^{-n})b\big(X(m2^{-n})\big) \quad \text{for } m2^{-n} < t \leq (m+1)2^{-n}.$$

Clearly,

$$X_n(t) = x + \int_0^t b\big(X_n([\tau]_n)\,d\tau,$$

where[6] $[\tau]_n = 2^{-n}[2^n\tau]$ is the largest dyadic number $m2^{-n}$ dominated by τ. Hence, if we can show that $\{X_n : n \geq 0\}$ is relatively compact in the space $C\big([0,\infty);\mathbb{R}^N\big)$, with the topology of uniform convergence on compacts, then we can take $t \rightsquigarrow X(t)$ to be any limit of the X_n's.

To simplify matters, assume for the moment that b is bounded. In that case, it is clear that $|X_n(t) - X_n(s)| \leq \|b\|_{\mathrm{u}}|t-s|$, and so the Ascoli–Arzela Theorem guarantees the required compactness. To remove the boundedness assumption, choose a $\psi \in C_{\mathrm{c}}^\infty\big(B(0,2);[0,1]\big)$ so that $\psi = 1$ on $\overline{B(0,1)}$ and, for each $k \geq 1$, replace b by b_k, where $b_k(x) = \psi(k^{-1}x)$. Next, let $t \rightsquigarrow X_k(t)$ be an integral curve of b_k starting at x, and observe that

$$\frac{d}{dt}|X_k(t)|^2 = 2\big(X_k(t), b_k(X_k(t))\big)_{\mathbb{R}^N} \leq \Lambda\big(1 + |X_k(t)|^2\big),$$

from which it is an easy step to the conclusion that $|X_k(t)| \leq R(T) \equiv (1 + |x|^2)e^{t\Lambda}$. But this means that, for each $T > 0$, $|X_k(t) - X_k(s)| \leq C(T)|t-s|$ for s, $t \in [0,T]$, where $C(T)$ is the maximum value of $|b|$ on the closed ball of radius $R(T)$ centered at the origin, and so we again can invoke the Ascoli–Arzela Theorem to see that $\{X_k : k \geq 1\}$ is relatively compact and therefore has a limit which is an integral curve of b.

In view of the preceding, it should be clear that our first task is to find an appropriate replacement for the Ascoli–Arzela Theorem. The one which we will choose is the following variant of Lévy's Continuity Theorem (cf. Exercise 3.1.19 in [53]), which states that if $\{\mu_n : n \geq 0\} \subseteq \mathbf{M}_1(\mathbb{R}^N)$ and $\hat{\mu}_n$ is the characteristic function (i.e., the Fourier transform) of μ_n, then $\mu = \lim_{n\to\infty} \mu_n$ exists in $\mathbf{M}_1(\mathbb{R}^N)$ if and only if $\hat{\mu}_n(\xi)$ converges for each ξ and uniformly in a neighborhood of 0, in which case $\mu_n \longrightarrow \mu$ in $\mathbf{M}_1(\mathbb{R}^N)$ where $\hat{\mu}(\xi) = \lim_{n\to\infty} \hat{\mu}_n(\xi)$.

In the following, and elsewhere, we say that $\{\varphi_k : k \geq 1\} \subseteq C_{\mathrm{b}}(\mathbb{R}^N;\mathbb{C})$ *converges to* φ *in* $C_{\mathrm{b}}(\mathbb{R}^N;\mathbb{C})$ and write $\varphi_k \longrightarrow \varphi$ in $C_{\mathrm{b}}(\mathbb{R}^N;\mathbb{C})$ if $\sup_k \|\varphi_k\|_{\mathrm{u}}$

[6] We use $[\tau]$ to denote the integer part of a number $\tau \in \mathbb{R}$

$< \infty$ and $\varphi_k(x) \longrightarrow \varphi(x)$ uniformly for x in compact subsets of $\mathbb{R}^N$. Also, we say that $\{\mu_k : k \geq 1\} \subseteq C\big(\mathbb{R}^M; \mathbf{M}_1(\mathbb{R}^N)\big)$ *converges to* μ *in* $C\big(\mathbb{R}^M; \mathbf{M}_1(\mathbb{R}^N)\big)$ and write $\mu_k \longrightarrow \mu$ in $C\big(\mathbb{R}^M; \mathbf{M}_1(\mathbb{R}^N)\big)$ if, for each $\varphi \in C_{\mathrm{b}}(\mathbb{R}^N; \mathbb{C})$, $\langle \varphi, \mu_k(z)\rangle \longrightarrow \langle \varphi, \mu(z)\rangle$ uniformly for z in compact subsets of $\mathbb{R}^M$.

THEOREM 1.1.13. *If $\mu_k \longrightarrow \mu$ in $C\big(\mathbb{R}^M; \mathbf{M}_1(\mathbb{R}^N)\big)$, then*

$$\langle \varphi_k, \mu_k(z_k)\rangle \longrightarrow \langle \varphi, \mu(z)\rangle$$

whenever $z_k \longrightarrow z$ in $\mathbb{R}^M$ and $\varphi_k \longrightarrow \varphi$ in $C_{\mathrm{b}}(\mathbb{R}^N; \mathbb{C})$. Moreover, if $\{\mu_n : n \geq 0\} \subseteq C\big(\mathbb{R}^M; \mathbf{M}_1(\mathbb{R}^N)\big)$ and $f_n(z, \xi) = \widehat{\mu_n(z)}(\xi)$, then $\{\mu_n : n \geq 0\}$ is relatively compact in $C\big(\mathbb{R}^M; \mathbf{M}_1(\mathbb{R}^N)\big)$ if $\{f_n : n \geq 0\}$ is equicontinuous at each $(z, \xi) \in \mathbb{R}^M \times \mathbb{R}^N$. In particular, $\{\mu_n : n \geq 0\}$ is relatively compact if, for each $\xi \in \mathbb{R}^N$, $\{f_n(\,\cdot\,, \xi) : n \geq 0\}$ is equicontinuous at each $z \in \mathbb{R}^N$ and, for each $r \in (0, \infty)$,

$$\lim_{R \to \infty} \sup_{n \geq 0} \sup_{|z| \leq r} \mu_n\big(z, \mathbb{R}^N \setminus B(0, R)\big) = 0.$$

PROOF: To prove the first assertion, suppose $\mu_k \longrightarrow \mu$ in $C\big(\mathbb{R}^M; \mathbf{M}_1(\mathbb{R}^N)\big)$, $z_k \longrightarrow z$ in $\mathbb{R}^M$, and $\varphi_k \longrightarrow \varphi$ in $C_{\mathrm{b}}(\mathbb{R}^N; \mathbb{C})$. Then, for every $R > 0$,

$$\begin{aligned}
&\overline{\lim_{k \to \infty}} \big|\langle \varphi_k, \mu_k(z_k)\rangle - \langle \varphi, \mu(z)\rangle\big| \\
&\leq \overline{\lim_{k \to \infty}} \Big(\big|\langle \varphi - \varphi_k, \mu_k(z_k)\rangle\big| + \big|\langle \varphi, \mu_k(z_k)\rangle - \langle \varphi, \mu(z_k)\big| \\
&\qquad\qquad + \big|\langle \varphi, \mu(z_k)\rangle - \langle \varphi, \mu(z)\rangle\big|\Big) \\
&\leq \overline{\lim_{k \to \infty}} \sup_{y \in B(0,R)} |\varphi_k(y) - \varphi(y)| + \sup_k \|\varphi_k\|_{\mathrm{u}} \overline{\lim_{k \to \infty}} \mu_k\big(z_k, B(0,R)\complement\big) \\
&\leq \sup_k \|\varphi_k\|_{\mathrm{u}} \mu\big(z, B(0,R)\complement\big)
\end{aligned}$$

since $\overline{\lim}_{k \to \infty} \mu_k(z_k, F) \leq \mu(z, F)$ for any closed $F \subseteq \mathbb{R}^N$. Hence, the required conclusion follows after one lets $R \to \infty$.

Turning to the second assertion, apply the Arzela–Ascoli Theorem to produce an $f \in C_{\mathrm{b}}(\mathbb{R}^M \times \mathbb{R}^N; \mathbb{C})$ and a subsequence $\{n_k : k \geq 0\}$ such that $f_{n_k} \longrightarrow f$ uniformly on compacts. By Lévy's Continuity Theorem, there is, for each $z \in \mathbb{R}^M$, a $\mu(z) \in \mathbf{M}_1(\mathbb{R}^N)$ for which $f(z, \cdot\,) = \widehat{\mu(z)}$. Moreover, if $z_k \longrightarrow z$ in $\mathbb{R}^M$, then, because $f_{n_k}(z_k, \cdot\,) \longrightarrow f(z, \cdot\,)$ uniformly on compact subsets of $\mathbb{R}^N$, another application of Lévy's Theorem shows that $\mu_{n_k}(z_k) \longrightarrow \mu(z)$ in $\mathbf{M}_1(\mathbb{R}^N)$, and from this it is clear that $\mu_{n_k} \longrightarrow \mu$ in $C\big(\mathbb{R}^M; \mathbf{M}_1(\mathbb{R}^N)\big)$.

It remains to show that, under the conditions in the final assertion, $\{f_n : n \geq 0\}$ is equicontinuous at each (z, ξ). But, by assumption, for each

$\xi \in \mathbb{R}^N$, $\{f_n(\,\cdot\,,\xi) : n \geq 0\}$ is equicontinuous at every $z \in \mathbb{R}^M$. Thus, it suffices to show that if $\xi_k \longrightarrow \xi$ in $\mathbb{R}^N$, then, for each $r > 0$,

$$\lim_{k\to\infty} \sup_{n\geq 0} \sup_{|z|\leq r} \big|f_n(z,\xi_k) - f_n(z,\xi)\big| = 0.$$

To this end, note that, for any $R > 0$,

$$\big|f_n(z,\xi_k) - f_n(z,\xi)\big| \leq R|\xi_k - \xi| + 2\mu_n\big(z, B(0,R)\complement\big),$$

and therefore

$$\varlimsup_{k\to\infty} \sup_{n\geq 0} \sup_{|z|\leq r} \big|f_n(z,\xi_k) - f_n(z,\xi)\big| \leq 2 \sup_{n\geq 0} \sup_{|z|\leq r} \mu_n\big(z, B(0,R)\complement\big) \longrightarrow 0$$

as $R \to \infty$. □

Now that we have a suitable compactness criterion, the next step is to develop an Euler approximation scheme. To do so, we must decide what plays the role in $\mathbf{M}_1(\mathbb{R}^N)$ that linear translation plays in $\mathbb{R}^N$. A hint comes from the observation that if $t \rightsquigarrow X(t,x) = x + tb$ is a linear translation along the constant vector field b, then $X(s+t,x) = X(s,x) + X(t,0)$. Equivalently, $\delta_{X(s+t,x)} = \delta_x \star \delta_{X(s,0)} \star \delta_{X(t,0)}$, where "$\star$" denotes convolution. Thus, "linear translation" in $\mathbf{M}_1(\mathbb{R}^N)$ should be a path $t \in [0,\infty) \longmapsto \mu(t) \in \mathbf{M}_1(\mathbb{R}^N)$ given by $\mu(t) = \nu \star \lambda(t)$, where $t \rightsquigarrow \lambda(t)$ satisfies $\lambda(0) = \delta_0$ and $\lambda(s+t) = \lambda(s) \star \lambda(t)$. That is, in the terminology of classical probability theory, $\mu(t) = \nu \star \lambda(t)$, where $\lambda(t)$ is an *infinitely divisible flow*. Moreover, because L is local and therefore the associated process has continuous paths, the only infinitely divisible laws which can appear here must be Gaussian (cf. §§,III.3 and III.4 in [53]). With these hints, we now take $Q(t,x) \in \mathbf{M}_1(\mathbb{R}^N)$ to be the normal distribution with mean $x + tb(x)$ and covariance $ta(x)$. Equivalently, if

(1.1.14) $$\gamma(d\omega) \equiv (2\pi)^{-\frac{M}{2}} e^{-\frac{|\omega|^2}{2}}\, d\omega$$

is the standard normal distribution on $\mathbb{R}^M$ and $\sigma : \mathbb{R}^N \longrightarrow \mathrm{Hom}(\mathbb{R}^M;\mathbb{R}^N)$ is a square root[7] of a in the sense that $a(x) = \sigma(x)\sigma(x)^\top$, then $Q(t,x)$ is the distribution of $\omega \rightsquigarrow x + t^{\frac{1}{2}}\sigma(x)\omega + tb(x)$ under γ. To check that $Q(t,x)$ will play the role that $x + tb(x)$ played above, observe that if $\varphi \in C^2(\mathbb{R}^N;\mathbb{C})$ and φ together with its derivatives have at most exponential growth, then

(1.1.15)
$$\langle \varphi, Q(t,x)\rangle - \varphi(x) = \int_0^t \langle L^x\varphi, Q(\tau,x)\rangle\, d\tau,$$
$$\text{where } L^x\varphi(y) = \frac{1}{2}\sum_{i,j}^N a(x)\partial_{y_i}\partial_{y_j}\varphi(y) + \sum_{i=1}^N b_i(x)\partial_{y_i}\varphi(y).$$

[7] At the moment, it makes no difference which choice of square root one chooses. Thus, one might as well assume here that $\sigma(x) = a(x)^{\frac{1}{2}}$, the non-negative definite, symmetric square root $a(x)$. However, later on it will be useful to have kept our options open.

Cambridge University Press
978-0-521-88651-2 - Partial Differential Equations for Probabilists
Daniel W. Stroock
Excerpt
More information

To verify (1.1.15), simply note that

$$\frac{d}{dt}\langle\varphi, Q(t,x)\rangle = \frac{d}{dt}\int \varphi\big(x+\sigma(x)\omega + tb(x)\omega\big)\,\gamma_t(d\omega),$$

where $\gamma_t(\omega) = g(t,\omega)\,d\omega$ with $g(t,\omega) \equiv (2\pi t)^{-\frac{M}{2}} e^{-\frac{|\omega|^2}{2t}}$ is the normal distribution on $\mathbb{R}^M$ with mean 0 and covariance tI, use $\partial_t g(t,\omega) = \frac{1}{2}\Delta g(t,\omega)$, and integrate twice by parts to move the Δ off of g. As a consequence of either (1.1.15) or direct computation, we have

$$\int |y|^2\,Q(t,x,dy) = \big|x+tb(x)\big|^2 + t\mathrm{Trace}\big(a(x)\big). \tag{1.1.16}$$

Now, for each $n \geq 0$, define the Euler approximation $t \in [0,\infty) \longmapsto \mu_n(t) \in \mathbf{M}_1(\mathbb{R}^N)$ so that

$$\mu_n(0) = \nu \quad\text{and}\quad \mu_n(t) = \int Q\big(t - m2^{-n}, y\big)\,\mu_n(m2^{-n}, dy) \quad \text{for } m2^{-n} < t \leq (m+1)2^{-n}. \tag{1.1.17}$$

By (1.1.16), we know that

$$\int |y|^2\,\mu_n(t,dy) = \int\Big[\big|y + (t-m2^{-n})b(y)\big|^2 + (t-m2^{-n})\mathrm{Trace}\big(a(y)\big)\Big]\,\mu_n\big(m2^{-n},dy\big) \tag{1.1.18}$$

for $m2^{-n} \leq t \leq (m+1)2^{-n}$.

LEMMA 1.1.19. *Assume that*

$$\lambda \equiv \sup_{x\in\mathbb{R}^N} \frac{\mathrm{Trace}\big(a(x)\big) + 2|b(x)|^2}{1+|x|^2} < \infty. \tag{1.1.20}$$

Then

$$\sup_{n\geq 0}\int (1+|y|^2)\,\mu_n(t,dy) \leq e^{(1+\lambda)t}\int (1+|x|^2)\,\nu(dx). \tag{1.1.21}$$

In particular, if $\int |x|^2\,\nu(dx) < \infty$, then $\{\mu_n : n \geq 0\}$ is a relatively compact subset of $C\big([0,\infty);\mathbf{M}_1(\mathbb{R}^N)\big)$ with the topology of uniform convergence on compacts.

PROOF: Suppose that $m2^{-n} \leq t \leq (m+1)2^{-n}$, and set $\tau = t - m2^{-n}$. First note that

$$\begin{aligned}&\big|y+\tau b(y)\big|^2 + \tau\mathrm{Trace}\big(a(y)\big)\\ &= |y|^2 + 2\tau\big(y,b(y)\big)_{\mathbb{R}^N} + \tau^2|b(y)|^2 + \tau\mathrm{Trace}\big(a(y)\big)\\ &\leq |y|^2 + \tau\big[|y|^2 + 2|b(y)|^2 + \mathrm{Trace}\big(a(y)\big)\big] \leq |y|^2 + (1+\lambda)\tau(1+|y|^2),\end{aligned}$$

and therefore, by (1.1.18),

$$\int (1+|y|^2)\,\mu_n(t,dy) \le \big(1+(1+\lambda)\tau\big)\int (1+|y|^2)\,\mu_n\big(m2^{-n},dy\big).$$

Hence,

$$\begin{aligned}&\int (1+|y|^2)\,\mu_n(t,dy)\\ &\quad\le \big(1+(1+\lambda)2^{-n}\big)^m\big(1+(1+\lambda)\tau\big)\int(1+|y|^2)\,\nu(dy)\\ &\quad\le e^{(1+\lambda)t}\int(1+|x|^2)\,\nu(dx).\end{aligned}$$

Next, set $f_n(t,\xi) = [\widehat{\mu_n(t)}](\xi)$. Under the assumption that the second moment $S \equiv \int |x|^2\,\nu(dx) < \infty$, we want to show that $\{f_n : n \ge 0\}$ is equicontinuous at each $(t,\xi) \in [0,\infty)\times\mathbb{R}^N$. Since, by (1.1.21),

$$\mu_n\big(t, \overline{B(0,R)}\complement\big) \le S(1+R^2)^{-1}e^{(1+\lambda)t},$$

the last part of Theorem 1.1.13 says that it suffices to show that, for each $\xi \in \mathbb{R}^N$, $\{f_n(\,\cdot\,,\xi) : n \ge 0\}$ is equicontinuous at each $t \in [0,\infty)$. To this end, first observe that, for $m2^{-n} \le s < t \le (m+1)2^{-n}$,

$$\big|f_n(t,\xi) - f_n(s,\xi)\big| \le \int \big|[\widehat{Q(t,y)}](\xi) - [\widehat{Q(s,y)}](\xi)]\big|\,\mu_n\big(m2^{-n},dy\big)$$

and, by (1.1.15),

$$\begin{aligned}&\big|[\widehat{Q(t,y)}](\xi) - [\widehat{Q(s,y)}](\xi)]\big| = \left|\int_s^t \left(\int L^y e_\xi(y')\,Q(\tau,y,dy')\right)d\tau\right|\\ &\quad\le (t-s)\Big(\tfrac12\big(\xi,a(y)\xi\big)_{\mathbb{R}^N} + |\xi||b(y)|\Big) \le \tfrac12(1+\lambda)(1+|y|^2)(1+|\xi|^2)(t-s),\end{aligned}$$

where $e_\xi(y) \equiv e^{\sqrt{-1}\,\xi\cdot y}$. Hence, by (1.1.21),

$$\big|f_n(t,\xi) - f_n(s,\xi)\big| \le \frac{(1+\lambda)(1+|\xi|^2)}{2}e^{(1+\lambda)t}\int(1+|x|^2)\,\nu(dx)(t-s),$$

first for $s < t$ in the same dyadic interval and then for all $s < t$. □

With Lemma 1.1.19, we can now prove Theorem 1.1.9 under the assumptions that a and b are bounded and that $\int |x|^2\,\nu(dx) < \infty$. Indeed, because we know then that $\{\mu_n : n \ge 0\}$ is relatively compact in $C\big([0,\infty);\mathbf{M}_1(\mathbb{R}^N)\big)$, all that we have to do is show that every limit satisfies (1.1.11). For this purpose, first note that, by (1.1.15),

$$\langle\varphi,\mu_n(t)\rangle - \langle\varphi,\nu\rangle = \int_0^t\left(\int\langle L^y\varphi, Q(\tau-[\tau]_n,y)\rangle\mu_n([\tau]_n,dy)\right)d\tau$$

for any $\varphi \in C^2_{\mathrm{b}}(\mathbb{R}^N;\mathbb{C})$. Next, observe that, as $n \to \infty$,

$$\langle L^y\varphi, Q(\tau - [\tau]_n, y)\rangle \longrightarrow L\varphi(y)$$

boundedly and uniformly for (τ, y) in compacts. Hence, if $\mu_{n_k} \longrightarrow \mu$ in $C\big([0,\infty);\mathbf{M}_1(\mathbb{R}^N)\big)$, then, by Theorem 1.1.13,

$$\langle \varphi, \mu_{n_k}(t)\rangle \longrightarrow \langle \varphi, \mu(t)\rangle \quad \text{and}$$
$$\int_0^t \left(\int \langle L^y\varphi, Q(\tau - [\tau]_n, y)\rangle \mu_n([\tau]_n, dy)\right) d\tau \longrightarrow \int_0^t \langle L\varphi, \mu(\tau)\rangle\, d\tau.$$

Before moving on, we want to show that $\int |x|^2\,\nu(dx) < \infty$ implies that (1.1.11) continues to hold for $\varphi \in C^2(\mathbb{R}^N;\mathbb{C})$ with bounded second order derivatives. Indeed, from (1.1.21), we know that

$$(*) \qquad \int (1+|y|^2)\,\mu(t,dy) \le e^{(1+\lambda)t}\int (1+|y|^2)\,\nu(dy).$$

Now choose $\psi \in C^\infty_{\mathrm{c}}\big(\mathbb{R}^N;[0,1]\big)$ so that $\psi = 1$ on $\overline{B(0,1)}$ and $\psi = 0$ off of $B(0,2)$, define ψ_R by $\psi_R(y) = \psi(R^{-1}y)$ for $R \ge 1$, and set $\varphi_R = \psi_R\varphi$. Observe that[8]

$$\frac{|\varphi(y)|}{1+|y|^2} \vee \frac{|\nabla\varphi(y)|}{1+|y|} \vee \|\nabla^2\varphi(y)\|_{\mathrm{H.S.}}$$

is bounded independent of $y \in \mathbb{R}^N$, and therefore so is $\frac{|L\varphi(y)|}{1+|y|^2}$. Thus, by $(*)$, there is no problem about integrability of the expressions in (1.1.11). Moreover, because (1.1.11) holds for each φ_R, all that we have to do is check that

$$\langle \varphi, \mu(t)\rangle = \lim_{R\to\infty}\langle \varphi_R, \mu(t)\rangle$$
$$\int_0^t \langle L\varphi, \mu(\tau)\rangle\, d\tau = \lim_{R\to\infty}\int_0^t \langle L\varphi_R, \mu(\tau)\rangle\, d\tau.$$

The first of these is an immediate application of Lebesgue's Dominated Convergence Theorem. To prove the second, observe that

$$L\varphi_R(y) = \psi_R(y)L\varphi(y) + \big(\nabla\psi_R(y), a(y)\nabla\varphi\big)_{\mathbb{R}^N} + \varphi(y)L\psi_R(y).$$

Again the first term on the right causes no problem. To handle the other two terms, note that, because ψ_R is constant off of $\overline{B(0,2R)} \setminus B(0,R)$ and because $\nabla\psi_R(y) = R^{-1}\nabla\psi(R^{-1}y)$ while $\nabla^2\psi_R(y) = R^{-2}\nabla^2\psi(R^{-1}y)$, one

[8] We use $\nabla^2\varphi$ to denote the Hessian matrix of φ and $\|\sigma\|_{\mathrm{H.S.}}$ to denote the Hilbert–Schmidt norm $\sqrt{\sum_{ij}\sigma_{ij}^2}$ of σ.