

# CELEBRATIO MATHEMATICA

**S. R. S. Varadhan**

DANIEL W. STROOCK

**DIFFUSION THEORY**

2012



<http://celebratio.org/>

## DIFFUSION THEORY

DANIEL W. STROOCK

Published: July 2012

Given a second-order (possibly degenerate) elliptic differential operator

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

acting on  $C^2(\mathbb{R}; \mathbb{C})$  and a Borel probability measure  $\mu$  on  $\mathbb{R}^N$ , what does it mean for a Borel probability measure  $\mathbb{P}_\mu$  on  $C([0, \infty); \mathbb{R}^N)$  to be a diffusion process determined by  $L$  with initial value  $\mu$ ? When Varadhan and I asked ourselves this question in the mid-1960s, we found no answer which satisfied us.

It was not that there were no answers. Indeed, already in the 1930s Kolmogorov [1931] had mapped out a path which started from  $L$  and ended at a measure  $\mathbb{P}_\mu$ . Namely, he began by solving what is now called Kolmogorov's forward equation

$$(2) \quad \partial_t p(t, x, \cdot) = L^* p(t, x, \cdot) \quad \text{with } p(0, x, \cdot) = \delta_x,$$

where  $L^*$  is the formal adjoint of  $L$  given by

$$L^* \psi(y) = \frac{1}{2} \sum_{i,j=1}^N \partial_{y_i} \partial_{y_j} (a_{ij}(y) \psi(y)) - \sum_{i=1}^N \partial_{y_i} (b_i(y) \psi(y)).$$

He then showed that, under suitable regularity and non-degeneracy conditions on the coefficients  $a$  and  $b$ , there is one and only solution to (2) which is the density of a probability measure on  $\mathbb{R}^N$ . Further he showed that  $p(t, x, y)$  is a continuous function of  $(t, x, y) \in (0, \infty) \times \mathbb{R}^N \times \mathbb{R}^N$  and that it satisfies the Chapman–Kolmogorov equation

$$(3) \quad p(s+t, x, y) = \int p(t, \zeta, y) p(s, x, \zeta) d\zeta.$$

Knowing (3), he could check that, for any  $\mu \in \mathbf{M}(\mathbb{R}^N)$ , the family of probability measures given by

$$P_\mu((t_1, \dots, t_n); \Gamma) = \int_{\Gamma} \cdots \int_{\Gamma} \prod_{m=1}^n p(t_m - t_{m-1}, y_{m-1}, y_m) \mu(dy_0) dy_1 \cdots dy_n$$

for  $n \geq 0$ ,  $0 = t_0 < t_1 < \cdots < t_n$ , and  $\Gamma \in \mathcal{B}_{(\mathbb{R}^N)^{n+1}}$

is consistent, in the sense that if  $\{t'_1, \dots, t'_n\} \subset \{t_1, \dots, t_n\}$  then  $P_\mu((t'_1, \dots, t'_n); \cdot)$  is the marginal distribution of  $P_\mu((t_1, \dots, t_n); \cdot)$  on the coordinates corresponding to  $\{t'_1, \dots, t'_n\}$ . Having established

this consistency, Kolmogorov applied his Consistency Theorem to produce a measure  $\tilde{\mathbb{P}}_\mu$  on  $(\mathbb{R}^N)^{[0, \infty)}$  whose finite-dimensional marginals are given by

$$\tilde{\mathbb{P}}_\mu(\{\omega : (\omega(t_0), \dots, \omega(t_n)) \in \Gamma\}) = P_\mu((t_1, \dots, t_n); \Gamma).$$

Finally, using his continuity criterion, he showed that  $C([0, \infty); \mathbb{R}^N)$  has outer measure 1 under  $\tilde{\mathbb{P}}_\mu$  and therefore that the restriction of  $\tilde{\mathbb{P}}_\mu$  to  $C([0, \infty); \mathbb{R}^N)$  determines a unique Borel probability measure  $\mathbb{P}_\mu$  there.

The trail which Kolmogorov blazed was widened and smoothed by many people, including W. Feller [1937], J. Doob [1953], G. Hunt [1957; 1958], R. Blumenthal and R. Gettoor [1968], E. B. Dynkin [1965], and a host of others. What emerged from this line of research, especially after the work of Hunt, was an isomorphism between Markov processes and potential theory. However, none of these really addressed the question that was bothering Varadhan and me. Namely, from our point of view, the route mapped out by Kolmogorov was too circuitous: one started with  $L$ , constructed from it either a transition probability function or resolvent operator, and only then characterized the resulting Markov processes in terms of these ancillary objects. What we were looking for was a characterization of  $\mathbb{P}_\mu$  directly in terms of  $L$ , one from which it would be clear that, if (for each  $n \geq 1$ )  $\mathbb{P}_n$  is related to  $L_n$  with initial distribution  $\mu_n$ , and if  $\mu_n \rightarrow \mu$  weakly and  $L_n \varphi \rightarrow L\varphi$  for smooth  $\varphi$  with compact support, then  $\mathbb{P}_n$  tends weakly to  $\mathbb{P}_\mu$ . Obviously, the best way to achieve this goal is to phrase the characterization in terms of  $\mathbb{P}_\mu$ -integrals involving  $L\varphi$ .

Formulating such a characterization was relatively easy. Namely, no matter how one goes about its construction,  $\mathbb{P}_\mu$  should have this property:<sup>1</sup>

$$(4) \quad \mathbb{E}^{\mathbb{P}_\mu}[\varphi(\omega(t_2)) \mid \mathcal{B}_{t_1}] - \varphi(\omega(t_1)) = \mathbb{E}^{\mathbb{P}_\mu}\left[\int_{t_1}^{t_2} L\varphi(\omega(\tau)) d\tau \mid \mathcal{B}_{t_1}\right] \quad \text{for all } \varphi \in C_c^\infty(\mathbb{R}^N; \mathbb{C}),$$

where  $\mathcal{B}_t$  is the  $\sigma$ -algebra generated by  $\{\omega(\tau) : \tau \in [0, t]\}$ . More succinctly,

$$(5) \quad \left(\varphi(\omega(t)) - \int_0^t L\varphi(\omega(\tau)) d\tau, \mathcal{B}_t, \mathbb{P}_\mu\right) \quad \text{is a martingale for all } \varphi \in C_c^\infty(\mathbb{R}^N; \mathbb{C}).$$

Thus, we said that  $\mathbb{P}_\mu$  *solves the martingale problem for  $L$  with initial distribution  $\mu$*  if  $\mu$  is the  $\mathbb{P}_\mu$ -distribution of  $\omega \rightsquigarrow \omega(0)$  and (5) holds.

Having formulated the martingale problem, our first task was to prove that, in great generality, solutions exist. Of course, seeing as all roads lead to solutions to the martingale problem (in situations to which they applied), we could have borrowed results from earlier work (for example, work of either the Kolmogorov or Itô schools). However, not only would that have made no contribution to known existence results, it would have been very inefficient. At least when  $a$  and  $b$  are bounded and continuous, existence of solutions to the martingale problem is an easy application of compactness, completely analogous to the proof of existence of solutions to ordinary differential equations with bounded continuous coefficients. Our second task was to show that, under appropriate conditions, solutions are unique, and we knew that uniqueness posed a challenge of an entirely different order from the one posed by existence. Indeed, if

<sup>1</sup>: Note that, throughout, I will be somewhat casual about integrability questions. In the following, no integrability problems arise unless the coefficients are unbounded, in which case special considerations have to be taken into account.

all roads lead to solutions to the martingale problem, then uniqueness would prove that all roads must have the same destination.

Our approach to uniqueness depended on the assumptions being made about the coefficients  $a$  and  $b$ . One approach turned on the observation that any solution to the martingale problem is the distribution of a solution to an Itô stochastic integral equation. When the coefficients are reasonably smooth, this observation allowed us to prove uniqueness for the martingale problem as a consequence of uniqueness for the associated stochastic integral equation. However, when the coefficients are not smooth, Itô's theory cannot cope. Specifically, because his theory ignores ellipticity properties, it is difficult to see how one would go about bolstering it with powerful results from the vast and beautiful PDE literature.

The key to overcoming the problem just raised is to realize that uniqueness for the martingale problem is a purely distributional question, whereas the uniqueness found in Itô's theory is a path-by-path statement about a certain transformation (the "Itô map") of Brownian paths. Thus, to take advantage of the PDE existence theory, we abandoned Itô and adopted an easy duality argument which shows that uniqueness for the martingale problem requires only a sufficiently good existence statement about the initial-value problem for the backward equation  $\partial_t u = Lu$ . Namely, if  $u$  is a classical solution<sup>2</sup> to the backward equation with initial data  $f$ , then one can show that, for any solution  $\mathbb{P}$  to the martingale problem for  $L$ ,

$$\left( u((t_2 - t)^+, \omega(t)), \mathcal{B}_t, \mathbb{P} \right)$$

is a martingale and therefore that, for all  $0 \leq t_1 < t_2$ ,

$$\mathbb{E}^{\mathbb{P}}[f(\omega(t_2)) \mid \mathcal{B}_{t_1}] = u(t_2 - t_1, \omega(t_1)).$$

Hence, if such classical solutions exist for enough of the  $f$ , then the conditional distribution of  $\omega \rightsquigarrow \omega(t_2)$  under  $\mathbb{P}$  given  $\mathcal{B}_{t_1}$  is uniquely determined for all  $0 \leq t_1 < t_2$ . From this, together with the condition that the  $\mathbb{P}$ -distribution of  $\omega \rightsquigarrow \omega(0)$  is  $\mu$ , it follows that there is only one solution to the martingale problem. This argument works without a hitch when, for example,  $a$  and  $b$  are bounded and uniformly Hölder continuous and  $a$  is uniformly elliptic (that is,  $a \geq \varepsilon I$  for some  $\varepsilon > 0$ ), since in that case PDE theorists had long ago provided the necessary existence statement about the backward equation.

When uniform Hölder continuity is replaced by simple uniform continuity, the existence theory for the backward equation is less satisfactory. To be precise, although solutions continue to exist, they are no longer classical. Instead, their second derivatives exist only in  $L^p(\mathbb{R}^N; \mathbb{C})$  for  $p \in (1, \infty)$ . As a consequence, the preceding duality argument breaks down unless one can show a priori that every solution to the martingale problem is sufficiently regular to tolerate limit procedures involving  $L^p(\mathbb{R}^N)$ -convergence of the integrands. Without question, the proof of such an a priori regularity result was the single most intricate part of our program. In broad outline, our proof went as follows: First, we developed machinery which allowed us to reduce everything to the case when  $b = 0$  and  $a$  is an arbitrarily small perturbation of the identity. Second, we again used the connection with the Itô stochastic integral equations to show that any solution  $\mathbb{P}$  can be approximated by  $\mathbb{P}_n$ 's which, conditioned on  $\mathcal{B}_{m/n}$ , are Gaussian during the time interval  $[m/n, (m+1)/n]$ , and therefore have the required regularity properties. Finally, as an application of the Calderon–Zygmund theory of singular integral operators, for each  $p \in$

<sup>2</sup>: By "classical solution," I mean that the solution is twice continuously differentiable in space, and once in time.

(1,  $\infty$ ) we were able to show that, when the perturbation is sufficiently small, the  $L^p(\mathbb{R}^N)$ -regularity properties of the approximating  $\mathbb{P}_n$ 's can be controlled independent of  $n$  and are therefore enjoyed by  $\mathbb{P}$  itself.

The steps outlined above led us to a proof of the following theorem, which should be considered the centerpiece of [Stroock and Varadhan 1969a; 1969b].

**Theorem.** *Let  $L$  be given by (1), where  $a$  and  $b$  are bounded,  $a$  is continuous,  $b$  is Borel measurable, and  $a(x)$  is symmetric and strictly positive definite for each  $x \in \mathbb{R}^N$ . Then, for each Borel probability measure  $\mu$  on  $\mathbb{R}^N$ , there is precisely one solution to the martingale for  $L$  with initial distribution  $\mu$ . Moreover, if  $\delta_x$  is the unit mass at  $x$  and  $\mathbb{P}_x = \mathbb{P}_{\delta_x}$ , then  $x \rightsquigarrow \mathbb{P}_x$  is weakly continuous and  $\{\mathbb{P}_x : x \in \mathbb{R}^N\}$  is a strong Markov family.*

### Concluding remarks

There are several additional comments which may be helpful.

1. Although, with 20/20 hindsight, the characterization given by (4) is the most obvious one, it is not the one which we chose initially. Instead, because we were influenced at the time both by Henry McKean's treatment [2005] of Itô's theory of stochastic integral equations and by a possible analogy with characteristic functions, we chose to characterize  $\mathbb{P}_\mu$  by saying that

$$(6) \quad \left( \exp \left[ \left( \xi, \omega(t) - \int_0^t b(\omega(\tau)) d\tau \right)_{\mathbb{R}}^N - \frac{1}{2} \int_0^t (\xi, a(\tau)\xi)_{\mathbb{R}}^N d\tau \right], \mathcal{B}_t, \mathbb{P}_\mu \right)$$

is a martingale for all  $\xi \in \mathbb{C}^N$ .

The equivalence of (4) and (6) is a quite easy application of elementary Fourier analysis and the observation that, aside from integrability issues,

$$\left( M(t) V(t) - \int_0^t M(\tau) dV(\tau), \mathcal{B}_t, \mathbb{P} \right)$$

is a martingale when  $(M(t), \mathcal{B}_t, \mathbb{P})$  is a continuous martingale and  $V(t)$  is a continuous,  $\{\mathcal{B}_t : t \geq 0\}$ -adapted process of locally bounded variation.

2. The preceding theorem is not the one on which we had originally hoped to settle. When we started, we had hoped for a statement in which  $a$  needed only be Borel measurable and locally uniformly elliptic. Using special, heavily dimension-dependent arguments, we did prove such a result when  $N \in \{1, 2\}$ , and we somewhat naively assumed that it was only the limits on our own technical powers which prevented us from doing so in higher dimensions. However, a few years ago Nikolai Nadirashvili [1997] gave a beautiful example that showed that the limits on our powers were not responsible for our inability to go to higher dimensions. Namely, he produced a uniformly elliptic, bounded, Borel measurable  $a$  on  $\mathbb{R}^3$  for which the associated martingale problem admits more than one solution.

3. In retrospect, there are many aspects of this work which look rather crude and unnecessarily cumbersome. For instance, around the same time, first P. A. Meyer [1966] and then N. Kunita and S. Watanabe [1967] gave a much better and cleaner treatment of the stochastic integral calculus which we developed.

More significant, had we been aware of the spectacular work being done by N. Krylov, in particular the estimate in [Krylov 1966], we could have vastly simplified the regularity proof just described.

4. In a later work, we extended the martingale problem approach to cover diffusions with boundary conditions. The resulting paper [1971] has, for good reasons, been much less influential: it is nearly unreadable. On the other hand, it contains information which, so far as I know, is available nowhere else. In particular, I do not think the theorems there dealing with approximations by discrete Markov chains have ever been improved.

5. Varadhan and my joint work on diffusion theory culminated in our papers [1972b] and [1972a] on degenerate diffusions. The first of these contains our initial version of “the support theorem,” which has been given several other proofs over the years. Among other things, the second paper contains an extension of “the support theorem” to cover situations in which the diffusion matrix  $a$  is smooth but does not admit a smooth square root. It seems that most probabilists are not particularly impressed by this extension, but we were very proud of it at the time.

## References

- [Blumenthal and Gettoor 1968] R. M. Blumenthal and R. K. Gettoor, *Markov processes and potential theory*, Pure and Applied Mathematics **29**, Academic Press, New York, 1968. [MR 41 #9348](#) [Zbl 0169.49204](#)
- [Doob 1953] J. L. Doob, *Stochastic processes*, John Wiley & Sons, New York, 1953. [MR 15,445b](#) [Zbl 0053.26802](#)
- [Dynkin 1965] E. B. Dynkin, *Markov processes*, Die Grundlehren der Mathematischen Wissenschaften **121–122**, Academic Press, Publishers, New York, 1965. In two volumes. [MR 33 #1887](#)
- [Feller 1937] W. Feller, “Zur Theorie der stochastischen Prozesse”, *Math. Ann.* **113**:1 (1937), 113–160. [MR 1513083](#) [Zbl 0014.22201](#)
- [Hunt 1957] G. A. Hunt, “Markoff processes and potentials. I, II”, *Illinois J. Math.* **1** (1957), 44–93, 316–369. [MR 19,951g](#)
- [Hunt 1958] G. A. Hunt, “Markoff processes and potentials. III”, *Illinois J. Math.* **2** (1958), 151–213. [MR 21 #5824](#)
- [Kolmogoroff 1931] A. Kolmogoroff, “Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung”, *Math. Ann.* **104**:1 (1931), 415–458. [MR 1512678](#) [Zbl 0001.14902](#)
- [Krylov 1966] N. V. Krylov, “On quasi-diffusion processes”, *Teor. Veroyatnost. i Primenen.* **11** (1966), 424–443. [MR 34 #860](#) [Zbl 0202.47502](#)
- [Kunita and Watanabe 1967] H. Kunita and S. Watanabe, “On square integrable martingales”, *Nagoya Math. J.* **30** (1967), 209–245. [MR 36 #945](#) [Zbl 0167.46602](#)
- [McKean 2005] H. P. McKean, *Stochastic integrals*, Reprint ed., AMS Chelsea Publishing, Providence, RI, 2005. [MR 2006d:60003](#) [Zbl 1101.60002](#)
- [Meyer 1966] P.-A. Meyer, *Probability and potentials*, Blaisdell Publishing Co. Ginn and Co., Waltham, MA-Toronto, Ont.-London, 1966. [MR 34 #5119](#) [Zbl 0138.10401](#)
- [Nadirashvili 1997] N. Nadirashvili, “Nonuniqueness in the martingale problem and the Dirichlet problem for uniformly elliptic operators”, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **24**:3 (1997), 537–549. [MR 99b:35042](#) [Zbl 0907.35039](#)
- [Stroock and Varadhan 1969a] D. W. Stroock and S. R. S. Varadhan, “Diffusion processes with continuous coefficients. I”, *Comm. Pure Appl. Math.* **22** (1969), 345–400. [MR 40 #6641](#) [Zbl 0167.43903](#)
- [Stroock and Varadhan 1969b] D. W. Stroock and S. R. S. Varadhan, “Diffusion processes with continuous coefficients. II”, *Comm. Pure Appl. Math.* **22** (1969), 479–530. [MR 40 #8130](#) [Zbl 0175.44802](#)
- [Stroock and Varadhan 1971] D. W. Stroock and S. R. S. Varadhan, “Diffusion processes with boundary conditions”, *Comm. Pure Appl. Math.* **24** (1971), 147–225. [MR 43 #2774](#) [Zbl 0227.76131](#)
- [Stroock and Varadhan 1972a] D. W. Stroock and S. R. S. Varadhan, “On degenerate elliptic-parabolic operators of second order and their associated diffusions”, *Comm. Pure Appl. Math.* **25** (1972), 651–713. [MR 52 #8651](#) [Zbl 0344.35041](#)

[Stroock and Varadhan 1972b] D. W. Stroock and S. R. S. Varadhan, “On the support of diffusion processes with applications to the strong maximum principle”, pp. 333–359 in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (Univ. California, Berkeley, CA, 1970/1971), vol. III: Probability theory, Univ. California Press, Berkeley, CA, 1972. [MR 53 #4259](#) [Zbl 0255.60056](#)

Published July 2012.

*Daniel W. Stroock*

[dws@math.mit.edu](mailto:dws@math.mit.edu)

*Department of Mathematics  
Massachusetts Institute of Technology  
Cambridge, MA 02139  
USA*