JAMES R. NORRIS

Integration by parts for jump processes

Séminaire de probabilités (Strasbourg), tome 22 (1988), p. 271-315

<http://www.numdam.org/item?id=SPS_1988__22__271_0>

INTEGRATION BY PARTS FOR JUMP PROCESSES

by J.R. Norris

We give a simple derivation of the integration by parts formula used by Bismut [4], Bichteler, Gravereaux and Jacod [3], Léandre [7, 8] and Bass and Cranston [1] in the study of jump processes. This is done in the same spirit as Bismut [5], starting in a situation with only finitely many jumps and then taking limits. The framework within which we work is the theory of stochastic integrals and stochastic differential equations (SDEs) based on a Poisson random measure (see Jacod [6], Bichteler and Jacod [2]). We restrict to autonomous SDEs and thus to Markov processes. §1 is directed towards establishing Theorem 1.2, which asserts the validity of the integration by parts formula for solutions of a class of graded SDEs, examples of which arise naturally in §2. The development follows closely [12], §1.

In §2 we show how iterations of the formula can lead to regularity properties of the transition kernel of the process. For example, in Theorem 2.10, sufficient conditions are given on the coefficients of the SDE defining the process to ensure there is a smooth transition density satisfying

$$|D_y^\alpha p_t(x,y)| \leq \frac{C\, t^{-\gamma(n+d+1)}\,(1+|x|^\beta)}{(1+|y|^s)} \tag{0.1}$$

for $|\alpha| = n$, $t > 0$, $x, y \in \mathbb{R}^d$. The method is essentially the same as in [3], as is the nature of the results, but there are a few novel twists. Motivated by Léandre [9], Theorem 2.6 shows that it is possible to localize the main non-degeneracy hypothesis to a neighbourhood of the starting point. Lemma 2.2 which underlies this localization may be of independent interest. Theorem 2.7, taking up an idea of Bass and Cranston [1], gives conditions implying (0.1) which hold even when the flow of the defining SDE is degenerate. Theorem 2.8, by using Bismut's trick [4] involving doing each integration by parts on a separate interval, weakens some of the hypotheses for (0.1). These three theorems and their implications for the transition kernel are summarized in 2.f. Theorem 2.9 states conditions for regularity of the transition kernel in the backward variable. In 2.h, independently of the main development of §2, we recover for a specific example conditions for the existence of a resolvent density, due to Bass and Cranston.

This work is a development of [11]. I gratefully acknowledge the support of the Science and Engineering Research Council of Great Britain. I would like to thank D.A. Edwards for suggesting this topic and for his encouragement and advice. Thanks also to G.F. Vincent-Smith and D. Williams.

§1. Integration by parts

The SDE in $\mathbb{R}^d$

$$x_t = x_0 + \int_0^t X(x_{s-})ds + \int_0^t\int_E Y(x_{s-}, y)\,(\mu-\nu)\,(dy, ds)$$

provides a machine which takes a Poisson random measure μ on $E \times [0,\infty)$, and produces a Markov process x_t, starting from x_0, with generator

$$\mathcal{G}f(x) = Df(x).X(x) + \int_E \{f(x+Y(x,y)) - f(x) - Df(x).Y(x,y)\}\,G(dy).$$

Here ν is the compensator of μ and $\nu(dy,dt) = G(dy)dt$. We will use this machine, written from now on as

$$dx_t = X(x_{t-})dt + Y(x_{t-}, y)\,(\mu-\nu)\,(dy,dt),$$

to deduce from simple properties of μ an integration by parts formula for functions of the process x_t (Theorem 1.2).

Integration by parts requires a differential structure to be imposed on at least some part of E. We in fact assume that $E = \mathbb{R}^d\backslash\{0\}$, that G is a Radon measure on E and that there is an open set $E' \subseteq E$ and a function $g \in C^1(E')$ with

$$G(dy) = g(y)dy \quad \text{and} \quad g > 0 \quad \text{on} \quad E'. \tag{1.1}$$

1.a Derivation of the formula

We assume for now

$$\left.\begin{array}{ll}\text{(i)} & X,\ Y(.\,,y) \text{ are } C^1,\ Y(x\,,.) \text{ is } C^1 \text{ on } E';\\ & X(x),\ DX(x),\ Y(x,y),\ D_1Y(x,y) \text{ are uniformly bounded, and}\\ & D_2Y(x,y) \text{ is bounded on } \mathbb{R}^d\times K' \text{ for each compact } K'\subseteq E';\\ \text{(ii)} & \operatorname{supp} Y\subseteq \mathbb{R}^d\times K \text{ for some compact } K\subseteq E.\end{array}\right\}\quad(1.2)$$

These conditions ensure, in particular, that x_t has only finitely many jumps in any interval $0 \le t \le T$, and is between jumps just the solution of a first order ODE.

The integration by parts formula will concern a previsible function $\mathrm{v}(t,y)$ with values in $\mathbb{R}^d$, called the *perturbation.* Choices of this function in relation to the process x_t will be made in §2. We assume for now

$$\left.\begin{array}{ll}\text{(i)} & \mathrm{v}(t,.) \text{ is } C^1 \text{ for each } 0\le t<\infty;\ \mathrm{v} \text{ and } D_2\mathrm{v} \text{ are uniformly bounded;}\\ \text{(ii)} & \operatorname{supp}\mathrm{v}(.\,,.)\subseteq[0,\infty)\times K' \text{ for some compact } K'\subseteq E'.\end{array}\right\}\quad(1.3)$$

These conditions ensure, in particular, that, for sufficiently small $h \in \mathbb{R}$, $y \mapsto \theta^h(t,y) \equiv y + \mathrm{v}(t,y).h$ defines a diffeomorphism of E, leaving all but K' fixed.

(1.2) and (1.3) permit an easy derivation of the integration by parts formula. In 1.b we use approximations to show it remains valid under weaker conditions.

Define a perturbed random measure μ^h by

$$\int_0^t\int_E \phi(s,y)\,\mu^h(dy,ds) = \int_0^t\int_E \phi(s,\theta^h(s,y))\,\mu(dy,ds).$$

If μ has an atom at (y,t), μ^h has one at $(\theta^h(t,y),t)$: we have moved the places at which μ jumps. Set

$$\lambda^h(t,y) = \begin{cases}\det D_2\theta^h(t,y)\,.\,\dfrac{g(\theta^h(t,y))}{g(y)}, & y\in K',\\ 1, & y\notin K',\end{cases}$$

and set

$$Z_t^h = \exp\{\int_0^t\int_E \log\lambda^h(s,y)\,\mu(dy,ds) - \int_0^t\int_E (\lambda^h(s,y)-1)\,\nu(dy,ds)\}\,, \qquad (1.4)$$

then

$$dZ_t^h = Z_{t-}^h(\lambda^h(t,y)-1)\,(\mu-\nu)\,(dy,ds)\,,$$

so Z_t^h is a martingale and we may define a new probability measure $\mathbb{P}^h$ by

$$\frac{d\mathbb{P}^h}{d\mathbb{P}} = Z_t^h \quad \text{on} \quad \mathcal{F}_t\,.$$

We will show that, with the weighting Z_t^h, μ^h has the original law of μ. It suffices to check for test functions ϕ and for

$$\mathrm{U}_t^h = \exp\{\int_0^t\int_E \phi(s,y)\,\mu^h(s,y)\}\,Z_t^h$$

that $\mathbb{E}(\mathrm{U}_t^h)$ does not depend on h. We have

$$d\mathrm{U}_t^h = d(\text{mart}) + \mathrm{U}_{t-}^h\{\exp(\phi(t,\theta^h(t,y))) - 1\}\,\lambda^h(t,y)\,\nu(dy,dt)$$

so

$$\mathbb{E}(\mathrm{U}_t^h) = 1 + \mathbb{E}\int_0^t\int_{E'} \mathrm{U}_{s-}^h\{\exp(\phi(s,\theta^h(s,y))) - 1\}\,g(\theta^h(s,y))\det D_2\theta^h(s,y)\,dy\,ds$$

$$= 1 + \int_0^t \mathbb{E}(\mathrm{U}_s^h)\int_{E'}\{\exp(\phi(s,y)) - 1\}\,g(y)\,dy\,ds$$

by the Jacobian formula in $\mathbb{R}^d$, which determines $\mathbb{E}(\mathrm{U}_t^h)$ uniquely, showing in particular independence of h.

Consider the perturbed process x_t^h defined by

$$\left.\begin{aligned} dx_t^h &= X(x_{t-}^h)\,dt + Y(x_{t-}^h,y)\,(\mu^h-\nu)\,(dy,dt) \\ x_0 &= x \in \mathbb{R}^d\,. \end{aligned}\right\}\quad (1.5)$$

In replacing μ by μ^h we have altered the size of the jumps of x_t whilst preserving the times at which they occur. Since x_t^h is the same measurable function of μ^h for all h, the law of x_t^h under $\mathbb{P}^h$ does not depend on h and we have, for all test functions f,

$$\frac{\partial}{\partial h} \mathbb{E}\left[f(x_t^h)\, Z_t^h\right] = 0 . \tag{1.6}$$

Elementary results on the differentiability of an ODE in its starting point may be applied jump by jump to show that x_t^h is a.s. differentiable in h and indeed $Dx_t \equiv \left.\frac{\partial}{\partial h}\right|_{h=0} x_t^h$ satisfies the SDE obtained by differentiating (1.5) formally:

$$\begin{aligned} d\,Dx_t &= DX(x_{t-})\,Dx_{t-}\,dt + D_1Y(x_{t-},y)\,Dx_{t-}(\mu-\nu)\,(dy,dt) \\ &\quad + D_2Y(x_{t-},y)\,v(t,y)\,\mu(dy,dt) \\ Dx_0 &= 0 \in \mathbb{R}^d . \end{aligned}$$

We may differentiate (1.4) to obtain

$$R_t \equiv \left.\frac{\partial}{\partial h}\right|_{h=0} Z_t^h = \int_0^t\!\!\int_E \frac{\operatorname{div}(g.v)(s,y)}{g(y)}\, (\mu-\nu)\,(dy,ds) .$$

Finally, given that the jump times of μ restricted to K occur as a Poisson process of finite rate, it is not hard to justify differentiation of (1.6) under the expectation sign to obtain

$$\mathbb{E}\,[Df(x_t)\,Dx_t] + \mathbb{E}\,[f(x_t)\,R_t] = 0 . \tag{1.7}$$

1.b Extension of the formula

To obtain a useful result we set up a framework of conditions on X, Y and v, weaker than (1.2), (1.3) under which the integration by parts formula remains valid. First we define classes of *graded* coefficients X and Y (which include all C^1 Lipschitz coefficients) and note a result on the associated SDEs.

Definition of $\mathcal{C}_\alpha(d_1, ..., d_k)$

$\mathcal{C}_\alpha(d_1, ..., d_k)$ denotes the set of measurable functions $X : \mathbb{R}^d \to \mathbb{R}^d$ $(d = d_1 + \cdots + d_k)$ having a decomposition

$$X(x) = \begin{pmatrix} X^{(1)}(x^1) \\ \vdots \\ X^{(j)}(x^1, ..., x^j) \\ \vdots \\ X^{(k)}(x^1, ..., x^k) \end{pmatrix} \quad \text{for} \quad x = \begin{pmatrix} x^1 \\ \vdots \\ x^k \end{pmatrix}$$

such that

(i) $X^{(j)}$ is C^1 in $x^j \in \mathbb{R}^{d_j}$ $(j = 1, ..., k)$,

(ii) $\|X\|_{C_\alpha(d_1, ..., d_k)} \equiv \sup_{x \in \mathbb{R}^d} \left(\frac{|X(x)|}{(1+|x|^\alpha)} \vee \sup_{1 \le j \le k} |D_j X^{(j)}(x)| \right) < \infty .$

For a function $\rho : E \to (0,\infty)$, $C_\alpha(d_1, ..., d_k; \rho)$ denotes the set of measurable functions $Y : \mathbb{R}^d \times E \to \mathbb{R}^d$ satisfying

(i) $Y(.\,, y) \in C_\alpha(d_1, ..., d_k)$ for each $y \in E$,

(ii) $\|Y\|_{C_\alpha(d_1, ..., d_k; \rho)} \equiv \sup_{y \in E} \frac{1}{\rho(y)} \|Y(.\,, y)\|_{C_\alpha(d_1, ..., d_k)} < \infty .$

Notice that $C_\alpha(d_1, ..., d_k)$ and $C_\alpha(d_1, ..., d_k; \rho)$ increase with α and ρ. Any SDE which may be written in the form

$$dx_t = X(x_{t-})\, dt + Y(x_{t-}, y)\, (\mu - \nu)\, (dy, dt)$$

with $X \in C_\alpha(d_1, ..., d_k)$ and $Y \in C_\alpha(d_1, ..., d_k; \rho)$ will be called a $C_\alpha(d_1, ..., d_k; \rho)$-*system*, or simply a *graded system*.

Lemma 1.1

For $0 \le \alpha < \infty$, $d_1, ..., d_k \ge 1$ and $\rho \in \bigcap_{2 \le p < \infty} L^p(G)$, every $C_\alpha(d_1, ..., d_k; \rho)$-system has a unique solution x_t. For all $0 \le t < \infty$, $2 \le p < \infty$ there exist constants C and β depending only on p, t, α, $d_1, ..., d_k$, $\|\rho\|_2 + \|\rho\|_{\alpha^+ p}$, $\|X\|_{C_\alpha}$, $\|Y\|_{C_\alpha(\rho)}$ such that

$$\| \sup_{s \le t} |x_s| \|_{L^p(\mathbb{P})} \le C(1 + |x_0|^\beta).$$

Proof. By induction on k, using (for example) Bichteler and Jacod [2], Theorem A.6 for $k = 1$ and the inductive step. (The proof works because, whilst X and Y are not Lipschitz, $X^{(j)}$ and $Y^{(j)}$ are Lipschitz in x^j and one can show inductively that $\sup_{s \le t} |(x_t^1, ..., x_t^{j-1})| \in L^p(\mathbb{P})$.) ◊

Recall that G is a Radon measure on $E = \mathbb{R}^d \setminus \{0\}$, E' is an open subset of E, $g \in C^1(E')$ and we assume (1.1), that is $G(dy) = g(y)dy$ and $g > 0$ on E'.

Theorem 1.2 (Integration by parts formula)

Let $X \in C_\alpha(d_1, \dots, d_k)$ and $Y \in C_\alpha(d_1, \dots, d_k; \rho)$, for some $0 \le \alpha < \infty$, $d_1 + \cdots + d_k = d$, and some continuous, strictly positive $\rho \in \bigcap_{2 \le p < \infty} L^p(G)$. Let x_t be the unique solution of the graded system

$$\left.\begin{aligned} dx_t &= X(x_{t-})\,dt + Y(x_{t-},y)\,(\mu - \nu)\,(dy,dt) \\ x_0 &= x \in \mathbb{R}^d . \end{aligned}\right\} \qquad (1.8)$$

Suppose X and $Y(.\,, y)$ are C^1 with

$$\left.\begin{aligned} &\text{(i)} \quad |DX(x)| \le C(1 + |x|^\alpha) \\ &\text{(ii)} \quad |D_1Y(x,y)| \le (1 + |x|^\alpha)\rho(y) \end{aligned}\right\} \qquad (1.9)$$

and suppose $Y(x\,,.)$ is C^1 on E'.

Let $\nu(t,y)$ be a previsible function, vanishing off E', with values in $\mathbb{R}^d$, satisfying

$$\left.\begin{aligned} &\text{(i)} \quad |D_2Y(x,y)\,\nu(t,y)| \le A(t)\,(1 + |x|^\alpha)\rho(y)^2 \\ &\text{(ii)} \quad |\nu(t,y)| \le A(t)\,\mathrm{dist}\,(y, \partial E')\rho(y), \\ &\text{(iii)} \quad \nu(t,.) \text{ is } C^1 \text{ on } E' \text{ with } |\mathrm{div}(g.\nu)\,(t,y)| \le A(t)g(y)\rho(y), \end{aligned}\right\} \qquad (1.10)$$

for some increasing previsible process $A(t) \in \bigcap_{p < \infty} L^p(\mathbb{P})$.

Then

(a) The SDE

$$\left.\begin{aligned} d\,Dx_t &= DX(x_{t-})\,Dx_{t-}\,dt + D_1Y(x_{t-},y)\,Dx_{t-}(\mu - \nu)\,(dy,dt) \\ &\qquad + D_2Y(x_{t-},y)\,\nu(t,y)\,\mu(dy,dt) \\ Dx_0 &= 0 \in \mathbb{R}^d \end{aligned}\right\} \qquad (1.11)$$

has a unique solution $Dx_t \in \bigcap_{p < \infty} L^p(\mathbb{P})$, for $0 \le t < \infty$.

(b) The stochastic integral

$$R_t = \int_0^t \int_E \frac{\operatorname{div}(g.\nu)(s,y)}{g(y)} (\mu - \nu)(dy,ds) \tag{1.12}$$

is well defined and $R_t \in \bigcap_{p<\infty} L^p(\mathbb{P})$, for $0 \le t < \infty$.

(c) For all test functions f, and all $0 \le t < \infty$

$$\mathbb{E}[Df(x_t)\,Dx_t] + \mathbb{E}[f(x_t)\,R_t] = 0. \tag{1.13}$$

Proof. (a) The graded structure of X, Y permits an inductive proof, for $j = 1, ..., k$, that the $\mathbb{R}^{d_j}$-component Dx_t^j is well defined and lies in $\bigcap_{p<\infty} L^p(\mathbb{P})$.

(b) See for example Bichteler and Jacod [2], Lemma (A.14).

Before proving (c) we state for reference a general criterion for convergence in $L^p(\mathbb{P})$ of $x_t(n)$ to x_t, where x_t is the solution of

$$\left.\begin{aligned} dx_t &= X(x_{t-})\,dt + Y(x_{t-},y)\,(\mu - \nu)(dy,dt) \\ x_0 &= x \in \mathbb{R}^d \end{aligned}\right\}$$

and $x_t(n)$ the solution of an approximating SDE with coefficients X_n, Y_n and $x_0(n) = x$. Here, we allow all coefficients an implicit previsible dependence on ω and t. Fix $0 \le T < \infty$. Suppose

(i) *The approximating coefficients are C^1, Lipschitz, uniformly in n:*

$$|DX_n(x)| \le C, \quad |D_1 Y_n(x,y)| \le \rho(y)$$

for all n, for some $0 \le C < \infty$, $\rho \in \bigcap_{2 \le p < \infty} L^p(G)$;

(ii) *they converge in $L^p(\mathbb{P})$ for some $p \ge 2$:*

$$\mathbb{E}\left[\sup_{t\le T}|\int_0^t\{X(x_{s-})-X_n(x_{s-})\}\,ds|^p\right]\to 0,$$

$$\mathbb{E}\left[\sup_{t\le T}|\int_0^t\int_E\{Y(x_{s-},y)-Y_n(x_{s-},y)\}\,(\mu-\nu)(dy,ds)|^p\right]\to 0,$$

Then $$\mathbb{E}\left[\sup_{t\le T}|x_t-x_t(n)|^p\right]\to 0 \quad \text{as} \quad n\to\infty.$$

(c) Consider the conditions

(i) $\nu(t,y)$ and $D_2\nu(t,y)$ are uniformly bounded;

(ii) $\operatorname{supp} Y\subseteq \mathbb{R}^d\times K$ and $\operatorname{supp}\nu\subseteq[0,\infty)\times K'$ for some compacts $K\subseteq E$ and $K'\subseteq E'$;

(iii) $X(x)$, $Y(x,y)/\rho(y)$, $DX(x)$, $D_1Y(x,y)/\rho(y)$ are uniformly bounded, and $D_2Y(x,y)$ is bounded on $\mathbb{R}^d\times K'$ for all compacts $K'\subseteq E'$.

(1.14)

These certainly imply (1.2), (1.3) and thus the validity of the integration by parts formula (1.13). We complete the proof in three steps, showing in turn that, under the hypotheses of the theorem, each of the above conditions may be dispensed with. In each step approximating sequences $x_t(n)$, $Dx_t(n)$ and $R_t(n)$, corresponding to approximations to X, Y and ν for which (1.13) is known to hold, are shown to converge in $L^2(\mathbb{P})$ to x_t, Dx_t and R_t which thus also satisfy (1.13). In the first two steps it will be clear by (1.14)(iii) that the approximating coefficients are C^1 Lipschitz uniformly in n, so we have only to check part (ii) of our convergence criterion.

Step I. Suppose (1.14)(ii), (iii) hold. By (1.14)(ii) the sets

$$B_n \equiv \{t\ge 0: |\nu(t,y)|,\ |D_2\nu(t,y)|\le n \quad \text{for all} \quad y\in E\}$$

increase to $[0,\infty)$ as $n\to\infty$. Set $\nu_n(t,y)=\nu(t,y)1_{B_n}(t)$ then ν_n satisfies (1.14)(i). Write $B_n^c=[0,\infty)\backslash B_n$. To obtain the desired convergence of $Dx_t(n)$ and $R_t(n)$ it suffices that, as $n\to\infty$,

$$\mathbb{E}\left[\sup_{s\le t}|\int_0^s\int_E D_2Y(x_{r-},y)\,(\nu-\nu_n)\,(r,y)\,\mu(dy,dr)|^2\right]$$

$$\leq \mathbb{E}\left[\left(\int_0^t\int_E A(s)\,(1+|x_{s-})|^{\alpha})\rho(y)^2 1_{B_n^c}(s)\,\mu(dy,ds)\right)^2\right] \to 0$$

and

$$\mathbb{E}\left[\,|R_t - R_t(n)|^2\right] = \mathbb{E}\left[\left|\int_0^t\int_E \frac{\operatorname{div}(g.(\mathrm{v}-\mathrm{v}_n))(s,y)}{g(y)}\,(\mu-\nu)(dy,ds)\right|^2\right]$$

$$\leq \mathbb{E}\left[\int_0^t\int_E A(s)^2\rho(y)^2 1_{B_n^c}(s)\,G(dy)ds\right] \to 0\,.$$

Step II. Suppose (1.14)(iii) holds. Set

$$E_n = \{\frac{1}{n} < |y| < n\}$$

and

$$E'_n = \{y \in E' : \operatorname{dist}(y\,,\ \partial E') > \frac{1}{n} \quad \text{and} \quad |y| < n\}\,.$$

By (1.10)(ii) there exists a sequence (ϕ_n) in $C^1(E)$ with $1_{E'_n} \leq \phi_n \leq 1_{E'_{2n}}$ and $|<D\phi_n(y),\,\mathrm{v}(t,y)>| \leq 3A(t)\rho(y)1_{E'\setminus E'_n}(y)$ for all n. Choose also a sequence (ψ_n) in $C^1(E)$ with $1_{E_{2n}} \leq \psi_n \leq 1_{E_{3n}}$. Observe that $Y_n(x,y) \equiv Y(x,y)\psi_n(y)$, $\mathrm{v}_n(t,y) = \mathrm{v}(t,y)\phi_n(y)$ satisfy (1.14)(ii) (with $K = \overline{E}_{3n}$, $K' = \overline{E}'_{2n}$. To obtain the desired convergence of $x_t(n), Dx_t(n), R_t(n)$ it suffices that, as $n \to \infty$,

$$\mathbb{E}\left[\sup_{s\leq t}\left|\int_0^s\int_E (Y-Y_n)\,(x_{r-},y)\,(\mu-\nu)\,(dy,dr)\right|^2\right]$$

$$\leq \mathbb{E}\left[\sup_{s\leq t}\sup_{y\in E}\left|\frac{Y(x_{s-},y)}{\rho(y)}\right|^2\right] t.\int_E (1-\psi_n(y))^2\,\rho(y)^2\,G(dy) \to 0\,,$$

$$\mathbb{E}\left[\sup_{s\leq t}\left|\int_0^s\int_E D_1(Y-Y_n)\,(x_{r-},y)\,Dx_{r-}(\mu-\nu)\,(dy,dr)\right|^2\right]$$

$$\leq \mathbb{E}\left[\sup_{s\leq t}\sup_{y\in E}\left|\frac{D_1Y(x_{s-},y)\,Dx_{s-}}{\rho(y)}\right|^2\right] t.\int_E (1-\psi_n(y))^2\,\rho(y)^2\,G(dy) \to 0\,,$$

$$\mathbb{E}\left[\sup_{s\leq t}\left|\int_0^s\int_E (D_2Y(x_{r-},y)\,\nu(r,y) - D_2Y_n(x_{r-},y)\,\nu_n(r,y))\,\mu(dy,dr)\right|^2\right]$$

$$\leq \mathbb{E}\left[\left(\int_0^t\int_{E'} A(s)\,(1+|x_{s-}|^{\alpha})\,(1-\phi_n(y))\,\rho(y)^2\,\mu(dy,ds)\right)^2\right] \to 0\,,$$

$$\mathbb{E}\left[|R_t - R_t(n)|^2\right] = \mathbb{E}\left[\int_0^t\int_{E'}\left((1-\phi_n(y))\,\frac{\operatorname{div}(g.\nu)(s,y)}{g(y)} - \langle D\phi_n(y), \nu(s,y)\rangle\right)^2 G(dy)ds\right]$$

$$\leq \mathbb{E}\left[A(t)^2\right] t.\int_{E'} (1-\phi_n(y) + 3.1_{E'\setminus E'_n}(y))^2\,\rho(y)^2\,G(dy) \to 0\,.$$

Step III. Here we use the graded structure of X, Y to replace the uniform boundedness conditions (1.14)(iii). For $j = 1, \ldots, k$, choose a sequence (ψ_n^j) in $C^1(\mathbb{R}^{d_j})$ with $1_{\{|x^j|\leq n\}} \leq \psi_n^j \leq 1_{\{|x^j|\leq 2n\}}$ and $\sup_n \sup_{x^j\in\mathbb{R}^{d_j}} n\,|D\psi_n^j(x^j)| < \infty$.

Set
$$\psi_n^{(j)}(x) = \psi_n^1(x^1)\bullet\cdots\bullet\psi_n^j(x^j)\,,$$
$$X_n^{(j)}(x) = X^{(j)}(x)\,\psi_n^{(j-1)}(x)\,\psi_{n^\alpha}^j(x^j)\,,$$
$$Y_n^{(j)}(x,y) = Y^{(j)}(x,y)\,\psi_n^{(j-1)}(x)\,\psi_{n^\alpha}^j(x^j)\,.$$

Then

$$X_n \equiv \begin{pmatrix} X_n^{(1)} \\ \cdot \\ \vdots \\ X_n^{(k)} \end{pmatrix} \quad\text{and}\quad Y_n \equiv \begin{pmatrix} Y_n^{(1)} \\ \cdot \\ \vdots \\ Y_n^{(k)} \end{pmatrix}$$

satisfy (1.14)(iii) for all n. Since $X_n = X$ and $Y_n = Y$ on $\{|x| \leq n\}$ we have $x_t(n) \to x_t$ and $Dx_t(n) \to Dx_t$ a.s. Thus to establish convergence in $L^2(\mathbb{P})$ it suffices to show boundedness in $L^p(\mathbb{P})$, for some $p > 2$.

First we show $\sup_n \| \sup_{s \le t} |x_s(n)| \|_{L^p(\mathbb{P})} < \infty$ for all $1 \le p < \infty$. Observe that $|X_n| \le |X|$ and

$$|X^{(j)} \psi_n^{(j-1)} (x^1, ..., x^j)| \le |X^{(j)} \psi_n^{(j-1)}(x^1, ..., x^{j-1}, 0)| + |x^j| \, \|D_j X^{(j)}\|_\infty$$
$$\le (1 + (2n)^\alpha + |x^j|) \, \|X\|_{C_\alpha}$$

so

$$|D_j X_n^{(j)}| = |D_j X^{(j)}.\psi_n^{(j-1)}.\psi_{n^\alpha}^j + X^{(j)}.\psi^{(j-1)} D\psi_{n^\alpha}^j|$$
$$\le \|X\|_{C_\alpha} + (1 + 2(2n)^\alpha) \, \|D\psi_{n^\alpha}^j\|_\infty \, \|X\|_{C_\alpha}, \tag{1.15}$$

and so $\sup_n \|X_n\|_{C_\alpha} < \infty$. Similarly $\sup_n \|Y_n\|_{C_\alpha(\rho)} < \infty$. By Lemma 1.1 we are done.

Now we show inductively, for $j = 1, ..., k$, that $\sup_n \| \sup_{s \le t} |Dx_s^j(n)| \|_{L^p(\mathbb{P})} < \infty$ for all $2 \le p < \infty$. A slight generalisation of Lemma 1.1 to previsible coefficients would tell us what to check: instead we work directly. The sequences of coefficients of the SDEs

$$d\,Dx_t^j(n) = \sum_{i=1}^{j} D_i X_n^{(j)}(x_t(n)) \, Dx_t^i(n) \, dt$$
$$+ \sum_{i=1}^{j} D_{1,i} Y_n^{(j)}(x_t(n),y) \, Dx_t^i(n)(\mu - \nu)(dy,dt)$$
$$+ D_2 Y_n^{(j)}(x_t(n),y) \, \nu(t,y) \, \mu(dy,dt)$$

$$Dx_0^j(n) = 0 \in \mathbb{R}^{d_j}$$

are Lipschitz, uniformly in n, for each j, by (1.15). Thus it suffices to check, assuming the inductive hypothesis for $1, ..., j-1$, that they are bounded in $L^p(\mathbb{P})$. But we know

$$\sup_n \| \sup_{s \le t} |x_s(n)| \|_{L^p(\mathbb{P})} < \infty, \qquad \text{for all} \quad 2 \le p < \infty,$$

$$|DX_n^{(j)}(x)| = |DX^{(j)}.\psi_n^{(j-1)}.\psi_{n^\alpha}^j(x) + X^{(j)} D(\psi_n^{(j-1)}\psi_{n^\alpha}^j)(x)|$$
$$\le (C + \|X\|_{C_\alpha} \cdot \frac{j}{n})(1 + |x|^\alpha).$$

Similarly

$$|D_1 Y_n^{(j)}(x,y)| \le (1 + \|Y\|_{C_\alpha(\rho)} \cdot \frac{j}{n})(1 + |x|^\alpha) \rho(y)$$

and

$$|D_2 Y_n(x,y)\, v(t,y)| \;\leq\; A(t)\,(1 + |x|^{\alpha})\, \rho(y)^2 ,$$

so this is indeed the case. ◊

§2. Semigroup estimates

Our object of study remains the solution of an SDE of the form

$$\left.\begin{aligned} dx_t &= X(x_{t-})\,dt + Y(x_{t-},y)\,(\mu - \nu)\,(dy,dt) \\ x_0 &\in \mathbb{R}^d . \end{aligned}\right\} \tag{2.1}$$

We explain how, by N applications of the integration by parts formula of §1

$$\mathbb{E}\,[Df(x_t)\,Dx_t] + \mathbb{E}\,[f(x_t)\,R_t] = 0$$

we can derive expressions

$$\mathbb{E}\,[D^{\alpha} f(x_t)] = \mathbb{E}\,[f(x_t)\,Q_t^{\alpha}], \quad D_{x_0}^{\alpha}\,\mathbb{E}[f(x_t)] = \mathbb{E}[f(x_t)\tilde{Q}_t^{\alpha}] \tag{2.2}$$

for multi-indices α of length N. Any conditions on X, Y and the jump intensity G which justify this derivation and allow bounds on $\mathbb{E}[|Q_t^{\alpha}|^p]$, $\mathbb{E}[|\tilde{Q}_t^{\alpha}|^p]$ will then yield useful estimates on the semigroup $P_t = e^{t\mathcal{G}}$ through its representation $P_t f(x_0) = \mathbb{E}\,[f(x_t)]$.

Recall from §1 that $E = \mathbb{R}^d \backslash \{0\}$, $G(dy)$ is a Radon measure on E, E' is an open subset of E, $G(dy) = g(y)dy$ on E' with $g \in C^1(E')$, $g > 0$ on E'; μ is a Poisson random measure on $E \times [0,\infty)$ with intensity $\nu(dy,dt) = G(dy)dt$, and $X : \mathbb{R}^d \to \mathbb{R}^d$, $Y : \mathbb{R}^d \times E \to \mathbb{R}^d$ are nice enough to give (2.1) a unique solution $x_t \in \bigcap_{p<\infty} L^p(\mathbb{P})$.

The main results of §2 are contained in 2.e and summarized in 2.f. The four preceding subsections isolate and discuss some ideas necessary to these results. In 2.a we consider integration by parts in the special case of an independent increment process. This leads us to consider in 2.b conditions for the integrability of $\left(\int_0^t \int_E h(y)\,\mu(dy,ds)\right)^{-1}$ for $h \in L^1_+(G)$. In 2.c we define some good properties of (g,E') which distinguish those cases where we can prove results of the form (2.2), and give some examples. 2.d examines a simple example illustrating the phenomenon of degenerating flow and its effect on densities.

We are principally interested in the forward variable and in relations of the form

$$D_y^\alpha P_t(x,dy) = q_t^\alpha(x,y)\, P_t(x,dy) \qquad x,\, y \in \mathbb{R}^d$$

together with bounds

$$\int_{\mathbb{R}^d} |q_t^\alpha(x,y)|^p\, P_t(x,dy) \le C_t(1+|x|^\beta)$$

and sometimes even $C_t \le Ct^{-\gamma N p}$. Such is the conclusion of Theorem 2.5 in 2.e. Theorems 2.6, 2.7 and 2.8 examine ways in which the hypotheses of Theorem 2.5 on X, Y and (g, E') may be weakened. Theorem 2.6 examines the possibility of a *local* non-degeneracy hypothesis on the coefficient Y. Theorem 2.7 examines the case where the flow may degenerate. Theorem 2.8 incorporates Bismut's [4] trick of using perturbations of disjoint support to relax regularity conditions on G, X, Y. No two of the improvements afforded by Theorems 2.6, 2.7, 2.8 can be achieved simultaneously. The implications of these results for the regularity of $P_t(x,dy)$ in y are reviewed in Theorem 2.10.

It becomes clear that, with slight changes in the machinery, we can get estimates in the backward variable of the form

$$D_x^\alpha P_t(x,dy) = \tilde{q}_t^\alpha(x,y)\, P_t(x,dy) \qquad x,\, y \in \mathbb{R}^d$$

with similar bounds on $\tilde{q}$ as on q above. These are stated in Theorem 2.9.

In 2.g we consider whether integration by parts gives the correct rate for slowly regularizing densities and the correct singularity as $t \downarrow 0$. This is done by a comparison with results obtainable from the characteristic function when this is known. The final subsection 2.h illustrates how a careful analysis of the remainder term on integrating by parts can improve results: we take a specific one-dimensional example and recover a criterion of Bass and Cranston for the existence of a resolvent density.

2.a Example: Lévy processes

If we take $X(x) \equiv 0$, $Y(x,y) \equiv y$ we find as a special case the independent increment process $x_t = \int_0^t \int_E y\,(\mu - \nu)\,(dy,ds)$. (We assume for the moment that $\int_E |y|^p G(dy) < \infty$ for all $2 \le p < \infty$, so this makes sense.) Taking a perturbation of the form $\nu(t,y) = h(y) e_j$, where $h \in C^1_+(E')$, $h \equiv 0$ off E', and $e_1, \dots, e_d$ is the standard basis of $\mathbb{R}^d$, the integration by parts formula reads

$$\mathbb{E}\left[D_j f(x_t) H_t \right] + \mathbb{E}\left[f(x_t) R_t^j \right] = 0, \tag{2.3}$$

where

$$H_t = \int_0^t \int_E h(y)\, \mu(dy,ds),$$

$$R_t^j = \int_0^t \int_{E'} \frac{D_j(g.h)(y)}{g(y)} (\mu - \nu)(dy,ds).$$

The conditions

$$h \in C^1_b(E') \cap L^1_+(G), \quad \frac{h}{\operatorname{dist}(.,\partial E')} \text{ and } \frac{D(g.h)}{g} \in C_b(E') \cap L^2(G) \tag{2.4}$$

ensure the validity of (2.3) by Theorem 1.2, and indeed, *provided* $\mathbb{E}(H_t^{-p}) < \infty$ *for some* $p > 1$, permit integration by parts of $f(x_t) H_t^{-1}$ to obtain

$$\mathbb{E}\left[D_j f(x_t) \right] = \mathbb{E}\left[f(x_t) \left\{ \frac{D_j H_t}{H_t^2} - \frac{R_t^j}{H_t} \right\} \right] \tag{2.5}$$

where

$$D_j H_t = \int_0^t \int_E D_j h(y) . h(y)\, \mu(dy,ds).$$

(2.5) implies in particular that the law of x_t has a density with respect to Lebesgue measure, so is a formula worth having. We are thus led to investigate the integrability of H_t^{-1} in the next section. Of course the characteristic function

$$\mathbb{E}\,(e^{i<u,x_t>}) = \exp\{-t \int_E (1 - e^{i<u,y>})\,(G(dy))\}$$

gives a more direct and more powerful way to study the transition kernel of x_t (see 2.g).

However, we will see that integration by parts provides a more flexible tool, extending to cases where the characteristic function is unknown. To satisfy (2.4) h must not be too big, whereas for $\mathbb{E}(H_t^{-p}) < \infty$, h must not be too small; the existence of an h satisfying these conflicting requirements is a property of g and E', which determines whether x_t has a density. In 2.c we will consider further such properties.

2.b Integrability of H_t^{-1}

Fix $h \in L_+^1(E)$ and set $H_t = \int_0^t \int_E h(y)\, \mu(dy,ds)$.

Lemma 2.1

(i) $\mathbb{E}(H_t^{-p}) < \infty$ whenever $\frac{p}{t} < \lambda_h \equiv \liminf_{\varepsilon \downarrow 0} \dfrac{G(h \geq \varepsilon)}{\log(\frac{1}{\varepsilon})}$.

(ii) $\mathbb{E}(H_t^{-p}) \leq C t^{-p/\nu}$ for all $0 < t \leq 1$, provided $\liminf_{\varepsilon \downarrow 0} \dfrac{G(h \geq \varepsilon)}{\varepsilon^{-\nu}} > 0$ for some $\nu > 0$.

Proof.

$$\mathbb{E}(H_t^{-p}) = \Gamma(p)^{-1} \int_0^\infty \beta^{p-1} e^{-t\xi(\beta)}\, d\beta$$

where

$$\xi(\beta) = \int_E (1 - e^{-\beta h(y)})\, G(dy).$$

Choose $\gamma < 1$ with $p < \gamma \lambda_h t$. There are constants $c, \beta_0 < \infty$ such that $\xi(\beta) \geq \gamma \lambda_h \log \beta - c$ for all $\beta \geq \beta_0$, so

$$\int_0^\infty \beta^{p-1} e^{-t\xi(\beta)}\, d\beta \leq \frac{\beta_0^p}{p} + e^{ct} \int_{\beta_0}^\infty \beta^{p-1-\gamma\lambda_h t}\, d\beta < \infty.$$

We move to (ii). Suppose there are constants $\delta, \nu > 0$ and $c < \infty$ such that $G(h \geq \varepsilon) \geq \delta \varepsilon^{-\nu} - c$ for all $\varepsilon > 0$, then

$$\xi(\beta) \geq (1 - e^{-1})(\delta \beta^\nu - c) \equiv \delta' \beta^\nu - c',$$

so

$$\int_0^\infty \beta^{p-1}\, e^{-t\xi(\beta)}\, d\beta \;\le\; e^{c't} \int_0^\infty \beta^{p-1}\, e^{-\delta' t \beta^\nu}\, d\beta$$

$$= e^{c't}\, t^{-p/\nu} \int_0^\infty \beta^{p-1}\, e^{-\delta' \beta^\nu}\, d\beta \;<\; \infty . \qquad \Diamond$$

One can go back to the exact integral at the beginning of the proof for finer results on $\mathbb{E}(H_t^{-p})$. We are content with the crude but simple results of the lemma.

In 2.e we will try to localize to a neighbourhood of the starting point x_0 the non-degeneracy condition on Y implying that x_t has a (smooth) density. This leads us to consider integrability of H_T^{-1}, where for some $\mathbb{R}^d$-valued previsible function $u(t,y)$ and some $\delta > 0$,

$$M_t \;=\; \int_0^t \int_E u(s,y)\,(\mu - \nu)\,(dy,ds)\,,$$

$$T \;=\; \inf\{t \ge 0 \;:\; |M_t| \ge \delta\}\,.$$

Typically, for small δ, $\mathbb{E}(T^{-1}) = \infty$, so we cannot appeal to Lemma 2.1.

Lemma 2.2

We have

$$\mathbb{E}(H_T^{-p}) \;<\; \infty \quad \text{for all } p < \infty$$

provided

$$|u(t,y)|^2 \;\le\; h(y) \quad \text{and} \quad \mathbb{E}(H_t^{-p}) \;\le\; Ct^{-\gamma p}$$

$$\text{for all } 0 < t \le 1,\, y \in E\,,\; p < \infty, \quad \text{for some } \; C,\, \gamma < \infty\,.$$

Proof. Pick $q > \gamma$. We have

$$\{H_T < \varepsilon\} \;\subseteq\; \{H_{\varepsilon^{1/q}} < \varepsilon\} \;\cup\; \{H_T < \varepsilon \quad \text{and} \quad T < \varepsilon^{1/q}\}$$

and

$$\mathbb{P}(H_{\varepsilon^{1/q}} < \varepsilon) \;\subseteq\; \varepsilon^p\, \mathbb{E}(H_{\varepsilon^{1/q}}^{-p}) \;\le\; C\, \varepsilon^{(1-\gamma/q)p} \qquad \text{for all } \; 0 < \varepsilon \le 1,\, p < \infty\,.$$

To show $\mathbb{E}(H_T^{-p})$ for all $p < \infty$ it thus suffices to show $\mathbb{P}(H_T < \varepsilon \text{ and } T < \varepsilon^{1/q}) = 0(\varepsilon^p)$ as $\varepsilon \downarrow 0$ for all $p < \infty$. Write $M_t = M_t^\varepsilon + L_t^\varepsilon + K_t^\varepsilon$ with

$$M_t^\varepsilon = \int_0^t \int_E u(s,y)\, 1_{\{|u(s,y)| < \varepsilon^{1/2q}\}}\, (\mu - \nu)(dy,ds),$$

$$L_t^\varepsilon = \int_0^t \int_E u(s,y)\, 1_{\{|u(s,y)| \geq \varepsilon^{1/2q}\}}\, \mu(dy,ds),$$

$$K_t^\varepsilon = \int_0^t \int_E u(s,y)\, 1_{\{|u(s,y)| \geq \varepsilon^{1/2q}\}}\, \nu(dy,ds).$$

Set $S = \inf\{t \geq 0 : [M^\varepsilon]_t \geq \varepsilon\}$, where

$$[M^\varepsilon]_t = \int_0^t \int_E |u(s,y)|^2\, 1_{\{|u(s,y)| < \varepsilon^{1/2q}\}}\, \mu(dy,ds).$$

Note that $[M^\varepsilon]_S \leq \varepsilon + \varepsilon^{1/q}$ and $\{H_T < \varepsilon\} \subseteq \{T \leq S\}$. We have the Burkholder-Davis-Gundy inequality

$$\mathbb{E}\left(\sup_{t \leq S} |M_t^\varepsilon|^p\right) \leq C(p,d)\, \mathbb{E}\left([M^\varepsilon]_S^{p/2}\right).$$

Therefore

$$\mathbb{P}\left(\sup_{t \leq T} |M_t^\varepsilon| > \frac{\delta}{2} \text{ and } H_T < \varepsilon\right) \leq \mathbb{P}\left(\sup_{t \leq S} |M_t^\varepsilon| > \frac{\delta}{2}\right)$$

$$\leq \frac{2^p}{\delta^p} C(p,d) \left(\varepsilon + \varepsilon^{1/q}\right)^{p/2} \quad \text{for all } p < \infty.$$

But for $T < \varepsilon^{1/q}$ and $H_T < \varepsilon$, $|L_T^\varepsilon| < \varepsilon^{1-1/2q}$ and $|K_T^\varepsilon| \leq \varepsilon^{1/2q} \int_E h(y) G(dy)$, so for sufficiently small $\varepsilon > 0$,

$$\mathbb{P}\left(H_T < \varepsilon \text{ and } T < \varepsilon^{1/q}\right)$$

$$\leq \mathbb{P}\left(H_T < \varepsilon \text{ and } \sup_{t \leq T} |M_t^\varepsilon| > \frac{\delta}{2}\right) = 0(\varepsilon^p) \text{ for all } p < \infty. \qquad \Diamond$$

2.c Regularizing measures

E' is an open subset of $\mathbb{R}^d \setminus \{0\}$, $g \in C^1(E')$ with $g > 0$ and $G(dy) = g(y)\, dy$ on E'. Set

$$\mathcal{H}_g = \{h \in C_b^1(E') \cap L_+^1(G):$$

$$\frac{h}{\operatorname{dist}(.,\partial E')} \text{ and } \frac{D(g.h)}{g} \in C(E') \cap L^2(G)\} \ ,$$

$$\mathcal{H}_g^\infty = \{h \in C_b^\infty(E') \cap L_+^1(G):$$

$$\frac{h}{\operatorname{dist}(.,\partial E')} \in C(E') \cap L^2(G) \text{ and } \frac{D(g.h)}{g} \in C_b^\infty(E') \cap L^2(G)\}$$

(the b in C_b^r refers to all r derivatives). Write

$$H_t = \int_0^t \int_{E'} h(y)\, \mu(dy,ds) \ .$$

We say that (g,E') is *slowly regularizing of rate* λ or λ*-regularizing* if

$$\mathbb{E}(H_t^{-\lambda t}) < \infty \qquad \text{for all } 0 < t < \infty, \text{ for some } h \in \mathcal{H}_g \ .$$

(g,E') is ∞*-regularizing* if

$$\mathbb{E}(H_t^{-p}) < \infty \qquad \text{for all } 0 < t < \infty, p < \infty, \text{ for some } h \in \mathcal{H}_g \ .$$

(g,E') is (γ–)*super-regularizing* if

$$\mathbb{E}(H_t^{-p}) \le C t^{-\gamma p} \qquad \text{for all } 0 < t \le 1, p < \infty, \text{ for some } C, \gamma < \infty, h \in \mathcal{H}_g \ .$$

If $\mathcal{H}_g$ is replaced by $\mathcal{H}_g^\infty$ in these definitions we say (g,E') is *smoothly* λ-regularizing and so on. In each case we refer to h as the *regularizing function.*

We shall see that if (g,E') is λ-regularizing then the transition density of the process $x_t = \int_0^t \int_E y\, (\mu - \nu)(dy,ds)$ and those of many related processes acquire derivatives at rate λ as $t \to \infty$. It seems hard to provide workable general criteria for (g,E') to have regularizing properties. Instead we look at some examples.

Fix an open subset B of S^{d-1}, $\delta > 0$ and $g_0 \in C^1(\mathbb{R}^d)$ with $g_0 > 0$. We take for E' the sector $S = \{y \in \mathbb{R}^d : \ 0 < |y| < \delta, \ \ y/|y| \in B\}$.

Lemma 2.3

(i) $(|y|^{-d} g_0(y), S)$ is λ-regularizing for all $\lambda <$ area $(B).g_0(0)$. (In the case $d = 1$, area $(B) \equiv$ card (B).)

(ii) $(|y|^{-d} \log\frac{1}{|y|} g_0(y), S)$ is ∞-regularizing.

(iii) $(|y|^{-\alpha} g_0(y), S)$ is super-regularizing whenever $\alpha > d$. $(\int_0^t \int_{E'} y(\mu - \nu)(dy, ds)$ is well defined for $\alpha < d + 2$.)

Proof. Take open sets B', B'' in S^{d-1} with $\overline{B''} \subseteq B', \overline{B'} \subseteq B$. Set $S' = \{y \in \mathbb{R}^d : 0 < |y| < \frac{2\delta}{3}, \frac{y}{|y|} \in B'\}$, $S'' = \{y \in \mathbb{R}^d : 0 < |y| < \delta/3, \frac{y}{|y|} \in B''\}$. Take functions $\phi \in C^1(\mathbb{R}^d)$ with $1_{B''} \leq \phi \leq 1_{B'}$ on S^{d-1}, and $\psi \in C^1(0,\delta)$ with $1_{(0,\delta/3)} \leq \psi \leq 1_{(0,2\delta/3)}$. Set $h(y) = |y|^\beta \psi(|y|)\phi(y/|y|)$ with $\beta > 1$. Then

$$Dh(y) = \{\beta\frac{y}{|y|}\psi(|y|)\ \phi(\frac{y}{|y|}) + D\ \psi(|y|)\ y\phi(\frac{y}{|y|}) + \psi(|y|)\ D\phi(\frac{y}{|y|})(I - \frac{yy^T}{|y|^2})\ |y|^{\beta-1}$$

is bounded. Noting that $\frac{Dg_0(y)}{g_0(y)}$ is bounded on S, we can show in each case (i), (ii), (iii) there is a constant $C < \infty$ with $\frac{|D(g.h)(y)|}{g(y)} \leq C|y|^{\beta-1}$ for all $y \in S$. Also, there is a constant $\gamma > 0$ such that $\text{dist}(y, \partial S) \geq \gamma|y|$ for all $y \in S'$.

(i) Consider the case $g(y) = |y|^{-d} g_0(y)$. We have

$$\int_{E'} h(y)\,g(y)\,dy \leq \|g_0\|_\infty \text{ area }(B)\int_0^\delta r^\beta r^{-1}\,dr < \infty\ ,$$

$$\int_{E'}\left(\frac{h(y)}{\text{dist}(y,\partial S)}\right)^2 g(y)dy \leq \frac{\|g_0\|_\infty \text{ area}(B)}{\gamma^2}\int_0^\delta r^{2(\beta-1)}\,r^{-1}dr < \infty$$

$$\int_{E'}\left(\frac{|D(g.h)(y)|}{g(y)}\right)^2 g(y)dy \leq C^2\|g_0\|_\infty \text{ area}(B)\int_0^\delta r^{2(\beta-1)}r^{-1}dr < \infty$$

So $h \in \mathcal{H}_g$. But

$$\frac{G(h \geq \varepsilon)}{\log\frac{1}{\varepsilon}} \geq \frac{G(\{h \geq \varepsilon\} \cap S'')}{\log\frac{1}{\varepsilon}} = \frac{1}{\log\frac{1}{\varepsilon}}\int_{\varepsilon^{1/\beta}}^{\delta/3}\int_{B''}\frac{g_0(r,\theta)}{r}\,d\theta dr$$

$$\to \frac{g_0(0)\text{ area}(B'')}{\beta} \quad \text{as } \varepsilon \downarrow 0\ .$$

Since we may take $\frac{\text{area}(B'')}{\beta}$ arbitrarily close to area (B), (i) now follows from Lemma 2.1(i).

(ii) Consider the case $g(y) = |y|^{-d} \log \frac{1}{|y|} g_0(y)$. Then $h \in \mathcal{H}_g$, as in (i), because $\int_0^\delta r^{2(\beta-1)} r^{-1} \log \frac{1}{r} \, dr < \infty$. This time however

$$\frac{G(h \geq \varepsilon)}{(\log 1/\varepsilon)^2} \geq \frac{1}{(\log 1/\varepsilon)^2} \int_{\varepsilon^{1/\beta}}^{\delta/3} \int_{B''} \frac{g_0(r,\theta)\log 1/r}{r} \, d\theta dr$$

$$\to \frac{g_0(\theta)\,\text{area}(B'')}{2\beta} \quad \text{as } \varepsilon \downarrow \theta$$

So $\frac{G(h \geq \varepsilon)}{\log 1/\varepsilon} \to \infty$ and (g,S) is ∞-regularizing by Lemma 2.1 (i)

(iii) Suppose finally $g(y) = |y|^{-\alpha} g_0(y)$ with $\alpha > d$. In order that $h \in \mathcal{H}_g$ we must have $\int_0^\delta r^\beta r^{-\alpha+d-1} \, dr < \infty$ and $\int_0^\delta r^{2(\beta-1)} r^{-\alpha+d-1} \, dr < \infty$, that is $\beta > (\alpha - d) \vee \left(\frac{\alpha-d}{2} + 1\right)$.
But any such β will do, because

$$\frac{G(h \geq \varepsilon)}{\varepsilon^{-(\frac{\alpha-d}{\beta})}} \geq \varepsilon^{\frac{\alpha-d}{\beta}} \int_{\varepsilon^{1/\beta}}^{\delta/3} \int_{B''} \frac{g_0(r,\theta)}{r^{\alpha-d+1}} \, d\theta dr$$

$$\to (\alpha - d) g_0(0)\,\text{area}\,(B'') \quad \text{as } \varepsilon \downarrow \theta\,,$$

and we may apply Lemma 2.1 (ii) ◊

Fix $\delta > 0$ and $g_0 \in C^\infty(\mathbb{R}^d)$ with $g_0 > 0$. We take for E' the punctured ball $\{0 < |y| < \delta\}$

Lemma 2.4

(i) $(|y|^{-d} g_0(y), \{0 < |y| < \delta\})$ is smoothly λ-regularizing for all $\lambda < \frac{1}{2} g_0(0)$. area (S^{d-1}).

(ii) $(|y|^{-\alpha} g_0(y), \{0 < |y| < \delta\})$ is smoothly super-regularizing whenever $\alpha > d$.

Proof. It is clear from the proof of Lemma 2.3, dispensing with the cut-off ϕ, that for (i) we may take $h(y) = |y|^2 \psi(|y|)$, for (ii) $h(y) = |y|^\beta \psi(|y|)$ where β is any even integer $\beta > (\alpha - d) \vee \left(\frac{\alpha - d}{2} + 1\right)$, and $\psi \in C^\infty(0,\delta)$ with $1_{(0,\delta/3)} \leq \psi \leq 1_{(0,2\delta/3)}$. ◊

2.d Degeneracy in the flow of diffeomorphisms

Consider the following example: $G(E) = 1$, $X(x) = x$, $Y(x,y) = -x$, so that

$$dx_t = -x_{t-} \mu(dy,dt)$$

and give x_t a smooth initial density $\mathbb{P}(x_0 \in dy) = p_0(y)\, dy$ then

$$\mathbb{P}(x_t \in dy) = (1 - e^{-t})\, \varepsilon_0\, (dy) + e^{-t} p_0(y)\, dy\ ,$$

where ε_0 is th unit mass at 0. There is clearly some feature of this innocuous-looking SDE we will wish to avoid.

Under appropriate regularity conditions there is a version of the family of solutions of (2.1) as x_0 ranges over $\mathbb{R}^d$ (and μ remains fixed) which depends smoothly on x_0. Differentiating formally (2.1) with respect to x_0, one in fact arrives at the correct SDE for the derivative of the map $x_0 \mapsto x_t$. Define a process J_t in $\mathcal{L}(\mathbb{R}^d)$ by the SDE

$$\left.\begin{aligned} dJ_t &= DX(x_{t-}) J_{t-}\, dt + D_1 Y(x_{t-},y) J_{t-} (\mu - \nu)(dy,dt) \\ J_0 &= I \in \mathcal{L}(\mathbb{R}^d) \end{aligned}\right\} \tag{2.6}$$

An application of the Itô formula for jump processes shows that the matrix inverse J_t^{-1} should satisfy

$$\left.\begin{aligned} dJ_t^{-1} &= -J_{t-}^{-1}\left\{DX(x_{t-}) - \int_E (I + D_1 Y(x_{t-},y))^{-1} D_1 Y(x_{t-},y)^2\, G(dy)\right\} dt \\ &\quad - J_{t-}^{-1}(I + D_1 Y(x_{t-},y))^{-1} D_1 Y(x_{t-},y)\, (\mu - \nu)(dy,dt) \\ J_0^{-1} &= I \in \mathcal{L}(\mathbb{R}^d) \end{aligned}\right\} \tag{2.7}$$

and it is easy to check that when a unique solution to this SDE, K_t say, exists we have $d(J_tK_t) = 0$ and so $K_t = J_t^{-1}$. This assumes of course that $I + D_1Y(x,y)$ is invertible.

The example above behaves badly because whenever μ jumps, wherever the process is, it jumps to the same point 0 :'non-degeneracy is lost.' This can be prevented (locally) by insisting $x_0 \mapsto x_t$ is a local diffeomorphism, that is, J_t is invertible. By this discussion we motivate the assumption made in 2.e that

$$\sup_{x\in\mathbb{R}^d} \sup_{y\in E} |(I + D_1Y(x,y))^{-1}| < \infty$$

Theorem 2.7 will show however that this assumption is not always necessary.

2.e *N*th order integration by parts

We return to the general case where x_t is the solution of an SDE

$$dx_t = X(x_{t-})dt + Y(x_{t-},y)(\mu - \nu)(dy,dt) ,$$

and prove five results about the semigroup $P_tf(x_0) = \mathbb{E}(f(x_t))$, as described above.
To obtain the first result, Theorem 2.5, we choose a perturbation which gives the derived process Dx_t a simple form and perform the 'obvious' iterations of the integration by parts formula (1.13). Theorems 2.6, 2.7, 2.8, 2.9 are variations on the same theme.

To aid comparison of results we state most hypotheses relating to the coefficients X,Y in one go. E,G,μ,ν,E',g are defined at the beginning of §2 and the sorts of hypothesis to be made on (g,E') are considered in 2.c. Fix constants C, $\beta < \infty$ and a strictly positive function $\rho \in C(E') \cap \bigcap_{2\le p<\infty} L^p(G)$. Consider the conditions

(i) $X : \mathbb{R}^d \to \mathbb{R}^d$ is C^{N+1} with

$$|DX(x)| \le C$$

$$|D^nX(x)| \le C(1 + |x|^\beta), \quad n \le N+1 ;$$

(ii) $Y : \mathbb{R}^d \times E \to \mathbb{R}^d$ is measurable, C^{N+1} in its first argument, C^2 in its second on E' (resp. C^{N+1} on E'), with

$$|D_1Y(x,y)| \le \rho(y) \qquad (2.8)$$

$$|D_1^nY(x,y)| \le (1 + |x|^\beta)\rho(y), \quad n \le N+1, \qquad \text{(resp. 2.9)}$$

$$|D_2^m D_1^n Y(x,y)| \le C(1+|x|^\beta) \text{ on } E', \quad m=1,2, \quad n \le N,$$

$$\text{(resp. } |D_2^m D_1^n Y(x,y)| \le C(1+|x|^\beta) \text{ on } E', \quad m+n \le N+1),$$

(iii) $|D_2Y(x,y)^{-1}| \le C(1+|x|^\beta)$ on E',

(iv) $|(I+D_1Y(x,y))^{-1}| \le C$.

The hypothesis made in Theorem 2.5 is the stronger condition (2.9). Theorem 2.6 relaxes (2.9) (iii) and Theorem 2.7 relaxes (2.9) (iv). Theorem 2.8 is proved under the weaker condition (2.8).

Theorem 2.5

Suppose (g,E') is smoothly λ-regularizing $(0<\lambda\le\infty)$ and that X,Y satisfy (2.9). Then for all $t>0$ and all multi-indices α of length $N<\lambda t$ there exists a Borel function q_t^α on $\mathbb{R}^d\times\mathbb{R}^d$ such that,

$$D_y^\alpha P_t(x,dy) = q_t^\alpha(x,y)\,P_t(x,dy), \quad x,y\in\mathbb{R}^d,$$

$$\int_{\mathbb{R}^d} |q_t^\alpha(x,y)|^p\,P_t(x,dy) \le C_{t,N,p}\,(1+|x|^\beta),$$

whenever $Np<\lambda t$. If (g,E') is moreover smoothly γ-super-regularizing we can take $\lambda=\infty$, $C_{t,N,p}=\theta(t)\,.\,t^{-\gamma Np}$ where $\theta(t)$ increases with $t>0$.

Proof. Take any regularizing function h and consider the function $v:\ \Omega\times[0,\infty)\times E\to\mathcal{L}(\mathbb{R}^d)$ given by $v(t,y)=D_2Y(x_{t-},y)^{-1}\,(I+D_1Y(x_{t-},y))\,J_{t-}\,h(y)$. We associate with v and with the solution z_t of an autonomous SDE

$$dz_t = Z(z_{t-})dt + W(z_{t-},y)\,(\mu-\nu)\,(dy,dt), \quad z_0\in\mathbb{R}^m,$$

a derived process Dz_t solving the SDE

$$\begin{aligned} dDz_t &= DZ(z_{t-})\,Dz_{t-}dt + D_1W(z_{t-},y)\,Dz_{t-}(\mu-\nu)\,(dy,dt) \\ &\quad + D_2W(z_{t-},y)\,.\,v(t,y)\,\mu(dy,dt), \\ Dz_0 &= 0\in\mathcal{L}(\mathbb{R}^d,\mathbb{R}^m). \end{aligned}$$

We chose v so that $Dx_t=J_tH_t$, where

$$H_t = \int_0^t \int_E h(y)\,\mu(dy,ds)\,.$$

This may be checked by identifying J_tH_t as the unique solution of the SDE for Dx_t. Set

$$R_t\,.\,e = \int_0^t \int_E \frac{\operatorname{div}(g.\mathrm{v}.e)\,(s,y)}{g\,(y)}\,(\mu - \nu)\,(dy,ds), \qquad e \in \mathbb{R}^d\,.$$

Using each component $\mathrm{v}_j = \mathrm{v}.e_j$ as a perturbation, we obtain from (1.13) the formula in $\mathcal{L}(\mathbb{R}^d,\mathbb{R})$

$$\mathbb{E}[Df\,(z_t).Dz_t] + \mathbb{E}[f\,(z_t)R_t] = 0\,. \tag{2.10}$$

Set $z_t^{(0)} = (x_t,J_t,H_t)$, $z_t^{(1)} = (z_t^{(0)},Dz_t^{(0)},R_t)$ and define inductively $z_t^{(n)} = (z_t^{(n-1)},Dz_t^{(n-1)})$, $n \geq 2$. The process $z_t = z_t^{(n)}$, $n \geq 0$, satisfies an autonomous SDE. We apply (2.10) with $f\,(z_t)$ replaced by $f\,(x_t)\,J_t^{-1}H_t^{-1}\,q(z_t^{(n)})$, where f and q are test functions. This produces a formula in $\mathcal{L}(\mathbb{R}^d,\mathcal{L}(\mathbb{R}^d))$. On taking the trace we get

$$\mathbb{E}[Df\,(x_t)\,q\,(z_t^{(n)})] = \mathbb{E}[f\,(x_t)\,(\mathcal{A}q)\,(z_t^{(n+1)})], \tag{2.11}$$

where

$$\begin{aligned}(\mathcal{A}q)\,(z_t^{(n+1)}) = &-D.\,(J_t^{-1})\,H_t^{-1}q\,(z_t^{(n)}) + DH_t\,.\,J_t^{-1}H_t^{-2}q(z_t^{(n)})\\ &-Dq\,(z_t^{(n)})\,.\,Dz_t^{(n)}\,.\,J_t^{-1}H_t^{-1} - R_t\,.\,J_t^{-1}H_t^{-1}q\,(z_t^{(n)}),\end{aligned} \tag{2.12}$$

and where

$$D.\,(J_t^{-1}).e = \operatorname{trace}\,(e' \mapsto -J_t^{-1}(DJ_t.e')\,J_t^{-1}e), \quad e \in \mathbb{R}^d$$

Writing $\mathcal{A} = (\mathcal{A}_1,\ldots,\mathcal{A}_d)$ and $\mathcal{A}_{(j_1,\ldots,j_N)} = \mathcal{A}_{j_N} \circ \ldots \circ \mathcal{A}_{j_1}$ we have by N iterations of (2.11), for $|\alpha| = N$,

$$\mathbb{E}[D^{\alpha}f\,(x_t)] = \mathbb{E}[D^{\alpha}f(x_t)1(z_t^{(0)})] = \mathbb{E}[f\,(x_t)\,(\mathcal{A}_{\alpha}1)\,(z_t^{(N)})] \tag{2.13}$$

To complete the proof we must form an adequate impression of the SDE's for $z_t^{(n)}$, $n = 0,1,\ldots,N-1$ and of the functions $\mathcal{A}_{\alpha}1$ to justify the steps leading to (2.13). In preparation we set (for fixed n)

$$\xi(x) = (X\,(x),\,DX\,(x),\ldots,D^{n+1}X\,(x)),$$

$$\eta(x,y) = (Y\,(x,y),\,D_1Y\,(x,y),\ldots,D_1^{n+1}Y\,(x,y)),$$

$$\chi(x,y) = D_2Y(x,y)^{-1}(I + D_1Y\,(x,y)),$$

$$\zeta(x,y) = \left(\frac{D\,(g.h)\,(y)}{g\,(y)},\,D_2\chi(x,y)h\,(y),\ldots,D_1^{n-1}D_2\chi(x,y)h\,(y)\right),$$

$$\rho(x,y)=\left(D^m h(y),\ D^{m-1}\left(\frac{D(g.h)(y)}{g(y)}\right),\ D_2^m D_1^k Y(x,y),\ D_2Y(x,y)^{-1} : m+k\le n+1\right)$$

We will eventually show for $n=0,1,\ldots,N-1$ that

(A) *The process $z_t^{(n)}$ is well defined as the unique solution, of a graded system of SDE's*

$$dz_t = Z(z_{t-})dt + W_1(z_{t-},y)\,(\mu-\nu)(dy,dt) + W_2(z_{t-},y)h(y)\mu(dy,dt)$$

where for some polynomials $P(\xi,z)$ linear in ξ, $Q(\zeta,z)$ linear in ζ, and $R(\rho,z)$, we have

$$Z(z)=P(\xi(x),z),$$

$$W_1(z,y)=P(\eta(x,y),z)+Q(\zeta(x,y),z),$$

$$W_2(z,y)=R(\rho(x,y),z).$$

Moreover $\mathbb{E}[|z_t^{(n)}|^p]\le\theta(t)(1+|x|^\beta)$ for some increasing function θ, for all $1\le p<\infty$.

(B) *For $|\alpha|=n$, $(\mathcal{A}_\alpha 1)(z_t^{(n)})$ is a polynomial in $z_t^{(n)}$, J_t^{-1} and $D^m H_t/H_t$, $m=1,\ldots,n$, multiplied by H_t^{-n}.*

First let us see that (A) holds for $n=0$ ((B) is trivial). We have the $C(d,d^2,1;\rho)$-system

$$dx_t = X(x_{t-})\,dt + Y(x_{t-},y)\,(\mu-\nu)\,(dy,dt)$$

$$dJ_t = DX(x_{t-})\,J_{t-}dt + D_1Y(x_{t-},y)\,J_{t-}(\mu-\nu)\,(dy,dt)$$

$$dH_t = h(y)\,\mu\,(dy,dt)$$

which is manifestly of the desired form.

Now let us see that (A), (B) suffice to justify the formal steps leading to (2.13) and show $\mathbb{E}[|(\mathcal{A}_\alpha 1)(z_t^{(N)})|^p]\le C_t(1+|x_0|^\beta)$ for $Np<\lambda t$, with $C_t=\theta(t)\,t^{-\gamma Np}$ if (g,E') is γ-super-regularizing. The result then follows on setting

$$q_t^\alpha(x_0,y)=\mathbb{E}[(\mathcal{A}_\alpha 1)(z_t^{(N)})\,|\,x_t=y].$$

Since $h \in \mathcal{H}_g^\infty$ and X,Y satisfy (2.9), (A) enables us to check the hypothesis of Theorem 1.2 and show formula (2.10) is valid for $z_t = z_t^{(n)}$ and for all test functions f. Of course $f(x_t) \,.\, J_t^{-1} H_t^{-1}(\mathcal{A}_\alpha 1)(z_t^{(n)})$ is not a test function of $z_t^{(n)}$! However

- by (2.9) (iv), in particular, the SDE (2.7) has a solution in $\bigcap_{p<\infty} L^p(\mathbb{P})$ which may be identified as J_t^{-1};
- if h is λ-regularizing and $Np < \lambda t$, we have $H_t^{-Np} \in L^1(\mathbb{P})$;
- if h is γ-super-regularizing $\mathbb{E}[H_t^{-Np}] \le Ct^{-\gamma Np}$;
- since $h \in \mathcal{H}_g^\infty$ and by (2.9) (ii), (iii) we have $D^m H_t \le A(t) H_t$ for some increasing process $A(t) \in \bigcap_{p<\infty} L^p(\mathbb{P})$, for $m = 1,\dots,n$.

Thus (A), (B) imply $\mathbb{E}[\,|(\mathcal{A}_\alpha 1)(z_t^{(n)})|^p] \le C_t(1 + x^\beta)$ for $|\alpha| = n$. Approximating $z_t^{(n)} \mapsto f(x_t) J_t^{-1} H_t^{-1}(\mathcal{A}_\alpha 1)(z_t^{(n)})$ by test functions and passing to the limit, we obtain the desired formulae.

Finally, assuming (A), (B) hold for n, we must show they hold for $n+1$ $(n = 0,1,\dots,N-2)$. Note firstly from the SDE's at the beginning of this proof how the SDE for Dz_t is obtained from that for z_t. It is routine to check from our hypotheses that, if the SDE for $z_t^{(n)}$ is a $C(d_1,\dots d_k;\, \rho)$-system then, by linearity of $P(\xi,z)$ in ξ, the SDE for $z_t^{(n+1)}$ is a $C(d_1,\dots d_k, d.d_1,\dots,d.d_k;\rho)$-system $(n \ge 1)$. Moreover, the coefficients of the SDE for $z_t^{(n+1)}$ being obtained as derivatives of coefficients for $z_t^{(n)}$, it is clear that the polynomial description given in (A) holds for these new coefficients. The $L^p(\mathbb{P})$ bound comes from Lemma 1.1. The inductive step for (B) is clear on inspecting the formula (2.12) for $\mathcal{A}q$. ◊

Theorem 2.6

Suppose (g,E') is smoothly γ-super-regularizing with regularizing function h and that X,Y satisfy (2.9) (i) (ii) (iv). Suppose further that for some open set $U \subseteq \mathbb{R}^d$, for all $x \in U$, $y \in E'$,

$$|Y(x,y)|^2 \le h(y),$$

$$|D_2 Y(x,y)^{-1}| \le C(1 + |x|^\beta).$$

Then for all $t > 0$ and all multi-indices α of length N there exists a Borel function q_t^α on $U \times \mathbb{R}^d$ such that

$$D_y^\alpha P_t(x,dy) = q_t^\alpha(x,y)\, P_t(x,dy), \qquad x \in U,\ y \in \mathbb{R}^d,$$

$$\int_{\mathbb{R}^d} |q_t^\alpha(x,y)|^p \, P_t(x,dy) \le \theta(t,\delta)\, t^{-\gamma N p} \, (1 + |x|^\beta)$$

for all $x \in U^\delta \equiv \{x \in U : \text{dist}(x, \partial U) > \delta\}$, where $\theta(t,\delta)$ increases with $t > 0$.

Proof. Only a slight modification of the proof of Theorem 2.5 is necessary. Suppose $x \in U^\delta$. Take a C^∞ function $\phi : \mathbb{R}^d \to [0,1]$ with

$$\phi(x) = \begin{cases} 1 & \text{if } |x - x_0| \le \delta/2, \\ 0 & \text{if } |x - x_0| \ge \delta \end{cases}$$

and set

$$v_\phi(t,y) = \phi(x_{t-})\, D_2 Y(x_{t-},y)^{-1} \, (I + D_1 Y(x_{t-},y))\, J_{t-} h(y) \ .$$

With this new perturbation we have $Dx_t = J_t H_t^\phi$ where

$$H_t^\phi = \int_0^t \int_E \phi(x_{s-})\, h(y)\, \mu(dy,ds) \ .$$

The cut-off ϕ serves to eliminate any singularities in $D_2 Y(x,y)$ which may result from weakening (2.9) (iii). Note that by the hypotheses of this theorem

$$\mathbb{E}[(H_t^\phi)^{-p}] \le \mathbb{E}[H_T^{-p} + H_t^{-p}] \le \theta(t,\delta)\, t^{-\gamma p} \, (1 + |x|^\beta) \quad \text{for all } p < \infty$$

by Lemma 2.2, where

$$T = \inf\{s \ge 0 : |x_s - x_0| \ge \delta/2\} \ .$$

The proof follows that of Theorem 2.5 with minor modifications. ◊

Whilst we have localized one non-degeneracy condition (2.9) (iii), Theorem 2.6 still assumes (2.9) (iv) $(I + D_1 Y(x,y))^{-1} \le C$ *for all x,y,* so this attempt at localization is rather unsatisfactory. However (2.9) (iv) cannot be made local: degeneracy of the flow is an ongoing danger!

The next result applies to the case where the flow may degenerate, that is, where condition (2.9) (iv) $|(I + D_1 Y(x,y))^{-1}| \le C$ fails. Consider the SDE for the derived process Dx_t

$$dDx_t = DX(x_{t-})\, Dx_{t-}\, dt + D_1 Y(x_{t-},y)\, Dx_{t-} (\mu - \nu)\, (dy,dt)$$

$$+ D_2 Y(x_{t-},y)\, \nu(t,y)\, \mu(dy,dt) \ .$$

The third term in this SDE is somewhat at our disposal and the idea is to use it to cancel out those jumps coming from the second term which make the matrix Dx_t (close to) degenerate, that is those for which $|(I + D_1 Y(x_{t-},y))^{-1}|$ is large.

Theorem 2.7

Suppose (g,E') is smoothly λ-regularizing $(0 < \lambda \le \infty)$ and that X, Y satisfy (2.9) (i) (ii) (iii). Suppose further that we can find $\phi \in C^\infty(E',[0,1])$ such that $G(\text{supp}\,\phi) < \infty$, $\phi(y) \le \text{dist}(y, \partial E')$, $\frac{D(g.\phi)}{g} \in C_b^{N-1}(E')$ and, setting $\phi \equiv 0$ off E', $\psi \equiv 1 - \phi$, we have

$$|(I + \psi(y)\, D_1 Y(x,y))^{-1}| \le C \ .$$

Then for all $t > 0$ and all multi-indices α of length $N < \lambda t$ there exists a Borel function q_t^α on $\mathbb{R}^d \times \mathbb{R}^d$ such that

$$D_y^\alpha P_t(x,dy) = q_t^\alpha(x,y)\, P_t(x,dy) \ , \quad x,y \in \mathbb{R}^d \, ,$$

$$\int_{\mathbb{R}^d} |q_t^\alpha(x,y)|^p \, P_t(x,dy) \le C_{t,N,p}\, (1 + |x|^\beta) \, ,$$

whenever $Np < \lambda t$. If (g,E') is moreover smoothly γ-super-regularizing we can take $\lambda = \infty$, $C_{t,N,p} = \theta(t)\, .\, t^{-\gamma Np}$ where $\theta(t)$ increases with $t > 0$.

Proof. Set

$$U(x,y) = \psi(y) D_1 Y(x,y),$$

$$V(x,y) = \phi(y)\, D_1 Y(x,y),$$

$$S(x) = DX(x) - \int_E V(x,y)\, G(dy),$$

and consider the following modification of the SDE (2.6) for J_t

$$dK_t = S(x_{t-})\, K_{t-} dt + U(x_{t-},y)\, K_{t-} (\mu - \nu)(dy,dt),$$

$$K_0 = I \in \mathcal{L}(\mathbb{R}^d).$$

We find there is a unique solution $K_t \in \bigcap_{p<\infty} L^p(\mathbb{P})$ and moreover K_t^{-1} exists and lies in $\bigcap_{p<\infty} L^p(\mathbb{P})$ for all $0 \le t < \infty$. This is because K_t^{-1} satisfies

$$dK_t^{-1} = -K_{t-}^{-1}\{S(x_{t-}) - \int_E (I + U(x_{t-},y))^{-1}U(x_{t-},y)^2\, G(dy)\}\, dt$$

$$-K_{t-}^{-1}\,(I + U(x_{t-},y))^{-1}U(x_{t-},y)\,(\mu - \nu)\,(dy,dt),$$

$$K_0^{-1} = I \in \mathcal{L}(\mathbb{R}^d).$$

and by hypothesis $|(I + U(x,y))^{-1}| \le C$.

The proof now follows closely that of Theorem 2.5 but with K_t in place of J_t and with modified perturbation

$$\mathrm{v}(t,y) = D_2Y(x_{t-},y)^{-1}\{(I + U(x_{t-},y))\,K_{t-}h(y) - V(x_{t-},y)\,K_{t-}\,H_{t-}\}$$

where h is some regularizing function and where

$$H_t = \int_0^t \int_E h(y)\,\mu(dy,ds).$$

With this choice of v we have $Dx_t = K_tH_t$, as may be checked by substituting K_tH_t into the SDE (1.11) for Dx_t.

Set $z_t^{(0)} = (x_t,K_t,H_t)$, $z_t^{(1)} = (z_t^{(0)},Dz_t^{(0)},R_t)$ and inductively $z_t^{(n)} = (z_t^{(n-1)},Dz_t^{(n-1)})$, $n \ge 2$, as in the proof of Theorem 2.5. It is a little more complicated to describe the SDE's for $z_t^{(n)}$ this time: fixing n, set

$$\xi(x) = (X(x),S(x),DS(x),\ldots,D^nS(x)),$$

$$\eta(x,y) = (Y(x,y),U(x,y),D_1U(x,y),\ldots,D_1^nU(x,y)),$$

$$\chi(x,y) = D_2Y(x,y)^{-1}\,(I + U(x,y)),\quad \pi(x,y) = D_2Y(x,y)^{-1}D_1Y(x,y),$$

$$\zeta(x,y) = \left(\frac{D(g.h)(y)}{g(y)}, D_2\chi(x,y)\,h(y),\ldots,D_1^{n-1}D_2\chi(x,y)\,h(y);\right.$$

$$\left.\frac{D(g.\phi)(y)}{g(y)}, D_2\pi(x,y)\,\phi(y),\ldots,D_1^{n-1}D_2\pi(x,y)\phi(y)\right),$$

$$\rho(x,y) = \left(D^mh(y),\, D^{m-1}\left(\frac{D(g.h)(y)}{g(y)}\right), D^m\phi(y),\, D^{m-1}\left(\frac{D(g.\phi)(y)}{g(y)}\right),\right.$$

$$\left.D_1^k\,D_2^m\,Y(x,y),\, D_2Y(x,y)^{-1} : m + k \le n + 1\right).$$

With ξ, μ, ζ, ρ so redefined we can show that assertion (A) in the proof of Theorem 2.5 holds for the new $z_t^{(n)}$. It is now routine to check the validity of each of the N integrations by parts leading to (the analogue of) (2.13) thus completing the proof. ◊

Theorem 2.8

Suppose (g,E') is λ-regularizing $(0 < \lambda \leq \infty)$ and that X,Y satisfy (2.8). Then for all $t > 0$ and all multi-indices α of length $N < \lambda t$ there exists a Borel function q_t^α on $\mathbb{R}^d \times \mathbb{R}^d$ such that

$$D_y^\alpha P_t(x,dy) = q_t^\alpha(x,y)\, P_t(x,dy)\ ,\quad x,y \in \mathbb{R}^d\,,$$

$$\int_{\mathbb{R}^d} |q_t^\alpha(x,y)|^p\, P_t(x,dy) \leq C_{t,N,p}\,(1 + |x|^\beta)\,,$$

whenever $Np < \lambda t$. If (g,E') is moreover γ-super-regularizing we can take $\lambda = \infty$, $C_{t,N,p} = \theta(t)\,.\,t^{-\gamma Np}$ where $\theta(t)$ increases with $t > 0$.

Proof. Fix $p \geq 1$, $\tau > \frac{p}{\lambda}$. We follow closely the proof of Theorem 2.5 but perform the nth integration by parts, $n = 1,\dots,N$, with perturbation

$$v^{(n)}(t,y) = D_2 Y(x_{t-},y)^{-1}\,(I + D_1 Y(x_{t-},y))\, J_{t-} h(y) 1_{((n-1)\tau, n\tau]}(t)$$

where h is some regularizing function.

We associate with $v^{(n)}$ and with the solution z_t of an autonomous SDE

$$dz_t = Z(z_{t-})dt + W(z_{t-},y)\,(\mu - \nu)\,(dy,dt), \qquad z_0 \in \mathbb{R}^m,$$

a derived process $D^{(n)} z_t$ solving the SDE

$$dD^{(n)}z_t = DZ(z_{t-})\, D^{(n)} z_{t-} dt + D_1 W(z_{t-},y)\, D^{(n)} z_{t-} (\mu - \nu)\,(dy,dt)$$
$$+ D_2 W(z_{t-},y)\,.\, v^{(n)}(t,y)\,\mu\,(dy,dt)$$
$$D^{(n)} z_t = 0 \in \mathcal{L}(\mathbb{R}^d,\mathbb{R}^m).$$

Set

$$H_t^{(n)} = \int_{t\wedge(n-1)\tau}^{t\wedge n\tau} h(y)\,\mu(dy,ds).$$

We chose $v^{(n)}$ so that $D^{(n)}x_t = J_t H_t^{(n)}$. This may be checked by identifying $J_t H_t^{(n)}$ as the unique solution of the SDE for $D^{(n)}x_t$. Define a random variable $R_t^{(n)} \in \mathcal{L}(\mathbb{R}^d, \mathbb{R})$ by

$$R_t^{(n)}.e = \int_0^t \int_E \frac{\operatorname{div}(g.v^{(n)}.e)(s,y)}{g(y)} (\mu - \nu)(dy,ds), \quad e \in \mathbb{R}^d$$

Using each component $v_j^{(n)} = v^{(n)}.e_j$ as a perturbation we obtain from (1.13) the formula in $\mathcal{L}(\mathbb{R}^d, \mathbb{R})$

$$\mathbb{E}[Df(z_t).D^{(n)}z_t] + \mathbb{E}[f(z_t)\, R_t^{(n)}] = 0. \tag{2.14}$$

Let $z_t^{(0)} = (x_t, J_t)$ and define inductively

$$z_t^{(n)} = (z_t^{(n-1)}, D^{(n)}z_t^{(n-1)}, H_t^{(n)}, D^{(n)}H_t^{(n)}, R_t^{(n)}), \quad n \geq 1.$$

The process $z_t = (z_t^{(n-1)}, H_t^{(n)})$ satisfies an autonomous SDE. We apply (2.14) with $f(z_t)$ replaced by $f(x_t)\, J_t^{-1} (H_t^{(n)})^{-1}\, q(z_t^{(n-1)})$, where f and q are test functions. This produces a formula in $\mathcal{L}(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d))$. On taking the trace we get

$$\mathbb{E}[Df(x_t)\, q(z_t^{(n-1)})] = \mathbb{E}[f(x_t)\, (\mathcal{A}^{(n)}q)(z_t^{(n)})] \tag{2.15}$$

where

$$\begin{aligned}\mathcal{A}^{(n)}q(z_t^{(n)}) = &-D^{(n)}.(J_t^{-1})\,(H_t^{(n)})^{-1}\, q(z_t^{(n-1)}) \\ &+ D^{(n)}\, H_t^{(n)} J_t^{-1}\, (H_t^{(n)})^{-2} q(z_t^{(n-1)}) \\ &- Dq(z_t^{(n-1)}).D^{(n)}z_t^{(n-1)} J_t^{-1} (H_t^{(n)})^{-1} \\ &- R_t^{(n)} J_t^{-1}\, (H_t^{(n)})^{-1}\, q(z_t^{(n-1)})\end{aligned} \tag{2.16}$$

and

$$D^{(n)}.(J_t^{-1}).e = \operatorname{trace}(e' \mapsto -J_t^{-1}(D^{(n)}J_t.e')\, J_t^{-1}e), \quad e \in \mathbb{R}^d$$

Writing $\mathcal{A}^{(n)} = (\mathcal{A}_1^{(n)}, \ldots, \mathcal{A}_d^{(n)})$ and $\mathcal{A}_{(j_1,\ldots,j_N)} = \mathcal{A}_{j_N}^{(N)} \circ \ldots \circ \mathcal{A}_{j_1}^{(1)}$ we have by N iterations of (2.15), for $|\alpha| = N$,

$$\mathbb{E}[D^\alpha f(x_t)] = \mathbb{E}[D^\alpha f(x_t) 1(z_t^{(0)})] = \mathbb{E}[f(x_t)\, (\mathcal{A}_\alpha 1)(z_t^{(N)})] \tag{2.17}$$

To complete the proof we must form an adequate impression of the SDE's for $z_t^{(n)}$, $n = 0, 1, \ldots, N-1$ and of the functions $\mathcal{A}_\alpha 1$ to justify the steps leading to (2.17) and estimate $\mathbb{E}[\,|(\mathcal{A}_\alpha 1)(z_t^{(N)})|^p]$. Fixing n we set

$$\xi(x) = (X(x), DX(x), \ldots, D^{n+1}X(x)),$$

$$\eta(x,y) = (Y(x,y), D_1 Y(x,y), \ldots, D_1^{n+1} Y(x,y)).$$

We will show for $n = 0, 1, \ldots, N-1$ that:

(A) *The process $z_t^{(n)}$ is well defined as the unique solution of an SDE*

$$dz_t = Z(z_{t-})\,dt + W(x_{t-},y)\,(\mu - \nu)\,(dy,dt)$$

$$+ W_1(t,z_{t-},y)\,(\mu - \nu)\,(dy,dt) + W_2(t,z_{t-},y)\,\mu\,(dy,dt)$$

where, for some polynomial $P(\xi,z)$, linear in ξ, $Z(z) = P(\xi(x),z)$, $W(z,y) = P(\xi(x,y),z)$, where

$$W_j(t,z,y) = \sum_{i=1}^{n} W_{ji}(z,y)\, 1_{((i-1)\tau, i\tau]}(t), \quad j = 1,2,$$

and where for some k, $d_1, \ldots, d_k$, $\rho \in \bigcap_{2 \le p < \infty} L^p(G)$ and for $i = 1, \ldots, n$

$$Z \in C(d_1, \ldots, d_k), \quad W, W_{1i} \in C(d_1, \ldots, d_k; \rho), \quad W_{2i} \in C(d_1, \ldots, d_k; \rho^2).$$

Moreover $\mathbb{E}[|z_t^{(n)}|^p] \le \theta(t)\,(1 + |x_0|^\beta)$ for some increasing function θ, for all $1 \le p < \infty$.

(B) *For $|\alpha| = n$, $(\mathcal{A}_\alpha 1)(z_t^{(n)})$ is a polynomial in $z_t^{(n)}$, J_t^{-1} and $D^{(k)} H_t^{(k)} / H_t^{(k)}$, $k = 1, \ldots, n$, multiplied by $\left(\prod_{k=1}^{n} H_t^{(k)} \right)^{-1}$.*

First we show that (A) holds for $n = 0$. ((B) is vacuous.) We have SDE's

$$dx_t = X(x_{t-})\,dt + Y(x_{t-},y)\,(\mu - \nu)\,(dy,dt)$$

$$dJ_t = DX(x_{t-})\,J_{t-}\,dt + D_1 Y(x_{t-},y)\,J_{t-}(\mu - \nu)\,(dy,dt),$$

so (A) holds with $W_1 = W_2 = 0$, $k = 2$, $d_1 = d$, $d_2 = d^2$ by (2.8) and Lemma 1.1.

Now we show why (A), (B) suffice to justify (2.15) with $q = \mathcal{A}_\alpha 1$ ($|\alpha| = n - 1$, $n \le N$) whenever $t \ge n\tau$ and, on choosing τ as close as necessary to $\frac{p}{\lambda}$,

$$\mathbb{E}[\,|(\mathcal{A}_\alpha 1)(z_t^{(N)})|^p] \le C_t(1+|x_0|^\beta), \quad |\alpha| = N$$

whenever $Np < \lambda t$, C_t taking the form $\theta(t)\, t^{-\gamma Np}$ in the γ-super-regularizing case. From this the result will follow on setting

$$q_t^\alpha(x_0,y) = \mathbb{E}[(\mathcal{A}_\alpha 1)(z_t^{(N)})\,|\,x_t = y]$$

The main point to note, and the justification for the rigmarole of N different perturbations, is that, to integrate by parts after some time $(n-1)\,\tau$ say (i.e. using a perturbation vanishing for $0 \le t < (n-1)\,\tau$), we only need conditions (1.9), (1.10) of Theorem 1.2 to hold for $t \ge (n-1)\,\tau$. It would be easy to extend the proof of Theorem 1.2 to cover this case. On the other hand, since μ has independent increments and $z_t^{(n)}$ is Markov, it is also a *consequence* of Theorem 1.2 (by conditioning on $\mathcal{F}_{(n-1)\tau}$). It is easy to check from (A) that (after $(n-1)\,\tau$) (1.9), (1.10) are satisfied by the SDE for $(z_t^{(n-1)}, H_t^{(n)})$ and the perturbation $v^{(n)}$. Thus the integration by parts formula (2.14) is valid for $z_t = (z_t^{(n-1)}, H_t^{(n)})$ and for all test functions f. Of course $f(x_t) J_t^{-1} (H_t^{(n)})^{-1} (\mathcal{A}_\alpha 1)(z_t^{(n-1)})$, $|\alpha| = n-1$ is not a test function of $(z_t^{(n-1)}, H_t^{(n)})$. However we can show for all $1 \le p < \infty$,

- $\mathbb{E}[\,|J_t^{-1}|^p] \le \theta(t)\,(1+|x_0|^\beta)$;
- if h is λ-regularizing and $\frac{Np}{\lambda} < N\tau \le t$, then $(H_t^{(n)})^{-p} \in L^1(\mathbb{P})$ for $n = 1,\dots,N$;
- if h is γ-super-regularizing and $\tau = \frac{t}{N}$, $\mathbb{E}[(H_t^{(n)})^{-p}] \le C\,(\frac{t}{N})^{-\gamma p}$;
- $\mathbb{E}[(D^{(n)} H_t^{(n)}/H_t^{(n)})^p] \le \theta(t)\,(1+|x_0|^\beta)$.

Thus (B) and the independence of $H_t^{(1)},\dots,H_t^{(n)}$ show that $(\mathcal{A}_\alpha 1)(z_t^{(n)}) \in L^1(\mathbb{P})$, for $|\alpha| = n$. Approximating $(z_t^{(n-1)}, H_t^{(n)}) \mapsto f(x_t)\, J_t^{-1} (H_t^{(n)})^{-1} (\mathcal{A}_\alpha 1)(z_t^{(n-1)})$ by test functions ($|\alpha| = n-1$) and passing to the limit we obtain the desired formulae.

Finally, assuming (A) and (B) above hold for $n-1$ we must show they hold for n. We have the following SDE's for $z_t^{(n)} = (z_t^{(n-1)}, D^{(n)} z_t^{(n-1)}, H_t^{(n)}, D^{(n)} H_t^{(n)}, R_t^{(n)})$

$$dz_t^{(n-1)} = Z(z_{t-}^{(n-1)})\,dt + W(z_{t-}^{(n-1)}, y)\,(\mu - \nu)(dy,dt)$$
$$+ W_1(t, z_{t-}^{(n-1)}, y)\,(\mu-\nu)(dy,dt) + W_2(t, z_{t-}^{(n-1)}, y)\,\mu(dy,dt),$$

$$dD^{(n)} z_t^{(n-1)} = DZ(z_{t-}^{(n-1)})\,D^{(n)} z_{t-}^{(n-1)}\,dt + D_1 W(z_{t-}^{(n-1)}, y)\,D^{(n)} z_{t-}^{(n-1)}\,(\mu-\nu)(dy,dt)$$
$$+ D_2 W(z_{t-}^{(n-1)}, y)\,v^{(n)}(t,y)\,\mu(dy,dt),$$

$$dH_t^{(n)} = h(y)\,1_{((n-1)\tau, n\tau]}(t)\,\mu(dy,dt),$$

$$dD^{(n)}H_t^{(n)} = Dh(y)\, v^{(n)}(t,y)\, \mu(dy,dt),$$

$$dR_t^{(n)} = \frac{\operatorname{div}(g.v^{(n)})(t,y)}{g(y)}\, (\mu - \nu)(dy,dt).$$

Note

(i) $v^{(n)}(\omega,t,y)$ may be written $v^{(n)}(t,x_{t-},J_{t-},y)$ where

$$v^{(n)}(t,x,J,y) = D_2Y(x,y)\,(I + D_1Y(x,y))\,J\,h(y)\,1_{((n-1)\tau,n\tau]}\,;$$

(ii) all terms in W_1, W_2 disappear from the SDE for $D^{(n)} z_t^{(n-1)}$ because $D^{(n)} z_0^{(n-1)} = 0$ and $v^{(n)}(t,y) = 0$ for $0 \le t \le (n-1)\tau$ (this feature of the derived processes explains best why the use of disjoint perturbations is a useful trick).

It is now routine to check from our hypotheses and the known graded structure of the coefficients Z, W, W_1, W_2 that (A) holds for n. In particular Lemma 1.1 gives the existence of a unique solution $z_t^{(n)}$ with the stated $L^p(\mathbb{P})$ bound. Suppose (B) holds for $n-1$). We inspect the formula (2.16) for $\mathcal{A}^{(n)}$. The only term potentially not of the right form is the third: if $g = \mathcal{A}_\alpha 1$ $(|\alpha| = n-1)$ Dq will introduce terms multiplied by $(H_t^{(1)})^{-2}$ (for example) rather than $(H_t^{(1)})^{-1}$. But $D^{(n)} H_t^{(m)} = 0$ for $m < n$ so these terms do not figure in $\mathcal{A}^{(n)} \mathcal{A}_\alpha 1$. This completes the inductive step. ◊

Theorem 2.9

Suppose (g,E') is λ-regularizing $(0 < \lambda \le \infty)$ (in particular γ-super-regularizing will do) and that X, Y satisfy (2.8) (i) (ii) (iii). Then for all $t > 0$ and all multi-indices α of length $N < \lambda t$ there exists a Borel function $\tilde{q}_t^\alpha$ on $\mathbb{R}^d \times \mathbb{R}^d$ such that

$$D_x^\alpha P_t(x,dy) = \tilde{q}_t^\alpha(x,y)\, P_t(x,dy)\ , \quad x,y \in \mathbb{R}^d$$

$$\int_{\mathbb{R}^d} |\tilde{q}_t^\alpha(x,y)|^p\, P_t(x,dy) \le C_{t,N,p}\,(1 + |x|^\beta) \tag{2.18}$$

whenever $Np < \lambda t$. In the γ-super-regularizing case we can take $C_{t,N,p} = \theta(t)\, t^{-\gamma Np}$ with θ increasing.

Suppose now that (g,E') is smoothly super-regularizing with regularizing function h and that X, Y satisfy only (2.9) (i) (ii) but that for some open set $U \subseteq \mathbb{R}^d$

$$|Y(x,y)|^2 \le h(y)\,,$$

$$|D_2Y(x,y)^{-1}| \le C(1 + |x|^\beta)$$

for all $x \in U$, $y \in \mathbb{R}^d$. Then for all $t > 0$ and all multi-indices α of length N there exists a Borel function $\tilde{q}_t^\alpha$ on $U \times \mathbb{R}^d$ such that

$$D_x^\alpha P_t(x,dy) = \tilde{q}_t^\alpha(x,y)\, P_t(x,dy) \ , \quad x \in U, \ y \in \mathbb{R}^d$$

$$\int_{\mathbb{R}^d} |\, \tilde{q}_t^\alpha \,(x,y)\,|^{\,p}\, P_t(x,dy) \le \theta(t,\delta)\, t^{-\gamma Np}\, (1 + |x|^\beta)$$

for $x \in U^\delta \equiv \{x \in U : \text{dist}\,(x, \partial U) > \delta\}$, where θ is increasing in t.

Sketch proof

For an autonomous SDE

$$dz_t = Z(z_{t-})\, dt + W(z_t,y)\, (\mu - \nu)\, (dy,dt)$$

with initial condition depending smoothly on x_0 $(z_0 = \psi(x_0))$ we know there exists a version $z_t = \psi_t(x_0)$ of the family of solutions as x_0 ranges over $\mathbb{R}^d$ depending smoothly on x_0, a.s. and in $L^2(\mathbb{P})$ (under appropriate conditions). Suppose inductively we have for multi-indices α of length n, such a $z_t^{(n)}$ and functions q^α with

$$D^\alpha P_t f\,(x) = \mathbb{E}[f \cdot \phi_t(x) . q^\alpha \cdot \psi_t^{(n)}(x)]$$

ϕ_t being the flow associated with $x_t = \phi_t(x_0)$. Differentiating

$$DD^\alpha P_t f\,(x) = \mathbb{E}[Df \cdot \phi_t(x)\, .\, D\phi_t(x) . q^\alpha \cdot \psi_t^{(n)}(x)]$$

$$+\ \mathbb{E}[f \cdot \phi_t(x)\, Dq^\alpha \cdot \psi_t^{(n)}\,(x) D\psi_t^{(n)}\,(x)]\,.$$

We now apply integration by parts with the perturbation of Theorem 2.5 to the first term on the right: taking $x_0 = x$, the formula reads

$$\mathbb{E}\,[Df\,(x_t)\, J_t\, q^\alpha(z_t^{(n)})] \tag{2.19}$$

$$= \mathbb{E}\,[f\,(x_t)\{H_t^{-2} DH_t\, q^\alpha(z_t^{(n)}) - H_t^{-1} Dq^\alpha(z_t^{(n)})\, Dz_t^{(n)} - H_t^{-1} q^\alpha(z_t^{(n)})\, R_t\}]$$

Thus if we set

$$z_t^{(n+1)} = (z_t^{(n)}, Dz_t^{(n)}, D\psi_t^{(n)}\,(x_0)) \tag{2.20}$$

(assuming H_t, R_t, x_t, J_t already to be among the components of $z_t^{(n)}$) then $z_t^{(n+1)}$ satisfies an autonomous SDE and writing $\psi_t^{(n+1)}$ for the associated flow we get expressions

$$D^\alpha P_t f\,(x) = \mathbb{E}[f \cdot \phi_t(x)\, q^\alpha \cdot \psi_t^{(n+1)}\,(x)]$$

for multi-indices α of length $n+1$. With hindsight we must begin with $z_t^{(0)} = (x_t, J_t, H_t)$ $z_t^{(1)} = (z_t^{(0)}, Dz_t^{(0)}, D\psi_t^{(0)}(x_0), R_t)$, and use (2.20) for $n \geq 1$. To carry out a detailed proof would require an analysis of the nature of the inductively defined SDE's for $z_t^{(n)}$ and the functions q^{α}: see for example the proofs of Theorems 2.5 and 2.8. This we omit. It is clear by a comparison of the SDE's for Dz_t and $D\psi_t(x_0)$ that the processes $z_t^{(n)}$, $n \geq 0$ defined above will behave no worse than those in Theorems 2.5, 2.8, so the same hypotheses will do. Indeed we are able to drop the hypothesis (2.8) / (2.9) (iv)

$$|(I + D_1 Y(x,y))^{-1}| \leq C$$

completely because there is no need to 'remove' J_t from the left side of (2.19), so no mention is necessary of J_t^{-1}. The crucial factor in $q^{\alpha}(z_t^{(n)})$ is still H_t^{-n} (or $\prod_{k=1}^{n} (H_t^{(k)})^{-1}$ if disjointly supported perturbations are used) which explains the similarity of the form of this theorem to earlier results. ◊

Notice that without condition (2.8) / (2.9) (iv) the local version of this theorem is more satisfactory than Theorem 2.6, in that it does without any global non-degeneracy condition.

2.f Regularity of the transition function

We show how the estimates of 2.e for the forward variable translate into regularity properties. First here is a summary of Theorems 2.6, 2.7, 2.8. Consider the following conditions on (g, E') (see 2.b)

(A) (g, E') is λ-regularizing $(0 < \lambda \leq \infty)$,

(B) (g, E') is γ-super-regularizing $(0 \leq \gamma < \infty)$,

(C) (g, E') is smoothly λ-regularizing,

(D) (g, E') is smoothly γ-super-regularizing with regularizing function h

(we have $(D) \Rightarrow (C) \Rightarrow (A)$, $(D) \Rightarrow (B) \Rightarrow (A)$), and the following conditions on X, Y (see 2.e)

(I) X, Y satisfy (2.8),

(II) X, Y satisfy (2.9) (i) (ii) (iii) and there is a $\phi \in C^{\infty}(E', [0,1])$ with $G(\text{supp}\,\phi) < \infty$, $\phi(y) \leq \text{dist}(y, \partial E')$, $\dfrac{D(g.\phi)}{g} \in C_b^{N-1}(E')$ and such that, setting $\phi \equiv 0$ off E', $\psi \equiv 1 - \phi$, we have, for all $x \in \mathbb{R}^d$, $y \in E$,

$$|(I + \psi(y) D_1 Y(x,y))^{-1}| \le C ,$$

(III) X,Y satisfy (2.9) (i) (ii) (iv) and there is an open set $U \subseteq \mathbb{R}^d$ such that for all $x \in U, y \in E$,

$$|Y(x,y)|^2 \le h(y) ,$$

$$|D_2 Y(x,y)^{-1}| \le C(1 + |x|^\beta) .$$

We showed that (AI) or (CII) or (DIII) imply for $|\alpha| = N$ the existence of a Borel function $q_t^\alpha(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ such that

$$D_y^\alpha P_t(x,dy) = q_t^\alpha(x,y) P_t(x,dy) ,$$

$$\int_{\mathbb{R}^d} |q_t^\alpha(x,y)|^p P_t(x,dy) \le C_t(1 + |x|^\beta) \tag{2.21}$$

subject to the restrictions

$Np < \lambda t$ under (A) and (C),

$x \in U^\delta \equiv \{x \in U : \text{dist}(x, \partial U) > \delta\}$ under (III)

and with the extra information under (B) or (D)) that

$C_t = \theta(t)\, t^{-\gamma Np}$ with θ increasing in t.

(θ may change below but it will remain an increasing function of t.)

Fix an integer $s \ge 0$ and a multi-index α of length $n = N - d - 1$ and consider the distribution

$$\mu(dy) = (1 + |y|^2)^s D_y^\alpha P_t(x,dy) .$$

If α' is a multi-index of length $d + 1$ we have by (2.21), for some polynomials $k_{\alpha''}$,

$$D^{\alpha'} \mu(dy) = \sum_{|\alpha''| \le d+1} k_{\alpha''}(y)\, q_t^{(\alpha'',\alpha)}(x,y)\, P_t(x,dy) .$$

Set

$$\hat{\mu}(u) = \int_{\mathbb{R}^d} e^{-i<u,y>} \mu(dy)$$

then by (2.21) (provided $N < \lambda t$ under (A) and (C))

$$|u^{\alpha'}| \, . \, |\hat{\mu}(u)| = |(D^{\alpha'}\mu)\hat{}\,(u)| \le C_t(1 + |x|^{\beta})$$

(with $C_t = \theta(t)\, t^{-\gamma(n+d+1+\varepsilon)}$ under (B) or (D)). Since α' was arbitary, $\hat{\mu} \in L^1(\mathbb{R}^d)$ and by Fourier inversion there is a function $f \in C_b(\mathbb{R}^d)$ such that

$$\mu(dy) = f(y)\, dy, \qquad \|f\|_{\infty} \le C_t(1 + |x|^{\beta})$$

We have proved.

Theorem 2.10

Suppose (AI) or (BI) or (CII) or (DII) or (DIII) hold, and $|\alpha| = n = N - d - 1 \ge 0$ with

$$0 \le n < \lambda t - d - 1, \qquad \text{in cases (A) (C)},$$

$$x \in U^{\delta} \quad \text{for some } \delta > 0, \quad \text{in case (III)}.$$

Then $P_t(x, \bullet)$ has a density $p_t(x, \bullet) \in C_b^n(\mathbb{R}^d)$ with

$$|D_y^{\alpha}\, p_t(x,y)| \le \frac{C_t(1 + |x|^{\beta})}{(1 + |y|^{s})}$$

and in cases (B), (D) we may take

$$C_t = \theta(t)\, t^{-\gamma(n+d+1+\varepsilon)}$$

with θ an increasing function of t, for any $\varepsilon > 0$. ◊

2.g How good are the results?

We test the power of our results in one dimension against a simple analysis of the Fourier transform

$$\hat{P}_t(u) = \mathbb{E}(e^{-iuy_t}) = \exp\left\{ -t \int_0^{\infty} (1 - e^{-iuy} - iuy)\, G(dy) \right\}$$

which is available in the case of the independent increment process with, for simplicity, positive jumps

$$y_t = \int_0^t \int_0^\infty y\,(\mu - \nu)\,(dy,ds)\,.$$

We assume $\int_0^\infty y^p G(dy) < \infty$ for all $2 \le p < \infty$.

Case 1: G admits a decomposition

$$G(dy) = \frac{\lambda}{y}\, 1_{\{0<y<\delta\}}\, dy + \tilde{G}(dy)$$

for some positive measure $\tilde{G}$. Then

$$|\hat{P}_t(u)| = \exp\left\{-t \int_0^\infty (1-\cos uy)\, G(dy)\right\} \le \exp\left\{-t \int_{\frac{\pi}{2|u|}}^{\delta} \frac{\lambda}{y}\, dy\right\}$$

$$= \left(\frac{\pi}{2\delta}\right)^{\lambda t} |u|^{-\lambda t}\,.$$

So y_t has a $C_b^n(\mathbb{R})$ density whenever $0 \le n < \lambda t - 1$. On the other hand we could apply Theorem 2.10, with $E' = (0,\delta)$, $g(y) = \lambda/y$ and regarding $\tilde{G}$ as living on a disjoint copy of $\mathbb{R}^+$. We showed in Lemma 2.3 that $(\lambda/y, \{0 < y < \delta\})$ is λ'-regularizing for all $\lambda' < \lambda$. The remaining conditions of Theorem 2.10 are easy to check and we find that y_t has a $C_b^n(\mathbb{R})$ density whenever $0 \le n < \lambda t - 2$. The rate of regularization is correct but one degree of differentiability has been lost - because one can only integrate by parts a whole number of times.

Case 2: G admits a decomposition

$$G(dy) = \frac{1}{y^{\alpha+1}}\, 1_{\{0<y<\delta\}}\, dy + \tilde{G}(dy)$$

for some $\alpha > 0$ and some positive measure $\tilde{G}$. Then

$$|\hat{P}_t(u)| \le \exp\left\{-t \int_{\frac{\pi}{2|u|}}^{\delta} \frac{1}{y^{\alpha+1}}\, dy\right\} = \theta(t)\, e^{-t|u|^\alpha}$$

So y_t has a C^∞ density $p_t(y)$ for all $t > 0$. Moreover

$$|D^n p_t(y)| \le \theta(t) \int_{\mathbb{R}} |u|^n e^{-t|u|^\alpha} du$$

$$= \theta(t)\, t^{-\frac{1}{\alpha}(n+1)}$$

On the other hand we could apply Theorem 2.10. By Lemma 2.3, $(y^{-(\alpha+1)}, \{0 < y < \delta\})$ is γ-super-regularizing for all $\gamma > \frac{1}{2} + \frac{1}{\alpha}$. By Theorem 2.10, y_t then has a C^∞ density for all $t > 0$ with

$$|D^n p_t(y)| \le \theta(t)\, t^{-\gamma(n+2)} \tag{2.22}$$

The extra $\frac{1}{2}$ in γ was thrown away on disregarding the behaviour as $t \to 0$ of $\mathbb{E}[|(\mathcal{A}_\alpha 1)(z_t^{(N)})|]$ (for example) in the proof of Theorem 2.8, and may be recoverable (see 2.h). The change from $n + 1$ to $n + 2$ is inevitable as discussed in Case 1.

2.h Locally stable processes (after Bass and Cranston)

We recover a result of [1] that the 'locally stable' Markov process x_t, with generator

$$\mathcal{G}f(x) = \frac{1}{\alpha(x)} \int_0^1 \{f(x+y) - f(x) - Df(x).y\}\, y^{-(1+\frac{1}{\alpha(x)})}\, dy\,,$$

has a resolvent density provided

(i) $\frac{1}{2} < \alpha' \le \alpha(x) \le \alpha'' < 1$ for all $x \in \mathbb{R}$, for some α', α'',

(ii) $\alpha \in C_b^2(\mathbb{R})$.

In fact we prove a stronger result that if $\alpha'' + \frac{1}{2} < \gamma < \frac{3}{2}$ then, for all λ sufficiently large and $f \in C_b^1(\mathbb{R})$,

$$\mathbb{E} \int_0^\infty e^{-\lambda t} Df(x_t)\, dt \le C \lambda^{-\frac{3}{2}+\gamma} \|f\|_\infty\,. \tag{2.23}$$

Take $E' = (1,\infty)$, $G(dy) = g(y)\, dy = 1_{\{y>1\}}\, dy$, $Y(x,y) = y^{-\alpha(x)}\, 1_{\{y>1\}}$, then

$$\mathcal{G}f(x) = \int_1^\infty \{f(x + Y(x,y)) - f(x) - Df(x)\,.\,Y(x,y)\}\, dy\,.$$

So the process may be realised by the SDE

$$dx_t = Y(x_{t-},y)\,(\mu - \nu)\,(dy,dt)\,.$$

Take $\phi \in C^1(\mathbb{R})$ with $\phi(x) = x \vee (-\frac{1}{2})$ except in a small neighbourhood of $-\frac{1}{2}$, take $\psi(x) = x - \phi(x)$ and $h \in C_b^1(1,\infty)$ with $h \equiv 0$ near 1, $h(y) = y^{-\gamma}$ near ∞. The hypotheses of Theorem 1.2 are satisfied by the graded system

$$dx_t = Y(x_{t-},y)\,(\mu - \nu)\,(dy,dt)\,,$$

$$dK_t = S(x_{t-})\,K_{t-}\,dt + \phi(D_1 Y(x_{t-},y))\,K_{t-}(\mu - \nu)\,(dy,dt)\,,$$

$$dH_t = h(y)\,(\mu - \nu)\,(dy,dt)\,,$$

$$(x_0,K_0,H_0) = (x_0,1,0)\,,$$

where

$$S(x) = \int_E \psi(D_1 Y(x,y))\,dy\,,$$

and by the perturbation

$$v(t,y) = \frac{(1 + \phi(D_1 Y(x_{t-},y)))\,K_{t-}h(y) - \psi(D_1 Y(x_{t-},y))\,K_{t-}\,H_{t-}}{D_2 Y(x_{t-},y)}\,.$$

It is easily checked by substituting in the SDE for the derived process that $Dx_t = K_t H_t$. We arrange that $1 + \phi(D_1 Y(x,y)) \geq \dfrac{1}{2}$ for all x,y, so the SDE

$$dL_t = -L_{t-}\{S(x_{t-}) - \int_E (1 + \phi(D_1 Y(x_{t-},y)))^{-1}\,\phi(D_1 Y(x_{t-},y))^2\,G(dy)\}\,dt$$

$$-L_{t-}(1 + \phi(D_1 Y(x_{t-},y)))^{-1}\,\phi(D_1 Y(x_{t-},y))\,(\mu - \nu)\,(dy,dt)\,,$$

$$L_0 = 1\,,$$

has a unique solution with $L_t \in \bigcap_{p<\infty} L^p(\mathbb{P})$. Moreover $d(K_t L_t) = 0$, so $L_t = K_t^{-1}$. We have $G(h \geq \varepsilon) \sim \varepsilon^{-1/\gamma}$ as $\varepsilon \to 0$, so by Lemma 2.1

$$\mathbb{E}(H_t^{-p}) \leq C t^{-p\gamma} \qquad \text{for all } 0 < t \leq 1\,. \tag{2.24}$$

It follows by a limit argument that the function

$$(x,K,H) \mapsto f(x)\,K^{-1}H^{-1}$$

may be integrated by parts to give

$$\mathbb{E}[Df(x_t)] = \mathbb{E}[f(x_t)\,\{K_t^{-2}\frac{DK_t}{H_t} + K_t^{-1}\frac{DH_t}{H_t^2} - K_t^{-1}\frac{R_t}{H_t}\}], \tag{2.25}$$

where

$$dR_t = D_2 v(t,y)\,(\mu - v)\,(dy,dt)\,, \qquad R_0 = 0\,.$$

The proof of (2.23) is completed by finding suitable bounds in $L^1(\mathbb{P})$ for the three terms inside the curly brackets. There are constants $C,\ a < \infty$ such that for all $t \geq 0$

$$\mathbb{E}(K_t^{-4}\,|DK_t|^2 \leq Cte^{2at}$$

$$\mathbb{E}(K_t^{-2}\,R_t^2) \leq Cte^{2at}\,.$$

This together with (2.24) gives a bound of the form $Ct^{-\gamma+\frac{1}{2}}e^{at}$ for the first and third terms.

We consider the second term. There exists an increasing process $A(t)$ with

$$\mathbb{E}(A(t)^p) < C(p)\,e^{a(p)t}$$

for all $p < \infty$, such that

$$|Dh(y)\,v(t,y)| \leq A(t)\,y^{-2\gamma+\alpha''}\,.$$

So $|DH_t| \leq A(t)\,Q_t$ where

$$dQ_t = k(y)\,\mu\,(dy,dt)\,, \quad Q_0 = 0$$

and $k(y) = y^{-2\gamma+\alpha''}$. Fix q with $1 < q < \dfrac{2\gamma-\alpha''}{\gamma+\frac{1}{2}}$. We have

$$\mathbb{E}(Q_t H_t^{-2q}) = \Gamma(p)t\int_0^\infty \beta^{2q-1}\,e^{-\xi(\beta)t}\,\eta(\beta)\,d\beta$$

where

$$\xi(\beta) = \int_1^\infty (1 - e^{-\beta h(y)})\,dy\,,$$

$$\eta(\beta) = \int_1^\infty k(y)\,e^{-\beta h(y)}\,dy\,.$$

But $\xi(\beta) \geq \delta\beta^{\frac{1}{\gamma}} - c$ for some $\delta, c < \infty$ (see proof of Lemma 2.1) and $\eta(\beta) \leq \delta' \beta^{-2+\frac{1}{\gamma}(\alpha''+1)} + c'$ for some $\delta', c' < \infty$. So

$$\mathbb{E}(Q_t H_t^{-2q}) \leq \Gamma(p) t \int_0^\infty \beta^{2q-1} (\delta' \beta^{-2+\frac{1}{\gamma}(\alpha''+1)} + c') e^{-\delta\beta^{\frac{1}{\gamma}} t + ct} \, d\beta$$

so by a substitution $\beta' = t^\gamma \beta$ we get

$$\mathbb{E}(Q_t H_t^{-2q})^{1/q} \leq C e^{at} t^{-2\gamma + \frac{1}{q}(2\gamma - \alpha'')}$$

$$\leq C e^{at} t^{-2\gamma + \gamma + \frac{1}{2}}$$

by our choice of q. Hence, if $\frac{1}{p} + \frac{1}{q} = 1$

$$\mathbb{E}(|K_t^{-1} \frac{DH_t}{H_t^2}|) \leq \mathbb{E}(|K_t^{-1}|^p A(t)^p Q_t)^{1/p} \, \mathbb{E}\left(\frac{Q_t}{H_t^{2q}}\right)^{1/q}$$

$$\leq C e^{at} t^{-\gamma + 1/2}$$

for some C, $a < \infty$. Finally for $\lambda > a$ by (2.25)

$$\mathbb{E}\int_0^\infty e^{-\lambda t} Df(x_t) \, dt \leq 3C \|f\|_\infty \int_0^\infty e^{-(\lambda - a)t} t^{-\gamma + 1/2} \, dt$$

$$\leq 3C \|f\| (\lambda - a)^{-\gamma + \frac{3}{2}} \int_0^\infty e^{-s} s^{-\gamma + \frac{1}{2}} \, ds < \infty .$$

References

1. R.F. BASS and M. CRANSTON: The Malliavin calculus for pure jump processes and applications to local time, Ann. Prob. 14, 490-532, 1986.
2. K. BICHTELER and J. JACOD: Calcul de Malliavin pour les processus avec sauts, existence d'une densité dans le cas uni-dimensionnel, *Séminaire de Probabilités XVII*, Lecture Notes in Mathematics 986, Springer, Berlin, 132-157, 1983.
3. K. BICHTELER, J.B. GRAVEREAUX and J. JACOD: *Malliavin calculus for processes with jumps*, Gordon and Breach, 1987.
4. J.M. BISMUT: Calcul des variations stochastiques et processus de sauts. Z. Wahrs. 63, 147-235, 1983.
5. J.M. BISMUT: Jump processes and boundary processes, *Proc. Taniguchi Symp. on Stochastic Analysis 1982*, ed. K. Itô, North-Holland, Amsterdam, 1984.
6. J. JACOD: *Calcul stochastique et problèmes de martingales*, Lecture Notes in Mathematics 714, Springer, Berlin, 1979.
7. R. LEANDRE: Thèse de 3ème cycle, Besançon, 1984.
8. R. LEANDRE: Régularité des processus de sauts dégénérés, Ann. Inst. H. Poincaré 21, 125-146, 1985.
9. R. LEANDRE: Calcul des variations sur un brownien subordonné, dans ce volume.
10. R. LEANDRE: Densité en temps petit d'un processus de sauts, preprint.
11. J.R. NORRIS: D. Phil thesis, Oxford, 1985.
12. J.R. NORRIS: Simplified Malliavin calculus, *Séminaire de Probabilités XX*, Lecture Notes in Mathematics 1204, Springer, Berlin, 101-130, 1986.

J.R. Norris

Cambridge University
Department of Pure Mathematics and
Mathematical Statistics
16 Mill Lane
Cambridge CB2 1SB