DIFFUSIONS AND THE MARTINGALE PROBLEM OF STROOCK AND VARADHAN

S. Ramasubramanian
Stat. Math. Unit
Indian Statistical Institute
8th Mile, Mysore Road
Bangalore - 560 059

1 Introduction

In an article on Varadhan, his long time friend Varadarajan [Vj] writes "... he [Varadhan] used to tease me occassionally that I could not regard myself as a probabilist till I did something on Markov processes". Not surprisingly, Varadhan's research interests have centred around Markov processes in general, and diffusions in particular. Theory of diffusion processes, drawing inspiration from physics, having deep connections with the theory of partial differential equations, and serving as models / approximations in diverse fields like genetics, engineering, economics / finance, has been an outstanding success of probabilistic methodology; it has in turn played a key role in shaping developments in probability theory.

The paradigm of this theory is the Brownian motion, representing the movement of diffusing particle in a "homogeneous isotropic medium in the absence of external forces" - postulated from considerations of molecular / kinetic theory of heat by Einstein in 1905, and independently by Bachelier as a model for the fluctuations in stock market prices five years *before* Einstein. A rigorous mathematical basis for Brownian motion was given by Wiener about 20 years later.

Passing through the hands of Kolmogorov, Feller, Levy, Ito, Doob, ... it was left to S.R.S. Varadhan and his colleague Daniel Stroock to develop arguably the most comprehensive theory of diffusion processes in the form of their *martingale problem*; [SV1, SV2]. The technical generality achieved by the Stroock-Varadhan theory is just part of the story. The more profound contribution is the different way of looking at Markov processes making the appropriate family of probability measures on function spaces the centre-piece. In fact "probabilistic methods in differential equations" has come of age only since the advent of martingale problem. The deep impact the papers [SV1, SV2] have had on probability theory and related fields has earned for the authors the prestigious "Steele prize of the American Mathematical Society for Seminal Contribution to Research" in 1996, more than 25 years after the appearance of the path breaking papers.

In Section 2 we indicate how diffusion processes were looked at before martingale problem arrived on the scene. Section 3 gives a brief description of the formulation and solution of the martingale problem; the reader is urged to go through the extremely readable original papers [SV1, SV2] of the masters; a synopsis of the results is given in [SV3], and an exhaustive account also in their book [SV7]. We outline briefly in Section 4 some of the developments triggered by the martingale problem; we look at just four offshoots which have become fields of study in their own right. Our presentation is certainly influenced by some of the excellent monographs like [B], [EK], [KS], [RW] and of course [SV7].

In addition to martingale problem, Varadhan has helped create two other theories. *Large deviation principles* was formulated in a series of papers with Donsker, and *Hydrodynamic limits* was given shape in collaboration with various coworkers / students. These too have stemmed from the theory of Markov processes, with some of Varadhan's earlier papers on diffusions being the forerunners; they also have ramifications in other disciplines like statistics, physics, partial differential equations, etc. Accounts of these can be found in [V] and [DP] respectively.

To end this section we just recall a few basic definitions and notations.

Let $(S, \mathcal{F}, P)$ be a probability space. Let Y be an integrable random variable and $\mathcal{A} \subseteq \mathcal{F}$ a sub-σ-algebra. The *conditional expectation of Y given $\mathcal{A}$*, denoted by $E(Y|\mathcal{A})$, is the $\mathcal{A}$-measurable random variable such that

$$\int_A Y \, dP = \int_A E(Y|\mathcal{A}) dP \quad \text{for all } A \in \mathcal{A}.$$

If $Y = 1_F$ for some $F \in \mathcal{F}$ then we write $P(F|\mathcal{A}) = E(1_F|\mathcal{A})$ and call it the *conditional probability of the event F given $\mathcal{A}$.*

An $I\!R^d$-valued random variable V and a σ-algebra $\mathcal{A}$ are said to be *independent* if $V^{-1}(B)$ and A are independent events for any $A \in \mathcal{A}$ and Borel set $B \subseteq I\!R^d$.

A sequence $\{Q_n\}$ of probability measures on a complete separable metric space (S, d) is said to *converge weakly* to a probability measure Q, denoted $Q_n \Rightarrow Q$, if

$$\lim_{n\to\infty} \int_S f(y) dQ_n(y) = \int_S f(y) dQ(y)$$

for any bounded continuous function f on S. The space of probability measures on S becomes a complete separable metric space under this notion of convergence, which coincides with weak-$*$ convergence of functional analysis. See [EK], [P] for details.

For a family $\{V_\alpha\}$ of random variables, $\sigma\{V_\alpha\}$ denotes the smallest σ-algebra making each V_α measurable.

For a function $f : [0, \infty) \times I\!R^d \to I\!R$ we denote $f_0(t, z) = \frac{\partial f}{\partial t}(t, z), f_i(t, z) = \frac{\partial f}{\partial z_i}(t, z), f_{ij}(t, z) = \frac{\partial^2 f}{\partial z_i \partial z_j}(t, z), 1 \leq i, j \leq d; C_b^{1,2}([0, \infty) \times I\!R^d) = \{f : [0, \infty) \times I\!R^d \to I\!R : f, f_0, f_i, f_{ij}, 1 \leq i, j \leq d$ are bounded and continuous$\}$. $C_b^2(I\!R^d)$ is defined analogously; $C_0^\infty(I\!R^d) = \{f : I\!R^d \to I\!R :$ derivative of any order of f is continuous with compact support$\}$ Let $S_d(I\!R) = \{d \times d$ real symmetric matrices$\}$. For $\sigma \in S_d(I\!R), \sigma^*$ denotes its adjoint / transpose. For $b \in I\!R^d, |b|^2 = \sum_{i=1}^{d} b_i^2$. For a matrix $a, |a|^2 = \sum_{i,j} a_{ij}^2$. Subscript x in $L_x, \nabla_x, L_x^{(r)}$ denotes that the operators $L, \nabla, L^{(r)}$ act on functions of x.

2 Before Martingale Problem

The concentration function of a diffusing species in a dilute solution is known to be governed by the Fokker-Planck equation; see [Rf]. Larger the value of the concentration function at (t, z), greater is the probability that a diffusing particle is found in a neighbourhood of z at time t, and conversely. That is, probabilistic modelling of the diffusion phenomenon is quite natural and has proved very effective. We review briefly two ways of modelling diffusions prior to the formulation of the martingale problem.

2.1 Classical approach

The d-dimensional *Fokker-Planck equation (or the forward Kolmogorov equation)* is given by

$$\begin{aligned}\frac{\partial v(t,z)}{\partial t} &= \left(L_z^{(t)}\right)^* v(t,z) \\ &\equiv \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial z_i \partial z_j}(a_{ij}(t,z)v(t,z)) - \sum_{i=1}^{d}\frac{\partial}{\partial z_i}(b_i(t,z)v(t,z)), \quad (2.1)\end{aligned}$$

for $t > 0, z \in I\!R^d$, where $a(\cdot,\cdot) = ((a_{ij}(\cdot,\cdot)))_{1\le i,j\le d}, b(\cdot,\cdot) = (b_1(\cdot,\cdot),\dots,b_d(\cdot,\cdot))$ are respectively the *infinitesimal diffusion and drift* parameters; $b(t,z)$ is the instantaneous velocity of diffusing particle at (t,z) due to "fluid flow" and external forces, and $a(t,z)$ is a measure of instantaneous "variance" associated with flucutations caused by random collisions with solvent / fluid molecules.

Note that a_{ij}, b_i must be sufficiently smooth even to make sense of (2.1). To overcome this restriction one can look at the "formal adjoint" of (2.1), viz. consider the so called *backward Kolmogorov equation*

$$\begin{aligned}&\frac{\partial u}{\partial s}(s,x) + L_x^{(s)}u(s,x) \\ \equiv\ &\frac{\partial u}{\partial s}(s,x) + \frac{1}{2}\sum_{i,j=1}^{d}a_{ij}(s,x)\frac{\partial^2 u}{\partial x_i \partial x_j}(s,x) + \sum_{i=1}^{d}b_i(s,x)\frac{\partial u}{\partial x_i}(s,x) \\ &= 0 \qquad (2.2)\end{aligned}$$

for $0 \le s < t, x \in I\!R^d$, where $t > 0$ is fixed. Instead of the initial condition at 0 one considers terminal condition at t. When the coefficients a, b are sufficiently smooth study of (2.1) and (2.2) turn out to be equivalent. Clearly requirements on a, b are less stringent in (2.2). In this case we have

Theorem 2.1 *Suppose $a_{ij}(\cdot,\cdot) = a_{ji}(\cdot,\cdot), 1 \le i,j \le d$ and there exist numbers $0 < C_1 < C_2 < \infty, 0 < \alpha \le 1, \lambda > 0$ such that*

$$\textit{(boundedness)}\ |a(r,y)| + |b(r,y)| \le C_1 \qquad (2.3)$$

$$\begin{aligned}\textit{(Holder continuity)}\ |a(r,y) - a(r',y')| + |b(r,y) - b(r',y')| \\ \le C_2[|y-y'|^\alpha + |r-r'|^\alpha] \qquad (2.4)\end{aligned}$$

$$\textit{(nondegeneracy)}\ \langle\theta, a(r,y)\theta\rangle \geq \lambda|\theta|^2 \tag{2.5}$$

for all $y, y' \in I\!R^d, \theta \in I\!R^d, r, r' \in [0,\infty)$. *Then there is a unique positive continuous function* $(s,x;t,z) \mapsto p(s,x;t,z), 0 \leq s < t, x, z \in I\!R^d$ *such that the function* u *given by*

$$u(s,x) = \int\limits_{I\!R^d} f(z)p(s,x;t,z)dz, s < t, x \in I\!R^d$$

solves the backward Kolmogorov equation (2.2) with terminal value $u(t,\cdot) = f(\cdot)$ *for any bounded continuous function* f *on* $I\!R^d$. *Also* $z \mapsto p(s,x;t,z)$ *is a strictly positive probability density function. Moreover the Chapman-Kolmogorov equation holds, that is, for* $t_1 < t_2 < t_3, x, z \in I\!R^d$

$$p(t_1,x;t_3,z) = \int\limits_{I\!R^d} p(t_1,x;t_2,y)p(t_2,y;t_3,z)dy. \tag{2.6}$$

$\square$

For a proof of the above result from the theory of partial differential equations (PDE theory, for short) see [Fm], [SI]. The function p is known as the *fundamental solution* of (2.2). One can interpret $p(s,x;t,z)dz$ as the probability of finding the diffusing particle in a dz neighbourhood of z at time t if the particle starts from x at time s. It can be shown that p is in $C^{1,2}([0,\infty)\times I\!R^d)$ as a function of (s,x). As p can be regarded as a "transition probability density function" which is also well behaved, using probabilistic techniques it can be shown that there exists a continuous Markov process $\{Y(t) : t \geq 0\}$ with p as the transition probability density function; that is, there is a probability space $(S, \mathcal{F}, P)$ and a family $\{Y(t) : t \geq 0\}$ of $I\!R^d$-valued random variables on S such that $t \mapsto Y(t,w)$ is continuous for $P - a \cdot a \cdot w \in S$, and for any $0 \leq s < t, A \in \mathcal{B}(I\!R^d)$

$$\begin{aligned} P(Y(t) \in A|\mathcal{F}_s) &= P(Y(t) \in A|\sigma\{Y_s\}) \\ &= \int\limits_A p(s,Y(s);t,y)dy, \end{aligned} \tag{2.7}$$

where $\mathcal{F}_s = \sigma\{Y(r) : 0 \leq r \leq s\}$; the σ-algebra $\mathcal{F}_s$ can be considered as the "repository" of all information on $Y(\cdot)$ upto (and including) time s. The l.h.s. of (2.7) denotes "conditional probability of $Y(t) \in A$ given $\mathcal{F}_s$".

Equation (2.7) can be interpreted as: Based on the entire history upto time s, if one wants to predict the probability that the diffusing particle is in A at time t, the information concerning the past $[0, s)$ is irrelevant, only the knowledge of the present "s" matters and the best estimate is given by r.h.s. of (2.7). See [Dy]. It is also interpreted as "the conditional independence of the past and the future given the present"; see Chapter 2 of [KS]. The process $\{Y(t) : t \geq 0\}$ is called the *diffusion process corresponding to (2.2).*

When $a_{ij}(\cdot, \cdot) \equiv \delta_{ij}, b \equiv 0$, the Fokker-Planck equation is the well-known *heat (or diffusion) equation*; (in this case (2.2) is sometimes called backward heat equation). The corresponding diffusion is the celebrated *Brownian motion* process in a homogeneous isotropic medium; this models the movement of a diffusing particle purely due to the very large number of random collisions with "solvent" molecules and the consequent fluctuations. In this case the fundamental solution is the *heat (Gaussian) kernel*

$$(s, x; t, z) \mapsto \left(\frac{1}{2\pi(t-s)} \right)^{d/2} \exp \left\{ -\frac{|z-x|^2}{2(t-s)} \right\}. \tag{2.8}$$

For any measurable function f on $I\!R^d, 0 \leq s < t$ set

$$(T_{s,t}f)(x) = \int_{I\!R^d} f(z)p(s, x; t, z)dz = E[f(Y(t))|Y(s) = x] \tag{2.9}$$

whenever the r.h.s. makes sense; of course $T_{s,s}f = f$. Using the Chapman-Kolmogorov equation (2.6) it follows that $T_{t_2,t_3} \circ T_{t_1,t_2} = T_{t_1,t_3}, t_1 \leq t_2 \leq t_3$ on $L^\infty(I\!R^d)$ or $L^2(I\!R^d)$. Properties of p and integration by parts yield

Theorem 2.2 *Let a, b be as in Theorem 2.1; let $\{Y(t)\}$ be the associated diffusion in the above sense. For any $f \in C_b^{1,2}([0,\infty) \times I\!R^d), 0 \leq s < t$*

$$\begin{aligned} T_{s,t}f(t,x) - f(s,x) &= \int_{I\!R^d} f(t,z)p(s,x;t,z)dz - f(s,x) \\ &= \int_s^t \int_{I\!R^d} \left(f_0(r,z) + L_z^{(r)} f(r,z) \right) p(s,x;r,z)dz\, dr \end{aligned} \tag{2.10}$$

or equivalently

$$E[f(t, Y(t))|Y(s) = x] - f(s, x)$$

$$= \quad E\left[\int_s^t \{f_0(r,Y(r)) + L_z^{(r)} f(r,Y(r))\}dr | Y(s) = x\right] \quad (2.11)$$

where as in (2.2)

$$L_z^{(r)} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(r,z)\frac{\partial^2}{\partial z_i \partial z_j} + \sum_{i=1}^{d} b_i(r,z)\frac{\partial}{\partial z_i}. \quad (2.12)$$

□

Note that the operator $\left(\frac{\partial}{\partial r} + L_z^{(r)}\right)$ in (2.10) or (2.11) is just the formal adjoint of the operator $-\frac{\partial}{\partial t} + \left(L_z^{(t)}\right)^*$ of the Fokker-Planck equation.

In case a, b are independent of the time variable then $p(s,x;t,z) = p(0,x;t-s,z)$ and $\{T_t \equiv T_{0,t} = T_{s,s+t} : t \geq 0\}$ forms a semigroup of operators. This approach (also called the semigroup approach) was pioneered by Kolmogorov and Feller.

2.2 Stochastic differential equations

Since Brownian motion represents diffusion purely due to fluctuations, Paul Levy had suggested that a diffusion process could be represented as a *stochastic differential equation (s.d.e.)*

$$dZ(t) = \sigma(t,Z(t))dB(t) + b(t,Z(t))dt \quad (2.13)$$

where $dB(t)$ can be taken as an "infinitesimal increment" of a d-dimensional Brownian motion, $\sigma(\cdot,\cdot)$ representing "infinitesimal dispersion" and $b(\cdot,\cdot)$ as before "infinitesimal drift"; $(\sigma(\cdot,\cdot)\sigma^*(\cdot,\cdot))$ will correspond to infinitesimal diffusion $a(\cdot,\cdot)$ encountered earlier. The problem here is, of course, to give a meaning to $dB(t)$ in (2.13), since it is well known that almost all sample paths of the Brownian motion process have infinite variation over any finite interval [KS]. This problem was circumvented in an ingenious way by Kiyosi Ito in the 40's. He defined *stochastic (or Ito) integrals* of the form $\int h(s)dB(s)$ and interpreted (2.13) as a *stochastic integral equation.*

Let $\{B(t)\}$ be a Brownian motion on $(S, \mathcal{F}, P)$. Let $\{\mathcal{F}_t : t \geq 0\}$ be a *filtration* such that for $s \leq t, B(t) - B(s)$ is independent of $\mathcal{F}_s$; here filtration means: each $\mathcal{F}_t$ is a sub σ-algebra of $\mathcal{F}$ and $t_1 \leq t_2 \Rightarrow \mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}$. Let h be a jointly measurable function such that $\omega \mapsto h(t,\omega)$ is $\mathcal{F}_t$-measurable for all t; that is, h is $\{\mathcal{F}_t\}$-*adapted.* If h is an *elementary function* of the form

$$h(t,\omega) = \sum_j e_j(\omega) I_{[t_j,t_{j+1})}(t)$$

where $0 = t_0 < t_1 < \ldots < t_n < \ldots, e_j$ being $\mathcal{F}_{t_j}$-measurable then

$$\left(\int_0^t h(s) dB(s) \right)(\omega) = \sum_j e_j(\omega) \left[B(t_{j+1} \wedge t, \omega) - B(t_j \wedge t, \omega) \right] \qquad (2.14)$$

is the natural definition. If in addition $E \int_0^t |h(s)|^2 ds < \infty$ for all t, then independence of $B(r) - B(s)$ and $\mathcal{F}_s$ for $r > s$ ensures that $E \left| \int_0^t h(s) dB(s) \right|^2 = E \int_0^t |h(s)|^2 ds$. Because of this isometry one can take limits of stochastic integrals of the form $\int_0^t h(s) dB(s)$ in $L^2(S, \mathcal{F}, P)$. Thus $\int_0^t h(s) dB(s)$ can be defined for any $\{\mathcal{F}_t\}$-adapted function such that $E \int_0^t |h(s)|^2 ds < \infty$ for all t. Of course, h can be scalar, vector or matrix valued in a suitable way; also the class of functions for which stochastic integrals are meaningful can be enlarged considerably.

For any one-dimensional Brownian motion $\{B_1(t)\}$ it is easy to prove that $\sum_j |B_1(t_{j+1} \wedge t) - B_1(t_j \wedge t)|^2 \to t$ in mean square as $\max_j |t_{j+1} - t_j| \to 0$. This, together with the fact that we have taken the left end value e_j in each of the summand in (2.14) results in the following change of variables formula for stochastic calculus, quite different from the conventional one.

Theorem 2.3 *(Ito's formula) Let $\{B(t)\}$ be an $\{\mathcal{F}_t\}$-adapted d- dimensional Brownian motion. Let $\beta(\cdot), h(\cdot)$ be respectively vector, matrix valued $\{\mathcal{F}_t\}$-adapted processes such that*

$$U(t) = \int_0^t h(s) dB(s) + \int_0^t \beta(s) ds, t \geq 0 \qquad (2.15)$$

is a well defined process. Let $f \in C_b^{1,2}([0,\infty) \times I\!R^d)$. *Then*

$$\begin{aligned} f(t,U(t)) - f(s,U(s)) &= \int_s^t \sum_{i,j=1}^d f_i(r,U(r))h_{ij}(r)dB_j(r) \\ &+ \int_s^t f_0(r,U(r))dr + \int_s^t \sum_{i=1}^d f_i(r,U(r))\beta_i(r)dr \\ &+ \frac{1}{2}\int_s^t \sum_{i,j=1}^d f_{ij}(r,U(r))(hh^*)_{ij}(r)dr \end{aligned} \qquad (2.16)$$

holds P- *a.e. for any* $0 \le s < t$. □

Note that the last term involving second order differentials in (2.16) is the extra term making all the difference! Once a meaning is given for (2.13), using Ito's formula, the isometry and a Picard iteration, the following result can be proved

Theorem 2.4 *(K. Ito) Let* $\{B(t)\}$ *be a d-dimensional Brownian motion on* $(S,\mathcal{F},P)$. *Let* ξ *be an* $I\!R^d$*-valued square integrable random variable independent of* $B(\cdot)$. *Let* $\mathcal{F}_t = \sigma\{\{\xi, B(r) : 0 \le r \le t\} \cup \{P-\mathit{null\ sets}\}\}$, $t \ge 0$. *Let* $\sigma : [0,\infty) \times I\!R^d \to S_d(I\!R), b : [0,\infty) \times I\!R^d \to I\!R^d$ *be functions such that there are constants* $C_1, C_2 > 0$ *satisfying*

$$\text{(Linear growth) } |\sigma(t,z)|^2 + |b(t,z)|^2 \le C_1(1+|z|^2) \qquad (2.17)$$

(Lipschitz condition)

$$|\sigma(t,z) - \sigma(t,z')| + |b(t,z) - b(t,z')| \le C_2|z-z'| \qquad (2.18)$$

for all $z, z' \in I\!R^d, t \ge 0$. *Then there exists a unique* $\{\mathcal{F}_t\}$*-adapted continuous process* $\{Z(t) : t \ge 0\}$ *such that*

$$Z(t) = \xi + \int_0^t b(s,Z(s))ds + \int_0^t \sigma(s,Z(s))dB(s). \qquad (2.19)$$

Also $Z(t)$ *is square integrable. Moreover* $\{Z(t)\}$ *is a Markov process, that is for* $s < t$

$$P(Z(t) \in A|\mathcal{F}_s) = P(Z(t) \in A|\sigma\{Z(s)\}). \qquad (2.20)$$

In fact the process is strong Markov in the sense that s, t in (2.20) can even be replaced by certain random times called stopping times. $\square$

Note: A measurable function $\tau : S \to [0, \infty) \cup \{\infty\}$ is called a *stopping time* w.r.t. $\{\mathcal{F}_t\}$ if $\{\tau \leq t\} \in \mathcal{F}_t$ for all t. For a continuous stochastic process, first time of entering / hitting an open or a closed set is an example of a stopping time.

The process $\{Z(t)\}$ is called the *strong solution to the s.d.e. (2.13).*

Proposition 2.5 *Let σ, b be as in Theorem 2.4. Put $a(t, z) = \sigma(t, z)$ $\sigma^*(t, z)$. Suppose a, b satisfy the hypotheses of Theorem 2.1. Then $\{Z(t)\}$ with $\xi = x$ coincides in law with the process $\{Y(t)\}$ of Section 2.1 with $Y(0) = x$, in the sense that the finite dimensional distributions are the same.* $\square$

In an analogous way one can consider s.d.e.'s (2.13) for $t > s$ with initial value $Z(s) = \xi$, that is

$$Z(t) = \xi + \int_s^t b(r, Z(r))dr + \int_s^t \sigma(r, Z(r))dB(r), t \geq s. \qquad (2.21)$$

Remark 2.6 Assume the hypotheses of Theorem 2.4. Apply Theorem 2.3 with $Z(\cdot)$ in the place of $U(\cdot)$. Note that the operator occurring in the terms involving dr integrals on r.h.s. of (2.16) is precisely the same as $\frac{\partial}{\partial r} + L_x^{(r)}$ occurring in (2.10), (2.11). Also the stochastic integral in (2.16) has zero expectation. Thus under the hypotheses of Theorem 2.4, (2.11) holds with $Y(\cdot)$ replaced by $Z(\cdot)$. $\square$

The discussion above indicates that $\{Z(t)\}$ can be called the *diffusion process corresponding to the infinitesimal dispersion $\sigma(\cdot, \cdot)$ and infinitesimal drift* $b(\cdot, \cdot)$. Under the conditions of Theorem 2.4 this approach gives a pathwise description of the diffusion in terms of the stochastic equation (2.19).

Remark 2.7 Since fluctuations play a major role in the diffusion phenomenon, a probabilistic description will be desirable and very useful. The classical approach is analytic, and the probabilistic aspects are hidden. On the other hand Ito's method is elegantly probabilistic, does not require the nondegeneracy assumption (2.5), and when it works it does perfectly; however the Lipschitz condition (2.18) is a severe limitation; also when only $a(\cdot,\cdot), b(\cdot,\cdot)$ are given it requires a Lipschitz continuous square root $\sigma(\cdot,\cdot)$ of $a(\cdot,\cdot)$. $\square$

For extensions / nuances concerning stochastic integrals and stochastic differential equations an interested reader may consult [KI], [IW], [RW], [KS].

3 Martingale Problem

3.1 Statement of problem

We begin this section by recalling the definition of a martingale. Let $(S, \mathcal{F}, P)$ be a probability space, $\{\mathcal{F}_t : t \geq 0\}$ a filtration and $\{M(t) : t \geq 0\}$ a stochastic process; $\{M(t) : t \geq 0\}$ is called a *P-martingale* w.r.t. $\{\mathcal{F}_t\}$ if $M(t)$ is $\mathcal{F}_t$-measurable and integrable for each $t \geq 0$, and for $s \leq t$ $E(M(t)|\mathcal{F}_s) = M(s)$ a.s. Like many objects in probability theory, martingale has its humble origins in gambling; basic results in martingale theory are due to Doob and later Meyer; martingale theory is now a versatile tool for probabilists and analysts; see [EK], [RW], [IW], [KS] and the references therein. In case $\{M(t)\}$ is square integrable and a martingale, then the increments of $\{M(t)\}$ are uncorrelated.

We have already encountered an important class of martingales, viz. stochastic integrals $\int_0^t h(s)dB(s)$. Therefore Remark 2.6 can now be rephrased as

Proposition 3.1 *Assume the hypotheses of Theorem 2.4. Then for any* $f \in C_b^{1,2}([0,\infty) \times \mathbb{R}^d), s \geq 0,$

$$f(t, Z(t)) - f(s, Z(s)) - \int_s^t \left(\frac{\partial}{\partial r} + L_x^{(r)} \right) f(r, Z(r))dr, \quad t \geq s$$

is a P-martingale w.r.t. $\{\mathcal{F}_t : t \geq s\}$. $\square$

A similar result is true for the process $\{Y(t)\}$ of Section 2.1 also. Actually (2.11) occurring in Theorem 2.2 or Remark 2.6 is just an integrated version of Proposition 3.1.

For greater clarity later on, we will now reformulate the above. Let $\Omega \equiv C([0,\infty) : I\!R^d) = \{w : [0,\infty) \to I\!R^d : w \text{ continuous}\}$; endowed with the topology of uniform convergence on compacta, it is a complete separable metric space. Let $X(t,w) = w(t), t \geq 0, w \in \Omega$ denote the t-th coordinate projection. For $0 \leq s \leq t$ let $\mathcal{B}_t^s = \sigma\{X(r) : s \leq r \leq t\}$. Let $\{Z(t) : t \geq s\}$ be the solution of the s.d.e. (2.21) with $Z(s) \equiv \xi \equiv x \in I\!R^d$. Since $t \mapsto Z(t,\omega)$ is continuous for $\omega \in S$, it induces a probability measure on $(\Omega, \mathcal{B}_\infty^s)$; denote the measure by $P_{s,x}$. Then Proposition 3.1 means that for any $f \in C_b^{1,2}([0,\infty) \times I\!R^d)$

$$\begin{aligned} f(t,X(t)) - f(s,x) \quad &- \quad \int_s^t \left(\frac{\partial}{\partial r} + L_x^{(r)}\right) f(r,X(r))dr \\ &= \quad P(s,x)- \text{ martingale} \qquad\qquad (3.1) \end{aligned}$$

w.r.t. $\{\mathcal{B}_t^s : t \geq s\}$. It can be shown that (3.1) holds also when $P_{s,x}$ denotes the distribution of the process $\{Y(t) : t \geq s\}$ starting at $Y(s) = x$ encountered in Section 2.1. So $\{X(t) : t \geq s\}$ may be viewed upon as a diffusion process with drift $b(\cdot,\cdot)$, infinitesimal diffusion $a(\cdot,\cdot)$ under $P_{s,x}$.

A natural question is : Can the above class of martingales given by (3.1) characterize the diffusion? Such a characterization of Brownian motion in terms of martingales, due to Levy, is known, viz. if $\{B(t)\}$ is a continuous one-dimensional stochastic process such that $\{B(t)\}$ and $\{B^2(t) - t\}$ are martingales, then $\{B(t)\}$ is a Brownian motion.

For Gaussian (or normal) random variables, independence and uncorrelatedness are equivalent. From (2.13) it is seen that a diffusion process behaves "locally" like a Gaussian random variable. As observed earlier Markov property is a sort of "conditional independence", and martingale property is akin to uncorrelatedness. So it is not unnatural that a characterization of diffusions is sought in terms of martingales. Of course, this oversimplified heuristic has the benefit of hindsight!

Since the family $\{P_{s,x} : s \geq 0, x \in I\!R^d\}$ of probability measures contains all the relevant information about the diffusion, this family is the focus of attention in the Stroock-Varadhan theory. We now state the martingale problem of Stroock and Varadhan.

Martingale problem: Let $\Omega, \{\mathcal{B}_t^s\}, \{X(t)\}$ be as above. Let $a : [0,\infty) \times \mathbb{R}^d \to S_d(\mathbb{R}), b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ be measurable functions. Set

$$L^{(r)} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(r,\cdot)\frac{\partial^2}{\partial z_i \partial z_j} + \sum_{i=1}^{d} b_i(r,\cdot)\frac{\partial}{\partial z_i}. \tag{3.2}$$

For $s \geq 0, x \in \mathbb{R}^d$ a *solution to the martingale problem for* $L^{(r)}$ *(or for* a, b*) starting from* (s,x) is a probability measure $P_{s,x}$ on $(\Omega, \mathcal{B}_\infty^s)$ such that
(i) $P_{s,x}(\{w : w(r) = x, 0 \leq r \leq s\}) = 1$, and
(ii) for each $f \in C_b^{1,2}([0,\infty) \times \mathbb{R}^d)$

$$\begin{aligned} f(t, X(t)) - f(s,x) &- \int_s^t \left(\frac{\partial}{\partial r} + L^{(r)}\right) f(r, X(r))dr, t \geq s \\ &= P_{s,x} - \text{ martingale.} \end{aligned} \tag{3.3}$$

If for each $s \geq 0, x \in \mathbb{R}^d$ there is exactly one solution to the above, then we say that the martingale problem for a, b is *well-posed*.

The obvious questions are those of existence, uniqueness and consequences in case of well-posedness. Anticipating that these questions have satisfactory answers, the family $\{P_{s,x} : s \geq 0, x \in \mathbb{R}^d\}$ of probability measures can be called the *diffusion with drift coefficient* $b(\cdot,\cdot)$ *and diffusion coefficient* $a(\cdot,\cdot)$. Under $\{P_{s,x}\}$ the collection of random variables $\{X(t)\}$ has all the probabilistic properties of the diffusion process corresponding to a, b.

Remark 3.2 Anyone familiar with basic probability theory will recall that all the relevant information concerning a random variable / random characteristic is contained in its distribution, which is a probability measure on $\mathbb{R}$ or $\mathbb{R}^d$. The point of view here is similar. And the Stroock-Varadhan characterization is analogous to characterizing measures on $\mathbb{R}$ or $\mathbb{R}^d$ by their Fourier / Laplace transforms. In fact condition (ii) of martingale problem is equivalent to either of the following:

(ii)$'$ for any $\theta \in \mathbb{R}^d$

$$\exp\left[\langle \theta, X(t) - X(s) - \int_s^t b(r, X(r))dr\rangle - \frac{1}{2}\int_s^t \langle \theta, a(r, X(r))\theta\rangle dr\right]$$

$$= \quad P_{s,x} - \text{martingale}, \tag{3.4}$$

(ii)" for any $\theta \in I\!R^d$

$$\exp\left[i\langle \theta, X(t) - X(s) - \int_s^t b(r, X(r))dr \rangle + \frac{1}{2}\int_s^t \langle \theta, a(r, X(r))\theta \rangle dr \right]$$
$$= \quad P_{s,x} - \text{martingale}. \tag{3.5}$$

See [SV1], [SV7]. A main advantage of focussing on probability measures on $C([0,\infty) : I\!R^d)$ is that one immediately has the following well developed notions / techniques at our disposal:
a) theory of weak convergence of probability measures on complete separable metric spaces,
b) theory of regular conditional probabilities,
c) localization arguments of martingale theory.

The main result of [SV1] can now be stated.

Theorem 3.3 *(Stroock-Varadhan) Let $b : [0,\infty) \times I\!R^d \to I\!R^d$ be a bounded measurable function; let $a : [0,\infty) \times I\!R^d \to S_d(I\!R)$ be a bounded continuous function such that the nondegeneracy (or strict ellipticity) condition (2.5) holds. Then for each $s \geq 0, x \in I\!R^d$ there is a unique probability measure $P_{s,x}$ on $(\Omega, \mathcal{B}^s_\infty)$ solving the martingale problem for a, b. Also $(s,x) \mapsto P_{s,x}$ is continuous in the sense of weak convergence of probability measures. Moreover, under $\{P_{s,x} : s \geq 0, x \in I\!R^d\}$ the process $\{X(t)\}$ is a continuous strong Markov process.* □

We shall briefly indicate the ideas involved in establishing existence and uniqueness when $b \equiv 0$; the case of general bounded measurable drift b can then be handled using a so called Cameron-Martin-Girsanov transformation of measures; see Section 6 of [SV1] or Section 6.4 of [SV7].

3.2 Existence

Let $a_0 \in S_d(I\!R)$ be strictly positive definite; let σ_0 be its strictly positive definite square root. Define

$$g(s, x; t, z)$$

$$= \left(\frac{1}{2\pi(t-s)}\right)^{d/2} \frac{1}{\sqrt{\det a_0}} \exp\left\{-\frac{1}{2(t-s)}\langle(z-x), \sigma_0^{-1}(z-x)\rangle\right\} \quad (3.6)$$

which is the d-dimensional Gaussian density with mean x and covariance matrix $(t-s)a_0$. It is well known that (see [KS]) for $s \geq 0, x \in I\!R^d$ there is unique probability measure $W^s_{a_0,x} \equiv W(s,x,a_0)$ on $(\Omega, \mathcal{B}^s_\infty)$ such that for $s < t_1 < t_2 < \ldots < t_n, A_1, \ldots, A_n \in \mathcal{B}(I\!R^d), n \geq 1$

$$\begin{aligned} & W^s_{a_0,x}(X(t_1) \in A_1, \ldots, X(t_n) \in A_n) \\ = & \int_{A_1} \ldots \int_{A_n} g(s,x;t_1,y_1)g(t_1,y_1;t_2,y_2)\ldots g(t_{n-1},y_{n-1};t_n,y_n)dy_n \ldots dy_1 \quad (3.7) \end{aligned}$$

$W^s_{a_0,x}$ is the *d-dimensional Wiener measure* starting at (s,x) and covariance $(t-s)a_0$; that is, it is the distribution of $\{\sigma_0(B(t)-B(s))+x : t \geq s\}$ where $\{B(t)\}$ is a standard d-dimensional Brownian motion. By Proposition 3.1, $W^s_{a_0,x}$ solves the martingale problem for $a(\cdot,\cdot) \equiv a_0, b \equiv 0$.

To proceed with our discussion we may consider a more general version of martingale problem. Let $a : [s,\infty) \times \Omega \to S_d(I\!R)$ be such that (i) a is $\{\mathcal{B}^s_t\}$-adapted, (ii) a is continuous and bounded, and (iii) $\exists \lambda > 0$ such that

$$\langle \theta, a(t,w)\theta\rangle \geq \lambda|\theta|^2, \forall\, \theta \in I\!R^d, t \geq s, w \in \Omega. \quad (3.8)$$

A probability measure P on $(\Omega, \mathcal{B}^s_\infty)$ is said to *solve the martingale problem for a starting at* (s,x) if $P(\{w : w(s) = x\}) = 1$, and for every $f \in C^{1,2}_b([0,\infty) \times I\!R^d)$

$$\begin{aligned} & f(t,X(t,w)) - f(s,X(s,w)) \\ & - \int_s^t \left\{ f_0(r,X(r,w)) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(r,w) f_{ij}(r,X(r,w)) \right\} dr \\ \equiv\ & f(t,w(t)) - f(s,x) - \int_s^t \left\{ f_0(r,w(r)) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(r,w) f_{ij}(r,w(r)) \right\} dr \\ =\ & P - \text{martingale} \end{aligned} \quad (3.9)$$

For notational convenience we take $s = 0$. A function $a : [0,\infty)\times\Omega \to S_d(I\!R)$ as above is said to be *n-simple* (where $n \geq 1$) if $a(t,w) = a\left(\frac{j}{n}, w\right), \frac{j}{n} \leq t < \frac{j+1}{n}, j = 0,1,2,\ldots$ We first indicate a method of obtaining a solution to the above martingale problem when a is n-simple satisfying (i)-(iii) above.

Let $x \in I\!R^d$. Let $w_0 \in \Omega$ be such that $w_0(0) = x$. Take $Q^{(0)} = W(0, x, a_0)$ where $a_0 = a(0, w_0)$. By canonical arguments one can obtain the probability measure $Q^{(1)}$ on $(\Omega, \mathcal{B}^0_\infty)$ such that $Q^{(1)} = Q^{(0)}$ on $\mathcal{B}^0_{1/n}$ and

$$Q^{(1)}(A|\mathcal{B}^0_{1/n})(w) = W\left(\frac{1}{n}, w\left(\frac{1}{n}\right), a\left(\frac{1}{n}, w\right)\right)(A)$$

for $A \in \mathcal{B}^{1/n}_\infty$ for $Q^{(0)} - a \cdot a \cdot w$; that is, the (regular) conditional probability of $Q^{(1)}$ and given $\mathcal{B}^0_{1/n}$, on $\mathcal{B}^{1/n}_\infty$, is $W\left(\frac{1}{n}, w\left(\frac{1}{n}\right), a\left(\frac{1}{n}, w\right)\right)$ a.s. $Q^{(0)}$. In a similar fashion, once $Q^{(k)}$ is defined, $Q^{(k+1)}$ is the probability measure on $(\Omega, \mathcal{B}^0_\infty)$ such that $Q^{(k+1)} = Q^{(k)}$ on $\mathcal{B}^0_{k/n}$, and

$$Q^{(k+1)}(A|\mathcal{B}^0_{k/n})(w) = W\left(\frac{k}{n}, w\left(\frac{k}{n}\right), a\left(\frac{k}{n}, w\right)\right)(A)$$

for $A \in \mathcal{B}^{k/n}_\infty$ for $Q^{(k)} - a \cdot a \cdot w$. As the $Q^{(k)}$ are defined in a consistent fashion, it can be shown that there is a unique probability mesure Q on $(\Omega, \mathcal{B}^0_\infty)$ such that $Q = Q^{(k)}$ on $\mathcal{B}^0_{k/n}$ and that Q is a solution to the martingale problem for a starting at $(0, x)$.

Next let $a : [0, \infty) \times \Omega \to S_d(I\!R)$ satisfy (i) - (iii) above. For $n = 1, 2, \ldots$ define $a_n(t, w) = a(\frac{j}{n}, w), \frac{j}{n} \leq t < \frac{j+1}{n}, j = 0, 1, 2, \ldots$ Let P_n denote the solution to the martingale problem for a_n starting at $(0, x)$ as obtained above. It can be shown that $\{P_n\}$ is relatively compact in the topology of weak convergence of probability measures; hence there is a limit point P which is a probability measure on $(\Omega, \mathcal{B}^0_\infty)$; P can be shown to be a solution to the martingale problem for a starting at $(0, x)$.

For $a : [0, \infty) \times I\!R^d \to S_d(I\!R)$ satisfying the hypotheses of Theorem 3.3 (and $b \equiv 0$), a solution can be found by the above procedure setting $a(t, w) = a(t, w(t)), t \geq 0, w \in \Omega$.

3.3 Uniqueness

Markov property, rather strong Markov property, is an important characteristic of diffusions. Yet the statement of martingale problem is curiously silent on that; in fact, no mention is made about any relationship among $P_{s,x}$ as s, x vary. It turns out that, if the martingale problem for a, b is well posed then the strong Markov property is an immediate consequence. This is because if P is a solution to the martingale problem for a, b starting at (s, x) then the regular conditional probability of P given $\mathcal{B}_t^s$ solves the martingale problem for a, b starting at $(t, X(t))$ for $t \geq s$; a similar result holds also for stopping times $\tau \geq s$. Well posedness also ensures simultaneously that $(s, x) \mapsto P_{s,x}$ is continuous. Thus a lot depends on having a satisfactory uniqueness result.

We consider only the case when a is independent of the time variable and $b \equiv 0$; since $a(\cdot)$ is continuous, by a localization argument it can be shown that the crux of the matter is to prove well-posedness when $a(\cdot)$ is *nearly* a constant matrix; and to simplify notation suppose the constant matrix is $((\delta_{ij}))$. In other words let $a : I\!R^d \to S_d(I\!R)$ be a measurable function satisfying the strict ellipticity condition (2.5) such that

$$|a_{ij}(x) - \delta_{ij}| \equiv |c_{ij}(x)| < \eta \tag{3.10}$$

for all $x \in I\!R^d, 1 \leq i, j \leq d$ where η is sufficiently small.

Let $P_x, \hat{P}_x$ be solutions to the martingale problem for a starting at $(0, x)$. As probability measures on Ω are determined by the 'finite dimensional marginal distributions', by inversion of Laplace transforms, etc. it can be shown that $P_x = \hat{P}_x$ if

$$E_x \int_0^\infty e^{-\lambda t} \varphi(X(t)) dt = \hat{E}_x \int_0^\infty e^{-\lambda t} \varphi(X(t)) dt = u(x), \text{say} \tag{3.11}$$

for all $\varphi \in C_0^\infty(I\!R^d), \lambda > 0$.

If the function $x \mapsto u(x)$ above is sufficiently smooth, then using analogues of Ito's formula one can show that

$$(\lambda - L)u(x) \equiv \lambda u(x) - \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) = \varphi(x), x \in I\!R^d \tag{3.12}$$

$$\lim_{|x| \to \infty} u(x) = 0. \tag{3.13}$$

The two preceding observations have been used as a handle on the problem of uniqueness by Stroock and Varadhan. Suppose there is a solution $u(\cdot)$, in an appropriate sense, to the problem (3.12)-(3.13) such that whenever P is a solution to the martingale problem for a starting at $(0,x)$ one has

$$u(x) = E^P\left(\int_0^\infty e^{-\lambda t}\varphi(X(t))dt\right)$$

where E^P denotes expectation w.r.t. P; then one may infer that there is atmost one solution to the martingale problem for a starting at $(0,x)$. *A novel feature here is that the problem of uniqueness has been linked to that of existence of a solution to (3.12)-(3.13).*

To appreciate what is needed to carry through the above heuristics, if P_y is a solution to the martingale problem for a starting at $(0,y)$ then

$$\int_0^\infty e^{-\lambda t}E_y\left[f(X(t)) - f(y) - \int_0^t Lf(X(r))dr\right]dt = 0$$

for any $\lambda > 0, f \in C_0^\infty(I\!R^d)$. Since $Lf(\cdot) = \frac{1}{2}\sum a_{ij}(\cdot)\frac{\partial^2 f}{\partial x_i \partial x_j}(\cdot)$ is bounded, an integration by parts now yields

$$R_\lambda^y(\lambda - L)f = f(y), y \in I\!R^d \tag{3.14}$$

where

$$R_\lambda^z h = \int_0^\infty e^{-\lambda t}E_z h(X(t))dt, \quad h \in C_0^\infty(I\!R^d). \tag{3.15}$$

(If things work out well $\{R_\lambda : \lambda > 0\}$ will be the resolvent of the diffusion, that is, Laplace transform of the corresponding Markov semigroup; however, without well-posedness we still do not have Markov property or even regularity of the map $z \mapsto P_z$.)

So "formally"

$$\begin{aligned} R_\lambda &= (\lambda - L)^{-1} = \left(\lambda - \frac{1}{2}\Delta + \frac{1}{2}\Delta - L\right)^{-1} \\ &= \left(\lambda - \frac{1}{2}\Delta\right)^{-1}\left[I - \left(L - \frac{1}{2}\Delta\right)\left(\lambda - \frac{1}{2}\Delta\right)^{-1}\right]^{-1} \\ &= U_\lambda\left[I - \left(\frac{1}{2}\sum_{i,j} c_{ij}(\cdot)D_iD_j\right)U_\lambda\right]^{-1} \end{aligned} \tag{3.16}$$

where Δ is the d-dimensional Laplacian, $D_i = \frac{\partial}{\partial x_i}, U_\lambda = (\lambda - \frac{1}{2}\Delta)^{-1}$ is the *Brownian resolvent* given by

$$\begin{aligned} U_\lambda h(x) &= \int_0^\infty e^{-\lambda t} E_x^W h(X(t)) dt \\ &= \int_0^\infty e^{-\lambda t} \int_{\mathbb{R}^d} h(z) \left(\frac{1}{2\pi t}\right)^{d/2} \exp\left\{-\frac{|z-x|^2}{2t}\right\} dz\ dt. \end{aligned} \tag{3.17}$$

(Here E_x^W denotes expectation w.r.t. the standard d-dimensional Wiener measure starting at x.)

Now it is not difficult to show that U_λ given by (3.17) maps $L^p(\mathbb{R}^d)$ into itself for any $p \geq 1$. If $[I - \frac{1}{2}(\sum c_{ij}(\cdot) D_i D_j) U_\lambda]^{-1}$ can be made meaningful as an operator on suitable $L^p(\mathbb{R}^d)$ then the formal expression (3.16) can be used to define R_λ as an operator on $L^p(\mathbb{R}^d)$. It turns out that $D_i D_j U_\lambda$ is a bounded linear operator on $L^p(\mathbb{R}^d)$ for $p > 1$; this can be proved by PDE methods using a result of Jones (which itself is an extension of the so called Calderon-Zygmund estimates), or by probabilistic methods via Burkholder-Davis-Gundy inequalities. Consequently choosing η small enough one can ensure that r.h.s. of (3.16) is meaningful, and hence for any $y \in \mathbb{R}^d, R_\lambda^y$ can be considered a bounded linear functional on $L^p(\mathbb{R}^d)$; also this functional is determined by (3.14).

Therefore for any $g \in L^p(\mathbb{R}^d)$, taking $f = U_\lambda g$, from (3.14) we get for any $y \in \mathbb{R}^d$

$$f(y) = R_\lambda^y g - R_\lambda^y \left(\frac{1}{2}\sum_{i,j} c_{ij}(\cdot) D_i D_j\right) U_\lambda g \tag{3.18}$$

with R_λ^y defined by (3.15) as a functional on $L^p(\mathbb{R}^d)$.

If $\tilde{P}_y$ is another solution to the martingale problem for a starting at $(0, y)$ one can get a functional $\tilde{R}_\lambda^y$ similarly and an analogue of (3.18) will hold. So by (3.18) and its analogue

$$\|R_\lambda^y - \tilde{R}_\lambda^y\| \leq \frac{\eta}{2}\left(\max_{i,j} \|D_i D_j U_\lambda\|\right) \|R_\lambda^y - \tilde{R}_\lambda^y\|.$$

Thus, if η is small enough then $R_\lambda^y = \tilde{R}_\lambda^y$, thereby forcing $P_y = \tilde{P}_y$.

Remark 3.4 Unlike the SDE formulation of Ito in Section 2.2, Stroock-Varadhan theory gives a direct connection between the operator (that is, the infinitesimal generator) L and the diffusion without bringing in the "nuisance parameter" Brownian motion. And the class of coefficients considered is much larger than in Section 2.1; boundedness hypothesis in Theorem 3.3 can be considerably relaxed and replaced by suitable conditions to ensure nonexplosion in finite time; see [SV7].

Remark 3.5 Well posedness of martingale problem and the techniques used to prove well posedness yield rich dividends. We have already mentioned strong Markov property and continuity of $(s, x) \mapsto P_{s,x}$. When the non-degeneracy condition (2.5) is satisfied these results and methods imply existence of transition probability density function (that is, density of the measure $P_{s,x}X(t)^{-1}$) as element of $L^p(I\!R^d)$ for suitable p, and estimates on it. A nice "invariance principle" (viz. convergence of Markov chains to diffusions) is also established in [SV2], as well as continuity of solution w.r.t. the coefficients; see also [SV7].

4 After Martingale Problem

As mentioned earlier the Stroock-Varadhan papers [SV1, SV2] won for their authors the prestigious Steele Prize for Seminal contribution to Research. This is just an indication of the diverse developments spawned by the Stroock-Varadhan theory of diffusions; quite a few have become fields of study in their own right. In this section we briefly touch upon some of these.

4.1 Weak solutions of SDE's

A systematic analysis of the connection between stochastic differential equations and martingale problems has lead to the notions of weak solutions of s.d.e.'s, pathwise uniqueness, and uniqueness in distribution.

In the case of strong solution of s.d.e.'s, as indicated in Theorem 2.4, in addition to the coefficients $b(\cdot, \cdot), \sigma(\cdot, \cdot)$ note that the probability space and the Brownian motion are also specified a priori; the solution $Z(\cdot)$ is adapted w.r.t. the filtration $\sigma\{Z(0), B(r) : 0 \leq r \leq t\}, t \geq 0$. On the other hand if only the coefficients $b(\cdot,\cdot), \sigma(\cdot,\cdot)$ are given and one is free to choose a probability space $(S, \mathcal{F}, P)$, a filtration $\{\mathcal{F}_t\}$, a Brownian motion $B(\cdot)$ adapted to $\{\mathcal{F}_t\}$ and an $\mathcal{F}_t$-adapted process $Z(\cdot)$ such that (2.19) holds, then such a $Z(\cdot)$ is called a *weak solution* of the s.d.e. (2.13); in this case $Z(\cdot)$ need not be adapted w.r.t. the filtration generated by the Brownian motion $B(\cdot)$. This notion leads naturally to two concepts of uniqueness viz. pathwise uniqueness which is a generalization of the uniqueness encountered in Theorem 2.4, and uniqueness in distribution similar to the one concerning martingale problem. It is a comforting result that, given fairly general $\sigma(\cdot,\cdot), b(\cdot,\cdot)$, existence of a weak solution is equivalent to the existence of a solution to the martingale problem for $a(\cdot,\cdot), b(\cdot,\cdot)$ where $a(\cdot,\cdot) = \sigma(\cdot,\cdot)\sigma^*(\cdot,\cdot)$. Similary well posedness of martingale problem is equivalent to well posedness of weak solution in the sense of distribution. An elucidation of the equivalence requires extension of the notion of weak solution to functional s.d.e.'s where the coefficients are $\{\mathcal{B}_t^s\}$-adapted functions on $[0,\infty) \times \Omega$ as in Section 3.2. This line of investigation was pioneered by Yamada and Watanabe. For details and proofs see Sections 5.3, 5.4 of [KS]. Weak solutions of stochastic differential equations play a central role in control theory; see [Bo].

4.2 Markov processes and martingale problem

Markov processes have traditionally been described by their generators and the corresponding evolution equations (as in Section 2.1) and resolvents. If A is the "generator" of a well behaved Markov process $\{Y(t)\}$ with domain $\mathcal{D}(A)$, it is known that

$$f(Y(t)) - f(Y(0)) - \int_0^t Af(Y(s))ds = \text{ a martingale} \tag{4.1}$$

for all $f \in \mathcal{D}(A)$. Often $\mathcal{D}(A)$ is difficult to describe completely. Martingale problem has indicated a sort of converse, and that too with possibly a subspace much smaller than $\mathcal{D}(A)$.

Let E be a complete separable metric space and $B(E)$ be the Banach space of bounded measurable functions on E. Let A be a linear operator on $B(E)$ with domain $\mathcal{D}$. An E-valued stochastic process $\{Y(t) : t \geq 0\}$ is a solution to the martingale problem for $(A, \mathcal{D})$ if (4.1) holds for all $f \in \mathcal{D}$. We say that uniqueness in distribution / law holds if any two solutions with the same initial distribution have the same finite dimensional distributions. In addition one may demand that the process $Y(\cdot)$ be continuous or have r.c.l.l. sample paths, that is, sample paths are in $D([0,\infty):E) \equiv \{E$-valued right continuous functions on $[0,\infty)$ having left limits$\}$. If uniqueness holds and the sample paths are r.c.l.l. then the process is strong Markov whose generator will be an appropriate closed extension of A. Thus martingale problem gives a way of characterizing Markov processes. This has proved to be a very powerful technqiue, especially for proving "convergence theorems" (like diffusion approximations / invariance principles) for Markov proesses; also the analysis is simpler and more transparent. Moreover infinite dimensional Markov processes can be dealt with more effectively.

It may be mentioned that the martingale problem of Stroock-Varadhan provides not only a new perspective on Markov processes, but also a tool for dealing with uniqueness. As in the case of diffusions, uniqueness problem is transformed into an existence problem for a "dual" situation. An excellent discussion of the above circle of ideas is give in [EK], see expecially Chapter 4.

A natural question in this context is the uniqueness in distribution in different classes of processes. For example, in the Stroock-Varadhan set up uniqueness in law is proved in the class of continuous processes. Since any process with continuous sample paths can also be considered a progressively measurable process, one can ask if uniqueness in law holds in the class of progressively measurable processes; see [BK] for a discussion of such issues and the advantages if such a uniqueness holds.

4.3 Reflected diffusions

Let G be a bounded smooth domain in $\mathbb{R}^d$. Consider the problem, for a fixed $t > 0$,

$$\begin{aligned} \frac{\partial u}{\partial s}(s,x) + L_x^{(s)} u(s,x) &= 0, s < t, x \in G && (4.2) \\ \langle \nabla_x u(s,x), \gamma(s,x) \rangle &= 0, s < t, x \in \partial G && (4.3) \\ \lim_{s \uparrow t} u(s,x) &= f(x), x \in \bar{G} && (4.4) \end{aligned}$$

where $\gamma(\cdot,\cdot)$ is a vector field on $[0,\infty) \times \partial G$ such that $\inf_{s,x} \langle \gamma(s,x), n(x) \rangle > 0$ with $n(x)$ denoting the unit inward normal at $x \in \partial G$. When G, a, b, γ are sufficiently smooth it is known that there is a unique fundamental solution $p^{\gamma}(s,x;t,z)$ to the above problem; see [SI]. As in Section 2.1 one can associate a strong Markov process with state space $\bar{G}$ such that p^{γ} is its transition probability density function. This process is called the *reflected diffusion* associated with (G, a, b, γ). Note that (4.3) is an insulation-type condition, viz. heat or diffusion cannot flow past $\bar{G}$. Thanks to a result of Levy, a stochastic differential equation formulation for reflected diffusion was given by Skorokhod. In turn, this, together with the martingale problem, led to the submartingale problem of Stroock and Varadhan [SV5]: Given a, b as in Section 3, and an $\mathbb{R}^d$-valued function γ on $[0,\infty) \times \partial G$, for $s \geq 0, x \in \bar{G}$, a probability measure $P_{s,x}$ on $C([0,\infty) : \bar{G})$ is said to solve the *submartingale problem* if for any $f \in C_b^{1,2}([0,\infty) \times \mathbb{R}^d)$ such that $\langle \gamma(t,y), \nabla_y f(t,y) \rangle \geq 0, t \geq s, y \in \partial G$ one has

$$\begin{aligned} & f(t, X(t)) - f(s,x) - \int_s^t \left(\frac{\partial}{\partial r} + L^{(r)} \right) f(r, X(r)) dr \\ = \quad & P_{s,x} - \text{ submartingale.} \end{aligned} \tag{4.5}$$

(An $\{\mathcal{F}_t\}$-adapted process $\{M(t)\}$ is called a P-submartingale if for $r \leq t$ one has $E(M(t)|\mathcal{F}_r) \geq M(r), P$ - a.s.) As in the case of martingale problem, submartingale problem helps in defining / characterizing reflected diffusions for a very large class of coefficients. Moreover this formulation can also be extended to define semimartingale reflecting Brownian motion in a domain with corners like a wedge / orthant; for such domains no analogue from PDE-theory is available; such processes are very useful in queueing networks and exhibit interesting properties; see [VW]. An expository account of reflected diffusions is available to Indian readers in [R2].

4.4 Martingale problem and PDE's

From the preceding sections it is quite clear that there is a strong connection between parabolic / elliptic differential equations and Markov processes. Before the advent of martingale problem the connection has essentially been one-way, viz. probability theory used to borrow results from p.d.e. theory. A notable exception is the purely probabilistic solution of the classical

Dirichlet problem (for the Laplacian) in terms of Brownian motion, due to Kakutani; this led Doob to develop his theory of martingales resulting in the independent probabilistic formulation of potential theory. See [PS] for a lucid exposition of this beautiful interplay of probability and analysis.

If a classical solution to the Cauchy problem (2.2) with terminal condition $u(t,x) = f(x)$ exists, using (2.11) or Proposition 3.1 it is easy to obtain stochastic representation of the solution in terms of the diffusion. Similarly for boundary value problem with Dirichlet boundary condition, if a classical solution exists, again using Proposition 3.1 and the time of hitting the boundary of the domain (which is a stopping time), the stochastic representation can easily be derived. However existence of a classical solution is guaranteed by PDE theory only under a minimal hypothesis of nondegeneracy and Holder continuity of the coefficients as in Theorem 2.1; see [Fm], [SI]. On the other hand, in the more general weak analytic formulation, solutions are in distributional sense, again regularity of solutions entails assumptions of smoothness on coefficients.

Martingale problem offers a way of obtaining solutions which are continuous functions for a large class of coefficients. For example, consider the problem (with $T > 0$ fixed),

$$\begin{aligned} \frac{\partial u}{\partial s}(s,x) + L_x^{(s)} u(s,x) &= g(s,x), s < T, x \in I\!R^d \qquad (4.6) \\ \lim_{s \uparrow T} u(s,x) &= f(x), x \in I\!R^d. \qquad (4.7) \end{aligned}$$

Suppose the martingale problem for a, b is well posed; let $\{P_{s,x} : s \geq 0, x \in I\!R^d\}$ be the solution (that is, the diffusion) corresponding to a, b. A function $u : [0,T] \times I\!R^d \to I\!R$ is said to be a *probabilistic solution* to (4.6)-(4.7) if (4.7) holds and for each $s < T, x \in I\!R^d$ one has

$$u(t, X(t)) - u(s,x) - \int_s^t g(r, X(r))dr = P_{s,x} - \text{ martingale.} \qquad (4.8)$$

When the martingale problem is well posed, the definition immediately offers also a candidate for solution, viz.

$$u(s,x) = E_{s,x}\left[f(X(T)) - \int_s^T g(r, X(r))dr \right], s \leq T, x \in I\!R^d. \qquad (4.9)$$

Using Markov property one can show that u given by (4.9) is indeed a probabilistic solution. Feller continuity of the diffusion implies u is bounded continuous whenever f, g are; in some cases even measurability of f, g would do. Also uniqueness in the class of bounded measurable functions is immediate.

To illustrate the point, let $\sigma, b, a = \sigma\sigma^*, L^{(r)}$ be as in Theorem 2.4, Remark 2.6. Suppose the uniform ellipticity / nondegeneracy condition (2.5) *does not* hold. As σ, b are Lipschitz continuous there is a strong solution to the s.d.e. (2.13); in this case the martingale problem formulation does not achieve more than what the s.d.e. formulation does for defining the diffusion. However, it can be shown that the family of probability measures induced by the strong solution of s.d.e. (2.13) is the only solution to the martingale problem corresponding to a, b; that is, the martingale problem for a, b is well posed. Hence the preceding paragraph indicates how the Cauchy problem (4.6)-(4.7) for $\frac{\partial}{\partial s} + L^{(s)}$ can be handled and a probabilistic solution got. As not much is known about solving degenerate parabolic / elliptic PDE's, the importance of the above can be gauged.

Problems with Dirichlet-type boundary conditions can be similarly handled using hitting times; in the degenerate case there are two natural choices of hitting times, viz. hitting time of the boundary and hitting time of the exterior of the domain - giving rise to different problems! For a reader interested in the nuances, and the connections with the analytic works of Fichera, Oleinik, Nirenberg, etc. [SV6] is an excellent source. Cauchy problem for the operator $\frac{\partial}{\partial s} + L^{(s)} + q(\cdot,\cdot)$ and the consequent Feynman-Kac formula can be treated even for badly behaved potentials $q(\cdot,\cdot)$; see [AS], [Fl], [CFZ] and the references therein for more on similar themes.

Another related development is the characterizaton of the support of the diffusion, even when the ellipticity condition (2.5) does not hold; this again due to Stroock and Varadhan in a fundamental paper [SV4]. Given σ, b as in Theorem 2.4, the support of $P_{s,x}$ is the closure of $\mathcal{S}_{s,x}$ where

$$\begin{aligned} \mathcal{S}_{s,x} = \quad & \{\psi : [0,\infty) \to \mathbb{R}^d ; \psi \text{ solves} \\ & \quad \psi(t) = x + \int_s^t b(r, \psi(r))dr + \int_s^t \sigma(r, \psi(r))\dot{\phi}(r)dr \\ & \text{with } \phi \in C^\infty([0,\infty) : \mathbb{R}^d)\}. \end{aligned}$$

In the nondegenerate case, of course, $\bar{\mathcal{S}}_{s,x} = \{w \in \Omega : w(s) = x\}$. Strong maximum principles for parabolic / elliptic operators can be deduced from the above. Moreover existence of the so called Onsager-Machlup function of statistical physics can be established; see [IW] for details. Finding analogues of Stroock-Varadhan support theorem for various types of Makrov processes has since become a major line of investigation; see the recent paper [S] and the references therein.

Probabilistic treatment of Neumann and oblique derivative problems can be analogously dealt with using reflected diffusions; see [H], [R1]. Probabilistic solutions of PDE's are also connected with the so called viscosity solutions developed in the 80's; see [L] and the references therein.

Acknowledgment: The author thanks A. Sitaram for bringing [Vj] to his notice.

References

[AS] M. Aizenman and B. Simon : Brownian motion and Harnack inequality for Schrödinger operators. *Comm. Pure Appl. Math.* **35** (1982) 209-273.

[B] R. Bass : *Diffusions and elliptic operators.* Springer, New York, 1998.

[BK] A. Bhatt and R.L. Karandikar : Invariant measures and evolution equations for Markov processes characterized via martingale problems. *Ann. Probab.* **21** (1993) 2246-2268.

[Bo] V.S. Borkar : *Optimal control of diffusion processes.* Pitman Research notes in Math.# 203, Longman Scientific & Technical, New York, 1989.

[CFZ] M. Cranston, E. Fabes and Z. Zhao : Conditional gauge and potential theory for the Schrödinger operator. *Trans. Amer. Math. Soc.* **307** (1988) 171-194.

[DP] A. De Masi and E. Presutti : *Mathematical methods for hydrodynamic limits.* LNM # 1501, Springer, New York, 1991.

[Dy] E.B. Dynkin : *Markov processes*, 2 vols. Springer, New York, 1965.

[EK] S.N. Ethier and T.G. Kurtz : *Markov processes : characterization and convergence.* John Wiley, New York, 1986.

[Fl] M. Freidlin : *Functional integration and partial differential equations.* Princeton Univ. Press, 1985.

[Fm] A. Friedman : *Partial differential equations of parabolic type.* Robert E. Krieger Publ. Co., Malabar, Florida, 1983.

[H] P. Hsu : Probabilistic approach to the Neumann problem. *Comm. Pure Appl. Math.* **38** (1985) 445-472.

[IW] N. Ikeda and S. Watanabe : *Stochastic differential equations and diffusion processes.* North-Holland / Kodansha, 1981.

[KI] K. Ito : *Lectures on Stochastic processes*, Tata Institute of Fundamental Research, Bombay, 1961.

[SI] S. Ito : *Diffusion equations.* Amer. Math. Soc. 1992.

[KS] I. Karatzas and S.E. Shreve : *Brownian motion and stochastic calculus.* Springer, New York, 1991.

[L] P.L. Lions : Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, Part 2 : Viscosity solutions and uniqueness. *Comm. in P.D.E.* **8** (1983) 1229-1276.

[P] K.R. Parthasarathy : *Probability measures on metric spaces.* Academic Press, New York, 1967.

[PS] S.C. Port and C.J. Stone : *Brownian motion and classical potential theory.* Academic Press, New York, 1978.

[R1] S. Ramasubramanian : Diffusions and the Neumann problem in the half space. *Sankhya Ser. A* **54** (1992) 351-378.

[R2] S. Ramasubramanian : Reflecting diffusions. *Current Science* **69** (1995) 231-236.

[Rf] F. Reif : *Fundamentals of statistical and thermal physics.* Mc Graw-Hill, London, 1965.

[RW] L.C.G. Rogers and D. Williams : *Diffusions, Markov processes, and martingales, vol. 2 : Ito Calculus.* Cambridge Univ. Press, UK, 2000.

[S] T. Simon : Support theorem for jump processes. *Stochastic Proc. Appl.* **89** (2000) 1-30.

[SV1] D.W. Stroock and S.R.S. Varadhan : Diffusion processes with continuous coefficients, I. *Comm. Pure Appl. Math.* **22** (1969) 345-400.

[SV2] D.W. Stroock and S.R.S. Varadhan : Diffusion processes with continuous coefficients, II. *Comm. Pure Appl. Math.* **22** (1969) 479-530.

[SV3] D.W. Stroock and S.R.S. Varadhan : Diffusion processes. In *Proc. of the sixth Berkely symposium on Math. Statistics and Probability, vol. III*, pp. 361-368, 1970.

[SV4] D.W. Stroock and S.R.S. Varadhan : On the support of diffusion processes with applications to the strong maximum principle. In *Proc. of the sixth Berkeley symposium on Math. Statistics and Probability, vol. III*, pp. 333-359, 1970

[SV5] D.W. Stroock and S.R.S. Varadhan : Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.* **24** (1971) 147-225.

[SV6] 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.

[SV7] D.W. Stroock and S.R.S. Varadhan : *Multidimensional diffusion processes.* Springer, New York, 1979.

[Vj] V.S. Varadarajan : Some mathematical reminiscences. *Preprint*, 2001.

[V] S.R.S. Varadhan : *Large deviations and applications* SIAM, Philadelphia, 1984.

[VW] S.R.S. Varadhan and R.J. Williams : Brownian motion in a wedge with oblique reflection. *Comm. Pure Appl. Math.* **38** (1985) 405-443.