Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 1–11

Backward Regularity for some Infinite Dimensional Hypoelliptic Semi-groups

Hélène Airault and Paul Malliavin

CONTENTS

We dedicate this work to Kiyosi Itô, the "Newton of continuous Stochastic Dynamic", one of the most influential scholars of the last century; the topic of this paper underlines in itsef the deep influence that the 1976 Kyoto Symposium [12] has had on the whole subsequent carrier of the second author who is also deeply indebted to Kiyosi Itô for fifty years of warm personnal relations; his attentive support from the beginning to some of our scientific enterprises has been a key step towards their international recognition.

In classical Stochastic Analysis regularity properties are time independent : the Brownian motion is for all time Hölderian of order $(\frac{1}{2}-\epsilon)$ regular, the tangent space to the Wiener space (i.e. the Cameron-Martin space) is also time independent. The Stochastic Analysis on Loop groups have recently confirmed the paradigm that regularity properties are time independent.

It has been a surprise that regularity exponents for highly non linear infinite dimensionnal diffusion as the canonic diffusion above Virasoro algebra are time dependent [2],[9]. We shall discuss in this paper the status of tangent space to Virasoro diffusion; we shall exhibit a minimal tangent space which is time independent; it is conceivable that the maximal tangent space is time dependent, fact which will be established on a toy model. The finite dimensional root of of this phenomen lies in the fact that hypoelliptic diffusion on R^d does not satisfy simple scaling relation when the time goes to zero [4], [11].

Stability of interest models in Mathematical Finance are deeply affected by these infinite dimensional effects.

Received April 24, 2003.

1. Regularity of the canonical diffusion above Virasoro algebra.

The group of C^∞ diffeomorphism of the circle S^1, Diff(S^1), has for Lie algebra diff(S^1) the C^∞ vector fields on S^1; we identify a function $u(\theta)$ to the vector field $u(\theta)\frac{d}{d\theta}$; with this identication the bracket of vector fields becomes $[u,v] = \dot{v}u - \dot{u}v$. Complexifying the underlying real vector space we get the following expression for this bracket in the complex trigonometric basis :

$$[e^{in\theta}, e^{im\theta}] = i(m-n)e^{i(m+n)\theta}$$

Given a positive constant $c > 0$, define the bilinear antisymmetric form

$$\omega_c(f,g) := -\frac{c}{12}\int_{S^1} (f' + f^{(3)})g \, d\theta;$$

then

$$\omega_c([f_1,f_2],f_3) + \omega_c([f_2,f_3],f_1) + \omega_c([f_3,f_1],f_2) = 0,$$

$$\omega_c(e^{in\theta}, e^{-im\theta}) = i\delta_n^m \, \frac{c}{6}(n^3 - n), \quad n > 0.$$

Virasoro algebra is defined as $\mathcal{V}_c := R \oplus \text{diff}(S^1)$ with the following bracket :

$$[\xi\kappa + f, \eta\kappa + g] := \omega_c(f,g)\kappa + [f,g].$$

Brownian motion on Diff $(\mathbf{S^1})$.

Define the Hilbertian metric $\frac{3}{2}$ by :

$$\|\phi\|^2_{\mathcal{H}^{\frac{3}{2}}} = \sum_{n>1}(n^3-n)\,(a_n^2+b_n^2), \quad \phi(\theta) = \frac{1}{2}a_0 + \sum_{n=1}^{+\infty}(a_n\cos n\theta + b_n \sin n\theta);$$

define

$$e_n \; : \; R^2 \mapsto \text{diff}(S^1), \quad e_n(\xi) = \frac{1}{\sqrt{n^3-n}}(\xi^1\cos n\theta + \xi^2 \sin n\theta), \quad n > 1.$$

Let X_k be independent copies of Wiener space of the R^2-valued Brownian motion; define $X = \bigotimes X_k$ and consider the Stratonovich SDE :

$$dg_x(t) = \left(\sum_{k>1} e_k(dx_k(t))\right) \circ g_x(t), \quad g_x(0) = \text{Identity}$$

$$dg_x^r(t) = \left(\sum_{k>1} r^k \; e_k(dx_k(t))\right) o \; g_x^r(t), \;\; g_x^r(0) = \text{Identity};$$

then, $g_x^r(t) \in$Diff (S^1) $\forall r < 1$.
Theorem.[2],[9].

Denote $\mathcal{H}^\beta(S^1)$ the group of homemorphism of S^1, with an Hölderian modulus of continuity β, then

$$\lim_{r\to 1} g_x^r(t) := g_x(t) \in \mathcal{H}^{\beta(t)}(S^1), \;\; a.s.,$$

$$\beta(t) = \frac{1-\sqrt{1-e^{-\frac{t}{2}}}}{1+\sqrt{1-e^{-\frac{t}{2}}}}.$$

The laws ν_t of $g_x(t)$ satisfy $\nu_t * \nu_{t'} = \nu_{t+t'}$.
Remark. The composition of two homemorphisms of Hölderian exponents γ, γ' can have an Hölderian exponent as worst as $\gamma\gamma'$: this fact explains the exponential decrease of $\beta(t)$ when $\to +\infty$.

It is obvious that the metric used to construct the Brownian motion degenerates on the vector fields $\cos\theta$, $\sin\theta$, 1. The Lie subagebra generated by these three vector fields is isomorphic to sl$(2, R)$; the corresponding subgroup Γ of Diff(S^1) is the restriction to the circle of the group of Möbius transformations of the unit disk.

It had be shown [1] that $\mathcal{M}_1$:=Diff$(S^1)/\Gamma$ is an homogeneous Riemannian manifold, that the Hilbert transform on the circle pass to the quotient and defines an integrable almost complex structure for which $\mathcal{M}_1$ becomes an homogeneous Kähler manifold. Denote π : Diff$(S^1) \to \mathcal{M}_1$, then $\pi(g_x^{-1}(t))$ is the Brownian motion on $\mathcal{M}_1$ and defines the heat semi-group on function on $\mathcal{M}_1$. This section will prove the backward regularity of this heat semi-group.

Background of finite dimensional Stochastic Riemannian Geometry.

Denote by M a Riemannian manifold of dimension d; a *frame* r is a Euclidean isomorphism of R^d onto the tangent plane $T_{\pi(r)}(M)$; the collection of all frames on M is a smooth manifold $O(M)$ on which the orthogonal group operates on the right : this is the *bundle of orthonormal frames.* The Levi-Civita connection defines on $O(M)$ a *parallelism* that is a canonical differential form of degree 1, with values in $R^d \oplus R^d \otimes_a R^d$ let $\omega = (\dot\omega, \ddot\omega)$. Riemannian geometry is encompassed in the Darboux-Cartan structural equations :

$$< A \wedge B \; , \; d\dot\omega >= \ddot\omega(A)\dot\omega(B) - \ddot\omega(B)\dot\omega(A),$$

$$< A \wedge B \ , \ d\ddot{\omega} >= \ddot{\omega}(A)\ddot{\omega}(B) - \ddot{\omega}(B)\ddot{\omega}(A) + \Omega(\dot{\omega}(A), \dot{\omega}(B)),$$

where Ω is the Riemann curvature tensor.

Given an R^d valued brownian motion $x(\tau)$ the *horizontal diffusion* is defined by the Stratonovitch SDE

$$< dr_x, \dot{\omega} >= dx, \qquad < dr_x, \ddot{\omega} >= 0, \quad r_x(0) = r_0,$$

where $r_0 \in O(M)$ is fixed. The *Itô parallel transport* is the isometry

$t^x_{0\leftarrow\tau} : T_{\pi(r_x(\tau))}(M) \mapsto T_{\pi(r_0)}(M)$ defined by $t^x_{0\leftarrow\tau} = r_x(0) \ o \ (r_x(\tau))^{-1}$.

A variation induces $x \mapsto x + \epsilon\tilde{\zeta}$ induces a variation of the path (ζ, ρ) defined by

$$\zeta(\tau) :=< \frac{dr^\epsilon(\tau)}{d\epsilon_{=0}} \ , \ \dot{\omega} >, \quad \rho(\tau) :=< \frac{dr^\epsilon(\tau)}{d\epsilon_{=0}} \ , \ \ddot{\omega} >, \quad r^\epsilon(\tau) := r_{x+\epsilon}(\tau).$$

These two variations are linked by the two following key SDE [6], [10], [7], [14], the first being an Itô SDE, the second a Stratonovitch SDE :

$$(1.1) \qquad d\tilde{\zeta} = d\zeta - \frac{1}{2}\text{Ricci}(\zeta) \ d\tau - \rho \ dx, \quad d\rho = \Omega(\zeta, \ o \ dx).$$

Two parallel transports on $\mathcal{M}_1$.

We follow Bowick-Lahiri [5]. We have on $\mathcal{M}_1$ two connections : the Levi-Civita connection ∇_X and the connection $\mathcal{L}_X$ induced by the left invariant Maurer-Cartan form on Diff(S^1); we introduce a tensorial operator on $T_0(\mathcal{M}_1)$ defined by

$$\phi_X = \mathcal{L}_X - \nabla_X$$

The operator ϕ, extended to the complexification, has the following expression in the complex trigonometric basis :

$$(1.2) \qquad \phi_{e^{ir\theta}}(e^{iq\theta}) = i(r-q)\Theta(-q-r), \quad r > 1,$$

where $\Theta(t) := 1_{[0,+\infty[}$ is the Heaviside function. For $s < -1$ we prolongate ϕ_* by requiring hermitian symmetry : $\phi_{e^{is\theta}} := (\phi_{e^{-is\theta}})^*$.

Then the Riemannian curvature of $\mathcal{M}_1$ can be expressed in terms of the operator ϕ_* by

$$\Omega(X,Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]} = [\phi_X, \phi_Y] - \phi_{[X,Y]},$$

the last identity results from $[\mathcal{L}_X, \mathcal{L}_Y] - \mathcal{L}_{[X,Y]} = 0$ together with $[\mathcal{L}_X, \nabla_Y] = \nabla_{[X,Y]}$ identity coming from the invariance of the Kählerian metric under the left action of Diff (S^1). The curvature tensor is of trace class [5], and its trace is

$$\text{Ricci} = -\frac{13}{6} \times \text{Identity} \tag{1.3}$$

Lemma.

Denote V_q the space generated by $\cos k\theta$, $\sin k\theta$, $k \in [2, q]$ *then the operators ϕ_* preserve V_q and are nilpotent on V_q.*

Denote $\eta_n(\xi) = \tilde{\phi}_{e_n(\xi)}$, $\xi \in R^2$, where $\tilde{\phi}_X$ is the matrix associated to ϕ_X in the real trignometric basis $(n^3-n)^{-\frac{1}{2}} \cos n\theta$, $(n^3-n)^{-\frac{1}{2}} \sin n\theta$.
Theorem

The matrix Stratonovich SDE

$$d\mathcal{U}_t = \mathcal{U}_t \; o \left(-\sum_{k>1} \eta_k(dx_k(t)) \right), \quad \mathcal{U}_0 = \text{Identity} \tag{1.4}$$

has a unique solution and $\mathcal{U}_t$ is a unitary matrix.
Proof.

The restriction to V_q of this SDE is equivalent to an SDE which is driven only by $2q$ Brownian motion; this SDE which is solvable by the finite dimensional theory •

Backward regularity. (Minimal tangent space)
Theorem.

Given z such that $\|z\|_{H^{\frac{3}{2}}} < \infty$ then, for a generic test function Φ defined on $\mathcal{M}_1$,

$$\begin{aligned} &\left| \frac{d}{d\epsilon_{=0}} E((\pi^*\Phi)(\exp(\epsilon z) g_x(t))) \right|^2 \\ &\leq \frac{13}{6(1-\exp(-\frac{13}{6}t)} \; \|z\|^2_{H^{\frac{3}{2}}} \; E(|\pi^*\Phi(g_x(t))|^2). \end{aligned} \tag{1.5}$$

Proof.

We follow the strategy that Driver [8] developped in the case of Loop groups making the change of variables

$$y_t = \int_0^t \mathcal{U}_s \; dx(s);$$

then y_t is a new brownian motion to which we can apply the finite dimensional Riemannian geometry because the curvature operator preserves the V_q •

2. Infinite dimensional non autonomous Riemannian metrics.

Consider a group G of dimension finite or infinite; for instance G could be the group of diffeomorphism of a compact manifold, case which includes the theory of Stochastic Flows.

We consider a left invariant diffusion on G; denote by $\Delta = \frac{1}{2}\sum_{k\geq 1}\partial^2_{A_k} + \partial_{A_0}$ its infinitesimal operator where the A_k are left invariant vector field on G; denote by ∇ the corresponding gradient : $\nabla\phi * \nabla\psi := \Delta(\phi\psi) - \phi\Delta\psi - \psi\Delta\phi$.

We denote by $p_T(dg)$ the law of the process starting from the identity. Given a tangent vector at the identity z define the "logarithmic derivative" of p_T by the identity

$$\frac{d}{d\epsilon_{=0}} E(\Phi(\exp(\epsilon z)g_x(T)) = E(K_{z,T}(g_x(T))\Phi(g_x(T))), \tag{2.1}$$

where Φ is a generic test function.

For all $T > 0$ define a Hilbertian norm by

$$\|z\|_T^2 := E(|K_{z,T}(g_x(T))|^2)). \tag{2.2}$$

Theorem.

If $T < T'$ *then*

$$\|z\|_{T'} \leq \|z\|_T. \tag{2.3}$$

Proof.

For $\eta > 0$ define $\Psi(g) := E_{g_x(T)=g}(\Phi(g_x(T+\eta))$, then

$$\begin{aligned} E(\Phi(\exp(\epsilon z)g_x(T+\eta)) &= E(E^{\mathcal{N}_T}(\Phi(\exp(\epsilon z)g_x(T+\eta)) \\ &= E(\Psi(\exp(\epsilon z)g_x(T))); \end{aligned}$$

differentiating relatively to ϵ we obtain

$$E(K_{z,T+\eta}(g_x(T+\eta))\ \Phi(g_x(T+\eta))) = E(K_{z,T}(g_x(T))\Psi(g_x(T))),$$

letting $\eta \to 0$ we write $\simeq$ equalities modulo $o(\epsilon)$; then by Itô calculus :

$$K_{z,T+\eta}(g_x(T+\eta) - K_{z,T}(g_x(T) \simeq \eta(\frac{\partial K}{\partial T} + \Delta K) + \nabla K * (x(T+\eta) - x(T))$$

$$\begin{aligned} &\Psi(g) - \Phi(g) \simeq \eta\Delta\Phi(g), \\ &\Phi(g_x(T+\eta)) \simeq \Phi(g_x(T)) + \eta(\Delta\Phi(g_x(T))) + \nabla\Phi * (x(T+\eta) - x(T)) \end{aligned}$$

$$\frac{1}{\eta}E^{\mathcal{N}_T}(K_{z,T+\eta}(g_x(T+\eta))\ \Phi(g_x(T+\eta))) - (K_{z,T}(g_x(T))\Phi(g_x(T)))$$

$$\simeq \Phi(\frac{\partial K}{\partial T} + \Delta K) + K\Delta\Phi + \nabla\Phi * \nabla K;$$

$$\frac{1}{\eta}E\Big(K_{z,T+\eta}(g_x(T+\eta))\ \Phi(g_x(T+\eta))) - (K_{z,T}(g_x(T))\Phi(g_x(T))\Big)$$

$$\simeq \Phi(\frac{\partial K}{\partial T} + \Delta K) + K\Delta\Phi + \nabla\Phi * \nabla K) - K\Delta\Phi$$

$$E\Big(\Phi(\frac{\partial K}{\partial T} + \Delta(K)) + \nabla K * \nabla\Phi\Big) = 0. \tag{2.4}$$

From the other hand

$$\frac{\partial}{\partial T}E[(K_T(g))^2)] = E[\Delta(K_T^2) + \frac{\partial K_T^2}{\partial T}]$$

$$= E[2K_T(\frac{\partial K}{\partial T} + \Delta(K_T)) + \nabla K_T * \nabla K_T)] = -E[\nabla K_T * \nabla K_T] < 0,$$

the last equality is obtained by applying (2.4) with $\Phi = K_T$ •

Consider now the free Lie algebra $\mathcal{G}$ generated by d vector fields $A_1, \dots A_d$; denote G the infinite dimensional group associated. Denote x a d-dimensional Brownian motion and define on G the process by the following Stratanovitch SDE

$$dg_x(t) = g_x(t)\ o\ \sum_{k=1}^{d} A_k\ dx^k(t),\quad g_x(0) = \text{Identity}$$

denote $\mathcal{H}_T$ the completion of $\mathcal{G}$ for the norm $\|z\|_T$.
Theorem. For $T \neq T'$, we have
(2.5)
$\mathcal{H}_T \neq \mathcal{H}_{T'}$, which means the inequivalence of the corresponding norms.

Proof.

We shall use the Ben-Arous expansion [3] (see Theorem 15)

$$g_x(t) = \exp\Big(\sum_{m=1}^{\infty}\sum_{J\in\sigma_m} M_J(t)U^J\Big)$$

where $A^J := [A_{j_1}, [A_{j_2}, \dots, [A_{j_{n-1}}, A_{j_n}]$, where σ_m denotes a maximal subset of $[1,d]^m$ such that the A^J are linearly independent in $\mathcal{G}$ and where iterated integrals M_J have been constructed by Meyer and are mutually orthogonal in L^2. We decompose

$$z = \sum_{m=1}^{\infty} z_m, \quad z_m = \sum_{J \in \sigma_m} c_J A^J.$$

Lemma.

$$\|z\|_T^2 = \sum_{m=1}^{\infty} \|z_m\|_T^2 \tag{2.6}$$

By the rescaling of Meyer integrals we have

$$\|z_m\|_{T'}^2 = \left[\frac{T}{T'}\right]^m \|z_m\|_T^2$$

relation which shows the inequivalence of the two norms •

3. Instability of Heath-Jarrow-Merton model of interest rates.

All long terms loans (States bounds, mortgages, companies bounds) are appearing on a single market, the " zero coupon default free bonds market". Every day it is possible to buy bonds at any maturity between 1 up to 360 months; for each maturity the market gives a price; all these prices can be summarized by a single positive function $r_t(x)$ *the instantaneous forward rate* such that the discount price today of a 1 dollar bound paid in five years is equal to

$$\exp(-\int_0^{60} r_t(x)\, dx).$$

The associated configuration space $\mathcal{C}$ is $(R^+)^{360}$.

The HJM model replace the $\mathcal{C}$ by the space of continuous positive functions $r_t(x)$, $x \in [0, 360]$ and propose that "for the risk free measure" the interest rate curve dynamic can be described by the following Itô SDE, driven by q independent Brownian motion $W^*(t)$,

$$dr_t(x) = \left(\frac{\partial r_t(x)}{\partial x} + Z_t(x)\right) dt + \sum_{k=1}^{q} \phi_{k,t}(x)\, dW^k(t),$$

$$Z_t(x) = \sum_{k=1}^{q} \phi_{k,t}(x) \int_0^x \phi_{k,t}(s)\, ds. \tag{3.1}$$

This HJM modell can be mathematically established under the two general assumptions : market where an agent cannot increase his wealth without risk (*arbitrage free)* and market variations free from jumps.

A practical fact is that the variance injected in the equation is very low : $q \leq 4$. This means that the operator associated with the SDE (3.1) is an *hypoellitic operator driven by at most four vectors fields in a Euclidean space of large dimension.*

Consider the Stochastic flow $U^{W}_{t\leftarrow t_0}$ defined as $U^{W}_{t\leftarrow t_0}(r_0)$ being the solution of (3.1) for $r_W(t_0) = r_0$. Denote by $J^{W}_{t\leftarrow t_0}$ the Jacobian of the flow $U^{W}_{t\leftarrow t_0}$ which is defined by solving the linearized SDE.

Greeks means the reaction of the market at an infinitesimal pertubation δ_0 of r_0 appearing at time t_0, $W^*(s) - W^*(t_0)$, $s \geq t_0$ being fixed :

$$\frac{d}{d\epsilon_{\epsilon=0}} U^{W}_{t\leftarrow t_0}(r_0 + \epsilon\delta_0) = J^{W}_{t\leftarrow t_0}(\delta_0) := \delta^W(t),$$

is called the *Greek propagation.*

Every trader can buy or sell *european options* which is a contract by which the seller obliges himself to pay at maturity T an amount of money equal to $F(r_T)$. The option is called *digital* if the function F is discontinuous.

Sensitivities at the option F is defined

$$\frac{d}{d\epsilon_{\epsilon=0}} E(F(U^{W}_{T\leftarrow t_0}(r_0 + \epsilon\delta_0)) = E(< dF \ , \ J^{W}_{T\leftarrow t_0}(\delta_0) >).$$

Sensitivities regularization for digital european options

Denote $\mathcal{C}$ the vector space of all possible infinitesimal pertubation δ_0 of the market at time t_0; consider the Hilbertian norm $\|\delta\|_{T,t_0}$ defined in (2.2) and denote $\mathcal{C}_{t_0,T}$ the corresponding Hilbert space then

$$|\frac{d}{d\epsilon_{\epsilon=0}} E(F(U^{W}_{T\leftarrow t_0}(r_0 + \epsilon\delta))| \leq \|\delta\|_{T,t_0}(E(|F(r_W(T)|^2))^{\frac{1}{2}}$$

Compartimentage Principle.

"Generically" the sequence of Hilbert spaces $\mathcal{C}_{T,t_0}$ is strictly increasing relatively the parameter T and strictly decreasing relatively to the parameter t_0.

Hedging

The Clark-Ocone-Karatzas formula

(3.2)

$$F(r_W(T)) - E(F(r_W(t_0))) = \sum_{k=1}^{q} \int_{t_0}^{T} E^{\mathcal{F}_s}(D_{s,k}(F(r_W(T))) \, dW^k(s)$$

gives a realization of the option along each trajectory. The corresponding strategy of *replication,* consist for the trader to balance at each time t his portfolio according the infinitesimal observed variation of the driving Brownian $W^k(t+\epsilon) - W^k(t)$, multiply by $E^{\mathcal{F}_s}(D_{s,k}(F(r_W(T)))$.

The formula (3.2) is a specialization of the general Itô theorem saying that any random variable of zero expectation is representable by a Stochastic integral; at this level of generality the integrand is only in $L^2([t_0, T])$ on each trajectory. As the *financial replication* of the option is given by this integrand, it is impossible to realize this replication if this integrand is not at least continuous; otherwise instabilities appear.

3.3. Theorem [13].

Denote Θ the stopping time such that

$$J^W_{t_0 \leftarrow t}(\Phi_k(r_W(t))) \in \mathcal{C}_{T,t} \quad \forall\, t \leq \Theta, \quad \forall k \in [1, q];$$

then $E^{\mathcal{F}_\Theta}(F(r_W(T)))$ is replicable by a stable Clark-Ocone-Karatzas formula.

Proof.

$$E^{\mathcal{F}_s}(D_{s,k}(F(r_W(T))) = E(< dF\ ,\ J^W_{T \leftarrow s}(\Phi_k(r_W(s)) >) \quad \bullet$$

Consequence : Traders must try to sale digital options before the stopping time Θ.

BIBLIOGRAPHY

[1] Airault(H.), Malliavin(P.) and Thalmaier(A), *Support of Virasoro unitarizing measures : C.R. Acad.Sci. Paris Ser. I 335 (2002) 621-626.*

[2] Airault(H.) and Ren(J.), *Modulus of continuity of the canonic Brownian motion "on the group of diffeomorphism of the circle" : J. Funct. Analysis 196 (2002) 325-426.*

[3] Ben-Arous (G.), *Flots et séries de Taylor stochastiques : Prob. Theory Rel. Fields 81 (1989) 29-77.*

[4] Ben-Arous (G.) and Léandre (R.), *Décroissance exponentielle du noyau de la chaleur sur la diagonale : Prob. Th. Rel. Fields 90 (1991) 175-202 and 377-402.*

[5] Bowick(M.J.) and Lahiri(A.), *The Ricci curvature of* $\mathrm{Diff}(S^1)/\mathrm{SL}(2, R)$ *: J.Math. Phys. 29(1988) 1979-1981.*

[6] Bismut(J.M.), *Large deviations and the Malliavin Calculus : Birkhäuser, Prog. Math. 45 (1984) 216 pp.*

[7] Cruzeiro(A.B.) and Malliavin(P.), *Renormalized differential geometry on path spaces : Structural equation, curvature : J. Funct. Analysis 139 (1996) 119-181.*

[8] Driver (B), *Integration by parts and quasi-invariance for heat kernel measure on Loop groups : J.Funct.Analysis 149 (1997) 470-547* and *155 (1998) 297-301.*

[9] Fang(S.), *Canonical brownian motion on the diffeomorphism group of the circle : J. Funct. Analysis 196 (2002) 162-179.*

[10] Fang(S.) and Malliavin (P.), *Stochastic Analysis on the path space of a Riemannian manifold : J. Funct. Analysis 118 (1993) 249-274.*

[11] Kusuoka(S.) and Stroock(D.), *Precise asymptotics of certain Wiener functionals : J. Funct. Analysis 99 (1991) 1-74.*

[12] Malliavin(P.), *Stochastic Calculus of Variations and hypoelliptic operators : In Proc. Int. Symp. on SDE Kyoto 1976, K. Itô ed., John Wiley 1978, 195-263.*

[13] Malliavin(P.) and Thalmaier(A.), *Stochastic Calculus of Variations in Mathematical Finance : Springer 2003, 120 pp.*

[14] Stroock(D.), *An introduction to the Analysis of Paths on a Riemannian manifold : American Math. Society 2000, pp. 269.*

H.A.
INSSET, Université de Picardie, 48 rue Raspail, 02100 Saint-Quentin, France;
`hairault@insset.u-picardie.fr`

P.M.
10 rue Saint Louis en l'Isle, 75004 Paris, France;
`sli@ccr.jussieu.fr`

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 13–29

Invariant Measures for a Stochastic Porous Medium Equation

Giuseppe Da Prato and Michael Röckner

Abstract.

We prove the existence of (infinitesimally) invariant measures for a stochastic version of the porous medium equation (of exponent $m = 3$) with Dirichlet Laplacian on an open set in $\mathbb{R}^d$.

§1. Introduction

The porous medium equation

$$\frac{\partial X}{\partial t} = \Delta(X^m), \quad m \in \mathbb{N}, \tag{1.1}$$

on a bounded open set $D \subset \mathbb{R}^d$ has been studied extensively. We refer to [1] for both the mathematical treatment and the physical background and also to [2, Section 4.3] for the general theory of equations of such type.

In this paper we are interested in a stochastic version of (1.1). Throughout this paper we assume

$$(H1) \qquad m = 3.$$

We believe our approach can be extended for other odd values of m, but this would require a technically much more complicated proof. To avoid the latter and to explain the main idea we restrict to the above case.

We consider Dirichlet boundary conditions for the Laplacian Δ. So, the stochastic partial differential equation we would like to analyze for suitable initial conditions is the following:

$$dX(t) = \Delta(X^3(t))dt + \sqrt{C}\, dW(t), \quad t \geq 0. \tag{1.2}$$

Received April 15, 2003.
2000 Mathematics Subject Classification AMS : 76S05, 35J25, 37L40 .

As in [3], where similar equations were studied (but with $x \to x^3$ replaced by some $\beta : \mathbb{R} \to \mathbb{R}$ of linear growth, satisfying, in particular, $\beta' \geq c > 0$), it turns out that the appropriate state space is $H^{-1}(D)$, i.e. the dual of the Sobolev space $H_0^1 := H_0^1(D)$. Below we shall use the standard $L^2(D)$ dualization $\langle \cdot, \cdot \rangle$ between $H_0^1(D)$ and $H = H^{-1}(D)$ induced by the embeddings

$$H_0^1(D) \subset L^2(D)' = L^2(D) \subset H^{-1}(D) = H$$

without further notice. Then for $x \in H$

$$|x|_H^2 = \int_D (-\Delta)^{-1} x(\xi)\, x(\xi) d\xi$$

and for the dual H' of H we have $H' = H_0^1$.

$(W_t)_{t \geq 0}$ is a cylindrical Brownian motion in H and C is a positive definite bounded operator on H of trace class. To be more concrete below we assume:

(H2) *There exists λ_k, $k \in [0, +\infty)$, $k \in \mathbb{N}$, such that for the eigenbasis $\{e_k |\ k \in \mathbb{N}\}$ of Δ (with Dirichlet boundary conditions) we have* $Ce_k = \sqrt{\lambda_k}\, e_k$ *for all* $k \in \mathbb{N}$.

(H3) *For* $\alpha_k := \sup_{\xi \in D} |e_k(\xi)|^2$, $k \in \mathbb{N}$, *we have*
$$K := \sum_{k=1}^{\infty} \alpha_k \lambda_k < +\infty.$$

Our aim in this paper is to construct invariant measures for (1.2). Existence of solutions to (1.2) will be studied in another paper. To formulate what is meant by "invariant measure" without refering to a solution of (1.2) we need to consider the generator, also called Kolmogorov operator, corresponding to (1.2).

Applying Itô's formula (on a heuristic level) to (1.2) one finds what the corresponding Kolmogorov operator, let us call it N_0, should be, namely

$$N_0\varphi(x) = \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k D^2\varphi(e_k, e_k) + D\varphi(x)(\Delta(x^3)), \quad x \in H, \tag{1.3}$$

where $D\varphi$, $D^2\varphi$ denote the first and second Fréchet derivatives of $\varphi : H \to \mathbb{R}$. So, we take $\varphi \in C_b^2(H)$.

In order to make sense of (1.3) one needs that $\Delta(x^3) \in H$ at least for "relevant" $x \in H$. Here one clearly sees the difficulties since x^3 is, of course, not defined for any Schwartz distribution in $H = H^{-1}$, not to mention that it will not be in $H_0^1(D)$. An invariant measure for (1.2) is now defined "infinitesimally" (cf.[4]), without having a solution to (1.2), as the solution to the equation

$$N_0^* \mu = 0 \tag{1.4}$$

with the property that μ is supported by those $x \in H$ for which x^3 makes sense and $\Delta(x^3) \in H$. (1.4) is a short form for

$$N_0 \varphi \in L^1(H, \mu) \text{ and } \int_H N_0 \varphi d\mu = 0 \text{ for all } \varphi \in C_b^2(H). \tag{1.5}$$

Any invariant measure for any solution of (1.2) in the classical sense will satisfy (1.4).

In §2 we construct a solution μ to (1.4) and prove the necessary support properties of μ, more precisely, that for all $M \in \mathbb{N}$, $M \geq 2$,

$$\mu\left(\{x \in L^2(D) | \, x^M \in H_0^1\}\right) = 1,$$

so that N_0 in (1.3) is μ–a.e. well defined for all $\varphi \in C_b^2(H)$. We rely on results in [3] which we apply to suitable approximations, i.e. the function $x \mapsto x^3$ is replaced by

$$\beta_\varepsilon(x) := \frac{x^3}{1 + \varepsilon x^2} + \varepsilon x, \quad \varepsilon \in (0, 1],$$

to which the results in [3] apply.

§2. Existence of an infinitesimal invariant measure

Throughout this section (H1)–(H3) are still in force. So, we first consider the following approximations for the Kolmogorov operator N_0. For $\varepsilon \in (0, 1]$ we define for $\varphi \in C_b^2(H)$, $x \in L^2(D)$ such that $\beta_\varepsilon(x) \in H_0^1$

$$N_\varepsilon \varphi(x) := \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k D^2 \varphi(x)(e_k, e_k) + D\varphi(x)(\Delta \beta_\varepsilon(x)), \tag{2.1}$$

where

$$\beta_\varepsilon(r) := \frac{r^3}{1 + \varepsilon r^2} + \varepsilon r, \quad r \in \mathbb{R}. \tag{2.2}$$

We note that β_ε is Lipschitz and recall the following result from [3] which is crucial for our further analysis, see [3, Theorems 3.1, 3.9, Remark 3.1].

Theorem 2.1. *Let $\varepsilon \in (0,1]$. Then there exists a probability measure μ_ε on H such that*

$$\mu_\varepsilon(H_0^1) = 1, \tag{2.3}$$

$$\int_H |x|^2_{H_0^1} \mu_\varepsilon(dx) < +\infty, \tag{2.4}$$

$$\int_H |\beta_\varepsilon|^2_{H_0^1} d\mu_\varepsilon = \int_H |\Delta\beta_\varepsilon|^2_H d\mu_\varepsilon < +\infty \tag{2.5}$$

and

$$\int_H N_\varepsilon \varphi d\mu_\varepsilon = 0 \quad \text{for all } \varphi \in C_b^2(H). \tag{2.6}$$

Remark 2.2. (i). In [3] only

$$\mu_\varepsilon\left(\{x \in L^2(D) |\ \beta_\varepsilon(x) \in H_0^1\}\right) = 1$$

was proved. But since $\beta_\varepsilon(0) = 0$, $\beta_\varepsilon(\mathbb{R}) = \mathbb{R}$, and

$$\beta_\varepsilon'(r) = r^2\, \frac{3 + \varepsilon r^2}{(1+\varepsilon r^2)^2} + \varepsilon \geq \varepsilon \quad \text{for all } r \in \mathbb{R}, \tag{2.7}$$

it follows that the inverse β_ε^{-1} of β_ε is Lipschitz with $\beta_\varepsilon^{-1}(0) = 0$, so $\beta_\varepsilon(x) \in H_0^1$ is equivalent to $x \in H_0^1$ and (2.4) follows from (2.5), since

$$|\nabla x| = |\nabla \beta_\varepsilon^{-1}(\beta_\varepsilon(x))| \leq \varepsilon^{-1} |\nabla \beta_\varepsilon(x)|.$$

We thank V. Barbu for pointing this out to us.

(ii) By Theorem 2.1 we have that $N_\varepsilon\varphi(x)$ is well defined for μ_ε–a.e. $x \in H$.

For $N \in \mathbb{N}$ we define

$$P_N x = \sum_{k=1}^N \langle x, e_k\rangle_k e_k, \quad x \in H.$$

Note that, since $\{e_k |\ k \in \mathbb{N}\}$ is the eigenbasis of the Laplacian we have that the respective restriction P_N is also an orthogonal projection on $L^2(D)$ and H_0^1 and on both spaces $(P_N)_{N\in\mathbb{N}}$ also converges strongly to the identity.

The first new result on μ_ε, $\varepsilon \in (0,1]$, is the following:

Proposition 2.3. $\{\mu_\varepsilon,\ \varepsilon \in (0,1]\}$ *is tight on H. For any weak limit point μ*

$$\int_H |x|^2_{L^2(D)}\mu(dx) \le \int_D 1\, d\xi + \frac{1}{2}\operatorname{Tr} C.$$

In particular, $\mu(L^2(D)) = 1$.

Proof. For $n \in \mathbb{N}$ let $\chi_n \in C^\infty(\mathbb{R})$, $\chi_n(x) = x$ on $[-n, n]$, $\chi_n(x) = (n+1)\text{sign } x$, for $x \in \mathbb{R}\backslash[-(n+2), n+2]$, $0 \le \chi_n' \le 1$ and $\sup_{n\in\mathbb{N}}|\chi_n''| < +\infty$. Define for $n, N \in \mathbb{N}$

$$\varphi_{N,n}(x) := \frac{1}{2}\,\chi_n(|P_n x|^2_H).$$

Then $\varphi_{N,n} \in C_b^2(H)$ and for $x \in H$

$$\begin{aligned} N_\varepsilon \varphi_{N,n}(x) \quad &= \quad \frac{1}{2}\sum_{k=1}^{N}\lambda_k\left[2\chi_n''(|P_n x|^2_H)\langle P_N x, e_k\rangle_H^2 + \chi_n'(|P_n x|^2_H)\right] \\ &\quad + \chi_n'(|P_n x|^2_H)\langle P_N x, \Delta\beta_\varepsilon(x)\rangle_H. \end{aligned}$$

Hence integrating with respect to μ_ε, by (2.6) we find

$$\begin{aligned} &\int_H \chi_n'(|P_n x|^2_H)\langle P_N x, \beta_\varepsilon(x)\rangle_{L^2(D)}\mu_\varepsilon(dx) \\ &= \frac{1}{2}\sum_{k=1}^{N}\lambda_k\int_H \left[2\chi_n''(|P_n x|^2_H)\langle P_N x, e_k\rangle_H^2 + \chi_n'(|P_n x|^2_H)\right]\mu_\varepsilon(dx) \\ &\le \frac{1}{2}\sum_{k=1}^{N}\lambda_k + \sup_{k\in\mathbb{N}}\lambda_k\int_H |\chi_n''(|P_n x|^2_H)|\ |P_N x|^2_H\mu_\varepsilon(dx). \end{aligned}$$

For all $n \in \mathbb{N}$ the integrand in the left hand side is bounded by

$$1_{\{|P_n x|^2_H \le n+2\}}\ |P_N x|_H\ |\beta_\varepsilon(x)|_{H_0^1},$$

and similar bounds for the integrand in the right hand side hold. Therefore, (2.5) and Lebesgue's dominated convergence theorem allow us to take $N \to \infty$ and obtain

$$\int_H \chi_n'(|x|_H^2)\langle x, \beta_\varepsilon(x)\rangle_{L^2(D)}\mu_\varepsilon(dx)$$

$$\leq \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k + \sup_{k\in\mathbb{N}}\lambda_k \int_H |\chi_n''(|x|_H^2)|\ |x|_H^2\ \mu_\varepsilon(dx).$$

$$\leq \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k + \sup_{k\in\mathbb{N}}\lambda_k \int_{\{|x|_H^2\geq n\}} |x|_H^2\ \mu_\varepsilon(dx).$$

Hence taking $n \to \infty$ by (2.4) and using the definition (2.2) of β_ε we arrive at

$$\int_H\int_D \left(\frac{x^4(\xi)}{1+\varepsilon x^2(\xi)} + \varepsilon x^2(\xi)\right) d\xi\mu_\varepsilon(dx) \leq \frac{1}{2}\operatorname{Tr} C.$$

Since $\varepsilon \in (0,1]$, this implies

$$\begin{aligned}(2.8)\qquad \int_H |x|_{L^2(D)}^2\mu_\varepsilon(dx) &\leq \int_D \left(1+\frac{x^4(\xi)}{1+x^2(\xi)}\right) d\xi\mu_\varepsilon(dx)\\ &\leq \int_D 1\, d\xi + \frac{1}{2}\operatorname{Tr} C.\end{aligned}$$

Since $L^2(D) \subset H$ is compact, this implies that $\{\mu_\varepsilon |\ \varepsilon \in (0,1]\}$ is tight on H. Since the map $x \to |x|_{L^2(D)}^2$ is lower semicontinuous and nonnegative on H all assertions follow. □

Later we need better support properties of μ. Therefore, our next aim is to prove the following:

Theorem 2.4. *Let $(H1) - (H3)$ hold. Then:*

(i) *For all $M \in \mathbb{N}$, $M \geq 2$, there exists a constant $C_M = C_M(D,K) > 0$ such that*

$$\sup_{\varepsilon\in(0,1]}\int_H\int_D x^{2(M-1)}(\xi)|\nabla x(\xi)|^2 d\xi\mu_\varepsilon(dx) \leq C_M.$$

(ii) *For all $M \in \mathbb{N}$, $M \geq 2$ and any limit point μ as in Proposition 2.3*

$$\int_H\int_D |\nabla(x^M)(\xi)|^2 d\xi\mu(dx) \leq C_M.$$

In particular, setting

$$H^1_{0,M} := \{x \in L^2(D) |\; x^M \in H^1_0\}$$

we have

$$\mu(H^1_{0,M}) = 1 \quad \text{for all } M \geq 2.$$

In order to prove Theorem 2.4 we need some preparation, i.e. more precise information about the μ_ε, $\varepsilon \in (0,1]$. This can be deduced from (2.6), i.e. from the fact that μ_ε is an infinitesimally invariant measure for N_ε. So, we fix $\varepsilon \in (0,1]$ and for the rest of this section we assume that $(H1)-(H3)$ hold.

We need to apply (2.6) with φ replaced by $\varphi_M : L^2(D) \to [0,\infty]$, $M \in \mathbb{N}$, given by

$$\varphi_M(x) := \int_D x^{2M}(\xi)\, d\xi, \quad x \in L^2(D).$$

Clearly, such functions are not in $C_b^2(H)$ so we have to construct proper approximations. So, define for $\delta \in (0,1]$

$$f_{M,\delta}(r) := \frac{r^{2M}}{1+\delta r^2}, \quad r \in \mathbb{R}. \tag{2.9}$$

Then for $r \in \mathbb{R}$

$$f'_{M,\delta}(r) = (1+\delta r^2)^{-2}[2Mr^{2M-1} + 2\delta(M-1)r^{2M+1}] \tag{2.10}$$

and

(2.11)
$$f''_{M,\delta}(r) = 2(1+\delta r^2)^{-3}[M(2M-1)r^{2M-2} + \delta(4M^2-6M-1)r^{2M} + \delta^2(M-1)(2M-3)r^{2M+2}].$$

We have chosen this approximation since below (cf. Lemma 2.7) it will be crucial that $f''_{M,\delta}$ is nonnegative if $M \geq 2$. More precisely we have

$$\begin{aligned} & 0 \leq f_{M,\delta}(r) \leq \tfrac{1}{\delta}\, r^{2M-2} \\ & 0 \leq f'_{M,\delta}(r) \leq \tfrac{2M}{\delta}\, |r|^{2M-3} \\ & 0 \leq f''_{M,\delta}(r) \leq 16M^2\, |r|^{2M-4} \inf\{r^2, 1/\delta\}. \end{aligned} \tag{2.12}$$

Remark 2.5. The following will be used below: if $x \in H_0^1$ is such that for $M \in \mathbb{N}$

$$\int_H x^{2(M-1)}(\xi)|\nabla x(\xi)|^2 d\xi < \infty, \tag{2.13}$$

then $x^M \in H_0^1$ and $x^{M-1}\nabla x = \frac{1}{M}\,\nabla x^M$, or using the notation introduced in Theorem 2.4–(ii) equivalently $x \in H^1_{0,M}$. The proof is standard by approximation. So, we omit it. We also note that by Poincaré's inequality, $H^1_{0,M} \subset L^{2M}(D)$. More precisely, there exists $C(D) \in (0,\infty)$ such that

$$\begin{aligned} C(D)\int_D x^{2M}(\xi)d\xi &\le \int_D |\nabla x^M(\xi)|^2 d\xi \\ &= M\int_D x^{2(M-1)}(\xi)|\nabla x^M(\xi)|^2 d\xi, \end{aligned} \tag{2.14}$$

for all x as above.

The following lemma is a consequence of (2.6) and crucial for our analysis of μ_ε, $\varepsilon \in (0,1]$ and their limit points.

Lemma 2.6. *Let $M \in \mathbb{N}$, $\delta \in (0,1]$. Assume that*

$$\int_H\int_D x^{2(M-2)}(\xi)|\nabla x(\xi)|^2 d\xi\mu_\varepsilon(dx) < \infty \quad \text{if } M \ge 3. \tag{2.15}$$

Then

$$\begin{aligned} &\frac{1}{2}\sum_{k=1}^{\infty}\lambda_k\int_H\int_D f''_{M,\delta}(x(\xi))e_k^2(\xi)d\xi\mu_\varepsilon(dx) \\ &= \int_H\int_D f''_{M,\delta}(x(\xi))\beta'_\varepsilon(x(\xi))|\nabla x(\xi)|^2 d\xi\mu_\varepsilon(dx). \end{aligned} \tag{2.16}$$

Proof. We first note that (2.15) holds for $M = 2$ by (2.3). For $\kappa \in (0,1]$ we define

$$f_{M,\delta,\kappa}(r) := f_{M,\delta}(r)e^{-\frac{1}{2}\,\kappa r^2}, \quad r \in \mathbb{R} \quad \text{if } M \ge 2$$

and $f_{1,\delta,\kappa} := f_{1,\delta}$. Then (2.11) implies that $f_{M,\delta,\kappa} \in C_b^2(\mathbb{R})$. Define

$$\varphi_{M,\delta,\kappa}(x) := \int_D f_{M,\delta,\kappa}(x(\xi))d\xi, \quad x \in L^2(D).$$

Then it is easy to check that $\varphi_{M,\delta,\kappa}$ is Gâteaux differentiable on $L^2(D)$ and that for all $y,z \in L^2(D)$

$$\varphi'_{M,\delta,\kappa}(x)(y) = \int_D f'_{M,\delta,\kappa}(x(\xi))y(\xi)d\xi, \tag{2.17}$$

$$\varphi''_{M,\delta,\kappa}(x)(y,z) = \int_D f''_{M,\delta,\kappa}(x(\xi))y(\xi)z(\xi)d\xi. \tag{2.18}$$

Hence

$$\varphi_{M,\delta,\kappa} \circ P_N \in C_b^2(H)$$

and for all $x \in H_0^1$ (hence $\beta_\varepsilon(x) \in H_0^1$)

$$\begin{aligned} N_\varepsilon(\varphi_{M,\delta,\kappa} \circ P_N)(x) &= \frac{1}{2}\sum_{k=1}^N \lambda_k \int_D f''_{M,\delta,\kappa}(P_N x(\xi))e_k^2(\xi)d\xi \\ &\quad + \int_D f'_{M,\delta,\kappa}(P_N x(\xi))P_N(\Delta\beta_\varepsilon(x))(\xi)d\xi. \end{aligned}$$

Since $P_N\Delta = \Delta P_N$, integrating by parts we obtain

$$\begin{aligned} N_\varepsilon(\varphi_{M,\delta,\kappa} \circ P_N)(x) &= \frac{1}{2}\sum_{k=1}^N \lambda_k \int_D f''_{M,\delta,\kappa}(P_N x(\xi))e_k^2(\xi)d\xi \\ &\quad - \int_D f''_{M,\delta,\kappa}(P_N x(\xi))\langle\nabla(P_N x)(\xi), \nabla(P_N\beta_\varepsilon(x))(\xi)\rangle_{\mathbb{R}^d} d\xi. \end{aligned}$$

Since $(P_N)_{N\in\mathbb{N}}$ converges strongly to the identity in H_0^1, we conclude by (H3) that

$$\begin{aligned} \lim_{N\to\infty} N_\varepsilon(\varphi_{M,\delta,\kappa} \circ P_N)(x) &= \frac{1}{2}\sum_{k=1}^\infty \lambda_k \int_D f''_{M,\delta,\kappa}(x(\xi))e_k^2(\xi)d\xi \\ &\quad - \int_D f''_{M,\delta,\kappa}(x(\xi))\beta'_\varepsilon(x)(\xi)|\nabla x(\xi)|^2 d\xi. \end{aligned}$$

Since β_ε is Lipschitz, by (2.3)–(2.5) and (H3) this convergence also holds in $L^1(H,\mu_\varepsilon)$. Hence (2.6) implies that

$$\begin{aligned} &\frac{1}{2}\sum_{k=1}^\infty \lambda_k \int_H\int_D f''_{M,\delta,\kappa}(x(\xi))e_k^2(\xi)d\xi\mu_\varepsilon(dx) \\ &= \int_H\int_D f''_{M,\delta,\kappa}(x(\xi))\beta'_\varepsilon(x)(\xi)|\nabla x(\xi)|^2 d\xi\mu_\varepsilon(dx). \end{aligned} \tag{2.19}$$

So, for $M = 1$ the assertion is proved. If $M \geq 2$, an elementary calculation shows that by (2.12) there exists a constant $C(M,\delta) > 0$ (only depending on M and δ) such that

$$|f''_{M,\delta,\kappa}(r)| \leq C(M,\delta) r^{2(M-2)}, \quad r \in \mathbb{R}. \tag{2.20}$$

Hence by (H3), Remark 2.5 and assumption (2.15) we can apply Lebesgue's dominated convergence theorem to (2.19) and letting $\kappa \to 0$ we obtain the assertion. □

Lemma 2.7. *Let $M \in \mathbb{N}$ and assume that* (2.15) *holds if $M \geq 3$.*
(i) We have

$$\begin{aligned} &\frac{K}{2} \int_H \int_D x^{2(M-1)}(\xi)\, d\xi \mu_\varepsilon(dx) \\ &\geq \int_H \int_D x^{2(M-1)}(\xi) \left(\frac{x^2(\xi)}{1+x^2(\xi)} + \varepsilon \right) |\nabla x(\xi)|^2 d\xi \mu_\varepsilon(dx). \end{aligned} \tag{2.21}$$

(ii) If $M \geq 2$, we have

$$\begin{aligned} &\frac{K}{2} \int_H \int_D \left(x^{2(M-1)}(\xi) + x^{2(M-2)}(\xi) \right) d\xi \mu_\varepsilon(dx) \\ &\geq \int_H \int_D x^{2(M-1)}(\xi)\, |\nabla x(\xi)|^2 d\xi \mu_\varepsilon(dx) \\ &= \frac{1}{M^2} \int_H \int_D |\nabla x^M(\xi)|^2 d\xi \mu_\varepsilon(dx). \end{aligned} \tag{2.22}$$

(iii)

$$\int_H \int_D |\nabla x(\xi)|^2 d\xi \mu_\varepsilon(dx) \leq \frac{K}{2\varepsilon}.$$

Proof. (i) By (H3) the left hand side of (2.16) is dominated by

$$\frac{K}{2} \int_H \int_D f''_{M,\delta}(x(\xi)) d\xi \mu_\varepsilon(dx).$$

If $M \geq 2$, by assumption (2.15) and Remark 2.5 we know that

$$\int_H \int_D x^{2(M-1)}(\xi) d\xi \mu_\varepsilon(dx) < \infty$$

which trivially also holds for $M = 1$. So, by (2.11), (2.12) and Lebesgue's dominated convergence theorem we obtain that for $M \geq 2$

$$\frac{K}{2} \int_H \int_D 2M(2M-1)x^{2(M-1)}(\xi)d\xi\mu_\varepsilon(dx)$$

$$\geq \liminf_{\delta\to 0} \int_H \int_D f''_{M,\delta}(x(\xi))\beta'_\varepsilon(x(\xi))|\nabla x(\xi)|^2 d\xi\mu_\varepsilon(dx).$$

Since $f''_{M,\delta} \geq 0$ for $M \geq 2$ and

$$\beta'_\varepsilon(r) \geq \frac{r^2}{1+r^2} + \varepsilon \geq 0 \quad \text{for all } r \in \mathbb{R},$$

we can apply Fatou's lemma to prove the assertion. If $M = 1$ we conclude in the same way by (2.3) and Lebesgue's dominated convergence theorem which applies since β'_ε is bounded and $|f''_{1,\delta}| \leq 6$ for all $\delta \in (0,1]$.

(ii) Since (2.15) holds for $M = 2$, by Hölder's inequality (2.15) holds with $M-1$ replacing M, since by assumption it holds for M. So, the inequality in (i) also holds with $M-2$ replacing $M-1$. Estimating ε on the right hand sides from below by 0 and adding both resulting inequalities we obtain the inequality in (2.22). The equality in (2.22) follows by Remark 2.5.

(iii) The assertion follows from (2.21) setting $M = 1$. □

By an induction argument we shall now prove that the integrals in (2.22) are all finite and at the same time prove the bounds claimed in Theorem 2.4.

Proof of Theorem 2.4. (i). If $M = 2$, then the left hand side of (2.22) is finite by (2.8) and moreover (2.22) applies, so that by (2.8) we have

(2.23)
$$\int_H \int_D x^2(\xi)|\nabla(x(\xi))|^2 d\xi\mu_\varepsilon(dx) \leq \frac{K}{2}\left(\frac{1}{2}\operatorname{Tr} C + 2\int_D 1 d\xi\right) < \infty.$$

Suppose the left hand side of (2.22) is finite for $M \in \mathbb{N}, M \geq 2$, and (2.15) holds. Then (2.22) holds and by Remark 2.5

$$\infty > \int_H \int_D x^{2(M-1)}(\xi)|\nabla(x(\xi))|^2 d\xi\mu_\varepsilon(dx)$$

$$= \frac{1}{M^2}\int_H \int_D |\nabla(x^M(\xi))|^2 d\xi\mu_\varepsilon(dx) \tag{2.24}$$

$$\geq \frac{C(D)^2}{M^2}\int_H \int_D x^{2M}(\xi)d\xi\mu_\varepsilon(dx).$$

Hence (2.15) holds with $M-1$ replacing $M-2$ and the left hand side of (2.22) is finite for $M+1$ replacing M, hence by induction for all $M \in \mathbb{N}$. Furthermore, for all M first applying (2.22) and then applying (2.24) first with $M-1$ replacing M and then with $M-2$ replacing M respectively we obtain

$$\int_H \int_D x^{2(M-1)}(\xi)|\nabla(x(\xi))|^2 d\xi \mu_\varepsilon(dx)$$

$$\leq \frac{K}{2}\left[\left(\frac{M-1}{C(D)}\right)^2 \int_H \int_D x^{2(M-2)}(\xi)|\nabla(x(\xi))|^2 d\xi \mu_\varepsilon(dx) + \int_H \int_D x^{2(M-2)}(\xi) d\xi \mu_\varepsilon(dx)\right] \tag{2.25}$$

$$\leq \frac{K}{2C(D)^2}\left[(M-1)^2 \int_H \int_D x^{2(M-2)}(\xi)|\nabla(x(\xi))|^2 d\xi \mu_\varepsilon(dx) + (M-2)^2 \int_H \int_D x^{2(M-3)}(\xi)|\nabla(x(\xi))|^2 d\xi \mu_\varepsilon(dx)\right]. \tag{2.26}$$

If $M=3$ we cannot use (2.26) since for the second summand we have no bound which is independent of ε, but from (2.25) we obtain by (2.23) and (2.8) that

$$\int_H \int_D x^4(\xi)|\nabla(x(\xi))|^2 d\xi \mu_\varepsilon(dx)$$

$$\leq \frac{K}{2}\left[\left(\frac{2}{C(D)}\right)^2 \frac{K}{2}\left(\frac{1}{2}\operatorname{Tr} C + 2\int_D 1 d\xi\right) + \frac{1}{2}\operatorname{Tr} C + \int_D 1 d\xi\right].$$

Now assertion (i) follows from (2.26) by induction.

To prove (ii) we start with the following
Claim: For all $M \in \mathbb{N}$

$$\Theta_M(x) := 1_{H^1_{0,M}}(x) \int_D |\nabla x^M(\xi)|^2 d\xi + \infty \cdot 1_{H \setminus H^1_{0,M}}(x), \quad x \in H \tag{2.27}$$

is a lower semicontinuous function on H.

Since μ is a weak limit point of $\{\mu_\varepsilon | \ \varepsilon \in (0,1]\}$ and $\Theta_M \geq 0$, the claim immediately implies assertion (ii).

To prove the claim let $\alpha > 0$ and $x_n \in \{\Theta_M \leq \alpha\}$, $n \in \mathbb{N}$, such that $x_n \to x$ in H as $n \to \infty$. By Poincaré's inequality $\{x_n | \ n \in \mathbb{N}\}$ is a bounded set in $L^{2M}(D)$. So $x_n \to x$ as $n \to \infty$ also weakly in $L^2(D)$, in particular $x \in L^2(D)$. Since $\{x_n^M | \ n \in \mathbb{N}\}$ is bounded in H_0^1, there exists a subsequence $(x_{n_k}^M)_{k \in \mathbb{N}}$ and $y \in H_0^1$ such that $x_{n_k}^M \to y$ as $k \to \infty$ weakly in H_0^1 and

$$\int_D |\nabla y(\xi)|^2 d\xi \leq \alpha.$$

Since the embedding $H_0^1 \subset L^2(D)$ is compact, $x_{n_k}^M \to y$ as $k \to \infty$ in $L^2(D)$. Selecting another subsequence if necessary, this convergence is $d\xi$–a.e., hence

$$x_{n_k} \to y^{\frac{1}{M}} \quad \text{as } k \to \infty,\ d\xi\text{–a.e.}$$

Since (selecting another subsequence if necessary) we also know that the Cesaro mean of $(x_{n_k})_{k \in \mathbb{N}}$ has x as an accumulation point in the topology of $d\xi$–a.e. convergence, we must have $x^M = y$, so $x \in \{\Theta_M \leq \alpha\}$. □

As a consequence of the previous proof we obtain:

Corollary 2.8. *Let $M \in \mathbb{N}$. Then Θ_M has compact level sets in H.*

Proof. We already know from the previous proof that Θ_M is lower semicontinuous. The relative compactness of their level sets is, however, clear by Poincaré's inequality since $L^{2M}(D) \subset H$ is compact. □

Since for $M \in \mathbb{N}$ and $x \in H_{0,M}^1$

$$|\Delta x^M|_H = \int_D |\nabla x^M(\xi)|^2 d\xi, \tag{2.28}$$

so $\Delta x^M \in H$, we can define the Kolmogorov operator in (1.3) rigorously for $x \in H_{0,3}^1$. So, for $\varphi \in C_b^2(H)$

$$N_0\varphi(x) := \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k D^2\varphi(x)(e_k, e_k) + D\varphi(x)(\Delta x^3). \tag{2.29}$$

We note that by Theorem 2.4–(ii) and (2.28), $N_0\varphi \in L^2(H, \mu)$ for any weak limit point μ of $\{\mu_\varepsilon | \ \varepsilon \in (0,1]\}$ on H. Now we can prove our main result, namely that any such μ is an infinitesimally invariant measure for N_0 in the sense of [4], i.e. satisfies (1.4).

Theorem 2.9. *Assume that* $(H1)$*–*$(H3)$ *hold. Let* μ *be as in Proposition* 2.3. *Then*

$$\int_H N_0 \varphi d\mu = 0 \quad \text{for all } \varphi \in C_b^2(H).$$

Proof. Let $\varphi \in C_b^2(H)$. For $N \in \mathbb{N}$ define $\varphi_N := \varphi \circ P_N$. Then for $x \in H_{0,3}^1$

$$\begin{aligned} N_0 \varphi_N(x) &= \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k D^2 \varphi(P_N x)(P_N e_k, P_N e_k) + D\varphi_N(x) \ (\Delta x^3) \\ &= \frac{1}{2} \sum_{k=1}^{N} \lambda_k D^2 \varphi(P_N x)(e_k, e_k) + D\varphi(P_N x)(P_N(\Delta x^3)). \end{aligned}$$

If we can prove that

$$\int_H N_0 \varphi_N d\mu = 0 \quad \text{for all } N \in \mathbb{N}, \tag{2.30}$$

the same is true for φ by Lebesgue's dominated convergence theorem. So, fix $N \in \mathbb{N}$. Then by (2.6)

$$\begin{aligned} \int_H N_0 \varphi_N d\mu &= \lim_{\varepsilon \to 0} \int_H \frac{1}{2} \sum_{k=1}^{\infty} \lambda_k D^2 \varphi_N(x)(e_k, e_k) \mu_\varepsilon(dx) \\ &\quad + \int_H D\varphi_N(x)(\Delta x^3) \mu(dx) \\ &= -\lim_{\varepsilon \to 0} \int_H D\varphi_N(x)(\Delta \beta_\varepsilon(x)) \mu_\varepsilon(dx) \\ &\quad + \int_H D\varphi_N(x)(\Delta x^3) \mu(dx) \\ &= \lim_{\varepsilon \to 0} \sum_{i=1}^{N} \int_H \Big[D\varphi(P_N x)(e_i) \langle e_i, \Delta x^3 \rangle_H \mu(dx) \\ &\quad - D\varphi(P_N x)(e_i) \langle e_i, \Delta \beta_\varepsilon(x) \rangle_H \mu_\varepsilon(dx) \Big]. \end{aligned} \tag{2.31}$$

For $i \in \{1, ..., N\}$ fixed we have

$$
\begin{aligned}
&\Big| \int_H D\varphi(P_N x)(e_i)\langle e_i, \Delta x^3\rangle_H \mu(dx) \\
&- \int_H D\varphi(P_N x)(e_i)\langle e_i, \Delta \beta_\varepsilon(x)\rangle_H \mu_\varepsilon(dx)\Big| \\
&\leq \left| \int_H D\varphi(P_N x)(e_i)\langle e_i, \Delta x^3\rangle_H (\mu - \mu_\varepsilon)(dx)\right| \\
&+ \left| \int_H D\varphi(P_N x)(e_i)\langle e_i, \Delta (x^3 - \beta_\varepsilon(x))\rangle_H \mu_\varepsilon(dx)\right| .
\end{aligned}
\tag{2.32}
$$

The right hand side's second summand is bounded by
(2.33)

$$
|e_i|_{L^2(D)} \sup_{x\in H} |D\varphi(x)|_{H_0^1} \int_H \left(\int_D |x^3(\xi) - \beta_\varepsilon(x(\xi))|^2 \, d\xi \right)^{1/2} \mu_\varepsilon(dx).
$$

We have

$$
|r^3 - \beta_\varepsilon(r)| = \left| \frac{\varepsilon r^5}{1+\varepsilon r^2} - \varepsilon r \right| \leq \varepsilon(|r|^5 + |r|), \quad r \in \mathbb{R}.
$$

So, the term in (2.33) is dominated by

$$
\varepsilon |e_i|_{L^2(D)} \sup_{x\in H} |D\varphi(x)|_{H_0^1} \int_H \left(||x|^5|_{L^2(D)} + |x|_{L^2(D)} \right) \mu_\varepsilon(dx)
$$

which by Theorem 2.4–(i), Remark 2.5 and Poincaré's inequality converges to 0 as $\varepsilon \to 0$.

Now we estimate the first summand in the right hand side of (2.32). So, we define

$$
f(x) := D\varphi(P_N x)(e_i)\langle e_i, \Delta x^3\rangle_H .
$$

Then since $\langle e_i, \Delta(x^3)\rangle_H = \langle e_i, x^3\rangle_{L^2(D)}$, it follows by the proof of the lower semicontinuity of Θ_3 that f is continuous on the level sets of Θ_3 (with Θ_3 defined as in (2.27)). Furthermore, since

$$
|f(x)| \leq \sup_{x\in H} |D\varphi(x)|_{H_0^1} \, |x^3|_{L^2(D)},
$$

it follows that

$$
\lim_{R\to\infty} \sup_{\{\Theta_3 \geq R\}} \frac{|f(x)|}{1+\Theta_3(x)} = 0.
$$

Furthermore, by Corollary 2.8 the function $1+\Theta_3$ has compact level sets. Hence by [8, Theorem 5.1 (ii)], there exist $f_n \in C_b(H)$, $n \in \mathbb{N}$, such that

$$\lim_{n\to\infty} \sup_{x\in H} \frac{|f(x)-f_n(x)|}{1+\Theta_3(x)} = 0. \tag{2.34}$$

But

$$\begin{aligned}&\left|\int_H D\varphi(P_N x)(e_i)\langle e_i, \Delta x^3\rangle_H(\mu-\mu_\varepsilon)(dx)\right|\\ &\leq \int_H |f(x)-f_n(x)|(\mu+\mu_\varepsilon)(dx) + \left|\int_H f_n(x)(\mu-\mu_\varepsilon)(dx)\right|.\end{aligned}$$

For fixed n the second summand tends to 0 as $\varepsilon \to 0$ and the first is dominated by

$$\sup_{x\in H} \frac{|f(x)-f_n(x)|}{1+\Theta_3(x)} \sup_{\varepsilon>0} \int_H (1+\Theta_3)d(\mu+\mu_\varepsilon),$$

which in turn by Theorem 2.4 and (2.34) tends to zero as $n \to \infty$. So, also the first summand in (2.32) tends to zero as $\varepsilon \to 0$. Hence the right hand side of (2.31) is zero and (2.30) follows which completes the proof. □

Aknowledgement. The first author would like to thank the University of Bielefeld for its kind hospitality and financial support. This work was also supported by the research program "Equazioni di Kolmogorov" from the Italian "Ministero della Ricerca Scientifica e Tecnologica".

The second named author would like to thank the Scuola Normale Superiore di Pisa for the hospitality and the financial support during a very pleasant stay in Pisa when most of this work was done. Financial support of the BiBoS-Research Centre and the DFG-Forschergruppe "Spectral Analysis, Asymptotic Distributions, and Stochastic Dynamics" is also gratefully acknowledged.

References

[1] D. G. Aronson, *The porous medium equation*, in Lect. Notes Math. Vol. 1224, (A. Fasano and al. editors), Springer, Berlin, 1–46, 1986.

[2] V. Barbu, *Analysis and control of nonlinear infinite dimensional systems*, Academic Press, San Diego, 1993.

[3] V. Barbu and G. Da Prato, *The two phase stochastic Stefan problem*, Probab. Theory Relat. Fields, **124**, 544–560, 2002.

[4] V. Bogachev and M. Röckner, *Elliptic equations for measures on infinite dimensional spaces and applications,* Probab. Theory Relat. Fields, 120, 445–496, 2001.

[5] G. Da Prato and M. Röckner, *Singular dissipative stochastic equations in Hilbert spaces,* Probab. Theory Relat. Fields, **124**, 2, 261–303, 2002.

[6] A. Eberle, *Uniqueness and non–uniqueness of singular diffusion operators,* Lecture Notes in Mathematics 1718, Berlin, Springer–Verlag, 1999.

[7] M. Röckner, *L^p-analysis of finite and infinite dimensional diffusions*, Lecture Notes in Mathematics, **1715**, G. Da Prato (editor), Springer-Verlag, 65-116, 1999.

[8] M. Röckner and Z. Sobol, *Kolmogorov equations in infinite dimensions: well-posedness, regularity of solutions, and applications to generalized stochastic Burgers equations,* Preprint 2003.

[9] W. Stannat, *The theory of generalized Dirichlet forms and its applications in Analysis and Stochastics,* Memoirs AMS, 678, 1999.

[10] D.W. Stroock and S.R.S. Varadhan, *Multidimensional Diffusion Processes*, Springer–Verlag, 1979.

Giuseppe Da Prato
Scuola Normale Superiore,
Piazza dei Cavalieri 7, 56126 Pisa,
Italy

Michael Röckner
Fakultät für Mathematik, Universität Bielefeld,
Postfach 100131, D–33501 Bielefeld,
Germany

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 31–47

Equivariant Diffusions on Principal Bundles

K. David Elworthy, Yves Le Jan and Xue-Mei Li

Let $\pi : P \to M$ be a smooth principal bundle with structure group G. This means that there is a C^∞ right multiplication $P \times G \to P$, $u \mapsto u \cdot g$ say, of the Lie group G such that π identifies the space of orbits of G with the manifold M and π is locally trivial in the sense that each point of M has an open neighbourhood U with a diffeomorphism

$$\begin{tikzcd} \tau_U : \pi^{-1}(U) \arrow[rr] \arrow[dr] & & U \times G \arrow[dl] \\ & U & \end{tikzcd}$$

over U, which is equivariant with respect to the right action of G, i.e. if $\tau_u(b) = (\pi(b), k)$ then $\tau_u(b \cdot g) = (\pi(b), kg)$. Assume for simplicity that M is compact. Set $n = dimM$. The fibres, $\pi^{-1}(x)$, $x \in M$ are diffeomorphic to G and their tangent spaces $VT_uP (= kerT_u\pi)$, $u \in P$, are the 'vertical' tangent spaces to P. A *connection* on P, (or on π) assigns a complementary 'horizontal' subspace HT_uP to VT_uP in T_uP for each u, giving a smooth horizontal sub-bundle HTP of the tangent bundle TP to P. Given such a connection it is a classical result that for any C^1 curve: $\sigma : [0, T] \to M$ and $u_0 \in \pi^{-1}(\sigma(0))$ there is a unique horizontal $\tilde{\sigma} : [0, T] \to P$ which is a lift of σ, i.e. $\pi(\tilde{\sigma}(t)) = \sigma(t)$ and has $\tilde{\sigma}(0) = u_0$.

In his startling ICM article [8] Itô showed how this construction could be extended to give horizontal lifts of the sample paths of diffusion processes. In fact he was particularly concerned with the case when M is given a Riemannian metric $\langle , \rangle_x$, $x \in M$, the diffusion is Brownian motion on M, and P is the orthonormal frame bundle $\pi : OM \to M$. Recall that each $u \in OM$ with $u \in \pi^{-1}(x)$ can be considered as an isometry $u : \mathbb{R}^n \to T_xM$, $\langle , \rangle_x$ and a

Received March 12, 2003.
Research supported by EPSRC grants GR/NOO 845 and GR/S40404/01, NSF grant DMS 0072387, and EU grant ERBF MRX CT 960075 A.

horizontal lift $\tilde{\sigma}$ determines parallel translation of tangent vectors along σ

$$/\!/_t \equiv /\!/(\sigma)_t : \qquad T_{\sigma(\cdot)}M \to T_{\sigma(t)}M$$
$$v \mapsto \tilde{\sigma}(t)(\tilde{\sigma}(0))^{-1}v.$$

The resulting parallel translation along Brownian paths extends also to parallel translation of forms and elements of $\wedge^p TM$. This enabled Itô to use his construction to obtain a semi-group acting on differential forms

$$P_t\phi = \mathbb{E}(/\!/_t^{-1})_*(\phi) = \mathbb{E}\phi(/\!/_t -).$$

As he pointed out this is not the semi-group generated by the Hodge-Kodaira Laplacian, Δ. To obtain that generated by the Hodge-Kodaira Laplacian, Δ, some modification had to be made since the latter contains zero order terms, the so called Weitzenbock curvature terms. The resulting probabilistic expression for the heat semi-groups on forms has played a major role in subsequent development.

In [5] we go in the opposite direction starting with a diffusion with smooth generator $\mathcal{B}$ on P, which is G-invariant and so projects to a diffusion generator $\mathcal{A}$ on M. We assume the symbol $\sigma_{\mathcal{A}}$ has constant rank so determining a sub-bundle E of TM, (so $E = TM$ if $\mathcal{A}$ is elliptic). We show that this set-up induces a 'semi-connection' on P over E (a connection if $E = TM$) with respect to which $\mathcal{B}$ can be decomposed into a horizontal component $\mathcal{A}^H$ and a vertical part $\mathcal{B}^V$. Moreover any vertical diffusion operator such as $\mathcal{B}^V$ induces only zero order operators on sections of associated vector bundles.

There are two particularly interesting examples. The first when π: $GLM \to M$ is the full linear frame bundle and we are given a stochastic flow $\{\xi_t : 0 \leq t < \infty\}$ on M, generator $\mathcal{A}$, inducing the diffusion $\{u_t : 0 \leq t < \infty\}$ on GLM by

$$u_t = T\xi_t(u_0).$$

Here we can determine the connection on GLM in terms of the LeJan-Watanabe connection of the flow [12], [1], as defined in [6], [7], in particular giving conditions when it is a Levi-Civita connection. The zero order operators arising from the vertical components can be identified with generalized Weitzenbock curvature terms.

The second example slightly extends the above framework by letting $\pi : P \to M$ be the evaluation map on the diffeomorphism group Diff$M$ of M given by $\pi(h) := h(x_0)$ for a fixed point x_0 in M. The group G corresponds to the group of diffeomorphisms fixing x_0. Again we take a flow $\{\xi_t(x) : x \in M, t \geq 0\}$ on M, but now the process on Diff$M$ is just the right invariant process determined by $\{\xi_t : 0 \leq t < \infty\}$. In this case the horizontal lift to the diffeomorphism group of the diffusion $\{\xi_t(x_0) : 0 \leq t < \infty\}$ on M is obtained by 'removal of redundant noise', c.f. [7] while the vertical process is a flow of diffeomorphisms preserving x_0, driven by the redundant noise.

Here we report briefly on some of the main results to appear in [5] and give details of a more probabilistic version Theorem 2.5 below: a skew product decomposition which, although it has a statement not explicitly mentioning connections, relates to Itô's pioneering work on the existence of horizontal lifts. The derivative flow example and a simplified version of the stochastic flow example are described in § 3.

The decomposition and lifting apply in much more generality than with the full structure of a principal bundle, for example to certain skew products and invariant processes on foliated manifolds. This will be reported on later. Earlier work on such decompositions includes [4] [13].

§1. Construction

A. If $\mathcal{A}$ is a second order differential operator on a manifold X, denote by $\sigma^{\mathcal{A}} : T^*X \to TX$ its symbol determined by

$$df\left(\sigma^{\mathcal{A}}(dg)\right) = \frac{1}{2}\mathcal{A}(fg) - \frac{1}{2}\mathcal{A}(f)\,g - \frac{1}{2}f\mathcal{A}(g),$$

for C^2 functions f, g. The operator is said to be *semi-elliptic* if $df\left(\sigma^{\mathcal{A}}(df)\right) \geq 0$ for each $f \in C^2(X)$, and *elliptic* if the inequality holds strictly. Ellipticity is equivalent to $\sigma^{\mathcal{A}}$ being onto. It is called a *diffusion operator* if it is semi-elliptic and annihilates constants, and is *smooth* if it sends smooth functions to smooth functions.

Consider a smooth map $p : N \to M$ between smooth manifolds M and N. By a *lift* of a diffusion operator $\mathcal{A}$ on M over p we mean a diffusion operator $\mathcal{B}$ on N such that

$$\mathcal{B}(f \circ p) = (\mathcal{A}f) \circ p \tag{1}$$

for all C^2 functions f on M. Suppose $\mathcal{A}$ is a smooth diffusion operator on M and $\mathcal{B}$ is a lift of $\mathcal{A}$.

Lemma 1.1. *Let $\sigma^{\mathcal{B}}$ and $\sigma^{\mathcal{A}}$ be respectively the symbols for $\mathcal{B}$ and $\mathcal{A}$. The following diagram is commutative for all $u \in p^{-1}(x)$, $x \in M$:*

$$\begin{tikzcd}
T_u^*N \arrow[r, "\sigma_u^{\mathcal{B}}"] & T_uN \arrow[d, "Tp"] \\
T_x^*M \arrow[u, "(Tp)^*"] \arrow[r, "\sigma_x^{\mathcal{A}}"'] & T_xM.
\end{tikzcd}$$

B. Semi-connections on principal bundles. Let M be a smooth finite dimensional manifold and $P(M, G)$ a principal fibre bundle over M with structure group G a Lie group. Denote by $\pi : P \to M$ the projection and R_a the right translation by a.

Definition 1.2. *Let E be a sub-bundle of TM and $\pi : P \to M$ a principal G-bundle. An E semi-connection on $\pi : P \to M$ is a smooth sub-bundle $H^E TP$ of TP such that*

(i) *$T_u\pi$ maps the fibres $H^E T_u P$ bijectively onto $E_{\pi(u)}$ for all $u \in P$.*
(ii) *$H^E TP$ is G-invariant.*

Notes.
(1) Such a semi-connection determines and is determined by, a smooth horizontal lift:

$$h_u : E_{\pi(u)} \to T_u P, \qquad u \in P$$

such that

(i) $T_u\pi \circ h_u(v) = v$, for all $v \in E_x \subset T_x M$;
(ii) $h_{u \cdot a} = T_u R_a \circ h_u$.

The *horizontal subspace* $H^E T_u P$ at u is then the image at u of h_u, and the composition $h_u \circ T_u P$ is a projection onto $H^E T_u P$.
(2) Let $F = P \times V/\sim$ be an associated vector bundle to P with fibre V. An element of F is an equivalence class $[(u, e)]$ such that $(ug, g^{-1}e) \sim (u, e)$. Set $\tilde{u}(e) = [(u, e)]$. An E semi-connection on P gives a covariant derivative on F. Let Z be a section of F and $w \in E_x \subset T_x M$, the covariant derivative $\nabla_w Z \in F_x$ is defined, as usual for connections, by

$$\nabla_w Z = u(d\tilde{Z}(h_u(w))), \qquad u \in \pi^{-1}(x) = F_x.$$

Here $\tilde{Z} : P \to V$ is $\tilde{Z}(u) = \tilde{u}^{-1} Z(\pi(u))$ considering $\tilde{u}$ as an isomorphism $\tilde{u} : V \to F_{\pi(u)}$. This agrees with the 'semi-connections on E' defined in Elworthy-LeJan-Li [7] when P is taken to be the linear frame bundle of TM and $F = TM$. As described there, any semi-connection can be completed to a genuine connection, but not canonically.

Consider on P a diffusion generator $\mathcal{B}$, which is *equivariant*, i.e.

$$\mathcal{B}f \circ R_a = \mathcal{B}(f \circ R_a), \qquad \forall f, g \in C^2(P, R),\ a \in G.$$

The operator $\mathcal{B}$ induces an operator $\mathcal{A}$ on the base manifold M by setting

$$\mathcal{A}f(x) = \mathcal{B}(f \circ \pi)(u), \qquad u \in \pi^{-1}(x), f \in C^2(M), \tag{2}$$

which is well defined since

$$\mathcal{B}(f \circ \pi)(u \cdot a) = \mathcal{B}((f \circ \pi))(u).$$

Let $E_x := Image(\sigma_x^{\mathcal{A}}) \subset T_xM$, the image of $\sigma_x^{\mathcal{A}}$. Assume the dimension of $E_x = p$, independent of x. Set $E = \cup_x E_x$. Then $\pi : E \to M$ is a sub-bundle of TM.

Theorem 1.3. *Assume $\sigma^{\mathcal{A}}$ has constant rank. Then $\sigma^{\mathcal{B}}$ gives rise to a semi-connection on the principal bundle P whose horizontal map is given by*

$$h_u(v) = \sigma^{\mathcal{B}}\left((T_u\pi)^*\alpha\right) \tag{3}$$

where $\alpha \in T^*_{\pi(u)}M$ satisfies $\sigma_x^{\mathcal{A}}(\alpha) = v$.

Proof. To prove h_u is well defined we only need to show $\psi(\sigma^{\mathcal{B}}(T_u\pi^*(\alpha))) = 0$ for every 1-form ψ on P and for every α in $\ker \sigma_x^{\mathcal{A}}$. Now $\sigma^{\mathcal{A}}\alpha = 0$ implies by Lemma 1.1 that

$$0 = \alpha\sigma^{\mathcal{A}}(\alpha) = (T\pi)^*(\alpha)\sigma^{\mathcal{B}}((T\pi)^*(\alpha)).$$

Thus $T\pi^*(\alpha)\sigma^{\mathcal{B}}(T\pi^*(\alpha)) = 0$. On the other hand we may consider $\sigma^{\mathcal{B}}$ as a bilinear form on T^*P and then for all $\beta \in T_u^*P$,

$$\begin{aligned}
&\sigma^{\mathcal{B}}(\beta + t(T\pi)^*(\alpha), \beta + t(T\pi)^*(\alpha)) \\
&= \sigma^{\mathcal{B}}(\beta, \beta) + 2t\sigma^{\mathcal{B}}(\beta, (T\pi)^*(\alpha)) + t^2\sigma^{\mathcal{B}}((T\pi)^*\alpha, (T\pi)^*\alpha) \\
&= \sigma^{\mathcal{B}}(\beta, \beta) + 2t\sigma^{\mathcal{B}}(\beta, (T\pi)^*(\alpha)).
\end{aligned}$$

Suppose $\sigma^{\mathcal{B}}(\beta, (T\pi)^*(\alpha)) \neq 0$. We can then choose t such that

$$\sigma^{\mathcal{B}}(\beta + t(T\pi)^*(\alpha), \beta + t(T\pi)^*(\alpha)) < 0,$$

which contradicts the semi-ellipticity of $\mathcal{B}$.

We must verify (i) $T_u\pi \circ h_u(v) = v$, $v \in E_x \subset T_xM$ and (ii) $h_{u \cdot a} = T_uR_a \circ h_u$. The first is immediate by Lemma 1.1 and for the second use the fact that $T\pi \circ TR_a = T\pi$ for all $a \in G$ and the equivariance of $\sigma^{\mathcal{B}}$. ∎

§2. Horizontal lifts of diffusion operators and decompositions of equivariant operators

A. Denote by $C^\infty\Omega^p$ the space of smooth differential p-forms on a manifold M. To each diffusion operator $\mathcal{A}$ we shall associate a unique operator $\delta^{\mathcal{A}}$. The horizontal lift of $\mathcal{A}$ can be defined to be the unique operator such that the associated operator $\bar{\delta}$ vanishes on vertical 1-forms and such that $\bar{\delta}$ and $\delta^{\mathcal{A}}$ are intertwined by the lift map π^* acting on 1-forms.

Proposition 2.1. *For each smooth diffusion operator $\mathcal{A}$ there is a unique smooth differential operator $\delta^{\mathcal{A}} : C^\infty(\Omega^1) \to C^\infty\Omega^0$ such that*

(1) $\delta^{\mathcal{A}}(f\phi) = df\sigma^{\mathcal{A}}(\phi)_x + f \cdot \delta^{\mathcal{A}}(\phi)$

(2) $\delta^{\mathcal{A}}(df) = \mathcal{A}(f)$.

For example if $\mathcal{A}$ has Hörmander representation

$$\mathcal{A} = \frac{1}{2}\sum_{j=1}^{m} \mathcal{L}_{X^j}\mathcal{L}_{X^j} + \mathcal{L}_A$$

for some C^1 vector fields X^i, A then

$$\delta^{\mathcal{A}} = \frac{1}{2}\sum_{j=1}^{m} \mathcal{L}_{X^j}\iota_{X^j} + \iota_A$$

where ι_A denotes the interior product of the vector field A acting on differential forms.

Definition 2.2. *Let S be a C^∞ sub-bundle of TN for some smooth manifold N. A diffusion operator $\mathcal{B}$ on N is said to be along S if $\delta^{\mathcal{B}}\phi = 0$ for all 1-forms ϕ which vanish on S; it is said to be strongly cohesive if $\sigma^{\mathcal{B}}$ has constant rank and $\mathcal{B}$ is along the image of $\sigma^{\mathcal{B}}$.*

To be along S implies that any Hörmander form representation of $\mathcal{B}$ uses only vector fields which are sections of S.

Definition 2.3. *When a diffusion operator $\mathcal{B}$ on P is along the vertical foliation VTP of the $\pi : P \to M$ we say $\mathcal{B}$ is vertical, and when the bundle has a semi-connection and $\mathcal{B}$ is along the horizontal distribution we say $\mathcal{B}$ is horizontal.*

If $\pi : P \to M$ has an E semi-connection and $\mathcal{A}$ is a smooth diffusion operator along E it is easy to see that $\mathcal{A}$ has a unique horizontal lift $\mathcal{A}^H$, i.e. a smooth diffusion operator $\mathcal{A}^H$ on P which is horizontal and is a lift of $\mathcal{A}$ in the sense of (1). By uniqueness it is equivariant.

B. The action of G on P induces a homomorphism of the Lie algebra $\mathfrak{g}$ of G with the algebra of right invariant vector fields on P: if $\alpha \in \mathfrak{g}$,

$$A^\alpha(u) = \left.\frac{d}{dt}\right|_{t=0} u\exp(t\alpha),$$

and A^α is called the fundamental vector field corresponding to α. Take a basis $A_1, \dots, A_k$ of $\mathfrak{g}$ and denote the corresponding fundamental vector fields by $\{A_i^*\}$.

We can now give one of the main results from [5]:

Theorem 2.4. *Let $\mathcal{B}$ be an equivariant operator on P with $\mathcal{A}$ the induced operator on the base manifold. Assume $\mathcal{A}$ is strongly cohesive. Then there is a unique semi-connection on P over E for which $\mathcal{B}$ has a decomposition*

$$\mathcal{B} = \mathcal{A}^H + \mathcal{B}^V,$$

where $\mathcal{A}^H$ is horizontal and $\mathcal{B}^V$ is vertical. Furthermore $\mathcal{B}^V$ has the expression $\sum \alpha^{ij}\mathcal{L}_{A_i^*}\mathcal{L}_{A_j^*} + \sum \beta^k \mathcal{L}_{A_k^*}$, where α^{ij} and β^k are smooth functions on P, given by $\alpha^{k\ell} = \tilde{\omega}^k\left(\sigma^{\mathcal{B}}(\tilde{\omega}^\ell)\right)$, and $\beta^\ell = \delta^{\mathcal{B}}(\tilde{\omega}^\ell)$ for $\tilde{\omega}$ any connection 1-form on P which vanishes on the horizontal subspaces of this semi-connection.

We shall only prove the first part of Theorem 2.4 here. The semi-connection is the one given by Theorem 1.3, and we define $\mathcal{A}^H$ to be the horizontal lift of $\mathcal{A}$. The proof that $\mathcal{B}^V := \mathcal{B} - \mathcal{A}^H$ is vertical is simplified by using the fact that a diffusion operator $\mathcal{D}$ on P is vertical if and only if for all C^2 functions f_1 on P and f_2 on M

$$\mathcal{D}(f_1(f_2 \circ \pi)) = (f_2 \circ \pi)\mathcal{D}(f_1). \tag{4}$$

Set $\tilde{f}_2 = f_2 \circ \pi$. Note

$$\left(\mathcal{B} - \mathcal{A}^H\right)(f_1\tilde{f}_2) = \tilde{f}_2(\mathcal{B} - \mathcal{A}^H)f_1 + f_1(\mathcal{B} - \mathcal{A}^H)\tilde{f}_2 + 2(df_1)\sigma^{\mathcal{B}-\mathcal{A}^H}(d\tilde{f}_2).$$

Therefore to show $(\mathcal{B} - \mathcal{A}^H)$ is vertical we only need to prove

$$f_1(\mathcal{B} - \mathcal{A}^H)\tilde{f}_2 + 2(df_1)\sigma^{\mathcal{B}-\mathcal{A}^H}(d\tilde{f}_2) = 0.$$

Recall Lemma 1.1 and use the natural extension of $\sigma^{\mathcal{A}}$ to $\sigma^{\mathcal{A}} : E^* \to E$ and the fact that by (3) $h \circ \sigma_x^{\mathcal{A}} = \sigma^{\mathcal{B}}(T_u\pi)^*$ to see

$$\begin{aligned}\sigma^{\mathcal{A}^H}(d\tilde{f}_2) &= \left(h \circ \sigma^{\mathcal{A}} h^*\right)(df_2 \circ T\pi) = h \circ \sigma^{\mathcal{A}} df_2 \\ &= \sigma^{\mathcal{B}}(df_2 \circ T\pi) = \sigma^{\mathcal{B}}(d\tilde{f}_2),\end{aligned}$$

and so $\sigma^{(\mathcal{B}-\mathcal{A}^H)}(d\tilde{f}_2) = 0$. Also by equation (1)

$$(\mathcal{B} - \mathcal{A}^H)\tilde{f}_2 = \mathcal{A}f_2 - \mathcal{A}^H\tilde{f}_2 = 0.$$

This shows that $\mathcal{B} - \mathcal{A}^H$ is vertical. ∎

Define $\alpha : P \to \mathfrak{g} \otimes \mathfrak{g}$ and $\beta : P \to \mathfrak{g}$ by

$$\alpha(u) = \sum \alpha^{ij}(u) A_i \otimes A_j$$

$$\beta(u) = \sum \beta^k(u) A_k.$$

It is easy to see that $\mathcal{B}^V$ depends only on α, β and the expression is independent of the choice of basis of $\mathfrak{g}$. From the invariance of $\mathcal{B}$ we obtain

$$\begin{aligned} \alpha(ug) &= (ad(g) \otimes ad(g))\, \alpha(u), \\ \beta(ug) &= ad(g)\beta(u) \end{aligned}$$

for all $u \in P$ and $g \in G$.

C. Theorem 2.4 has a more directly probabilistic version. For this let $\pi : P \longrightarrow M$ be as before and for $0 \leq l < r < \infty$ let $C(l, r; P)$ be the space of continuous paths $y : [l, r] \longrightarrow P$ with its usual Borel σ-algebra. For such write $l_y = l$ and $r_y = r$. Let $C(*, *; P)$ be the union of such spaces. It has the standard additive structure under concatenation: if y and y' are two paths with $r_y = l_{y'}$ and $y(r_y) = y'(l_{y'})$ let $y + y'$ be the corresponding element in $C(l_y, r_{y'}; P)$. The *basic* σ-algebra of $C(*, *, P)$ is defined to be the pull back by π of the usual Borel σ-algebra on $C(*, *; M)$.

Consider the laws $\{\mathbb{P}_a^{l,r} : 0 \leq l < r, a \in P\}$ of the process running from a between times l and r, associated to a smooth diffusion operator $\mathcal{B}$ on P. Assume for simplicity that the diffusion has no explosion. Thus $\{\mathbb{P}_a^{l,r}, a \in P\}$ is a kernel from P to $C(l, r; P)$. The right action R_g by g in G extends to give a right action, also written R_g, of G on $C(*, *, P)$. Equivariance of $\mathcal{B}$ is equivalent to

$$\mathbb{P}_{ag}^{l,r} = (R_g)_* \mathbb{P}_a^{l,r}$$

for all $0 \leq l \leq r$ and $a \in P$. If so $\pi_*(\mathbb{P}_a^{l,r})$ depends only on $\pi(a)$, l, r and gives the law of the induced diffusion $\mathcal{A}$ on M. We say that such a diffusion $\mathcal{B}$ is *basic* if for all $a \in P$ and $0 \leq l < r < \infty$ the basic σ-algebra on $C(l, r; P)$ contains all Borel sets up to $\mathbb{P}_a^{l,r}$ negligible sets, i.e. for all $a \in P$ and Borel subsets B of $C(l, r; P)$ there exists a Borel subset A of $C(l, r, M)$ s.t. $\mathbb{P}_a(\pi^{-1}(A) \Delta B) = 0$.

For paths in G it is more convenient to consider the space $C_{id}(l, r; G)$ of continuous $\sigma : [l, r] \longrightarrow G$ with $\sigma(l) = id$ for 'id' the identity element. The corresponding space $C_{id}(*, *, G)$ has a multiplication

$$C_{id}(s, t; G) \times C_{id}(t, u; G) \longrightarrow C_{id}(s, u; G)$$

$$(g, g') \mapsto g \times g'$$

where $(g \times g')(r) = g(r)$ for $r \in [s, t]$ and $(g \times g')(r) = g(t)g'(r)$ for $r \in [t, u]$.

Given probability measures $\mathbb{Q}$, $\mathbb{Q}'$ on $C_{id}(s, t; G)$ and $C_{id}(t, u; G)$ respectively this determines a convolution $\mathbb{Q} * \mathbb{Q}'$ of $\mathbb{Q}$ with $\mathbb{Q}'$ which is a probability measure on $C_{id}(s, u; G)$.

Theorem 2.5. *Given the laws* $\{\mathbb{P}_a^{l,r} : a \in P, 0 \le l < r < \infty\}$ *of an equivariant diffusion* $\mathcal{B}$ *as above with* $\mathcal{A}$ *strongly cohesive there exist probability kernels* $\{\mathbb{P}_a^{H,l,r} : a \in P\}$ *from* P *to* $C(l, r; P)$, $0 \le l < r < \infty$ *and* $\mathbb{Q}_y^{l,r}$, *defined* $\mathbb{P}^{l,r}$ *a.s. from* $C(l, r, P)$ *to* $C_{id}(l, r; G)$ *such that*

(i) $\{\mathbb{P}_a^{H,l,r} : a \in P\}$ *is equivariant, basic and determining a strongly cohesive generator.*

(ii) $y \mapsto \mathbb{Q}_y^{l,r}$ *satisfies*

$$\mathbb{Q}_{y+y'}^{l_y, r_{y'}} = \mathbb{Q}_y^{l_y, r_y} * \mathbb{Q}_{y'}^{l_{y'}, r_{y'}}$$

for $\mathbb{P}^{l_y,r_y} \otimes \mathbb{P}^{l_{y'},r_{y'}}$ *almost all* y, y' *with* $r_y = l_{y'}$.

(iii) *For* U *a Borel subset of* $C(l, r, P)$,

$$\mathbb{P}_a^{l,r}(U) = \int\int \chi_U(y_\cdot \cdot g_\cdot)\mathbb{Q}_y^{l,r}(dg)\mathbb{P}_a^{H,l,r}(dy).$$

The kernels $\mathbb{P}_a^{H,l,r}$ *are uniquely determined as are the* $\{\mathbb{Q}_y^{l,r} : y \in \mathbb{R}\}$, $\mathbb{P}_a^{H,l,r}$ *a.s. in* y *for all* a *in* P. *Furthermore* $\mathbb{Q}_y^{l,r}$ *depends on* y *only through its projection* $\pi(y)$ *and its initial point* y_l.

Proof. Fix a in P and let $\{b_t : l \le r \le t\}$ be a process with law $\mathbb{P}_a^{l,r}$. By Theorem 2.4 we can assume that $b_\cdot$ is given by an s.d.e. of the form

$$db_t = \tilde{X}(b_t) \circ dB_t + \tilde{X}^0(b_t)dt + A(b_t) \circ d\beta_t + V(b_t)dt \tag{5}$$

where $\tilde{X} : P \times \mathbb{R}^p \to TP$ is the horizontal lift of some $X : M \times \mathbb{R}^p \to E$, $\tilde{X}^0$ is the horizontal lift of a vector field X^0 on M, while $A : P \times \mathbb{R}^1 \to TP$ and the vector field V are vertical and determine $\mathcal{B}^V$. Here $B_\cdot$ and $\beta_\cdot$ are independent Brownian motions on $\mathbb{R}^p$ and $\mathbb{R}^q$ respectively, some q, and we are using the semi-connection on P induced by $\mathcal{B}$ as in Theorem 1.3.

Let $\{\tilde{x}_t : l \le t \le r\}$ satisfy

$$\begin{aligned} d\tilde{x}_t &= \tilde{X}(\tilde{x}_t) \circ dB_t + \tilde{X}^0(\tilde{x}_t)dt \\ \tilde{x}_l &= a \end{aligned} \tag{6}$$

so $\tilde{x}_\cdot$ is the horizontal lift of $\{\pi(b_t) : l \le t \le r\}$. Then there is a unique continuous process $\{g_t : l \le t \le r\}$ in G with $g_l = id$ such that

$$\tilde{x}_t g_t = b_t.$$

We have to analyse $\{g_t : l \le t \le r\}$. Using local trivialisations of $\pi : P \to M$ we see it is a semi-martingale. As in [9], Proposition 3.1 on page 69,

$$db_t = TR_{g_t}(\circ d\tilde{x}_t) + A^{g_t^{-1} \circ dg_t}(b_t)$$

giving

$$\tilde{\omega}(\circ db_t) = \tilde{\omega}\left(A^{g_t^{-1}\circ dg_t}(b_t)\right) = g_t^{-1} \circ dg_t$$

for any smooth connection form $\tilde{\omega} : P \to \mathfrak{g}$ on P which vanishes on $H^E TP$. Thus

$$\begin{aligned} dg_t &= TL_{g_t}\tilde{\omega}\big(A(\tilde{x}_t g_t) \circ d\beta_t + V(\tilde{x}_t g_t)dt\big) \\ g_l &= id, \qquad l \le t \le r. \end{aligned} \tag{7}$$

For $y \in C(l, r : P)$ let $\{g_t^y : l \le t \le r\}$ be the solution of

$$\begin{aligned} dg_t^y &= TL_{g_t^y}\tilde{\omega}\big(A(y_t g_t^y) \circ d\beta_t + V(y_t g_t^y)dt\big) \\ g_l^y &= id \end{aligned} \tag{8}$$

(where the Stratonovich equation is interpreted with '$dy_t d\beta_t = 0$'). Since $\beta.$ and $B.$ and hence $\beta.$ and $\tilde{x}.$ are independent we see $g = g^{\tilde{x}}$ almost surely. For a discussion of some technicalities concerning skew products, see [16].

For $y.$ in $C(*, *; P)$ let $\{h(y)_t : l_y \le t \le r_y\}$ be the horizontal lift of $\pi(y).$, starting at y_{l_y}. This exists for almost all y as can be seen either by the extension of Itô's result to general principal bundles, e.g. using (6), or by the existence of measurable sections using the fact that $\mathcal{A}^H$ is basic. Define $\mathbb{P}_a^{H,l,r}$ to be the law of $\tilde{x}.$ above and $Q_y^{l,r}$ to be that of $g^{h(y)}$. Clearly conditions (i) is satisfied.

To check (ii) take y and y' with $r_y = l_{y'}$. Then

$$h(y + y') = h(y) + h(y')\left(g_{r_y}^{h(y)}\right)^{-1},$$

writing $y = h(y)g^{h(y)}$ and $y' = h(y')g^{h(y')}$. For $r_y \le t \le r_{y'}$ this shows

$$(y + y')_t = h(y')_t\left(g_{r_y}^{h(y)}\right)^{-1} g_t^{h(y+y')}.$$

But $(y + y')_t = y'_t = h(y')_t g_t^{h(y')}$ and so we have $g_t^{h(y+y')} = g_{r_y}^{h(y)} g_t^{h(y')}$ for $t \ge r_y$, giving $g^{h(y+y')} = g^{h(y)} \times g^{h(y')}$ almost surely. This proves (ii).

For uniqueness suppose we have another set of probability measures denoted $\tilde{\mathbb{Q}}_y^{l,r}$ and $\tilde{P}_a^{H,l,r}$ which satisfy (i), (ii), (iii). Since $\{\tilde{\mathbb{P}}_a^{H,l,r}\}_a$ is equivariant and induces $\mathcal{A}$ on M we can apply the preceding argument to it in place of $\{\mathbb{P}_a^{l,r}\}_a$. However since it is basic the term involving β in the stochastic differential equation (6) must vanish. Since it is also strongly cohesive the vertical part V must vanish also and we have $\tilde{\mathbb{P}}_a^{H,l,r} = \mathbb{P}_a^{H,l,r}$. On the other hand in the decomposition $b_t = \tilde{x}_t g_t^{\tilde{x}_t}$ the law of $g^{\tilde{x}}_{\cdot}$ is determined by those of $b.$ and $\tilde{x}.$ but $\mathbb{Q}_y^{l,r}$ is the conditional law of $g^{\tilde{x}}_{\cdot}$ given $\tilde{x} = y.$ and so is uniquely determined as described. ■

In fact $\mathbb{Q}_y^{l,r}$ is associated to the time dependent generator which at $g \in G$ and $t \in [l,r]$ is $\sum \alpha^{ij}(h(y)_t g)\mathcal{L}_{A_i}\mathcal{L}_{A_j} + \sum \beta^k(h(y)_t g)\mathcal{L}_{A_k}$ for α^{ij} and β^k as defined in Theorem 2.4 while $\mathbb{P}^{H,l,r}$ is associated to $\mathcal{A}^H$.

§3. Stochastic flows and derivative flows

A. Derivative flows. Let $\mathcal{A}$ on M be given in Hörmander form

$$\mathcal{A} = \frac{1}{2}\sum_{j=1}^{m} \mathcal{L}_{X^j}\mathcal{L}_{X^j} + \mathcal{L}_A$$

for some vector fields $X^1, \dots X^m$, A. As before let $E_x = span\{X^1(x), \dots, X^m(x)\}$ and assume $dimE_x$ is constant, p, say, giving a sub-bundle $E \subset TM$. The $X^1(x), \dots, X^m(x)$ determine a vector bundle map of the trivial bundle $\underline{\mathbb{R}}^m$

$$X : \underline{\mathbb{R}}^m \longrightarrow TM$$

with $\sigma^{\mathcal{A}} = X(x)X(x)^*$. We can, and will, consider X as a map $X : \underline{\mathbb{R}}^m \to E$.

As such it determines (a) a Riemannian metric $\{\langle , \rangle_x : x \in M\}$ on E (the same as that determined by $\sigma^{\mathcal{A}}$) and (b) a metric connection $\breve{\nabla}$ on E uniquely defined by the requirement that for each x in M,

$$\breve{\nabla}_v X(e) = 0$$

for all $v \in T_xM$ whenever e is orthogonal to the kernel of T_xM. Then for any differentiable section U of E,

$$\breve{\nabla}_v U = Y(x)d\left(Y_\cdot(U(\cdot))\right)(v), \qquad v \in T_xM, \tag{9}$$

where Y is the $\mathbb{R}^m$ valued 1-form on M given by

$$\langle Y_x(v), e\rangle_{\mathbb{R}^m} = \langle X(x)(e), v\rangle_x, \qquad e \in \mathbb{R}^m, v \in E_x, x \in M$$

e.g. [7] where it is referred to as the LeJan-Watanabe connection in this context. By a theorem of Narasimhan and Ramanan [14] all metric connections on E arise this way, see [15], [7].

For $\{B_t : 0 \le t < \infty\}$ a Brownian motion on $\mathbb{R}^m$, the stochastic differential equation

$$dx_t = X(x_t) \circ dB_t + A(x_t)dt \tag{10}$$

determines a Markov process with differential generator $\mathcal{A}$. Over each solution $\{x_t : 0 \le t < \rho\}$, where ρ is the explosion time, there is a 'derivative' process $\{v_t : 0 \le t < \rho\}$ in TM which we can write as $\{T\xi_t(v_0) : 0 \le t < \rho\}$

with $T\xi_t : T_{x_0}M \to T_{x_t}M$ linear. This would be the derivative of the flow $\{\xi_t : 0 \le t < \rho\}$ of the stochastic differential equation when the stochastic differential equation is strongly complete. In general it is given by a stochastic differential equation on the tangent bundle TM, or equivalently by a covariant equation along $\{x_t : 0 \le t < \rho\}$:

$$Dv_t = \nabla X(v_t) \circ dB_t + \nabla A(v_t)dt$$

with respect to any torsion free connection. Take P to be the linear frame bundle $GL(M)$ of M, treating $u \in GL(M)$ as an isomorphism $u : \mathbb{R}^n \to T_{\pi(u)}M$. For $u_0 \in GLM$ we obtain a process $\{u_t : 0 \le t < \rho\}$ on GLM by

$$u_t = T\xi_t \circ u_0.$$

Let $\mathcal{B}$ be its differential generator. Clearly it is equivariant and a lift of $\mathcal{A}$.

A proof of the following in the context of stochastic flows, is given later. For $w \in E_x$, set

$$Z^w(y) = X(y)Y(x)(w). \tag{11}$$

Theorem 3.1. *The semi-connection ∇ induced by $\mathcal{B}$ is the adjoint connection of the LeJan-Watanabe connection $\breve{\nabla}$ determined by X, as defined by (9),* [7]. *Consequently $\nabla_w V = L_{Z^w} V$ for any vector field V and $w \in E$ also $\nabla_{V(x)} Z^w$ vanishes if $w \in E_x$.*

In the case of the derivative flow the α, β of Theorem 2.4 have an explicit expression: for $u \in GLM$,

$$\begin{cases} \alpha(u) = \frac{1}{2}\sum \left(u^{-1}(-)\breve{\nabla}_{u(-)}X^p\right) \otimes \left(u^{-1}(-)\breve{\nabla}_{u(-)}X^p\right) \\ \beta(u) = -\frac{1}{2}\sum u^{-1}\breve{\nabla}_{\breve{\nabla}_{u(-)}X^p}X^p - \frac{1}{2}u^{-1}\text{Ric}^{\#}u(-). \end{cases} \tag{12}$$

Here $\breve{R}$ is the curvature tensor for $\breve{\nabla}$ and $\breve{Ric}^{\#} : TM \to E$ the Ricci curvature defined by $\breve{Ric}^{\#}(v) = \sum_{j=1}^{p} \breve{R}(v, e^j)e^j$, $v \in T_xM$.

Equivariant operators on GLM determine operators on associated bundles, such as $\wedge^q TM$. If the original operator was vertical this turns out to be a zero order operator (as is shown in [5] for general principal bundles) and in the case of $\wedge^q TM$ these operators are the generalized Weitzenbock curvature operators described in [7]. In particular for differential 1-forms the operator is $\phi \mapsto \phi(Ric^{\#}-)$. To see this, as an illustrative example, given a 1-form ϕ

define $\tilde{\phi} : GLM \to L(\mathbb{R}^n; \mathbb{R})$ by $\tilde{\phi}(u) = \phi_{\pi u} u$ so $\tilde{\phi}(ug) = \phi_{\pi u}(ug-)$. Then

$$\begin{aligned} L_{A_j^*}(\tilde{\phi})(u) &= \frac{d}{dt}\tilde{\phi}(u \cdot e^{A_j t})|_{t=0} \\ &= \frac{d}{dt}\phi_{\pi u}(u \cdot e^{A_j t})|_{t=0} \\ &= \phi_{\pi u}(uA_j-) = \tilde{\phi}(u)(A_j-). \end{aligned}$$

Iterating we have

$$\begin{aligned} \mathcal{B}^V(\tilde{\phi})(u) &= \sum_{i,j} \alpha^{i,j}(u)\phi_{\pi u}(uA_jA_i-) + \sum_k \beta^k(u)\phi_{\pi u}(uA_k-) \\ &= -\frac{1}{2}\tilde{\phi}(u)(u^{-1}Ric^{\#}(u-)) \end{aligned}$$

as required, by using the map $gl(n) \otimes gl(n) \to gl(n)$, $S \otimes T \mapsto S \circ T$, and equation (12).

B. Stochastic flows. In fact Theorem 3.1 can be understood in the more general context of stochastic flows as diffusions on the diffeomorphism groups. For this assume that M is compact and for $r \in \{1, 2, \dots\}$ and $s > r + \dim(M)/2$ let $\mathcal{D}^s = \mathcal{D}^s M$ be the C^∞ manifold of diffeomorphisms of M of Sobolev class H^s, (for example see Ebin-Marsden [2] or Elworthy [3].) Alternatively we could take the space $\mathcal{D}^\infty$ of C^∞ diffeomorphisms with differentiable structure as in [11]. Fix a base point x_0 in M and let $\pi : \mathcal{D}^s \to M$ be evaluation at x_0. This makes $\mathcal{D}^s$ into a principal bundle over M with group the manifold $\mathcal{D}^s_{x_0}$ of H^s- diffeomorphisms θ with $\theta(x_0) = x_0$, acting on the right by composition (although the action of $\mathcal{D}^{s+r}$ is only C^r, for $r = 0, 1, 2, \dots$).

Let $\{\xi_t^s : 0 \leq s \leq t < \infty\}$ be the flow of (10) starting at time s. Write ξ_t for ξ_t^0. The more general case allowing for infinite dimensional noise is given in [5]. We define probability measures $\{\mathbb{P}_\theta^{s,t} : \theta \in \mathcal{D}^s\}$ on $C([s.t]; M)$ be letting $\mathbb{P}_\theta^{s,t}$ be the law of $\{\xi_r^s \circ \theta : s \leq r \leq t\}$ (These correspond to the diffusion process on $\mathcal{D}^s$ associated to the right-invariant stochastic differential equation on $\mathcal{D}^s$ satisfied by $\{\xi_t : 0 \leq t < \infty\}$ as in [3].) These are equivariant and project by π to the laws given by the stochastic differential equation on M. Assuming that these give a strongly cohesive diffusion on M we are essentially in the situation of Theorem 2.5.

Let $K(x) : \mathbb{R}^m \to \mathbb{R}^m$ be the orthogonal projection onto the kernel of $X(x)$, each $x \in M$. set $K^\perp(x) = id - K(x)$. Consider the $\mathcal{D}^\infty$-valued process $\{\theta_t : 0 \leq t < \infty\}$ given by (or as the flow of)

(13) $d\theta_t(x) = X(\theta_t(x))K^\perp(\theta_t(x_0)) \circ dB_t + X(\theta_t(x))Y(\theta_t(x_0))A(\theta_t(x_0))$

for given θ_0 in $\mathcal{D}^\infty$ and, define a $\mathcal{D}^\infty_{x_0}$-valued process $\{g_t : 0 \le t < \infty\}$ by

$$
\begin{aligned}
(14) \qquad dg_t &= T\theta_t^{-1}\{X(\theta_t g_t-)K(\theta_t x_0) \circ dB_t \\
&\quad + A(\theta_t g_t-)dt - X(\theta_t g_t-)Y(\theta_t x_0)A(\theta_t x_0)dt\} \\
g_0 &= id.
\end{aligned}
$$

Set $x_t^\theta = \xi_t(\theta_0(x_0))$. Note that $\pi(\theta_t) = \theta_t(x_0) = x_t^\theta$ since

$$X(\theta_t(x_0))K^\perp(\theta_t(x_0)) = X(\theta_t(x_0))$$

and

$$X(\theta_t(x_0))Y(\theta_t(x_0))A(\theta_t(x_0)) = A(\theta_t(x_0)).$$

Thus $\{\theta_t : 0 \le t < \infty\}$ is a lift of $\{x_t^\theta, 0 \le t < \infty\}$. It can be considered to be driven by the 'relevant noise', (from the point of view of $\xi_\cdot(\theta_0(x_0))$, i.e. by the Brownian motion $\tilde{B}_\cdot$ given by

$$\tilde{B}_t = \int_0^t \tilde{/\!/}(x_\cdot^\theta)_s^{-1} K^\perp(x_s^\theta)dB_s$$

where $\{\tilde{/\!/}(x_\cdot^\theta), 0 \le s < \infty\}$ is parallel translation along $x_\cdot^\theta$ with respect to the connection on the trivial bundle $M \times \mathbb{R}^m \to M$ determined by K and $K^\perp$, so that

$$\tilde{/\!/}(x_\cdot^\theta)_s : \mathbb{R}^m \to \mathbb{R}^m$$

is orthogonal and maps the kernel of $X(\theta_\cdot(x_0))$ onto the kernel of $X(x_s^\theta)$ for $0 \le s < \infty$, see [7](chapter 3).

Correspondingly there is the 'redundant noise', the Brownian motion $\{\beta_t : 0 \le t < \infty\}$ given by

$$\beta_t = \int_0^t \tilde{/\!/}(x_\cdot^\theta)_s^{-1} K(x_s^\theta)dB_s.$$

Then, as shown in [7](chapter 3),

(i) $\tilde{B}_\cdot$ has the same filtration as $\{x_s^\theta : 0 \le s < \infty\}$

(ii) $\beta_\cdot$ and $\tilde{B}_\cdot$ are independent

(iii) $dB_t = \tilde{/\!/}_t d\beta_t + \tilde{/\!/}_t d\tilde{B}_t$.

We wish to see how $g_\cdot$ is driven by $\beta_\cdot$. For this observe

$$\int_0^t K(x_s^\theta) \circ dB_s = \int_0^t K(x_s^\theta)dB_s + \int_0^t \Lambda(x_s^\theta)ds$$

for $\Lambda : M \to \mathbb{R}$ given by the Stratonovich correction term. By (iii)

$$\int_0^t K(x_s^\theta)dB_s = \int_0^t \tilde{/\!/}_s d\beta_s = \int_0^t \tilde{/\!/}_s \circ d\beta_s$$

since $\tilde{/\!/}.$ is independent of β by (i) and (ii). Thus equation (14) for $g.$ can be written as

$$\begin{aligned} dg_t &= T\theta_t^{-1}\left\{X(\theta_t g_t-)\tilde{/\!/}(\theta.(x_0))_t \circ d\beta_t + X(\theta_t g_t-)\Lambda(\theta_t(x_0))dt \right. \\ &\quad \left. + A(\theta_t g_t-)dt - X(\theta_t g_t-)Y(\theta_t x_0)A(\theta_t x_0)dt\right\} \end{aligned}$$

and if we define

$$\begin{aligned} dg_t^y &= Ty_t^{-1}\left\{X(y_t g_t-)\tilde{/\!/}(y.(x_0))_t \circ d\beta_t + X(y_t g_t-)\Lambda(y_t(x_0))dt \right. \\ &\quad \left. + A(y_t g_t-)dt - X(y_t g_t-)Y(y_t x_0)A(y_t x_0)dt\right\} \\ g_0 &= id \end{aligned}$$

for any continuous $y : [0,\infty) \to \mathcal{D}^\infty$, we see, by the independence of β and θ that $g. = g^\theta_.$.

By Itô's formula on $\mathcal{D}^s$, for $x \in M$,

$$d(\theta_t g_t(x_0)) = (\circ d\theta_t)(g_t(x)) + T\theta_t(\circ dg_t^\theta(x)).$$

Now

$$\begin{aligned} T\theta_t(\circ dg_t^\theta(x)) &= \{X(\theta_t g_t(x))K(\theta_t x_0) \circ dB_t \\ &\quad + A(\theta_t g_t(x))dt - X(\theta_t g_t(x))Y(\theta_t x_0)A(\theta_t x_0)dt\} \end{aligned}$$

and so by (13) we see that $\theta_t g_t = \xi_t \circ \theta_0$, a.s.

Taking $\theta_0 = id$ we have

Proposition 3.2. *The flow $\xi.$ has the decomposition*

$$\xi_t = \theta_t g_t^{\theta.}, \qquad 0 \le t < \infty$$

for θ and $g^\theta_. \equiv g.$ given by (13) and (14) above. For almost all $\sigma : [0,\infty) \to M$ with $\sigma(0) = x_0$ and bounded measurable $F : C(0,\infty;\mathcal{D}^\infty) \to \mathbb{R}$

$$\mathbb{E}\{F(\xi.)\,|\xi.(x_0) = \sigma\} = \mathbb{E}\{F(\tilde{\sigma} g^{\tilde{\sigma}}_.)\}$$

where $\tilde{\sigma} : [0,\infty) \to \mathcal{D}^\infty$ is the horizontal lift of σ with $\tilde{\sigma}(0) = id$.

To define the 'horizontal lift' above we can use the fact, from (i) above, that $\theta.$ has the same filtration as $\xi.(x_0)$ and so furnishes a lifting map.

In terms of the semi-connection induced on $\pi : \mathcal{D}^s \to M$ over E, from above, by uniqueness or directly, we see the horizontal lift

$$\begin{aligned} h_\theta &: E_{\theta(x_0)} \longrightarrow T_\theta \mathcal{D}^s \\ h_\theta(v) &: M \longrightarrow TM \end{aligned}$$

is given by $h_\theta(v) = X(\theta(x))Y(\theta(x_0))v$ and the horizontal lift $\tilde{\sigma}.$ from $\tilde{\sigma}_0$ of a C^1 curve σ on M with $\tilde{\sigma}_0(x_0) = \sigma_0$ and $\dot{\sigma}(t) \in E_{\sigma(t)}$, all t, is given by

$$\frac{d}{dt}\tilde{\sigma}_t = X(\tilde{\sigma}_t-)Y(\sigma_t)\dot{\sigma}_t$$

for $\tilde{\sigma}_0 = id$. The lift is the solution flow of the differential equation

$$\dot{y}_t = Z^{\dot{\sigma}}(y_t)$$

on M.

For each frame $u : \mathbb{R}^n \to T_{x_0}M$ there is a homomorphism of principal bundles

$$\begin{array}{ll} \mathcal{D}^s & \to GL(M) \\ \theta & \mapsto T_{x_0}\theta \circ u. \end{array} \tag{15}$$

This sends $\{\xi_t : t \geq 0\}$ to the derivative process $T_x\xi_t \circ u$. (If the latter satisfies the strongly cohesive condition we could apply our analysis to this submersion $\mathcal{D}^s \to GLM$ and get another decomposition of $\xi.$.)

Results in Kobayashi-Nomizu [9] (Proposition 6.1 on page 79) apply to the homomorphism $\mathcal{D}^s \to GL(M)$ of (15). This gives a relationship between the curvature and holonomy groups of the semi-connection $\hat{\nabla}$ on GLM determined by the derivative flow and those of the connection induced by the diffusion on $\mathcal{D}^s \xrightarrow{\pi} M$. It also shows that the horizontal lift $\{\tilde{x}_t : t \geq 0\}$ through u of $\{x_t : t \geq 0\}$ to $GL(M)$ is just $T_{x_0}\theta_t \circ u$ for $\{\theta_t : t \geq 0\}$ the flow given by (13) with $\theta_0 = id$, i.e. the solution flow of the stochastic differential equation

$$dy_t = Z^{\circ dx_t}(y_t).$$

From this and Lemma 1.3.4 of [7] we see that $\hat{\nabla}$ is the adjoint of the LeJan-Watanabe connection determined by the flow, so proving Theorem 3.1 above. However the present construction applies with GLM replaced by any natural bundle over M (e.g. jet bundles, see Kolar-Michor-Slovak [10]), to give semi-connections on these bundles.

References

[1] L. Accardi, A. Mohari, and Centro V. Volterra. On the structure of classical and quantum flows. *J. Funct. Anal.*, 135(2):421–455, 1996.

[2] D. G. Ebin and J. Marsden. Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. Math.*, pages 102–163, 1970.

[3] K. D. Elworthy. *Stochastic Differential Equations on Manifolds*. LMS Lecture Notes Series 70, Cambridge University Press, 1982.

[4] K. D. Elworthy and W. S. Kendall. Factorization of harmonic maps and Brownian motions. In *From local times to global geometry, control and physics (Coventry, 1984/85), Pitman Res. Notes Math. Ser., 150*, pages 75–83. Longman Sci. Tech., Harlow, 1986.

[5] K. D. Elworthy, Y. LeJan, and Xue-Mei Li. In preparation.

[6] K. D. Elworthy, Y. LeJan, and Xue-Mei Li. Concerning the geometry of stochastic differential equations and stochastic flows. In *'New Trends in stochastic Analysis', Proc. Taniguchi Symposium, Sept. 1994, Charingworth, ed. K. D. Elworthy and S. Kusuoka, I. Shigekawa*. World Scientific Press, 1996.

[7] K. D. Elworthy, Y. LeJan, and Xue-Mei Li. *On the geometry of diffusion operators and stochastic flows, Lecture Notes in Mathematics 1720*. Springer, 1999.

[8] K. Itô. The Brownian motion and tensor fields on Riemannian manifold. In *Proc. Internat. Congr. Mathematicians (Stockholm, 1962) Inst. Mittag-Leffler, Djursholm*, pages 536–539, 1963.

[9] S. Kobayashi and K. Nomizu. *Foundations of differential geometry, Vol. I*. Interscience Publishers, 1969.

[10] I. Kolar, P. W. Michor, and J. Slovak. *Natural operations in differential geometry*. Springer-Verlag, Berlin, 1993.

[11] A. Kriegl and P. W. Michor. *The convenient setting of global analysis. Mathematical Surveys and Monographs. 53*. American Mathematical Society, 1997.

[12] Y. LeJan and S. Watanabe. Stochastic flows of diffeomorphisms. In *Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, 32,*, pages 307–332. North-Holland, Amsterdam, 1984.

[13] M. Liao. Factorization of diffusions on fibre bundles. *Trans. Amer. Math. Soc.*, 311(2):813–827, (1989).

[14] M. S. Narasimhan and S. Ramanan. Existence of universal connections. *American J. Math.*, 83, 1961.

[15] D. Quillen. Superconnection character forms and the Cayley transform. *Topology*, 27(2):211–238, 1988.

[16] J. C. Taylor. Skew products, regular conditional probabilities and stochastic differential equation: a technical remark. In *Sḿinaire de probabilités, XXVI, Lecture Notes in Mathematics 1526, ed. J. Azéma, P. A. Meyer and M. Yor*, pages 113–126, 1992.

K. D. Elworthy
Mathematics Institute, Warwick University,
Coventry CV4 7AL, UK

Y. Le Jan
Départment de Mathématique, Université Paris Sud,
91405 Orsay, France

Xue-Mei Li
Department of Computing and Mathematics, The Nottingham Trent University,
Nottingham NG7 1AS, UK. e-mail address: xuemei.li@ntu.ac.uk

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 49–74

Monge-Kantorovitch Measure Transportation, Monge-Ampère Equation and the Itô Calculus

Denis Feyel and Ali Süleyman Üstünel

Dedicated to Professor Kiyosi Itô for his 88th birthday

Abstract.

Let (W, μ, H) be an abstract Wiener space assume two ν_i, $i = 1, 2$ probabilities on $(W, \mathcal{B}(W))$. Assume that the Wasserstein distance between ν_1 and ν_2 with respect to the Cameron-Martin norm

$$d_H(\nu_1, \nu_2) = \left\{ \inf_\beta \int_{W \times W} |x - y|_H^2 d\beta(x, y) \right\}^{1/2}$$

is finite, where the infimum is taken on the set of probability measures β on $W \times W$ whose first and second marginals are ν_1 and ν_2 and that ν_1 has regular disintegration along a sequence of finite dimensional projections. Then there exists a unique (cyclically monotone) map $T = I_W + \xi$, with $\xi : W \to H$, such that T maps ν_1 to ν_2 and the measure $\gamma = (I \times T)\nu_1$ is the unique solution of the Monge-Kantorovitch problem. Besides, if $\nu_2 \ll \mu$[1], then T is stochastically invertible, i.e., there exists $S : W \to W$ such that $S \circ T = I_W$ ν_1 a.s. and $T \circ S = I_W$ ν_2 a.s. If $\nu_1 = \mu$, then there exists a 1-convex function ϕ in the Gaussian Sobolev space $\mathbb{D}_{2,1}$, such that $\xi = \nabla\phi$. These results imply that the quasi-invariant transformations of the Wiener space with finite Wasserstein distance from μ can be written as the composition of a transport map T and a rotation, i.e., a measure preserving map. We give also 1-convex sub-solutions using by calculating the Gaussian jacobian. Finally the full solutions of the Monge-Ampère equation on W are given with the help of the Itô calculus.

§1. Introduction

In 1781, Gaspard Monge has published his celebrated memoire about the most economical way of earth-moving [20]. The configurations of

Received March 6, 2003.

[1] In fact this hypothesis is too strong, cf. Theorem 4.1.

excavated earth and remblai were modelized as two measures of equal mass, say ρ and ν, that Monge had supposed absolutely continuous with respect to the volume measure. Later Ampère has studied an analogous question about the electricity current in a media with varying conductivity. In modern language of measure theory we can express the problem in the following terms: let W be a Polish space on which are given two positive measures ρ and ν, of finite, equal mass. Let $c(x, y)$ be a cost function on $W \times W$, which is, usually, assumed to be positive. Does there exist a map $T : W \to W$ such that $T\rho = \nu$ and T minimizes the integral

$$\int_W c(x, T(x))d\rho(x)$$

between all such maps? The problem has been further studied by Appell [1, 2] and by Kantorovitch [16]. Kantorovitch has succeeded to transform this highly nonlinear problem of Monge into a linear problem by replacing the search for T with the search of a measure γ on $W \times W$ with marginals ρ and ν such that the integral

$$\int_{W\times W} c(x, y)d\gamma(x, y)$$

is the minimum of all the integrals

$$\int_{W\times W} c(x, y)d\beta(x, y)$$

where β runs in the set of probability measures on $W \times W$ whose marginals are ρ and ν. Since then the problem addressed above is called the Monge problem and the quest of the optimal measure is called the Monge-Kantorovitch problem.

In this paper we survey and complete the recent results (cf. [12, 13]) about the Monge-Kantorovitch and the Monge problem in the frame of an abstract Wiener space with a (infinitesimal) cost function c on $W \times W$, which is singular with respect to the natural Fréchet topology of this space. Let us explain all this more rigourously: let W be a separable Fréchet space with its Borel sigma algebra $\mathcal{B}(W)$ and assume that there is a separable Hilbert space H which is injected densely and continuously into W, thus the topology of H is, in general, stronger than the topology induced by W. The cost function $c : W \times W \to \mathbb{R}_+ \cup \{\infty\}$ is defined as

$$c(x, y) = |x - y|_H^2 ,$$

we suppose that $c(x, y) = \infty$ if $x - y$ does not belong to H. Clearly, this choice of the function c is not arbitrary, in fact it is closely related to Itô Calculus, hence also to the problems originating from Physics, quantum chemistry, large deviations, etc. Since for all the interesting measures on W, the Cameron-Martin space is a negligeable set, the cost function will be infinity very frequently. Let $\Sigma(\rho, \nu)$ denote the set of probability measures on $W \times W$ with given marginals ρ and ν. It is a convex, compact set under the weak topology $\sigma(\Sigma, C_b(W \times W))$. As explained above, the problem of Monge consists of finding a measurable map $T : W \to W$, called the optimal transport of ρ to ν, i.e., $T\rho = \nu$[2)] which minimizes the total cost

$$U \to \int_W |x - U(x)|_H^2 d\rho(x)\,,$$

between all the maps $U : W \to W$ such that $U\rho = \nu$. The Monge-Kantorovitch problem will consist of finding a measure on $W \times W$, which minimizes the function $\theta \to J(\theta)$, defined by

$$J(\theta) = \int_{W \times W} |x - y|_H^2 d\theta(x, y)\,, \tag{1.1}$$

where θ runs in $\Sigma(\rho, \nu)$. Note that $\inf\{J(\theta) : \theta \in \Sigma(\rho, \nu)\}$ is the square of Wasserstein metric $d_H(\rho, \nu)$ with respect to the Cameron-Martin space H.

Any solution γ of the Monge-Kantorovitch problem will give a solution to the Monge problem provided that its support is included in the graph of a map. Hence our work consists of realizing this program. Although in the finite dimensional case this problem is well-studied in the path-breaking papers of Brenier [4] and McCann [18, 19], cf. also [25, 26], the things do not come up easily in our setting and the difficulty is due to the fact that the cost function is not continuous with respect to the Fréchet topology of W, for instance the weak convergence of the probability measures does not imply the convergence of the integrals of the cost function. In other words the function $|x - y|_H^2$ takes the value plus infinity "very often". On the other hand the results we obtain seem to have important applications to several problems of stochastic analysis that we shall explain while enumerating the contents of the paper.

In Section 2, are given the basic results of functional analysis on the Wiener space (cf., for instance [10, 28]) and the related probabilistic theory of convex functions developed in [11]. Section 3 deals with the derivation of some inequalities which control the Wasserstein distance.

---

2) We denote the push-forward of ρ by T, i.e., the image of ρ under T, by $T\rho$.

In particular, with the help of the Girsanov theorem and the Itô calculus, we give a very simple proof of an inequality, initially discovered by Talagrand ([24]), some simple applications are also illustrated. The facility with which one obtains these results gives an idea about the efficiency of the infinite dimensional techniques, namely the Itô calculus for the Monge-Kantorovitch like problems.

In Section 4 we give the full statement for the existence and the uniqueness of solution of the Monge problem and the uniqueness of the solution of the Monge-Kantorovitch problem under the hypothesis that the Wasserstein distance is finite. We have avoided to give the corresponding proofs which are quite technical (cf. [13]), however all the applications are provided with proofs or explanations in such a way that the reader can have an idea about how to do it.

Section 5 studies the Monge-Ampère equation for the measures which are absolutely continuous with respect to the Wiener measure. First we define the Alexandrov versions of the Ornstein-Uhlenbeck operator and the second order Sobolev derivatives for 1-convex Wiener maps. With the help of these, we write the corresponding Jacobian using the modified Carleman-Fredholm determinant which is natural in the infinite dimensional case (cf., [29]). Here we have a major difficulty which originates from the pathology of the Radon-Nikodym derivatives of the vector measures with respect to a scalar measure: in fact even if the second order Sobolev derivative of a Wiener function is a vector measure with values in the space of Hilbert-Schmidt operators, its absolutely continuous part has no reason to be Hilbert-Schmidt. Hence the Carleman-Fredholm determinant may not exist, however due to the 1-convexity, the detereminants of the approximating sequence are all with values in the interval $[0, 1]$. Consequently we can construct the subsolutions with the help of the Fatou lemma.

Last but not the least, in section 6, we remark that all these difficulties can be overcome thanks to the natural renormalization of the Itô stochastic calculus. In fact using the Itô representation theorem and the Wiener space analysis extended to the distributions [27], we can give the explicit solution of the Monge-Ampère equation. This is a remarkable result in the sense that such techniques do not exist in the finite dimensional case.

§2. Preliminaries and notations

Let W be a separable Fréchet space equipped with a Gaussian measure μ of zero mean whose support is the whole space. The corresponding Cameron-Martin space is denoted by H. Recall that the injection $H \hookrightarrow W$ is compact and its adjoint is the natural injection $W^\star \hookrightarrow H^\star \subset L^2(\mu)$. The triple (W, μ, H) is called an abstract Wiener space. Recall that $W = H$ if and only if W is finite dimensional. A subspace F of H is called regular if the corresponding orthogonal projection has a continuous extension to W, denoted again by the same letter. It is well-known that there exists an increasing sequence of regular subspaces $(F_n, n \geq 1)$, called total, such that $\cup_n F_n$ is dense in H and in W. Let $\sigma(\pi_{F_n})$[3] be the σ-algebra generated by π_{F_n}, then for any $f \in L^p(\mu)$, the martingale sequence $(E[f|\sigma(\pi_{F_n})], n \geq 1)$ converges to f (strongly if $p < \infty$) in $L^p(\mu)$. Observe that the function $f_n = E[f|\sigma(\pi_{F_n})]$ can be identified with a function on the finite dimensional abstract Wiener space (F_n, μ_n, F_n), where $\mu_n = \pi_n \mu$.

Since the translations of μ with the elements of H induce measures equivalent to μ, the Gâteaux derivative in H direction of the random variables is a closable operator on $L^p(\mu)$-spaces and this closure will be denoted by ∇ cf., for example [10, 28]. The corresponding Sobolev spaces (the equivalence classes) of the real random variables will be denoted as $\mathbb{D}_{p,k}$, where $k \in \mathbb{N}$ is the order of differentiability and $p > 1$ is the order of integrability. If the random variables are with values in some separable Hilbert space, say Φ, then we shall define similarly the corresponding Sobolev spaces and they are denoted as $\mathbb{D}_{p,k}(\Phi)$, $p > 1$, $k \in \mathbb{N}$. Since $\nabla : \mathbb{D}_{p,k} \to \mathbb{D}_{p,k-1}(H)$ is a continuous and linear operator its adjoint is a well-defined operator which we represent by δ. In the case of classical Wiener space, i.e., when $W = C(\mathbb{R}_+, \mathbb{R}^d)$, then δ coincides with the Itô integral of the Lebesgue density of the adapted elements of $\mathbb{D}_{p,k}(H)$ (cf. [28]).

For any $t \geq 0$ and measurable $f : W \to \mathbb{R}_+$, we note by

$$P_t f(x) = \int_W f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) \mu(dy),$$

it is well-known that $(P_t, t \in \mathbb{R}_+)$ is a hypercontractive semigroup on $L^p(\mu), p > 1$, which is called the Ornstein-Uhlenbeck semigroup (cf. [10, 28]). Its infinitesimal generator is denoted by $-\mathcal{L}$ and we call $\mathcal{L}$ the Ornstein-Uhlenbeck operator (sometimes called the number operator by the physicists). The norms defined by

$$\|\phi\|_{p,k} = \|(I + \mathcal{L})^{k/2}\phi\|_{L^p(\mu)} \tag{2.2}$$

are equivalent to the norms defined by the iterates of the Sobolev derivative ∇. This observation permits us to identify the duals of the space

[3] For the notational simplicity, in the sequel we shall denote it by π_{F_n}.

$\mathbb{D}_{p,k}(\Phi); p > 1$, $k \in \mathbb{N}$ by $\mathbb{D}_{q,-k}(\Phi')$, with $q^{-1} = 1 - p^{-1}$, where the latter space is defined by replacing k in (2.2) by $-k$, this gives us the distribution spaces on the Wiener space W (in fact we can take as k any real number). An easy calculation shows that, formally, $\delta \circ \nabla = \mathcal{L}$, and this permits us to extend the divergence and the derivative operators to the distributions as linear, continuous operators. In fact $\delta : \mathbb{D}_{q,k}(H \otimes \Phi) \to \mathbb{D}_{q,k-1}(\Phi)$ and $\nabla : \mathbb{D}_{q,k}(\Phi) \to \mathbb{D}_{q,k-1}(H \otimes \Phi)$ continuously, for any $q > 1$ and $k \in \mathbb{R}$, where $H \otimes \Phi$ denotes the completed Hilbert-Schmidt tensor product (cf., for instance [28]).

Let us recall some facts from the convex analysis. Let K be a Hilbert space, a subset S of $K \times K$ is called cyclically monotone if any finite subset $\{(x_1, y_1), \ldots, (x_N, y_N)\}$ of S satisfies the following algebraic condition:

$$\langle y_1, x_2 - x_1\rangle + \langle y_2, x_3 - x_2\rangle + \cdots + \langle y_{N-1}, x_N - x_{N-1}\rangle + \langle y_N, x_1 - x_N\rangle \leq 0\,,$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product of K. It turns out that S is cyclically monotone if and only if

$$\sum_{i=1}^{N} (y_i, x_{\sigma(i)} - x_i) \leq 0\,,$$

for any permutation σ of $\{1, \ldots, N\}$ and for any finite subset $\{(x_i, y_i) : i = 1, \ldots, N\}$ of S. Note that S is cyclically monotone if and only if any translate of it is cyclically monotone. By a theorem of Rockafellar, any cyclically monotone set is contained in the graph of the subdifferential of a convex function in the sense of convex analysis ([22]) and even if the function may not be unique its subdifferential is unique.

Let now (W, μ, H) be an abstract Wiener space; a measurable function $f : W \to \mathbb{R} \cup \{\infty\}$ is called 1-convex if the map

$$h \to f(x+h) + \frac{1}{2}|h|_H^2 = F(x,h)$$

is convex on the Cameron-Martin space H with values in $L^0(\mu)$. Note that this notion is compatible with the μ-equivalence classes of random variables thanks to the Cameron-Martin theorem. It is proven in [11] that this definition is equivalent the following condition: Let $(\pi_n, n \geq 1)$ be a sequence of regular, finite dimensional, orthogonal projections of H, increasing to the identity map I_H. Denote also by π_n its continuous extension to W and define $\pi_n^\perp = I_W - \pi_n$. For $x \in W$, let $x_n = \pi_n x$ and $x_n^\perp = \pi_n^\perp x$. Then f is 1-convex if and only if

$$x_n \to \frac{1}{2}|x_n|_H^2 + f(x_n + x_n^\perp)$$

is $\pi_n^\perp\mu$-almost surely convex.

§3. Some Inequalities

Definition 3.1. *Let ξ and η be two probabilities on $(W,\mathcal{B}(W))$. We say that a probability γ on $(W\times W,\mathcal{B}(W\times W))$ is a solution of the Monge-Kantorovitch problem associated to the couple (ξ,η) if the first marginal of γ is ξ, the second one is η and if*

$$\begin{aligned} J(\gamma) &= \int_{W\times W}|x-y|_H^2 d\gamma(x,y) \\ &= \inf\left\{\int_{W\times W}|x-y|_H^2 d\beta(x,y) : \beta\in\Sigma(\xi,\eta)\right\}, \end{aligned}$$

where $\Sigma(\xi,\eta)$ denotes the set of all the probability measures on $W\times W$ whose first and second marginals are respectively ξ and η. We shall denote the Wasserstein distance between ξ and η, which is the positive square-root of this infimum, with $d_H(\xi,\eta)$.

Remark: By the weak compacteness of probability measures on $W\times W$ and the lower semi-continuity of the strictly convex cost function, the infimum in the definition is attained even if the functional J is identically infinity.
The following result, whose proof is outlined below (cf. also[12, 13]) is an extension of an inequality due to Talagrand [24] and it gives a sufficient condition for the Wasserstein distance to be finite:

Theorem 3.1. *Let $L\in\mathbb{L}\log\mathbb{L}(\mu)$ be a positive random variable with $E[L]=1$ and let ν be the measure $d\nu=Ld\mu$. We then have*

$$d_H^2(\nu,\mu)\le 2E[L\log L]\,. \tag{3.3}$$

Proof: Let us remark first that we can take W as the classical Wiener space $W=C_0([0,1])$ and, using the stopping techniques of the martingale theory, we may assume that L is upper and lower bounded almost surely. Then a classical result of the Itô calculus implies that L can be represented as an exponential martingale

$$L_t=\exp\left\{-\int_0^t \dot u_\tau dW_\tau-\frac{1}{2}\int_0^t|\dot u_\tau|^2 d\tau\right\},$$

with $L=L_1$. Let us define $u:W\to H$ as $u(t,x)=\int_0^t\dot u_\tau d\tau$ and $U:W\to W$ as $U(x)=x+u(x)$. The Girsanov theorem implies that

$x \to U(x)$ is a Browian motion under ν, hence $\beta = (U \times I)\nu \in \Sigma(\mu, \nu)$. Let γ be any optimal measure, then

$$\begin{aligned} J(\gamma) &= d_H^2(\nu, \mu) \leq \int_{W\times W} |x-y|_H^2 d\beta(x,y) \\ &= E[|u|_H^2 L] \\ &= 2E[L \log L], \end{aligned}$$

where the last equality follows also from the Girsanov theorem and the Itô stochastic calculus. □

Combining Theorem 3.1 with the triangle inequality for the Wasserstein distance gives:

Corollary 3.1. *Assume that ν_i $(i = 1, 2)$ have Radon-Nikodym densities L_i $(i = 1, 2)$ with respect to the Wiener measure μ which are in $\mathbb{L}\log\mathbb{L}$. Then*

$$d_H(\nu_1, \nu_2) < \infty .$$

Let us give a simple application of the above result in the lines of [17]:

Corollary 3.2. *Assume that $A \in \mathcal{B}(W)$ is any set of positive Wiener measure. Define the H-gauge function of A as*

$$q_A(x) = \inf(|h|_H : h \in (A - x) \cap H) .$$

Then we have

$$E[q_A^2] \leq 2 \log \frac{1}{\mu(A)},$$

in other words

$$\mu(A) \leq \exp\left\{ -\frac{E[q_A^2]}{2} \right\} .$$

Similarly if A and B are H-separated, i.e., if $A_\varepsilon \cap B = \emptyset$, for some $\varepsilon > 0$, where $A_\varepsilon = \{x \in W : q_A(x) \leq \varepsilon\}$, then

$$\mu(A_\varepsilon^c) \leq \frac{1}{\mu(A)} e^{-\varepsilon^2/4}$$

and consequently

$$\mu(A)\,\mu(B) \leq \exp\left(-\frac{\varepsilon^2}{4} \right) .$$

Remark 3.1. *We already know that, from the $0-1$–law, q_A is almost surely finite, besides it satisfies $|q_A(x+h) - q_A(x)| \leq |h|_H$, hence $E[\exp \lambda q_A^2] < \infty$ for any $\lambda < 1/2$ (cf. the Appendix B.8 of [29]). In fact all these assertions can also be proved with the technique used below.*

Proof: Let ν_A be the measure defined by

$$d\nu_A = \frac{1}{\mu(A)} 1_A d\mu .$$

Let γ_A be the solution of the Monge-Kantorovitch problem, it is easy to see that the support of γ_A is included in $W \times A$, hence

$$|x-y|_H \geq \inf\{|x-z|_H : z \in A\} = q_A(x) ,$$

γ_A-almost surely. This implies in particular that q_A is almost surely finite. It follows now from the inequality (3.3)

$$E[q_A^2] \leq -2\log\mu(A) ,$$

hence the proof of the first inequality follows. For the second let $B = A_\varepsilon^c$ and let γ_{AB} be the solution of the Monge-Kantorovitch problem corresponding to ν_A, ν_B. Then we have from the Corollary 3.1,

$$d_H^2(\nu_A, \nu_B) \leq -4\log\mu(A)\mu(B) .$$

Besides the support of the measure γ_{AB} is in $A \times B$, hence γ_{AB}-almost surely $|x-y|_H \geq \varepsilon$ and the proof follows. □

§4. Construction of the transport map

In this section we call optimal every probability measure[4] γ on $W \times W$ such that $J(\gamma) < \infty$ and that $J(\gamma) \leq J(\theta)$ for every other probability θ having the same marginals as those of γ. We recall that a finite dimensional subspace F of W is called regular if the corresponding projection is continuous. Similarly a finite dimensional projection of H is called regular if it has a continuous extension to W.

The proof of the next theorem, for which we refer the reader to [13], can be done by choosing a proper disintegration of any optimal measure in such a way that the elements of this disintegration are the solutions of finite dimensional Monge-Kantorovitch problems. The latter is proven with the help of the section-selection theorem [6].

Theorem 4.1 (General case). *Suppose that ρ and ν are two probability measures on W such that*

$$d_H(\rho, \nu) < \infty .$$

[4] In fact the results of this section are essentially true for the bounded, positive measures.

Let $(\pi_n, n \geq 1)$ be a total increasing sequence of regular projections (of H, converging to the identity map of H). Suppose that, for any $n \geq 1$, the regular conditional probabilities $\rho(\cdot\,|\pi_n^\perp = x^\perp)$ vanish $\pi_n^\perp\rho$-almost surely on the subsets of $(\pi_n^\perp)^{-1}(W)$ with Hausdorff dimension $n-1$. Then there exists a unique solution of the Monge-Kantorovitch problem, denoted by $\gamma \in \Sigma(\rho,\nu)$ and γ is supported by the graph of a Borel map T which is the solution of the Monge problem. $T : W \to W$ is of the form $T = I_W + \xi$, where $\xi \in H$ almost surely. Besides we have

$$\begin{aligned} d_H^2(\rho,\nu) &= \int_{W\times W} |T(x)-x|_H^2 d\gamma(x,y) \\ &= \int_W |T(x)-x|_H^2 d\rho(x)\,, \end{aligned}$$

and for $\pi_n^\perp\rho$-almost almost all $x_n^\perp$, the map $u \to \xi(u + x_n^\perp)$ is cyclically monotone on $(\pi_n^\perp)^{-1}\{x_n^\perp\}$, in the sense that

$$\sum_{i=1}^N \left(u_i + \xi(x_n^\perp + u_i), u_{i+1} - u_i\right)_H \leq 0$$

$\pi_n^\perp\rho$-almost surely, for any cyclic sequence $\{u_1,\ldots,u_N,u_{N+1} = u_1\}$ from $\pi_n(W)$. Finally, if, for any $n \geq 1$, $\pi_n^\perp\nu$-almost surely, $\nu(\cdot\,|\pi_n^\perp = y^\perp)$ also vanishes on the n–1-Hausdorff dimensional subsets of $(\pi_n^\perp)^{-1}(W)$, then T is invertible, i.e, there exists $S : W \to W$ of the form $S = I_W + \eta$ such that $\eta \in H$ satisfies a similar cyclic monotononicity property as ξ and that

$$\begin{aligned} 1 &= \gamma\left\{(x,y) \in W\times W : T\circ S(y) = y\right\} \\ &= \gamma\left\{(x,y) \in W\times W : S\circ T(x) = x\right\}\,. \end{aligned}$$

In particular we have

$$\begin{aligned} d_H^2(\rho,\nu) &= \int_{W\times W} |S(y)-y|_H^2 d\gamma(x,y) \\ &= \int_W |S(y)-y|_H^2 d\nu(y)\,. \end{aligned}$$

Remark 4.1. *In particular, for all the measures ρ which are absolutely continuous with respect to the Wiener measure μ, the second hypothesis is satisfied, i.e., the measure $\rho(\cdot\,|\pi_n^\perp = x_n^\perp)$ vanishes on the sets of Hausdorff dimension $n-1$.*

The case where one of the measures is the Wiener measure and the other is absolutely continuous with respect to μ is the most important one for the applications. Consequently we give the related results separately in the following theorem where the tools of the Malliavin calculus give more information about the maps ξ and η of Theorem 4.1:

Theorem 4.2 (Gaussian case). *Let ν be the measure $d\nu = Ld\mu$, where L is a positive random variable, with $E[L] = 1$. Assume that $d_H(\mu, \nu) < \infty$ (for instance $L \in \mathbb{L}\log\mathbb{L}$). Then there exists a 1-convex function $\phi \in \mathbb{D}_{2,1}$, unique upto a constant, such that the map $T = I_W + \nabla\phi$ is the unique solution of the original problem of Monge. Moreover, its graph supports the unique solution of the Monge-Kantorovitch problem γ. Consequently*

$$(I_W \times T)\mu = \gamma$$

In particular T maps μ to ν and T is almost surely invertible, i.e., there exists some T^{-1} such that $T^{-1}\nu = \mu$ and that

$$\begin{aligned} 1 &= \mu\left\{x : T^{-1} \circ T(x) = x\right\} \\ &= \nu\left\{y \in W : T \circ T^{-1}(y) = y\right\} . \end{aligned}$$

Remark 4.2. *Assume that the operator ∇ is closable with respect to ν, then we have $\eta = \nabla\psi$. In particular, if ν and μ are equivalent, then we have*

$$T^{-1} = I_W + \nabla\psi\,,$$

where is ψ is a 1-convex function.

Remark 4.3. *Let $(e_n, n \in \mathbb{N})$ be a complete, orthonormal in H, denote by V_n the sigma algebra generated by $\{\delta e_1, \ldots, \delta e_n\}$ and let $L_n = E[L|V_n]$. If $\phi_n \in \mathbb{D}_{2,1}$ is the function constructed in Theorem 4.2, corresponding to L_n, then, using the inequality (3.3) we can prove that the sequence $(\phi_n, n \in \mathbb{N})$ converges to ϕ in $\mathbb{D}_{2,1}$.*

Remark 4.4. *Assume that $L \in \mathbb{L}^1_+(\mu)$, with $E[L] = 1$ and let $(D_k, k \in \mathbb{N})$ be a measurable partition of W such that on each D_k, L is bounded. Define $d\nu = L\,d\mu$ and $\nu_k = \nu(\cdot|D_k)$. It follows from Theorem 3.1, that $d_H(\mu, \nu_k) < \infty$. Let then T_k be the map constructed in Theorem 4.2 satisfying $T_k\mu = \nu_k$. Define $n(dk)$ as the probability distribution on $\mathbb{N}$ given by $n(\{k\}) = \nu(D_k)$, $k \in \mathbb{N}$. Then we have*

$$\int_W f(y)d\nu(y) = \int_{W\times\mathbb{N}} f(T_k(x))\mu(dx)n(dk)\,.$$

A similar result is given in [9], the difference with that of above lies in the fact that we have a more precise information about the probability space on which T is defined.

Let us give some applications of the above theorem to the factorization of the absolutely continuous transformations of the Wiener measure.

Assume that $V = I_W + v : W \to W$ be an absolutely continuous transformation and let $L \in \mathbb{L}^1_+(\mu)$ be the Radon-Nikodym derivative of $V\mu$ with respect to μ. Let $T = I_W + \nabla\phi$ be the transport map such that $T\mu = L.\mu$. Then it is easy to see that the map $s = T^{-1} \circ V$ is a rotation, i.e., $s\mu = \mu$ (cf. [29]) and it can be represented as $s = I_W + \alpha$. In particular we have

$$\alpha + \nabla\phi \circ s = v\,. \tag{4.4}$$

Since ϕ is a 1-convex map, we have $h \to \frac{1}{2}|h|_H^2 + \phi(x+h)$ is almost surely convex (cf. [11]). Let $s' = I_W + \alpha'$ be another rotation with $\alpha' : W \to H$. By the 1-convexity of ϕ, we have

$$\frac{1}{2}|\alpha'|_H^2 + \phi \circ s' \geq \frac{1}{2}|\alpha|_H^2 + \phi \circ s + \Big(\alpha + \nabla\phi \circ s, \alpha' - \alpha\Big)_H\,,$$

μ-almost surely. Taking the expectation of both sides, using the fact that s and s' preserve the Wiener measure μ and the identity (4.4), we obtain

$$E\left[\frac{1}{2}\,|\alpha|_H^2 - (v,\alpha)_H\right] \leq E\left[\frac{1}{2}|\alpha'|_H^2 - (v,\alpha')_H\right]\,.$$

Hence we have proven the existence part of the following

Proposition 4.1. *Let $\mathcal{R}_2$ denote the subset of $L^2(\mu, H)$ whose elements are defined by the property that $x \to x + \eta(x)$ is a rotation, i.e., it preserves the Wiener measure. Then α is the unique element of $\mathcal{R}_2$ which minimizes the functional*

$$\eta \to M_v(\eta) = E\left[\frac{1}{2}|\eta|_H^2 - (v,\eta)_H\right]\,.$$

Proof: The only claim to prove is the uniqueness and it follows easily from Theorem 4.2. □

The following theorem, whose proof is rather easy, gives a better understanding of the structure of absolutely continuous transformations of the Wiener measure:

Theorem 4.3. *Assume that $U : W \to W$ be a measurable map and $L \in \mathbb{L}\log\mathbb{L}$ a positive random variable with $E[L] = 1$. Assume that the measure $\nu = L \cdot \mu$ is a Girsanov measure for U, i.e., that one has*

$$E[f \circ U\, L] = E[f]\,,$$

for any $f \in C_b(W)$. *Then there exists a unique map* $T = I_W + \nabla\phi$ *with* $\phi \in \mathbb{D}_{2,1}$ *is* 1*-convex, and a measure preserving transformation* $R : W \to W$ *such that* $U \circ T = R$ μ*-almost surely and* $U = R \circ T^{-1}$ ν*-almost surely.*

Another version of Theorem 4.3 can be announced as follows:

Theorem 4.4. *Assume that* $Z : W \to W$ *is a measurable map such that* $Z\mu \ll \mu$, *with* $d_H(Z\mu, \mu) < \infty$. *Then* Z *can be decomposed as*

$$Z = T \circ s\,,$$

where T *is the unique transport map of the Monge-Kantorovitch problem for* $\Sigma(\mu, Z\mu)$ *and* s *is a rotation.*

Although the following result is a translation of the results of this section, it is interesting from the point of view of stochastic differential equations:

Theorem 4.5. *Let* (W, μ, H) *be the standard Wiener space on* $\mathbb{R}^d$, *i.e.,* $W = C(\mathbb{R}_+, \mathbb{R}^d)$. *Assume that there exists a probability* $P \ll \mu$ *which is the weak solution of the stochastic differential equation*

$$dy_t = dW_t + b(t, y)dt\,,$$

such that $d_H(P, \mu) < \infty$. *Then there exists a process* $(T_t, t \in \mathbb{R}_+)$ *which is a pathwise solution of some (anticipative) stochastic differential equation whose law is equal to* P.

Proof: Let T be the transport map constructed in Theorem 4.2 corresponding to $dP/d\mu$. Then it has an inverse T^{-1} such that $\mu\{T^{-1} \circ T(x) = x\} = 1$. Let ϕ be the 1-convex function such that $T = I_W + \nabla\phi$ and denote by $(D_s\phi, s \in \mathbb{R}_+)$ the representation of $\nabla\phi$ in $L^2(\mathbb{R}_+, ds)$. Define $T_t(x)$ as the trajectory $T(x)$ evaluated at $t \in \mathbb{R}_+$. Then it is easy to see that $(T_t, t \in \mathbb{R}_+)$ satisfies the stochastic differential equation

$$T_t(x) = W_t(x) + \int_0^t l(s, T(x))ds\ ,\ t \in \mathbb{R}_+\,,$$

where $W_t(x) = x(t)$ and $l(s, x) = D_s\phi \circ T^{-1}(x)$ if $x \in T(W)$ and zero otherwise. □

§5. The Monge-Ampère equation

Assume that $W = \mathbb{R}^n$ and take a density $L \in \mathbb{L}\log\mathbb{L}$. Let $\phi \in \mathbb{D}_{2,1}$ be the 1-convex function such that $T = I + \nabla\phi$ maps μ to $L \cdot \mu$. Let $S = I + \nabla\psi$ be its inverse with $\psi \in \mathbb{D}_{2,1}$. Let now $\nabla_a^2\phi$ be the second Alexandrov derivative of ϕ, i.e., the Radon-Nikodym derivative of the absolutely continuous part of the vector measure $\nabla^2\phi$ with respect to the Gaussian measure μ on $\mathbb{R}^n$. Since ϕ is 1-convex, it follows that $\nabla^2\phi \geq -I_{\mathbb{R}^n}$ in the sense of the distributions, consequently $\nabla_a^2\phi \geq -I_{\mathbb{R}^n}$ μ-almost surely. Define also the Alexandrov version $\mathcal{L}_a\phi$ of $\mathcal{L}\phi$ as the Radon-Nikodym derivative of the absolutely continuous part of the distribution $\mathcal{L}\phi$. Since we are in finite dimensional situation, we have the explicit expression for $\mathcal{L}_a\phi$ as

$$\mathcal{L}_a\phi(x) = (\nabla\phi(x), x)_{\mathbb{R}^n} - \operatorname{trace}\left(\nabla_a^2\phi\right) .$$

Let Λ be the Gaussian Jacobian

$$\Lambda = \det{}_2\left(I_{\mathbb{R}^n} + \nabla_a^2\phi\right)\exp\left\{-\mathcal{L}_a\phi - \frac{1}{2}|\nabla\phi|^2_{\mathbb{R}^n}\right\} .$$

Remark 5.1. In this expression as well as in the sequel, the notation $\det_2(I_H + A)$ denotes the modified Carleman-Fredholm determinant of the operator $I_H + A$ on a Hilbert space H. If A is an operator of finite rank, then it is defined as

$$\det{}_2\left(I_H + A\right) = \prod_{i=1}^{n}(1 + l_i)e^{-l_i} ,$$

where $(l_i,\, i \leq n)$ denotes the eigenvalues of A counted with respect to their multiplicity. In fact this determinant has an analytic extension to the space of Hilbert-Schmidt operators on a separable Hilbert space, cf. [7] and Appendix A.2 of [29]. As explained in [29], the modified determinant exists for the Hilbert-Schmidt operators while the ordinary determinant does not, since the latter requires the existence of the trace of A. Hence the modified Carleman-Fredholm determinant is particularly useful when one studies the absolute continuity properties of the image of a Gaussian measure under non-linear transformations in the setting of infinite dimensional Banach spaces (cf., [29] for further information).

It follows from the change of variables formula given in Corollary 4.3 of [19], that, for any $f \in C_b(\mathbb{R}^n)$,

$$E[f \circ T\, \Lambda] = E\left[f\, 1_{\partial\Phi(M)}\right] ,$$

where M is the set of non-degeneracy of $I_{\mathbb{R}^n} + \nabla_a^2\phi$,

$$\Phi(x) = \frac{1}{2}|x|^2 + \phi(x)$$

and $\partial\Phi$ denotes the subdifferential of the convex function Φ. Let us note that, in case $L > 0$ almost surely, T has a global inverse S, i.e., $S \circ T = T \circ S = I_{\mathbb{R}^n}$ μ-almost surely and $\mu(\partial\Phi(M)) = \mu(S^{-1}(M))$. Assume now that $\Lambda > 0$ almost surely, i.e., that $\mu(M) = 1$. Then, for any $f \in C_b(\mathbb{R}^n)$, we have

$$\begin{aligned} E[f \circ T] &= E\left[f \circ T \frac{\Lambda}{\Lambda \circ T^{-1} \circ T}\right] \\ &= E\left[f \frac{1}{\Lambda \circ T^{-1}} 1_{\partial\Phi(M)}\right] \\ &= E[f\,L]\,, \end{aligned}$$

where T^{-1} denotes the left inverse of T whose existence is guaranteed by Theorem 4.2. Since $T(x) \in \partial\Phi(M)$ almost surely, it follows from the above calculations

$$\frac{1}{\Lambda} = L \circ T\,,$$

almost surely. Take now any $t \in [0, 1)$, the map $x \to \frac{1}{2}|x|_H^2 + t\phi(x) = \Phi_t(x)$ is strictly convex and a simple calculation implies that the mapping $T_t = I + t\nabla\phi$ is $(1-t)$-monotone (cf. [29], Chapter 6), consequently it has a left inverse denoted by S_t. Let us denote by Ψ_t the Legendre transformation of Φ_t:

$$\Psi_t(y) = \sup_{x \in \mathbb{R}^n} \{(x, y) - \Phi_t(x)\}\,.$$

A simple calculation shows that

$$\begin{aligned} \Psi_t(y) &= \sup_x \left[(1-t)\left\{(x, y) - \frac{|x|^2}{2}\right\} + t\left\{(x, y) - \frac{|x|^2}{2} - \phi(x)\right\}\right] \\ &\le (1-t)\frac{|y|^2}{2} + t\Psi_1(y)\,. \end{aligned}$$

Since Ψ_1 is the Legendre transformation of $\Phi_1(x) = |x|^2/2 + \phi(x)$ and since $L \in \mathbb{L}\log\mathbb{L}$, it is finite on a convex set of full measure, hence it is finite everywhere. Consequently $\Psi_t(y) < \infty$ for any $y \in \mathbb{R}^n$. Since a finite, convex function is almost everywhere differentiable, $\nabla\Psi_t$ exists almost everywhere on and it is equal almost everywhere on $T_t(M_t)$ to the left inverse T_t^{-1}, where M_t is the set of non-degeneracy of $I_{\mathbb{R}^n} + t\nabla_a^2\phi$.

Note that $\mu(M_t) = 1$. The strict convexity implies that T_t^{-1} is Lipschitz with a Lipschitz constant $\frac{1}{1-t}$. Let now Λ_t be the Gaussian Jacobian

$$\Lambda_t = \det_2\left(I_{\mathbb{R}^n} + t\nabla_a^2\phi\right)\exp\left\{-t\mathcal{L}_a\phi - \frac{t^2}{2}|\nabla\phi|^2_{\mathbb{R}^n}\right\}.$$

Since the domain of ϕ is the whole space $\mathbb{R}^n$, $\Lambda_t > 0$ almost surely, hence, as we have explained above, it follows from the change of variables formula of [19] that $T_t\mu$ is absolutely continuous with respect to μ and that

$$\frac{1}{\Lambda_t} = L_t \circ T_t\,,$$

μ-almost surely.
Let us come back to the infinite dimensional case: we first give an inequality which may be useful.

Theorem 5.1. *Assume that (W,μ,H) is an abstract Wiener space, assume that $K, L \in \mathbb{L}^1_+(\mu)$ with $K > 0$ almost surely and denote by $T: W \to W$ the transfer map $T = I_W + \nabla\phi$, which maps the measure $Kd\mu$ to the measure $Ld\mu$. Then the following inequality holds:*

$$\frac{1}{2}E[|\nabla\phi|^2_H] \leq E[-\log K + \log L\circ T]\,. \tag{5.5}$$

Proof: Let us define k as $k = K\circ T^{-1}$, then for any $f \in C_b(W)$, we have

$$\begin{aligned}\int_W f(y)L(y)d\mu(y) &= \int_W f\circ T(x)K(x)d\mu(x)\\ &= \int_W f\circ T(x)k\circ T(x)d\mu(x)\,,\end{aligned}$$

hence

$$T\mu = \frac{L}{k}\,.\mu\,.$$

It then follows from the inequality (3.3) that

$$\begin{aligned}\frac{1}{2}E\left[|\nabla\phi|^2_H\right] &\leq E\left[\frac{L}{k}\log\frac{L}{k}\right]\\ &= E\left[\log\frac{L\circ T}{k\circ T}\right]\\ &= E[-\log K + \log L\circ T]\,.\end{aligned}$$

□

In case K and L are given as e^{-U} and e^{-V} we have another inequality containing the Fisher information (cf. [21] for the finite dimensional case):

Theorem 5.2. *Assume that $U, V \in \mathbb{D}_{2,1}$ are such that $E[e^{-U}] = E[e^{-V}] = 1$ and that $E[e^{-U}|\nabla U|_H^2] + E[e^{-V}|\nabla V|_H^2] < \infty$. Define $d\rho = e^{-U}d\mu$ and $d\nu = e^{-V}d\mu$. Then we have*

$$d_H(\rho,\nu) \le E_\rho[|\nabla U|_H^2]^{1/2} + E_\nu[|\nabla V|_H^2]^{1/2}. \tag{5.6}$$

The equation (5.6) can be refined in the following way: let

$$\kappa_\pm(U,V) = E_\rho[|\nabla U|_H^2]^{1/2} \pm \left\{E_\rho[|\nabla U|_H^2] + 2(E_\rho(U) - E_\nu(V))\right\}^{1/2}$$

and let

$$\kappa_\pm(V,U) = E_\nu[|\nabla V|_H^2]^{1/2} \pm \left\{E_\nu[|\nabla V|_H^2] + 2(E_\nu(V) - E_\rho(U))\right\}^{1/2}.$$

We then have

$$\max\{\kappa_-(U,V), \kappa_-(V,U)\} \le d_H(\rho,\nu) \le \min\{\kappa_+(U,V), \kappa_+(V,U)\}. \tag{5.7}$$

Proof: By taking the conditional expectations with respect to the sigma algebras generated by a finite number of elements of the first Wiener chaos and using the Jensen inequality, we can reduce the problem to the finite dimensional case. Let now $T = I + \nabla\phi$ be the transport map sending ρ to ν, let also $S = I + \nabla\psi$ be its inverse. As we have seen before, we can write $\Lambda = \exp(-U + V \circ T)$ as

$$\Lambda = \det\nolimits_2(I + \nabla_a^2\phi)\exp\left\{-\mathcal{L}_a\phi - \frac{1}{2}|\nabla\phi|_H^2\right\}.$$

Solving $|\nabla\phi|_H^2$ from this expression and using the fact that $-\mathcal{L}_a\phi \le -\mathcal{L}\phi$ in the sense of the distributions, we obtain

$$\frac{1}{2}|\nabla\phi|_H^2 \le U - V \circ T - \mathcal{L}\phi.$$

Taking the expectation of both sides of this inequality with respect to ρ, we obtain

$$\frac{1}{2}d_H^2(\rho,\nu) \le E_\rho[U] - E_\nu[V] + E_\rho[(\nabla\phi, \nabla U)]. \tag{5.8}$$

Interchanging T and S and ρ and ν in the inequality (5.8), we obtain also

$$\frac{1}{2}d_H^2(\rho,\nu) \le E_\nu[V] - E_\rho[U] + E_\nu[(\nabla\psi, \nabla V)]. \tag{5.9}$$

Adding (5.8) to (5.9) and using the Cauchy-Schwarz inequality completes the proof of the first part. For the second, using the inequalities (5.8) and (5.9), with the help of the Cauchy-Schwarz inequality and the general expression of solutions of the second order polynomial equation we get the claim at once. □

Suppose that $\phi \in \mathbb{D}_{2,1}$ is a 1-convex Wiener functional. Let V_n be, as usual, the sigma algebra generated by $\{\delta e_1, \ldots, \delta e_n\}$, where $(e_n,\, n \geq 1)$ is an orthonormal basis of the Cameron-Martin space H. Then $\phi_n = E[\phi|V_n]$ is again 1-convex (cf. [11]), hence $\mathcal{L}\phi_n$ is a measure as it can be easily verified. However the sequence $(\mathcal{L}\phi_n,\, n \geq 1)$ converges to $\mathcal{L}\phi$ only in $\mathbb{D}'$. Consequently, there is no reason for the limit $\mathcal{L}\phi$ to be a measure. In case this happens, we shall denote the Radon-Nikodym density with respect to μ, of the absolutely continuous part of this measure by $\mathcal{L}_a\phi$.

Lemma 5.1. *Let $\phi \in \mathbb{D}_{2,1}$ be 1-convex and let V_n be defined as above and define $F_n = E[\phi|V_n]$. Then the sequence $(\mathcal{L}_a F_n, n \geq 1)$ is a submartingale, where $\mathcal{L}_a F_n$ denotes the μ-absolutely continuous part of the measure $\mathcal{L}F_n$.*

Proof: Note that, due to the 1-convexity, we have $\mathcal{L}_a F_n \geq \mathcal{L}F_n$ for any $n \in \mathbb{N}$. Let $X_n = \mathcal{L}_a F_n$ and $f \in \mathbb{D}$ be a positive, V_n-measurable test function. Since $\mathcal{L}E[\phi|V_n] = E[\mathcal{L}\phi|V_n]$, we have

$$
\begin{aligned}
E[X_{n+1}\, f] &\geq \langle \mathcal{L}F_{n+1}, f\rangle \\
&= \langle \mathcal{L}F_n, f\rangle\,,
\end{aligned}
$$

where $\langle\cdot,\cdot\rangle$ denotes the duality bracket for the dual pair $(\mathbb{D}', \mathbb{D})$. Consequently

$$
E[f\, E[X_{n+1}|V_n]] \geq \langle \mathcal{L}F_n, f\rangle\,,
$$

for any positive, V_n-measurable test function f, it follows that the absolutely continuous part of $\mathcal{L}F_n$ is also dominated by the same conditional expectation and this proves the submartingale property. □

Lemma 5.2. *Assume that $L \in \mathbb{L}\log\mathbb{L}$ is a positive random variable whose expectation is one. Assume further that it is lower bounded by a constant $a > 0$. Let $T = I_W + \nabla\phi$ be the transport map such that $T\mu = L\,.\mu$ and let $T^{-1} = I_W + \nabla\psi$. Then $\mathcal{L}\psi$ is a Radon measure on $(W, \mathcal{B}(W))$. If L is upper bounded by $b > 0$, then $\mathcal{L}\phi$ is also a Radon measure on $(W, \mathcal{B}(W))$.*

Proof: Let $L_n = E[L|V_n]$, then $L_n \geq a$ almost surely. Let $T_n = I_W + \nabla\phi_n$ be the transport map which satisfies $T_n\mu = L_n\,.\mu$ and let

$T_n^{-1} = I_W + \nabla\psi_n$ be its inverse. We have

$$L_n = \det_2\left(I_H + \nabla_a^2\psi_n\right)\exp\left[-\mathcal{L}_a\psi_n - \frac{1}{2}|\nabla\psi_n|_H^2\right].$$

By the hypothesis $-\log L_n \leq -\log a$. Since ψ_n is 1-convex, it follows from the finite dimensional results that $\det_2\left(I_H + \nabla_a^2\psi_n\right) \in [0,1]$ almost surely. Therefore we have

$$\mathcal{L}_a\psi_n \leq -\log a\,,$$

besides $\mathcal{L}\psi_n \leq \mathcal{L}_a\psi_n$ as distributions, consequently

$$\mathcal{L}\psi_n \leq -\log a$$

as distributions, for any $n \geq 1$. Since $\lim_n \mathcal{L}\psi_n = \mathcal{L}\psi$ in $\mathbb{D}'$, we obtain $\mathcal{L}\psi \leq -\log a$, hence $-\log a - \mathcal{L}\psi \geq 0$ as a distribution, hence $\mathcal{L}\psi$ is a Radon measure on W, c.f., [10], [28]. This proves the first claim. Note that whenever L is upperbounded, $\Lambda = 1/L\circ T$ is lowerbounded, hence the proof of the second claim is similar to that of the first one. □

Theorem 5.3. *Assume that L is a strictly positive bounded random variable with $E[L]=1$. Let $\phi \in \mathbb{D}_{2,1}$ be the 1-convex Wiener functional such that*

$$T = I_W + \nabla\phi$$

is the transport map realizing the measure $L\,.\mu$ and let $S = I_W + \nabla\psi$ be its inverse. Define $F_n = E[\phi|V_n]$, then the submartingale $(\mathcal{L}_aF_n, n\geq 1)$ converges almost surely to $\mathcal{L}_a\phi$. Let $\lambda(\phi)$ be the random variable defined as

$$\begin{aligned}\lambda(\phi) &= \liminf_{n\to\infty} \Lambda_n\\ &= \left(\liminf_n \det_2\left(I_H + \nabla_a^2F_n\right)\right)\exp\left\{-\mathcal{L}_a\phi - \frac{1}{2}\,|\nabla\phi|_H^2\right\}\end{aligned}$$

where

$$\Lambda_n = \det_2\left(I_H + \nabla_a^2F_n\right)\exp\left\{-\mathcal{L}_aF_n - \frac{1}{2}\,|\nabla F_n|_H^2\right\}.$$

Then it holds true that

$$E[f\circ T\,\lambda(\phi)] \leq E[f] \tag{5.10}$$

for any $f \in C_b^+(W)$, *in particular* $\lambda(\phi) \leq \frac{1}{L \circ T}$ *almost surely. If* $E[\lambda(\phi)] = 1$, *then the inequality in (5.10) becomes an equality and we also have*

$$\lambda(\phi) = \frac{1}{L \circ T}.$$

Proof: Let us remark that, due to the 1-convexity, $0 \leq \det_2(I_H + \nabla_a^2 F_n) \leq 1$, hence the lim inf exists. Now, Lemma 5.2 implies that $\mathcal{L}\phi$ is a Radon measure. Let $F_n = E[\phi|V_n]$, then we know from Lemma 5.1 that $(\mathcal{L}_a F_n, n \geq 1)$ is a submartingale. Let $\mathcal{L}^+\phi$ denote the positive part of the measure $\mathcal{L}\phi$. Since $\mathcal{L}^+\phi \geq \mathcal{L}\phi$, we have also $E[\mathcal{L}^+\phi|V_n] \geq E[\mathcal{L}\phi|V_n] = \mathcal{L}F_n$. This implies that $E[\mathcal{L}^+\phi|V_n] \geq \mathcal{L}_a^+ F_n$. Hence we find that

$$\sup_n E[\mathcal{L}_a^+ F_n] < \infty$$

and this condition implies that the submartingale $(\mathcal{L}_a F_n, n \geq 1)$ converges almost surely. We shall now identify the limit of this submartingale. Let $\mathcal{L}_s G$ be the singular part of the measure $\mathcal{L}G$ for a Wiener function G such that $\mathcal{L}G$ is a measure. We have

$$\begin{aligned} E[\mathcal{L}\phi|V_n] &= E[\mathcal{L}_a\phi|V_n] + E[\mathcal{L}_s\phi|V_n] \\ &= \mathcal{L}_a F_n + \mathcal{L}_s F_n, \end{aligned}$$

hence

$$\mathcal{L}_a F_n = E[\mathcal{L}_a\phi|V_n] + E[\mathcal{L}_s\phi|V_n]_a$$

almost surely, where $E[\mathcal{L}_s\phi|V_n]_a$ denotes the absolutely continuous part of the measure $E[\mathcal{L}_s\phi|V_n]$. Note that, from the Theorem of Jessen (cf., for example Theorem 1.2.1 of [29]), $\lim_n E[\mathcal{L}_s^+\phi|V_n]_a = 0$ and $\lim_n E[\mathcal{L}_s^-\phi|V_n]_a = 0$ almost surely, hence we have

$$\lim_n \mathcal{L}_a F_n = \mathcal{L}_a\phi,$$

μ-almost surely. To complete the proof, an application of the Fatou lemma implies that

$$\begin{aligned} E[f \circ T\, \lambda(\phi)] &\leq E[f] \\ &= E\left[f \circ T \frac{1}{L \circ T}\right], \end{aligned}$$

for any $f \in C_b^+(W)$. Since T is invertible, it follows that

$$\lambda(\phi) \leq \frac{1}{L \circ T}$$

almost surely. Therefore, in case $E[\lambda(\phi)] = 1$, we have

$$\lambda(\phi) = \frac{1}{L \circ T},$$

and this completes the proof. □

Corollary 5.1. *Assume that K, L are two positive random variables with values in a bounded interval $[a, b] \subset (0, \infty)$ such that $E[K] = E[L] = 1$. Let $T = I_W + \nabla\phi$, $\phi \in \mathbb{D}_{2,1}$, be the transport map pushing $Kd\mu$ to $Ld\mu$, i.e, $T(Kd\mu) = Ld\mu$. We then have*

$$L \circ T \, \lambda(\phi) \leq K,$$

μ-almost surely. In particular, if $E[\lambda(\phi)] = 1$, then T is the solution of the Monge-Ampère equation.

Proof: Since $a > 0$,

$$\frac{dT\mu}{d\mu} = \frac{L}{K \circ T} \leq \frac{b}{a}.$$

Hence, Theorem 5.10 implies that

$$\begin{aligned} E[f \circ T \, L \circ T \, \lambda(\phi)] &\leq E[f \, L] \\ &= E[f \circ T \, K], \end{aligned}$$

consequently

$$L \circ T \, \lambda(\phi) \leq K,$$

the rest of the claim is now obvious. □

For later use we give also the folowing result:

Theorem 5.4. *Assume that L is a positive random variable of class $\mathbb{L}\log\mathbb{L}$ such that $E[L] = 1$. Let $\phi \in \mathbb{D}_{2,1}$ be the 1-convex function corresponding to the transport map $T = I_W + \nabla\phi$. Define $T_t = I_W + t\nabla\phi$, where $t \in [0, 1]$. Then, for any $t \in [0, 1]$, $T_t\mu$ is absolutely continuous with respect to the Wiener measure μ.*

Proof: Let ϕ_n be defined as the transport map corresponding to $L_n = E[P_{1/n}L_n|V_n]$ and define T_n as $I_W + \nabla\phi_n$. For $t \in [0, 1)$, let $T_{n,t} = I_W + t\nabla\phi_n$. It follows from the finite dimensional results which are summarized in the beginning of this section, that $T_{n,t}\mu$ is absolutely continuous with respect to μ. Let $L_{n,t}$ be the corresponding Radon-Nikodym density and define $\Lambda_{n,t}$ as

$$\Lambda_{n,t} = \det_2\left(I_H + t\nabla_a^2\phi_n\right)\exp\left\{-t\mathcal{L}_a\phi_n - \frac{t^2}{2}|\nabla\phi_n|_H^2\right\}.$$

Besides, for any $t \in [0,1)$,

$$\left((I_H + t\nabla_a^2\phi_n)h, h\right)_H > 0\,, \tag{5.11}$$

μ-almost surely for any $0 \neq h \in H$. Since ϕ_n is of finite rank, (5.11) implies that $\Lambda_{n,t} > 0$ μ-almost surely and we have shown at the beginning of this section

$$\Lambda_{n,t} = \frac{1}{L_{n,t} \circ T_{n,t}}$$

μ-almost surely. An easy calculation shows that $t \to \Lambda_{n,t}$ is logarithmically concave. Consequently

$$\begin{aligned} E\left[L_{t,n}\log L_{t,n}\right] &= E\left[\log L_{n,t} \circ T_{n,t}\right] \\ &= -E\left[\log \Lambda_{t,n}\right] \\ &\leq E\left[L_n \log L_n\right] \\ &\leq E[L\log L]\,, \end{aligned}$$

by the Jensen inequality. Therefore

$$\sup_n E[L_{n,t}\log L_{n,t}] < \infty$$

and this implies that the sequence $(L_{n,t}, n \geq 1)$ is uniformly integrable for any $t \in [0,1]$. Consequently it has a subsequence which converges weakly in $L^1(\mu)$ to some L_t. Since, from Theorem 4.2 and from Remark 4, $\lim_n \phi_n = \phi$ in $\mathbb{D}_{2,1}$, where ϕ is the transport map associated to L, for any $f \in C_b(W)$, we have

$$\begin{aligned} E[f \circ T_t] &= \lim_k E\left[f \circ T_{n_k,t}\right] \\ &= \lim_k E\left[f\, L_{n_k,t}\right] \\ &= E[f\, L_t]\,, \end{aligned}$$

and this completes the proof. □

Let us give an application of this result:

Proposition 5.1. *Assume that the hypothesis of Theorem 5.4 are valid. Let $\nu_t = T_t\mu$ with $\nu_1 = \nu$, $t \in [0,1]$. Then*

$$d_H(\nu_s, \nu_t) = |t-s| d_H(\mu,\nu) \text{ for } s,t \in [0,1].$$

In particular $T_t \circ S_s$ is the optimal transport of ν_s to ν_t and the Wiener functional $\nabla\psi_s + t\nabla\phi \circ S_s$ is an exact form.

Proof: It suffices to prove the claim for the case $s = 1$. Let $\nu_t = T_t \circ S\nu$, then

$$\begin{aligned} d_H^2(\nu_t,\nu) &\le \int |T_t \circ S(y) - y|_H^2 d\nu(y) \\ &= \int |T_t(x) - T(x)|_H^2 d\mu(x) \\ &= (1-t)^2 E[|\nabla\phi|_H^2], \end{aligned}$$

which means that

$$d_H(\nu_t,\nu) \le (1-t) d_H(\mu,\nu).$$

Moreover, from the triangle inequality

$$\begin{aligned} d_H(\mu,\nu) &\le d_H(\nu,\nu_t) + d_H(\nu_t,\mu) \\ &= d_H(\nu,\nu_t) + t d_H(\mu,\nu), \end{aligned}$$

we obtain that $(1-t)d_H(\mu,\nu) \le d_H(\nu,\nu_t)$. The rest is obvious from Theorem 4.2. □

§6. Solution of the Monge-Ampère equation with the Itô calculus

To have a better understanding of what will follow, let us give an interpretation of the Monge-Ampère equation. Assume that we are given two probability densities K and L and we look for a map $T : W \to W$ such that

$$L \circ T \, J(T) = K$$

almost surely, where $J(T)$ is a kind of Jacobian to be written in terms of T. In Corollary 5.1, we have shown the existence of some $\lambda(\phi)$ which gives an inequality instead of the equality. The reason for that originates from the singularity of the second derivative of the potential function ϕ. Although in the finite dimensional case there are some regularity results about the transport map (cf., [5]), in the infinite dimensional case such techniques do not work. All these difficulties can be circumvented using the miraculous renormalization of the Itô calculus. In fact assume that K and L satisfy the hypothesis of the corollary. First let us indicate that we can assume $W = C_0([0,1],\mathbb{R})$ (cf., [29], Chapter II, to see how one can pass from an abstract Wiener space to the standard one) and in this case the Cameron-Martin space H becomes $H^1([0,1])$, which is the space of absolutely continuous functions on $[0,1]$, with a square integrable Sobolev derivative. Let now

$$\Lambda = \frac{K}{L \circ T},$$

where T is as constructed above. Then $\Lambda.\mu$ is a Girsanov measure for the map T. This means that the law of the stochastic process $(t,x) \to T_t(x)$ under $\Lambda.\mu$ is equal to the Wiener measure, where $T_t(x)$ is defined as the evaluation of the trajectory $T(x)$ at $t \in [0,1]$. In other words the process $(t,x) \to T_t(x)$ is a Brownian motion under the probability $\Lambda.\mu$. Let $(\mathcal{F}_t^T, t \in [0,1])$ be its filtration, the invertibility of T implies that

$$\bigvee_{t \in [0,1]} \mathcal{F}_t^T = \mathcal{B}(W).$$

Λ is upper and lower bounded μ-almost surely, hence also $\Lambda.\mu$-almost surely. The Itô representation theorem implies that it can be represented as

$$\Lambda = E[\Lambda^2] \exp\left\{ -\int_0^1 \dot{\alpha}_s dT_s - \frac{1}{2}\int_0^1 |\dot{\alpha}_s|^2 ds \right\},$$

where $\alpha(\cdot) = \int_0^{\cdot} \dot{\alpha}_s ds$ is an H-valued random variable. In fact α can be calculated explicitly using the Itô-Clark representation theorem (cf., [28]), and it is given as

$$\dot{\alpha}_t = \frac{E_\Lambda[D_t\Lambda | \mathcal{F}_t^T]}{E_\Lambda[\Lambda | \mathcal{F}_t^T]} \tag{6.12}$$

$dt \times \Lambda d\mu$-almost surely, where E_Λ denotes the expectation operator with respect to $\Lambda.\mu$ and $D_t\Lambda$ is the Lebesgue density of the absolutely continuous map $t \to \nabla\Lambda(t,x)$. From the relation (6.12), it follows that α is a function of T, hence we have obtained the strong solution of the Monge-Ampère equation. Let us announce all this as

Theorem 6.1. *Assume that K and L are upper and lower bounded densities, let T be the transport map constructed in Theorem 4.1. Then T is also the strong solution of the Monge-Ampère equation in the Itô* *sense, namely*

$$E[\Lambda^2]\,L\circ T\exp\left\{-\int_0^1 \dot\alpha_s dT_s - \frac{1}{2}\int_0^1 |\dot\alpha_s|^2 ds\right\} = K\,,$$

μ-almost surely, where α is given with (6.12).

References

[1] P. Appell: "Mémoire sur déblais et les remblais des systèmes continus ou discontinus". Mémoires présentées par divers savants à l'Académie des Sciences de l'Institut de France. Paris, I. N. **29**, 1-208, 1887.

[2] P. Appell: "Le problème géométrique des déblais et des remblais". Mémorial des Sciences Mathématiques, fasc. **XXVII**, Paris, 1928.

[3] P. J. Bickel and D. A. Freedman: "Some asymptotic theory for the bootstrap". The Annals of Statistics, Vol. **9**, No. 6, 1196-1217, 1981.

[4] Y. Brenier: "Polar factorization and monotone rearrangement of vector valued functions". Comm. pure Appl. Math, **44**, 375-417, 1991.

[5] L. A. Caffarelli: "The regularity of mappings with a convex potential". Jour. Amer. Math. Soc., **5**, 99-104, 1992.

[6] C. Dellacherie and P. A. Meyer: *Probabilités et Potentiel, Ch. I à IV.* Paris, Hermann, 1975.

[7] N. Dunford and J. T. Schwartz: *Linear Operators* **2**, Interscience 1963.

[8] X. Fernique: "Extension du théorème de Cameron-Martin aux translations aléatoires", Comptes Rendus Mathématiques, Vol. **335**, Issue 1, 65-68, 2002.

[9] X. Fernique: "Comparaison aux mesures gaussiennes, espaces autoreproduisants. Une application des propriétés isopérimétriques". Preprint.

[10] D. Feyel and A. de La Pradelle: "Capacités gaussiennes". Annales de l'Institut Fourier, t.41, f.1, 49-76, 1991.

[11] D. Feyel and A. S. Üstünel: "The notion of convexity and concavity on Wiener space". Journal of Functional Analysis, **176**, 400-428, 2000.

[12] D. Feyel and A. S. Üstünel: "Transport of measures on Wiener space and the Girsanov theorem". Comptes Rendus Mathématiques, Vol. **334**, Issue 1, 1025-1028, 2002.

[13] D. Feyel and A. S. Üstünel: "Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space". Preprint, 2003.

[14] W. Gangbo and R. J. McCann: "The geometry of optimal transportation". Acta Mathematica, **177**, 113-161, 1996.

[15] K. Itô and M. Nisio: "On the convergence of sums of independent Banach space valued random variables". Osaka Journ. Math. **5**, 35-48, 1968.

[16] L. V. Kantorovitch: "On the transfer of masses". Dokl. Acad. Nauk. SSSR **37**, 227-229, 1942.

[17] K. Marton: "Bounding $\bar{d}$-distance by informational divergence: a method to prove measure concentration". Annals of Probability, **24**, no.2, 857-866, 1996.

[18] R. J. McCann: "Existence and uniqueness of monotone measure-preserving maps". Duke Math. Jour., **80**, 309-323, 1995.

[19] R. J. McCann: "A convexity principle for interacting gases". Advances in Mathematics, **128**, 153-179, 1997.

[20] G. Monge: "Mémoire sur la théorie des déblais et des remblais". Histoire de l'Académie Royale des Sciences, Paris, 1781.

[21] F. Otto and C. Villani: "Generalization of an inequality by Talagrand and links with logarithmic Sobolev inequality". Journal of Functional Analysis, **173**, 361-400, 2000.

[22] T. Rockafellar: *Convex Analysis.* Princeton University Press, Princeton, 1972.

[23] V. N. Sudakov: "Geometric problems in the theory of infinite dimensional probability distributions". Proc. Steklov Inst. Math., **141**, 1-178, 1979.

[24] M. Talagrand: "Transportation cost for Gaussian and other product measures". Geom. Funct. Anal., **6**, 587-600, 1996.

[25] H. Tanaka "An inequality for a functional of probability distributions and its applications to Kac's one-dimensional model of a Maxwellian gas". Z. für Wahrsch. verw. Geb., **27**, 47-52, 1973.

[26] H. Tanaka "Certain functional of probability measures on Hilbert spaces". Hiroshima Math. Jour., **6**, 421-428, 1976.

[27] A. S. Üstünel: "Representation of distributions on Wiener space and Stochastic Calculus of Variations". Journal of Functional Analysis, **70**, 126-139, 1987.

[28] A. S. Üstünel: *Introduction to Analysis on Wiener Space.* Lecture Notes in Math. Vol. **1610**. Springer, 1995.

[29] A. S. Üstünel and M. Zakai: *Transformation of Measure on Wiener Space.* Springer Monographs in Mathematics. Springer Verlag, 1999.

Denis Feyel
Université d'Evry-Val-d'Essone,
91025 Evry Cedex,
France.
E-mail: `feyel@maths.univ-evry.fr`

Ali Süleyman Üstünel
ENST,
Dépt. Infres,
46, rue Barrault, 75634 Paris Cedex 13,
France.
E-mail: `ustunel@enst.fr`

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 75–89

Function Spaces and Symmetric Markov Processes

Masatoshi Fukushima

Dedicated to Professor Kiyosi Itô on the occasion of his 88th birthday

Abstract.

We exhibit some mutual interactions between potential theory for concrete function spaces on $\mathbb{R}^n$ and the Dirichlet space theory associated with symmetric Markov processes. Our first concern is the role of the Dirichlet form version of the capacitary strong type inequality in the study of the ultracontractivity of the transition semigroup of time changed symmetric Markov processes. In particular, we study time changes of symmetric stable processes in relation to d-bounds of measures. We next show how the theory on capacity and the spectral synthesis for the Dirichlet space can be well inherited to a general function space with contractive p-norm. A link connecting those two topics is a contractive Besov space over a d-set of $\mathbb{R}^n$.

§1. Introduction

Since the publication of the seminal work of Beurling and Deny [5], their axiomatic potential theory of the Dirichlet space $(\mathcal{F}, \mathcal{E})$ has been unified under one roof with the theory of the symmetric Markov process $\mathbf{M}$. In particular, any σ-finite positive measure μ charging no set of zero capacity can now be studied in relation to the trace Dirichlet space $(\check{\mathcal{F}}, \check{\mathcal{E}})$ on the support F of μ and the time changed process $\check{\mathbf{M}}$ on F of $\mathbf{M}$ by means of the positive continuous additive functional associated with μ ([13]).

In §2, we shall see for $\kappa \in (0,1)$ that a simple *capacitary isoperimetric inequality*

$$\mu(K)^\kappa \leq \Theta \ \mathrm{Cap}(K), \qquad \forall K(\text{compact}), \tag{1}$$

Received February 28, 2003.

is equivalent to the *ultracontractivity*

$$(2) \qquad \check{p}_t(x,y) \le \left(\frac{H}{t}\right)^{\frac{1}{1-\kappa}}, \qquad t>0,$$

of the transition function $\check{p}_t$ of $\check{\mathbf{M}}$, with the isoperimetric constant Θ for the measure μ and the heat constant H for the process $\check{\mathbf{M}}$ controlling each other. When the (extended) Dirichlet space is the Riesz potential space $\dot{L}^{\alpha,2}(\mathbb{R}^n)$ and $\mathbf{M}$ is the symmetric 2α-stable process $(0<\alpha<1)$, we shall also see in §3 that the isoperimetric constant can be replaced by the *d-bound*

$$(3) \qquad v_d(\mu) = \sup_{x\in\mathbb{R}^n, r>0} \frac{\mu(B(x,r))}{r^d},$$

of the measure μ. Detailed proof of theorems in §2 and §3 can be found in [16].

An important ingredient in proving the above equivalence is the *capacitary strong type inequality*

$$(4) \quad \int_0^\infty \mathrm{Cap}(\{x\in X : |u(x)|\ge t\})d(t^2) \le 4\mathcal{E}(u,u) \quad \forall u \in \mathcal{F}\cap C_0(X),$$

which readily ensures the equivalence of (1) to a Sobolev type imbedding of the trace Dirichlet space $\check{\mathcal{F}}$. We can then invoke the works by Carlen,Kusuoka and Stroock[8] and Bakry,Coulhon,Ledoux and Saloff-Coste[2] to relate (1) and (2).

In the meantime, potential theory have advanced being modelled on concrete function spaces like Sobolev spaces $W^{r,p}$, Bessel potential spaces $L^{\alpha,p}$, Besov spaces $B^{p,q}_\alpha$ and so on. Imbedding theorems and spectral synthesis have been among important issues in potential theory ([1], [4], [23]).

Actually the capacitary strong type inequality was first established by Maz'ya[22] for the Sobolev space $W^{1,p}(\mathbb{R}^n)$, $1<p<\infty$. It was then extended to a large class of function spaces on $\mathbb{R}^n$ including the Riesz and Bessel potential spaces. It has been also proved in [21], [14] for a general function space with *contractive p-norm* $(1\le p<\infty)$

$$(5) \qquad \begin{cases} |||u|||_p^p &= \int_{X\times X\setminus d} |u(x)-u(y)|^p N(x,dy)m(dx) \\ \mathcal{F}_p &= \{u\in L^p(X;m) : |||u|||_p^p < \infty\}, \end{cases}$$

which include as an important example the *contractive Besov space* $B^{p,p}_\alpha(F)$, $0<\alpha<1, 1\le p<\infty$, over a d-set $F\subset\mathbb{R}^n$ defined as

(6)
$$\begin{cases} \|u; B_\alpha^{p,p}(F)\| & = \|u\|_{L^p(F;\mu)} + \left(\iint_{F\times F} \frac{|u(x)-u(y)|^p}{|x-y|^{d+\alpha p}}\mu(dx)\mu(dy)\right)^{1/p} \\ B_\alpha^{p,p}(F) & = \{u \text{ is measurable } : \ \|u; B_\alpha^{p,p}(F)\| < \infty\}. \end{cases}$$

μ being taken to be the restriction to F of the d-dimensional Hausdorff measure.

Its Dirichlet space version (4) accompanied by the best constant 4 was proved rather recently by Vondraček [25]. [16] provides an alternative simple proof of (4).

When $p = 2$, the contractive Besov space on a d-set is a regular Dirichlet space on $L^2(F;\mu)$ and the properties of the associated jump type Markov process on F have been studied in [14], [6] and [9]. As we shall see in §3, this space is closely related to the Dirichlet space $(\check{\mathcal{F}}, \check{\mathcal{E}})$ on $L^2(F;\mu)$ of the time changed process of a symmetric stable process on $\mathbb{R}^n$ in the sense that the former is continuously imbedded into the latter, although these two spaces are generally different because the latter may involve a killing term in general.

Even when $p \neq 2$, the function space (5) with contractive p-norm shares with the Dirichlet space a common feature that every normal contraction operates on it and deserves to be studied on its own light. We shall see in §4 that the well known theory on capacity and spectral synthesis for the Dirichlet space ([5], [10], [13]) can be well inherited to the function space (5).

In particular, the spectral synthesis is possible for the contractive Besov space on a d-set $F \subset \mathbb{R}^n$ for $1 < p < \infty$. As an application, we shall get in §4 the following criterion for an relatively open set $H \subset F$ such that $F \setminus H$ has a locally finite positive $\tilde{d}$-dimensional Hausdorff measure with $\tilde{d} < d$:

$$B_{\alpha,0}^{p,p}(H) = B_\alpha^{p,p}(F) \Longleftrightarrow \alpha \leq \frac{d-\tilde{d}}{p}, \tag{7}$$

$B_{\alpha,0}^{p,p}(H)$ being the closure of $B_\alpha^{p,p}(F) \cap C_0(H)$ in the space $B_\alpha^{p,p}(F)$.

This completes and extends the corresponding results by Caetano [7] and Farkas and Jacob [11]. When $p = 2$, $d = n$, $F = \overline{D}$, $H = D$, for an open set $D \subset \mathbb{R}^n$, (7) has been shown by Bogdan,Burdzy and Chen [6] giving a complete characterization for almost no sample path of the censored 2α-stable process on D to approach the boundary ∂D in finite time. Detailed proof of theorems in §4 can be found [15].

§2. Capacitary bounds of measures and time changed processes

Let $(X, m, \mathcal{E}, \mathcal{F})$ be a regular transient Dirichlet space. By this, we mean that X is a locally compact separable metric space, m is an everywhere dense positive Radon measure on X, and that $(\mathcal{E}, \mathcal{F})$ is a regular transient Dirichlet form on $L^2(X; m)$. The 0-order capacity of a compact set $K \subset X$ is then defined by

$$\text{(8)} \qquad \mathrm{Cap}(K) = \inf \{\mathcal{E}(u,u) : u \in \mathcal{F} \cap C_0(X),\ u(x) \geq 1,\ x \in K\}$$

and extended to any subsets of X as a Choquet capacity. $\mathcal{F}_e$ denotes the extended Dirichlet space. In what follows, any function $u \in \mathcal{F}_e$ will be always taken to be quasi-continuous (cf. [13]).

Owing to Vondraček [25], we then have the capacitary strong type inequality (4), which in turn implies the following (cf. [1, §7.2]):

Theorem 1. *Let μ be a Borel measure on X and $\kappa \in (0,1]$.*
(i) *If the capacitary isoperimetric inequality* (1) *holds for some positive constant Θ, then μ is a smooth Radon measure and*

$$\text{(9)} \qquad ||u||^2_{L^{2/\kappa}(X;\mu)} \leq S\mathcal{E}(u,u), \quad \forall\, u \in \mathcal{F}_e,$$

for some positive constant $S \leq (4/\kappa)^\kappa \Theta$.
(ii) *Conversely, if* (9) *holds for any $u \in \mathcal{F} \cap C_0(X)$ and for some positive constant S, then* (1) *holds for some positive constant $\Theta \leq S$.*

For a measure μ on X, we introduce its *isoperimetric constant* and *Sobolev constant* respectively by

$$\text{(10)} \qquad \Theta_\kappa(\mu) = \sup_K \frac{\mu(K)^\kappa}{\mathrm{Cap}(K)} \qquad \kappa \in (0,1],$$

$$\text{(11)} \qquad S_\eta(\mu) = \sup_{u \in \mathcal{F} \cap C_0(X)} \frac{||u||^2_{L^\eta(\mu)}}{\mathcal{E}(u,u)} \qquad \eta \in [2, \infty).$$

The supremum in (11) can be taken for all $u \in \mathcal{F}_e$. $S_2(\mu)$ may be called the *Poincaré constant* of μ. Theorem 1 can be rephrased as follows:

Corollary 1. *For a measure μ on X and for $\kappa \in (0,1]$, $0 < \Theta_\kappa(\mu) < \infty$ if and only if $0 < S_{2/\kappa}(\mu) < \infty$. Moreover,*

$$\text{(12)} \qquad \Theta_\kappa(\mu) \leq S_{2/\kappa}(\mu) \leq (4/\kappa)^\kappa\, \Theta_\kappa(\mu), \qquad \kappa \in (0,1].$$

Let $\mathbf{M} = \{X_t, P_x\}$ be an m-symmetric Hunt process on X associated with the Dirichlet form $\mathcal{E}$ and $A = A_t$ be a PCAF of $\mathbf{M}$ whose Revuz measure is a given smooth Radon measure μ. Denote by F and $\tilde{F}$ the support of μ and A respectively. Then $\tilde{F} \subset F$ q.e., $\mu(F \setminus \tilde{F}) = 0$ and further $\tilde{F}$ is a quasi-support of μ, namely, if quasi-continuous functions coincide μ-a.e., then they coincide q.e. on $\tilde{F}$. Recall that each element $u \in \mathcal{F}_e$ is taken to be quasi-continuous.

We consider the time changed process $\check{\mathbf{M}} = (\check{X}_t, P_x)_{x \in \tilde{F}}$ defined by

$$\check{X}_t = X_{\tau_t} \qquad \tau_t = \inf\{s > 0 : A_s > t\}.$$

$\check{\mathbf{M}}$ is a μ-symmetric transient right process, whose Dirichlet form $(\check{\mathcal{E}}, \check{\mathcal{F}})$ on $L^2(F;\mu)$ and the extended Dirichlet space $\check{\mathcal{F}}_e$ can be described as follows (cf. [13, §6.2]) :

$$\check{\mathcal{F}}_e = \{\varphi = u|_F \, \mu - a.e. : u \in \mathcal{F}_e\} \qquad \check{\mathcal{F}} = \check{\mathcal{F}}_e \cap L^2(F;\mu) \tag{13}$$

$$\check{\mathcal{E}}(\varphi,\varphi) = \mathcal{E}\left(H_{\tilde{F}}u, H_{\tilde{F}}u\right) \quad \varphi = u|_F \in \check{\mathcal{F}}_e, \tag{14}$$

where

$$H_{\tilde{F}}u(x) = E_x\left(u\left(X_{\sigma_{\tilde{F}}}\right)\right) \quad x \in X,$$

E_x denoting the expectation with respect to P_x and $\sigma_{\tilde{F}}$ being the hitting time of the set $\tilde{F}$ by the sample path X_t. Two elements of $\check{\mathcal{F}}_e$ are regarded identical if they coincides μ-a.e. Since $\tilde{F}$ is a quasi-support of μ, the definition (14) of $\check{\mathcal{E}}$ makes sense.

We can restate (14) as follows (the Dirichlet principle):

$$\check{\mathcal{E}}(\varphi,\varphi) = \inf\{\mathcal{E}(u,u) : u \in \mathcal{F}_e,\ u = \varphi \ \mu\text{-a.e. on } F\}, \quad \varphi \in \check{\mathcal{F}}_e. \tag{15}$$

The first half of the next theorem is immediate from (9) and (15).

Theorem 2. *Suppose a measure μ satisfies $\Theta_\kappa(\mu) \in (0,\infty)$ for some $\kappa \in (0,1)$.*
Then the following holds for $S = S_{2/\kappa}(\mu)(\in (\Theta_\kappa(\mu), (4/\kappa)^\kappa \Theta_\kappa(\mu)))$.
(i)

$$||\varphi||^2_{L^{2/\kappa}(F;\mu)} \le S\ \check{\mathcal{E}}(\varphi,\varphi) \qquad \forall \varphi \in \check{\mathcal{F}}_e. \tag{16}$$

(ii) *The transition function $\check{p}_t$ of the time changed process $\check{\mathbf{M}}$ on F satisfies the ultracontractivity* (2) *for $\mu \times \mu$*–a.e. $(x,y) \in F \times F$, *where H is some positive constant with*

$$H \le \frac{1}{1-\kappa} \cdot S. \tag{17}$$

We know that (1) and (16) are equivalent by Voropoulos [24]. But we are more concerned with dependence of the isoperimetric constant Θ_κ and the heat constant H.

Simple mutual dependence of H and the constant N appearing in the Nash type inequality has been well studied in [8]. The Sobolev inequality (16) can be readily converted by a Hölder inequality into the Nash type inequality with $N = S$ and we can get the bound (17) easily. On the other hand, we know that the Sobolev inequality can be derived from the Nash type inequality under a certain control of S by N in view of [2, Cor, 4.4, Cor. 7.3], and we can get the following converse to Theorem 2.

Theorem 3. *Suppose that μ is a smooth Radon measure with support F and that the transition function $\check{p}_t$ of the time changed process $\check{\mathbf{M}}$ on F with respect to the PCAF with Revuz measure μ satisfies the ultracontractivity* (2) *for some $\kappa \in (0,1)$, $H > 0$. Then*
(i) *The Sobolev inequality* (16) *holds for some positive constant S with*

$$S \leq 48\, e^2\, \frac{1}{\kappa} \left(\frac{2-\kappa}{1-\kappa}\right)^{\frac{2-\kappa}{1-\kappa}} \cdot H. \tag{18}$$

(ii) *μ admits an isoperimetric constant $\Theta_\kappa(\mu)$ with a bound*

$$(4/\kappa)^{-\kappa}\, S \leq \Theta_\kappa(\mu) \leq S \tag{19}$$

by the constant S of (i).

Tierry Coulhon has called author's attention to the relevance of the capacitary isoperimetric inequality (1) to the Faber-Krahn inequality.

For an open set $G \subset X$, we put

$$\mathcal{F}_G = \{u \in \mathcal{F} : u = 0 \quad \text{q.e. on } X \setminus G\}.$$

Due to the spectral synthesis theory for the Dirichlet space, $\mathcal{E}$ with domain $\mathcal{F}_G$ can be considered as a regular Dirichlet form on $L^2(G; m)$ which is called the part of $(\mathcal{E}, \mathcal{F})$ on G ([13, §4.4]). For a measure μ on X, we let

$$\lambda_1(\mu; G) = \inf_{u \in \mathcal{F}_G} \frac{\mathcal{E}(u,u)}{\|u\|^2_{L^2(\mu)}} \left(= \inf_{u \in \mathcal{F} \cap C_0(G)} \frac{\mathcal{E}(u,u)}{\|u\|^2_{L^2(\mu)}} \right),$$

which may be regarded, on account of the Dirichlet principle (15), as the first eigenvalue for the part of the trace Dirichlet space $(\check{\mathcal{F}}, \check{\mathcal{E}})$ on the relatively open sunset $F \cap G$ of F. Since $\lambda_1(\mu; G)$ is the reciprocal of the

Poincaré constant $S_2(\mu; G)$ defined by (11) for the part form $(\mathcal{E}, \mathcal{F}_G)$, we get from (12)

$$\frac{1}{\lambda_1(\mu; G)} \leq 4 \sup_{K \subset G} \frac{\mu(K)}{\text{Cap}(K; G)}, \tag{20}$$

where $\text{Cap}(K; G)$ is defined by (8) with X being replaced by G.

Let us assume that $\Theta_\kappa(\mu)$ is finite for some $\kappa \in (0,1)$. Since $\text{Cap}(K; G) \geq \text{Cap}(K)$ for $K \subset G$, we have

$$\frac{1}{\text{Cap}(K; G)} \leq \frac{\Theta_\kappa(\mu)}{\mu(K)^\kappa}. \tag{21}$$

(20) and (21) lead us to

$$\frac{1}{\lambda_1(\mu; G)} \leq 4 \sup_{K \subset G} \Theta_\kappa(\mu) \cdot \mu(K)^{1-\kappa} = 4\Theta_\kappa(\mu) \cdot \mu(G)^{1-\kappa}$$

and

$$\lambda_1(\mu; G) \geq \frac{1}{4\Theta_\kappa(\mu)} \cdot \frac{1}{\mu(G)^{1-\kappa}} \tag{22}$$

for any open set $G \subset X$ of finite μ-measure.

(22) is called the *Faber-Krahn inequality* and the above procedure of getting (22) from (1) using the capacitary strong type inequality has been indicated by Grigor'yan [18]. Very intimate relationship among the Faber-Krahn inequality, ultracontractivity (2) and the Nash type inequality has been studied in [19]. However, in order to recover the capacitary isoperimetric inequality (1) from the ultracontractivity (2), one may need to path through Nash type inequality to Sobolev's one as being done in this section.

§3. Application to time changes of symmetric stable processes on d-sets

In this section, we consider the symmetric 2α-stable process $\mathbf{M} = (X_t, P_x)$ on $\mathbb{R}^n$ for $0 < \alpha \leq 1, 2\alpha < n$. The transition function of $\mathbf{M}$ is a convolution semigroup $\{\nu_t, t > 0\}$ of symmetric probability measures on $\mathbb{R}^n$ with

$$\hat{\nu}_t(x) \left(= \int_{\mathbb{R}^n} e^{i(x,y)} \nu_t(dy) \right) = e^{-tc|x|^{2\alpha}},$$

c being a fixed positive constant. For simplicity, we take $c = 1$. In case that $\alpha = 1$, $\mathbf{M}$ is the n-dimensional Brownian motion with variance of

μ_t being equal to $2t$. $\mathbf{M}$ is transient. The Dirichlet form $(\mathcal{E},\mathcal{F})$ of $\mathbf{M}$ on $L^2(\mathbb{R}^n)$ is given by

$$(23)\qquad \begin{cases} \mathcal{E}(u,u) &= \displaystyle\int_{\mathbb{R}^n} \hat{u}(x)\bar{\hat{v}}(x)|x|^{2\alpha}dx \\ \mathcal{F} &= \{u \in L^2(\mathbb{R}^n) : \displaystyle\int_{\mathbb{R}^n} |\hat{u}(x)|^2|x|^{2\alpha}dx < \infty\}. \end{cases}$$

The extended Dirichlet space $(\mathcal{F}_e,\mathcal{E})$ of $\mathbf{M}$ can then be identified with the Riesz potential space $\dot{L}^{\alpha,2}(\mathbb{R}^n) = \{I_\alpha * f : f \in L^2(\mathbb{R}^n)\}$, where the Riesz potential of a measure ν on $\mathbb{R}^n$ is defined by

$$I_\alpha * \nu(x) = \gamma_\alpha \int_{\mathbb{R}^n} |x-y|^{-(n-\alpha)}\nu(dy), \quad \gamma_\alpha = \frac{\Gamma((n-\alpha)/2))}{\pi^{n/2}2^\alpha\Gamma(\alpha/2)}.$$

The capacity defined by (8) for the present Dirichlet form coincides with the Riesz capacity defined for any compact set $K \subset \mathbb{R}^n$ by

$$(24)\quad \dot{C}_{\alpha,2}(K) = \inf\{||f||^2_{L^2(\mathbb{R}^n)} : f \in L^2_+(\mathbb{R}^n),\ I_\alpha * f(x) \geq 1\ \forall x \in K\}.$$

We call a closed subset F of $\mathbb{R}^n$ a (semi global) *d-set* for $0 < d \leq n$ if there exists a positive measure μ supported by F satisfying, for some constants $0 < c_1 \leq c_2$,

$$c_1 r^d \leq \mu(B(x,r)) \quad \forall x \in F, \forall r \in (0,1)$$
$$\mu(B(x,r)) \leq c_2 r^d \quad \forall x \in F, \forall r \in (0,\infty),$$

where $B(x,r)$ denotes the n-dimensional ball with center x and radius r. Such a measure is called a *d-measure.* It is known that the restriction of the d-dimensional Hausdorff measure to a d-set F is a d-measure (cf.[20]).

For a d-measure μ, we will be concerned with its *d-bound* defined by (3). We consider a d-measure μ on a d set F with

$$n - 2\alpha < d \leq n.$$

Otherwise, $\dot{C}_{\alpha,2}(F) = 0$ and μ can not satisfy the isoperimetric inequality with respect to the present Dirichlet form. Since

$$\dot{C}_{\alpha,2}(B(x,r)) = \dot{c}_{\alpha,2}\ r^{n-2\alpha}, \qquad \dot{c}_{\alpha,2} = \dot{C}_{\alpha,2}(B(0,1)),$$

we can immediately obtain a lower bound of the isoperimetric constant for μ by its d-bound:

$$(25)\qquad \dot{c}^{-1}_{\alpha,2}\ v_d(\mu)^{\frac{n-2\alpha}{d}} \ \leq\ \Theta_{\frac{n-2\alpha}{d}}(\mu).$$

We can also obtain an inequality in the opposite direction:

Theorem 4. *For any Radon measure μ with finite d-bound, it holds that*

$$\Theta_{\frac{n-2\alpha}{d}}(\mu) \leq c(n,\alpha,d)\, v_d(\mu)^{\frac{n-2\alpha}{d}} \tag{26}$$

for

$$c(n,\alpha,d) = \frac{4d^2\, \gamma_\alpha^2\, v_n\, (n-\alpha)^2}{(n-2\alpha)^2\{d-(n-2\alpha)\}^2}, \tag{27}$$

where v_n is the volume of the n dimensional unit ball.

By setting $\kappa = \frac{n-2\alpha}{d}$ in Corollary 1 and using (25) and (26), we get the bound of the Sobolev constant $S = S_{\frac{2d}{n-2\alpha}}(\mu)$ for μ in terms of its d-bound $v_d(\mu)$:

$$\dot{c}_{\alpha,2}^{-1}\, v_d(\mu)^{\frac{n-2\alpha}{d}} \leq S \leq (4d/(n-2\alpha))^{\frac{n-2\alpha}{d}}\, c(n,\alpha,d)\, v_d(\mu)^{\frac{n-2\alpha}{d}} \tag{28}$$

for the constant $c(n,\alpha,d)$ of (27).

By setting $\kappa = \frac{n-2\alpha}{d}$ in Theorem 1 and Theorem 2, we have

Theorem 5. *Suppose μ is a d-measure on $\mathbb{R}^n$ with $n-2\alpha < d \leq n$. Then we have the following for S satisfying the bounds* (28)*:*
(i)

$$||u||^2_{L^{\frac{2d}{n-2\alpha}}(\mathbb{R}^n;\mu)} \leq S\, \mathcal{E}(u,u) \qquad \forall u \in \dot{L}^{\alpha,2}(\mathbb{R}^n). \tag{29}$$

(ii) *Let $\check{\mathbf{M}}$ be the time changed process on the support F of μ of $\mathbf{M}$ by the PCAF with Revuz measure μ. Then its transition function $\check{p}_t$ satisfies*

$$\check{p}_t(x,y) \leq \left(\frac{H}{t}\right)^{\frac{d}{d-(n-2\alpha)}}, \qquad t>0, \tag{30}$$

for $\mu\times\mu$-a.e. $(x,y)\in F\times F$, where H is some positive constant with

$$H \leq \frac{d}{d-(n-2\alpha)}\, S. \tag{31}$$

Actually inequality (29) together with the bounds

$$c_3\, v_d(\mu)^{\frac{n-2\alpha}{d}} \leq S \leq c_4\, v_d(\mu)^{\frac{n-2\alpha}{d}}$$

holding for some positive constants c_3, c_4 independent of μ goes back to the work of Adams ([23, 1.4.1]). Here we have made these contants c_3 and c_4 more explicit in (28).

We can also derive from Theorem 3 the following converse to Theorem 4.

Theorem 6. *Suppose that μ is a smooth Radon measure on $\mathbb{R}^n$ with support F and that the transition function $\tilde{p}_t$ of the time changed process $\tilde{\mathbf{M}}$ on F with respect to the PCAF with Revuz measure μ satisfies the bound* (30) *for some $d \in (n-2\alpha, n]$ and $H > 0$. Then*
(i) *The inequality* (29) *holds for some positive constant S with*

$$S \leq \frac{48de^2}{n-2\alpha}\left(\frac{2d-(n-2\alpha)}{d-(n-2\alpha)}\right)^{\frac{2d-(n-2\alpha)}{d-(n-2\alpha)}} \cdot H. \tag{32}$$

(ii) *μ is a d-measure whose d-bound $v_d(\mu)$ satisfies*

$$\frac{n-2\alpha}{4d}\left(\frac{S}{c(n,\alpha,d)}\right)^{\frac{d}{n-2\alpha}} \leq v_d(\mu) \leq (\dot{c}_{\alpha,2}\ S)^{\frac{d}{n-2\alpha}} \tag{33}$$

for the constant S of (i) *and for $c(n,\alpha,d)$ of* (27).

Let $\mu, F, \check{\mathbf{M}}$ be as in Theorem 5 and $(\check{\mathcal{E}}, \check{\mathcal{F}})$ be the Dirichlet form of $\check{\mathbf{M}}$ on $L^2(F;\mu)$ the trace Dirichlet form of (23) on the d-set F. Put

$$\delta = \alpha - \frac{n-d}{2} \in (0,1] \tag{34}$$

and consider the Besov space $B_\delta^{2,2}(F)$ over F defined by
(35)

$$\begin{cases} (\varphi,\psi)_{B_\delta^{2,2}(F)} &= \displaystyle\int_{F\times F\backslash d} \frac{(\varphi(x)-\varphi(y))(\psi(x)-\psi(y))}{|x-y|^{d+2\delta}}\mu(dx)\mu(dy) \\ B_\delta^{2,2}(F) &= \{\varphi \in L^2(F;\mu) : (\varphi,\varphi)_{B_\delta^{2,2}(F)} < \infty\}. \end{cases}$$

$B_\delta^{2,2}(F)$ is a Dirichlet form on $L^2(F;\mu)$ equipped with the norm

$$||\varphi; B_\delta^{2,2}(F)||^2 = (\varphi,\varphi)_{L^2(F;\mu)} + (\varphi,\varphi)_{B_\delta^{2,2}(F)}.$$

By virtue of a Jonsson-Wallin trace theorem [20, chap. V], this space is related to the Bessel potential space $L_{\alpha,2}(\mathbb{R}^n)$ as

$$B_\delta^{2,2}(F) = L_{\alpha,2}(\mathbb{R}^n)|_F \tag{36}$$

both the restriction and extension operators involved being continuous. Since the present Dirichlet space (23) equipped with $\mathcal{E}_1$-norm is known to be equivalent to the Bessel potential space, we are led from the Dirichlet principle (15) and (36) to the following continuous embedding:

$$(37) \quad B_\delta^{2,2}(F) \subset \check{\mathcal{F}}_e, \qquad \check{\mathcal{E}}(\varphi,\varphi) \le C\ \|\varphi; B_\delta^{2,2}(F)\|^2,\ \forall\varphi \in B_\delta^{2,2}(F),$$

for some positive constant C.

Nevertheless 0-oder forms $\check{\mathcal{E}}$ and $(\cdot,\cdot)_{B_\delta^{2,2}(F)}$ are not necessarily equivalent. For instance, let $\mathbf{M}$ be the standard Brownian motion on $\mathbb{R}^n$ with $n \ge 3$, F be the unit sphere Σ contered at the origin and μ be the surface measure σ on Σ. Then we have the following expression of the trace Dirichlet form $\check{\mathcal{E}}(f,f)$ for $f \in \check{\mathcal{F}}$ ([17]):

(38)

$$\check{\mathcal{E}}(f,f) = \frac{1}{\Omega}\int_{\Sigma\times\Sigma\setminus d}(f(\xi)-f(\eta))^2\frac{1}{|\xi-\eta|^n}\sigma(d\xi)\sigma(d\eta)+v_0\int_\Sigma f(\xi)^2\sigma(d\xi),$$

where Ω is the area of Σ and $v_0 = \frac{n-2}{2}$. The first term on the right hand side correponds to the form (35) for $d = n-1, \delta = 1/2$. But the additional second term appears due to the transience of the Brownian motion.

§4. Spectral synthesis for contractive p-norms and Besov spaces

Let X be a locally compact separable metric space and m a positive Radon measure on X with $\mathrm{supp}[m] = X$. Let $N(x,dy)$ be a positive kernel on $(X,\mathcal{B}(X))$ such that $N(x,\{x\}) = 0,\ x \in X$, and $N(x,dy)m(dx)$ is a symmetric measure over $X \times X - d$, where $d = \{(x,x) : x \in X\}$. For a fixed $1 \le p < \infty$, we introduce the pseudo-norm $|||\cdot|||_p$ and the function space $\mathcal{F}_p$ by (5). Denoting the norm of the space $L^p(X;m)$ by $||\cdot||_p$, we equip $\mathcal{F}_p$ with the norm

$$(39) \qquad |||u|||_{p,1} = |||u|||_p + ||u||_p \qquad u \in \mathcal{F}_p \cap L^p(X;m).$$

We assume the regularity of this space in the sense that $\mathcal{F}_p \cap C_0(X)$ is dense in $\mathcal{F}_p$ with norm (39) and in $C_0(X)$ with uniform norm.

Denote by $\mathcal{O}$ the family of all open sets in X. We define the *p-capacity* of $A \in \mathcal{O}$ by

(40)

$$\mathrm{Cap}_p(A) = \inf\{|||u|||_p^p + ||u||_p^p : u \in \mathcal{F}_p,\ u \ge 1\ m\text{-a.e. on } A\} \quad A \in \mathcal{O},$$

and extend it to any set $B \subset X$ by

$$\mathrm{Cap}_p(B) = \inf\{\mathrm{Cap}_p(A) : A \in \mathcal{O},\ B \subset A\}.$$

'q.e.' will mean *'except for a set of zero p-capacity'*. Cap_p-quasicontinuous function will be called simply *quasicontinuous*. In what follows, we also assume that $1 < p < \infty$.

Although the space $(\mathcal{F}_p, |||\cdot|||_{p,1}$ is slightly more complicated than the ordinary L^p space, we can well adopt the uniform convexity argument to ensure the unique existence of the equilibrium potential for any $A \in \mathcal{O}$ with finite p-capacity. Thus Cap_p on open sets can be seen to enjoy the continuity along the increasing limit as in [12]. It is also strongly subadditive as in [21]. Hence Cap_p is a Choquet capacity, each element $u \in \mathcal{F}_p$ has a quasicontinuous version $\tilde{u}$, each set of finite p-capacity has a unique equilibrium potential just as in the case of the Dirichlet space. We also have the following nice property:

(41)

u is quasi-continuous and $u = 0$ m-a.e. on $G(\in \mathcal{O}) \Longrightarrow u = 0$ q.e. on G.

For $G \in \mathcal{O}$, we let

$$\mathcal{F}_{p,0}^G = \overline{\mathcal{F}_p \cap C_0(G)}^{|||\cdot|||_{p,1}}, \tag{42}$$

where $C_0(G)$ denotes the family of continuous functions on X whose support is compact and contained in G. We say that *the spectral synthesis is possible for* $G \in \mathcal{O}$ if

$$\mathcal{F}_{p,0}^G = \{u \in \mathcal{F}_p : \tilde{u} = 0 \ q.e. \text{ on } X \setminus G\}. \tag{43}$$

Following the method of [1, §9.2] for the space $W^{1,p}(\mathbb{R}^n)$ and using the contraction property of the space $\mathcal{F}_p$ together with the above mentioned properties of Cap_p, we can prove the next theorem.

Theorem 7. (i) *The spectral synthesis is possible for* $G \in \mathcal{O}$ *if* $X \setminus G$ *is compact.*

(ii) *The spectral synthesis is possible for any* $G \in \mathcal{O}$ *under the next assumption:*

(A) *There exist non-negative functions* $w_n \in C_0(X)$ *increasing to* 1 *such that*

$$\sup_n \sup_{x\in X} W_n(x) < \infty, \ \lim_{n\to\infty} \sup_{x\in K} W_n(x) = 0 \text{ for any compact } K \subset X,$$

where

$$W_n(x) = \int_X |w_n(x) - w_n(y)|^p N(x, dy) \qquad x \in X. \tag{44}$$

As a consequence of Theorem 7 (i), the next useful identity holds for any compact set $K \subset X$:

$$(45) \quad \mathrm{Cap}_p(K) = \inf\{|||u|||_p^p + ||u||_p^p : u \in \mathcal{F}_p \cap C_0(X),\ u \geq 1 \text{ on } K\}.$$

We now let

$$0 < d \leq n, \quad < \alpha < 1, \quad 1 < p < \infty,$$

and consider the contractive Besov space $B_\alpha^{p,p}(F)$ on a d-set $F \subset \mathbb{R}^n$ defined by (6). This is a special example of the space $\mathcal{F}_p$ with contractive p-norm $||| \cdot |||_{p,1}$. The associated p-capacity of a set $A \subset F$ is denoted by $\mathrm{Cap}_{\alpha,p}(A; F)$. It can be shown that condition **A** is satisfied by this space. By Theorem 7, the spectral synthesis is therefore possible for any relatively open set $H \subset F$ with respect to $B_\alpha^{p,p}(F)$, which immediately implies the equivalence

$$(46) \qquad B_{\alpha,0}^{p,p}(H) = B_\alpha^{p,p}(F) \Longleftrightarrow \mathrm{Cap}_{\alpha,p}(F \setminus H; F) = 0,$$

where $B_{\alpha,0}^{p,p}(H)$ denotes the closure of $B_\alpha^{p,p}(F) \cap C_0(H)$ in the space $B_\alpha^{p,p}(F)$.

On the other hand, the next implications have been proved in [14] by making use of the property (45) of $\mathrm{Cap}_{\alpha,p}(\cdot; F)$, a Jonsson-Wallin trace theorem ([20]) and the metric properties of the Bessel capacity on $\mathbb{R}^n$ ([1]):

$$(47) \qquad \mathrm{Cap}_{\alpha,p}(\Lambda; F) = 0 \Longrightarrow \mathcal{H}_{dim}(\Lambda) \leq d - \alpha p,$$

$$(48) \qquad H_{d-\alpha p}(\Lambda) < \infty \Longrightarrow \mathrm{Cap}_{\alpha,p}(\Lambda; F) = 0.$$

Here $\mathcal{H}_{dim}$ and H_γ denote the Hausdorff dimension and γ-dimensional Hausdorff measure respectively.

(46),(47) and (48) lead us to the next desired theorem.

Theorem 8. *Assume that H is a relatively open subset of F and that $F \setminus H$ has a locally finite positive $\tilde{d}$-dimensional Hausdorff measure with $\tilde{d} < d$. Then $B_{\alpha,0}^{p,p}(H) = B_\alpha^{p,p}(F)$ if and only if $\alpha \leq \dfrac{d - \tilde{d}}{p}$.*

References

[1] D. R. Adams and L. I. Hedberg, *Function Spaces and Potential Theory*, Springer-Verlag, 1996

[2] D. Bakry, T. Coulhon, M. Ledoux and L. Saloff-Coste, Sobolev inequalities in disguise, *Indiana Univ. Math. J.* 44(1995), 1033-1074

[4] A. Beurling, *The Collected Works of Arne Beurling,* eds L. Carleson, P. Malliavin, J. Neuberger, and J. Wermer, Birkhäuser, 1989

[5] A. Beurling and J. Deny, Dirichlet spaces, *Proc. Nat. Acad. Sci. U.S.A.* 45(1959),208-215

[6] K. Bogdan, K. Burdzy and Z.Q. Chen, Censored stable processes, *Probab. Theory Relat. Fields* to appear

[7] A. Caetano, Approximation by functions of compact support in Besov-Triebel-Lizorkin spaces on irregular domains, *Studia Math.,* 142(2000), 47-63

[8] E. Carlen, S. Kusuoka and D. Stroock, Upper bounds for symmetric Markov transition functions, *Ann. Inst. H. Poincaré, Prob.Stat.* 23(1987), 245-287

[9] Z.Q Chen and T. Kumagai, Heat kernel estimate for stable-like processes on d-sets, Preprint

[10] J. Deny, Méthodes Hilbertiennes et théorie du potentiel, in *Potnetial Theory,* Centro Internazionale Mathematico Estovo, Edizioni Cremonese, Romea, 1970

[11] W. Farkas and N. Jacob, Sobolev spaces on non-smooth domains and Dirichlet forms related to subordinate reflecting diffusions, *Math. Nachr.,* 224(2001), 75-104

[12] M. Fukushima and H. Kaneko, On (r, p)-capacities for general Markovian semi-groups, in *Stochastic process and infinite dimensional analysis,* ed. by S. Albeverio, Pitman, 1985

[13] M. Fukushima, Y. Oshima and M. Takeda, *Dirichlet Forms and Symmetric Markov Processes*, Walter de Gruyter, 1994

[14] M. Fukushima and T. Uemura, On Sobolev and capacitary inequalities for contractive Besov spaces over d-sets, *Potential Anal.* 18(2003), 59-77

[15] M. Fukushima and T. Uemura, On spectral synthesis for contractive p-norms and Besov spaces, *Potential Anal.,* to appear

[16] M. Fukushima and T. Uemura, Capacitary boundes of measures and ultracontractivity of time changed processes, *J. Math. Pures Appl.* 82(2003), 553-572

[17] M. Fukushima, P. He and J. Ying, Time changes of symmetric diffusions and Feller measures, Preprint

[18] A. Grigor'yan, Isoperimetric inequalities and capacities on Riemannian manifolds, *Operator Theory; Advances and Applications,* 109(1999), 139-153

[19] A. Grigor'yan, Heat kernel upper bounds on a complete non-compact manifold, *Rev. Mat. Iberoamericana,* 10-2(1994), 395-452

[20] A. Jonsson and H. Wallin, *Function spaces on subsets of* $\mathbb{R}^n$, Harwood Academic Pub., 1984

[21] V. Kaimanovich, Dirichlet norms, capacities and generalized isoperimetric inequalities for Markov operators, *Potential Anal.* 1(1992), 61-82

[22] V.G. Maz'ya, On certain integral inequalities for functions of many variables, *J. Soviet Math.* 1(1973) 205-237

[23] V.G. Maz'ya, *Sobolev spaces*, Springer-Verlag, 1985

[24] N. Th. Varopoulos, Hardy-Littlewood theory for semigroups, *J. Func. Anal.* 63(1985), 240-260

[25] Z. Vondraček, An estimate for the L^2-norm of a quasi continuous function with respect to a smooth measure, *Arch. Math.* 67(1996), 408-414

Faculty of Engeneering
Kansai University
Suita, Osaka 564-8680
Japan

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 91–109

Gauge Theorems for Stieltjes Exponentials

Ronald Getoor

Abstract.

A gauge theorem for the Stieltjes exponential of a right continuous additive functional satisfying a general Kato type condition is established. Results for the ordinary exponential are then obtained as corollaries.

§1. Introduction

In a series of recent papers Chen and Song [CS02], [CS03] and Chen [C02] have made remarkable progress in establishing the gauge and conditional gauge theorem under quite general hypotheses. The gauge theorem has the following structure: Given a multiplicative function $M = (M_t)$ and a terminal time τ of a strong Markov process X the gauge function $g(x) := E^x(M_\tau)$ is bounded on $\{g < \infty\}$ under suitable hypotheses on X and M. Usually X is assumed to satisfy some irreducibility hypothesis which then implies that g is either bounded or identically infinite. Also M usually is of the form $M_t = \exp(A_t)$ where A is an additive functional. See the above cited papers and also the book of Chung and Zhao [CZ95] for some history of the subject. Before the above cited papers A was usually assumed to be continuous and often of the form $A_t = \int_0^t q(X_s)ds$ where q is a function on the state space of X. See however [CR88], [So93] and [St91] for notable exceptions.

In their papers Chen and Song and Chen consider both continuous and a class of discontinuous additive functionals. The arguments in the two cases are similar in structure but somewhat different in detail. It turns out that by modifying slightly their approach one can prove a gauge theorem for arbitrary right continuous additive functionals in a unified way, assuming only that the underlying process X is a Borel right

Received January 27, 2003.
Keywords and Phrases: gauge theorem, additive functional, Stieltjes exponential.
Mathematics Subject Classification (2000): Primary 60J40, Secondary 60555, 60557

process. The present paper is devoted to spelling this out in some detail. We obtain a slightly sharper result even in the case considered by Chen and Song, but our emphasis is the generality and the simplicity of the result obtained. We work directly with Stieltjes exponentials of additive functionals obtaining a gauge theorem for such exponentials. Results for ordinary exponentials then appear as simple corollaries. Our results are general enough to apply to infinite dimensional processes since we do not assume any absolute continuity condition.

One of the main purposes of the Chen, Song papers was to prove a gauge theorem in enough generality that it could be applied to prove a conditional gauge theorem. A critical hypothesis for their result for a conditional gauge theorem is that X be in strong duality with another Borel right process X. Under this duality hypothesis, the hypotheses and argument in section 3 of [CS00] can be adapted to prove a conditional gauge result for Stieltjes exponentials along the lines of this paper. We leave the precise formulation to the interested reader.

We close this introduction with some words on notation. If $(F, \mathcal{F}, \mu)$ is a measure space, then we use $\mathcal{F}$ also to denote the class of all $\overline{\mathbb{R}} = [-\infty, \infty]$ valued $\mathcal{F}$ measurable functions. If $\mathcal{M} \subset \mathcal{F}$, then $b\mathcal{M}$(resp.$p\mathcal{M}$) denotes the class of bounded (resp. $[0, \infty]$ valued) functions in $\mathcal{M}$. For $f \in \mathcal{F}, \mu(f)$ denotes the integral $\int f d\mu$. If $(E, \mathcal{E})$ is a second measurable space and $K = K(x, dy)$ is a kernel from $(F, \mathcal{F})$ to $(E, \mathcal{E})$ (i.e. $x \to K(x, A)$ is in $\mathcal{F}$ for each $A \in \mathcal{E}$ and $K(x, \cdot)$ is a measure on $(E, \mathcal{E})$ for each $x \in F$), then we write μK for the measure $A \to \int \mu(dx) K(x, A)$ and Kf for the function $x \to \int K(x, dy) f(y)$. The symbol ":=" stands for "is defined to be".

§2. Preliminaries

Throughout this paper $(P_t, t \geq 0)$ will denote a Borel right semigroup on a Lusin state space $(E, \mathcal{E})$, and $X = (X_t, P^x)$ will denote the canonical realization of (P_t) as a right continuous strong Markov process. A (positive) σ-finite measure m on $(E, \mathcal{E})$ is *excessive* provided $mP_t \leq m$ for all $t \geq 0$. Since (P_t) is a right semigroup, it follows that $mP_t \uparrow m$ setwise as $t \downarrow 0$. See [DM87; XII, 36-37]. We fix an excessive measure m to serve as a background measure. In general we shall use the standard notation for Markov processes without special mention. See, for example, [BG68], [DM87], [S88] and [G90]. In particular, $U^q := \int_0^\infty e^{-qt} P_t dt, q \geq 0$, denotes the resolvent and $U := U^0$ the potential kernel of (P_t) or X. We assume only that (P_t) is sub-Markovian and so a point Δ is adjoined to E as an isolated point to serve as a cemetary and $\zeta := \inf\{t : X_t = \Delta\}$ is the lifetime of X and $P^x(\zeta > t) = P_t 1(x)$.

As usual a function f on E is extended to Δ by $f(\Delta) = 0$ unless explicitly stated otherwise. Thus for example, if $f \geq 0$

$$U^q f(x) = E^x \int_0^\zeta e^{-qt} f(X_t) dt = E^x \int_0^\infty e^{-qt} f(X_t) dt.$$

We shall assume throughout that X is transient or equivalently that U is proper. More precisely we shall assume:

(2.1) Transience Assumption. *There exists a function $b \in \mathcal{E}, 0 < b \leq 1$ with $Ub \leq 1$; reducing b if necessary we may also suppose that $m(b) < \infty$.*

Recall that a set $B \in \mathcal{E}^n$ is m-polar (resp. m-semipolar) provided $\{t : X_t \in B\}$ is empty (resp. at most countable) P^m a.s. Here $\mathcal{E}^n$ denotes the σ-algebra of nearly Borel sets. A nearly Borel set N is m-**inessential** provided it is m-polar and $E \backslash N$ is absorbing for X. By [GS 84; (6.12)] any m-polar set is contained in a Borel m-inessential set. A property or statement $P(x)$ is said to hold quasi-everywhere (q.e.) or for quasi-every x provided it holds for all x outside some m-polar set. The exceptional set may then be supposed to be m-inessential. We also write a.e. for m-a.e.

(2.2) Definition. *A positive additive functional (PAF) is an $(\mathcal{F}_t)$-adapted increasing process $A = (A_t; t \geq 0)$ with values in $[0, \infty]$, for which there exist a defining set $\Omega_A \in \mathcal{F}$ and a Borel m-inessential set N_A (called an exceptional set for A) such that*

- *(i) $P^x(\Omega_A) = 1$ for all $x \notin N_A$;*
- *(ii) $\theta_t \Omega_A \subset \Omega_A$ for all $t \geq 0$;*
- *(iii) For all $\omega \in \Omega_A$ the mapping $t \to A_t(\omega)$ is right continuous on $[0, \infty[$, finite valued on $[0, \zeta(\omega)[$ with $A_0(\omega) = 0$;*
- *(iv) For all $\omega \in \Omega_A$ and $s, t \geq 0, A_{s+t}(\omega) = A_t(\omega) + A_s(\theta_t \omega)$;*
- *(v) For all $t \geq 0, A_t([\Delta]) = 0$ where $[\Delta]$ is the dead path identically equal to Δ.*
- *(vi) For $\omega \in \Omega_A, A_t(\omega) = A_{\zeta(\omega)-}(\omega)$ for $t \geq \zeta(\omega)$.*

We let $\widetilde{\mathcal{A}}^+$ denote the class of all PAFs. If in (2.2-iii) right continuous is replaced by continuous, then A is a positive continuous additive functional (PCAF) and we write $\widetilde{\mathcal{A}}_c^+$ for the class of such functionals. Two PAFs A and B are m-equivalent provided $P^m(A_t \neq B_t) = 0$ for all $t \geq 0$. One can check that A and B are m-equivalent if and only if they have a common defining set Λ and a common exceptional set N such that $A_t(\omega) = B_t(\omega)$ for all $t \geq 0$ and $\omega \in \Lambda$. See the argument just below Definition 3.1 in [FG 96]. Equality between elements of $\widetilde{\mathcal{A}}^+$ will mean m-equivalence unless explicitly mentioned otherwise. An $A \in \widetilde{\mathcal{A}}^+$ may be decomposed as $A = A^c + A^d$ where A^c is a PCAF and $A_t^d = \sum_{0<s\leq t} \Delta A_s$ is the sum of the jumps of $A, \Delta A_s = A_s - A_{s-}$. Because of (vi), $\Delta A_\zeta = 0$ on Ω_A. Of course $\Omega_A = \Omega$ and $N_A = \phi$ are allowed in Definition 2.2. By restricting X to the Borel absorbing set $E \backslash N_A$ one may usually reduce to the case in which N_A is empty. In order to keep this exposition as simple possible I will restrict attention to this case. To be precise we define

$$\mathcal{A}^+ = \{A \in \widetilde{\mathcal{A}}^+ : N_A = \phi\}. \tag{2.3}$$

Our results will be stated for $A \in \mathcal{A}^+$, but the interested reader should have no difficulty in formulating them for $A \in \widetilde{\mathcal{A}}^+$. If $A \in \mathcal{A}^+$, then $P^x(\Omega_A) = 1$ for all x. Hence A is almost perfect as defined in [S88, p.173].

If $A \in \mathcal{A}^+$, then its characteristic (Revuz) measure μ_A is defined by

$$\mu_A(f) :=\uparrow \lim_{t\to 0} \frac{1}{t} E^m \int_0^t f(X_s)dA_s \tag{2.4}$$

for $f \in p\mathcal{E}^n$. Moreover $\mu_A(f) =\uparrow \lim_{q\to\infty} q \cdot mU_A^q f$ where

$$U_A^q f(x) := E^x \int_0^\zeta e^{-qt} f(X_t)dA_t \tag{2.5}$$

is the q-potential operator of A. As usual $U_A := U_A^0$. Clearly μ_A does not charge m-polars.

We shall also need the Stieltjes exponential $\mathrm{Exp}(A)$ of $A \in A^+$:

$$\mathrm{Exp}(A_t) := e^{A_t^c} \prod_{0<s\leq t} (1 + \Delta A_s). \tag{2.6}$$

Clearly $\mathrm{Exp}(A)$ is increasing, right continuous and finite on $\{(t,\omega) : 0 \leq t < \zeta(\omega), \omega \in \Omega_A\}$ and $\mathrm{Exp}(A_0) = 1$. On Ω_A the compensated powers $A^{(n)}$ of A are defined by $A_t^{(0)} = 1$ and for $t < \zeta$

$$A_t^{(n)} = n \int_{]0,t]} A_{s-}^{(n-1)} dA_s; \quad n \geq 1.$$

It is well-known that $\mathrm{Exp}(A_t) = \sum_{n\geq 0} \frac{1}{n!} A_t^{(n)}$, see [DD70, p189]. Recall that a terminal time τ for X is a stopping time which satisfies $t + \tau \circ \theta_t = \tau$ on $\{t < \tau\}$. A straightforward induction argument yields the following result. See [SS00] for much more general results.

(2.7) Lemma. *Let $A \in \mathcal{A}^+$ and let σ be either a terminal time or a constant time. Let $\tau = \sigma \wedge \zeta$. If $E^x(A_{\tau-}) \le C < 1$ for all x, then $E^x[\mathrm{Exp}(A_{\tau-})] \le (1-C)^{-1}$ for all x.*

If $B \in \mathcal{E}^n, \tau(B) := \inf\{t > 0 : X_t \notin B]$ denotes the exit time from B. Evidently $\tau(B) \le \zeta$.

(2.8) Proposition. *Let $A \in \mathcal{A}^+$ and suppose that A has bounded jumps; that is, there exists $0 \le c < \infty$ such that $\sup_t \Delta A_t \le c$ a.s. Then there exists an increasing sequence (G_n) of finely open nearly Borel sets with $E = \cup G_n, \tau(G_n) \uparrow \zeta$ a.s. and for each $n, \mu_A(G_n) < \infty$ and $U_A 1_{G_n} \le (2c+1)n$.*

Proof. Suppose first that $c < 1$. Define

$$M_t := \mathrm{Exp}(-A_t) = e^{-A_t^c} \prod_{0<s\le t} (1 - \Delta A_s).$$

Then a.s., $t \to M_t$ is right continuous and decreasing, $M_t > 0$ if $t < \zeta$ and $M_0 = M_{0+} = 1$. It is well-known and easily verified that $M_{t+s} = M_t M_s \circ \theta_t$ and $d(M_t^{-1}) = M_t^{-1} dA_t$. Define

$$g(x) := E^x \int_0^\zeta M_t b(X_t) dt \tag{2.9}$$

where b is the function in (2.1). Clearly $g > 0$, and

$$\begin{aligned} U_A g &= E^{\cdot} \int_0^\zeta g(X_t) dA_t = E^{\cdot} \int_0^\zeta M_t^{-1} \int_t^\zeta M_s b(X_s) ds\, dA_t \\ &= E^{\cdot} \int_0^\zeta M_s b(X_s) \int_0^s M_t^{-1} dA_t ds = Ub - g. \end{aligned}$$

Hence $0 < g \le Ub$ and $U_A g \le Ub$. It is easily checked that g is excessive relative to (X, M) – the M subprocess of X– and consequently g is nearly Borel and finely continuous. Thus the sets $G_n := \{g > \frac{1}{n}\}$ form an increasing sequence of finely open nearly Borel subsets of E with $\cup G_n = E$. Let τ_n be the exit time from G_n. Since G_n^c is finely closed, $g(X_{\tau_n}) \le \frac{1}{n}$, a.s. on $\{\tau_n < \zeta\}$. Hence

$$\frac{1}{n} \ge E^{\cdot}[g(X_{\tau_n})] = E^{\cdot}[M_{\tau_n}^{-1} \int_{\tau_n}^\zeta M_t b(X_t) dt] \ge E^{\cdot}\left[\int_{\tau_n}^\zeta M_t b(X_t) dt\right].$$

But $b > 0$ and $M_t > 0$ on $[0, \zeta[$ a.s., so we must have $\lim_n \tau_n = \zeta$ a.s. Since $U_A g \le Ub$ one has $U_A 1_{G_n} \le nUb \le n$. Therefore $\mu_A(G_n) \le n\mu_A Ub \le n \cdot m(b) < \infty$ where the second inequality comes from (the proof of) the lemma at the bottom of page 508 in [Re 70]. This establishes (2.8) when $c < 1$. If $c \geq 1$, let $A_t^* = (2c)^{-1}A_t$ so that A^* has jumps bounded by $\frac{1}{2}$. Let g be defined as in (2.9) but with A replaced by A^*. Since $U_A = 2cU_{A^*}$ and $\mu_A = 2c\mu_{A^*}, G_n := \{g > \frac{1}{n}\}$ has the desired properties. ■

The proof of (2.8) is easily modified to prove the following:

(2.10) Proposition. *Let $A \in \widetilde{\mathcal{A}}^+$ and suppose that $\sup_t \Delta A_t \leq c < \infty$ a.s. P^x for $x \in E\backslash N_A$. Then there exists an increasing sequence (G_n) of finely open nearly Borel sets such that $E\backslash \cup G_n$ is m-polar, $\tau(G_n) \uparrow \zeta$ a.s. P^x for $x \in E\backslash N_A$, and for each $n, \mu_A(G_n) < \infty$ and $U_A 1_{G_n} \leq (2c+1)n$ on $E\backslash N_A$.*

From time to time we will spell out the situation when $\widetilde{\mathcal{A}}^+$ replaces $\mathcal{A}^+$. Usually it is just a matter of keeping track of the exceptional set N_A as illustrated in (2.10).

§3. Kato Classes

In this section we introduce some Kato classes of additive functionals. The definitions are modifications of those in [C02]. Let $\|\cdot\|_\infty$ denote the norm in $L^\infty(m)$; that is for $f \in \mathcal{E}^n$, $\|f\|_\infty = m\text{-}\operatorname{ess\,sup}_x |f(x)|$. We shall also use the q.e. supremum norm for $f \in \mathcal{E}^n$; that is $\|f\|_{qe} = \inf\{\beta : |f| \leq \beta \, q.e.\}$. Clearly $\|f\|_{qe} \geq \|f\|_\infty$. Recall that $f \in \mathcal{E}^n$ is quasi-finely continuous (qfc) provided it is finely continuous off an m-polar set which may be assumed to be m-inessential. Since an m-null finely open set is m-polar, it follows that if f is qfc, then $\|f\|_\infty = \|f\|_{qe}$.

(3.1) Definition. *Let $0 < \beta < \infty$. Then $\widetilde{K}_\beta$ (resp. $\widetilde{K}_\beta^*$) consists of those $A \in \widetilde{\mathcal{A}}^+$ that have bounded jumps as defined in (2.10) and such that there exist a positive measure ν on E, a set $K \in \mathcal{E}$ with $\nu(K) < \infty$ and $\delta = \delta(\nu, K) > 0$ with the following property:*

(3.2) *If $B \subset \mathcal{E}^n$ with $B \subset K$ and $\nu(B) < \delta$, then*
$$\|U_A 1_{B \cup K^c}\|_\infty \leq \beta \text{ (resp. } \|E^\cdot(A_{\tau(B\cup K^c)-}\|_{qe} \leq \beta).$$

Remarks. For and $D \in \mathcal{E}^n, U_A 1_D$ is excessive for X restricted to $E\backslash N_A$ and hence qfc. But $E^\cdot(A_{\tau(D)-}) \leq E^\cdot \int 1_D(X_t)dA_t = U_A 1_D$ and therefore $\widetilde{K}_\beta \subset \widetilde{K}_\beta^*$. Replacing ν by $\nu|_K$ one may suppose that ν is finite when convenient.

Once again to keep the exposition simple we are going to eliminate the exceptional sets.

(3.3) Definition. *If $0 < \beta < \infty$, then $A \in K_\beta$ (resp. K^*_β) provided $A \in \mathcal{A}^+$ and has bounded jumps as in (2.8) and there exist ν, K and δ as in (3.1) such that $B \subset K, \nu(B) < \delta$ imply*

$$\sup_{x \in E} U_A 1_{B \cup K^c}(x) \leq \beta \ (\textit{resp.} \sup_{x \in E} E^x(A_{T(B \cup K^c)-}) \leq \beta).$$

We are going to work under (3.3) and leave the straightforward extension to the more general situation to the interested reader. Obviously if $0 < \beta < \gamma$, then $K_\beta \subset K_\gamma$ and $K^*_\beta \subset K^*_\gamma$. It is convenient to define

$$K_0 := \cap_{\beta > 0} K_\beta \quad \text{and} \quad K^*_0 := \cap_{\beta > 0} K^*_\beta. \tag{3.4}$$

It will turn out that the K^*_β are the appropriate classes for the gauge theorem. Moreover in an important special case a sufficient condition that $A \in K^*_\beta$ is that $A^p \in K^*_\beta$ where A^p is the dual predictable projection of A. We now describe this result. Let $A \in \mathcal{A}^+$ have bounded jumps. Then there exists a unique predictable element $A^p \in \mathcal{A}^+$– the dual predictable projection of A– such that for any positive predictable process (Z_t)

$$E^x \int_0^\infty Z_t \, dA_t = E^x \int_0^\infty Z_t \, dA^p_t. \tag{3.5}$$

See [S88, §31].

(3.6) Proposition. *Let $A \in \mathcal{A}^+$ have bounded jumps with bound c as in (2.8). If the dual predictable projection A^p of A is continuous, $A^p \in K^*_\beta$ if and only if $A \in K^*_\beta$.*

Proof. If T is a stopping time, $1_{]0,T]}(t)$ is predictable. Therefore since A^p is continuous

$$E^x(A_T) - c \leq E^x(A_{T-}) \leq E^x(A_T) = E^x(A^p_T) = E^x(A^p_{T-}),$$

and this establishes (3.6). ∎

The next two propositions are taken from [C02]. We give proofs for the convenience of the reader. For $A \in \mathcal{A}^+, u_A := U_A 1 = E^\cdot(A_{\zeta-})$ is the potential (function) of A.

(3.7) Proposition. *Suppose $A \in K_\beta, \beta > 0$. Then $\sup_{x \in E} u_A(x) < \infty$.*

Proof. Let $\nu, K \in \mathcal{E}$ and δ be as in Definition 3.3. Then K contains at most a finite number of points $\{x_1, \dots, x_n\}$ with $\nu(\{x_j\}) > \delta$. It follows from a result of Saks (see [DS58, p308]) that $K \backslash \{x_1, \dots x_n\}$ can be written as the disjoint union of a finite number $B_1, \dots, B_k$ of sets in $\mathcal{E}$ with $\nu(B_j) \le \delta$. From (2.8) there exists (G_j) with $G_j \uparrow E$ and $U_A 1_{G_j} \le (2j+1)c$ for each j. Let $F = \{x_1, \dots, x_n\}$. Since $E = \cup G_j$, there exists an ℓ with $1_F \le 1_{G_\ell}$. Hence

$$u_A = U_A 1_{K^c} + U_A 1_F + \sum_{j=1}^{k} U_A 1_{B_j} \le (k+1)\beta + (2\ell+1)c. \qquad \blacksquare$$

The definition of K_β depends on what appears to be an arbitrary choice of the measure ν. The next proposition gives an intrinsic criterion for A to be in K_β, at least when m is a reference measure. If $f: E \to \overline{\mathbb{R}}$, let $\|f\| = \sup_{x \in E} |f(x)|$.

(3.8) Proposition. *Let $A \in \mathcal{A}^+$ and $\beta > 0$. (i) If $A \in K_\beta$, then for every decreasing sequence $(D_n) \subset \mathcal{E}^n$ with $\cap D_n = \phi, \lim_n \|U_A 1_{D_n}\| \le \beta$. (ii) If m is a reference measure and if for every decreasing sequence $(D_n) \subset \mathcal{E}^n$ with $\cap D_n = \phi, \lim_n \|U_A 1_{D_n}\| < \frac{\beta}{2}$ and A has bounded jumps, then $A \in K_\beta$.*

Proof. (i) Suppose $\beta > 0$ and $A \in K_\beta$. Let $D_n \downarrow \phi$. Then there exists an N such that $\nu(K \cap D_n) \le \delta$ for $n \ge N$. Thus for $n \ge N, U_A 1_{D_n} \le U_A 1_{(D_n \cap K) \cup K^c}$ and so $\|U_A 1_{D_n}\| \le \beta$ for $n \ge N$. (ii) Let $A \in \mathcal{A}^+$ have bounded jumps. By (2.8) there exists an increasing sequence (G_n) of finely open sets in $\mathcal{E}^n$ with $\mu_A(G_n) < \infty$ and $\cup G_n = E$. Let $D_n = E \backslash G_n$. Then $\cap D_n = \phi$ and so $\lim_n \|U_A 1_{D_n}\| < \frac{\beta}{2}$. Fix an n with $\|U_A 1_{D_n}\| < \frac{\beta}{2}$ and put $K = G_n$. We claim that there exists a $\delta > 0$ such that if $B \subset K$ and $\mu_A(B) < \delta$, then $\|U_A 1_{B \cup K^c}\| \le \beta$. Suppose no such $\delta > 0$ exists. Then for each n there exists $B_n \subset K$ with $\mu_A(B_n) \le 2^{-n-1}$ and $\|U_A 1_{B_n \cup K^c}\| > \beta$. Let $F_n = \cup_{k \ge n} B_k$. Then (F_n) is a decreasing sequence with $\mu_A(F_n) \le 2^{-n}$. If $F := \cap F_n$, then $\mu_A(F) = 0$. Hence $0 = \mu_A(F) = \uparrow \lim_{q \to \infty} q m U_A^q 1_F$ and so $m U_A^q 1_F = 0$ for $q > 0$. Letting $q \downarrow 0$ we see that $U_A 1_F = 0$ a.e. m and thus everywhere since $U_A 1_F$ is excessive and m is a reference measure. Consequently $U_A 1_{F_n} = U_A 1_{F_n \backslash F}$ and since $F_n \backslash F \downarrow \phi, \lim_n \|U_A 1_{F_n}\| < \frac{\beta}{2}$. Choose n with $\|U_A 1_{F_n}\| < \frac{\beta}{2}$. Now $B_n \subset F_n$ so

$$\beta < \|U_A 1_{F_n \cup K^c}\| \le \|U_A 1_{F_n}\| + \|U_A 1_{K^c}\| < \beta$$

and this contradiction establishes (3.8). $\blacksquare$

Remarks. We emphasize that the measure constructed in the proof of (3.8) is $\nu = \mu_A$. The only place in the proof that the fact that m is a reference measure is used is concluding that $U_A 1_F = 0$ from $U_A 1_F = 0$ a.e. and of course then q.e. Consequently the proof is easily adapted using (2.10) in place of (2.8) to show:

(3.9) Proposition. *Let $A \in \widetilde{\mathcal{A}}^+$ have bounded jumps as in (2.10). If for every decreasing sequence $(D_n) \subset \mathcal{E}^n$ with $\cap D_n = \phi, \lim_n \|U_A 1_{D_n}\|_\infty < \frac{\beta}{2}$, then $A \in \widetilde{K}_\beta$.*

(3.10) Proposition. *Suppose $A^j \in K_{\beta_j}$ (resp. $K^*_{\beta_j}$) for $j = 1, 2$. Then $A^1 + A^2 \in K_{\beta_1+\beta_2}$ (resp. $K^*_{\beta_1+\beta_2}$).*

Proof. Let ν_j, K_j and δ_j serve for $A^j \in K_{\beta_j}, j = 1, 2$. We may suppose that ν_j is carried by $K_j, j = 1, 2$. Define $\nu = \nu_1 + \nu_2, K = K_1 \cup K_2$ and $\delta = \delta_1 \wedge \delta_2$. Then

$$\begin{aligned} \nu(K) &= \nu(K_1 \cap K_2^c) + \nu(K_1^c \cap K_2) + \nu(K_1 \cap K_2) \\ &\leq 2[\nu_1(K_1) + \nu_2(K_2)] < \infty. \end{aligned}$$

Suppose $B \subset K$ with $\nu(B) < \delta$. Then $\nu_j(B) \leq \delta_j$ for $j = 1, 2$. Note that $B \cup K^c \subset (B \cap K_j) \cup K_j^c$ for $j = 1, 2$ and so

$$U_{A_1+A_2} 1_{B \cup K^c} \leq \sum_{j=1}^{2} U_{A_j} 1_{(B \cap K_j) \cup K_j^c} \leq \beta_1 + \beta_2.$$

Of course $A_1 + A_2$ has bounded jumps. The same argument works when the K_β are replaced by K^*_β. ■

§4. Gauge Theorems

Gauge theorems are usually stated for fluctuating additive functionals. Formally let $\mathcal{A} := \mathcal{A}^+ - \mathcal{A}^-$ and introduce the obvious notion of equality: If $A_j = A_j^+ - A_j^-, A_j^\pm \in \mathcal{A}^+$ for $j = 1, 2$, then $A_1 = A_2$ provided $A_1^+ + A_2^- = A_1^- + A_2^+$ in $\mathcal{A}^+$. Then it is known that $A \in \mathcal{A}$ can be written uniquely as $A = A^+ - A^-$ with $A^+, A^- \in \mathcal{A}^+$ having a common defining set Ω_A and such that the measures $dA_t^+(\omega)$ and $dA_t^-(\omega)$ on $[0, \zeta(\omega)[$ are orthogonal for $\omega \in \Omega_A$. Actually the only thing that we shall use is that a.s., A^+ and A^- have no common discontinuities. Of course $A \in \mathcal{A}$ can be decomposed as $A = A^c + A^d$ where $A^c \in \mathcal{A}$ is continuous and $A^d \in \mathcal{A}$ is purely discontinuous,

$$A_t^d = \sum_{0<s\leq t} \Delta A_s = \sum_{0<s\leq t} \Delta A_s^+ - \sum_{0<s\leq t} \Delta A_s^-$$

with $\sum_{0<s\leq t} |\Delta A_s| < \infty$ if $t < \zeta$. In particular $t \to A_t$ is of bounded variation on compact intervals in $[0, \zeta[$ a.s. Fix $A = A^+ - A^- \in \mathcal{A}$. Define for $t < \zeta$,

$$L_t^+ := \operatorname{Exp}(A_t^+),\ L_t^- := \operatorname{Exp}(A_t^-).$$

It turns out that the appropriate multiplicative functional to consider is

$$(4.1) \qquad L_t := L_t^+/L_t^- = e^{A_t^c} \prod_{0<s\leq t} \frac{(1+\Delta A_s^+)}{(1+\Delta A_s^-)}.$$

Clearly a.s., $t \to L_t^{\pm}$ is increasing and finite on $[0, \zeta[$ and $\Delta A_\zeta^{\pm} = 0$. Hence $L_{\zeta-}^{\pm} = L_\zeta^{\pm}$ and so $L_{\zeta-} = L_\zeta$ where $\infty/\infty = 0$ by convention. Moreover $t \to L_t$ is right continuous on $[0, \zeta[$ and is of bounded variation on compact subintervals of $[0, \zeta[$, a.s. Henceforth we shall omit the qualifier "a.s" in places where it is obviously required such as in the preceding two sentences. Note that $L_0 = L_{0+} = 1$. The function

$$(4.2) \qquad g(x) := E^x[L_\zeta] = E^x[L_{\zeta-}]$$

is called the **gauge** of A (or L).

(4.3) Proposition. *The gauge g is nearly Borel measurable and finely continuous. If $F := \{g < \infty\}$, then F is absorbing.*

Remark. Since g may take the value $+\infty, g$ is finely continuous as a map from E to $[0, \infty]$.

Proof. Since $L_\zeta \in \mathcal{F}, g$ is universally measurable and because L is a multiplicative functional, $L_\zeta \circ \theta_t = L_{\zeta-} \circ \theta_t = L_{\zeta-}/L_t$ if $t < \zeta$. Let $M_t = (L_t^-)^{-1}$. Then M is a decreasing, right continuous multiplicative functional that is strictly positive for $t < \zeta$. Now

$$(4.4) \qquad \begin{aligned} E^x[g(X_t)M_t] &= E^x(M_t L_{\zeta-}/L_t\,; t < \zeta) \\ &= E^x[L_{\zeta-}/L_t^+; t < \zeta] \uparrow g(x) \end{aligned}$$

as $t \downarrow 0$. Therefore g is excessive relative to (X, M)– the M-subprocess of X– and hence g is nearly Borel and finely continuous. In particular $F = \{g < \infty\}$ is finely open and nearly Borel. The computation in (4.4), aside from taking the limit as $t \downarrow 0$, holds with t replaced by a stopping time T. Hence $E^x(M_T g(X_T)) \leq g(x)$ for any stopping time T. Let $D = \{g = \infty\} = E \backslash F$. Then D is finely closed and so $g \circ X_{T(D)} = \infty$ a.s. on $\{T(D) < \zeta\}$ where $T(D) := \inf\{t > 0 : X_t \in D\}$ is the hitting time of D. Hence if $x \in F$

$$\infty > g(x) \geq E^x(M_{T(D)} g \circ X_{T(D)}; T(D) < \zeta)$$

and since $M_{T(D)} > 0$ on $\{T(D) < \zeta\}$, this forces $P^x(T(D) < \zeta) = 0$. Hence F is absorbing. ∎

Of course $g(x) \leq E^x(L_\zeta^+)$ and so we may obtain bounds on g by estimating $E^x(L_\zeta^+)$. Thus in what follows we are going to assume that $A \in \mathcal{A}^+$. For $A \in \mathcal{A}^+$ let $c(A)$ denote the infimum of the c such that $\sup_{t<\zeta} \Delta A_t \leq c$ a.s. Then $A \in \mathcal{A}^+$ has bounded jumps provided $c(A) < \infty$. We come now to the main result of this section. The proof is borrowed from Chen and Song [CS02].

(4.5) Theorem. *Suppose $A \in K_\beta^*$ with $\beta < 1$. Then the gauge $g(x) = E^x(L_\zeta)$ is bounded on $\{g < \infty\}$.*

Proof. Let ν, K and δ be as in (3.3) for $A \in K_\beta^*$. Since $\nu(K) < \infty$ we may choose M large enough that $\nu(K \cap \{M < g < \infty\}) < \delta$. Let $B := K^c \cup \{M < g < \infty\} = K^c \cup (K \cap \{M < g < \infty\})$. Then $E^\cdot(A_{\tau(B)-}) \leq \beta$. Consequently by (2.7)

$$E^\cdot(L_{\tau(B)-}) \leq \gamma := (1-\beta)^{-1} < \infty. \tag{4.6}$$

Fix an x. Then

$$\begin{aligned} g(x) &= E^x[L_{\tau(B)-}; \tau(B) = \zeta] + E^x[L_{\zeta-}; \tau(B) < \zeta] \\ &\leq \gamma + E^x[L_{\tau(B)} g(X_{\tau(B)}); \tau(B) < \zeta]. \end{aligned}$$

Let $F = \{g < \infty\}$ and suppose that $x \in F$. But F is absorbing and so $g(X_t) < \infty$ a.s. P^x on $[0, \zeta[$. Hence $g \circ X_{\tau(B)} \leq M$ on $\{\tau(B) < \zeta\}$ a.s. P^x since g is finely continuous. Therefore $g(x) \leq \gamma + M E^x[L_{\tau(B)}; \tau(B) < \zeta]$ on F. Since a.s., $L_{\tau(B)} \leq (1+c) L_{\tau(B)-}$ where $c = c(A)$, we see that $g \leq \gamma + M\gamma(1+c)$ on $F = \{g < \infty\}$. ∎

(4.7) Remark. If one only assumes that $A \in \widetilde{K}_\beta^*$ with $\beta < 1$, then g is only defined on $E \backslash N_A$. It follows that g is finely continuous on $E \backslash N_A$ and g is bounded on $\{g < \infty\} \cap (E \backslash N_A)$. Under the hypotheses of (4.5), $\{g = \infty\} = \{g > M\}$ where $M = \sup\{g(x) : g(x) < \infty\}$. Thus both $\{g < \infty\}$ and $\{g = \infty\}$ are finely open. Therefore if E cannot be written as the disjoint union of two finely open nearly Borel sets one of which is absorbing, in particular if E is finely connected, then g is either bounded or identically infinite. This is the classical form of a gauge theorem.

(4.8) Corollary. *Let $A \in \mathcal{A}^+$ and define $B_t := A_t^c + \sum_{0<s\leq t} (e^{\Delta A_s} - 1)$. Then $B \in \mathcal{A}^+$. If $B \in K_\beta^*$ for some $\beta < 1$, then $g_A(x) := E^x(e^{A_\zeta})$ is finely continuous and bounded on $\{g_A < \infty\}$.*

Proof. Since $\sum_{0<s\leq t} \Delta A_s < \infty$ on $[0,\zeta[$ on Ω_A it is clear that $B \in \mathcal{A}^+$ and has the same defining set as A. Now (4.8) is evident because $e^{A_t} = \mathrm{Exp}(B_t)$. ∎

Remark. If $A = A^+ - A^- \in \mathcal{A}$ and one defines $B_t^{\pm} := A_t^{\pm c} + \sum_{0<s\leq t} (e^{\Delta A_s^{\pm}} - 1)$, then $B_t := B_t^+ - B_t^- \in \mathcal{A}$ and $e^{A_t} = \mathrm{Exp}(B_t^+)/\mathrm{Exp}(B_t^-)$. Therefore $g_A(x) := E^x(e^{A_\zeta})$ is nearly Borel, finely continuous and $\{g_A < \infty\}$ is absorbing. Here $e^{A_\zeta} = e^{A_\zeta^+} e^{-A_\zeta^-}$ with $\infty \cdot 0 = 0$ as customary.

Suppose $A \in \mathcal{A}^+$ is purely discontinuous ($A = A^d$) and all of its jumps are totally inaccessible. In this case the dual predictable projection A^p of A has an especially nice form which we now describe. Let

$$J := \{(t,\omega) : X_{t-}(\omega) \neq X_t(\omega), X_{t-}(\omega) \in E\} \tag{4.9}$$

be the set of totally inaccessible discontinuities of X. Here X_{t-} denotes the left limit in the Ray topology. A Lévy system (N, H) for X consists of a kernel $N = N(x, dy)$ on E with $N(x, \{x\}) = 0$ and a PCAF, H, with empty exceptional set and bounded one potential such that if $F \in (\mathcal{E}^n \otimes \mathcal{E}^n)$ with $F \geq 0$ and $Z = (Z_t)$ with $Z_t \geq 0$ is predictable, then

$$E^x \sum_{s\in J} Z_s F(X_{s-}, X_s) = E^x \int_0^\infty Z_t NF(X_t) dH_t \tag{4.10}$$

where $NF(x) = \int F(x,y)N(x,dy)$. If $A \in \mathcal{A}^+$ is purely discontinuous with totally inaccessible jumps, then there exists such an F vanishing on the diagonal such that

$$A_t = A_t^F = \sum_{s\in J, s\leq t} F(X_{s-}, X_s). \tag{4.11}$$

See §73 of [S88]. If X is a special standard process and X_{t-}^0 is the left limit in the original topology of E, then X_{t-} and X_{t-}^0 are indistinguishable on $[0,\zeta[$ and so X_{s-} may be replaced by X_{s-}^0 in (4.11) and s is automatically in J when $X_{s-} (= X_{s-}^0) \neq X_s, s < \zeta$. See [S88, (47.10)]. Moreover if $A = A^F$ then

$$(NF * H)_t := \int_0^t NF(X_s) dH_s$$

is the dual optional projection of A^F. Suppose F is bounded. Since $NF * H$ is continuous, (3.6) implies that $A^F \in K_\beta^*$ whenever $NF * H \in K_\beta^*$. In particular if $NF * H \in K_\beta$. The next proposition treats an important special case. It is the case considered in [C02]. It is an immediate consequence of (3.6), (3.10), (4.5) and (4.8).

(4.12) Proposition. *Let $A \in \mathcal{A}^+$ have the form $A = A^c + A^F$ where A^c is continuous and A^F is as in (4.11) with F bounded. If $A^c \in K^*_\beta$ and $NF * H \in K^*_\gamma$ with $\beta + \gamma < 1$, then g is bounded on $\{g < \infty\}$ where g is defined in (4.2). If $G = e^F - 1$ and $NG * H \in K^*_\gamma$ with $\beta + \gamma < 1$, then g_A is bounded on $\{g_A < \infty\}$ where g_A is defined (4.8).*

§5. Additional Conditions for the Gauge to be Bounded

In this section we develop some conditions that are equivalent to the boundedness of the gauge function. We follow a well-trodden path that was originally broken by Chung and Rao [CR88] and explored further by Chen and Song and Chen in their papers. The direct results using the Stieltjes exponential appear to be new. In what follows $A = A^+ - A^- \in \mathcal{A}$ as in the beginning of section 4 and L_t is defined in (4.1). The gauge g is defined in (4.2). We begin with the following proposition which is the part of Lemma 9 in [CR88] and Lemma 7 in [C02] that carries over to the present situation.

(5.1) Proposition. *Suppose that $A^+ \in K^*_\beta$ for $\beta < 1$ with $A^+_\zeta < \infty$ a.s. and that $E^\cdot(A^-_\zeta)$ is bounded. Let $\epsilon > 0$. Define*

$$\tau_n := \inf\{t : A^+_t > n\epsilon\}, \ n \geq 1$$

where as usual the infimum of the empty set is $+\infty$. If the gauge g is bounded, then

$$\lim_n \sup_x E^x[L_{\tau_n}; \tau_n < \zeta] = 0$$

Remark. If $A^+ \in K_\beta$, then $E^\cdot(A^+_\zeta)$ is bounded according to (3.7) and so $A^+_\zeta < \infty$ a.s. in this case.

Proof. Since the proof is the same for all $\epsilon > 0$, we shall give it for $\epsilon = 1$, which is the only case used later. Let K, ν, δ be as in the requirement that $A^+ \in K^*_\beta$. Thus $E^\cdot[A^+_{\tau(B\cup K^c)-}] \leq \beta$ when $B \subset K$ with $\nu(B) < \delta$. Since $A^+_\zeta < \infty$ it follows that $\{\tau_n < \zeta\} \downarrow \phi$. Here and in the remainder of this section we omit the qualifier "a.s." where it is obviously required. Therefore by Egorov's theorem, since $E^x(L_\zeta) < \infty, E^x[L_\zeta; \tau_n < \zeta] \downarrow 0$ almost uniformly on K with respect to ν. Recall $\nu(K) < \infty$. Hence given $\epsilon > 0$ there exists a closed set $D \subset K$ with $\nu(K \backslash D) < \delta$ and an N such that if $n \geq N$, then

$$\sup_{x \in D} E^x[L_\zeta; \tau_n < \zeta] < \epsilon. \tag{5.2}$$

Now $D^c = K^c \cup K\backslash D$ and $\tau(D^c) = T(D) \wedge \zeta$ where $T(D)$ is the hitting time of D. But $A_t^+ = A_{\zeta-}^+$ if $t \geq \zeta$ and so $E^{\cdot}(A_{T(D)-}^+) = E^{\cdot}(A_{\tau(D^c)-}^+) \leq \beta < 1$. Thus by (2.7)

$$(5.3) \qquad E^{\cdot}[L_{T(D)-}^+] \leq (1-\beta)^{-1} < \infty.$$

Define $B_t = A_t^{+c} + \sum_{0<s\leq t} \log(1 + \Delta A_s^+)$. Note that $e^{B_t} = L_t^+$ and so $E^x[\exp(B_{T(D)-})] \leq (1-\beta)^{-1}$. Since A^+, and hence, B has bounded jumps, $\sup_x E^x[\exp(B_{T(D)})] < \infty$. Consequently it follows from Corollary 4.2 in [SS00] that there exists $p > 1$ with $\sup_x E^x[\exp(pB_{T(D)})] < \infty$. Finally $e^{pB_t} = [\operatorname{Exp}(A_t^+)]^p = (L_t^+)^p$ and so

$$(5.4) \qquad \sup_x E^x[(L_{T(D)}^+)^p] = M' < \infty.$$

Let $n > \max(N, c(A^+))$ where $c(A^+)$ is defined in the paragraph above (4.5). Then

$$\begin{aligned} E^x[L_\zeta; \tau_{3n} < \zeta] = E^x[L_\zeta; \tau_n \leq T(D), \tau_{3n} < \zeta] \\ +E^x[L_\zeta; T(D) < \tau_n, \tau_{3n} < \zeta] = I + II. \end{aligned}$$

If $T(D) < \tau_n$, then $T(D) + \tau_n \circ \theta_{T(D)} \leq \tau_{3n}$ since $\sup_s \Delta A_s^+ < n$ and $\tau_n < \tau_{3n}$ on $\{\tau_{3n} < \zeta\}$. Therefore writing $T_D = T(D)$ when convenient,

$$\begin{aligned} (5.5) \qquad II &= E^x[L_\zeta; T(D) < \tau_n < \tau_{3n} < \zeta] \\ &\leq E^x[L_{T(D)} L_\zeta \circ \theta_{T(D)}; T(D) + \tau_n \circ \theta_{T(D)} < \zeta, T(D) < \zeta] \\ &= E^x[L_{T(D)} E^{X(T_D)}[L_\zeta; \tau_n < \zeta]; T(D) < \zeta]. \end{aligned}$$

Noting that $L_\zeta = L_{T(D)} L_\zeta \circ \theta_{T(D)}$ even if $T(D) \geq \zeta$ because $\zeta \circ \theta_{T(D)} = 0$ in that case, and that $L_{T(D)} \leq L_{T(D)}^+$ one obtains from (5.4),

$$I \leq \|g\| E^x[L_{T(D)}^+; \tau_n \leq T(D),] \leq \|g\| E^x[(L_{T(D)}^+)^p]^{1/p} P^x[\tau_n \leq T(D)]^{1/q}$$

where $1/p + 1/q = 1$. Moreover $\{\tau_n \leq T(D)\} \subset \{A_{T(D)}^+ \geq n\}$ and so

$$P^x[\tau_n \leq T(D)] \leq \frac{1}{n}[E^x[A_{T(D)-}^+] + c(A^+)] \leq \frac{1}{n}[\beta + c(A^+)].$$

Consequently I approaches zero uniformly in x as $n \to \infty$. On the otherhand $X(T_D) \in D$ on $\{T(D) < \zeta\}$ since D is closed. Thus from (5.2), (5.3) and (5.5)

$$II \leq \epsilon E^x[L_{T(D)}^+] \leq \epsilon(1-\beta)^{-1}[1 + c(A^+)].$$

Combining these estimates we find that

$$\lim_n \sup_x E^x[L_\zeta; \tau_n < \zeta] = 0. \tag{5.6}$$

Next observe that $L_\zeta \geq (L_\zeta^-)^{-1} \geq e^{-A_\zeta^-}$. Therefore if $c = \sup_x E^x(A_\zeta^-)$, Jensen's inequality implies that $E^x(L_\zeta) \geq e^{-c}$. Hence

$$E^x[L_{\tau_n}; \tau_n < \zeta] \leq e^c E^x[L_{\tau_n} E^{X(\tau_n)}(L_\zeta); \tau_n < \zeta] = e^c E^x[L_\zeta; \tau_n < \zeta]$$

and combining this with (5.6) completes the proof of Proposition 5.1 ∎

We come now to the main result of this section. This should be compared with Corollary 2.16 in [C02].

(5.7) Theorem. *Let A^+ and A^- satisfy the hypotheses of (5.1) and suppose in addition that A^- has bounded jumps. Then the following are equivalent:*

(i) $\sup_x E^x[L_\zeta] < \infty$;
(ii) $\sup_x E^x \int_0^\zeta L_{t-} dA_t^+ < \infty$;
(iii) $\sup_x E^x \int_0^\zeta L_t dA_t^+ < \infty$;
(iv) $\sup_x E^x[\sup_{t<\zeta} L_t] < \infty$;

Proof. Since $L_t = L_{t-}(1 + \Delta A_t^+)(1 + \Delta A_t^-)^{-1} \leq [1 + c(A^+)]L_{t-}$ and $L_{t-} \leq [1 + c(A^-)]L_t$, the equivalence of *(ii)* and *(iii)* is clear. Also $dL_t^+ = L_{t-}^+ dA_t^+$ and so

$$\int_{[0,t[} L_{s-} dA_s^+ = \int_{[0,t[} (L_{s-}^-)^{-1} dL_s^+ \geq (L_{t-}^-)^{-1}[L_{t-}^+ - 1] \geq [L_{t-} - 1]$$

for $t \leq \zeta$. Taking $t = \zeta$, *(ii)* implies *(i)*. Also taking the supremum over $t \in [0, \zeta[$,

$$\sup_{t<\zeta} L_t = \sup_{t<\zeta} L_{t-} \leq 1 + \int_{[0,\zeta[} L_{t-} dA_t^+, \tag{5.8}$$

and so *(ii)* implies *(iv)*. Clearly *(iv)* implies *(i)*. Thus it suffices to show that *(i)* implies *(ii)* to complete the proof of (5.7). Therefore suppose that *(i)* holds. Using (5.1) with $\epsilon = 1$ choose $N > c(A^+)$ such that

$$\lambda := \sup_x E^x[L_{\tau_N}; \tau_N < \zeta] < 1.$$

Define $\tau_N^0 = 0$ and $\tau_N^{k+1} = \tau_N^k + \tau_N \circ \theta_{\tau_N^k}$ for $k \geq 0$. Then using the strong Markov property, $\sup_x E^x[L_{\tau_N^k}; \tau_N^k < \zeta] \leq \lambda^k$. We claim that $\tau_{kN} \leq \tau_N^k$ and hence $\lim_k \tau_N^k \geq \zeta$. This is obvious when $k = 1$. Assume that it holds for a fixed $k \geq 0$. Writing $A^+(t) = A_t^+$ for typographical simplicity we have

$$\tau_N^{k+1} = \tau_N^k + \sup\{t : A^+(\tau_N^k + t) - A^+(\tau_N^k) > N\}.$$

If $\tau_{kN} \leq \tau_N^k, A^+(\tau_N^k) \geq kN$ and so $\tau_N^{k+1} \geq \tau_{(k+1)N}$. This establishes the claim. Now

$$\begin{aligned} E^x \int_0^\zeta L_{t-} dA_t^+ &= \sum_{k\geq 0} E^x \int_{[\tau_N^k, \tau_N^{k+1}[} L_{t-} dA_t^+ ; \tau_N^k < \zeta] \\ &= \sum_{k\geq 0} E^x [L_{\tau_N^k} E^{X(\tau_N^k)} \int_{[0,\tau_N[} L_{t-} dA_t^+ ; \tau_N^k < \zeta]. \end{aligned}$$

But

$$\begin{aligned} E^x \int_{[0,\tau_N[} L_{t-} dA_t^+ &\leq E^x \int_{[0,\tau_N[} L_{t-}^+ dA_t^+ \\ &= E^x [L_{\tau_N -}^+ - 1] \leq E^x e^{A^+(\tau_N -)} \leq e^N, \end{aligned}$$

and combining these estimates we obtain

$$E^x \int_0^\zeta L_{t-} dA_t^+ \leq e^N \sum_{k\geq 0} E^x [L_{\tau_N^k} ; \tau_N^k < \zeta] \leq e^N (1 - \lambda)^{-1}.$$

Hence *(i)* implies *(ii)* establishing (5.7). ∎

(5.9) Remark. Suppose in addition to the hypotheses in (5.7), that E can not be written as the disjoint union of two finely open sets one of which is absorbing as in (4.7). Then the condition (5.7 - *i*) is equivalent to $E^x(L_\zeta) < \infty$ for at least one $x \in E$. But in view of (5.8), the remaining conditions in (5.7) are equivalent to the corresponding condition with $\sup_x$ replaced by for at least one $x \in E$.

We next give a sufficient condition that the gauge g is bounded. The integral condition in the following result should be compared to the conditions in Theorem 5.7.

(5.10) Theorem. *Let $A^+ \in K_\beta$ with $\beta < 1$ and suppose that $\zeta < \infty$ a.s. If $\sup_x E^x \int_0^\zeta L_t dt < \infty$, then g is bounded.*

In the course of the proof we shall need the following lemma.

(5.11) Lemma. *Let $A \in K_\beta$ with $\beta < \infty$. Then $\lim_{t\downarrow 0} \sup_x E^x(A_t) \leq \beta$. If $\beta < 1$, then there exist $C < \infty$ and $\lambda > 0$ such that $\sup_x E^x[L_t] \leq Ce^{t\lambda}$.*

Proof. Proposition 2.3 in [C02] asserts that $\lim_{\alpha\to\infty} \sup_x E^x \int_0^\infty e^{-\alpha t} dA_t \le \beta$. The proof in [C02] works perfectly well for right continuous A. If $\eta > 0$,

$$\begin{aligned} E^x \int_0^\infty e^{-\alpha t} dA_t &= \alpha E^x \int_0^\infty e^{-\alpha t} A_t dt \\ &= \int_0^\infty e^{-t} E^x(A_{t/\alpha}) dt \ge E^x[A_{\eta/\alpha}] e^{-\eta}. \end{aligned}$$

This implies that $\lim_{t\to 0} \sup_x E^x(A_t) \le \beta e^{\eta}$ and letting η fall to zero yields the first assertion in (5.11). If $\beta < 1$, then it follows from (2.7) that there exists $t > 0$ such that $\sup_x E^x(L_t) < \infty$. Since $Q_t f(x) := E^x[f(X_t)L_t]$ defines a semigroup, it is well-known and easily checked that this implies the final assertion in (5.11). ∎

We now turn to the proof of (5.10). By (5.11), there exist $C < \infty$ and $\lambda > 0$ such that $E^x[L_t^+] \le Ce^{\lambda t}$. Since $\zeta < \infty$,

$$\begin{aligned} g(x) &= \sum_{n\ge 0} E^x[L_\zeta; n < \zeta \le n+1] \\ &= \sum_{n\ge 0} E^x[L_n E^{X(n)}[L_\zeta; \zeta \le 1];; n < \zeta] \\ &\le \sum_{n\ge 0} E^x[L_n E^{X(n)}[L_1^+];; n < \zeta] \le Ce^{\lambda} \sum_{n\ge 0} E^x[L_n; n < \zeta]. \end{aligned}$$

If $n \le t < n+1$, then writing $c = Ce^{\lambda}$

$$\begin{aligned} &E^x[L_{n+1}; n+1 < \zeta] \le E^x[L_{n+1}; t < \zeta] \\ &\quad = E^x[L_t E^{X(t)}[L_{n+1-t}]; t < \zeta] \le cE^x[L_t; t < \zeta]. \end{aligned}$$

Consequently

$$\sum_{n\ge 0} E^x[L_n; n < \zeta] \le 1 + cE^x \int_0^\zeta L_t dt,$$

and hence g is bounded. ∎

Remarks. The proof of (5.10) is just the argument on page 831 of [CR88]. Under the hypotheses in the first sentence of (5.10), the proof shows that for x fixed, $E^x \int_0^\zeta L_t dt < \infty$ implies that $g(x) < \infty$. Note that in (5.10) it is not assumed that $E^{\cdot}(A_\zeta^-)$ is bounded. If one assumes in addition that $E^{\cdot}(\zeta)$ and $E^{\cdot}(A_\zeta^-)$ are bounded, then the proof of Theorem 6 in [CR88] may be modified to show that $\sup_x E^x \int_0^\zeta L_{t-} dt < \infty$

is necessary for g to be bounded. This requires showing first that $\lim_{n\to\infty} E^x[L_n; n < \zeta] = 0$ uniformly in x, which may be proved by an argument that is similar to, but simpler than, the proof of (5.1). We leave the details to the interested reader.

References

[BG68] R.M. Blumenthal and R. K. Getoor, *Markov Processes and Potential Theory*, Academic Press, New York, 1968.

[C02] Z-Q. Chen, *Gaugeability and conditional gaugeability*, Trans. Amer. Math. Soc., **354**(2002), 4639-4679.

[CS02] Z-Q. Chen and R. Song, *General gauge and conditional gauge theorems*, Ann. Probab., **30**(2002), 1313-1339.

[CS03] Z-Q. Chen and R. Song, *Conditional gauge theorem for non-local Feynman-Kac transforms*, Probab. Theory Related Fields, **125**(2003), 45-72.

[CR88] K. L. Chung and K. M. Rao, *General gauge theorem for multiplicative functionals*, Trans. Amer. Math. Soc., **306**(1988), 819-836.

[CZ95] K. L. Chung and Z. Zhao, *From Brownian Motion to Schrödinger's equation*, Springer, Berlin, 1995.

[DD70] C. Doléans-Dade, *Quelques applications de la formule de changement de variables pour les semimartingales*, Z. Wahrsch, **16**(1970), 181-194.

[DM87] C. Dellacherie and P. -A. Meyer, *Probabilités et Potentiel, Chapitres XII - XVI*, Hermann, Paris, 1987.

[DS58] N. Dunford and J. T. Schwartz, *Linear Operators: Part I*, Interscience, New York, 1958.

[FG96] P.J. Fitzsimmons and R.K. Getoor, *Smooth measures and continuous additive functionals of right Markov processes.* In Itô's Stochastic Calculus and probability Theory, Springer, Tokyo, 1996, pp.31-49.

[G90] R. K. Getoor, *Excessive Measures*, Birkhäuser, Boston, 1990.

[GS84] R. K. Getoor and M. J. Sharpe, *Naturality, standardness and weak duality*, Z. Wahrsch., **67**(1984), 1-62.

[Re70] D. Revuz, *Measures aassociées aux fonctionnelles additives de Markov I*, Trans. Amer. Math. Soc., **148**(1970), 501-531.

[S88] M. Sharpe, *General Theory of Markov Processes*, Academic Press, Boston, 1988.

[So93] R. Song, *Probabilistic approach to the Dirichlet probem of perturbed stable processes*, Prob. Th. Rel. Fields **95**(1993), 371-389.

[SS00] W. Stummer and K. -Th Sturm, *On exponentials of additive functionals*, Stochastic Process. Appl., **85** (2000), 45-60.

[St91] K.-Th. Sturm, *Gauge theorems for resolvents with application to Markov processes*, Probab. Theory Related Fields, **89**(1991), 387-406.

Department of Mathematics
University of California
La Jolla, CA 92093–0112
U.S.A.
email: rgetoor@ucsd.edu

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 111–119

A Frontier of White Noise Analysis, in line with Itô Calculus

Takeyuki Hida

Dedicated to Professor K. Itô

Abstract.

This note discusses two topics; one is the notion of the multiple Wiener integral and the other is the Lévy-Itô decomposition of a Lévy process.

Both have been taken up by Professor K. Itô showing the significance in stochastic analysis. The extensive development of the stochastic analysis at present largely depends on these discoveries by him.

It seems to be a good time to remind his profound ideas and to discuss some of future directions in probability theory.

AMS Subject Classification: 60H40

§1. Introduction

The White Noise Analysis has extensively developed in the last quarter of a century and now it is the time to be in search of further directions of research which will be still in line with the Itô's original contribution toward stochastic analysis.

Two directions are proposed in this line.

(1) Analysis of white noise functionals parameterized by a contour or a surface, namely some kinds of random fields.
Those fields are assumed to live in the Hilbert space (L^2) of white noise functionals (or of Brownian functionals). In order to analyze those functionals we shall appeal to the variational calculus for the random fields, and there is requested to introduce a suitable class of generalized white noise functionals. A good tool from the stochastic analysis for this purpose is the direct sum decomposition of the Hilbert space (L^2) in terms of

Received May 2, 2003.

the subspaces of *multiple Wiener integrals* (see [2]) or of those involving homogeneous chaos.
We then come to introduce classes of generalized white noise functionals, where we should note that our generalization can be done for each subspace given by the decomposition mentioned above. Representations of those functionals can be given by the so-called S-transform which looks like an infinite dimensional analogue of the Laplace transform.
To discuss a variational calculus for random fields in question some more probabilistic interpretation is necessary; however the decomposition of the basic Hilbert space is the milestone of the further study of random fields formed by white noise functionals.

(2) Reduction of random complex systems.
Given a random, evolutional complex system, we first try to obtain the *innovation* of the system. Under some reasonable conditions, we may assume that the innovation comes from a Lévy process. Some concrete results on this topic can be seen in [5].
Then, we are naturally led to the decomposition of the Lévy process established by Itô (see [1]). The decomposition itself was obtained earlier, by taking a relaxed view of rigor. When nonlinear functionals of innovation are considered, finer and rigorous results on the decomposition are necessary, so that the results in [1] are quite helpful. This fact can be seen as soon as we come to the analysis of nonlinear functionals of Poisson process (which is an elemental process) or of compound Poisson noise. It is noted that interesting results are obtained by viewing dissimilarity to the Gaussian case.

§2. Topic (1)

2.1. Decomposition of the white noise space

Let μ be the standard white noise measure on a space E^* of generalized functions on $\mathbb{R}$, and let $L^2(E^*, \mu) = (L^2)$ be the space of white noise functionals. One of the basic tools for the analysis on (L^2) is the direct sum decomposition due to Itô and Wiener:

$$(L^2) = \oplus_{n \geq 0} H_n,$$

where H_n is the space of multiple Wiener integrals or homogeneous chaos of degree n.

Another tool is the so-called S-transform (see e.g. I. Kubo and S. Takenaka [8] or [4, Chapt. 2]) defined by

$$(S\varphi)(\xi) = \exp[-\frac{1}{2}\|\xi\|^2] \int_{E^*} \exp[\langle x, \xi\rangle]\varphi(x) d\mu(x),$$

where φ is an (L^2)-functional. The S-transform gives a nice representation of white noise functionals. Indeed, the transform plays fundamental roles in two ways; one is that a visualized representation of φ is given, and the other is that it helps to introduce generalized white noise functionals which are very important in advanced white noise analysis.

2.2. The isomorphism

$$H_n \cong L^2(\mathbb{R}^n)^\wedge$$

(up to $\sqrt{n!}$), where $L^2(\mathbb{R}^n)^\wedge$ denotes the subspace of $L^2(\mathbb{R}^n)$ containing only symmetric functions. This isomorphism extends to

$$H_n^- \cong H^{-\frac{n+1}{2}}(\mathbb{R}^n)^\wedge,$$

where $H^m(\mathbb{R}^n)^\wedge$ is the subspace of the Sobolev space of order m involving symmetric functions. Define a space $(L^2)^-$ of generalized functionals by

$$(L^2)^- = \oplus c_n H_n^-,$$

where c_n is a certain sequence of positive numbers such that $c_n \to 0$. For details, we refer to the literature [4, Chapt.3.A]. It is noted that the choice of the $\{c_n\}$ depends on the singularity of functionals to be discussed.

Another space $(S)^*$ of generalized functional is defined by a Gel'fand triple

$$(S) \subset (L^2) \subset (S)^*,$$

which is an infinite dimensional analogue of the triple to define the Schwartz distributions. The Rigorous definition and the properties of $(S)^*$ are found in [4, Chapt.4]. Also see [9, Chapt.4].

The spaces $(L^2)^-$ and $(S)^*$ of generalized white noise functionals are basic concept of the white noise analysis; indeed, the choice of these spaces is one of the advantages of our analysis. By using these spaces $(L^2)^-$ and $(S)^*$, we can carry on the analysis of white noise functionals in sufficiently large class and find significant applications in other fields like quantum dynamics. Actually, to be surprising enough, we can find an interesting application to the non-commutative geometry (e.g. [10]).

§3. Topic (2)

3.1. Elemental processes involved in innovation

We shall discuss the analysis of stochastic processes and random fields. The approach is done in line with

$$\textit{Reduction} \to \textit{Synthesis} \to \textit{Analysis}.$$

The step of the reduction is realized by taking the innovation for the random complex phenomena in question. Then, we are given a (Gaussian) white noise and/or a Poisson noise, both of which are *elemental* generalized stochastic processes with independent values at every time point, as is illustrated just below. The two cases have much similarity in the analysis on L^2-spaces; however dissimilarity is also interesting. Hence, they are discussed separately, except those properties in common.

In the multi-dimensional parameter case, say $\mathbb{R}^d$-parameter in general, innovations should still be generalized stochastic processes having independent values at every point in $\mathbb{R}^d$.

To fix the idea, we first take the one-dimensional parameter case. Then, under mild, and in fact reasonable assumptions, we may assume that innovation is the time derivative of some Lévy process with stationary independent increments. Let it be denoted by $L(t)$. We may assume it has no non-random part. Then, the Lévy-Itô decomposition gives us an expression of the form

$$L(t) = cB(t) + X(t),$$

where c is a constant, $B(t)$ is a Brownian motion, and $X(t)$ is a compound Poisson process which consists of independent Poisson processes with different heights of jumps (see [1]). Thus, we can conclude that a Brownian motion and each Poisson process being a component of the $X(t)$ are all *elemental* Lévy processes.

After the reduction follows the step of synthesis which means the construction of the functionals of (Gaussian) white noise and Poisson noises of different heights of jumps. It is, formally speaking, easy to form a general space of functionals of the innovation, however we need profound background. Actually rigorous interpretation can be seen e.g. in [6].

The collection of S-transforms of (L^2)-functionals $\{\varphi\}$ forms a Reproducing Kernel Hilbert space $\mathbf{F}$ which is isomorphic to (L^2). In short, a φ has a representation in terms of a functional (indeed, non-random functional) of a smooth function ξ. We can therefore appeal to the classical theory of functional analysis in order to carry on the calculus on (L^2) .

3.2. Random fields

Coming to random fields, a certain class of them is introduced. Let $\mathbf{C}$ be a class of manifolds C in $\mathbb{R}^d$:

$$\mathbf{C} = \{C;\ \text{convex},\ C^\infty\text{-manifold},\ \approx\ S^{d-1}\},$$

where $\approx$ means a diffeomorphism.

The topology is introduced to $\mathbf{C}$ by using the Euclidean distance.

Given a random field $X(C) = X(C, x), x \in E^*(\mu)$, indexed by $C \in \mathbf{C}$. The S-transform is now of the form

$$(SX(C))(\xi) = \exp[-\frac{1}{2}\|\xi\|^2] \int_{E^*} \exp[\langle x, \xi\rangle] X(C, x) d\mu(x),$$

and it defines a U-functional:

$$(SX(C))(\xi) = U(C, \xi),\ \xi \in E.$$

Suppose that $X(C)$ is a linear functional of a multiple Wiener integral of degree n. Then, the associated U-functional is expressed in the form

$$U(C, \xi) = \int F(C, \mathbf{u}) \xi(u)^{n\otimes} du^n.$$

Thus, we are ready to apply the classical theory of calculus of variations to the functional $U(C, \xi)$ of variable C (see [5, Chapt.6]).

We then come to the case of a Poisson noise.

The case of functionals of a single Poisson process $P(t)$, we introduce the U-transform following the paper by K. Saitô and A. Tsoi [14]. It is given by the formula

$$(U\varphi)(\xi) = C_P(\xi)^{-1} \int \exp[i < x, \xi >] \varphi(x) d\mu_P(x),$$

where μ_P is the probability distribution of Poisson noise $\dot{P}$, and where $C_P(\xi)$ is the characteristic functional of $\dot{P}$. Namely,

$$C_P(\xi) = \exp[\lambda \int (e^{i\xi(u)} - 1) du].$$

For simplicity, the intensity λ is taken to be 1 in this subsection.

Fact. Under the U-transform, a discrete chaos (in Wiener's sense) φ of degree n has a representation of the form

$$\int \cdots \int F(u)\Pi_1^n(e^{i\xi(u_j)} - 1)du^n,$$

where $u = (u_1, \cdots, u_n)$. (For proof see [14].)

Generalized Poisson noise functionals can be introduced by using this representation. We can further play a similar game for the analysis of them to the Gaussian case.

Generalization to the multi-dimensional parameter case is just straightforward. Also, a random field $X(C)$ of functional of Poisson noise is defined, and its variation can be discussed in a similar manner to the Gaussian case.

3.3. Stochastic variational equation

First note that, as a generalization of the infinitesimal equation, there is given a stochastic variational equation for a random field $X(C)$:

$$\delta X(C) = \Phi(X(C'), C' < C, Y(s), s \in C, C, \delta C),$$

where $\{Y(s), s \in C\}$ is the innovation, and where $C' < C$ denotes that C' lies inside of C.

As was explained before, we apply S-transform in the Gaussian case, and we are given a variational equation for U-functional. When we come back to a random function by using S^{-1}, a difficulty arises. This can be illustrated as follows.

To fix the idea, we consider a Gaussian random field $X(C)$ which has a *causal* representation in terms of white noise. Namely, it is a linear functional of x expressed in the form:

$$X(C) = \int_{(C)} F(C, u)x(u)du^d,$$

where $F(C, u)$ is smooth in (C, u), and where (C) denotes the domain enclosed by the ovaloid C. Then, U-functional can be expressed as

$$U(C, \xi) = \int_{(C)} F(C, u)\xi(u)du^d.$$

The variation of U is easily computed to have

$$\delta U(C,\xi) = \int_C F(C,s)\xi(s)\delta n(s)ds + \int_{(C)} \delta F(C,u)\xi(u)du^d,$$

where the δF is the variation of F in the variable C which is defined in the classical sense, and where ds is the surface element over C.

In what follows, the representation of $X(C)$ in the above form is assumed to be *canonical.* Namely, the sigma-field $\mathbf{B}_C(X)$ generated by $X(C')$ with C' inside of C is equal to the sigma-field $\mathbf{B}_C(x)$ generated by $x(u)$, u being inside of C for every C.

Theorem. *The variation $\delta X(C)$ is given by*

$$\delta X(C) = \int_C F(C,s)x(s)\delta n(s)ds + \int_{(C)} \delta F(C,u)x(u)du^d.$$

The innovation is obtained from the first term of the right side.

Proof. This result may be thought of as an easy consequence of the definition of $\delta X(C,x)$, but not quite. A rigorous proof needs, in addition to the functional analysis, the following considerations on the restriction of the parameter and the choice of C. Hence, what follow in Subsection 3.4 are to be included in the proof. For details, see [5].

3.4. Other variations

Restriction of the parameter.

i) Gaussian case.

Given an $\mathbb{R}^d$ parameter white noise (E^*,μ). For $f \in L^2(\mathbb{R}^d)$ the stochastic bilinear form $\langle x, f\rangle, x \in E^*$, is defined which is subject to a Gaussian distribution $N(0,\|f\|^2)$. Take f to be an indicator function such that

$$f_t(u) = \chi_{I(t)}(u),\ t = (t_1,t_2,\cdots,t_d), I(t) = \Pi_j[0,t_j].$$

Then, $X(I(t)) = \langle x, f_t\rangle$ is a Brownian sheet, that is,

$$E[X(I(t))] = 0, E[X(I(t))\cdot X(I(s))] = \Pi_j(t_j \wedge s_j).$$

Set $t_d = 1$. Then, we are given an $\mathbb{R}^{d-1}$-dimensional parameter Brownian sheet. It is now ready to have an $\mathbb{R}^{d-1}$-dimensional parameter white noise by applying partial derivatives. Similarly, much lower dimensional parameter Brownian sheet and white noise can be derived.

The idea is the same for the restriction of the parameter to a hypersurface C which is a smooth ovaloid.

Proposition. *The restriction of the parameter of white noise can be done with the help of Brownian sheet.*

ii) Poisson case

For the Poisson noise the same trick can be applied as is easily shown. Just remind that the characteristic functional $C_P(\xi)$ of a Poisson noise with $\mathbb{R}^d$-parameter is

$$C_P(\xi) = \exp[\lambda \int (e^{i\xi(t)} - 1)dt^d],$$

where $\lambda > 0$ is the intensity. The associated measure on E^* is denoted by μ_p.

Take ξ to be $I(t)$ as above, and form a stochastic bilinear form $\langle x, f_t \rangle, x \in E^*(\mu_p)$, which is to be a Poisson sheet. A restriction of the parameter to a hyperplane defines a lower dimensional parameter Poisson noise. Also a restriction to a hypersurface is given.

As in the Gaussian case we can state a proposition. Since it is similar, so that we omit.

It is noted that observation of on a Poisson sheet, we can see invariance of Poisson noise under some transformations of the parameter space. This will be reported later.

Choice of the C.

One may ask why the parameter C for a random field should be taken to be a smooth ovaloid (convex, closed, without boundary). For one thing, the deformation of C will be done by members of a known transformation group acting on $\mathbb{R}^d$. For another reason, a white noise parameterized by a point in C should be defined. Indeed, as soon as the variational calculus is applied we are naturally led to define an innovation or white noise with parameter set C, as we have seen in Subsection 3.3. For this purpose a Gel'fand triple of functions spaces on C should be defined. Our assumptions for C guarantees the existence of a Gel'fand triple to be requested.

With the considerations mentioned above, we can deal with variational equations, not only for a Gaussian type field reviewed in this section, but also more general non-Gaussian fields and even in the case where $X(C)$ is a nonlinear functional of Poisson noise.

The above considerations are more significant if functionals of Poisson noise are discussed. For example we take a causal and linear functional of a homogeneous chaos, sample function of which is denoted by

x. It is expressed in the form

$$X(C) = \int_{(C)} F(C,u)x(u)du^d.$$

Note that, the *canonical property* (which is an easy generalization of the notion on canonical property discussed in [3, II [8]]) of the above representation follows easily under the assumption that the kernel $F(C,u)$ never vanishes. Hence, the innovation can immediately be obtained.

References

[1] K. Itô, *On stochastic processes (I)*, Japanese J. of Math. 18 (1942), 261-301.
[2] K. Itô, *Multiple Wiener integrals*, J. Math. Soc. Japan, 3 (1951), 157-169.
[3] L. Accardi, H.-H. Kuo, N. Obata, K. Saitô, Si Si and L. Streit eds, *Selected papers of Takeyuki Hida*. World Sci. Pub. Co. Pte. Ltd, 2001.
[4] T. Hida, H.-H. Kuo, J. Potthoff and L. Streit, *White noise. An infinite dimensional calculus.* Kluwer Acad. Pub. 1993.
[5] T. Hida and Si Si, *An innovation approach to random fields. Application of white noise theory.* World Scientific Pub. Co. Pte. Ltd. 2003.
[6] T. Hida and Si Si, *Lévy processes, revisited.* Levico Conf. 2003.
[7] John R. Klauder and E.C.G. Sudarshan, *Fundamentals of quantum optics*, Benjamin Inc, 1968.
[8] I. Kubo and S. Takenaka, *Calculus on Gaussian white noise.* I,II,III, IV. Proc. Japan Academy. 56 (1980), 376-380, 411-416; 57 (1981) 433-437; 58 (1982), 186-189.
[9] H.-H. Kuo, *White noise distribution theory.* CRC Press, 1996.
[10] R. Léandre and H. Ouerdiane, *Connes-Hida calculus and Bismut-Quillen superconnections.* Les prépublications de l'Institut Élie Cartan. 2003/no. 17.
[11] P. Lévy, *Théorie de l'addition des variables aléatoires.* Gauthier-Villars, 1937.
[12] P. Lévy, *Processus stochastiques et mouvement brownien.* Gauthier-Villars, 1948. 2ème éd. 1965.
[13] P. Lévy, *Problèmes concrets d'analyse fonctionelle.* Gauthier-Villars, 1951.
[14] K. Saitô and A. Tsoi, *The Lévy Laplacian acting on Poisson noise functionals.* Infinite Dimensional Analysis, Quantum Probability and Related Topics, 2 (1999) 503-510.

Meijo University
Nagoya 468-8502, Japan

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 121–140

Integral Representation of Linear Functionals on Vector Lattices and its Application to BV Functions on Wiener Space

Masanori Hino

Abstract.

We consider vector lattices $\mathbb{D}$ generalizing quasi-regular Dirichlet spaces and give a characterization for bounded linear functionals on $\mathbb{D}$ to have a representation by an integral with respect to smooth measures. Applications to BV functions on Wiener space are also discussed.

§1. Introduction

Let X be a compact Hausdorff space and $C(X)$ the Banach space of all continuous functions on X with supremum norm. The Riesz representation theorem says that every bounded linear operator on $C(X)$ is realized by an integral with respect to a certain finite signed measure on X. As a variant of this fact, Fukushima [7] proved that, for any quasi-regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ and for $u \in \mathcal{F}$, $\mathcal{E}(u, \cdot)$ is represented as an integral by a smooth signed measure if and only if

$$|\mathcal{E}(u,v)| \le C_k \|v\|_{L^\infty} \qquad \text{for all } v \in \mathcal{F}_{b,F_k},\ k \in \mathbb{N}$$

holds for some nest $\{F_k\}_{k\in\mathbb{N}}$ and some constants C_k, $k \in \mathbb{N}$. As its applications, he gave a characterization for additive functionals of function type for $(\mathcal{E}, \mathcal{F})$ to be semimartingales, and also proved the smoothness of the measures associated with BV functions ([7, 8, 9]).

In this paper, we show a corresponding result in the framework of vector lattices generalizing quasi-regular Dirichlet spaces. The proof is similar to that in [7] but based on a purely analytical argument, unlike [7] where probabilistic methods are used together. Typical examples

Received February 27, 2003.
Partially supported by Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Encouragement of Young Scientists, 15740089, 2003.

which lie in this framework are first order Sobolev spaces derived from a gradient operator and fractional order Sobolev spaces by a real interpolation method with differentiability index between 0 and 1. Using such results, we can improve the smoothness of the measure associated with BV functions on Wiener space, discussed in [8, 9].

We can find several studies closely related to this article regarding the Riesz representation theorem, e.g., in [22, 14]. Their frameworks are based on Markovian semigroups and the function spaces are derived from their generators, which seems to be suitable for complex interpolation spaces. Ours is based on the lattice property instead and fits for real interpolation spaces.

The organization of this paper is as follows. In Section 2, we give a general framework and preparatory lemmas, which are slight modifications of what have been developed already in the case of Dirichlet spaces or in the framework of the nonlinear potential theory. We also give some examples there. In Section 3, the representation theorems are proved. In Section 4, we discuss some applications to BV functions on Wiener space.

§2. Framework and main results

Let X be a topological space and λ a Borel measure on X. Let $L^0(X)$ be the space of all λ-equivalence classes of real-valued Borel measurable functions on X. We will adopt a standard notation to describe function spaces and their norms, such as $L^p(X)$ (or simply L^p) and $\|\cdot\|_{L^p}$.

We suppose that a subspace $\mathbb{D}$ of $L^0(X)$ equipped with norm $\|\cdot\|_{\mathbb{D}}$ satisfies the following.

(A1) $(\mathbb{D}, \|\cdot\|_{\mathbb{D}})$ is a separable and uniformly convex Banach space.
(A2) (Consistency condition) If a sequence in $\mathbb{D}$ converges to 0 in $\mathbb{D}$, then its certain subsequence converges to 0 λ-a.e.

Since $\mathbb{D}$ is assumed to be uniformly convex, it is reflexive and the Banach-Saks property holds: every bounded sequence in $\mathbb{D}$ has a subsequence whose arithmetic means converge strongly in $\mathbb{D}$ (see [16, 19, 13] for the proof). The following lemma is proved by a standard argument.

Lemma 2.1. *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence bounded in $\mathbb{D}$. Then, there exists a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ such that f_{n_k} converges to some f weakly in $\mathbb{D}$ and the arithmetic means $(1/k)\sum_{j=1}^{k} f_{n_j}$ converge to f strongly in $\mathbb{D}$. Moreover, $\|f\|_{\mathbb{D}} \le \liminf_{k\to\infty} \|f_{n_k}\|_{\mathbb{D}}$. If furthermore f_n converges to some g λ-a.e., then it holds that $g \in \mathbb{D}$ and f_n weakly converges to g in $\mathbb{D}$.*

Proof. By virtue of the Banach-Alaoglu theorem and the reflexivity of $\mathbb{D}$, $\{f_n\}_{n\in\mathbb{N}}$ is weakly relatively compact in $\mathbb{D}$. Using the Banach-Saks property together, we can prove the first claim. The second one follows from the Hahn-Banach theorem. The last one is a consequence of the consistency condition (A2). Q.E.D.

We further assume the following.

(A3) (Normal contraction property) For every $f \in \mathbb{D}$, $\check{f} := 0 \vee f \wedge 1$ belongs to $\mathbb{D}$ and $\|\check{f}\|_{\mathbb{D}} \leq \|f\|_{\mathbb{D}}$.

(A4) For every f in $\mathbb{D}_b := \mathbb{D} \cap L^\infty$, f^2 belongs to $\mathbb{D}$. Moreover, $\sup\{\|f^2\|_{\mathbb{D}} \mid \|f\|_{\mathbb{D}} + \|f\|_{L^\infty} \leq 1\}$ is finite.

The condition (A4) is equivalent to the following:

(A4)' for every f and g in $\mathbb{D}_b$, fg belongs to $\mathbb{D}$. Moreover, $\sup\{\|fg\|_{\mathbb{D}} \mid \|f\|_{\mathbb{D}} + \|f\|_{L^\infty} \leq 1,\ \|g\|_{\mathbb{D}} + \|g\|_{L^\infty} \leq 1\}$ is finite.

Indeed, it is clear that (A4)' implies (A4). To show the converse implication, use the identity $fg = ((f+g)/2)^2 - ((f-g)/2)^2$ and the subadditivity of the norm $\|\cdot\|_{\mathbb{D}}$.

We introduce a sufficient condition for (A3) and (A4).

Lemma 2.2. *Under* (A1) *and* (A2), *the following condition* (C) *implies* (A3) *and* (A4):

(C) *when χ is a bounded and infinitely differentiable function on $\mathbb{R}$ with $\chi(0) = 0$ and $\|\chi'\|_\infty \leq c$ for some $c > 0$, then $\chi \circ v \in \mathbb{D}$ and $\|\chi \circ v\|_{\mathbb{D}} \leq c\|v\|_{\mathbb{D}}$ for every $v \in \mathbb{D}$.*

Proof. To show (A3), apply Lemma 2.1 with a sequence $\{\chi_n \circ v\}_{n\in\mathbb{N}}$ so that $\|\chi_n\|_\infty \leq 1$ and $\|\chi_n'\|_\infty \leq 1$ for every n, and χ_n converges pointwise to $\chi(x) = 0 \vee x \wedge 1$. (A4) is similarly proved. Q.E.D.

For $f \in L^0(X)$, we set $f_+(z) = f(z) \vee 0$ and $f_-(z) = -(f(z) \wedge 0)$.

Lemma 2.3. *Let $f \in \mathbb{D}$. Then $f_+ \in \mathbb{D}$ and $\|f_+\|_{\mathbb{D}} \leq \|f\|_{\mathbb{D}}$.*

Proof. Define $f_n := 0 \vee f \wedge n = n(0 \vee (f/n) \wedge 1)$, $n \in \mathbb{N}$. Then $\|f_n\|_{\mathbb{D}} \leq n\|f/n\|_{\mathbb{D}} = \|f\|_{\mathbb{D}}$ by (A3) and $f_n \to f_+$ pointwise. Lemma 2.1 finishes the proof. Q.E.D.

Lemma 2.4. *For every $f \in \mathbb{D}$, $(-a) \vee f \wedge a \to 0$ weakly in $\mathbb{D}$ as $a \downarrow 0$ and $(-a) \vee f \wedge a \to f$ weakly in $\mathbb{D}$ as $a \to \infty$.*

Proof. It is enough to notice that

$$\|(-a) \vee f \wedge a\|_{\mathbb{D}} = \|0 \vee f \wedge a - 0 \vee (-f) \wedge a\|_{\mathbb{D}} \leq 2\|f\|_{\mathbb{D}},$$
$$(-a) \vee f \wedge a \to 0 \text{ pointwise as } a \downarrow 0,$$
$$(-a) \vee f \wedge a \to f \text{ pointwise as } a \to \infty$$

and to use Lemma 2.1. Q.E.D.

For a measurable set A, we let $\mathbb{D}_A := \{f \in \mathbb{D} \mid f = 0\ \lambda\text{-a.e. on } X \setminus A\}$ and $\mathbb{D}_{b,A} := \mathbb{D}_A \cap L^\infty$. A sequence $\{F_k\}_{k\in\mathbb{N}}$ of increasing sets in X is called a *nest* if each F_k is closed and $\bigcup_{k=1}^\infty \mathbb{D}_{F_k}$ is dense in $\mathbb{D}$. A nest $\{F_k\}_{k\in\mathbb{N}}$ is called (λ-)*regular* if, for all k, any open set O with $\lambda(O \cap F_k) = 0$ satisfies $O \subset X \setminus F_k$. A subset N of X is called *exceptional* if there is a nest $\{F_k\}_{k\in\mathbb{N}}$ such that $N \subset \bigcap_{k=1}^\infty (X \setminus F_k)$. When A is a subset of X, we say that a statement depending on $z \in A$ holds *quasi everywhere* (q.e. in abbreviation) if it does for every $z \in A \setminus N$ for a certain exceptional set N. For a nest $\{F_k\}_{k\in\mathbb{N}}$, we denote by $C(\{F_k\})$ the set of all functions f on X such that f is continuous on each F_k. A function f on X is said to be *quasi-continuous* if there is a nest $\{F_k\}_{k\in\mathbb{N}}$ such that $f \in C(\{F_k\})$. We say that a Borel measure μ on X is *smooth* if it does not charge any exceptional Borel sets and there exists a nest $\{F_k\}_{k\in\mathbb{N}}$ such that $\mu(F_k) < \infty$ for all k. A set function ν on X which is given by $\nu = \nu_1 - \nu_2$ for some smooth measures ν_1 and ν_2 with finite total mass is called a *finite signed smooth measure.* A *signed smooth measure* ν with attached nest $\{F_k\}_{k\in\mathbb{N}}$ is a map from $\mathcal{R} := \{A \subset X \mid A \text{ is a Borel set of some } F_k\}$ to $\mathbb{R}$ such that ν is represented as $\nu(A) = \nu_1(A) - \nu_2(A)$, $A \in \mathcal{R}$, for some smooth Borel measure ν_1 and ν_2 satisfying $\nu_i(F_k) < \infty$ for each $i = 1, 2$ and $k \in \mathbb{N}$. When we want to emphasize the dependency of $\mathbb{D}$, we write $\mathbb{D}$-nest, $\mathbb{D}$-smooth, and so on.

We further assume the following quasi-regularity conditions.

(QR1) There exists a nest consisting of compact sets.

(QR2) There exists a dense subset of $\mathbb{D}$ whose elements have quasi-continuous λ-modifications.

(QR3) There exists a countable subset $\{\varphi_n\}_{n\in\mathbb{N}}$ in $\mathbb{D}$ and an exceptional set N such that each φ_n has a quasi-continuous λ-modification $\tilde{\varphi}_n$ and $\{\tilde{\varphi}_n\}_{n\in\mathbb{N}}$ separates the points of $X \setminus N$.

Every quasi-regular symmetric Dirichlet form $(\mathcal{E}, \mathcal{F})$ satisfies all conditions (A1)–(A4) and (QR1)–(QR3) (and (C)) when letting $\mathbb{D} = \mathcal{F}$ and $\|f\|_{\mathbb{D}} = (\mathcal{E}(f,f) + \|f\|_{L^2}^2)^{1/2}$. We give other examples in the last part of this section.

Lemma 2.5. *There exist some $\rho \in \mathbb{D}$ and some countable subset $\mathcal{C} = \{h_n\}_{n\in\mathbb{N}}$ of $\mathbb{D}_b$ such that $0 \le \rho \le 1$ λ-a.e., $h_n \ge 0$ λ-a.e. for all n, $\mathcal{C} - \mathcal{C} := \{h - \hat{h} \mid h,\ \hat{h} \in \mathcal{C}\}$ is dense in $\mathbb{D}$, and for each $n \in \mathbb{N}$, $\rho \ge c_n h_n$ λ-a.e. for some $c_n \in (0, \infty)$.*

Proof. Let $\{f_m\}_{m\in\mathbb{N}}$ be a countable dense subset of $\mathbb{D}$. Denote by $\mathcal{C} = \{h_n\}_{n\in\mathbb{N}}$ the set of all arithmetic means of finite number of functions in $\{(f_m)_+ \wedge M,\ (f_m)_- \wedge M\}_{m\in\mathbb{N},\, M\in\mathbb{N}}$. Then $\mathcal{C} - \mathcal{C}$ is dense in $\mathbb{D}$ by

Lemma 2.4 and the Banach-Saks property. Define

$$\rho = \sum_{n=1}^{\infty} \frac{h_n}{2^n(\|h_n\|_{\mathbb{D}} + \|h_n\|_{L^\infty} + 1)}.$$

Then ρ and $\mathcal{C}$ satisfy the conditions in the claim. Q.E.D.

We will fix ρ and $\mathcal{C}$ satisfying the statement in the lemma above. Note that we can always take $\rho \equiv 1$ if $1 \in \mathbb{D}$.

Take a strictly increasing and right-continuous function $\xi : [0, \infty) \to [0, \infty)$ with $\xi(0) = 0$. For an open set $O \subset X$, we define

(1) $\quad \mathrm{cap}_\xi(O) = \inf\{\xi(\|f\|_{\mathbb{D}}) \mid f \in \mathbb{D} \text{ and } f \geq \rho \ \lambda\text{-a.e. on } O\}.$

For any subset A of X, we define the capacity of A by

$$\mathrm{cap}_\xi(A) = \inf\{\mathrm{cap}_\xi(O) \mid O \supset A,\ O : \text{open}\}.$$

It should be noted that $\mathrm{cap}_\xi(A) \leq \xi(\|\rho\|_{\mathbb{D}}) < \infty$ for every $A \subset X$.

The following lemma is proved in the same way as in [10].

Lemma 2.6. *For every open set O, there exists a unique function e_O in $\mathbb{D}$ attaining the infimum in* (1). *Moreover,* $0 \leq e_O \leq 1$ λ-a.e.

Proof. The uniqueness follows from the uniform convexity of $\mathbb{D}$. The existence is deduced by the Banach-Saks property and (A2). The last claim is a consequence of (A3). Q.E.D.

We will discuss some basic properties of the capacity.

Lemma 2.7. *Let $\{O_n\}_{n\in\mathbb{N}}$ be a sequence of open sets such that* $\mathrm{cap}_\xi(O_n) \to 0$. *Then, there exists a sequence $\{n_k\} \uparrow \infty$ such that* $e_{O_{n_k}} \to 0$ λ-a.e.

Proof. Since $\|e_{O_n}\|_{\mathbb{D}} \to 0$, the claim is clear from (A2). Q.E.D.

Lemma 2.8. *If* $\mathrm{cap}_\xi(A) = 0$, *then $A \cap \{\rho > 0\}$ is a λ-null set.*

Proof. Take a decreasing open sets $\{O_n\}_{n\in\mathbb{N}}$ such that $A \subset O_n$ and $\mathrm{cap}_\xi(O_n) \to 0$. Since $e_{O_n} \geq \rho$ λ-a.e. on $\bigcap_{k=1}^\infty O_k$, we have $\rho = 0$ λ-a.e. on $\bigcap_{k=1}^\infty O_k$ by virtue of Lemma 2.7. This implies the assertion. Q.E.D.

Lemma 2.9. *Let $\{A_k\}_{k\in\mathbb{N}}$ be a sequence of increasing closed sets. Then the following are equivalent.*

(i) *$\{A_k\}_{k\in\mathbb{N}}$ is a nest.*
(ii) $\lim_{k\to\infty} \mathrm{cap}_\xi(X \setminus A_k) = 0$.

Proof. Suppose (i) holds. Take a sequence $\{f_k\}_{k\in\mathbb{N}}$ in $\mathbb{D}$ such that $f_k \in \mathbb{D}_{A_k}$ and $f_k \to \rho$ in $\mathbb{D}$. Since $\rho - f_k = \rho$ λ-a.e. on $X \setminus A_k$, we have $\mathrm{cap}_\xi(X \setminus A_k) \le \xi(\|\rho - f_k\|_{\mathbb{D}}) \to 0$ as $k \to \infty$.

Next, suppose (ii) holds. It suffices to prove that each $h \in \mathcal{C}$ can be approximated in $\mathbb{D}$ by functions in $\bigcup_k \mathbb{D}_{A_k}$. Since $\rho \ge ch$ λ-a.e. for some $c > 0$, it holds that $e_{X\setminus A_k} \ge ch$ λ-a.e. on $X \setminus A_k$ for each k. Let $f_k = (h - c^{-1}e_{X\setminus A_k})_+ \in \mathbb{D}_{A_k}$. By Lemma 2.7, there exists a sequence $\{k'\}$ diverging to infinity such that $e_{X\setminus A_{k'}} \to 0$ λ-a.e. Therefore, $f_{k'} \to h$ λ-a.e. as $k' \to \infty$. On the other hand, $\|f_k\|_{\mathbb{D}} \le \|h\|_{\mathbb{D}} + c^{-1}\|e_{X\setminus A_k}\|_{\mathbb{D}}$, which is bounded in k. From Lemma 2.1, we can take arithmetic means of some subsequence of $\{f_{k'}\}$, which belong to $\bigcup_k \mathbb{D}_{A_k}$, so that they converge to h in $\mathbb{D}$. Q.E.D.

As is seen from this lemma, any choices of $\mathcal{C}$, ρ and ξ are consistent with the notion of nest. From now on, we treat only the case $\xi(t) = t$ and write cap in place of cap_ξ.

Lemma 2.10. *For any sequence of subsets $\{A_k\}_{k\in\mathbb{N}}$ in X, it follows that* $\mathrm{cap}\left(\bigcup_{k=1}^\infty A_k\right) \le \sum_{k=1}^\infty \mathrm{cap}(A_k)$.

Proof. When $O_1, \dots, O_k$ are open sets, it is easy to see the inequality $\mathrm{cap}\left(\bigcup_{j=1}^k O_j\right) \le \sum_{j=1}^k \mathrm{cap}(O_j)$. Indeed, since $\sum_{j=1}^k e_{O_j} \ge \rho$ λ-a.e. on $\bigcup_{j=1}^k O_j$, we have

$$\mathrm{cap}\left(\bigcup_{j=1}^k O_j\right) \le \left\|\sum_{j=1}^k e_{O_j}\right\|_{\mathbb{D}} \le \sum_{j=1}^k \|e_{O_j}\|_{\mathbb{D}} = \sum_{j=1}^k \mathrm{cap}(O_j).$$

Now, let $\varepsilon > 0$. Take an open set O_k for each $k \in \mathbb{N}$ such that $O_k \supset A_k$ and $\mathrm{cap}(O_k) < \mathrm{cap}(A_k) + \varepsilon 2^{-k}$. Let $U_k = \bigcup_{j=1}^k O_j$. Since $\|e_{U_k}\|_{\mathbb{D}} \le \|\rho\|_{\mathbb{D}} < \infty$, Lemma 2.1 assures the existence of a subsequence $\{e_{U_{k'}}\}$ of $\{e_{U_k}\}$ and $e \in \mathbb{D}$ such that $e_{U_{k'}}$ converges to e weakly in $\mathbb{D}$ and the arithmetic means of $\{e_{U_{k'}}\}$ converge to e in $\mathbb{D}$. Since $e \ge \rho$ λ-a.e. on $\bigcup_{k=1}^\infty O_k$ by using (A2), we have

$$\begin{aligned}\mathrm{cap}\left(\bigcup_{k=1}^\infty A_k\right) &\le \mathrm{cap}\left(\bigcup_{k=1}^\infty O_k\right) \le \|e\|_{\mathbb{D}} \le \liminf_{k'\to\infty} \|e_{U_{k'}}\|_{\mathbb{D}} \\ &= \lim_{k\to\infty} \mathrm{cap}(U_k) \le \lim_{k\to\infty} \sum_{j=1}^k \mathrm{cap}(O_j) \le \varepsilon + \sum_{k=1}^\infty \mathrm{cap}(A_k).\end{aligned}$$

Since ε is arbitrary, we obtain the claim. Q.E.D.

The following series of lemmas are now proved in a standard way as in the case of quasi-regular Dirichlet spaces; see e.g. [11, 15] for the proof.

Lemma 2.11. *Suppose that $f \in \mathbb{D}$ has a quasi-continuous λ-modification $\tilde{f}$. Then, we have* $\mathrm{cap}(\{\tilde{f} > \lambda\}) \le \lambda^{-1}\|f\|_{\mathbb{D}}$ *for each $\lambda > 0$.*

Lemma 2.12. (i) *When $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of quasi-continuous functions, there exists a nest $\{F_k\}_{k\in\mathbb{N}}$ such that $f_n \in C(\{F_k\})$ for every n.*

(ii) *If $f_n \in \mathbb{D}$ has a quasi-continuous λ-modification $\tilde{f}_n$ and converges to f in $\mathbb{D}$ as $n \to \infty$, then f has a quasi-continuous λ-modification $\tilde{f}$ and there exists a sequence $\{n_l\} \uparrow \infty$ and a nest $\{F_k\}_{k\in\mathbb{N}}$ such that every $\tilde{f}_n$ belongs to $C(\{F_k\})$ and $\tilde{f}_{n_l}$ converges to $\tilde{f}$ uniformly on each F_k. In particular, $\tilde{f}_{n_l}$ converges to $\tilde{f}$ q.e.*

(iii) *Every $f \in \mathbb{D}$ has a quasi-continuous λ-modification $\tilde{f}$.*

Lemma 2.13. *There exists a regular nest $\{K_k\}_{k\in\mathbb{N}}$ such that K_k is a separable and metrizable compact space with respect to the relative topology for any k.*

Lemma 2.14. *Suppose that $\{F_k\}_{k\in\mathbb{N}}$ is a regular nest and $f \in C(\{F_k\})$. If $f \ge 0$* λ-a.e. *on an open set O, then $f \ge 0$ on $O \cap \bigcup_k F_k$.*

Lemma 2.15. *If u_1 and u_2 are quasi-continuous functions and $u_1 = u_2$* λ-a.e., *then $u_1 = u_2$ q.e.*

In what follows, $\tilde{f}$ always means a quasi-continuous λ-modification of a function f, a particular version of which is sometimes chosen to suit the context.

We can also prove the next two propositions as in [10] (see also [21, Section 2]) by using Lemma 3.1 below together, though they are not used later in this article.

Proposition 2.16. *For any subset A of X, there exists a unique element e_A in the set $\{f \in \mathbb{D} \mid \tilde{f} \ge \tilde{\rho}$ q.e. on $A\}$ minimizing the norm $\|f\|_{\mathbb{D}}$. Moreover, $0 \le e_A \le 1$* λ-a.e. *and* $\mathrm{cap}(A) = \|e_A\|_{\mathbb{D}}$.

Proposition 2.17. cap *is a Choquet capacity.*

We remark that the assumption (A4) is not necessary so far. The following are our main theorems, which are stated in [7] in the case of quasi-regular Dirichlet spaces.

Theorem 2.18. *Under* (A1)–(A4) *and* (QR1)–(QR3), *for a bounded linear functional T on $\mathbb{D}$, the next two conditions are equivalent.*

(i) *There exist a nest $\{F_k\}_{k\in\mathbb{N}}$ and positive constants $\{C_k\}_{k\in\mathbb{N}}$ such that for each $k \in \mathbb{N}$,*

$$|T(v)| \le C_k \|v\|_{L^\infty(X)} \quad \text{for all } v \in \mathbb{D}_{b,F_k}.$$

(ii) *There exists a signed smooth measure ν with some attached nest $\{F'_k\}_{k\in\mathbb{N}}$ such that*

$$T(v) = \int_X \tilde{v}(z)\,\nu(dz) \quad \text{for all } v \in \bigcup_{k=1}^{\infty} \mathbb{D}_{b,F'_k}.$$

Moreover, the measure ν is uniquely determined.

Theorem 2.19. *Under* (A1)–(A4) *and* (QR1)–(QR3), *for a bounded linear functional T on $\mathbb{D}$ and a positive constant C, the next two conditions are equivalent.*

(i) $|T(v)| \le C\|v\|_{L^\infty(X)}$ *for all* $v \in \mathbb{D}_b$.

(ii) *There exists a finite signed smooth measure ν on X such that the total variation of ν is dominated by C and*

$$T(v) = \int_X \tilde{v}(z)\,\nu(dz) \quad \text{for all } v \in \mathbb{D}_b.$$

In addition, ν is uniquely determined. Moreover, if (C) *in Lemma 2.2 holds, we may replace $\mathbb{D}_b$ in* (i) *by $\mathcal{L}$ that satisfies the following:*

($\mathcal{L}$) *$\mathcal{L}$ is a $\mathbb{D}$-dense subspace of $\mathbb{D}_b$ such that, for each $\varepsilon > 0$, there is a C^∞ function χ on $\mathbb{R}$ with $|\chi| \le 1+\varepsilon$, $0 \le \chi' \le 1$, $\chi(x) = x$ on $[-1,1]$, and $\chi \circ v \in \mathcal{L}$ for every $v \in \mathcal{L}$.*

Before ending this section, we give a few examples of $\mathbb{D}$ other than quasi-regular Dirichlet spaces. Suppose that X is a separable Banach space and H a separable Hilbert space which is continuously and densely imbedded to X. The inner product and the norm of H will be denoted by $\langle\cdot,\cdot\rangle_H$ and $\|\cdot\|_H$, respectively. The topological dual X^* of X is identified with a subspace of H. Let λ be a finite Borel measure on X. When K is a separable Hilbert space, we denote by $L^p(X \to K)$ the L^p space consisting of K-valued functions on the measure space (X,λ).

Define function spaces $\mathcal{F}C_b^1$ and $(\mathcal{F}C_b^1)_{X^*}$ on X by

$$\mathcal{F}C_b^1 = \left\{ u : X \to \mathbb{R} \;\middle|\; \begin{array}{l} u(z) = f(\ell_1(z),\dots,\ell_m(z)), \\ \ell_1,\dots,\ell_m \in X^*,\ f \in C_b^1(\mathbb{R}^m) \\ \text{for some } m \in \mathbb{N} \end{array} \right\},$$

$$(\mathcal{F}C_b^1)_{X^*} = \left\{ G : X \to X^* \;\middle|\; \begin{array}{l} G(z) = \sum_{j=1}^m g_j(z)\ell_j, \\ g_1,\dots,g_m \in \mathcal{F}C_b^1, \\ \ell_1,\dots,\ell_m \in X^* \text{ for some } m \in \mathbb{N} \end{array} \right\},$$

where $C_b^1(\mathbb{R}^m)$ is the set of all bounded functions f on $\mathbb{R}^m$ that have bounded and continuous first-order derivatives. Let $u \in \mathcal{F}C_b^1$ and $\ell \in X^* \subset H \subset X$. We define $\partial_\ell u$ by $\partial_\ell u(z) = \lim_{\varepsilon\to 0}(u(z+\varepsilon\ell) - u(z))/\varepsilon$.

The H-derivative ∇u is a unique map from X to H that satisfies the relation

$$\langle \nabla u(z), \ell \rangle_H = \partial_\ell u(z), \qquad \ell \in X^* \subset H.$$

We assume that, if $u \in \mathcal{F}C_b^1$ and $v \in \mathcal{F}C_b^1$ coincide on a measurable set A, then $\nabla u = \nabla v$ λ-a.e. on A. Let $p \geq 1$. We also assume that $(\nabla, \mathcal{F}C_b^1)$ is closable as a map from L^p to $L^p(X \to H)$. We denote by $W^{1,p}$ the domain of the closure of $(\nabla, \mathcal{F}C_b^1)$ and extend the domain of ∇ to $W^{1,p}$ naturally. The space $W^{1,p}$ is a separable Banach space with norm $\|f\|_{W^{1,p}} = \|f\|_{L^p} + \|\nabla f\|_{L^p(X \to H)}$.

Proposition 2.20. *Suppose also* $p > 1$. *Then, when we regard* $W^{1,p}$ *as* $\mathbb{D}$, *the conditions* (A1)–(A4), (QR1)–(QR3), *and* (C) *are satisfied.*

Proof. From the results of [21], (A1) and (QR1) hold. Since X is separable, X^* is also separable with respect to the weak* topology (see Corollary after Proposition 8 in Chapter IV of [5] for the proof). When $\{\ell_n\}_{n \in \mathbb{N}}$ is a countable dense set of X^*, $\varphi_n(\cdot) = \arctan \ell_n(\cdot)$ and $N = \emptyset$ assure the validity of (QR3). The remaining conditions are easily checked. Q.E.D.

In order to give another example, we introduce real interpolation spaces. Let B_0 and B_1 be separable Banach spaces. We assume that B_0 is continuously imbedded to B_1 for simplicity. For parameters $q \in (1, \infty)$ and $\theta \in (0,1)$, we define the space $(B_0, B_1)_{\theta,q}$ by all elements $f \in B_1$ such that there exist some B_j-valued measurable functions $u_j(t)$ on $[0, \infty)$ $(j = 0, 1)$ satisfying

$$(2) \quad u_0(t) + u_1(t) = f \text{ a.e. } t, \ \int_0^\infty (t^{j-\theta} \|u_j(t)\|_{B_j})^q \, \frac{dt}{t} < \infty \ (j = 0, 1).$$

We set the norm of $f \in (B_0, B_1)_{\theta,q}$ by

$$\|f\|_{(B_0,B_1)_{\theta,q}} = \inf_{u_0, u_1} \left[\max_{j=0,1} \left(\int_0^\infty (t^{j-\theta} \|u_j(t)\|_{B_j})^q \, \frac{dt}{t} \right)^{1/q} \right],$$

where the infimum is taken over all pairs u_0 and u_1 satisfying (2). From the general theory of real interpolation, $(B_0, B_1)_{\theta,q}$ is a Banach space, we have continuous imbeddings $B_0 \hookrightarrow (B_0, B_1)_{\theta,q} \hookrightarrow B_1$, and B_0 is dense in $(B_0, B_1)_{\theta,q}$. Keeping the notation in the previous example, we have the following proposition.

Proposition 2.21. *Let* $p \in (1, \infty)$, $q \in (1, \infty)$, *and* $\theta \in (0, 1)$. *Then* $\mathbb{D} := (W^{1,p}, L^p)_{\theta,q}$ *satisfies* (A1)–(A4), (QR1)–(QR3) *and* (C).

Proof. In general, we can prove that $(B_0, B_1)_{\theta,q}$ is uniformly convex if B_0 or B_1 is, in the same way as Proposition V.1 of [3]. Therefore, $\mathbb{D}$ is uniformly convex. The separability, (QR1) and (QR2) come from those of $W^{1,p}$. (QR3) is proved in the same way as the case of $W^{1,p}$. (A2) is clearly true. We will prove (C). Let χ be as in (C) in Lemma 2.2. Let $f \in \mathbb{D}$ and take u_0 and u_1 satisfying (2). Set $v_0(t) = \chi \circ u_0(t)$ and $v_1(t) = \chi \circ f - \chi \circ u_0(t)$. Then $v_0(t) + v_1(t) = \chi \circ f$ and it is easy to see that $\|v_0(t)\|_{W^{1,p}} \le c\|u_0(t)\|_{W^{1,p}}$ and $\|v_1(t)\|_{L^p} \le c\|u_1(t)\|_{L^p}$. This implies that $\chi \circ f \in \mathbb{D}$ and $\|\chi \circ f\|_{\mathbb{D}} \le c\|f\|_{\mathbb{D}}$. Q.E.D.

§3. Proof of Theorems 2.18 and 2.19

First, we will prove that (ii) implies (i) in Theorem 2.18. We take $F_k = F'_k$ and $C_k = |\nu|(F_k) < \infty$. Let $v \in \mathbb{D}_{b,F_k}$ and $M = \|v\|_{L^\infty(X)}$. We can take a quasi-continuous λ-modification $\tilde{v}$ so that $|\tilde{v}| \le M$ everywhere. Then $|T(v)| \le M|\nu|(F_k) = C_k M$. Therefore, (i) holds.

Next, we will prove that (i) implies (ii) in Theorem 2.18. Take a nest $\{E_k^{(1)}\}_{k\in\mathbb{N}}$ so that $\tilde{\rho} \in C(\{E_k^{(1)}\})$. Define $E_k^{(2)} = E_k^{(1)} \cap \{\tilde{\rho} \ge 1/k\}$.

Lemma 3.1. *$\{E_k^{(2)}\}_{k\in\mathbb{N}}$ is a nest.*

Proof. Clearly, $\{E_k^{(2)}\}_{k\in\mathbb{N}}$ is a sequence of increasing closed sets. Define $\rho_k = \rho \wedge (1/k)$, $k \in \mathbb{N}$. Then $\rho_k \to 0$ weakly in $\mathbb{D}$ by Lemma 2.4. Take a sequence $\{k_j\} \uparrow \infty$ so that $\hat{\rho}_m := (1/m)\sum_{j=1}^m \rho_{k_j}$ converges to 0 in $\mathbb{D}$ as $m \to \infty$. Since $\hat{\rho}_m + e_{X\setminus E_{k_m}^{(1)}} \ge \rho$ λ-a.e. on $X \setminus E_{k_m}^{(2)}$, we have $\mathrm{cap}(X \setminus E_{k_m}^{(2)}) \le \|\hat{\rho}_m\|_{\mathbb{D}} + \|e_{X\setminus E_{k_m}^{(1)}}\|_{\mathbb{D}} \to 0$ as $m \to \infty$. Therefore, $\{E_k^{(2)}\}_{k\in\mathbb{N}}$ is a nest. Q.E.D.

Define $E_k^{(3)} = F_k \cap K_k \cap E_k^{(2)}$, $k \in \mathbb{N}$, where K_k is what appeared in Lemma 2.13. Then $\{E_k^{(3)}\}_{k\in\mathbb{N}}$ is a regular nest consisting of separable and metrizable compact sets. Given $k \in \mathbb{N}$, let $\{U_{k,n}\}_{n\in\mathbb{N}}$ be a countable open basis of $E_k^{(3)}$. The totality of every union of finite elements in $\{U_{k,n}\}_{n\in\mathbb{N}}$ will be denoted by $\{V_{k,n}\}_{n\in\mathbb{N}}$. Take a countable family $\mathcal{O}$ of open sets in X such that for every $k \in \mathbb{N}$, $n \in \mathbb{N}$ and $\varepsilon > 0$, some $O \in \mathcal{O}$ satisfies that $O \supset V_{k,n}$ and $\mathrm{cap}(O) < \mathrm{cap}(V_{k,n}) + \varepsilon$.

Recall the condition (QR3) and set $\mathcal{S} = \{(-M) \vee \tilde{\varphi}_n \wedge M \mid n \in \mathbb{N},\ M \in \mathbb{N}\}$. The functions of $\mathcal{S}$ separate the points of $X \setminus N$. Fix a countable subset $\mathcal{D}$ of $\mathbb{D}$ such that $\{f \in \mathcal{D} \mid \|f\|_{L^\infty(X)} \le M\}$ is a dense set of $\{f \in \mathbb{D} \mid \|f\|_{L^\infty(X)} \le M\}$ for each $M \in \mathbb{N}$. Take a nest $\{\hat{F}_k\}_{k\in\mathbb{N}}$ so that $\hat{F}_k \subset E_k^{(3)}$ for each k, $N \subset \bigcap_{k\in\mathbb{N}}(X \setminus \hat{F}_k)$ and the quasi-continuous

λ-modifications of all elements in $\mathcal{S} \cup \mathcal{D} \cup \{e_O \mid O \in \mathcal{O}\}$ belong to $C(\{\hat{F}_k\})$. Denote by $\mathcal{A}$ the algebra generated by $\mathcal{S} \cup \{1 \wedge M\tilde{\rho} \mid M \in \mathbb{N}\}$. Note that all functions in $\mathcal{A}$ and $\tilde{\rho}$ belong to $C(\{\hat{F}_k\})$. From the Stone-Weierstrass theorem, $\{f|_{\hat{F}_k} \mid f \in \mathcal{A}\}$ is dense in $C(\hat{F}_k)$ with uniform topology for any k.

Lemma 3.2. *There exist a nest $\{F'_k\}_{k\in\mathbb{N}}$ and functions $\{\psi_n\}_{n\in\mathbb{N}}$ in $\bigcup_k \mathbb{D}_{\hat{F}_k}$ satisfying the following:*

(i) $F'_k \subset \hat{F}_k$ *for all* k;

(ii) *the quasi-continuous λ-modification $\tilde{\psi}_n$ belongs to $C(\{F'_k\})$ for all n;*

(iii) $0 \le \psi_n \le 1$ λ-a.e. *on X and $\tilde{\psi}_n = 1$ on F'_n for all n.*

Proof. Take a sequence $\{\eta_n\}_{n\in\mathbb{N}} \subset \bigcup_k \mathbb{D}_{\hat{F}_k}$ such that $\|\eta_n - \rho\|_{\mathbb{D}} < 1/(n2^{n+1})$, $n \in \mathbb{N}$. By Lemma 2.11, there exists an open set G_n so that $G_n \supset \{|\tilde{\eta}_n - \tilde{\rho}| > 1/(2n)\}$ and $\operatorname{cap}(G_n) < 2^{-n}$ for each n. Take a nest $\{E_k\}_{k\in\mathbb{N}}$ such that $E_k \subset \hat{F}_k$ and $\{\tilde{\eta}_n\}_{n\in\mathbb{N}} \subset C(\{E_k\})$. Then $\tilde{\eta}_n \ge 1/(2n)$ on $E_n \setminus G_n$ since $\tilde{\rho} \ge 1/n$ on E_n. Define $F'_k = E_k \setminus \bigcup_{n=k}^{\infty} G_n$ and $\psi_n = 0 \vee 2n\eta_n \wedge 1$. Then $\psi_n \in \bigcup_k \mathbb{D}_{\hat{F}_k}$, $\tilde{\psi}_n = 1$ on F'_n, $\{F'_k\}_{k\in\mathbb{N}}$ is a sequence of increasing closed sets, and by Lemma 2.10, we have $\operatorname{cap}(X \setminus F'_k) \le \operatorname{cap}(X \setminus E_k) + \sum_{n=k}^{\infty} \operatorname{cap}(G_n) \to 0$ as $k \to \infty$. Q.E.D.

Now, fix $n \in \mathbb{N}$ and take $m \in \mathbb{N}$ so that $\psi_n \in \mathbb{D}_{\hat{F}_m}$. Define $T_n : \mathbb{D}_b \to \mathbb{R}$ by $T_n(f) = T(\psi_n f)$. Since $\psi_n f \in \mathbb{D}_{b,\hat{F}_m} \subset \mathbb{D}_{b,F_m}$, the statement (i) of Theorem 2.18 implies $|T_n(f)| \le C_m \|\psi_n f\|_{L^\infty(X)} \le C_m \|f\|_{L^\infty(X)}$.

For an arbitrary $f \in C(\hat{F}_m)$, we can take $\{f_j\}_{j\in\mathbb{N}} \subset \mathcal{A}$ such that $\lim_{j\to\infty} \|f_j - f\|_{C(\hat{F}_m)} = 0$. Then,

$$|T_n(f_i) - T_n(f_j)| \le C_m \|\psi_n(f_i - f_j)\|_{L^\infty(X)} \le C_m \|f_i - f_j\|_{L^\infty(\hat{F}_m)} \to 0$$

as $i \ge j \to \infty$. The limit of $\{T_n(f_j)\}_{j\in\mathbb{N}}$, denoted by $\hat{T}_n(f)$, satisfies $|\hat{T}_n(f)| \le C_m \|f\|_{C(\hat{F}_m)}$. Therefore, $\hat{T}_n$ is a bounded linear functional on $C(\hat{F}_m)$. On account of the Riesz representation theorem, there exists an associated finite signed measure ν_n on $\hat{F}_m$ such that $\hat{T}_n(f) = \int_{\hat{F}_m} f\, d\nu_n$ for every $f \in C(\hat{F}_m)$. We extend ν_n to a measure on X by letting $\nu_n(A) := \nu_n(A \cap \hat{F}_m)$.

Lemma 3.3. *The measure ν_n charges no exceptional sets.*

Proof. Since the measure $|\nu_n|$ restricted on $\hat{F}_m$ is regular, it is enough to prove that $|\nu_n|(K) = 0$ for any compact set $K \subset \hat{F}_m$ of null capacity. Take such K. Then we can take a sequence $\{O_j\}_{j\in\mathbb{N}}$ from $\mathcal{O}$

so that $K \subset O_j$ for all j and $\lim_{j\to\infty} \mathrm{cap}(O_j) = 0$. Indeed, for each j, take an open set O such that $K \subset O$ and $\mathrm{cap}(O) < 1/j$. Since K is compact, there is a set V in $\{V_{k,m}\}_{k\in\mathbb{N}}$ such that $K \subset V \subset O$. Choose $O_j \in \mathcal{O}$ so that $V \subset O_j$ and $\mathrm{cap}(O_j) \le \mathrm{cap}(V) + 1/j$.

Let $f_j := e_{O_j}$. Since $\{\hat{F}_k\}$ is a regular nest, Lemma 2.14 implies that $\tilde{f}_j = 1$ on $K \subset O_j \cap \hat{F}_m$. We may also assume that $0 \le \tilde{f}_j \le 1$ everywhere. Since $\lim_{j\to\infty} \|f_j\|_{\mathbb{D}} = 0$, we can suppose $f_j \to 0$ λ-a.e. as $j \to \infty$ by taking a subsequence if necessary. Since $\{\tilde{f}_j\}_{j\in\mathbb{N}}$ is bounded in $L^2(|\nu_n|)$, the arithmetic means $\{\hat{f}_j\}_{j\in\mathbb{N}}$ of a further subsequence of $\{\tilde{f}_j\}_{j\in\mathbb{N}}$ converge strongly in $L^2(|\nu_n|)$. Take a sequence $\{j_l\} \uparrow \infty$ such that $\hat{f}_{j_l}$ converges $|\nu_n|$-a.e. as $l \to \infty$. Define $f(z) = \liminf_{l\to\infty} \hat{f}_{j_l}(z)$. Then $0 \le f \le 1$ on X, $f = 1$ on K, and $f = 0$ λ-a.e. by the way of construction.

Given $h \in \mathcal{A}$, we have

$$\int_{\hat{F}_m} \hat{f}_{j_l} h \, d\nu_n = \hat{T}_n(\hat{f}_{j_l} h|_{\hat{F}_m}) = T_n(\hat{f}_{j_l} h) = T(\psi_n \hat{f}_{j_l} h). \tag{3}$$

When l tends to ∞, the left-hand side of (3) converges to $\int_{\hat{F}_m} fh \, d\nu_n$ by the dominated convergence theorem. On the other hand, $\{\psi_n \hat{f}_{j_l} h\}_{l\in\mathbb{N}}$ is bounded in $\mathbb{D}$ by (A4)'. Since they converge to 0 λ-a.e., they also converge weakly to 0 in $\mathbb{D}$ by Lemma 2.1. Therefore, the right-hand side of (3) converges to 0 as $l \to \infty$. Namely, $\int_{\hat{F}_m} fh \, d\nu_n = 0$. Since $\{h|_{\hat{F}_m} \mid h \in \mathcal{A}\}$ is dense in $C(\hat{F}_m)$, we conclude that $f \, d\nu_n = 0$, therefore, $|\nu_n|(K) = 0$. Q.E.D.

Lemma 3.4. *For all $f \in \mathbb{D}_b$, $T_n(f) = \int_X \tilde{f} \, d\nu_n$.*

Proof. We can take a sequence $\{f_j\}_{j\in\mathbb{N}}$ from $\mathcal{D}$ so that $\{f_j\}_{j\in\mathbb{N}}$ is bounded in $L^\infty(X)$, f_j converges to f in $\mathbb{D}$ and $\tilde{f}_j$ converges to $\tilde{f}$ outside some Borel exceptional set N_0. Note that $\tilde{f}_j|_{\hat{F}_m} \in C(\hat{F}_m)$. Then, $T_n(f_j) \to T_n(f)$ as $j \to \infty$, while

$$\begin{aligned} T_n(f_j) \quad &= \quad \hat{T}_n(\tilde{f}_j|_{\hat{F}_m}) = \int_X \tilde{f}_j \, d\nu_n = \int_{X\setminus N_0} \tilde{f}_j \, d\nu_n \\ &\xrightarrow{j\to\infty} \int_{X\setminus N_0} \tilde{f} \, d\nu_n = \int_X \tilde{f} \, d\nu_n \end{aligned}$$

by means of the dominated convergence theorem. Q.E.D.

For any k, $l \in \mathbb{N}$, we have $\tilde{\psi}_k \, d\nu_l = \tilde{\psi}_l \, d\nu_k$. Indeed, For $f \in \mathcal{A}$,

$$\int_X f\tilde{\psi}_k \, d\nu_l = T_l(f\psi_k) = T(\psi_l f \psi_k) = T_k(f\psi_l) = \int_X f\tilde{\psi}_l \, d\nu_k.$$

Therefore, we can define a signed smooth measure ν by $\nu = \nu_n$ on F'_n $(n = 1, 2, \ldots)$, which is well-defined by the fact that $\tilde{\psi}_n = 1$ on F'_n. Then for any $f \in \mathbb{D}_{b,F'_k}$, we have $\psi_k f = f$ and

$$T(f) = T(\psi_k f) = T_k(f) = \int_X \tilde{f}\, d\nu_k = \int_X \tilde{f}\, d\nu.$$

Thus, (ii) holds.

In order to prove the uniqueness of ν, it is enough to show that $\nu \equiv 0$ if $\int_X \tilde{v}\, d\nu = 0$ for all $v \in \bigcup_k \mathbb{D}_{b,F_k}$, where $\{F_k\}_{k \in \mathbb{N}}$ is a nest attached with ν. Following the same procedure as in the proof of (i)$\Rightarrow$(ii), take the nests $\{\hat{F}_k\}_{k \in \mathbb{N}}$ and $\{F'_k\}_{k \in \mathbb{N}}$, the function space $\mathcal{A}$, and the sequence of functions $\{\psi_n\}_{n \in \mathbb{N}}$. For any $n \in \mathbb{N}$ and $f \in \mathcal{A}$, we have $f\psi_n \in \mathbb{D}_{b,\hat{F}_n} \subset \mathbb{D}_{b,F_n}$, therefore $\int_X f\tilde{\psi}_n\, d\nu = 0$. Since $\{f|_{\hat{F}_n} \mid f \in \mathcal{A}\}$ is dense in $C(\hat{F}_n)$, we have $\tilde{\psi}_n\, d\nu = 0$. In particular, $\nu = 0$ on F'_n because $\tilde{\psi}_n = 1$ on F'_n. This implies that $\nu \equiv 0$.

The implication (ii)$\Rightarrow$(i) of Theorem 2.19 is proved in the same way as in Theorem 2.18. Because of the result and the proof of (i)$\Rightarrow$(ii) of Theorem 2.18, Theorem 2.19 (i) implies that there exists a finite signed smooth measure ν with some attached nest $\{F'_k\}_{k \in \mathbb{N}}$ such that the total variation is dominated by C and $T(v) = \int_X \tilde{v}(z)\, \nu(dz)$ for all $v \in \bigcup_k \mathbb{D}_{b,F'_k}$. It is easy to show that this identity holds for all $v \in \mathbb{D}_b$ by an approximation argument.

The uniqueness of ν is clear from the corresponding result of Theorem 2.18. The final claim is also deduced by an approximation argument and the use of Lemma 2.1.

This completes the proof of Theorems 2.18 and 2.19.

§4. Application to BV functions on Wiener space

First, we will review some results of [9]. Let E be a separable Banach space and H a separable Hilbert space which is continuously and densely imbedded to E. We use the notations in the end of Section 2 with letting $X = E$. Define a Gaussian measure μ on E by the following identity:

$$\int_E \exp(\sqrt{-1}\,\ell(z))\, \mu(dz) = \exp(-\|\ell\|_H^2/2), \qquad \ell \in E^* \subset H.$$

When Y is a separable Hilbert space and ρ is a nonnegative measurable function on E, we denote by $L^p(E \to Y; \rho)$ in this section the L^p space consisting of Y-valued functions on E with underlying measure $\rho\, d\mu$. We omit $E \to Y$ and ρ from the notation when $Y = \mathbb{R}$ and $\rho \equiv 1$, respectively, and write simply L^p for $L^p(E \to \mathbb{R}; 1)$. We also set

$L^{\infty-} = \bigcap_{p>1} L^p$ and denote by L^p_+ the set of all nonnegative functions in L^p.

If $u \in \mathcal{F}C_b^1$ and $v \in \mathcal{F}C_b^1$ coincide on a measurable set A, then $\nabla u = \nabla v$ μ-a.e. on A. See Proposition I.7.1.4 of [4] for the proof.

For $p \geq 1$, $Cl_p(E)$ denotes the set of all functions ρ in L^1_+ such that $(\nabla, \mathcal{F}C_b^1)$ is closable as a map from $L^p(\rho)$ to $L^p(E \to H; \rho)$. A simple example for such ρ is a function which is uniformly away from 0. Suppose $\rho \in Cl_p(E)$. We write $W^{1,p}(\rho)$ instead of $W^{1,p}$ when regarding $(E, \rho\, d\mu)$ as (X, λ) in Section 2. When $p > 1$, $W^{1,p}(\rho)$ satisfies all the conditions (A1)–(A4), (QR1)–(QR3) and (C).

Let F^ρ be the topological support of the measure $\rho\, d\mu$. Since $L^0(E \to Y; \rho)$ is identified with $L^0(F^\rho \to Y; \rho)$, we abuse the notation and $W^{1,p}(\rho)$ is also regarded as a function space on F^ρ. When $\rho \in Cl_2(E)$, an associated Dirichlet form $(\mathcal{E}^\rho, W^{1,2}(\rho))$ on $L^2(F^\rho; \rho)$ is defined by

$$\mathcal{E}^\rho(f, g) = \int_{F^\rho} \langle \nabla f, \nabla g \rangle_H \,\rho\, d\mu, \qquad f,\, g \in W^{1,2}(\rho).$$

This is a quasi-regular Dirichlet form and a finite signed measure ν on F^ρ is smooth with respect to $\mathcal{E}^\rho$ if and only if ν is $W^{1,2}(\rho)$-smooth.

For each $G \in (\mathcal{F}C_b^1)_{E^*}$, the (formal) adjoint $\nabla^* G$ is defined by the following identity:

$$\int_E (\nabla^* G) u \, d\mu = \int_E \langle G, \nabla u \rangle_H \, d\mu \qquad \text{for all } u \in \mathcal{F}C_b^1.$$

Denote by $L(\log L)^{1/2}$ the space of all functions f on E such that $\Phi \circ |f| \in L^1$, where $\Phi(x) = x((\log x) \vee 0)^{1/2}$. We say that a real measurable function ρ on E is of bounded variation ($\rho \in BV(E)$) if $\rho \in L(\log L)^{1/2}$ and

$$V(\rho) := \sup_G \int_E (\nabla^* G) \rho \, d\mu < \infty,$$

where G is taken over all functions in $(\mathcal{F}C_b^1)_{E^*}$ such that $\|G(z)\|_H \leq 1$ for every $z \in E$.

Let $\{T_t\}_{t>0}$ be the Ornstein-Uhlenbeck semigroup, which is associated with $\mathcal{E}^1$. It is strongly continuous, analytic and contractive on L^p for any $p \in (1, \infty)$.

We recall some results discussed in [9].

Theorem 4.1. (i) *For $\rho \in BV(E)$, $\|\nabla T_t \rho\|_{L^1} \leq V(\rho)$ for every $t > 0$.*

(ii) *$BV(E)$ is a vector lattice. Namely, it is a vector space, and for each $\rho \in BV(E)$, ρ_+ also belongs to $BV(E)$.*

(iii) *A function ρ belongs to $BV(E)$ if and only if $\rho \in L^1$ and there exists a sequence $\{\rho_n\}_{n\in\mathbb{N}}$ in $W^{1,1}(:= W^{1,1}(1))$ such that $\|\rho_n\|_{W^{1,1}}$ is bounded in n and $\rho_n \to \rho$ in L^1 as $n \to \infty$.*

(iv) *Each $\rho \in BV(E)$ has a unique finite Borel measure ν and a unique H-valued Borel function σ on E such that $\|\sigma\|_H = 1$ ν-a.e. and for every $G \in (\mathcal{F}C_b^1)_{E^*}$,*

$$\int_E (\nabla^* G)\rho \, d\mu = \int_E \langle G, \sigma\rangle_H \, d\nu.$$

The measure ν is $W^{1,2}(|\rho|+1)$-smooth. If moreover $\nu \in Cl_2(E)$, then $\nu|_{E\setminus F^\rho} = 0$ and ν is $W^{1,2}(\rho)$-smooth.

In what follows, we will write ν_ρ for ν in the theorem above. In this section, we improve the result for the smoothness of ν_ρ. In view of the proof of Theorem 4.1 (iv) (Theorem 3.9 of [9]), the smoothness of ν_ρ is derived from the smoothness of ν_ℓ for each $\ell \in E^*$, where ν_ℓ is a unique finite signed measure on E satisfying

$$\int_E \partial_\ell u(z)\rho(z)\,\mu(dz) = -2\int_E u(z)\,\nu_\ell(dz), \qquad u \in \mathcal{F}C_b^1.$$

Therefore, applying Theorem 2.19 with $\mathcal{L} = \mathcal{F}C_b^1$, if we show that the functional

$$I_\ell : \mathcal{F}C_b^1 \ni u \mapsto \int_E \partial_\ell u(z)\rho(z)\,\mu(dz) \in \mathbb{R} \tag{4}$$

extends continuously on $\mathbb{D}$, where $\mathbb{D}$ satisfies (A1)–(A4), (QR1)–(QR3), and (C), and has $\mathcal{F}C_b^1$ as a dense set, then we can say that ν_ℓ, hence ν_ρ, is $\mathbb{D}$-smooth. It is obvious that I_ℓ extends to a continuous functional on $W^{1,p}(|\rho|+1)$ (and $W^{1,p}(\rho)$ if furthermore $\rho \in Cl_p$) for every $p \geq 1$. Also, if $\rho \in L^q$ for some $q \in (1,\infty)$, then I_ℓ extends to a continuous functional on $W^{1,q/(q-1)}(:= W^{1,q/(q-1)}(1))$ by Hölder's inequality. Therefore, we have the following results.

Proposition 4.2. *Let $\rho \in BV(E)$. Then, ν_ρ is $W^{1,p}(|\rho|+1)$-smooth for every $p > 1$. If moreover $\rho \in Cl_p$, then ν is $W^{1,p}(\rho)$-smooth.*

Proposition 4.3. *Let $\rho \in BV(E) \cap L^q$ for some $q \in (1,\infty)$. Then, ν_ρ is $W^{1,q/(q-1)}$-smooth.*

In Proposition 4.2, the smaller p is, the stronger the claim is.

Now, we will give other examples of $\mathbb{D}$ so that ν_ρ is $\mathbb{D}$-smooth. Let us recall the Sobolev spaces in the context of Malliavin calculus. We give several notations in somewhat informal way. We refer to [12] for precise definitions. Let $L = -\nabla^*\nabla$ be the Ornstein-Uhlenbeck operator, which is regarded as a generator of $\{T_t\}_{t>0}$. The Sobolev space $\mathbb{D}^{\alpha,p}$, $\alpha \in \mathbb{R}$, $1 < p < \infty$, is given by $\mathbb{D}^{\alpha,p} = (1-L)^{-\alpha/2}(L^p)$. Each $\mathbb{D}^{\alpha,p}$ is a separable Banach space with norm $\|f\|_{\mathbb{D}^{\alpha,p}} := \|(1-L)^{\alpha/2}f\|_{L^p}$. The topological dual of $\mathbb{D}^{\alpha,p}$ is identified with $\mathbb{D}^{-\alpha,q}$, $q = p/(p-1)$. When $n \in \mathbb{N}$, by Meyer's equivalence, ∇^n is defined as a continuous operator from $\mathbb{D}^{n,p}$ to $L^p(E \to H^{\otimes n})$ and $\|\cdot\|_{L^p} + \|\nabla^n \cdot\|_{L^p(E\to H^{\otimes n})}$ gives a norm on $\mathbb{D}^{n,p}$ which is equivalent to $\|\cdot\|_{\mathbb{D}^{n,p}}$. In particular, $W^{1,p}(:= W^{1,p}(1))$ is identical with $\mathbb{D}^{1,p}$ as a set and their norms are mutually equivalent.

We define another Sobolev space $\mathbb{E}^{\alpha,p}$, $\alpha \in \mathbb{R}$, $1 < p < \infty$, firstly introduced in [24], by

$$\mathbb{E}^{\alpha,p} = \begin{cases} \mathbb{D}^{\alpha,p} & \text{if } \alpha \in \mathbb{Z}, \\ (\mathbb{D}^{k+1,p}, \mathbb{D}^{k,p})_{k+1-\alpha,p} & \text{if } k < \alpha < k+1,\ k \in \mathbb{Z}. \end{cases}$$

The general theory of real interpolation implies that $(\mathbb{E}^{\alpha,p})^*$ is identified with $\mathbb{E}^{-\alpha,q}$, where $q = p/(p-1)$ (see also [24]). When $0 < \alpha < 1$ and $1 < p < \infty$, $\mathbb{E}^{\alpha,p}$ satisfies conditions (A1)–(A4), (QR1)–(QR3), and (C) by virtue of Proposition 2.21, if $\mathbb{E}^{\alpha,p}$ is equipped with a norm deduced by $(W^{1,p}, L^p)_{1-\alpha,p}$. For such indices, $\mathcal{F}C_b^1$ is dense in $\mathbb{E}^{\alpha,p}$ since $W^{1,p}$ is dense in $\mathbb{E}^{\alpha,p}$. For later use, following [1, 2], we introduce another equivalent norm on $\mathbb{E}^{\alpha,p}$ based on the K-method by

$$\|f\|'_{\mathbb{E}^{\alpha,p}} = \left(\int_0^1 (\varepsilon^{-\alpha}K(\varepsilon,f))^p \frac{d\varepsilon}{\varepsilon}\right)^{1/p},$$

where

$$K(\varepsilon, f) = \inf\left\{\|f_1\|_{L^p} + \varepsilon\|f_2\|_{W^{1,p}} \mid f = f_1 + f_2,\ f_1 \in L^p,\ f_2 \in \mathbb{D}^{1,p}\right\}.$$

The connection between $BV(E)$ and $\mathbb{E}^{\alpha,p}$ is given as follows.

Theorem 4.4. *Let $q > 1$. Then $BV(E) \cap L^q \subset \mathbb{E}^{\alpha,p}$ if $1 < p < q$ and $\alpha < (1/p - 1/q)/(1 - 1/q)$. Also, this inclusion is continuous when $BV(E) \cap L^q$ is equipped with norm $\|f\|_{BV(E)\cap L^q} = V(f) + \|f\|_{L^q}$. In particular, $BV(E) \cap L^{\infty-} \subset \mathbb{E}^{\alpha,p}$ if $p > 1$ and $\alpha p < 1$.*

For the proof, we need the following estimates.

Lemma 4.5. (i) *When $\theta/a + (1-\theta)/b = 1/p$ with $0 < \theta < 1$, $a, b, p \ge 1$, we have $\|f\|_{L^p} \le \|f\|_{L^a}^{\theta}\|f\|_{L^b}^{1-\theta}$.*

(ii) *For each $r \ge 0$ and $p \in (1,\infty)$, there exists some C such that $\|(1-L)^r T_t f\|_{L^p} \le Ct^{-r}\|f\|_{L^p}$ for every $t \in (0,1]$ and $f \in L^p$.*

Proof. The claim (i) follows from a simple application of Hölder's inequality. The claim (ii) is a consequence of Theorem 6.13 (c) of Chapter 2 in [18], since $\{T_t\}_{t>0}$ is an analytic semigroup on L^p. Q.E.D.

Proof of Theorem 4.4. Let $f \in BV(E) \cap L^q$ with $V(f) + \|f\|_{L^q} \le 1$. In the following, c_i denotes a constant depending only on p and q. By Theorem 4.1 (i), $\|\nabla T_t f\|_{L^1} \le V(f) \le 1$ for any $t > 0$. By virtue of Meyer's equivalence and Lemma 4.5 (ii), for $t \in (0, 1]$,

$$\|\nabla T_t f\|_{L^q} \le c_1 \|(1-L)^{1/2} T_t f\|_{L^q} \le c_2 t^{-1/2}.$$

Applying Lemma 4.5 (i) with $a = 1$ and $b = q$, that is, $\theta = (1/p - 1/q)/(1-1/q)$, we have $\|\nabla T_t f\|_{L^p} \le (c_2 t^{-1/2})^{1-\theta}$ for $t \in (0, 1]$, therefore,

$$\|T_t f\|_{W^{1,p}} \le c_3 t^{-(1-\theta)/2}. \tag{5}$$

From the identity

$$\begin{aligned} f - T_t f &= -\int_0^t \frac{d}{ds} T_s f\, ds = -\int_0^t L T_s f\, ds \\ &= \int_0^t \{((1-L)^{1/2} T_{s/2})^2 f - T_s f\}\, ds, \end{aligned}$$

we obtain, for $t \in (0, 1]$,

$$\begin{aligned} \|f - T_t f\|_{L^p} &\le \int_0^t \|((1-L)^{1/2} T_{s/2})^2 f\|_{L^p}\, ds + t\|f\|_{L^p} \\ &\le \int_0^t c_4 s^{-1/2} \|(1-L)^{1/2} T_{s/2} f\|_{L^p}\, ds + t \\ &\le \int_0^t c_5 s^{-1/2} \|T_{s/2} f\|_{W^{1,p}}\, ds + t \\ &\le \int_0^t c_6 s^{-1/2} s^{-(1-\theta)/2}\, ds + t \le c_7 t^{\theta/2}. \end{aligned} \tag{6}$$

Here we used Lemma 4.5 (ii) in the second line and (5) in the last line. By combining (5) and (6), for each $\varepsilon \in (0, 1]$,

$$K(\varepsilon, f) \le \|f - T_{\varepsilon^2} f\|_{L^p} + \varepsilon \|T_{\varepsilon^2} f\|_{W^{1,p}} \le c_8 \varepsilon^\theta,$$

and, if $\alpha \in (0, \theta)$,

$$\left(\int_0^1 (\varepsilon^{-\alpha} K(\varepsilon, f))^p \frac{d\varepsilon}{\varepsilon} \right)^{1/p} \le c_8 \{p(\theta - \alpha)\}^{-1/p} < \infty.$$

This proves the claim. Q.E.D.

Using Theorem 4.4, we obtain the $\mathbb{E}^{\alpha,p}$-smoothness of ν_ρ by the following proposition.

Proposition 4.6. *Let $\rho \in BV(E) \cap L^q$, $q > 1$. Then the map I_ℓ in (4) extends continuously on $\mathbb{E}^{\alpha,p}$ if $p > q/(q-1)$ and $\alpha p > q/(q-1)$. Therefore, ν_ρ is $\mathbb{E}^{\alpha,p}$-smooth for such α and p with $\alpha \in (0,1)$. In particular, if $\rho \in BV(E) \cap L^{\infty-}$, then ν_ρ is $\mathbb{E}^{\alpha,p}$-smooth for any α, p with $\alpha \in (0,1)$ and $\alpha p > 1$.*

Proof. Due to Meyer's equivalence, the map $u \mapsto \partial_\ell u$ is continuous from $\mathbb{D}^{1,p}$ to L^p and from L^p to $\mathbb{D}^{-1,p}$, respectively. By the real interpolation theorem, it is continuous from $\mathbb{E}^{\alpha,p}$ to $\mathbb{E}^{\alpha-1,p}$ for any $\alpha \in (0,1)$. The claim follows from the fact $(\mathbb{E}^{1-\alpha,p/(p-1)})^* = \mathbb{E}^{\alpha-1,p}$ and $BV(E) \cap L^q \subset \mathbb{E}^{1-\alpha,p/(p-1)}$ by the assumption and Theorem 4.4. Q.E.D.

REMARK 4.7. (i) In [24], it is proved that $\mathbb{D}^{\alpha+\varepsilon,p} \hookrightarrow \mathbb{E}^{\alpha,p} \hookrightarrow \mathbb{D}^{\alpha-\varepsilon,p}$ for every $\alpha \in \mathbb{R}$, $1 < p < \infty$ and $\varepsilon > 0$. Therefore, Theorem 4.4 and Proposition 4.6 remain valid if we replace $\mathbb{E}^{\alpha,p}$ by $\mathbb{D}^{\alpha,p}$.

(ii) When $\rho \in BV(E)$ is an indicator function of some set A, ν_ρ can be regarded as a surface measure of A. The smoothness of ν_ρ that is proved in the proposition above is consistent with Theorem 9 of [6] saying that the Hausdorff measure of codimension n on Wiener space does not charge any set of (α,p)-capacity as long as $p > 1$ and $\alpha p > n$.

Lastly, we give a few nontrivial examples of BV functions, referring to the work [2]. Note that by combining Theorem 4.8 and Theorem 4.4 we recover a part of the results in [2].

Theorem 4.8. (i) *Let F be a function such that $F \in \mathbb{D}^{2,p}$ and $\|\nabla F\|_H^{-1} \in L^q$ for some $p > 1$ and $q > 1$ with $1/p + 1/q < 1$. Let $A = \{F < x\}$ with $x \in \mathbb{R}$. Then $1_A \in BV(E)$.*

(ii) *Suppose that (E, H, μ) is a classical Wiener space on $[0,1]$. For $x > 0$, set $A = \{w \in E \mid \max_{0\le s\le 1} |w(s)| < x\}$. Then $1_A \in BV(E)$.*

Proof. (i): From the assumptions, we have $\nabla^*(\nabla F/\|\nabla F\|_H) \in L^a$ for some $a > 1$. Indeed, keeping in mind the fact $\|\nabla F\|_H$, $\|\nabla^2 F\|_{H\otimes H} \in L^p$ due to Meyer's equivalence, let ε tend to 0 in the identity

$$\nabla^*\left(\frac{\nabla F}{\sqrt{\|\nabla F\|_H^2 + \varepsilon}}\right) = -\frac{LF}{\sqrt{\|\nabla F\|_H^2 + \varepsilon}} + \frac{\langle \nabla F \otimes \nabla F, \nabla^2 F\rangle_{H\otimes H}}{(\|\nabla F\|_H^2 + \varepsilon)^{3/2}}.$$

Now, set $\psi_n(y) = n 1_{[x-1/n,x]}(y)$, $\varphi_n(y) = \int_y^\infty \psi_n(z)\,dz$, and $\rho_n = \varphi_n(F)$. Then we have

$$\|\nabla \rho_n\|_{L^1(E\to H)} = \int_E \psi_n(F)\|\nabla F\|_H\,d\mu$$

$$
\begin{aligned}
&= -\int_E \left\langle \nabla \rho_n, \frac{\nabla F}{\|\nabla F\|_H} \right\rangle_H d\mu \\
&= -\int_E \rho_n \nabla^* \left(\frac{\nabla F}{\|\nabla F\|_H} \right) d\mu \\
&\leq \left\| \nabla^* \left(\frac{\nabla F}{\|\nabla F\|_H} \right) \right\|_{L^1},
\end{aligned}
$$

which is bounded in n. Since $\{\rho_n\}_{n\in\mathbb{N}}$ is uniformly bounded and converges to 1_A pointwise, Theorem 4.1 (iii) completes the proof.

(ii): Set $\rho_n(w) = 0 \vee n\,(1 - \max_{0\leq s\leq 1} |w(s)|/x) \wedge 1$, $w \in E$ for each $n \in \mathbb{N}$. By the calculation in the proof of Theorem 3.1 of [2], we have $\rho_n \in W^{1,1}$ and $\|\rho_n\|_{W^{1,1}}$ is bounded in n. Since $\{\rho_n\}_{n\in\mathbb{N}}$ is uniformly bounded and tends to 1_A pointwise, Theorem 4.1 (iii) completes the proof. Q.E.D.

References

[1] H. Airault, P. Malliavin and J. Ren, Smoothness of stopping times of diffusion processes, J. Math. Pures Appl. (9) **78** (1999), 1069–1091.

[2] H. Airault and X. Zhang, Smoothness of indicator functions of some sets in Wiener spaces, J. Math. Pures Appl. (9) **79** (2000), 515–523.

[3] B. Beauzamy, Espaces d'interpolation réels: topologie et géométrie, Lecture Notes in Math. **666**, Springer, Berlin, 1978.

[4] N. Bouleau and F. Hirsch, Dirichlet forms and analysis on Wiener space, de Gruyter Studies in Mathematics **14**, Walter de Gruyter, Berlin, 1991.

[5] N. Bourbaki, Topological vector spaces, Chapters 1–5, translated from the French by H. G. Eggleston and S. Madan, Springer, Berlin, 1987.

[6] D. Feyel and A. de La Pradelle, Hausdorff measures on the Wiener space, Potential Anal. **1** (1992), 177–189.

[7] M. Fukushima, On semi-martingale characterizations of functionals of symmetric Markov processes, Electron. J. Probab. **4** (1999), no. 18, 32 pp. (electronic).

[8] M. Fukushima, BV functions and distorted Ornstein Uhlenbeck processes over the abstract Wiener space, J. Funct. Anal. **174** (2000), 227–249.

[9] M. Fukushima and M. Hino, On the space of BV functions and a related stochastic calculus in infinite dimensions, J. Funct. Anal. **183** (2001), 245–268.

[10] M. Fukushima and H. Kaneko, (r,p)-capacities for general Markovian semigroups, in Infinite-dimensional analysis and stochastic processes (Bielefeld, 1983), 41–47, Res. Notes in Math. **124**, Pitman, Boston, MA, 1985.

[11] M. Fukushima, Y. Oshima, and M. Takeda, Dirichlet forms and symmetric Markov processes, de Gruyter Studies in Mathematics **19**, Walter de Gruyter, Berlin, 1994.
[12] N. Ikeda and S. Watanabe, Stochastic differential equations and diffusion processes, 2nd ed., North-Holland Mathematical Library **24**, North-Holland, Amsterdam, 1989.
[13] S. Kakutani, Weak convergence in uniformly convex spaces, Tôhoku Math. J. **45** (1938), 188–193.
[14] T. Kazumi and I. Shigekawa, Measures of finite (r,p)-energy and potentials on a separable metric space, in Séminaire de Probabilités, XXVI, 415–444, Lecture Notes in Math. **1526**, Springer, Berlin, 1992.
[15] Z. M. Ma and M. Röckner, Introduction to the theory of (nonsymmetric) Dirichlet forms, Universitext, Springer-Verlag, Berlin, 1992.
[16] D. P. Milman, On some criteria for the regularity of spaces of the type (B), Doklady Acad. Nauk SSSR (N.S.) **20** (1938), 243–246.
[17] D. Nualart, The Malliavin calculus and related topics, Probability and its Applications, Springer-Verlag, New York, 1995.
[18] A. Pazy, Semigroups of linear operators and applications to partial differential equations, Applied Mathematical Sciences **44**, Springer-Verlag, New York, 1983.
[19] B. J. Pettis, A proof that every uniformly convex space is reflexive, Duke Math. J. **5** (1939), 249–253.
[20] J. Ren and M. Röckner, Ray Hölder-continuity for fractional Sobolev spaces in infinite dimensions and applications, Probab. Theory Related Fields **117** (2000), 201–220.
[21] M. Röckner and B. Schmuland, Tightness of general $C_{1,p}$ capacities on Banach space, J. Funct. Anal. **108** (1992), 1–12.
[22] H. Sugita, Positive generalized Wiener functions and potential theory over abstract Wiener spaces, Osaka J. Math. **25** (1988), 665–696.
[23] H. Triebel, Interpolation theory, function spaces, differential operators, North-Holland Mathematical Library **18**, North-Holland, Amsterdam, 1978.
[24] S. Watanabe, Fractional order Sobolev spaces on Wiener space, Probab. Theory Related Fields **95** (1993), 175–198.

Graduate School of Informatics
Kyoto University
Kyoto, 606-8501
Japan

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 141–166

Least-Squares Approximation of Random Variables by Stochastic Integrals

Chunli Hou and Ioannis Karatzas

Dedicated to Professor K. Itô on the occasion of his 88th Birthday

Abstract.

This paper addresses the problem of approximating random variables in terms of sums consisting of a real constant and of a stochastic integral with respect to a given semimartingale X. The criterion is minimization of $\mathbf{L}^2$−distance, or "least-squares". This problem has a straightforward and well-known solution when X is a Brownian motion or, more generally, a square-integrable martingale, with respect to the underlying probability measure P. We address the general, semimartingale case by means of a *duality approach*; the adjoint variables in this duality are *signed measures*, absolutely continuous with respect to P, under which X behaves like a martingale. It is shown that this duality is useful, in that the value of an appropriately formulated dual problem can be computed fairly easily; that it "has no gap" (i.e., the values of the primal and dual problems coincide); that the signed measure which is optimal for the dual problem can be easily identified whenever it exists; and that the duality is also "strong", in the sense that one can then identify the optimal stochastic integral for the primal problem. In so doing, the theory presented here both simplifies and extends the extant work on the subject. It has also natural connections and interpretations in terms of the theory of "variance-optimal" and "mean-variance efficient" portfolios in Mathematical Finance, pioneered by H. Markowitz and then greatly extended by H. Föllmer, D. Sondermann and most notably M. Schweizer.

Received November 21, 2002.
This work was supported in part by the National Science Foundation under Grant NSF-DMS-00-99690.

§1. Introduction

Suppose we are given a square-integrable, d-dimensional process $X = \{X(t);\ 0 \le t \le T\}$ defined on the finite time-horizon $[0, T]$, which is a semimartingale on the filtered probability space $(\Omega, \mathcal{F}, P)$, $\mathbf{F} = \{\mathcal{F}(t)\}_{0 \le t \le T}$. *How closely can we approximate in the sense of least-squares a given, square-integrable and $\mathcal{F}(T)$-measurable random variable H, by a linear combination of the form $c + \int_0^T \vartheta' dX$?* Here c is a real number and ϑ a predictable d-dimensional process for which the stochastic integral $\int_0^{\cdot} \vartheta' dX \equiv \sum_{i=1}^d \int_0^{\cdot} \vartheta_i dX_i$ is well-defined and is itself a square-integrable semimartingale.

In other words, if we denote by Θ the space of all such processes ϑ, how do we compute

$$(1.1) \qquad V(c) \triangleq \inf_{\vartheta \in \Theta} E\left(H - c - \int_0^T \vartheta' dX\right)^2$$

if $c \in \mathbb{R}$ is given and we have the freedom to choose ϑ over the class Θ as above? How do we find

$$(1.2) \qquad V \triangleq \inf_{(c,\vartheta) \in \mathbb{R} \times \Theta} E\left(H - c - \int_0^T \vartheta' dX\right)^2 = \inf_{c \in \mathbb{R}} V(c)$$

when we have the freedom to select both c and ϑ? And how do we characterize, or even compute, the process $\vartheta^{(c)}$ and the pair $(\hat{c}, \hat{\vartheta})$ that attain the infimum in (1.1) and (1.2), respectively, whenever these exist? To go one step further: How does one

$$(1.3) \qquad \text{minimize the variance} \quad \mathrm{Var}\left(H - \int_0^T \vartheta' dX\right)$$

over all $\vartheta \in \Theta$ as above? Or even more interestingly, how does one

$$(1.4) \qquad \left\{ \begin{array}{c} \text{minimize the variance } \mathrm{Var}\left(H - \int_0^T \vartheta' dX\right) \\ \text{over } \vartheta \in \Theta \text{ with } E\left[\int_0^T \vartheta' dX\right] = \mu \end{array} \right\}$$

for some given $\mu \in \mathbb{R}$?

Questions such as (1.3) and (1.4) can be traced back to the pioneering work of H. Markowitz (1952, 1959), and have been studied more recently by Föllmer & Sondermann (1986), Föllmer & Schweizer

(1991), Duffie & Richardson (1991), Schäl (1992), in the modern context of Mathematical Finance. Most importantly, problems (1.1)-(1.4) have received an exhaustive and magisterial treatment in a series of papers by Schweizer (1992, 1994, 1995.a,b, 1996) and his collaborators (cf. Rheinländer & Schweizer (1997), [Ph.R.S.] (1998), [DMSSS] (1997), as well as Hipp (1993), [G.L.Ph.] (1996), Laurent & Pham (1999), Grandits (1999), Arai (2002)). In this context, the components $X_i(\cdot)$, $i = 1, \ldots, d$ of the semimartingale X are interpreted as the (discounted) stock-prices in a financial market, and H as a contingent claim, or liability, that one is trying to replicate as faithfully as possible at time T, starting with initial capital c and trading in this market. Such trading is modelled by the predictable *portfolio process* ϑ, whose component $\vartheta_i(t)$ represents the number of shares being held at time t in the i^{th} stock, for $i = 1, \ldots, d$. Then $\int_0^T \vartheta' dX \equiv \sum_{i=1}^d \int_0^T \vartheta_i(s) dX_i(s)$ corresponds to the (discounted) *gains from trading* accrued by the terminal time T, with which one tries to approximate the contingent claim H, and one might be interested in minimizing the variance of this approximation over all admissible portfolio choices (problem of (1.3)), or just over those portfolios that guarantee a given mean-rate-of-return (problem of (1.4)).

It turns out that solving the problem of (1.1) provides the key to answering all these questions. For instance, if $\vartheta^{(c)}$ attains the infimum in (1.1) and $\hat{c} \equiv \mathrm{argmin}_{c \in \mathbb{R}} V(c)$, then $\hat{\vartheta} \equiv \vartheta^{(\hat{c})}$ is optimal for the problem of (1.3); the pair $(\hat{c}, \hat{\vartheta})$ is optimal for the problem of (1.2); and the process $\vartheta^{(c_\mu)}$ with $c_\mu = \left(E[\pi(H)] - E(H) + \mu\right) / \left(E[\pi(1)] - 1\right)$ is optimal for the problem of (1.4). Here π denotes the projection operator from the Hilbert space $\mathbf{L}^2(P)$ onto the orthogonal complement of its linear subspace $\left\{ \int_0^T \vartheta' dX \,/\, \vartheta \in \Theta \right\}$.

The problem of (1.1) has a very simple solution, if X is a (square-integrable) *martingale*; then every H as above has the so-called Kunita-Watanabe decomposition

$$H = E(H) + \int_0^T (\zeta^H)' dX + L^H(T),$$

where $\zeta^H \in \Theta$ and $L^H(\cdot)$ is a square-integrable martingale strongly orthogonal to $\int_0^{\cdot} \vartheta' dX$ for every $\vartheta \in \Theta$. Then the infimum in (1.1) is computed as $V(c) = (E[H] - c)^2 + E\left(L^H(T)\right)^2$ and is attained by $\zeta^H \in \Theta$, which also attains the infimum $V = E\left(L^H(T)\right)^2$ in (1.2).

In order to deal with a general semimartingale X we develop a simple *duality approach*, which in a sense tries to reduce the problem to the "easy" martingale case just described. This approach is the main contribution of the present paper. The dual or "adjoint" variables in this duality are signed measures Q, absolutely continuous with respect to P and with $dQ/dP \in \mathbf{L}^2(P)$, under which X behaves like a martingale (Definition 2.1 and Remark 2.2). A simple observation, described in (3.1)-(3.7), leads to a dual *maximization problem.* The resulting duality is useful because, as it turns out, the dual problem is relatively straightforward to solve (Proposition 3.1); its value is easily computed as $E[\pi^2(H-c)]$ and coincides with the value $V(c)$ of the original problem (1.1), so *there is no "duality gap"*; and furthermore the duality is "strong", in that one can identify the optimal integrand $\vartheta^{(c)}$ of (1.1) rather easily, under suitable conditions (Theorem 4.1 and Remark 4.1). Several examples are presented in Section 5.

We follow closely the notation and the setting of Schweizer (1996), our great debt to which should be clear to anyone familiar with this excellent work. Indeed, the present paper can be considered as complementing and extending the results of this work, by means of our simple duality approach.

§2. The Problem

On a given complete probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $\mathbf{F} = \{\mathcal{F}(t)\,;\, 0 \le t \le T\}$ that satisfies the usual conditions, consider a process

$$X(t) = X(0) + M(t) + B(t), \qquad 0 \le t \le T, \tag{2.1}$$

defined on the finite time-horizon $[0,T]$ and belonging to the space $\mathcal{S}^2 \equiv \mathcal{S}^2(P)$ of square-integrable d-dimensional semimartingales. This means that each $X_i(0)$ is in $\mathbf{L}^2(\Omega, \mathcal{F}(0), P)$; that each $M_i(\cdot)$ belongs to the space $\mathcal{M}_0^2(P)$ of square-integrable $\mathbf{F}$-martingales with $M_i(0) = 0$ and RCLL paths; and that we have $B_i(\cdot) = A_i^+(\cdot) - A_i^-(\cdot)$, where $A_i^{\pm}(\cdot)$ are increasing, right-continuous and predictable processes with $A_i^{\pm}(0) = 0$ and $E\left(A_i^{\pm}(T)\right)^2 < \infty$, for every $i = 1, \cdots, d$. We denote by Θ the space of "good integrands" for the square-integrable semimartingale $X = \{X(t), 0 \le t \le T\}$, namely, those $\mathbf{F}-$predictable processes whose stochastic integrals with respect to X are themselves square-integrable semimartingales:

$$\Theta \triangleq \left\{ \vartheta \in \mathcal{L}(X) \;/\; G_{\cdot}(\vartheta) \equiv \int_0^{\cdot} \vartheta' dX \in \mathcal{S}^2(P) \right\}. \tag{2.2}$$

Here $\mathcal{L}(X)$ stands for the space of all $\mathbb{R}^d$-valued and predictable processes, whose stochastic integrals

$$G_t(\vartheta) \triangleq \int_0^t \vartheta'(s)\, dX(s) = \sum_{i=1}^d \int_0^t \vartheta_i(s)\, dX_i(s), \quad 0 \le t \le T \tag{2.3}$$

with respect to X are well-defined.

Suppose now we are given a random variable H in the space $\mathbf{L}^2(P) \equiv \mathbf{L}^2(\Omega, \mathcal{F}(T), P)$. The following problem will occupy us in this paper.

Problem 2.1. *Given $H \in \mathbf{L}^2(P)$, compute*

$$V \triangleq \inf_{(c,\vartheta)\in\mathbb{R}\times\Theta} E\big(H - c - G_T(\vartheta)\big)^2 \tag{2.4}$$

and try to find a pair $(\hat{c}, \hat{\vartheta}) \in \mathbb{R} \times \Theta$ that attains the infimum, if such a pair exists. □

In other words, we are looking to find the *least-squares approximation* of H, as the sum of a constant $c \in \mathbb{R}$ and of the stochastic integral $G_T(\vartheta)$, for some process $\vartheta \in \Theta$.

This problem has a rather obvious solution, if it is known that the random variable H is of the form

$$H = h + G_T(\zeta^H) \tag{2.5}$$

for some $h \in \mathbb{R}$ and $\zeta^H \in \Theta$; because then we can take $\hat{c} \equiv h$, $\hat{\vartheta} \equiv \zeta^H$, and deduce that $V = 0$ in (2.4). Now it is a classical result (e.g. Karatzas & Shreve (1991), pp. 181-185 for a proof) that *every* $H \in \mathbf{L}^2(P)$ can be written in the form (2.5), in fact with $h = E(H)$, *if $X(\cdot)$ is Brownian motion* and if $\mathbf{F}$ is the (augmentation of the) filtration $\mathbf{F}^X$ generated by $X(\cdot)$ itself. One can then also describe the integrand ζ^H in terms of the famous Clark (1970) formula, under suitable conditions on the random variable $H \in \mathbf{L}^2(P) \equiv \mathbf{L}^2(\Omega, \mathcal{F}^X(T), P)$. Thus, in this special case, we can take $\hat{c} = E(H)$, $\hat{\vartheta} = \zeta^H$, and have $V = 0$ in (2.4).

A little more generally, suppose that $X(\cdot) \in \mathcal{M}_0^2(P)$ is a square-integrable *martingale* (i.e., $X(0) = 0$ and $A(\cdot) \equiv 0$ in (2.1)). Then again it is well-known that every $H \in \mathbf{L}^2(P)$ admits the so-called *Kunita-Watanabe (1967) decomposition*

$$H = h + G_T(\zeta^H) + L^H(T) \tag{2.6}$$

for $h = E(H)$, some $\zeta^H \in \Theta$, and some square-integrable martingale $L^H(\cdot) \in \mathcal{M}_0^2(P)$ which is strongly orthogonal to $G_\cdot(\vartheta)$ for every $\vartheta \in \Theta$;

in particular,

$$E\left[L^H(T) \cdot G_T(\vartheta)\right] = 0, \qquad \forall\, \vartheta \in \Theta. \tag{2.7}$$

Then

$$E\left(H - c - G_T(\vartheta)\right)^2 = (h-c)^2 + E\left(G_T(\zeta^H - \vartheta)\right)^2 + E\left(L^H(T)\right)^2,$$

and it is clear that Problem 2.1 admits again the solution $\hat{c} = E(H)$, $\hat{\vartheta} = \zeta^H$, but now with $V = E\left(L^H(T)\right)^2$ in (2.4).

What happens for a *general, square-integrable semimartingale* $X(\cdot) \in \mathcal{S}^2(P)$? In view of the above discussion it is tempting to try and "reduce" this general problem to the case where $X(\cdot)$ is a martingale. This can be accomplished by absolutely continuous change of the probability measure P. We formalize this idea as in Schweizer (1996).

Definition 2.1. A signed measure Q on $(\Omega, \mathcal{F})$ is called a *signed Θ-martingale measure*, if $Q(\Omega) = 1$, $Q \ll P$ with $(dQ/dP) \in \mathbf{L}^2(P)$, and

$$E\left[\frac{dQ}{dP} \cdot G_T(\vartheta)\right] = 0, \qquad \forall\, \vartheta \in \Theta. \tag{2.8}$$

We shall denote by $\mathbf{P}_s(\Theta)$ the set of all such signed Θ-martingale measures, and introduce the closed, convex set

$$\begin{aligned} \mathcal{D} &\triangleq \{D \in \mathbf{L}^2(P) \,/\, D = (dQ/dP), \text{ some } Q \in \mathbf{P}_s(\Theta)\} \\ &= \{D \in \mathbf{L}^2(P) \,/\, E(D) = 1 \text{ and } E(DG_T(\vartheta)) = 0, \quad \forall\, \vartheta \in \Theta\}. \end{aligned} \tag{2.9}$$

□

We shall assume from now onwards, that

$$\mathbf{P}_s(\Theta) \neq \emptyset \qquad (\text{equivalently, } \mathcal{D} \neq \emptyset). \tag{2.10}$$

Remark 2.1 : The linear subspace $G_T(\Theta) \triangleq \{\int_0^T \vartheta'(s)\, dX(s) \,/\, \vartheta \in \Theta\}$ of the Hilbert space $\mathbf{L}^2(P)$ is not necessarily closed for a general semimartingale $X(\cdot)$ (it is, if $B(\cdot) \equiv 0$ in (2.1), or equivalently if $X(\cdot)$ is a square-integrable martingale, since then the stochastic-integral of (2.3) is an isometry). For a semimartingale $X(\cdot)$ with continuous paths, necessary and sufficient conditions for the closedness of $G_T(\Theta)$ have been provided by Delbaen, Monat, Schachermayer, Schweizer & Stricker [DMSSS] (1997); see also Theorem 2 in Rheinländer & Schweizer (1997), as well as Corollary 4 in Pham, Rheinländer & Schweizer [Ph.R.S.] (1998) and section 5, equation (5.11) of the present paper, for sufficient conditions. As noted by W. Schachermayer (see Schweizer (1996), p. 210;

Lemma 4.1 in Schweizer (2001)), the assumption (2.10) is equivalent to the requirement that

$$(2.10)' \qquad \left\{ \begin{array}{l} \text{the closure in } \mathbf{L}^2(P) \text{ of } G_T(\Theta) \\ \text{does not contain the constant } 1 \end{array} \right\}.$$

On the other hand, the orthogonal complement

$$(2.11) \qquad (G_T(\Theta))^{\perp} \triangleq \{D \in \mathbf{L}^2(P) \ / \ E(DG_T(\vartheta)) = 0, \ \forall\, \vartheta \in \Theta\}$$

of $G_T(\Theta)$ is a *closed* linear subspace of $\mathbf{L}^2(P)$, and its orthogonal complement $(G_T(\Theta))^{\perp\perp}$ is the smallest closed, linear subspace of $\mathbf{L}^2(P)$ that contains $G_T(\Theta)$. Clearly, $(G_T(\Theta))^{\perp}$ includes the set $\mathcal{D}$ of (2.9), and the requirement (2.10)$'$ amounts to

$$(2.10)'' \qquad 1 \notin (G_T(\Theta))^{\perp\perp}.$$

Remark 2.2 : The notion of signed Θ-martingale measure in Definition 2.1 depends on the space Θ itself, as well as on the definition of the stochastic integral $G_T(\vartheta)$, $\vartheta \in \Theta$. In many cases of interest, though, every $Q \in \mathbf{P}_s(\Theta)$ belongs also to the space $\mathbf{P}_s^2(X)$ of signed *martingale measures for* X, namely those signed measures Q on $(\Omega, \mathcal{F})$ with $Q \ll P$, $dQ/dP \in \mathbf{L}^2(P)$, $Q(\Omega) = 1$ and

$$(2.12) \qquad E\left[\frac{dQ}{dP} \cdot \big(X(t) - X(s)\big) \,\Big|\, \mathcal{F}(s) \right] = 0, \quad \text{a.s.}$$

for any $0 \le s \le t \le T$. (If Q is a true probability measure, as opposed to a signed measure with $Q(\Omega) = 1$, then (2.12) amounts to the martingale property of X under Q.) See Müller (1985), Lemma 12(b) in Schweizer (1996), as well as conditions (5.1)-(5.4) and the paragraph immediately following them in the present paper.

In addition to Problem 2.1, it is useful to consider also the following question, which is interesting in its own right.

Problem 2.2. *Given $H \in \mathbf{L}^2(P)$ and $c \in \mathbb{R}$, compute*

$$(2.13) \qquad V(c) \triangleq \inf_{\vartheta \in \Theta} E\Big(H - c - G_T(\vartheta)\Big)^2$$

and try to find $\vartheta^{(c)} \in \Theta$ that attains the infimum in (2.13), *if such exists.*

Clearly, $\inf_{c \in \mathbb{R}} V(c)$ coincides with the quantity V of (2.4); and if this last infimum is attained by some $\hat{c} \in \mathbb{R}$, then the pair $\big(\hat{c}, \vartheta^{(\hat{c})}\big)$ attains the infimum in (2.4). In the next Sections we shall try to solve Problems 2.2

and 2.1 using very elementary *duality* ideas. In this effort, the elements of the set $\mathcal{D}$ in (2.9) will play the role of *adjoint* or *dual variables.* For the duality methodology to work in any generality it is critical to allow, as we did in Definition 2.1, for *signed* measures Q with $Q(\Omega) = 1$, as opposed to just standard probability measures.

§3. The Duality

The duality approach to Problem 2.2 is simple, and is based on the elementary observation

$$\begin{aligned} \min_{x\in\mathbb{R}} \left[(H-x)^2 + yx\right] &= (H-(H-y/2))^2 + y\,(H-y/2) \\ &= yH - y^2/4, \qquad \forall\, y \in \mathbb{R}. \end{aligned} \tag{3.1}$$

The key idea now, is to read (3.1) with $x = c + G_T(\vartheta)$, $y = 2kD$ for given $c \in \mathbb{R}$ and arbitrary $\vartheta \in \Theta$, $D \in \mathcal{D}$ as in (2.2) and (2.9), and with arbitrary $k \in \mathbb{R}$, to obtain

$$(H - c - G_T(\vartheta))^2 + 2kD\,(c + G_T(\vartheta)) \;\geq\; 2kDH - k^2D^2. \tag{3.2}$$

Note also that (3.2) holds as equality for some $\vartheta \equiv \vartheta^{(c)} \in \Theta$, $D \equiv D^{(c)} \in \mathcal{D}$ and $k \equiv k^{(c)} \in \mathbb{R}$, if and only if

$$c + G_T(\vartheta^{(c)}) \;=\; H - k^{(c)}D^{(c)}. \tag{3.3}$$

Now let us take expectations in (3.2) to obtain, in conjunction with the properties of (2.9):

$$E\,(H - c - G_T(\vartheta))^2 \;\geq\; -k^2E(D^2) + 2k\,[E(DH) - c]\,, \tag{3.4}$$

for every $k \in \mathbb{R}, D \in \mathcal{D}$ and $\vartheta \in \Theta$. Clearly,

$$E(D^2) - 1 \;=\; \mathrm{Var}(D) \;\geq\; 0, \qquad \forall\, D \in \mathcal{D} \tag{3.5}$$

and the mapping $k \mapsto -k^2E(D^2) + 2k\,[E(DH) - c]$ attains over $\mathbb{R}$ its maximal value $(E(DH) - c)^2 / E(D^2)$ at the point

$$k_{D,c} \;\triangleq\; \frac{E(DH) - c}{E(D^2)}\,. \tag{3.6}$$

Thus, we obtain from (3.4) the inequality

$$
\begin{aligned}
V(c) &\triangleq \inf_{\vartheta\in\Theta} E\left(H-c-G_T(\vartheta)\right)^2 \\
&\geq \sup_{D\in\mathcal{D}} \sup_{k\in\mathbb{R}} \left[-k^2E(D^2)+2k\left(E(DH)-c\right)\right] \qquad (3.7)\\
&= \sup_{D\in\mathcal{D}} \frac{\left(E(DH)-c\right)^2}{E(D^2)} =: \tilde{V}(c),
\end{aligned}
$$

which is the basis of our *duality approach.* Here $V(c)$ is the value of our original ("primal") optimization Problem 2.2, whereas $\tilde{V}(c)$ is the value of an auxiliary ("dual") optimization problem.

This kind of duality is useful, only if the dual problem is easier to solve than the primal Problem 2.2 and if there is no "duality gap" (i.e., equality holds in (3.7)), so that by computing the value of the dual problem we also compute the value of the primal. Both these features hold for our setting, as we are about to show. Furthermore, the duality is "strong", in the sense that we can identify an optimal $\tilde{D}_c \in \mathcal{D}$ for the dual problem, namely

$$
\tilde{V}(c) = \frac{\left(E(\tilde{D}_c H)-c\right)^2}{E(\tilde{D}_c^2)} \qquad (3.8)
$$

for all but a critical value of the parameter c, and then obtain from this an optimal process $\vartheta^{(c)}$ for the primal problem via (3.3).

In order to make headway with this program, let us start by introducing the *projection operator* $\pi : \mathbf{L}^2(P) \to (G_T(\Theta))^{\perp}$ with the property

$$
E\left[(H-\pi(H))\cdot D\right]=0, \quad \forall\, H\in\mathbf{L}^2(P), \quad \forall\, D\in (G_T(\Theta))^{\perp}. \qquad (3.9)
$$

In particular,

$$
E\left[(H_1-\pi(H_1))\cdot\pi(H_2)\right] = 0, \qquad \forall\, H_1, H_2 \in \mathbf{L}^2(P), \qquad (3.9)'
$$

and from (3.9)′ and (2.10)′ we have

$$
E[\pi(1)] = E[\pi^2(1)] > 0. \qquad (3.10)
$$

Proposition 3.1. *The value of the dual problem in (3.7), namely*

$$
\tilde{V}(c) \triangleq \sup_{D\in\mathcal{D}} \frac{\left(E(DH)-c\right)^2}{E(D^2)}, \qquad (3.11)
$$

can be computed as

$$\tilde{V}(c) \;=\; E\left[\pi^2(H-c)\right], \qquad \forall\, c \in \mathbb{R}. \tag{3.12}$$

The supremum in (3.11) is attained by

$$\tilde{D}_c \;\triangleq\; \frac{\pi(H-c)}{E\left[\pi(H-c)\right]}, \qquad \forall\, c \neq \hat{c} \;\triangleq\; \frac{E[\pi(H)]}{E[\pi(1)]} \quad ; \tag{3.13}$$

it is not attained for $c = \hat{c}$.

For every $c \neq \hat{c}$, we shall call the random variable $\tilde{D}_c \in \mathcal{D}$ of (3.13) the "density of the *dual-optimal* signed martingale measure" in $\mathbf{P}_s(\Theta)$, namely

$$\tilde{Q}_c(A) \;\triangleq\; \int_A \tilde{D}_c\, dP\,, \qquad A \in \mathcal{F}\,. \tag{3.14}$$

Remark 3.1 : Suppose that for some $h \in \mathbb{R}$ we have

$$E(DH) \;=\; h, \qquad \forall\, D \in \mathcal{D}\,. \tag{3.15}$$

(For instance, this is the case when H is of the form (2.5).) Then the dual value function of (3.11) becomes

$$\tilde{V}(c) \;=\; \frac{(h-c)^2}{\inf_{D\in\mathcal{D}} E(D^2)} \tag{3.16}$$

and, for $c \neq h$, the dual-optimal $\tilde{D}_c$ of (3.13) coincides with

$$\tilde{D} \;\triangleq\; \arg\min_{D\in\mathcal{D}} E(D^2) \;=\; \frac{\pi(1)}{E[\pi(1)]} \;=\; E(\tilde{D}^2) + R \;\in\; \mathcal{D} \tag{3.17}$$

for some $R \in (G_T(\Theta))^{\perp\perp}$, and $E(\tilde{D}^2) = 1/E[\pi(1)] \geq 1$, as we shall establish below. Following Schweizer (1996), we shall call $\tilde{D}$ the "density of the *variance-optimal* signed martingale measure"

$$\tilde{Q}(A) \;\triangleq\; \int_A \tilde{D}\, dP\,, \qquad A \in \mathcal{F} \tag{3.18}$$

in $\mathbf{P}_s(\Theta)$. This terminology should be clear from (3.5) and the definition in (3.17).

• *Proof of Proposition 3.1* : For every $D \in \mathcal{D}$, we have

$$E(DH) - c \;=\; E[D(H-c)] \;=\; E[D\cdot\pi(H-c)]$$

thanks to (3.9). Thus, from the Cauchy-Schwarz inequality,

$$(E(DH)-c)^2 \;=\; (E[D\cdot\pi(H-c)])^2 \;\le\; E(D^2)\cdot E\left[\pi^2(H-c)\right],$$

which implies

$$\tilde{V}(c) \;\le\; E\left[\pi^2(H-c)\right].$$

Now these last two inequalities are valid as equalities, if and only if we can find $\tilde{D}_c \in \mathcal{D}$ of the form

$$\tilde{D}_c \;=\; \text{const}\cdot\pi(H-c), \tag{3.13$'$}$$

where the constant has to be chosen so that $E(\tilde{D}_c)=1$. This is impossible to do if $E[\pi(H-c)] = E[\pi(H) - c\cdot\pi(1)]$ is equal to zero, i.e., if $c=\hat{c}$ as in (3.13); in other words, the supremum of (3.11) cannot be attained in this case. But for $c\neq\hat{c}$, the normalizing constant in (3.13)$'$ can be taken as $1/E[\pi(H-c)]$, leading to the expression of (3.13) and to (3.12) as well.

It remains to show that (3.12) holds even for $c=\hat{c}$. For this, let $c_n \triangleq c-1/n$, $n\in\mathbb{N}$ and

$$\varphi_n \;\triangleq\; \pi(H-c_n), \quad \varphi \;\triangleq\; \pi(H-c), \quad \tilde{D}_{c_n} \;=\; \frac{\varphi_n}{E(\varphi_n)} \;\in\; \mathcal{D}$$

so that

$$\begin{aligned}
\frac{\left(E(\tilde{D}_{c_n}H)-c\right)^2}{E(\tilde{D}^2_{c_n})} &= \frac{\left(E(\tilde{D}_{c_n}H)-c_n-1/n\right)^2}{E(\tilde{D}^2_{c_n})}\\
&= \frac{\left(E(\tilde{D}_{c_n}H)-c_n\right)^2}{E(\tilde{D}^2_{c_n})} + \frac{1/n^2}{E(\tilde{D}^2_{c_n})} - \frac{2}{n}\cdot\frac{E\left[\tilde{D}_{c_n}(H-c_n)\right]}{E(\tilde{D}^2_{c_n})}\\
&= E\left[\pi^2(H-c_n)\right] + \frac{1/n^2}{E(\tilde{D}^2_{c_n})} - \frac{2}{n}E[\pi(H-c_n)]\\
&= E(\varphi_n^2) + \frac{1/n^2}{E(\tilde{D}^2_{c_n})} - \frac{2}{n}E(\varphi_n)\\
&\to E\left(\varphi^2\right) \;=\; E\left[\pi^2(H-c)\right]
\end{aligned}$$

as $n\to\infty$. We have used the inequality $0 < 1/E\left(\tilde{D}^2_{c_n}\right) \le 1$; the facts $\varphi_n - \varphi = \pi(1)/n \to 0$ a.s., $|\varphi_n| \le |\varphi| + \pi(1) \in \mathbf{L}^2(P)$; the Dominated

Convergence Theorem; and the observation that, for $c \neq \hat{c}$, we have

(3.19)

$$\begin{aligned}\frac{E\left[\tilde{D}_c \cdot (H-c)\right]}{E(\tilde{D}_c^2)} &= \frac{E\left[\tilde{D}_c \cdot \pi(H-c)\right]}{E(\tilde{D}_c^2)} \\ &= \frac{E\left[\pi^2(H-c)\right]}{E\left[\pi(H-c)\right]} \cdot \frac{(E[\pi(H-c)])^2}{E\left[\pi^2(H-c)\right]} = E[\pi(H-c)]\end{aligned}$$

from (3.9)′, (3.13). □

• *Proof of (3.17)* : For any $D \in \mathcal{D}$, we have

$$1 \;=\; (E(D\cdot 1))^2 \;=\; (E(D\cdot\pi(1)))^2 \;\le\; E\left(D^2\right)\cdot E\left[\pi^2(1)\right],$$

from (3.9) and the Cauchy-Schwarz inequality. Equality holds if and only if

$$D \;=\; \tilde{D} \;\triangleq\; \text{const}\cdot\pi(1),$$

and the normalizing constant has to be chosen so that $E(\tilde{D}) = 1$, namely, equal to $1/E[\pi(1)]$. We conclude that $\tilde{D} = \pi(1)/E[\pi(1)]$ satisfies

$$E\left(D^2\right) \;\ge\; \frac{1}{E\left[\pi^2(1)\right]} \;=\; \frac{1}{E\left[\pi(1)\right]} \;=\; E(\tilde{D}^2)\;, \qquad \forall\; D \in \mathcal{D}. \tag{3.20}$$

On the other hand, since

$$1 = \pi(1) + \eta \qquad \text{for some} \quad \eta \in (G_T(\Theta))^{\perp\perp}\;, \tag{3.21}$$

we have $\tilde{D} = (1-\eta)/E[\pi(1)] = E(\tilde{D}^2) + R$, with $R = -\eta/E[\pi(1)]$. □

Remark 3.2 : If $G_T(\Theta)$ is closed in $\mathbf{L}^2(P)$, then (3.21) becomes

$$1 \;=\; \pi(1) + G_T(\xi^1), \qquad \text{for some} \quad \xi^1 \in \Theta \tag{3.21$'$}$$

and (3.17), (3.20) give

$$\tilde{D} \;=\; E(\tilde{D}^2) + G_T(\tilde{\xi}), \qquad \text{with} \quad \tilde{\xi} \;\triangleq\; -E(\tilde{D}^2)\cdot\xi^1 \;\in\; \Theta. \tag{3.22}$$

§4. Results

We are now in a position to use the duality developed in the previous section in order to provide solutions to Problems 2.1 and 2.2. First, a lemma from Schweizer (1996), pp. 230-231; we provide the proof for completeness.

Lemma 4.1. *Suppose that the infimum in (2.13) is attained by some $\vartheta^{(c)} \in \Theta$. Then this process satisfies*

$$(4.1) \qquad E\left[H - c - G_T\left(\vartheta^{(c)}\right)\right] = \frac{E(\tilde{D}H) - c}{E(\tilde{D}^2)} \qquad \textit{and}$$

$$(4.2) \qquad E\left[H - c - G_T\left(\vartheta^{(c)}\right)\right]^2 = \frac{c^2 - 2c \cdot E(\tilde{D}H)}{E(\tilde{D}^2)} + E\left[\pi^2(H)\right],$$

in the notation of (3.9) and (3.17).

Proof of (4.1) : The assumption implies that, for any given $\xi \in \Theta$, the function

$$\begin{aligned} f(\varepsilon) &\triangleq E\left[H - c - G_T\left(\vartheta^{(c)} + \varepsilon\xi\right)\right]^2 \\ &= \varepsilon^2 \cdot E\left[G_T^2(\xi)\right] - 2\varepsilon \cdot E\left[\left(H - c - G_T\left(\vartheta^{(c)}\right)\right) \cdot G_T(\xi)\right] + V(c) \end{aligned}$$

attains its minimum over $\mathbb{R}$ at $\varepsilon = 0$. This gives $f'(0) = 0$, or equivalently

$$(4.3) \qquad E\left[\left(H - c - G_T\left(\vartheta^{(c)}\right)\right) \cdot G_T(\xi)\right] = 0, \qquad \forall\, \xi \in \Theta.$$

Let us also notice that the mapping

$$(4.4) \qquad D \mapsto E(\tilde{D}D) \text{ is constant on } \mathcal{D},$$

since we have

$$\begin{aligned} E(\tilde{D}^2) - E(\tilde{D}D) = E\left[\tilde{D}(\tilde{D} - D)\right] &= E\left[\pi(1)(\tilde{D} - D)\right]/E\left[\pi(1)\right] \\ &= E(\tilde{D} - D)/E\left[\pi(1)\right] = 0 \end{aligned}$$

thanks to (3.9).

Now denote $\gamma \triangleq E\left[H - c - G_T\left(\vartheta^{(c)}\right)\right]$. If $\gamma = 0$, then the random variable

$$D_1 \triangleq \tilde{D} + \left(H - c - G_T\left(\vartheta^{(c)}\right)\right)$$

belongs to $\mathcal{D}$ by virtue of (4.3), and (4.4) implies

$$0 = E\left[\tilde{D}(D_1 - \tilde{D})\right] = E\left[\tilde{D}\left(H - c - G_T\left(\vartheta^{(c)}\right)\right)\right] = E(\tilde{D}H) - c,$$

so that (4.1) holds. If $\gamma \neq 0$, then $D_2 \triangleq \left[H - c - G_T\left(\vartheta^{(c)}\right)\right]/\gamma$ is in $\mathcal{D}$, and by (4.4) once again:

$$E(\tilde{D}^2) = E(\tilde{D}D_2) = \frac{1}{\gamma}\left(E(\tilde{D}H) - c\right),$$

and so (4.1) holds in this case too. □

• *Proof of (4.2)* : From (4.3), the random variable $H - c - G_T\left(\vartheta^{(c)}\right)$ belongs to the closed subspace $(G_T(\Theta))^\perp$ of (2.11), so we have

$$\begin{aligned}
&E\left[H - c - G_T\left(\vartheta^{(c)}\right)\right]^2 \\
&= E\left(\left[H - c - G_T\left(\vartheta^{(c)}\right)\right] \cdot \left[H - \pi(H) + \pi(H) - c - G_T\left(\vartheta^{(c)}\right)\right]\right) \\
&= E\left(\left[H - c - G_T\left(\vartheta^{(c)}\right)\right] \cdot \left[\pi(H) - c\right]\right) \\
&= E\left[\pi^2(H)\right] - c\,E\left[\pi(H)\right] - c\,\frac{E(\tilde{D}H) - c}{E(\tilde{D}^2)}
\end{aligned}$$

thanks to (4.3), (3.9) and (4.1). The equation (4.2) now follows from

$$E(\tilde{D}H) = E(\tilde{D}^2) \cdot E\left[\pi(H)\right]. \tag{4.5}$$

In order to check (4.5), recall (3.17), (3.20) and use (3.9)$'$ repeatedly, to wit:

$$\frac{E(\tilde{D}H)}{E(\tilde{D}^2)} = E\left[H\pi(1)\right] = E\left[\pi(H)\pi(1)\right] = E\left[\pi(H)\right]. \tag{4.6}$$

□

Theorem 4.1. *(i) Suppose that there exists some $\vartheta^{(c)} \in \Theta$ which attains the infimum in (2.13). Then this $\vartheta^{(c)}$ satisfies*

$$H - c - G_T\left(\vartheta^{(c)}\right) = \pi(H - c), \tag{4.7}$$

and there is no duality gap in (3.7), namely

(4.8)
$$\begin{aligned}
V(c) = \tilde{V}(c) &= E\left[\pi^2(H - c)\right] \\
&= \frac{\left(E(\tilde{D}H) - c\right)^2}{E(\tilde{D}^2)} + E\left[\pi^2(H)\right] - \frac{\left(E\pi(H)\right)^2}{E\left[\pi(1)\right]}.
\end{aligned}$$

(ii) Conversely, suppose there exists some $\vartheta^{(c)} \in \Theta$ that satisfies (4.7); then this $\vartheta^{(c)}$ attains the infimum in (2.13), and the equalities of (4.8) hold.

• *Proof of (4.8), Part (i)* : Under the assumption of (i), we claim that

$$
\begin{aligned}
V(c) &= E\left(H - c - G_T\left(\vartheta^{(c)}\right)\right)^2 \\
&= \frac{\left(E(\tilde{D}H) - c\right)^2}{E(\tilde{D}^2)} + E\left[\pi^2(H)\right] - (E\pi(H))^2 \cdot E(\tilde{D}^2) \qquad (4.9)\\
&= E\left[\pi^2(H-c)\right] \;=\; \tilde{V}(c),
\end{aligned}
$$

which clearly proves (4.8) in light of the last equality in (3.20). Indeed, the first equality in (4.9) holds by assumption, whereas the second is a consequence of (4.2), (4.5). The third equality is a consequence of (4.6), (3.20) and (4.2), thanks to the simple computation

$$
\begin{aligned}
E\left[\pi^2(H-c)\right] &= E\left(\pi(H) - c \cdot \pi(1)\right)^2 \\
&= E\left[\pi^2(H)\right] + c^2 \cdot E\left[\pi^2(1)\right] - 2c \cdot E\left[\pi(1)\pi(H)\right] \\
&= E\left[\pi^2(H)\right] + \frac{c^2 - 2c \cdot E(\tilde{D}H)}{E(\tilde{D}^2)} .
\end{aligned}
$$

Finally, the last equality in (4.9) is just (3.12).

• *Proof of (4.7), Part (i); $c \neq \hat{c}$* : Let us write (3.2) with $\vartheta \equiv \vartheta^{(c)}$, $D \equiv \tilde{D}_c$ as in (3.13), and

$$
k \;\equiv\; k_c \;\triangleq\; k_{\tilde{D}_c, c} \;=\; \frac{E(\tilde{D}_c H) - c}{E(\tilde{D}_c^2)}
$$

as in (3.6): namely,

$$
\begin{aligned}
&\left(H - c - G_T\left(\vartheta^{(c)}\right)\right)^2 + 2k_c\tilde{D}_c\left(c + G_T\left(\vartheta^{(c)}\right)\right) \qquad (4.10)\\
&\qquad\qquad \geq\; 2k_c\tilde{D}_c H - \left(k_c\tilde{D}_c\right)^2, \text{ a.s.}
\end{aligned}
$$

Taking expectations in (4.10), and recalling the optimality of $\vartheta^{(c)}$ as well as Proposition 3.1, we obtain

$$
\begin{aligned}
V(c) &= E\left(H - c - G_T\left(\vartheta^{(c)}\right)\right)^2 \\
&\geq -k_c^2\, E(\tilde{D}_c^2) + 2k_c\left(E(\tilde{D}_c H) - c\right) \qquad (4.11)\\
&= \frac{\left(E(\tilde{D}_c H) - c\right)^2}{E(\tilde{D}_c^2)} = \tilde{V}(c).
\end{aligned}
$$

But from (4.8) we know that (4.11) actually holds as equality, which means that the left-hand side and the right-hand side of (4.10) have the same expectation. In other words, (4.10) must hold as equality, which we know happens only if (3.3) holds, namely only if

$$
H - c - G_T\left(\vartheta^{(c)}\right) = k_c\,\tilde{D}_c = \frac{E(\tilde{D}_c H) - c}{E(\tilde{D}_c^2)}\,\frac{\pi(H-c)}{E\left[\pi(H-c)\right]} = \pi(H-c)\,,
$$

holds a.s., thanks to (3.19).

• *Proof of (4.7), part (i); $c = \hat{c}$* : In this case we shall need a new kind of duality, namely with

$$
\mathcal{L} \triangleq \left\{L \in (G_T(\vartheta))^{\perp} \;/\; E(L) = 0\right\} \qquad (4.12)
$$

replacing the space $\mathcal{D}$ of (2.9); the elements of $\mathcal{L}$ will be the dual (adjoint) variables in this new duality. We begin by writing (3.1) with $x = c + G_T(\vartheta)$, $y = 2L$ for arbitrary $\vartheta \in \Theta$, $L \in \mathcal{L}$:

$$
(H - c - G_T(\vartheta))^2 + 2L\,(c + G_T(\vartheta)) \;\geq\; 2LH - L^2, \qquad (4.13)
$$

with equality if and only if

$$
H - c - G_T(\vartheta) = L \qquad (4.14)
$$

holds a.s. Taking expectations in (4.13), we obtain

$$
\begin{aligned}
E\,(H - c - G_T(\vartheta))^2 &\geq E\left[2L(H-c) - L^2\right] \\
&= E\left[2L\cdot\pi(H-c) - L^2\right] \qquad (4.15)\\
&= E\left[\pi^2(H-c)\right] - E\,(L - \pi(H-c))^2 .
\end{aligned}
$$

This suggests that we should read (4.13)-(4.15) with $\vartheta \equiv \vartheta^{(c)}$, the element of Θ that attains the infimum in (2.13) and is thus optimal for

Problem 2.2, and $L \equiv \tilde{L} \triangleq \pi(H-c)$, noting that $E(\tilde{L}) = 0$ since $c = \hat{c} = E[\pi(H)]/E[\pi(1)]$. With these choices, the left-most member of (4.15) becomes

$$E\left(H - c - G_T\left(\vartheta^{(c)}\right)\right)^2 = V(c),$$

whilst its right-most member is $E\left[\pi^2(H-c)\right] = \tilde{V}(c)$. From (4.8) we know that these two quantities are equal, so the two sides of (4.13) have the same expectation. This means that (4.13) must holds as equality, which happens only if (4.14) is valid, namely

$$H - c - G_T\left(\vartheta^{(c)}\right) = \pi(H-c), \quad \text{a.s.}$$

• *Proof of Part (ii)* : Suppose there exists some $\vartheta^{(c)} \in \Theta$ that satisfies (4.7); then

$$E\left(H - c - G_T\left(\vartheta^{(c)}\right)\right)^2 = E\left[\pi^2(H-c)\right] = \tilde{V}(c)$$

from Proposition 3.1. But we also have

$$\tilde{V}(c) \leq V(c) \leq E\left(H - c - G_T\left(\vartheta^{(c)}\right)\right)^2,$$

thanks to (3.7) and (2.13). In other words, these last two inequalities are valid as equalities, $\vartheta^{(c)}$ attains the infimum in (2.13), and (4.8) holds.

□

Remark 4.1 : The case of $G_T(\Theta)$ closed.
If $G_T(\Theta)$ is closed in $\mathbf{L}^2(P)$, then the infimum in (2.13) is attained, as was assumed in Lemma 4.1 and in Theorem 4.1(i). In this case we have of course $G_T(\Theta) = (G_T(\Theta))^{\perp\perp}$, and every $H \in \mathbf{L}^2(P)$ admits a decomposition of the form

$$H = \pi(H) + G_T(\xi^H) \qquad \text{for some } \xi^H \in \Theta. \tag{4.16}$$

In particular, there exists $\xi^1 \in \Theta$ so that (3.21) holds, and thus

$$H - c - \pi(H-c) = [H - \pi(H)] - c\,[1 - \pi(1)] = G_T\left(\xi^H - c\xi^1\right).$$

Comparing this expression with (4.7), we conclude that (4.7) is satisfied with the choice

$$\vartheta^{(c)} = \xi^H - c\xi^1. \tag{4.17}$$

According to Theorem 4.1(ii), *the process $\vartheta^{(c)} \in \Theta$ of (4.17) is then optimal for Problem 2.2, and (4.8) holds.*

Example 4.1. Föllmer-Schweizer decomposition. The assertion at the end of Remark 4.1 remains valid even if $G_T(\Theta)$ is not closed in $\mathbf{L}^2(P)$, provided that (3.21) and (4.16) still hold. Consider, for example, the case of a semimartingale $X(\cdot) \in \mathcal{S}^2(P)$ and of a random variable $H \in \mathbf{L}^2(P)$ which admits the so-called "Föllmer-Schweizer decomposition"; this means that H can be written in the form $H = h + G_T(\zeta^H) + L^H(T)$ of (2.6), for $h = E(H)$, some $\zeta^H \in \Theta$, and some martingale $L^H(\cdot) \in \mathcal{M}_0^2(P)$ that satisfies the property (2.7).

Suppose that (3.21) is also satisfied; then $\pi(H) = h\pi(1) + L^H(T)$ and we have $H - \pi(H) = h\big(1 - \pi(1)\big) + G_T\big(\zeta^H\big) = G_T\big(h\xi^1 + \zeta^H\big)$, so we may take $\xi^H \equiv h\xi^1 + \zeta^H$ in (4.16) and

$$\vartheta^{(c)} \equiv (h-c)\cdot\xi^1 + \zeta^H \tag{4.17$'$}$$

in (4.7). The process $\vartheta^{(c)}$ of (4.17)$'$ is then optimal for the Problem 2.2, i.e., attains the infimum in (2.13), which can be readily computed as

$$V(c) = (h-c)^2 \cdot E[\pi(1)] + E\left(L^H(T)\right)^2 .$$

This simple result may be compared with Theorem 3 and Proposition 18 in Schweizer (1994).

We are now in a position to discuss the solution of Problem 2.1 as well.

Theorem 4.2. *Suppose that $G_T(\Theta)$ is closed in $\mathbf{L}^2(P)$. Then the value of Problem 2.1 is given as*

$$V = V(\hat{c}) = E\left[\pi^2(H)\right] - \frac{\left(E\left[\pi(H)\right]\right)^2}{E\left[\pi(1)\right]} \tag{4.18}$$

with the notation

$$\hat{c} = \frac{E\left[\pi(H)\right]}{E\left[\pi(1)\right]} = E(\tilde{D}H) = E\left(\frac{d\tilde{Q}}{dP}H\right) \tag{4.19}$$

of (3.13). Furthermore, the infimum in (2.4) is attained by the pair $(\hat{c}, \hat{\vartheta})$ with $\hat{c}$ as in (4.19) and with

$$\hat{\vartheta} \triangleq \vartheta^{(\hat{c})} = \xi^H - \hat{c}\xi^1 . \tag{4.20}$$

Proof : Immediate from Theorem 4.1 and Remark 4.1, when it is observed that the number $\hat{c}$ of (4.19) minimizes the expression of (4.8) over $c \in \mathbb{R}$. □

Note that when the signed measure $\tilde{Q}$ of (3.8) is a probability measure (i.e., when $P[\pi(1) \geq 0] = 1$), the quantity of (4.19) is just *the expectation of the random variable* H *under the dual-optimal measure* $\tilde{Q}$. Sufficient conditions are spelled out in the next section.

Remark 4.2 : Variance Minimization. If $G_T(\Theta)$ is closed in $\mathbf{L}^2(P)$, then the process $\hat{\vartheta}$ of (4.20) also

$$\text{minimizes} \quad \mathrm{Var}\,(H - G_T(\vartheta))\,, \quad \text{over all } \vartheta \in \Theta. \tag{4.21}$$

This is because for any $\vartheta \in \Theta$, and with $c_\vartheta \triangleq E\,[H - G_T(\vartheta)]$, we have:

$$\begin{aligned} \mathrm{Var}\,(H - G_T(\vartheta)) &= E\,(H - c_\vartheta - G_T(\vartheta))^2 \\ &\geq E\left(H - \hat{c} - G_T(\hat{\vartheta})\right)^2 = \mathrm{Var}\left(H - G_T(\hat{\vartheta})\right), \end{aligned}$$

from Theorem 4.2.

More generally, for any given $c \in \mathbb{R}$, the process $\vartheta^{(c)} \in \Theta$ of (4.17) has the *"mean-variance efficiency" property*

$$\left\{ \begin{array}{c} \mathrm{Var}\left(H - G_T(\vartheta^{(c)})\right) \leq \mathrm{Var}\,(H - G_T(\vartheta))\,, \quad \text{for any} \\ \vartheta \in \Theta \ \text{ that satisfies } \ E\,[H - G_T(\vartheta)] = E\left[H - G_T\left(\vartheta^{(c)}\right)\right] \end{array} \right\}. \tag{4.22}$$

Indeed, let $\mu_c \triangleq E\left[H - G_T\left(\vartheta^{(c)}\right)\right]$ and observe that, for any $\vartheta \in \Theta$ with $E\,[H - G_T(\vartheta)] = \mu_c$, we have

$$\begin{aligned} &\mathrm{Var}\,(H - G_T(\vartheta)) \\ &= \mathrm{Var}\,(H - c - G_T(\vartheta)) = E\,(H - c - G_T(\vartheta))^2 - (\mu_c - c)^2 \\ &\geq E\left(H - c - G_T(\vartheta^{(c)})\right)^2 - \left(E\left(H - c - G_T(\vartheta^{(c)})\right)\right)^2 \\ &= \mathrm{Var}\left(H - c - G_T\left(\vartheta^{(c)}\right)\right) = \mathrm{Var}\left(H - G_T\left(\vartheta^{(c)}\right)\right). \end{aligned}$$

Remark 4.3 : Mean-Variance Frontier. Suppose that $G_T(\Theta)$ is closed in $\mathbf{L}^2(P)$, that we have $P\left[\pi(1) \neq E(\pi(1))\right] > 0$; this implies $E(\tilde{D}^2) > 1$ in (3.5), hence $E[\pi(1)] < 1$. For some given $m \in \mathbb{R}$, consider the following problem:

$$\left\{ \begin{array}{c} \text{To minimize the variance } \mathrm{Var}\,(H - G_T(\vartheta))\,, \\ \text{over } \ \vartheta \in \Theta \ \text{ with } \ E\,[H - G_T(\vartheta)] = m. \end{array} \right\} \tag{4.23}$$

In view of the property (4.22), it suffices to show that we can find $c \equiv c_m$ such that

$$E\left[H - G_T\left(\vartheta^{(c)}\right)\right] = m. \tag{4.24}$$

Then the solution of the problem (4.23) will be given by $\vartheta^{(c)} \in \Theta$ as in (4.17), with $c \equiv c_m$. Now from (4.1) we have $E\left[H - G_T\left(\vartheta^{(c)}\right)\right] = c + \frac{E(\tilde{D}H)-c}{E(\tilde{D}^2)}$, so that (4.24) amounts to

$$c \;=\; c_m \;\triangleq\; \frac{m \cdot E(\tilde{D}^2) - E(\tilde{D}H)}{E(\tilde{D}^2) - 1} \;=\; \frac{m - E\left[\pi(H)\right]}{1 - E\left[\pi(1)\right]},$$

thanks to (4.6) and (3.20). We take these two Remarks 4.2, 4.3 from Schweizer (1994, 1996).

§5. A Mathematical Finance Interpretation

The Problems 2.1, 2.2 have an interesting interpretation in the context of Mathematical Finance, when one interprets the components $X_i(\cdot)$ of the semimartingale in (2.1) as the (discounted) prices of several risky assets in a financial market. In this context, $\vartheta_i(t)$ represents the number of shares in the corresponding i^{th} asset, held by an investor at time $t \in [0,T]$. The resulting process $\vartheta \in \Theta$ stands then for the investor's (self-financing) *trading strategy*, and $G_t(\vartheta) = \int_0^t \vartheta'(s)\,dX(s) = \sum_{i=1}^d \int_0^t \vartheta_i(s)\,dX_i(s)$ for the (discounted) *gains-from-trade* associated with the strategy ϑ by time t.

Suppose now that the investor faces a contingent claim (liability) H at the end T of the time-horizon $[0,T]$. Starting with a given initial capital c, and using a trading strategy $\vartheta \in \Theta$, the investor seeks to replicate this contingent claim H *as faithfully as possible*, in the sense of minimizing the expected squared-error loss $E\left(H - c - G_T(\vartheta)\right)^2$. This leads us to Problem 2.2. When the determination of the "right" initial capital c is also part of the problem, one is led naturally to the formulation of Problem 2.1. Similarly, one may consider minimizing the variance of the discrepancy $H - c - G_T(\vartheta)$ over all trading strategies $\vartheta \in \Theta$ (problem of (4.21)), or just over those strategies that guarantee a given "mean-gains-from-trade" level $E\left[G_T(\vartheta)\right] = E(H) - m$ (problem of (4.23)).

If one decides to stick with this interpretation, it makes sense to ask whether the model for the financial assets represented by (2.1) admits *arbitrage opportunities*; these are trading strategies $\vartheta \in \Theta$ with $P[G_T(\vartheta) \geq 0] = 1$ and $P[G_T(\vartheta) > 0] > 0$. To this effect, let A be an increasing, predictable and RCLL process with values in $[0,\infty)$ and $A(0) = 0$, $\langle M \rangle_i \ll A$ for $i = 1, \cdots, d$, and suppose that the processes $M = (M_1, \ldots, M_d)$ and $B = (B_1, \ldots, B_d)$, in the decomposition

$X = X(0) + M + B$ of (2.1) for the semimartingale $X = (X_1, \ldots, X_d)$, satisfy

$$B_i(\cdot) \ll \langle M \rangle_i(\cdot) \tag{5.1}$$

$$B_i(t) = \int_0^t \gamma_i(s)\, dA(s), \qquad 0 \le t \le T \tag{5.2}$$

$$\langle M_i, M_j \rangle(t) = \int_0^t \sigma_{ij}(s)\, dA(s), \qquad 0 \le t \le T \tag{5.3}$$

for $i = 1, \cdots, d$ and $j = 1, \cdots, d$. Here $\gamma(\cdot) = (\gamma_1(\cdot), \cdots, \gamma_d(\cdot))'$ and $\sigma(\cdot) = \{\sigma_{ij}(\cdot)\}_{1 \le i,j \le d}$ are suitable predictable processes that satisfy

$$\sigma(t)\lambda(t) = \gamma(t), \qquad \text{a.e. } t \in [0,T] \tag{5.4}$$

almost surely, for some $\lambda(\cdot) = (\lambda_1(\cdot), \cdots, \lambda_d(\cdot))'$ in the space

(5.5)

$$\mathcal{L}^2(M) \triangleq \left\{ \vartheta : [0,T] \to \mathbb{R}^d \text{ predictable} \Big/ \right.$$
$$\left. E\int_0^T \vartheta'(s)\sigma(s)\vartheta(s)\, dA(s) = E\left(\int_0^T \vartheta'(s)\, dM(s)\right)^2 < \infty \right\}.$$

Following Schweizer (1996), we shall refer to (5.1)-(5.4) as the *structure conditions* on the semimartingale $X(\cdot)$.

Under these conditions, it can be shown that the semimartingale $X(\cdot)$ does not admit arbitrage opportunities, and that we have equality $\mathbf{P}_s(\Theta) = \mathbf{P}_s^2(X)$ in Remark 2.2 (cf. Ansel & Stricker (1992); Schweizer (1995); and Schweizer (1996), Lemma 12). If, in addition, $X(\cdot)$ has continuous paths, then it can be shown that the variance-optimal martingale measure $\tilde{Q}$ of (3.18) is nonnegative, namely a probability measure:

$$P[\tilde{D} \ge 0] = 1 \quad \text{and} \quad \tilde{Q}(\Omega) = E(\tilde{D}) = 1 \tag{5.6}$$

in (3.17), (3.18). This $\tilde{Q}$ is in fact equivalent to P (i.e., $P[\pi(1) > 0] = 1$), under the extra assumption

$$\left\{ \begin{array}{c} X(\cdot) \text{ is a } Q\text{−local martingale under some} \\ \text{probability measure } Q \sim P \text{ with } (dQ/dP) \in \mathbf{L}^2(P) \end{array} \right\} \tag{5.7}$$

(cf. Schweizer (1996), Theorem 13; Delbaen & Schachermayer (1996), Theorem 1.3). This condition (5.7) also implies that

$$\text{the mapping } \vartheta \mapsto G_T(\vartheta) \text{ is injective;} \tag{5.8}$$

cf. [DMSSS], Lemma 3.5.

Always under the structure conditions (5.1)-(5.4) on the semimartingale $X(\cdot)$ of (2.1), consider now the so-called *mean-variance tradeoff* process

$$K(t) \triangleq \int_0^t \lambda'(s)\,dB(s) = \int_0^t \lambda'(s)\sigma(s)\lambda(s)\,dA(s) = \left\langle \int \lambda'\,dM \right\rangle(t), \quad 0 \le t \le T.$$

If this process is P-a.s. *bounded*, then

$$\Theta = \mathcal{L}^2(M) \tag{5.9}$$

in the notation of (2.2), (5.5) and

$$G_T(\Theta) \text{ is closed in } \mathbf{L}^2(P) \tag{5.10}$$

([Ph.R.S.], Corollary 4 and below; Schweizer (1996), Lemma 12). See also [DMSSS] where conditions both necessary and sufficient for (5.10) are presented.

Remark 5.1 : In the one-dimensional case $d = 1$, the structure conditions (5.1)-(5.4) are satisfied if there exists a process $\lambda \in \mathcal{L}^2(M)$ with

$$X(t) = X(0) + M(t) + \int_0^t \lambda(s)\,d\langle M\rangle(s), \qquad 0 \le t \le T. \tag{5.11}$$

In this case, the mean-variance tradeoff process of (5.9) takes the form

$$K(t) = \int_0^t \lambda^2(s)\,d\langle M\rangle(s). \tag{5.12}$$

Remark 5.2 : Suppose that $X(\cdot)$ has continuous paths, and that (5.11) and (5.7) hold. If the random variable $H \in \mathbf{L}^2(P)$ is of the form $H = h + G_T(\zeta^H)$ in (2.5), then

$$H - \pi(H) = G_T(h\xi^1 + \zeta^H)$$

and the injectivity property (5.8) allows us to make the identifications

$$\xi^H = h\xi^1 + \zeta^H, \qquad \vartheta^{(c)} = (h - c)\cdot\xi^1 + \zeta^H \tag{5.13}$$

in (4.16) and (4.17), respectively.

More generally, under the assumptions of Remark 5.2 but now for *any* $H \in \mathbf{L}^2(P)$, we have the Kunita-Watanabe decomposition under the variance-optimal probability measure $\tilde{Q}$ of (3.18), namely

$$\tilde{E}[H \mid \mathcal{F}(t)] = \tilde{h} + G_t(\tilde{\zeta}^H) + \tilde{L}^H(t), \quad 0 \le t \le T \tag{5.14}$$

(cf. Ansel & Stricker (1993), or Theorem 3 in Rheinländer & Schweizer (1997)). Here $\tilde{E}$ denotes expectation with respect to the probability measure $\tilde{Q}$, $\tilde{h} \triangleq \tilde{E}(H) = E(\tilde{D}H)$, $\tilde{\zeta}^H \in \Theta$, and $\tilde{L}^H \in \mathcal{S}^2(P)$ is a $\tilde{Q}$-martingale with $\tilde{L}^H(0) = 0$ and

$$\left\langle \tilde{L}^H, X_i \right\rangle (\cdot) \equiv 0, \qquad \forall\, i = 1, \cdots, d. \tag{5.15}$$

On the other hand, since $\tilde{L}^H(T)$ belongs to the space $\mathbf{L}^2(P)$, we also have its decomposition

$$\tilde{L}^H(T) = G_T(\tilde{\vartheta}^H) + \pi(\tilde{L}^H(T)) \tag{5.16}$$

from the closedness of $G_T(\Theta)$, where $\tilde{\vartheta}^H \in \Theta$ is such that

$$E\left[\left(\tilde{L}^H(T) - G_T(\tilde{\vartheta}^H)\right) \cdot G_T(\vartheta)\right] = 0, \qquad \forall\, \vartheta \in \Theta. \tag{5.17}$$

Back in (5.14), this gives $H = \tilde{h} + G_T(\tilde{\zeta}^H + \tilde{\vartheta}^H) + \pi(\tilde{L}^H(T))$, so that

$$\begin{aligned}(H - c) - \pi(H - c) &= (\tilde{h} - c)(1 - \pi(1)) + G_T(\tilde{\zeta}^H + \tilde{\vartheta}^H)\\ &= G_T\left((\tilde{h} - c)\xi^1 + \tilde{\zeta}^H + \tilde{\vartheta}^H\right)\end{aligned}$$

from (3.21)′. Again, the injectivity property (5.8) allows us to make the identification

$$\vartheta^{(c)} = (\tilde{h} - c)\xi^1 + \tilde{\zeta}^H + \tilde{\vartheta}^H \tag{5.18}$$

in (4.17). Finally, let us consider the positive $\tilde{Q}$-martingale

$$\begin{aligned}\tilde{D}(t) \triangleq \tilde{E}[\tilde{D} \,|\, \mathcal{F}(t)] &= E(\tilde{D}^2) + G_t(\tilde{\xi})\\ &= E(\tilde{D}^2)\left[1 - G_t(\xi^1)\right], \quad 0 \le t \le T\end{aligned} \tag{5.19}$$

obtained by taking conditional expectations in (3.22) under $\tilde{Q}$.

We are now in a position to identify the process $\tilde{\vartheta}^H$ appearing in (5.16), (5.18) and state the following result, which simplifies and generalizes Theorems 5, 6 of Rheinländer & Schweizer (1997).

Theorem 5.1. *Suppose that the semimartingale $X(\cdot) \in \mathcal{S}^2(P)$ has continuous paths, and that (5.7), (5.11) hold. Then the optimal process $\vartheta^{(c)} \in \Theta$ for Problem 2.2 takes the form*

$$\vartheta^{(c)} = \tilde{\zeta}^H + \left[(E(\tilde{D}H) - c) + E(\tilde{D}^2)\int_0^{\cdot} \frac{d\tilde{L}^H(s)}{\tilde{D}(s)}\right] \cdot \xi^1 \tag{5.20}$$

in the notation of (5.14), (5.19).

Sketch of Proof: The $\tilde{Q}$-martingales $\tilde{N}(t) \triangleq \int_0^t (1/\tilde{D}(s))\, d\tilde{L}^H(s)$, $0 \le t \le T$ and $\tilde{D}(\cdot)$ of (5.19) are orthogonal, since

$$\langle \tilde{N}, \tilde{D}\rangle(\cdot) = \sum_{i=1}^{d} \int_0^{\cdot} \frac{\tilde{\xi}_i(s)}{\tilde{D}(s)}\, d\langle \tilde{L}^H, X_i\rangle(s) \equiv 0$$

from (5.15) and (5.19). Thus, integration by parts gives

(5.21)
$$\begin{aligned} \tilde{L}^H(T) &= \int_0^T \tilde{D}(s)\, d\tilde{N}(s) = \tilde{D}(T)\, \tilde{N}(T) - \int_0^T \tilde{N}(s)\tilde{\xi}'(s)\, dX(s) \\ &= \tilde{D}\tilde{N}(T) - G_T(\tilde{N}\tilde{\xi}), \end{aligned}$$

and transforms (5.17) into

$$E\left[G_T(\tilde{\vartheta}^H + \tilde{N}\tilde{\xi}) \cdot G_T(\vartheta)\right] = \tilde{E}\left[\tilde{N}(T) G_T(\vartheta)\right], \qquad \forall\, \vartheta \in \Theta. \tag{5.22}$$

But the right-hand-side of (5.22) vanishes, since

$$\begin{aligned} \tilde{E}\left[\tilde{N}(T)G_T(\vartheta)\right] &= \tilde{E}\left[\int_0^T \frac{d\tilde{L}^H(s)}{\tilde{D}(s)} \cdot \sum_{i=1}^{d} \int_0^T \vartheta_i(s)\, dX_i(s)\right] \\ &= \tilde{E} \sum_{i=1}^{d} \int_0^T \frac{\vartheta_i(s)}{\tilde{D}(s)}\, d\langle \tilde{L}^H, X_i\rangle(s) = 0 \end{aligned}$$

thanks to (5.15). Thus the left-hand-side of (5.22) also vanishes for every $\vartheta \in \Theta$, which suggests

$$\tilde{\vartheta}^H = -\tilde{N}\tilde{\xi} = E(\tilde{D}^2)\, \xi^1 \int_0^{\cdot} \frac{d\tilde{L}^H(s)}{\tilde{D}(s)}$$

and leads to (5.20) after substitution into (5.18). □

In order to justify the legitimacy of the above argument, particularly the steps leading to (5.22), one needs to show that the random variable $\sup_{0 \le t \le T} \left|\tilde{D}(t) \int_0^t (1/\tilde{D}(s))\, d\tilde{L}^H(s)\right|$ belongs to $\mathbf{L}^2(P)$; this is carried out on pp. 1820-1823 of Rheinländer & Schweizer (1997).

References

[1] ANSEL, J.-P. & STRICKER, C. (1992) Lois de martingale, densités et décomposition de Föllmer-Schweizer. *Ann. Inst. Henri Poincaré* **28**, 375-392.

[2] ARAI, T. (2002) Mean-Variance Hedging for Discontinuous Semimartingales. *Tokyo J. Mathematics,* to appear.

[3] ARAI, T. (2002) Mean-Variance Hedging for General Semimartingales. *Preprint,* Tokyo University of Science.

[4] CLARK, J.M.C. (1970) The representation of functionals of Brownian motion as stochastic integrals. *Annals of Mathematical Statistics* **41**, 1282-1295. Correction, 1778.

[5] DUFFIE, D. & RICHARDSON, H.R. (1991) Mean-variance hedging in continuous time. *Annals of Applied Probability* **1**, 1-15.

[6] DELBAEN, F., MONAT P., SCHACHERMAYER, W., SCHWEIZER, M. & STRICKER, C. [DMSSS] (1997) Weighted norm inequalities and hedging in incomplete markets. *Finance & Stochastics* **1**, 181-227.

[7] DELBAEN, F. & SCHACHERMAYER, W. (1996) The variance-optimal martingale measure for the continuous processes. *Bernoulli* **2**, 81-105.

[8] FÖLLMER, H. & SCHWEIZER, M. (1991) Hedging of contingent claims under incomplete information. In *Applied Stochastic Analysis* (M.H.A. Davis and R.J. Elliott, eds.), *Stochastics Monographs* **5**, 389-414. Gordon & Breach, New York.

[9] FÖLLMER, H. & SONDERMANN, D. (1986) Hedging of non-redundant contingent claims. In *Contributions to Mathematical Economics: Essays in Honor of G. Debreu* (W. Hildenbrand and A. MasColell, eds.), 205-223. North-Holland, Amsterdam.

[10] GOURIÉROUX C., LAURENT J.P. & PHAM, H. [G.L.Ph.] (1998) Mean-Variance Hedging and Numéraire. *Mathematical Finance* **8**, 179-200.

[11] GRANDITS, P. (1999) The p-optimal martingale measure and its asymptotic relation with the minimal-entropy martingale measure. *Bernoulli* **5**, 225-247.

[12] HIPP, C. (1993) Hedging general claims. *Proceedings of the 3rd AFIR Colloquium, Rome* **2**, 603-613.

[13] KARATZAS, I. & SHREVE, S. E. (1991) *Brownian Motion and Stochastic Calculus.* Second Edition, Springer-Verlag, New York.

[14] KUNITA, H. & WATANABE, S. (1967) On square integrable martingales. *Nagoya Mathematical Journal* **30**, 209-245.

[15] LAURENT, J. & PHAM, H. (1999) Dynamic programming and mean-variance hedging. *Finance & Stochastics* **3**, 83-110.

[16] MARKOWITZ, H. (1952) Portfolio Selection. *Journal of Finance* **8**, 77-91.

[17] MARKOWITZ, H. (1959) *Portfolio Selection: Efficient Diversification of Investment.* J. Wiley & Sons, New York. Second Edition by B. Blackwell, Oxford, 1991.

[18] MONAT, P. & STRICKER, C. (1995) Föllmer-Schweizer decomposition and mean-variance hedging for general claims. *Annals of Probability* **23**, 605-628.

[19] MÜLLER, S.M. (1985) Arbitrage Pricing of Contingent Claims. *Lecture Notes in Economics and Mathematical Systems* **254**. Springer-Verlag, New York.

[20] PHAM, H., RHEINLÄNDER, T. & SCHWEIZER, M. [Ph.R.S] (1998) Mean-variance hedging for continuous processes : new proofs and examples. *Finance & Stochastics* **2**, 173-198.

[21] RHEINLÄNDER, T. & SCHWEIZER, M. (1997) On $\mathbf{L}^2$-projections on a space of stochastic integrals, *Annals of Probability* **25**, 1810-1831.

[22] SCHÄL, M. (1992) On quadratic cost criteria for option hedging. *Mathematics of Operations Research* **19**, 121-131.

[23] SCHWEIZER, M. (1992) Mean-variance hedging for general claims. *Annals of Applied Probability* **2**, 171-179.

[24] SCHWEIZER, M. (1994) Approximating random variables by stochastic integrals. *Annals of Probability* **22**, 1536-1575.

[25] SCHWEIZER, M. (1995a) On the minimal martingale measure and the Föllmer-Schweizer decomposition. *Stochastic Analysis & Applications* **13**, 573-599.

[26] SCHWEIZER, M. (1995b) Variance-optimal hedging in discrete time. *Mathematics of Operations Research* **20**, 1-32.

[27] SCHWEIZER, M. (1996) Approximation pricing and the variance-optimal martingale measure. *Annals of Probability* **24**, 206-236.

[28] SCHWEIZER, M. (2001) A guided tour through quadratic hedging approaches. In *Handbook of Mathematical Finance: Option Pricing, Interest Rates and Risk Management* (E.Jouini, J.Cvitanic and M.Musiela, eds.), 632-670. Cambridge University Press.

Chunli Hou
Nomura Securities International
2 World Financial Center, Building B
New York, NY 10281
`chou@us.nomura.com`

Ioannis Karatzas
Departments of Mathematics and Statistics
619 Mathematics Building
Columbia University
New York, NY 10027
`ik@math.columbia.edu`

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 167–187

Quadratic Wiener Functionals, Kalman-Bucy Filters, and the KdV Equation

Nobuyuki Ikeda and Setsuo Taniguchi

Dedicated to Professor Kiyosi Itô on his 88th birthday

Abstract.

Soliton solutions and the tau function of the KdV equation are studied within the stochastic analytic framework. A key role is played by the Itô formula and the Cameron-Martin transformation.

§ Introduction

In this paper, we investigate the Korteweg-de Vries (KdV) equation within the framework of stochastic analysis. We shall study soliton solutions with the help of the Itô formula, whose original form was achieved in 1942 ([9]). The Cameron-Martin transformation, which was established in the early 1940's ([2, 3]), also plays a key role.

Let $x > 0$ and $\mathcal{W}^n$ be the space of $\mathbf{R}^n$-valued continuous functions on $[0, x]$ starting at the origin, and let P be the Wiener measure on $\mathcal{W}^n$. Following the idea of Cameron-Martin [3], we can show that

$$(1)\quad \begin{aligned} I(x,t) &:= \int_{\mathcal{W}^1} \exp\left[-\frac{a^2}{2}\int_0^x w(y)^2 dy - \frac{a}{2}\tanh(a^3 t) w(x)^2\right] P(dw) \\ &= \big(\cosh(a^3 t)\big)^{1/2}\big(\cosh(ax + a^3 t)\big)^{-1/2} \quad \text{for any } a > 0, \end{aligned}$$

where $w(y) \in \mathbf{R}$ denotes the position of $w \in \mathcal{W}^1$ at time y (see §4 and [8]). Then $u(x,t) = -4\partial_x^2 \log I(x,t)$, where $\partial_x = \partial/\partial x$, is a 1-soliton solution of the KdV equation

$$(2)\qquad \frac{\partial u}{\partial t} = \frac{3}{2} u \frac{\partial u}{\partial x} + \frac{1}{4}\frac{\partial^3 u}{\partial x^3}.$$

Received February 27, 2003.
Revised June 16, 2003.
The second author is partially supported by Grant-in-Aid for Scientific Research (A)(1) 14204010.

A reflectionless potential with scattering data $\eta_j, m_j > 0$, $1 \le j \le n$, is by definition a function

$$q(x) = -2\frac{d^2}{dx^2}\log\det(I + A(x)), \tag{3}$$

where

$$A(x) = \left(\frac{\sqrt{m_i m_j}}{\eta_i + \eta_j} e^{-(\eta_i+\eta_j)x}\right)_{1\le i,j\le n}. \tag{4}$$

We denote by $\mathcal{Q}_n$ the totality of all reflectionless potentials with scattering data consisting of $2n$ positive numbers. Let Σ be the set of all pairs $\sigma = (\sigma_+, \sigma_-)$ of non-negative measures $\sigma_\pm$ on $(-\infty, 0]$ such that $\int_{(-\infty,0]} e^{\lambda\sqrt{-z}}\sigma_\pm(dz) < \infty$ for any $\lambda > 0$. For $\sigma \in \Sigma$, set

$$\begin{aligned} G(u,v;\sigma) = {} & \frac{1}{4}\int_{-\infty}^{0} \frac{1}{\sqrt{-z}}\left(e^{\sqrt{-z}(u+v)} - e^{\sqrt{-z}|u-v|}\right)\sigma_+(dz) \\ & + \frac{1}{4}\int_{-\infty}^{0} \frac{1}{\sqrt{-z}}\left(e^{-\sqrt{-z}|u-v|} - e^{-\sqrt{-z}(u+v)}\right)\sigma_-(dz). \end{aligned}$$

We consider a family $\mathcal{G}$ of all Gaussian processes X^σ with mean 0 and covariance function $G(u,v;\sigma)$, $\sigma \in \Sigma$. We also consider the totality $\mathcal{Q}$ of all functions q^σ, $\sigma \in \Sigma$, defined by

$$q^\sigma(x) = -4\frac{d^2}{dx^2}\log E\left[\exp\left(-\frac{1}{2}\int_0^x |X^\sigma(y)|^2 dy\right)\right], \quad X^\sigma \in \mathcal{G},$$

where E stands for the expectation with respect to the underlying probability measure. In [12], Kotani showed that $\mathcal{Q}$ includes all $\mathcal{Q}_n$, $n = 1, 2, \dots$, and any element of $\mathcal{Q}$ is obtained as a limit of reflectionless potentials in the topology of uniform convergence on compacts.

Furthermore, it is well known that $q(x,t)$ defined by (3) and (4) with $m_j(t) = m_j \exp[-2\eta_j^3 t]$ instead of m_j, $1 \le j \le n$, gives a rise of an n-soliton solution $u(x,t) = -q(x,t)$ of the KdV equation (2).

The facts mentioned above indicate that soliton solutions of the KdV equation may be represented in terms of Gaussian processes. In this paper, we shall establish such an expression of n-soliton solutions and the tau function, which plays a fundamental role in the study of the KdV hierarchy (see [14, 16, 17]), in the Wiener space.

If both components $\sigma_\pm$ of $\sigma \in \Sigma$ are discrete measures, then the corresponding Gaussian process belongs to $\bigcup_{n\in\mathbf{N}} \mathcal{G}_n$, where $\mathcal{G}_n$ is a set of Gaussian processes obtained as superpositions of n independent 1-dimensional Ornstein-Uhlenbeck processes (for the definition of $\mathcal{G}_n$, see

§1.2). In this case, the correspondence between $\mathcal{G}$ and $\mathcal{Q}$ is given concretely; for every $n \in \mathbf{N}$, we shall give a mapping from $\mathcal{G}_n$ to $\mathcal{Q}_n$. See §2. Moreover, not only reflectionless potentials but also n-soliton solutions and the tau function of the KdV equation can be represented in terms of Gaussian processes in $\bigcup_{n\in\mathbf{N}} \mathcal{G}_n$. See §4. These expressions show that "a superposition" to make an n-soliton solution out of 1-soliton ones can be realized in the Wiener space. Further, we can explicitly see how speed parameters of 1-solitons reflect on those of the n-solitons obtained as superpositions. See §4.

An exact expression of Wiener integrals of Wiener functionals of the form $\exp[-(a^2/2)\int_0^x X(y)^2 dy + R(x)]$ of $X \in \bigcup_{n\in\mathbf{N}} \mathcal{G}_n$ plays a basic role in this paper, where $R(x)$ is a Wiener functional which varies according as we deal with reflectionless potentials, n-soliton solutions, and the tau function. Such exact expressions are achieved with the help of the Itô formula and the Cameron-Martin transformation. The Cameron-Martin transformation we deal with is determined by a second order ordinary differential equation. When a 1-dimensional Wiener process, which is in $\mathcal{G}_1$, is considered, the equation is the Sturm-Liouville one employed in [3]. The ordinary differential equations in this paper appear in different features from place to place, while they correspond to the same Wiener integral. Namely, we encounter several types of $n \times n$-matrix Riccati equations and second order $n \times n$-matrix and first order $2n \times n$-matrix linear ordinary differential equations. These different features are unified in terms of Grassmannians. See §1.1. It should be also mentioned that the above Riccati equations play an important role in the theory of the linear filtering problem by Kalman-Bucy. See §3.

Before closing this section, we note that the class $\mathcal{G}$ of Gaussian processes is closely related to the one studied by Hida-Streit [5] and Okabe [15].

The authors are grateful to Professor S. Kotani for his helpful comments and informing us of his recent achievement privately. They also thank Professor M. Yor for stimulating discussions with him. The subject of the paper grew out when the first author was visiting the Mathematics Research Centre, the University of Warwick in summer in 2001 at the invitation of Professor K.D. Elworthy, and discussed with Professor M. Kleptsyna. The second author had a chance to talk about the prototype of this paper with Professor Elworthy while he visited Japan in 2001. It is a great pleasure of the authors to thank both of them.

§1. Cameron-Martin transformation — Ornstein-Uhlenbeck process

1.1. Ordinary differential equations

We first recall several known facts about ordinary differential equations. For $n \in \mathbf{N}$, we set $\mathcal{A}_n = \mathcal{P}_n \times \mathcal{C}_n$, where

$$\mathcal{P}_n = \{\boldsymbol{p} = {}^t(p_1, \ldots, p_n) \in \mathbf{R}^n : p_i \neq p_j \text{ for } i \neq j\},$$
$$\mathcal{C}_n = \{\boldsymbol{c} = {}^t(c_1, \ldots, c_n) \in \mathbf{R}^n : c_i > 0 \text{ for } 1 \leq i \leq n\}.$$

For $(\boldsymbol{p}, \boldsymbol{c}) \in \mathcal{A}_n$ and $a > 0$, we define $n \times n$-matrices

$$D_{\boldsymbol{p}} = \operatorname{diag}[p_1, \ldots, p_n] \quad \text{and} \quad E_{\boldsymbol{p},\boldsymbol{c}}(a) = D_{\boldsymbol{p}}^2 + a^2 \boldsymbol{c} \otimes \boldsymbol{c}.$$

We shall often write simply D and $E(a)$ for $D_{\boldsymbol{p}}$ and $E_{\boldsymbol{p},\boldsymbol{c}}(a)$, respectively.

Let $\Phi_a(y)$ be a solution of a first order $2n \times n$-matrix ordinary differential equation

$$\Phi' + M_{\boldsymbol{p},\mathbf{c},a}\Phi = 0, \quad \text{where } M_{\boldsymbol{p},\mathbf{c},a} = \begin{pmatrix} D_{\boldsymbol{p}} & I \\ a^2 \boldsymbol{c} \otimes \boldsymbol{c} & -D_{\boldsymbol{p}} \end{pmatrix} \tag{5}$$

and f' stands for the derivative of f. For $n \times n$-matrices A and B, we often write $\Phi_a(y; A, B)$ to emphasize the initial condition $\Phi_a(0) = \begin{pmatrix} A \\ B \end{pmatrix}$. Denote by $\phi_a(y)$ and $\psi_a(y)$ the upper and the lower half $n \times n$-submatrices of $\Phi_a(y)$, respectively;

$$\Phi_a(y) = \begin{pmatrix} \phi_a(y) \\ \psi_a(y) \end{pmatrix}.$$

Then $\phi_a(y)$ obeys a second order ordinary differential equation

$$\phi'' - E(a)\phi = 0. \tag{6}$$

In the cases we deal with in this paper, $\phi_a(y)$ is always shown to be invertible for any $y \geq 0$. Moreover, if $\Phi_a(y) = \Phi_a(y; I, 0)$, which is the case investigated in §2 and §3, then $\psi_a(z)$ is also invertible for $z > 0$ (see a paragraph after Theorem 2.1). Hence, in the remainder of this subsection, we assume that $\phi_a(y)$ and $\psi_a(z)$ are both invertible for any $y \geq 0$ and $z > 0$. Then $\Phi_a(y)$ determines an n-frame in a $2n$-dimensional vector space $V(2n)$ over $\mathbf{R}$, and hence gives a rise of a dynamics on the Grassmannian $GM(n, V(2n))$ consisting of all n-dimensional vector subspaces of $V(2n)$. Moreover, $\Phi_a(y)$ is identified in $GM(n, V(2n))$ with

$$\begin{pmatrix} I \\ \psi_a(y)\phi_a^{-1}(y) \end{pmatrix} = \begin{pmatrix} I \\ \gamma_a(y) \end{pmatrix},$$

where $\phi_a^{-1}(y) = (\phi_a(y))^{-1}$ and $\gamma_a(y) = \psi_a(y)\phi_a^{-1}(y) = -\phi'(y)\phi_a^{-1}(y) - D$. Due to the Cole-Hopf transformation, γ_a obeys the $n \times n$-matrix Riccati equation

$$\gamma' - \gamma D - D\gamma - \gamma^2 + a^2 \boldsymbol{c} \otimes \boldsymbol{c} = 0$$

(see [18]). We next consider an n-frame obtained by reversing the time of $\Phi_a(\cdot)$; $\widetilde{\Phi}_a(y) = \Phi_a(x-y)$. Set $\mu(y) = \gamma_a(x-y)$ and $\nu(y) = \mu(y)^{-1}$, $y < x$. Then, for $y < x$, $\widetilde{\Phi}_a(y)$ determines a point in $GM(n, V(2n))$ identified with $\begin{pmatrix} I \\ \mu(y) \end{pmatrix}$ and $\begin{pmatrix} \nu(y) \\ I \end{pmatrix}$. Thus a dynamics of $\widetilde{\Phi}_a(y)$, $y < x$, in the Grassmannian is expressed in two different ways by Riccati equations

$$\begin{aligned} &\mu' + \mu D + D\mu + \mu^2 - a^2 \boldsymbol{c} \otimes \boldsymbol{c} = 0, \\ &\nu' - \nu D - D\nu + \nu(a^2 \boldsymbol{c} \otimes \boldsymbol{c})\nu - I = 0. \end{aligned}$$

The second equation is a Riccati equation which an error matrix appearing in the linear filtering theory obeys (see §3 and [1]).

Let $\alpha_{ij}(y)$ be the (i,j)-component of the n-frame $\Phi_a(y)$, $0 \le i \le 2n-1$, $0 \le j \le n-1$. The Plücker coordinate of $\Phi_a(y)$ is given by

$$\alpha_I(y) = \det\big[\big(\alpha_{i_k j}(y)\big)_{0 \le k,j \le n-1}\big], \qquad I = (i_0, \dots, i_{n-1}) \in \mathcal{I},$$

where $\mathcal{I}$ is the totality of $I = (i_0, \dots, i_{n-1}) \in \mathbf{Z}^n$ with $0 \le i_0 < \cdots < i_{n-1} \le 2n-1$. We set $F = \begin{pmatrix} -D & -I \\ -a^2 \boldsymbol{c} \otimes \boldsymbol{c} & D \end{pmatrix}$ and define a $\binom{2n}{n} \times \binom{2n}{n}$-matrix G by

$$G\alpha_I = \sum_{k=0}^{n-1} \sum_{j=0}^{2n-1} F_{i_k j} \alpha_{i_0, \dots, i_{k-1}, j, i_{k+1}, \dots, i_{n-1}}, \quad I = (i_0, \dots, i_{n-1}) \in \mathcal{I},$$

where, for $0 \le k_0, \dots, k_{n-1} \le 2n-1$, $\alpha_{k_0, \dots, k_{n-1}}$ is defined in the same manner as α_I. We then have a dynamics on the Grassmannian in terms of the Plücker coordinate;

$$\frac{d}{dy}\big(\alpha_I(y)\big)_{I \in \mathcal{I}} = G\big(\alpha_I(y)\big)_{I \in \mathcal{I}}.$$

It should be mentioned that

$$\alpha_I(y) = \det \phi_a(y) \qquad \text{for } I = (0, 1, \dots, n-1).$$

For related results, see also [4].

1.2. Cameron-Martin transformation

For $\boldsymbol{p} \in \mathcal{P}_n$, let $\xi_{\boldsymbol{p}}(y) = {}^t(\xi_{\boldsymbol{p}}^1(y), \dots, \xi_{\boldsymbol{p}}^n(y))$ be the unique solution of the $\mathbf{R}^n$-valued stochastic differential equation

$$d\xi(y) = dw(y) + D\xi(y)dy, \quad \xi(0) = 0. \tag{7}$$

We set $\mathcal{OU}_n = \{\xi_{\boldsymbol{p}} : \boldsymbol{p} \in \mathcal{P}_n\}$. For $(\boldsymbol{p}, \boldsymbol{c}) \in \mathcal{A}_n$, we define a superposition of 1-dimensional Ornstein-Uhlenbeck processes $\xi_{\boldsymbol{p}}^1(y), \dots, \xi_{\boldsymbol{p}}^n(y)$ by

$$X_{\boldsymbol{p},\boldsymbol{c}}(y) = \langle \boldsymbol{c}, \xi_{\boldsymbol{p}}(y) \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on $\mathbf{R}^n$. $X_{\boldsymbol{p},\boldsymbol{c}}(y)$, $0 \le y \le x$, is a continuous Gaussian process with mean 0 and covariance function

$$R(u,v) = \sum_{j=1}^{n} \frac{c_j^2}{2p_j} \left(e^{p_j(u+v)} - e^{p_j|u-v|} \right), \quad 0 \le u, v \le x. \tag{8}$$

We set

$$\mathcal{G}_n = \{X_{\boldsymbol{p},\boldsymbol{c}} : (\boldsymbol{p}, \boldsymbol{c}) \in \mathcal{A}_n\} = \{\langle \boldsymbol{c}, \xi_{\boldsymbol{p}} \rangle : (\boldsymbol{p}, \boldsymbol{c}) \in \mathcal{A}_n\}.$$

Obviously, $aX_{\boldsymbol{p},\boldsymbol{c}} = X_{\boldsymbol{p},a\boldsymbol{c}}$ for $a > 0$, and hence $\mathcal{G}_n$ is closed under the multiplication by positive numbers. Moreover, $X_{\boldsymbol{p},\boldsymbol{c}} \in \mathcal{G}_n$ is invariant under permutation in the sense that $X_{\boldsymbol{p}',\boldsymbol{c}'} = X_{\boldsymbol{p},\boldsymbol{c}}$ if $(\boldsymbol{p}, \boldsymbol{c}) \in \mathcal{A}_n$, σ is a permutation of $\{1, \dots, n\}$, and $\boldsymbol{p}' = {}^t(p_{\sigma(1)}, \dots, p_{\sigma(n)})$, $\boldsymbol{c}' = {}^t(c_{\sigma(1)}, \dots, c_{\sigma(n)})$.

Given an $n \times n$-matrix valued C^1 function κ on $[0, x]$ with $\det \kappa(y) \neq 0$ for each $y \in [0, x]$, we define two Cameron-Martin transformations K and L on $\mathcal{W}^n$ by

$$K[w](y) = w(y) - \int_0^y \kappa'(u)\kappa^{-1}(u)w(u)du, \tag{9}$$

$$L[w](y) = w(y) - \kappa(y) \int_0^y (\kappa^{-1})'(u)w(u)du, \quad w \in \mathcal{W}^n. \tag{10}$$

By a change of variables formula on $[0, x]$, we see that

$$K[L[w]] = L[K[w]] = w \quad \text{for any } w \in \mathcal{W}^n. \tag{11}$$

Set $\widetilde{\theta}(y) = \kappa'(y)\kappa^{-1}(y)$ and let $\widehat{\theta}$ be an $n \times n$-matrix valued continuous function on $[0, x]$. We then have

Lemma 1.1. *For any measurable $f : \mathcal{W}^n \to [0,\infty)$, it holds that*

$$\begin{aligned}(12)\quad &\int_{\mathcal{W}^n} f(\xi_{\boldsymbol{p}}) \exp\Big[-\frac{a^2}{2}\int_0^x X_{\boldsymbol{p},\boldsymbol{c}}(y)^2 dy \\ &\qquad + \frac{1}{2}\langle(\widetilde{\theta}(x) - D)\xi_{\boldsymbol{p}}(x), \xi_{\boldsymbol{p}}(x)\rangle - \frac{1}{2}\int_0^x |\widehat{\theta}(y)\xi_{\boldsymbol{p}}(y)|^2 dy\Big]\, dP \\ &= \int_{\mathcal{W}^n} f(w) \exp\Big[-\frac{1}{2}\int_0^x \langle E(a)w(y), w(y)\rangle dy \\ &\qquad + \frac{1}{2}\langle\widetilde{\theta}(x)w(x), w(x)\rangle - \frac{1}{2}\int_0^x |\widehat{\theta}(y)w(y)|^2 dy\Big]\, dP\, e^{-(x/2)\mathrm{tr}D}.\end{aligned}$$

Proof. By using the Maruyama-Girsanov transformation ([8, 13, 19]), we obtain that

the left hand side of (12)

$$\begin{aligned}&= \int_{\mathcal{W}^n} f(w) \exp\Big[-\frac{a^2}{2}\int_0^x \langle \boldsymbol{c}, w(y)\rangle^2 dy \\ &\qquad + \frac{1}{2}\langle(\widetilde{\theta}(x) - D)w(x), w(x)\rangle - \frac{1}{2}\int_0^x |\widehat{\theta}(y)w(y)|^2 dy \\ &\qquad + \int_0^x \langle Dw(y), dw(y)\rangle - \frac{1}{2}\int_0^x |Dw(y)|^2 dy\Big] P(dw),\end{aligned}$$

where the identity may hold as $\infty = \infty$. Applying the Itô formula ([8, 13]) to $\langle Dw(x), w(x)\rangle$, it is easily seen that this implies (12). Q.E.D.

Suppose that a solution $\phi_a(y)$, $y \in [0,x]$, of the ordinary differential equation (6) satisfies the condition that $\det \phi_a(y) \neq 0$, $0 \le y \le x$, where the initial condition is not specified. We set

$$(13)\qquad \beta_{a,x}(y) = -(\phi_a'\phi_a^{-1})(x - y),$$

and denote by $\widetilde{\beta}_{a,x}(y)$ and $\widehat{\beta}_{a,x}(y)$ its symmetric and skew-symmetric parts, respectively. Let $\kappa_{a,x}(y)$ be the solution of the differential equation

$$\kappa'(y) = \widetilde{\beta}_{a,x}(y)\kappa(y), \quad \kappa(x) = I,$$

and define linear transformations $K_{a,x}, L_{a,x} : \mathcal{W}^n \to \mathcal{W}^n$ by (9) and (10) with $\kappa = \kappa_{a,x}$.

Proposition 1.1. *For any bounded measurable $f : \mathcal{W}^n \to [0,\infty)$, it holds that*

$$(14)\quad \int_{\mathcal{W}^n} f(\xi_{\boldsymbol{p}}) \exp\Big[-\frac{a^2}{2}\int_0^x X_{\boldsymbol{p},\boldsymbol{c}}(y)^2 dy + \frac{1}{2}\langle(\widetilde{\beta}_{a,x}(x)-D)\xi_{\boldsymbol{p}}(x), \xi_{\boldsymbol{p}}(x)\rangle - \frac{1}{2}\int_0^x |\widehat{\beta}_{a,x}(y)\xi_{\boldsymbol{p}}(y)|^2 dy\Big] dP = \big(\det\phi_a(0)\big)^{1/2}\big(e^{x\,\mathrm{tr} D}\det\phi_a(x)\big)^{-1/2}\int_{\mathcal{W}^n} f\circ L_{a,x}\, dP.$$

Proof. For the sake of simplicity, we write β, $\widetilde{\beta}$, and $\widehat{\beta}$ for $\beta_{a,x}$, $\widetilde{\beta}_{a,x}$, and $\widehat{\beta}_{a,x}$, respectively. It follows from (6) that

$$(15)\qquad \beta' = E(a) - \beta^2 \quad\text{and}\quad \widetilde{\beta}' = E(a) - \widetilde{\beta}^2 - \widehat{\beta}^2.$$

Since $\mathrm{tr}\widetilde{\beta} = \mathrm{tr}\beta$, by virtue of the Itô formula, we then have

$$\begin{aligned}
&\frac{1}{2}\langle\widetilde{\beta}(x)w(x), w(x)\rangle \\
&= \frac{1}{2}\int_0^x \langle E(a)w(y), w(y)\rangle dy + \frac{1}{2}\int_0^x |\widehat{\beta}(y)w(y)|^2 dy \\
&\quad + \int_0^x \langle\widetilde{\beta}(y)w(y), dw(y)\rangle - \frac{1}{2}\int_0^x |\widetilde{\beta}(y)w(y)|^2 dy + \frac{1}{2}\int_0^x \mathrm{tr}\beta(y)dy.
\end{aligned}$$

Combining this with Lemma 1.1, we have

$$\begin{aligned}
(16)\quad &\int_{\mathcal{W}^n} f(\xi_{\boldsymbol{p}})\exp\Big[-\frac{a^2}{2}\int_0^x X_{\boldsymbol{p},\boldsymbol{c}}(y)^2 dy \\
&\qquad + \frac{1}{2}\langle(\widetilde{\beta}(x)-D)\xi_{\boldsymbol{p}}(x), \xi_{\boldsymbol{p}}(x)\rangle - \frac{1}{2}\int_0^x |\widehat{\beta}(y)\xi_{\boldsymbol{p}}(y)|^2 dy\Big] dP \\
&= \int_{\mathcal{W}^n} f(w)\exp\Big[\int_0^x \langle\widetilde{\beta}(y)w(y), dw(y)\rangle - \frac{1}{2}\int_0^x |\widetilde{\beta}(y)w(y)|^2 dy\Big] P(dw) \\
&\qquad\qquad \times \exp\Big[-\frac{x}{2}\mathrm{tr}D + \frac{1}{2}\int_0^x \mathrm{tr}\beta(y)dy\Big].
\end{aligned}$$

Applying the Maruyama-Girsanov transformation to the equation

$$dz(y) = dw(y) + \widetilde{\beta}(y)z(y)dy,$$

and noting that $\widetilde{\beta} = \kappa'_{a,x}\kappa_{a,x}^{-1}$, we obtain that

$$(17)\quad \int_{\mathcal{W}^n} (g \circ K_{a,x})(w) \exp\left[\int_0^x \langle \widetilde{\beta}(y)w(y), dw(y)\rangle - \frac{1}{2}\int_0^x |\widetilde{\beta}(y)w(y)|^2 dy\right] P(dw) = \int_{\mathcal{W}^n} g\, dP$$

for any bounded measurable $g : \mathcal{W}^n \to [0,\infty)$. By the definition of β, we have

$$(18)\qquad \exp\left[\int_0^x \mathrm{tr}\beta(y)dy\right] = \det\phi_a(0)\big(\det\phi_a(x)\big)^{-1}.$$

By (11), combining (17) and (18) with (16), we obtain (14). Q.E.D.

1.3. Eigenvalues of $E(a)$

Let $(\boldsymbol{p}, \boldsymbol{c}) \in \mathcal{A}_n$ and $a > 0$. We shall specify the eigenvalues of $E(a) = E_{\boldsymbol{p},\boldsymbol{c}}(a)$. As for $\boldsymbol{p} = {}^t(p_1, \ldots, p_n)$, rearranging if necessary, we may assume that there exist m and $1 \le j(1) < \cdots < j(m) \le n$ such that

$(\mathrm{H})_m$ $\quad |p_j| \le |p_{j+1}|$ for $j = 1, 2, \ldots, n-1$, $p_{j(\ell)} > 0$ and $p_{j(\ell)+1} = -p_{j(\ell)}$ for $\ell = 1, \ldots, m$, and $\#\{|p_1|, \ldots, |p_n|\} = n - m$.

When $m = 0$, this conditions means that $|p_1| < |p_2| < \cdots < |p_n|$. If $n = 1$, then $(\mathrm{H})_0$ holds. We define a Herglotz function $h_{\boldsymbol{p},\boldsymbol{c},a}$ on $\mathbf{C}^+ = \{z \in \mathbf{C} : \mathrm{Im} z > 0\}$ by

$$h_{\boldsymbol{p},\boldsymbol{c},a}(z) = \frac{1}{2}\int_0^\infty \frac{1}{u - z}\{\sigma_+ + \sigma_-\}(-du),$$

where

$$(19)\quad \sigma_+(du) = 2a^2 \sum_{i:p_i \ge 0} c_i^2 \delta_{-p_i^2}(du), \quad \sigma_-(du) = 2a^2 \sum_{i:p_i<0} c_i^2 \delta_{-p_i^2}(du)$$

(cf.[11]). Then $h_{\boldsymbol{p},\boldsymbol{c},a}(\lambda + t\sqrt{-1})$ converges to

$$h_{\boldsymbol{p},\boldsymbol{c},a}(\lambda + 0\sqrt{-1}) = a^2 \sum_{j=1}^n \frac{c_j^2}{p_j^2 - \lambda} \qquad \text{as } t \downarrow 0.$$

Under $(\mathrm{H})_m$, the equation $h_{\boldsymbol{p},\boldsymbol{c},a}(r + 0\sqrt{-1}) = -1$ possesses $n - m$ roots $0 < r_1 < \cdots < r_{n-m}$ such that $r_j^{1/2} \notin \{p_{j(1)}, \ldots, p_{j(m)}\}$, $j = 1, \ldots, n-m$. Take $\eta_1, \ldots, \eta_n \in \mathbf{R}$ so that

$$\{|\eta_1|, \ldots, |\eta_n|\} = \{p_{j(1)}, \ldots, p_{j(m)}, r_1^{1/2}, \ldots, r_{n-m}^{1/2}\}.$$

Define an $n \times n$ matrix $U = (U_{ij})_{1\le i,j\le n}$ by

$$(20) \qquad U_{ij} = \begin{cases} \dfrac{c_i}{|(D^2 - r_k I)^{-1}\boldsymbol{c}|(p_i^2 - r_k)}, & \text{if } \eta_j^2 = r_k, \\ \dfrac{\delta_{i,j(\ell)+1}c_{j(\ell)} - \delta_{i,j(\ell)}c_{j(\ell)+1}}{(c_{j(\ell)}^2 + c_{j(\ell)+1}^2)^{1/2}}, & \text{if } \eta_j^2 = p_{j(\ell)}^2. \end{cases}$$

Lemma 1.2. *$U \in O(n)$ and it holds that*

$$E(a) = UR^2U^{-1}, \qquad \textit{where } R = \mathrm{diag}[\eta_1, \dots, \eta_n].$$

Proof. It is easily checked that $p_{j(\ell)}^2$ is an eigenvalue of $E(a)$ with eigenvector

$$u = {}^t(\underbrace{0, \dots, 0}_{j(\ell)-1}, -c_{j(\ell)+1}, c_{j(\ell)}, 0, \dots, 0).$$

Noting that $D^2 - r_j I$ is invertible, we see that r_j is an eigenvalue with eigenvector $u = (D^2 - r_j I)^{-1}\boldsymbol{c}$. Thus we have obtained n distinct eigenvalues of $E(a)$ and the associated eigenvectors. In conjunction with the symmetry of $E(a)$, we obtain the desired assertion. Q.E.D.

§2. Reflectionless potential

For $(\boldsymbol{p}, \boldsymbol{c}) \in \mathcal{A}_n$ and $a > 0$, define $\sigma = (\sigma_+, \sigma_-) \in \Sigma$ by (19). Then $G(u, v; \sigma)$ coincides with $a^2R(u, v)$, the covariance function described in (8). Hence we can identify X^σ with $aX_{\boldsymbol{p},\boldsymbol{c}}$, and $\mathcal{G}_n \subset \mathcal{G}$. We shall spell out a correspondence between $\mathcal{G}_n$ and $\mathcal{Q}_n$.

Assuming $(\mathrm{H})_m$, we define $0 < r_1 < \cdots < r_{n-m}$ as described before Lemma 1.2. Define $0 < \eta_1 < \cdots < \eta_n$ and $m_1, \dots, m_n > 0$ by

$$\{\eta_1, \dots, \eta_n\} = \{p_{j(1)}, \dots, p_{j(m)}, r_1^{1/2}, \dots, r_{n-m}^{1/2}\}.$$

$$m_i = \begin{cases} 2\eta_i \dfrac{c_{j(\ell)+1}^2}{c_{j(\ell)}^2} \displaystyle\prod_{k\ne i} \frac{\eta_k + \eta_i}{\eta_k - \eta_i} \prod_{k\ne j(\ell), j(\ell)+1} \frac{p_k + \eta_i}{p_k - \eta_i}, & \text{if } i = j(\ell), \\ -2\eta_i \displaystyle\prod_{k\ne i} \frac{\eta_k + \eta_i}{\eta_k - \eta_i} \prod_{k=1}^{n} \frac{p_k + \eta_i}{p_k - \eta_i}, & \text{otherwise.} \end{cases}$$

Mention that $\eta_{j(\ell)} = p_{j(\ell)}$, $\ell = 1, \dots, m$. We set

$$(21) \qquad I_{\boldsymbol{p},\boldsymbol{c},a}(x) = \int_{\mathcal{W}^n} \exp\left[-\frac{a^2}{2}\int_0^x X_{\boldsymbol{p},\boldsymbol{c}}(y)^2 dy\right] dP.$$

Theorem 2.1. *Define $A(x)$ by (4) with the above scattering data $\eta_i, m_i > 0$, $i = 1, \ldots, n$. Then it holds that*

$$\log(I_{\boldsymbol{p},\mathbf{c},a}(x)) = -\frac{1}{2}\log\det(I + A(x)) + \frac{1}{2}\log\det(I + A(0)) - \frac{x}{2}\sum_{i=1}^{n}(p_i + \eta_i). \tag{22}$$

In particular,

$$4\frac{d^2}{dx^2}\log(I_{\boldsymbol{p},\mathbf{c},a}(x)) = -2\frac{d^2}{dx^2}\log\det(I + A(x)).$$

Note that $I_{\boldsymbol{p},\mathbf{c},a}(x)$ is invariant under the permutation of parameters $(\boldsymbol{p}, \boldsymbol{c})$ in the sense as stated after (8). Thus, so is the reflectionless potential associated with $I_{\boldsymbol{p},\mathbf{c},a}(x)$.

The proof of the theorem is divided into two steps, each step being a lemma. In the sequel, we write R for $\mathrm{diag}[\eta_1, \ldots, \eta_n]$ and let $\phi_a(y)$ be the upper half of $\Phi_a(y; I, 0)$. Then

$$\phi_a(y) = U\{\cosh(yR) - \sinh(yR)R^{-1}U^{-1}DU\}U^{-1}, \tag{23}$$

where $U \in O(n)$ is defined by (20), and, for $n \times n$-matrix A,

$$\cosh(A) = (e^A + e^{-A})/2 \quad \text{and} \quad \sinh(A) = (e^A - e^{-A})/2.$$

By computing the product matrix $\phi_a\psi_a$, we see that $\phi_a(y)$ and $\psi_a(z)$ are both invertible for $y \geq 0$, $z > 0$. Then we can define $\beta_{a,x}$ by (13) with this ϕ_a. Obviously $\beta_{a,x}(x) = D$. Since $\beta_{a,x}$ obeys the Riccati equation (15), it turns out to be symmetric. Applying Proposition 1.1, we obtain

$$I_{\boldsymbol{p},\mathbf{c},a}(x) = \big(\det\phi_a(0)\big)^{1/2}\big(\det\phi_a(x)\big)^{-1/2}e^{-(x/2)\mathrm{tr}D}. \tag{24}$$

In this expression, for the latter use, we left $\det\phi_a(0)$ while it is equal to one. If $(\mathrm{H})_0$ holds, then we set

$$X_{ij} = \big(p_j + r_i^{1/2}\big)^{-1}, \quad Y_{ij} = \big(p_j - r_i^{1/2}\big)^{-1}, \quad 1 \leq i, j \leq n,$$
$$X = \big(X_{ij}\big)_{1\leq i,j\leq n}, \quad Y = \big(Y_{ij}\big)_{1\leq i,j\leq n},$$
$$V(\boldsymbol{c}) = \mathrm{diag}\big[|(D^2 - r_1 I)^{-1}\mathbf{c}|^{-1}, \ldots, |(D^2 - r_n I)^{-1}\mathbf{c}|^{-1}\big],$$
$$\sigma(i) = \mathrm{sgn}\left[\prod_{\beta=1}^{n}(p_\beta - \eta_i)\right], \quad b(i) = \sigma(i)\left\{-2\eta_i\frac{\prod_{\alpha\neq i}(\eta_\alpha^2 - \eta_i^2)}{\prod_{\beta=1}^{n}(p_\beta^2 - \eta_i^2)}\right\}^{1/2},$$

and $B = \mathrm{diag}[b(1), \ldots, b(n)]$.

Lemma 2.1. *Suppose that* $(\mathrm{H})_0$ *holds. Then it holds that*

$$\phi_a(y) = -\frac{1}{2}UV(\boldsymbol{c})R^{-1}B\bigl(I+A(y)\bigr)e^{yR}B^{-1}XC(\boldsymbol{c}), \tag{25}$$

where $C(\boldsymbol{c}) = \mathrm{diag}[c_1,\dots,c_n]$. *Moreover, the identity* (22) *holds.*

Proof. Due to (23), we have

$$\phi_a(y) = \frac{1}{2}UR^{-1}\bigl\{e^{yR}(RU^{-1}-U^{-1}D) + e^{-yR}(RU^{-1}+U^{-1}D)\bigr\}.$$

Set $Z = (Z_{ij})_{1\le i,j\le n}$, where

$$Z_{ij} = (p_i^2 - r_j)^{-1}, \quad 1 \le i,j \le n.$$

Then it holds that

$$U = C(\boldsymbol{c})ZV(\boldsymbol{c}), \quad U^{-1} = V(\boldsymbol{c})\,{}^tZC(\boldsymbol{c}).$$

Since $R, D, C(\boldsymbol{c})$, and $V(\boldsymbol{c})$ are all diagonal matrices, we have that

$$RU^{-1} - U^{-1}D = -V(\boldsymbol{c})XC(\boldsymbol{c}), \quad RU^{-1} + U^{-1}D = V(\boldsymbol{c})YC(\boldsymbol{c}).$$

Hence we obtain

$$\phi_a(y) = -\frac{1}{2}UR^{-1}V(\boldsymbol{c})\bigl\{I - e^{-yR}YX^{-1}e^{-yR}\bigr\}e^{yR}XC(\boldsymbol{c}). \tag{26}$$

We now compute YX^{-1}. Applying Cauchy's identity (cf. [14])

$$\det\biggl(\Bigl(\frac{1}{\alpha_i+\beta_j}\Bigr)_{1\le i,j\le n}\biggr) = \frac{\prod_{1\le i<j\le n}(\alpha_i-\alpha_j)(\beta_i-\beta_j)}{\prod_{i,j=1}^n(\alpha_i+\beta_j)} \tag{27}$$
$$\text{for } \alpha_1,\dots,\alpha_n,\beta_1,\dots,\beta_n \in \mathbf{C},$$

to the cofactor matrix of X, we obtain

$$\bigl(X^{-1}\bigr)_{k\ell} = \frac{\prod_{\alpha\ne\ell}(p_k+\eta_\alpha)\prod_{\beta=1}^n(p_\beta+\eta_\ell)}{\prod_{\beta\ne k}(p_\beta-p_k)\prod_{\alpha\ne\ell}(\eta_\alpha-\eta_\ell)}, \quad 1\le k,\ell\le n.$$

Using Lagrange's interpolation formula

$$\sum_{k=1}^n \frac{\prod_{j=1}^{n-1}(a_k+b_j)\prod_{\beta\ne k}(a_\alpha - z)}{\prod_{\beta\ne k}(a_\alpha-a_k)} = \prod_{j=1}^{n-1}(z+b_j), \quad z\in\mathbf{C},$$

we have

$$(YX^{-1})_{ij} = \frac{\prod_{\alpha\ne i}(\eta_\alpha+\eta_i)}{\prod_{\beta=1}^n(p_\beta-\eta_i)}\,\frac{2\eta_i}{\eta_i+\eta_j}\,\frac{\prod_{\beta=1}^n(p_\beta+\eta_j)}{\prod_{\alpha\ne j}(\eta_\alpha-\eta_j)}, \quad 1\le i,j\le n.$$

Since $\operatorname{sgn}\big[\big(\prod_{\beta=1}^n(p_\beta+\eta_i)\big)/\big(\prod_{\alpha\neq i}(\eta_\alpha-\eta_i)\big)\big]=-\sigma(i)$, it holds that

$$b(i)\sqrt{m_i}=2\eta_i\frac{\prod_{\alpha\neq i}(\eta_\alpha+\eta_i)}{\prod_{\beta=1}^n(p_\beta-\eta_i)},\quad \frac{\sqrt{m_i}}{b(i)}=-\frac{\prod_{\beta=1}^n(p_\beta+\eta_i)}{\prod_{\alpha\neq i}(\eta_\alpha-\eta_i)},\quad 1\le i\le n.$$

Hence

$$(YX^{-1})_{ij}=-\frac{b(i)\sqrt{m_i}\sqrt{m_j}(b(j))^{-1}}{\eta_i+\eta_j},\quad 1\le i,j\le n.$$

Combining this with (26), we obtain (25).

The identity (25) implies

$$\det\phi_a(0)\big(\det\phi_a(x)\big)^{-1}=\det(I+A(0))\big(e^{x\operatorname{tr}R}\det(I+A(x))\big)^{-1}.$$

Thus the second assertion follows from this and (24). Q.E.D.

Lemma 2.2. *Let $m\ge 1$ and suppose that* $(\mathrm{H})_m$ *is satisfied. Then* (22) *holds.*

Proof. For $\varepsilon>0$, set

$$p_i^\varepsilon=\begin{cases}p_i-\varepsilon, & \text{if } i=j(\ell)+1 \text{ for some } \ell,\\ p_i, & \text{otherwise,}\end{cases}\qquad 1\le i\le n.$$

Choosing a sufficiently small $\varepsilon>0$, we may assume that

$$|p_1^\varepsilon|<|p_2^\varepsilon|<\cdots<|p_n^\varepsilon|.$$

Let $0<r_1^\varepsilon<\cdots<r_n^\varepsilon$ be roots of $a^2\sum_{i=1}^n c_i^2/\{(p_i^\varepsilon)^2-r\}=-1$. Then it holds that

$$(p_i^\varepsilon)^2<r_i^\varepsilon<(p_{i+1}^\varepsilon)^2<r_{i+1}^\varepsilon,\quad i=1,2,\dots,n-1. \tag{28}$$

Define scattering data $\eta_i^\varepsilon, m_i^\varepsilon>0$, $i=1,\dots,n$, with these p_i^ε's, r_i^ε's and c_i's as described before Theorem 2.1. By Lemma 2.1, we have

$$\begin{aligned}\log\big(I_{\boldsymbol{p}^\varepsilon,\mathbf{c},a}(x)\big)=&-\frac12\log\det(I+A^\varepsilon(x))\\&+\frac12\log\det(I+A^\varepsilon(0))-\frac{x}{2}\sum_{i=1}^n(p_i^\varepsilon+\eta_i^\varepsilon)\end{aligned}$$

where $\boldsymbol{p}^\varepsilon={}^t(p_1^\varepsilon,\dots,p_n^\varepsilon)$ and $A^\varepsilon(x)$ is defined by (4) with $\eta_i^\varepsilon, m_i^\varepsilon$, $i=1,\dots,n$. Since $p_i^\varepsilon\to p_i$ as $\varepsilon\downarrow 0$, $i=1,\dots,n$, we have that

$$\log\big(I_{\boldsymbol{p}^\varepsilon,\mathbf{c},a}(x)\big)\to\log\big(I_{\boldsymbol{p},\mathbf{c},a}(x)\big)\quad\text{as }\varepsilon\downarrow 0.$$

Moreover, recalling that $\eta_i^\varepsilon = (r_i^\varepsilon)^{1/2}$, $i = 1, \ldots, n$, it follows from (28) that $\eta_i^\varepsilon \to \eta_i$ as $\varepsilon \downarrow 0$, $i = 1, \ldots, n$. Hence

$$\sum_{i=1}^{n}(p_i^\varepsilon + \eta^\varepsilon) \to \sum_{i=1}^{n}(p_i + \eta_i), \quad \text{as } \varepsilon \downarrow 0.$$

Thus the proof completes once we have shown the convergence of $A^\varepsilon(y)$ to $A(y)$ as $\varepsilon \downarrow 0$.

To see the convergence of $A^\varepsilon(y)$, it suffices to show that m_i^ε tends to m_i as $\varepsilon \downarrow 0$ for every $i = 1, \ldots, n$. If $i \neq j(\ell)$ for any ℓ, then it is easily seen that $m_i^\varepsilon \to m_i$ as $\varepsilon \downarrow 0$. We now consider the case that $i = j(\ell)$. Since $\eta_{j(\ell)}^\varepsilon \to \eta_{j(\ell)} = p_{j(\ell)}$ as $\varepsilon \downarrow 0$,

$$\text{(29)} \qquad \frac{p_{j(\ell)}^\varepsilon + \eta_{j(\ell)}^\varepsilon}{p_{j(\ell)+1}^\varepsilon - \eta_{j(\ell)}^\varepsilon} \longrightarrow -1, \quad \text{as } \varepsilon \downarrow 0.$$

It follows from (28) and the identity $a^2 \sum_{j=1}^n c_j^2/\{(p_j^\varepsilon)^2 - r_{j(\ell)}^\varepsilon\} = -1$ that

$$c_{j(\ell)}^2\{(p_{j(\ell)} + \varepsilon)^2 - r_{j(\ell)}^\varepsilon\} + c_{j(\ell)+1}^2\{p_{j(\ell)}^2 - r_{j(\ell)}^\varepsilon\} = O(\varepsilon^2), \quad \text{as } \varepsilon \downarrow 0.$$

This yields that

$$\text{(30)} \qquad (p_{j(\ell)}^\varepsilon)^2 - (\eta_{j(\ell)}^\varepsilon)^2 = \frac{-2p_{j(\ell)}c_{j(\ell)}^2}{c_{j(\ell)}^2 + c_{j(\ell)+1}^2}\,\varepsilon + O(\varepsilon^2),$$

$$\text{(31)} \qquad (p_{j(\ell)+1}^\varepsilon)^2 - (\eta_{j(\ell)}^\varepsilon)^2 = \frac{2p_{j(\ell)}c_{j(\ell)+1}^2}{c_{j(\ell)}^2 + c_{j(\ell)+1}^2}\,\varepsilon + O(\varepsilon^2), \qquad \text{as } \varepsilon \downarrow 0.$$

Hence

$$\frac{p_{j(\ell)+1}^\varepsilon + \eta_{j(\ell)}^\varepsilon}{p_{j(\ell)}^\varepsilon - \eta_{j(\ell)}^\varepsilon} \longrightarrow c_{j(\ell)+1}^2 c_{j(\ell)}^{-2}, \quad \text{as } \varepsilon \downarrow 0.$$

Combining this with (29) and the definition of $m_{j(\ell)}^\varepsilon$, we see that $m_{j(\ell)}^\varepsilon \to m_{j(\ell)}$ as $\varepsilon \downarrow 0$. Q.E.D.

Remark 2.1. (i) The exponent of the integrand in the right hand side of (21) is a sum of a quadratic Wiener functional and a constant. Hence the right hand side of (24) can be expressed in terms of the Carleman-Fredholm determinant of the symmetric Hilbert-Schmidt operator determining the quadratic Wiener functional. Moreover, ϕ_a is a solution of the Jacobi equation associated with the Lagrangian related to the Wiener functional (cf. [6, 7]).
(ii) In Theorem 2.1, for each $n \in \mathbf{N}$, a mapping from $\mathcal{A}_n$ to the space

$\{(\eta_1,\ldots,\eta_n,m_1,\ldots,m_n) : 0<\eta_1<\cdots<\eta_n, m_1,\ldots,m_n>0\}$ of scattering data was established. If $n=2$, the mapping is invertible.

§3. Filtering theory

In this section, we shall see that the change of variables formula (24) relates to the filtering theory. On the $(n+1)$-dimensional Wiener space $\mathcal{W}^{n+1}$, consider the following filtering problem.

$$\begin{aligned} d\xi_{\boldsymbol{p}}(y) &= dw(y) + D\xi_{\boldsymbol{p}}(y)dy, & \xi(0)&=0, & &\text{(system)},\\ dY(y) &= db(y) + a\langle \boldsymbol{c}, \xi_{\boldsymbol{p}}(y)\rangle dy, & Y(0)&=0, & &\text{(observation)}, \end{aligned}$$

where $(\boldsymbol{p},\boldsymbol{c})\in\mathcal{A}_n$, $a>0$ and $(w,b)\in\mathcal{W}^n\times\mathcal{W}^1=\mathcal{W}^{n+1}$. Let $\mathcal{F}_y^Y$ be the σ-field generated by $Y(u)$, $u\le y$. A solution $\widehat{\xi}_{\boldsymbol{p}}(y)$ to the filtering problem with respect to $Y(u)$, $u\le y$, is realized as a function whose error matrix is minimal in the space of error matrices of $\mathbf{R}^n$-valued $\mathcal{F}_y^Y$-measurable functions, where the order is the one inherited from the non-negative definiteness ([1, Theorem 4.1]). In our case, $\widehat{\xi}_{\boldsymbol{p}}(y)$ coincides with the conditional expectation $E[\xi_{\boldsymbol{p}}(y)|\mathcal{F}_y^Y]$ of $\xi_{\boldsymbol{p}}(y)$ given $\mathcal{F}_y^Y$, which is called the Kalman-Bucy filter. The corresponding error matrix

$$P_a(y) = \int_{\mathcal{W}^{n+1}} (\xi_{\boldsymbol{p}}(y)-\widehat{\xi}_{\boldsymbol{p}}(y))^t(\xi_{\boldsymbol{p}}(y)-\widehat{\xi}_{\boldsymbol{p}}(y))\,dP$$

obeys the $n\times n$-matrix Riccati equation

$$P' = DP + PD - P(a^2\boldsymbol{c}\otimes\boldsymbol{c})P + I, \quad P(0)=0,$$

which we have already seen as an equation for $\nu(y)$ in §1.1. Let $\rho_{\boldsymbol{p},\mathbf{c},a}(y)$ be the error variance of $X_{\boldsymbol{p},\boldsymbol{c}}(y)$;

$$\rho_{\boldsymbol{p},\mathbf{c},a}(y) = \int_{\mathcal{W}^{n+1}} \left|X_{\boldsymbol{p},\boldsymbol{c}}(y) - E[X_{\boldsymbol{p},\boldsymbol{c}}(y)|\mathcal{F}_y^Y]\right|^2 dP.$$

It then holds that

$$\rho_{\boldsymbol{p},\mathbf{c},a}(y) = \mathrm{tr}\big[(\boldsymbol{c}\otimes\boldsymbol{c})P_a(y)\big].$$

Let $\Phi_a(y;I,0) = \begin{pmatrix}\phi_a(y)\\ \psi_a(y)\end{pmatrix}$. As was seen before (24), $\det\phi_a(y)\neq 0$, $y\ge 0$. We set $\gamma_a(y)=\psi_a(y)\phi_a^{-1}(y)$ and $\gamma_{a,x}(y)=\gamma_a(x-y)$. Note that

$$\gamma'_{a,x} = -D\gamma_{a,x} - \gamma_{a,x}D - \gamma_{a,x}^2 + a^2(\boldsymbol{c}\otimes\boldsymbol{c}) \quad \text{on } [0,x], \quad \gamma_{a,x}(x)=0.$$

Then it holds that

$$\{\det(I-\gamma_{a,x}P_a)\}' = \left(-\mathrm{tr}\left[a^2(\boldsymbol{c}\otimes\mathbf{c})P_a+\gamma_{a,x}\right]\right)\det(I-\gamma_{a,x}P_a).$$

Since $P(0)=\gamma_{a,x}(x)=0$,

$$1=\det(I-\gamma_{a,x}(x)P_a(x))=\exp\left[-\int_0^x \mathrm{tr}\left[a^2(\boldsymbol{c}\otimes\mathbf{c})P_a(y)+\gamma_{a,t}(y)\right]dy\right].$$

Combined with (18), this implies

$$\exp\left[-\int_0^x \mathrm{tr}\left[a^2(\boldsymbol{c}\otimes\mathbf{c})P_a(y)\right]dy\right]=\frac{e^{-x\mathrm{tr}D}}{\det\phi_a(x)}.$$

By (24), we obtain

$$\int_{\mathcal{W}^n}\exp\left[-\frac{a^2}{2}\int_0^x X_{\boldsymbol{p},\mathbf{c}}(y)^2dy\right]dP=\exp\left[-\frac{a^2}{2}\int_0^x \rho_{\boldsymbol{p},\mathbf{c},a}(y)dy\right],$$

which can be also shown by applying the result due to M.L. Kleptsyna and A. Le Breton [10].

Let $\delta(y,u)$ be the unique solution of the integral equation

$$\delta(y,u)=a^2R(y,u)-\int_0^y \delta(y,v)\delta(u,v)dv,\quad 0\le y,u\le x. \tag{32}$$

In [10], they also have shown that $a^2\rho_{\boldsymbol{p},\mathbf{c},a}(y)$ coincides with $\delta(y,y)$.

§4. KdV equation

Throughout this section, we assume $(\mathrm{H})_m$. Let scattering data $0<\eta_1<\cdots<\eta_n$, $m_1,\ldots,m_n>0$ and $n\times n$-matrices U and R be the ones stated before and after Theorem 2.1. Set

$$\mathbf{T}=\{\boldsymbol{t}=(t_1,t_2,\ldots,):t_j\in\mathbf{R},\#\{j:t_j\neq 0\}<\infty\}.$$

For $x\in\mathbf{R},\boldsymbol{t}\in\mathbf{T}$, define

$$\zeta_j(x,\boldsymbol{t})=x\eta_j+\sum_{\alpha=1}^{\infty}t_\alpha\eta_j^{2\alpha+1},\quad \zeta(x,\boldsymbol{t})=\mathrm{diag}[\zeta_1(x,\boldsymbol{t}),\ldots,\zeta_n(x,\boldsymbol{t})].$$

The tau function $\tau(x,\boldsymbol{t})$ of the KdV equation is of the form

$$\tau(x,\boldsymbol{t})=1+\sum_{p=1}^{n}\sum_{1\le i_1<\cdots<i_p\le n}\prod_{j=1}^{p}\frac{m_{i_j}}{2\eta_{i_j}}\prod_{1\le j<k\le p}\left(\frac{\eta_{i_j}-\eta_{i_k}}{\eta_{i_j}+\eta_{i_k}}\right)^2 \times\exp\left[-2\sum_{j=1}^{p}\zeta_{i_j}(x,\boldsymbol{t})\right]. \tag{33}$$

For details, see [14, 16]. If we set

$$A(x,\boldsymbol{t})=\left(\frac{\sqrt{m_im_j}}{\eta_i+\eta_j}e^{-\{\zeta_i(x,\boldsymbol{t})+\zeta_j(x,\boldsymbol{t})\}}\right)_{1\le i,j\le n},$$

then, with the help of Cauchy's identity (27), we can show that

$$\tau(x,\boldsymbol{t})=\det(I+A(x,\boldsymbol{t})),\quad x\in\mathbf{R},\ \boldsymbol{t}\in\mathbf{T}. \tag{34}$$

We shall show that $\tau(x,\boldsymbol{t})$ can be expressed in terms of Wiener integral. To do this, let

$$\phi_a(x,\boldsymbol{t})=U\big\{\cosh(\zeta(x,\boldsymbol{t}))-\sinh(\zeta(x,\boldsymbol{t}))R^{-1}U^{-1}DU\big\}U^{-1}. \tag{35}$$

Then, for each $\boldsymbol{t}\in\mathbf{T}$, $\phi_a(\cdot,\boldsymbol{t})$ obeys the differential equation (6) with initial condition

$$\begin{aligned}\phi_a(0)&=U\big\{\cosh(\zeta(0,\boldsymbol{t}))-\sinh(\zeta(0,\boldsymbol{t}))R^{-1}U^{-1}DU\big\}U^{-1},\\ \phi_a'(0)&=U\big\{R\sinh(\zeta(0,\boldsymbol{t}))-\cosh(\zeta(0,\boldsymbol{t}))U^{-1}DU\big\}U^{-1}.\end{aligned}$$

As we shall see in Lemma 4.1 below, $\det\phi_a(x,\boldsymbol{t})\neq 0$, and then we can define

$$\beta_{a,x,\boldsymbol{t}}(y)=-((\partial_x\phi_a)\phi_a^{-1})(x-y,\boldsymbol{t}).$$

We shall show that $\beta_{a,x,\boldsymbol{t}}(y)$ is symmetric. To do this, write $\beta(y,\boldsymbol{t})$ for $\beta_{a,x,\boldsymbol{t}}(y)$. For each $k=1,2,\dots$, the partial derivative $\partial_{t_k}\phi_a$ of ϕ_a with respect to t_k satisfies that

$$\partial_{t_k}\phi_a(x,\boldsymbol{t})=U\big\{R^{2k+1}\sinh(\zeta(x,\boldsymbol{t}))-R^{2k}\cosh(\zeta(x,\boldsymbol{t}))U^{-1}DU\big\}U^{-1}.$$

This implies that

$$\partial_{t_k}\phi_a(x,\boldsymbol{t})=E(a)^k\partial_x\phi_a(x,\boldsymbol{t})\ \text{ and }\ \partial_{t_k}^2\phi_a(x,\boldsymbol{t})=E(a)^{2k+1}\phi_a(x,\boldsymbol{t})$$

for any $k=1,2,\dots$ We then obtain that

$$\partial_{t_k}\beta(y,\boldsymbol{t})=-E(a)^{k+1}+\beta(y,\boldsymbol{t})E(a)^k\beta(y,\boldsymbol{t}),\quad k=1,2,\dots \tag{36}$$

Since $E(a)$ is symmetric, the transpose ${}^t\beta(y,\boldsymbol{t})$ of $\beta(y,\boldsymbol{t})$ satisfies the same identities in (36). Hence $\beta(y,\boldsymbol{t})$ is symmetric if and only if so is $\beta(y,\boldsymbol{t}[k])$ for some $k=1,2,\dots$, where $\boldsymbol{t}[k]$ is obtained from $\boldsymbol{t}$ by replacing t_k by 0. As was seen in the paragraph before (24), $\beta(y,\mathbf{0})$ is symmetric, where $\mathbf{0}=(0,0,\dots)\in\mathbf{T}$. In conjunction with the above observation, this implies that $\beta(y,\boldsymbol{t})$ is symmetric for $\boldsymbol{t}=(t_1,0,0,\dots)$, $t_1\in\mathbf{R}$. Apply the above observation again, it follows that $\beta(y,\boldsymbol{t})$ is symmetric for $\boldsymbol{t} = (t_1, t_2, 0, 0, \dots)$, $t_1, t_2 \in \mathbf{R}$. Repeating this argument successively, we can conclude that $\beta(y, \boldsymbol{t})$ is symmetric for any $(y, \boldsymbol{t}) \in \mathbf{R} \times \mathbf{T}$.

We set

$$I_{\boldsymbol{p},\mathbf{c},a}(x, \boldsymbol{t}) = \int_{\mathcal{W}^n} \exp\Big[-\frac{a^2}{2} \int_0^x X_{\boldsymbol{p},\mathbf{c}}(y)^2 dy + \frac{1}{2}\langle(\beta_{a,x,\boldsymbol{t}}(x) - D)\xi_{\boldsymbol{p}}(x), \xi_{\boldsymbol{p}}(x)\rangle\Big] dP.$$

To state our result, we introduce a set $J = \{(j(\ell) + 1, j(\ell)) : \ell = 1, \dots, m\}$ and a quantity

$$Z_m(\boldsymbol{p}, \mathbf{c}) = (-1)^m \prod_{i=1}^n \frac{c_i}{2\eta_i} \prod_{1 \le i < j \le n} (p_i - p_j)(\eta_i - \eta_j) \Big\{ \prod_{(i,j) \notin J} (p_i + \eta_j) \times \prod_{k=1}^{n-m} |(D_{\boldsymbol{p}}^2 - r_k I)^{-1} \mathbf{c}| \prod_{\ell=1}^m \frac{c_{j(\ell)+1}\big(c_{j(\ell)}^2 + c_{j(\ell)+1}^2\big)^{1/2}}{2p_{j(\ell)} c_{j(\ell)}} \Big\}^{-1},$$

where, if $m = 0$, then $J = \emptyset$, "$(i,j) \notin J$" means "$1 \le i, j \le n$", and $\prod_{\ell=1}^0(\cdots) = 1$.

Theorem 4.1. (i) *It holds that*

$$\text{(37)} \quad \det \phi_a(x, \boldsymbol{t}) = \tau(x, \boldsymbol{t}) e^{\operatorname{tr}\zeta(x,\boldsymbol{t})} Z_m(\boldsymbol{p}, \mathbf{c}),$$

$$\text{(38)} \quad \log\big(I_{\boldsymbol{p},\mathbf{c},a}(x, \boldsymbol{t})\big) = -\frac{1}{2}\log \tau(x, \boldsymbol{t}) + \frac{1}{2}\log \tau(0, \boldsymbol{t}) - \frac{x}{2}\sum_{i=1}^n (p_i + \eta_i).$$

(ii) *Let* $\boldsymbol{t} = (t, 0, \dots)$. *We write* t *for* $\boldsymbol{t}$.
(a) *Set*

$$q_{\boldsymbol{p},\mathbf{c},a}(x, t) = -4\partial_x^2 \log\big(I_{\boldsymbol{p},\mathbf{c},a}(x, t)\big).$$

Then $q_{\boldsymbol{p},\mathbf{c},a}(x, t)$ *solves the KdV equation* (2).
(b) *Both* $\det \phi_a(x, t)$ *and* $\big(I_{\boldsymbol{p},\mathbf{c},a}(x, t)\big)^{-2}$ *solve the Hirota equation:*

$$\text{(39)} \qquad (4D_t D_x - D_x^4) u \cdot u = 0,$$

where (D_x, D_t) *denotes the Hirota derivatives with respect to the variables* (x, t) ([14]).

While $q_{\boldsymbol{p},\mathbf{c},a}(x, t)$ is defined only on $[0, \infty) \times \mathbf{R}$ in our framework, by virtue of (37), it extends to $\mathbf{R} \times \mathbf{R}$ so that the extension also solves the KdV equation (2).

To prove Theorem 4.1, we first show a relation between $\phi_a(x, \boldsymbol{t})$ and $A(x, \boldsymbol{t})$.

Lemma 4.1. *It holds that*

$$\det \phi_a(y,\boldsymbol{t}) = \det(I + A(y,\boldsymbol{t}))e^{\mathrm{tr}\zeta(y,\boldsymbol{t})} Z_m(\boldsymbol{p},\boldsymbol{c}). \tag{40}$$

In particular, (37) *holds and* $\det \phi_a(y,t) \neq 0$, $y \geq 0$.

Proof. For $\varepsilon > 0$, define $\boldsymbol{p}^\varepsilon = {}^t(p_1^\varepsilon, \ldots, p_n^\varepsilon) \in \mathcal{P}_n$ as in the proof of Lemma 2.2. For sufficiently small ε, we see that $\boldsymbol{p}^\varepsilon$ satisfies the condition $(\mathrm{H})_0$. In the sequel, as in the proof of Lemma 2.2, we use the superscript ε to indicate the dependence on $\boldsymbol{p}^\varepsilon$; given a quantity f defined with $\boldsymbol{p}$, we write f^ε for the same quantity defined with $\boldsymbol{p}^\varepsilon$ instead of $\boldsymbol{p}$. Then, in repetition of the argument employed to prove Lemma 2.1, we can show that

$$\phi_a^\varepsilon(y,t) = -\frac{1}{2}U^\varepsilon(R^\varepsilon)^{-1}V^\varepsilon(\boldsymbol{c})B^\varepsilon\{I + A^\varepsilon(y,t)\}(B^\varepsilon)^{-1}e^{\zeta^\varepsilon(y,\boldsymbol{t})}X^\varepsilon C(\boldsymbol{c}).$$

Applying Cauchy's identity (27) to computing $\det X^\varepsilon$ and $\det U^\varepsilon$, we have

$$\det \phi_a^\varepsilon(y,t) = \det\bigl(I + A^\varepsilon(y,\boldsymbol{t})\bigr)e^{\mathrm{tr}\zeta^\varepsilon(y,\boldsymbol{t})} Z_0(\boldsymbol{p}^\varepsilon,\boldsymbol{c}). \tag{41}$$

As we have already seen in the proof of Lemma 2.2, as $\varepsilon \downarrow 0$, $\eta_i^\varepsilon \to \eta_i$ and $m_i^\varepsilon \to m_i$ for $i = 1, \ldots, n$. Moreover, taking the advantage of (30) and (31), we can show that, as $\varepsilon \downarrow 0$, $U^\varepsilon \to U$ and $Z_0(\boldsymbol{p}^\varepsilon,\boldsymbol{c}) \to Z_m(\boldsymbol{p},\boldsymbol{c})$. Hence, letting $\varepsilon \downarrow 0$ in (41), we obtain (40).

(37) is an immediate consequence of (34) and (40). The non-singularity of $\phi_a(y,\boldsymbol{t})$ follows from that of $I + A(y,\boldsymbol{t})$ and (40). Q.E.D.

Proof of Theorem 4.1. (i) We have already seen (37) in Lemma 4.1. By the same lemma, $\phi_a(y,\boldsymbol{t}) \neq 0$, $y \geq 0$. Applying Proposition 1.1 with $\phi_a(y) = \phi_a(y,\boldsymbol{t})$, we have

$$I_{\boldsymbol{p},\boldsymbol{c},a}(x,\boldsymbol{t}) = \bigl(\det \phi_a(0,\boldsymbol{t})\bigr)^{1/2}\bigl(e^{x\mathrm{tr}D} \det \phi_a(x,\boldsymbol{t})\bigr)^{-1/2}.$$

By (37), it holds that

$$\det \phi_a(0,\boldsymbol{t})\bigl(\det \phi_a(x,\boldsymbol{t})\bigr)^{-1} = \tau(0,\boldsymbol{t})\tau(x,\boldsymbol{t})^{-1}\exp\bigl[\mathrm{tr}\zeta(0,\boldsymbol{t}) - \mathrm{tr}\zeta(x,\boldsymbol{t})\bigr].$$

Since $\zeta(0,\boldsymbol{t}) - \zeta(x,\boldsymbol{t}) = -xR$, we obtain (38).

(ii)(a) The assertion follows form (34) and (38).

(b) It is well known ([14]) that $\tau(x,t)$ solves the Hirota equation (39). Since $\det \phi_a(x,t)$ and $\bigl(I_{\boldsymbol{p},\boldsymbol{c},a}(x,t)\bigr)^{-2}$ are both of the form $k(t)e^{cx}\tau(x,t)$ with a constant c and a function $k : \mathbf{R} \to \mathbf{R}$, they also obey the same Hirota equation (39) that $\tau(x,t)$ does. Q.E.D.

References

[1] R. Bucy and P. Joseph, Filtering for stochastic processes with applications to guidance, Intersci. Publ., 1968.

[2] R.H. Cameron and W.T. Martin, The Wiener measure of Hilbert neighbourhoods in the space of real continuous functions, J. Math. Phys. M.I.T., **23** (1944), 195–209.

[3] R.H. Cameron and W.T. Martin, Evaluation of various Wiener integrals by use of certain Sturm-Liouville differential equations, Bull. A.M.S., **51** (1945), 73–90.

[4] K. Hara and N. Ikeda, Quadratic Wiener functionals and dynamics on Grassmannians, Bull. Sci. math., **125** (2001), 481–528.

[5] T. Hida and L. Streit, On quantum theory in terms of white noise, Nagoya Math. J., **68** (1977), 21–34.

[6] N. Ikeda and S. Manabe, Asymptotic formulae for stochastic oscillatory integrals, in "Asymptotic problems in probability theory: Wiener functionals and asymptotics", ed. by K.D. Elworthy and N. Ikeda, Pitman Research Notes Math. **284**, Longman, 1993, 136–155.

[7] N. Ikeda and S. Manabe, Van Vleck-Pauli formula for Wiener integrals and Jacobi fields, in "Itô's stochastic calculus and probability theory", Springer-Verlag, 1996, 141–156.

[8] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, 2nd edition, North-Holland/Kodansha, 1989.

[9] K. Itô, Differential equations determining a Markoff process (in Japanese), Journ. Pan-Japan Math. Coll. No.1077 (1942). English transl. in "Kiyosi Itô Selected papers" ed. by D. Stroock and S.R.S. Varadhan, Springer, 1987, 42–75.

[10] M.L. Kleptsyna and A. Le Breton, A Cameron-Martin type formula for general Gaussian processes — A filtering approach, Stoch. Stoch. Rep., **72** (2002), 229–250.

[11] S. Kotani, One-dimensional random Schrödinger operators and Herglotz functions, in "Probabilistic methods in mathematical physics (Katata/Kyoto, 1985)" ed. by K. Itô and N. Ikeda, Academic Press, 1987, 219–250.

[12] S. Kotani, Probabilistic approach to reflectionless potentials (in Japanese), Symposium on random matrices and related topics at Tsukuba Univ., 2000.

[13] H. McKean, Stochastic integrals, Academic Press, New York, 1969.

[14] T. Miwa, M. Jimbo and E. Date, Solitons, Cambridge Univ. Press, Cambridge, 2000.

[15] Y. Okabe, On the stationary Gaussian process with T-positivity and its associated Langevin equation and S-matrix, Jour. Fac. Sci. Univ. Tokyo Sect. IA Math.,**26** (1979), 115–165.

[16] M. Sato, "Lectures of Mikio Sato (1984–1985)" noted by T. Umeda (in Japanese), RIMS Lecture Notes No.5, RIMS, Kyoto Univ., 1989.

[17] M. Sato and Y. Sato, Soliton equations as dynamical systems on infinite-dimensional Grassmann manifold, in "Nonlinear PDE in Applied Sci. US-Japan Seminar, Tokyo 1982", ed. by H. Fujita, P.D. Lax and G. Strang, North-Holland, 1983, 259–271.

[18] K. Takasaki, Integrable systems as deformations of $\mathcal{D}$-modules, Proc. Sympo. Pure Math., **49** (1989), 143–168.

[19] A.I. Yashin An extension of the Cameron-Martin result, J. Appl. Prob., **30** (1993), 247–251.

Nobuyuki Ikeda
Suzukakedai 2-40-3, Sanda 669-1322, Japan

Setsuo Taniguchi
Faculty of Mathematics, Kyushu University, Fukuoka 812-8581, Japan
e-mail: `taniguch@math.kyushu-u.ac.jp`

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 189–207

Homogenization on Finitely Ramified Fractals

Takashi Kumagai

Abstract.

Let X_t be a continuous time Markov chain on some finitely ramified fractal graph given by putting i.i.d. random resistors on each cell. We prove that under an assumption that a renormalization map of resistors has a non-degenerate fixed point, $\alpha^{-n}X_{\tau^n t}$ converges in law to a non-degenerate diffusion process on the fractal as $n \to \infty$, where α is a spatial scale and τ is a time scale of the fractal. Especially, when the fixed point of the renormalization map is unique, the diffusion is a constant time change of Brownian motion on the fractal. These results improve and extend our former results in [10].

§1. Introduction

In this paper, we consider the homogenization problem on uniform finitely ramified fractals, which is a class of finitely ramified fractals with a unique spatial scaling rate. We put random resistors on each cell of the fractal graph and set X_t be the corresponding continuous time Markov chain. Our aim is to show that $\alpha^{-n}X_{\tau^n t}$ converges in law to a non-degenerate diffusion process on the fractal as $n \to \infty$. Here α is the spatial scale and τ is a time scale of the fractal.

Homogenization of a diffusion process is interpreted as a limit theorem of a random process for changing scales. For $\mathbf{R}^d$ case, it is discussed that $\epsilon X_{t/\epsilon^2}$ converges to a constant time change of Brownian motion as $\epsilon \to 0$ under a condition that X_t has random diffusion coefficients or it moves in some random environment such as random scatterers. (See [8] for general references on homogenization of differential operators.) The martingale method has been well developed for this problem.

For the case of fractals, typical diffusions are sub-diffusive, in the sense that $E[|B_t|] \approx t^{1/d_w}$ as $t \to \infty$ for some $d_w > 2$. Those diffusions are not semi-martingales, thus we need a different approach. In [10], we

Received February 17, 2003.
Partially supported by Grant-in-Aid for Scientific Research (C)(2) 14540113.

developed a theory to be applied for homogenization problem on nested fractals, a class of finitely ramified fractals with good symmetries. In this paper, we inherits the basic approach of [10], which we now explain.

We first consider a Dirichlet form corresponding to the continuous time Markov chain with random resistors which are almost surely bounded from above and below by some non-degenerate resistor (this corresponds to the uniform ellipticity condition for the operator) and the distributions are *i.i.d.* for each cell. As a solution of a variation problem of the Dirichlet form, we induce a renormalization map F on the space of matrices, by which we can produce a new form that is a renormalization of the original one. We assume that there is a non-degenerate fixed point of the map. Then, what we should prove are the following:
1) Convergence of the iteration of the renormalization map to a fixed point and convergence of the forms.
2) Convergence of finite dimensional distributions and tightness.

In [10], we prove 1) under certain condition for an adjoint of a Fréchet derivative of the renormalization map at a fixed point (see Remark 3.6 1) for details). Unfortunately, the condition is not easy to check in general and there are examples (even for nested fractals) that the condition does not hold. On the other hand, the dynamics of the iteration of the renormalization map has been well studied recently for the finitely ramified fractals (see [16, 13, 15] etc.). In this paper, we apply the results to improve our former results in [10]. In Section 4, we will prove 1) under a very mild condition on the renormalization map (see Assumption 2.3). We note that the map F we study is on a infinite dimensional space since we have infinite number of resistors, whereas the renormalization map $\hat{F}$ studied in [16, 13, 15] is on a finite dimensional space. We thus make some efforts to show the stability of the map F from that of $\hat{F}$.

In general, non-degenerate fixed points of the renormalization map are not necessarily unique. When the uniqueness (up to constant multiples) is guaranteed, we can further show that the diffusion obtained as the limit is a constant time change of some special diffusion (which can be called Brownian motion) for the fractal. Especially, when we consider random resistors which are invariant under all reflection maps on a nested fractal, then the diffusion obtained is a constant time change of Brownian motion on the fractal.

For the proof of 2), uniform Harnack inequality and uniform heat kernel estimates of the Markov chains play important roles. Here we can adopt stability results of parabolic Harnack inequalities and heat kernel estimates which are actively studied recently for the fractal graph cases (see [2, 3, 4, 5]). In this paper, we skip the proof of 2) since we can apply the same argument as in [10], but note that we can shorten the proof by applying the results in [5].

The organization of the paper is as follows. In Section 2, we define uniform finitely ramified fractals (graphs), renormalization maps of resistors on them and briefly mention about Dirichlet forms and their heat kernel estimates on the fractals. In Section 3, we give our main theorem on homogenization. Section 4 is for the proof of 1) above. In Section 5, we state main propositions concerning 2) above. In Appendix, we give the proof of the stability results of the (finite dimensional) renormalization map $\hat{F}$ studied in [16, 13, 15].

§2. Uniform finitely ramified graphs and their Dirichlet forms

2.1. Uniform finitely ramified graphs

For $\alpha > 1$ and $I = \{1, 2, \cdots, N\}$, let $\{\Psi_i\}_{i\in I}$ be a family of α-*similitudes* on $\mathbf{R}^D$. An α-similitude is a map $\Psi_i \mathbf{x} = \alpha^{-1} U_i \mathbf{x} + \gamma_i$, $\mathbf{x} \in \mathbf{R}^D$ where U_i is a unitary map and $\gamma_i \in \mathbf{R}^D$. We will impose several assumption on this family. First, we assume

(H-0) $\{\Psi_i\}_{i\in I}$ satisfies the open set condition,

i.e., there is a non-empty, bounded open set W such that $\{\Psi_i(W)\}_{i\in I}$ are disjoint and $\cup_{i\in I} \Psi_i(W) \subset W$. As $\{\Psi_i\}_{i\in I}$ is a family of contraction maps, there exists a unique non-void compact set $\hat{K}$ such that $\hat{K} = \cup_{i\in I} \Psi_i(\hat{K})$. We assume

(H-1) $\hat{K}$ is connected.

Let Fix be the set of fixed points of the Ψ_i's, $i \in I$. A point $x \in Fix$ is called an *essential fixed point* if there exist $i, j \in I$, $i \neq j$ and $y \in Fix$ such that $\Psi_i(x) = \Psi_j(y)$. Let I_F be the set of $i \in I$ for which the fixed point of Ψ_i is an essential fixed point. We write $\hat{V}_0$ for the set of essential fixed points. Denote $\Psi_{i_1,\dots,i_n} = \Psi_{i_1} \circ \cdots \circ \Psi_{i_n}$. We further assume the following finitely ramified property.

(H-2) If $\{i_1, \dots, i_n\}, \{j_1, \dots, j_n\}$ are distinct sequences, then

$$\Psi_{i_1,\dots,i_n}(\hat{K}) \bigcap \Psi_{j_1,\dots,j_n}(\hat{K}) = \Psi_{i_1,\dots,i_n}(\hat{V}_0) \bigcap \Psi_{j_1,\dots,j_n}(\hat{V}_0).$$

Definition 2.1. ([5]) *A (compact) uniform finitely ramified fractal (u.f.r. fractal for short) $\hat{K}$ is a set determined by α-similitudes $\{\Psi_i\}_{i\in I}$ satisfying the assumption (H-0), (H-1), (H-2) and that $\sharp\hat{V}_0 \geq 2$.*

If we further assume the following symmetry condition, then $\hat{K}$ is called a (compact) nested fractals introduced in [12].

(SYM) If $x, y \in \hat{V}_0$, then the reflection in the hyperplane $H_{xy} = \{z \in \mathbf{R}^D : |z - x| = |z - y|\}$ maps $\hat{V}_n$ to itself, where

$$\hat{V}_n = \cup_{i_1,\cdots,i_n \in I} \Psi_{i_1,\ldots,i_n}(\hat{V}_0). \tag{2.1}$$

Thus, u.f.r. fractals form a class of fractals which is wider than nested fractals, and is included in the class of p.c.f. self-similar sets ([9]).

For each $n \geq 0$ and $i_1, \cdots, i_n \in I$, we call a set of the form $\Psi_{i_1,\cdots,i_n}(\hat{V}_0)$ an n-cell and $\Psi_{i_1,\cdots,i_n}(\hat{K})$ an n-complex. For $x, y \in \hat{K}$, $\{x_0, \cdots, x_m\}$ is called a n-chain from x to y if $x_0 = x, x_m = y$, $x_j \in \hat{V}_n$ for $1 \leq j \leq m-1$ and x_i, x_{i+1} are in the same n-complex for $0 \leq i \leq m-1$. We then have the following topological properties of u.f.r. fractals.

Lemma 2.2.
1) Each element in $\hat{V}_0$ belongs to only one n-cell for each $n \geq 0$.
2) Any 1-cell contains at most one element of $\hat{V}_0$.
3) For each $x \in \hat{V}_1$ and $y \in \hat{V}_0$, there exists a 1-chain $\{x_0, \cdots, x_m\}$ from x to y such that $x_1, \cdots, x_{m-1} \notin \hat{V}_0$.

Proof. 1) and 2) can be proved in the same way as [11] (Lemma 2.8 and Proposition 2.9) and [12] (Proposition *IV*.13 and Corollary *IV*.14). (They discuss for nested fractals, but the symmetry assumption is not used there.) For 3), we first note that any 1-junction is not an element of $\hat{V}_0$ due to 1), where $x \in \hat{V}_1$ is called a 1-junction if there exist $i \neq j \in I$ such that $x \in \Psi_i(\hat{V}_0) \cap \Psi_j(\hat{V}_0)$. Using (H-1) and (H-2), we can choose a 1-chain $\{x_0, \cdots, x_m\}$ such that $x_1, \cdots, x_{m-1}$ are 1-junctions. Since 1-junction is not an element of $\hat{V}_0$, we obtain the result. Q.E.D.

Next we define unbounded u.f.r. fractals. We assume without loss of generality that $\Psi_1(\mathbf{x}) = \alpha_1^{-1}\mathbf{x}$ and $\mathbf{0}$ belongs to $\hat{V}_0$. Set $K = \cup_{n=1}^{\infty} \alpha^n \hat{K}$. Then, clearly $\Psi_1(K) = K$. We call K an unbounded uniform finitely ramified fractal. Let $V = V_0 = \cup_{n=0}^{\infty} \alpha^n \hat{V}_n$ and $V_n = \alpha^{-n} V$ for $n \in \mathbf{Z}$. (Note that this labelling is the opposite to the one given in [5]. As n gets bigger, the graph distance between each vertex of V_n gets smaller and $V_{n-1} \subset V_n$.) Then, $K = Cl(\cup_{n \in \mathbf{Z}} V_n)$. For $n \in \mathbf{Z}$, we define n-cells and n-complexes similarly as the compact fractals.

We now introduce uniform finitely ramified graphs. These will be graphs with vertices V and a collection of edges B. In order to define the edges, we first define $\hat{B}_0 := \{\{x, y\} : x \neq y \in \hat{V}_0\}$. Then inside each 0-cell we place a copy of $\hat{B}_0$ and we denote by B the set of all the edges determined in this way. We call the graph (V, B) a uniform finitely ramified (u.f.r.) graph. If we construct the graph starting from a nested fractal, then it will be called a nested fractal graph. Let

$$\begin{aligned}\Omega &= \{\omega \in I^{\mathbf{Z}} : \text{ there is an } n \in \mathbf{Z} \text{ such that } \omega_k = 1, k \geq n\},\\ \Omega_+ &= \{\omega \in I^{\mathbf{N}} : \text{ there is an } n \in \mathbf{N} \text{ such that } \omega_k = 1, k \geq n\}.\end{aligned}$$

Then, there is a continuous map $\pi : \Omega \to \mathbf{R}^D$ such that $\pi(\omega) = \lim_{n\to\infty} \alpha^n \Psi_{\omega_n}(\Psi_{\omega_{n-1}}(\cdots(\Psi_{\omega_{-n}}(\mathbf{0}))\cdots))$. It is easy to see $K = \pi(\Omega)$. For any $\omega \in \Omega_+$ and $i \in I_F$, let $[\omega, i]$ denotes an element of Ω given by

$$[\omega, i](k) = \begin{cases} \omega_k, & k \geq 1 \\ i, & k \leq 0. \end{cases}$$

Then, $V = \{\pi([\omega, i]) : \omega \in \Omega_+, i \in I_F\}$.

2.2. Renormalization maps

Let $\mathcal{Q}$ be the set of $Q = \{Q_{ij} \in \mathbf{R} : i, j \in I_F \times I_F\}$ such that $Q_{ij} = Q_{ji}$ for any $i, j \in I_F$ and that $\sum_{j\in I_F} Q_{ij} = 0, i \in I_F$. $\mathcal{Q}$ is a vector space with an inner product $(\cdot,\cdot)_{\mathcal{Q}}$ given by

$$(Q, Q')_{\mathcal{Q}} = \sum_{j,k\in I_F} Q_{jk} Q'_{jk} = \text{Trace } Q^t Q', \qquad Q, Q' \in \mathcal{Q}.$$

For a set A, we denote $l(A) = \{f : A \to \mathbf{R}\}$. Let $\mathcal{Q}_+$ be the set of $Q \in \mathcal{Q}$ such that $\hat{S}_Q(\xi, \xi) \geq 0$ for any $\xi \in l(I_F)$, where

$$\hat{S}_Q(\xi, \xi) = -\sum_{i,j\in I_F} Q_{ij}\xi_i\xi_j = \frac{1}{2}\sum_{i,j\in I_F} Q_{ij}(\xi_i - \xi_j)^2.$$

Set $\|Q\|^2 = \sup_{\xi\in l(I_F)} \hat{S}_Q(\xi,\xi)/(\sum_{i\in I_F}\xi_i^2)$. Note that there exist $c_{2.1}$, $c_{2.2} > 0$ such that $c_{2.1}\|Q\|^2 \leq (Q, Q)_{\mathcal{Q}} \leq c_{2.2}\|Q\|^2$ for all $Q \in \mathcal{Q}_+$. We sometimes denote $\hat{S}_Q(\xi, \xi)$ as $\hat{S}_Q(\xi)$. Let $\mathcal{Q}_M$ be the set of $Q \in \mathcal{Q}$ such that $Q_{ij} \geq 0$ for any $i, j \in I_F$ with $i \neq j$. Also, let $\mathcal{Q}_{irr}$ be the set of $Q \in \mathcal{Q}_M$ such that $\hat{S}_Q(\xi, \xi) = 0$ if and only if ξ is constant. Note that $\mathcal{Q}_{irr} \subset \mathcal{Q}_M \subset \mathcal{Q}_+$.

Take $Q_* \in \text{Int}(\mathcal{Q}_M) := \{Q \in \mathcal{Q}_M : Q_{ij} > 0 \text{ for any } i \neq j \in I_F\}$ and let $\mathcal{X}_+ = \{X \in C(\Omega_+, \mathcal{Q}_+) : \text{there exists } C_0 > 0 \text{ such that } \hat{S}_{X(\omega)}(\xi,\xi) \leq C_0\hat{S}_{Q_*}(\xi,\xi) \text{ for any } \omega \in \Omega_+ \text{ and } \xi = (\xi_j)_{j\in I_F}\}$. Also, let $\mathcal{X}_M = \mathcal{X}_+ \cap C(\Omega_+, \mathcal{Q}_M)$, $\mathcal{X}_{irr} = \mathcal{X}_+ \cap C(\Omega_+, \mathcal{Q}_{irr})$ and $\text{Int}(\mathcal{X}_M) = \mathcal{X}_+ \cap C(\Omega_+, \text{Int}(\mathcal{Q}_M))$. Then $\mathcal{X}_+$ and $\mathcal{X}_M$ are convex cones. For any $X \in \mathcal{X}_+$, let S_X denote a non-negative definite bilinear form on $\mathbf{L}^2(V, d\nu_0)$ given by

$$S_X(u, u) = \frac{1}{2}\sum_{\omega\in\Omega_+} \hat{S}_{X(\omega)}(u(\pi([\omega,\cdot])), u(\pi([\omega,\cdot]))), \ u \in \mathbf{L}^2(V, d\nu_0).$$

Here ν_0 is a measure on V so that $\nu_0(\{x\}) = 1/N$ for all $x \in V$. If $X \in \mathcal{X}_M$, then S_X is a Dirichlet form on $\mathbf{L}^2(V, d\nu_0)$. So there is a Markov process which we denote by $\{P_X^x : x \in V\}$. We introduce an order relation $\leq$ in $\mathcal{X}_+$ as follows.

$$X \leq Y \text{ if } S_X(u,u) \leq S_Y(u,u) \text{ for all } u \in \mathbf{L}^2(V, d\nu_0).$$

The norm on $\mathcal{X}_+$ is given by $\|X\|^2 = \sup_{u \in \mathbf{L}^2(V,d\nu_0)} S_X(u,u)/\|u\|^2_{\mathbf{L}^2(V,d\nu_0)}$.

For any $X \in \mathcal{X}_+$, let $S_X^{\bar{F}} : \mathbf{L}^2(V, d\nu_0) \to [0, \infty)$ be given by

$$S_X^{\bar{F}}(u) = \inf\{S_X(v,v) : v \in \mathbf{L}^2(V, d\nu_0),\ v(\alpha x) = u(x),\ x \in V\}.$$

Let $S_X^{\bar{F}}(u,v) = \frac{1}{2}(S_X^{\bar{F}}(u+v) - S_X^{\bar{F}}(u) - S_X^{\bar{F}}(v)),\ u, v \in \mathbf{L}^2(V, d\nu_0)$. Then we see that $S_X^{\bar{F}}$ is a Dirichlet form on $\mathbf{L}^2(V, d\nu_0)$. Moreover, by the self-similarity of K, we see that there is a renormalization map $\bar{F} : \mathcal{X}_+ \to \mathcal{X}_+$ such that $S_X^{\bar{F}}(u) = S_{\bar{F}(X)}(u,u)$ for all $X \in \mathcal{X}_+$ and $u \in \mathbf{L}^2(V, d\nu_0)$. Let $\iota : \mathcal{Q}_+ \to \mathcal{X}_+$ be such that $\iota(Q)(\omega) = Q$ for all $\omega \in \Omega_+$ and $Q \in \mathcal{Q}_+$. We define a renormalization map $\tilde{F} : \mathcal{Q}_+ \to \mathcal{Q}_+$ as $\tilde{F}(Q) = \bar{F}(\iota(Q))(\omega)$ for $\omega \in \Omega_+$ (it is independent of the choice of $\omega \in \Omega_+$). Note that $\bar{F}, \tilde{F}$ is in general a non-linear map. By Schauder's fixed point theorem, we know that there exists $Q_* \in \mathcal{Q}_M$ (with $(Q_*)_{ij} > 0$ for some $i \neq j$) and $\rho_{Q_*} > 0$ such that $\tilde{F}(Q_*) = \rho_{Q_*}^{-1} Q_*$. Throughout this paper, we assume the following.

Assumption 2.3. *1) For each $Q \in \mathcal{Q}_{irr}$, there exists $l = l(Q) \in \mathbf{N}$ such that $\tilde{F}^n(Q) \in Int(\mathcal{Q}_M)$ for all $n \geq l$.*
2) There exists $Q_0 \in Int(\mathcal{Q}_M)$ and $\rho_{Q_0} > 0$ such that $\tilde{F}(Q_0) = \rho_{Q_0}^{-1} Q_0$.

Remark 2.4. *1) By Corollary* 6.20 *of* [1], *$\rho_{Q_0} > 0$ is uniquely determined, i.e., if $Q_1, Q_2 \in \mathcal{Q}_{irr}$ satisfies $\tilde{F}(Q_j) = \rho_{Q_j}^{-1} Q_j$ $(j = 1, 2)$ with $\rho_{Q_1}, \rho_{Q_2} > 0$, then $\rho_{Q_1} = \rho_{Q_2} = \rho_{Q_0}$. In the class of fractal graphs we consider, we can prove $\rho_{Q_0} > 1$ (see* [9] *etc.).*
2) A sufficient condition for Assumption 2.3 1) is the following.
(H-3) There exists $l \in \mathbf{N}$ such that for each $x, y \in \hat{V}_0$, there is a l-chain $\{x_0, \cdots, x_m\}$ from x to y such that for each $1 \leq i \leq m-2$, there is a l-cell containing x_i and x_{i+1} that does not contain any element of $\hat{V}_0$.

Indeed, if (H-3) holds, it is easy to show $\tilde{F}^n(Q) \in Int(\mathcal{Q}_M)$ for $n \geq l$, $Q \in \mathcal{Q}_{irr}$ by observing the corresponding Markov chain on $\hat{V}_n$.
3) Every nested fractals satisfy Assumption 2.3 1) and 2). Indeed, (H-3) can be shown for nested fractals using (SYM) and [11] *Lemma* 2.10 *(*[12] *Proposition IV.*11*), so that 1) holds. 2) is proved in* [11] *Theorem* 3.10 *and in* [12] *Theorem V.*5.

Set $F = \rho_{Q_0}\bar{F} : \mathcal{X}_+ \to \mathcal{X}_+$ and $S_X^F(u) = \rho_{Q_0}S_X^{\bar{F}}(u)$ for $u \in \mathbf{L}^2(V, d\nu_0)$. Set $\hat{F} = \rho_{Q_0}\tilde{F}$ in the same way.

2.3. Dirichlet forms and heat kernel estimates

For $u, v \in l(\hat{V}_n)$, define

$$\hat{\mathcal{E}}^n_{Q_0}(u,v) = \rho^n_{Q_0} \sum_{i_1,\cdots,i_n \in I} \hat{S}_{Q_0}(u \circ \Psi_{i_1,\dots,i_n}, v \circ \Psi_{i_1,\dots,i_n}). \tag{2.2}$$

Let $\hat{\nu}$ be a normalized Hausdorff measure on $\hat{K}$. Then, the following is known (see for example, [6, 9]).

Theorem 2.5. *Let $Q_0 \in Int(\mathcal{Q}_M)$ be as in Assumption 2.3 2), i.e. $\hat{F}(Q_0) = Q_0$. Then, there is a local regular Dirichlet form $(\hat{\mathcal{E}}, \hat{\mathcal{F}})$ in $\mathbf{L}^2(\hat{K}, d\hat{\nu})$ satisfying the following,*

$$\begin{aligned} \hat{\mathcal{F}} &= \{u \in C(\hat{K}, \mathbf{R}) : \sup_n \hat{\mathcal{E}}^n_{Q_0}(u,u) < \infty\}, \\ \hat{\mathcal{E}}(u,v) &= \lim_{n\to\infty} \hat{\mathcal{E}}^n_{Q_0}(u,v) \qquad \textit{for } \ u, v \in \hat{\mathcal{F}}. \end{aligned}$$

For each $m \in \mathbf{N}$, let $K_m = \alpha^m \hat{K}$ and define $\sigma_m : C(K_m, \mathbf{R}) \to C(\hat{K}, \mathbf{R})$ by $\sigma_m u(x) = u(\alpha^m x)$ for $x \in \hat{K}$. Set $\mathcal{F}_{<m>} = \sigma_{-m}\hat{\mathcal{F}}$, $\mathcal{E}_{<m>}(u,v) = \rho_{Q_0}^{-m}\hat{\mathcal{E}}(\sigma_m u, \sigma_m v)$ for $u, v \in \mathcal{F}_{<m>}$. Let ν be a Hausdorff measure on K such that $\nu|_{\hat{K}} = \hat{\nu}$ and $N\nu = \nu \circ \Psi_1^{-1}$. Now let

$$\begin{aligned} \mathcal{F} &= \{u \in l(K) : u|_{K_m} \in \mathcal{F}_{<m>} \text{ for all } \ m \in \mathbf{N}, \\ &\qquad \lim_{m\to\infty} \mathcal{E}_{<m>}(u|_{K_m}, u|_{K_m}) < \infty\} \cap \mathbf{L}^2(K, d\nu) \\ \mathcal{E}(u,v) &= \lim_{m\to\infty} \mathcal{E}_{<m>}(u|_{K_m}, v|_{K_m}) \qquad \text{for } \ u, v \in \mathcal{F}. \end{aligned}$$

Then the following holds.

Theorem 2.6. *$(\mathcal{E}, \mathcal{F})$ is a local regular Dirichlet form on $\mathbf{L}^2(K, d\nu)$. $\mathcal{F} \subset C(K, \mathbf{R})$ and this form has the following scaling property,*

$$\mathcal{E}(u,v) = \rho_{Q_0}\mathcal{E}(u \circ \Psi_1, v \circ \Psi_1) \qquad \textit{for } \ u, v \in \mathcal{F}.$$

Finally, we will mention heat kernel estimates for Markov chains on u.f.r. graphs. For $X \in \mathrm{Int}(\mathcal{X}_M)$ and $x \neq y \in V$, define

$$R_X(x,y) = \left(\inf\{S_X(u,u) : u \in l(V), u(x) = 1, u(y) = 0\}\right)^{-1}.$$

Let $R_X(x,x) = 0$ for $x \in V$. Then, $R_X(\cdot,\cdot)$ is a metric which is called a resistance metric. By simple modifications of the proof of Corollary 4.12 in [5], the following holds (note that as we will mention later in Remark 3.6 4), Assumption 2.3 of [5] always holds under our Assumption 2.3).

Theorem 2.7. *For each $X \in Int(\mathcal{X}_M)$, let $p_k^X(\cdot,\cdot)$ be the heat kernel of the discrete time Markov chain which is induced from the continuous time Markov chain corresponding to $(S_X, \mathbf{L}^2(V, d\nu_0))$. Then, there exists $c_{2.3}, \cdots, c_{2.6} > 0$ (which depend on X) and $0 < \gamma_1 \le \gamma_2$ such that for each $x, y \in V$ and $k \ge d(x,y)$,*

$$\begin{aligned} p_k^X(x,y) &\le c_{2.3} k^{-\frac{S}{S+1}} \exp\left(-c_{2.4}\left(\frac{R_X(x,y)^{S+1}}{k}\right)^{\gamma_1}\right), \\ p_k^X(x,y) + p_{k+1}^X(x,y) &\ge c_{2.5} k^{-\frac{S}{S+1}} \exp\left(-c_{2.6}\left(\frac{R_X(x,y)^{S+1}}{k}\right)^{\gamma_2}\right), \end{aligned}$$

where $S = \log N / \log \rho_{Q_0}$ and $d(\cdot,\cdot)$ is a graph distance.

We note that similar heat kernel estimates for $(\hat{\mathcal{E}}, \hat{\mathcal{F}})$ and $(\mathcal{E}, \mathcal{F})$ (given in Theorem 2.5 and 2.6) can be also obtained (cf. [6]).

Let $\beta > 0$. We say $(S_X, \mathbf{L}^2(V, d\nu_0))$ satisfies $(PHI(\beta))$, a parabolic Harnack inequality of order β if whenever $u(n,x) \ge 0$ is defined on $[0,4N] \times \bar{B}(y,2r)$ and satisfies

$$u(n+1,x) - u(n,x) = \mathcal{L}u(n,x) \qquad (n,x) \in [0,4N] \times B(y,2r),$$

($\mathcal{L}$ is the corresponding difference operator), then

$$\max_{\substack{N \le n \le 2N \\ x \in B(y,r)}} u(n,x) \le c_{2.7} \min_{\substack{3N \le n \le 4N \\ x \in B(y,r)}} (u(n,x) + u(n+1,x)),$$

where $N \ge 2r$ and $c_{2.8} r^\beta \le N \le c_{2.9} r^\beta$ (cf. [2, 3, 4, 5]). By Theorem 2.7 and a standard argument, we can deduce the following.

Proposition 2.8.
$(S_X, \mathbf{L}^2(V, d\nu_0))$ satisfies $(PHI(S+1))$ w.r.t. the resistance metric.

§3. Homogenization

In this section, we will state our main theorem. First, we give some definition for later use. Let $\hat{V}_0 = \{a_i : i \in I_F\}$. For $Q^* \in \mathrm{Int}(\mathcal{Q}_M)$, we define a matrix $A_{k,Q^*} \in l(\hat{V}_0)$, $k \in I$ by

$$(3.1) \qquad (A_{k,Q^*})_{ij} = P_{Q^*}^{\Psi_k(a_i)}\big(w^1(\tau_{\hat{V}_0}) = a_j\big),$$

where w^1 is a Markov chain on $\hat{V}_1$ whose transition probability is determined by Q^* and $\tau_{\hat{V}_0} = \inf\{n \ge 0 : w_n^1 \in \hat{V}_0\}$. Then, by Lemma 2.2 3), the following clearly holds for u.f.r. graphs.

Lemma 3.1. $0 < (A_{k,Q^*})_{ij} < 1$ *if $k \ne i$ and $(A_{k,Q^*})_{kj} = \delta_{kj}$.*

For any $X \in \mathcal{X}_+$ and any $Q_* \in \mathrm{Int}(\mathcal{Q}_M)$ with $\hat{F}(Q_*) = Q_*$, let $S_X^{H_{Q_*}} : \mathbf{L}^2(V, d\nu_0) \to [0, \infty)$ be given by

$$S_X^{H_{Q_*}}(u) = \rho_{Q_0} S_X(v, v) \qquad \text{for } \ u \in \mathbf{L}^2(V, d\nu_0),$$

where $v \in \mathbf{L}^2(V, d\nu_0)$ satisfies $v(\alpha x) = u(x)$, $x \in V$, and v is Q_*-harmonic on $V \setminus (\alpha^{-1}V)$, i.e.,

$$v(\pi([\omega \cdot i, j])) = \sum_{k \in I_F} (A_{i,Q_*})_{jk} u(\pi([\omega, k])) \qquad \text{for } \ i \in I, j \in I_F.$$

Here $\omega \cdot i \in \Omega_+$ is given by $(\omega \cdot i)_n = \omega_{n-1}$, $n \geq 2$ and $(\omega \cdot i)_1 = i$. In the same way as we did for $S_X^{\bar{F}}$, we can define a Dirichlet form $S_X^{H_{Q_*}}(\cdot, \cdot)$ on $\mathbf{L}^2(V, d\nu_0)$. It is easy to see that $S_X^{H_{Q_*}}(u) = S_{H_{Q_*}(X)}(u, u)$ where $H_{Q_*}(X)(\omega) = \rho_{Q_0} \sum_{k \in I} {}^t A_{k,Q_*} X(\omega \cdot k) A_{k,Q_*}$ for all $X \in \mathcal{X}_+$ and $u \in \mathbf{L}^2(V, d\nu_0)$. We define a map $\hat{H}_{Q_*} : \mathcal{Q}_+ \to \mathcal{Q}_+$ as $\hat{H}_{Q_*}(Q) = H_{Q_*}(\iota(Q))(\omega)$ for $\omega \in \Omega_+$ (it is independent of the choice of $\omega \in \Omega_+$).

Definition 3.2.
Let μ be a probability measure on $\mathcal{X}_M$ satisfying the following.
1) $\{X(\omega) : \omega \in \Omega_+\}$ are independently identically distributed $\mathcal{Q}_M$-valued random variables under μ.
2) $\mu(\{X \in \mathcal{X}_M : X(\omega) \in \mathcal{Q}_{C_1 Q_0, C_2 Q_0}$ for all $\omega \in \Omega_+\}) = 1$ for some $C_1, C_2 > 0$, where $\mathcal{Q}_{C_1 Q_0, C_2 Q_0} := \{Q \in \mathcal{Q}_M : C_1 Q_0 \leq Q \leq C_2 Q_0\}$.

The following properties are easy, but important.

Proposition 3.3. *Let $Q_0, Q_* \in Int(\mathcal{Q}_M)$ be as above.*
1) $F : \mathcal{X}_M \to \mathcal{X}_M$ and $H_{Q_*} : \mathcal{X}_M \to \mathcal{X}_M$ are continuous maps.
2) $F(\iota(Q_0)) = \iota(Q_0)$, $F(\iota(Q_*)) = H_{Q_*}(\iota(Q_*)) = \iota(Q_*)$.
3) If $X, Y \in \mathcal{X}_+$ and $X \leq Y$, then $F(X) \leq F(Y)$ and $H_{Q_*}(X) \leq H_{Q_*}(Y)$.
4) $F(X) \leq H_{Q_*}(X)$ for all $X \in \mathcal{X}_+$.
5) For any $X, Y \in \mathcal{X}_+$ and $a, b \geq 0$, $F(aX + bY) \geq aF(X) + bF(Y)$ and $H_{Q_*}(aX + bY) = aH_{Q_*}(X) + bH_{Q_*}(Y)$.
6) $E_\mu[F(X)] \leq F(E^\mu[X])$ for all $X \in \mathcal{X}_+$.

Note that the same results hold for $\hat{F}$ and $\hat{H}_{Q_*}$.

Let F^n be the n-th iteration of F. Then we have the following key theorem.

Theorem 3.4. *Under Assumption 2.3, there exists $Q_\mu \in Int(\mathcal{Q}_M)$ such that for all $\omega \in \Omega_+$,*

$$Q_\mu = \lim_{n \to \infty} F^n(X)(\omega) \qquad \text{in } \ \mathbf{L}^1(\mathcal{Q}_M, \mu). \tag{3.2}$$

Since $\hat{F}(Q_\mu) = Q_\mu \in \mathrm{Int}(\mathcal{Q}_M)$, we can construct a local regular Dirichlet form on K using Q_μ (see Theorem 2.6). We denote the corresponding diffusion as $(X_\mu, \{P_\mu^x\}_{x\in K})$. We now state our main theorem.

Theorem 3.5. *Let μ be the probability measure on $\mathcal{X}_M$ as in Definition 3.2 and let $\tau_{Q_0} := \rho_{Q_0} N$. Under Assumption 2.3, the following holds.*

$$E^{P_X^{x_n}}[f(\alpha^{-n} w(\tau_{Q_0}^n \cdot))] \to E^{P_\mu^{x_\infty}}[f(w(\cdot))] \quad \text{as } n \to \infty$$

in probability under μ, for any $f \in C_b(D([0,\infty), K) \to \mathbf{R})$ and any sequence $\{x_n\}_{n=1}^\infty \subset V$ with $\alpha^{-n} x_n \to x_\infty \in K$ as $n \to \infty$. Here the expectations are taken over $\omega \in D([0,\infty), K)$.

Further, if there is a convex cone $\mathcal{X}_{sub} \subset \mathcal{X}_{irr}$ such that the following holds; a) $F(\mathcal{X}_{sub}) \subset \mathcal{X}_{sub}$, b) there exists a unique (up to constant multiples) $Q \in \mathcal{X}_{sub} \cap Int(\mathcal{X}_M)$ which satisfies $F(Q) = Q$, c) the support of μ is in $\mathcal{X}_{sub}$. Then P_μ is a constant time change of the diffusion constructed from Q on K.

Remark 3.6. *1) In [10], similar statement is given under the assumption that there exists $Q_1 \in \mathcal{Q}_{irr}$ such that $\hat{H}_{Q_0}^*(Q_1) = Q_1$ where $\hat{H}_{Q_0}^*$ is an adjoint operator of $\hat{H}_{Q_0}$ in $\mathcal{Q}$ ([10] Assumption 3.1). In general, it is not easy to check this assumption and there is a example in nested fractals that this does not hold.*
2) Let $\mathcal{I}_I$ be the set of all bijective maps σ on I such that $\sigma(I_F) = I_F$, and let G be a subgroup of $\mathcal{I}_I$. Then, as in [10] Section 7, we can obtain similar results for random resistors on $\mathcal{X}_M^G$, a subcone of $\mathcal{X}_M$ which consists of G-invariant elements, if a), b), c) in Theorem 3.5 holds for $\mathcal{X}_M^G$. Especially, we can prove the following; For nested fractals, let G_0 be a subgroup of $\mathcal{I}_I$ generated by all the reflection maps and suppose that the support of μ is in $\mathcal{X}_M^{G_0}$. Then P_μ in Theorem 3.5 is a constant time change of Brownian motion on the nested fractal. (This is because, it is known that a non-degenerate fixed point for G_0-invariant resistors on nested fractals is unique up to constant multiples; see [16, 13, 15].)
3) Note that non-degenerate fixed points of $\hat{F}$ is not necessarily unique even for nested fractals. In [1] Example 6.13, one parameter family of non-degenerate fixed points on the Vicsek set are given. The homogenization problem for this particular fractal is studied in [7].
4) By Theorem 3.4 (or Proposition 4.1), we see that Assumption 2.3 in [5] always holds under our Assumption 2.3.

§4. Convergence of Dirichlet forms

In this section, we will prove Theorem 3.4 and show a convergence of the corresponding forms (Proposition 4.4).

The next proposition is a restricted version of the result by Peirone ([15]), whose original ideas come from Sabot ([16]). For completeness, we give the proof in Appendix A.

Proposition 4.1. *Under Assumption 2.3, for each $M \in \mathcal{Q}_{irr}$, there exists $Q_M \in Int(\mathcal{Q}_M)$ such that*

$$Q_M = \lim_{n\to\infty} \hat{F}^n(M) \tag{4.1}$$

For the proof of Theorem 3.4, we use two lemmas in [10]. Let $H_{Q_*}^n$ be the n-th iteration of H_{Q_*}.

Lemma 4.2. *Assume that $Q_* \in Int(\mathcal{Q}_M)$ satisfies $\hat{F}(Q_*) = Q_*$. Then, there exist $c_{4.1} > 0$ and $0 < \epsilon < 1$ such that*

$$E^\mu[\|H_{Q_*}^n(X)(\omega) - H_{Q_*}^n(E^\mu[X])(\omega)\|^2] \le c_{4.1}(1-\epsilon)^n, \ \forall \omega \in \Omega_+, n \ge 1.$$

In particular,

$$\lim_{n\to\infty} \|H_{Q_*}^n(X)(\omega) - H_{Q_*}^n(E^\mu[X])(\omega)\| = 0, \ \mu\text{-a.e. } X, \ \forall \omega \in \Omega_+.$$

Proof. By the linearity of H_{Q_*}, $E^\mu[H_{Q_*}(X)(\omega)] = H_{Q_*}(E^\mu[X])(\omega)$. Then the proof is basically the same as that of Lemma 4.1 in [10]. Q.E.D.

Lemma 4.3. ([10], Lemma 4.2) *Let $\{Y_n\}_{n=1}^\infty$ be random variables such that $\sup_n E[Y_n^2] < \infty$. Let $Y = \limsup_{n\to\infty} Y_n$ and assume that $\lim_{n\to\infty} E[Y_n] = E[Y]$. Then $\lim_{n\to\infty} E[|Y - Y_n|] = 0$.*

Proof of Theorem 3.4. Let $R_m = E^\mu[F^m(X)(\omega)]$ (R_m is independent of ω). By Proposition 4.1, for each $m \in \mathbf{N}$, there exists $Q_m \in \mathrm{Int}(\mathcal{Q}_M)$ such that $\lim_{n\to\infty} \hat{F}^n(R_m) = Q_m$ and $\hat{F}(Q_m) = Q_m$. On the other hand, by Proposition 3.3 6) we see

$$\hat{F}^n(R_m) \ge R_{n+m} \qquad \forall m, n \in \mathbf{N} \cup \{0\}, \tag{4.2}$$

so that $Q_m \ge Q_{n+m}$. Denote the limit of $\{Q_m\}$ as Q_+, then $\hat{F}(Q_+) = Q_+$. In particular, $Q_+ \in \mathrm{Int}(\mathcal{Q}_M)$ due to Assumption 2.3 1). For any $\epsilon > 0$, there exists $N_\epsilon \in \mathbf{N}$ such that

$$(1+\epsilon)Q_+ \ge R_m \qquad \forall m \ge N_\epsilon. \tag{4.3}$$

Indeed, if this does not hold, then because $\mathcal{Q}_{(1+\epsilon)Q_+,C_2Q_0}$ is compact, there exists a subsequence $\{l_j\}$ such that $R_{l_j} \geq (1+\epsilon)Q_+$ and $\lim_{j\to\infty} R_{l_j}$ $=: \bar{R}$ exists. On the other hand, by (4.2), we have $\hat{F}^{l_{j'}-l_j}(R_{l_j}) \geq R_{l_{j'}}$ for all $j' \geq j$ so that $Q_+ \geq \bar{R}$, which is a contradiction. By definition of $\{Q_m\}$, for each m and $\epsilon > 0$, there exists $L_{m,\epsilon}$ such that $(1-\epsilon)Q_m \leq \hat{F}^n(R_m)$ for all $n \geq L_{m,\epsilon}$. Combining these facts and noting $\hat{H}^n_{Q_+}(R_m) \geq \hat{F}^n(R_m)$, we have

$$(4.4) \qquad (1-\epsilon)Q_+ \leq \hat{H}^n_{Q_+}(R_m) \leq (1+\epsilon)Q_+ \qquad \forall n \geq L_{m,\epsilon}, m \geq N_\epsilon.$$

On the other hand, by Lemma 4.2, we have

$$(4.5) \qquad \begin{aligned} &\lim_{n\to\infty} \|H^n_{Q_+}(F^m(X))(\omega) - H^n_{Q_+}(\iota(R_m))(\omega)\| \\ = \; &\lim_{n\to\infty} \|H^n_{Q_+}(F^m(X))(\omega) - \hat{H}^n_{Q_+}(R_m)\| = 0 \end{aligned}$$

μ-a.e. X and for all $\omega \in \Omega_+$. Since $H^n_{Q_+}(F^m(X))(\omega) \geq F^{n+m}(X)(\omega)$, we see that the following holds for some $N'_{\epsilon,\omega} \in \mathbf{N}$,

$$(4.6) \qquad (1+\epsilon)Q_+ \geq F^m(X)(\omega) \qquad \mu\text{-a.e. } X,\ \forall \omega \in \Omega_+, m \geq N'_{\epsilon,\omega}.$$

We now consider more about $\hat{H}_{Q_+}$. It is easy to see

$$\sup_n |||\hat{H}^n_{Q_+}||| := \sup_n \sup_{Q\in\mathcal{Q}_M, \|Q\|=1} \|\hat{H}^n_{Q_+}(Q)\| < \infty$$

(see Lemma 4.3 in [10]). Using this, we see that the size of each Jordan cell corresponding to the largest eigenvalue of $\hat{H}_{Q_+}$ is 1. We thus obtain that there exists an orthogonal projection $P_0 : \mathcal{Q}_M \to \mathcal{Q}_M$ so that for each $k \in \mathbf{N}$, there exists $n_k \in \mathbf{N}$ such that

$$(4.7) \qquad |||\hat{H}^{n_k}_{Q_+} - P_0||| \leq 2^{-k}.$$

By (4.4) and (4.7), we have $R_m \geq P_0 R_m \geq (1-\epsilon)Q_+$ for all $m \geq N_\epsilon$. Together with (4.3), we have

$$(4.8) \qquad \lim_{n\to\infty} R_n = Q_+.$$

Now, by Fatou's lemma and (4.8),

$$(4.9) \qquad E^\mu\big[\limsup_{n\to\infty} \hat{S}_{F^n(X)(\omega)}(u,u)\big] \geq \limsup_{n\to\infty} \hat{S}_{R_n}(u,u) = \hat{S}_{Q_+}(u,u),$$

for all $\omega \in \Omega_+$, $u \in l(V_\omega)$, where $V_\omega := \{\pi([\omega, i]) : i \in I_F\}$ is a 0-cell whose address is ω. By (4.6) and (4.9), we have

$$\limsup_{n\to\infty} \hat{S}_{F^n(X)(\omega)}(u,u) = \hat{S}_{Q_+}(u,u),$$

μ-a.e. X and for all $\omega \in \Omega_+$, $u \in l(V_\omega)$. Applying Lemma 4.3 with $Y_n = \hat{S}_{F^n(X)(\omega)}(u,u)$ and $Y = \hat{S}_{Q_+}(u,u)$ (Y is non-random), we have

$$\lim_{n\to\infty} E^\mu\left[|\hat{S}_{F^n(X)(\omega)}(u,u) - \hat{S}_{Q_+}(u,u)|\right] = 0 \qquad \forall \omega \in \Omega_+,\ u \in l(V_\omega).$$

Since $l(V_\omega)$ is finite dimensional, we obtain (3.2) where $Q_\mu = Q_+$. Q.E.D.

Using Theorem 3.4, we can prove the convergence of forms.

Proposition 4.4. *For all $u \in \mathbf{L}^2(V, d\nu_0)$,*

$$\lim_{n\to\infty} E^\mu[|S_{F^n(X)}(u,u) - S_{\iota(Q_\mu)}(u,u)|] \to 0.$$

Proof. When the support of u is in one 0-cell whose address is ω, then the result is clear by Theorem 3.4. When u is compactly supported, we can decompose the form into finite number of forms on 0-cells, so the result still holds. It is then a routine work to show the result for all $u \in \mathbf{L}^2(V, d\nu_0)$. Q.E.D.

We note here that there are several errors in [10] Section 4, where the proof of the theorem corresponding to our Theorem 3.4 is given. They can be fixed, but since we have given a proof of the improved theorem, we omit mentioning where the errors are and how to fix them.

§5. Proof of Theorem 3.5

Now that we obtain Proposition 4.4, the proof of Theorem 3.5 is basically the same as the proof of Theorem 3.6 in [10]. Here we will just state key propositions and briefly comments how to prove them. For detailed arguments, we refer to Section 5 in [10].

As before, define $K_m = \alpha^m \hat{K}$. We consider processes killed at $\alpha^m \hat{V}_0 \setminus \{0\}$. Set $\mathcal{F}_{n,m} = \{u \in l(V_n) : u|_{K \setminus K_m} = 0\}$. For $u, v \in \mathcal{F}_{n,m}$ and $X \in \mathcal{X}_M$, we set $\mathcal{E}_X^{n,m}(u,v) = \rho_{Q_0}^n S_X(u \circ \Psi_1^n, v \circ \Psi_1^n)$. Then, $(\mathcal{E}_X^{n,m}, \mathcal{F}_{n,m})$ is a regular Dirichlet form. We denote the corresponding process $(Y^{X,n,m}, \{P^x_{X,n,m}\}_{x \in V_n \cap K_m})$ and the corresponding generator $L^{(X,n,m)}$. Also, let $(Y^{\mu,m}, \{P^x_{\mu,m}\}_{x \in K_m})$ denote the process corresponding to $(\mathcal{E}_{Q_\mu}, \mathcal{F}_m)$, where $\mathcal{E}_{Q_\mu}$ is a form constructed from Q_μ in Theorem

3.4 and $\mathcal{F}_m = \{f \in \mathcal{F} : f|_{K \setminus K_m} = 0\}$. $L^{(m)}$ is the corresponding generator.

The first key proposition is the following convergence of finite dimensional distribution. This can be obtained by using (H-2), Theorem 2.7 and Proposition 4.4. See Proposition 5.7 in [10] for the proof.

Proposition 5.1. *Let $\{x_n\}_{n=1}^{\infty} \subset V$ be a sequence such that $\alpha^{-n}x_n \to x_\infty \in K_m$. Then, for all $0 < t_1 < \cdots < t_k$,*

$$E^{P^{\alpha^{-n}x_n}_{X,n,m}}[f_1(\omega_{t_1})\cdots f_k(\omega_{t_k})] \to E^{P^{x_\infty}_{\mu,m}}[f_1(\omega_{t_1})\cdots f_k(\omega_{t_k})],$$

in probability under μ, for any $f_1, \cdots, f_k \in C(K_m, \mathbf{R})$.

For a process Z on K let $T_0^r(Z) = \inf\{t \geq 0 : Z(t) \in V_r\}$ and define inductively $T_i^r(Z) = \inf\{t > T_{i-1}^r(Z) : Z(t) \in V_r \setminus Z(T_{i-1}^r(Z))\}$ for $i \in \mathbf{N}$. Then the following holds.

Lemma 5.2. *Let $\{x_n\}_n \subset V$ be as in Proposition 5.1. Then there exist $\gamma, c_{5.1}, c_{5.2} > 0$ such that the following holds for $s \geq 0$, μ-a.e. X.*

$$\limsup_{n\to\infty}\sup_{i\geq 0} P^{\alpha^{-n}x_n}_{X,n,m}(T_{i+1}^r(Y^{X,n,m}) - T_i^r(Y^{X,n,m}) \leq s) \leq c_{5.1}e^{-c_{5.2}(\tau_{Q_0}^r s)^{-\gamma}}.$$

To show this, uniform (elliptic) Harnack inequality for $L^{(X,n,m)}$ is important (see [10] Lemma 5.8). In our case, we can obtain it easily by using Proposition 2.8. See [10] Lemma 5.10 for the detailed proof (thanks to Proposition 2.8, the proof can be shortened).

Using Lemma 5.2, the following tightness is deduced by a standard argument (see Proposition 5.11 in [10]).

Proposition 5.3. *Let $\{x_n\}_n \subset V$ be as in Proposition 5.1. Then $\{P^{\alpha^{-n}x_n}_{X,n,m};\ n \geq 1\}$ is tight (pre-compact) in $D([0,\infty), K_m)$ for μ-a.e. X.*

By Proposition 5.1 and Proposition 5.3, we have the killed process version of Theorem 3.5. Using Lemma 5.2 again, it is easy to deduce the full version of Theorem 3.5.

§Appendix A. Proof of Proposition 4.1

To start with, we prepare several results for the proof. First, we define Hilbert's projective metric on $\mathcal{Q}_+$ (cf. [14]). For $X, Y \in \mathcal{Q}_+$, let

$$h_+(X,Y) = \inf\{\alpha > 0 : X \leq \alpha Y\},\ h_-(X,Y) = \sup\{\alpha > 0 : \alpha Y \leq X\}.$$

Clearly, $h_-(X,Y) \leq h_+(X,Y)$. Define

$$h(X,Y) = \log\frac{h_+(X,Y)}{h_-(X,Y)}.$$

Note that $h(aX, bY) = h(X, Y)$ for all $a, b > 0$ and $h(X, Y) = 0$ if and only if $X = aY$ for some $a > 0$, so that $h(\cdot,\cdot)$ is not a metric. But it is a metric on $\{X \in \mathcal{Q}_+ : \|X\| = 1\}$. Using Proposition 3.3, it is easy to prove the following (cf. [13] Section 3, [15] Remark 3.2 and [16] Proposition 3.3).

Proposition A.1.
1) $h_+(\hat{F}(X), \hat{F}(Y)) \leq h_+(X, Y)$ *and* $h_-(\hat{F}(X), \hat{F}(Y)) \geq h_-(X, Y)$ *for all* $X, Y \in \mathcal{Q}_+$. *In particular,* $h(\hat{F}(X), \hat{F}(Y)) \leq h(X, Y)$.
2) If $Q_* \in Int(\mathcal{Q}_M)$ *satisfies* $\hat{F}(Q_*) = Q_*$, *then for each* $n \in \mathbf{N} \cup \{0\}$ *and each* $X \in \mathcal{Q}_{irr}$, $h_-(X, Q_*)Q_* \leq \hat{F}^n(X) \leq h_+(X, Q_*)Q_*$.

For $X, Y \in \mathcal{Q}_+$, let $A^{\pm}(X, Y) = \{u \in l(\hat{V}_0) : u$ is non-constant, $\hat{S}_X(u, u) = h_{\pm}(X, Y)\hat{S}_Y(u, u)\}$. Also, for each $Q \in \mathcal{Q}_{irr}$ and $u \in l(\hat{V}_0)$, define $\mathcal{H}_{n,Q}(u)$ as a unique function on $\hat{V}_n$ so that

$$\hat{S}_{\hat{F}^n(Q)}(u, u) = \hat{\mathcal{E}}^n_Q(\mathcal{H}_{n,Q}(u), \mathcal{H}_{n,Q}(u)),$$

where $\hat{\mathcal{E}}^n_Q(\cdot,\cdot)$ is defined in (2.2). In other word, $\mathcal{H}_{n,Q}(u) \in l(\hat{V}_n)$ is a Q-harmonic extension of $u \in l(\hat{V}_0)$. By definition, $A_{j,Q}(u) = \mathcal{H}_{1,Q}(u) \circ \Psi_j$. Thus the following holds for all $m \geq 0$ and $l \geq n \geq 0$.

$$\begin{aligned} \text{(A.1)} \qquad & \mathcal{H}_{m+n,\hat{F}^{l-n}(Q)}(u) \circ \Psi_{i_1,\cdots,i_m,j_1,\cdots,j_n} \\ = \; & A_{j_n,\hat{F}^{l-n}(Q)} \circ \cdots \circ A_{j_1,\hat{F}^{l-1}(Q)}(\mathcal{H}_{m,\hat{F}^l(Q)}(u) \circ \Psi_{i_1,\cdots,i_m}). \end{aligned}$$

We have the following (cf. [15] Proposition 3.3, [16] Lemma 5.8).

Lemma A.2. *For* $X, Y \in \mathcal{Q}_{irr}$, *define* $h_{\pm,n} = h_{\pm}(\hat{F}^n(X), \hat{F}^n(Y))$. *Then. for each* $0 \leq m \leq n$,

$$\text{(A.2)} \qquad h_{+,n} \leq h_{+,m} \leq h_{+,0}, \qquad h_{-,n} \geq h_{-,m} \geq h_{-,0}.$$

$$\text{(A.3)} \qquad \textit{There exists } \lim_{n\to\infty} h_{\pm,n} \in (0, \infty).$$

Further, if $h_{\pm,n} = h_{\pm,0}$, *then for all* $u \in A^{\pm}(\hat{F}^n(X), \hat{F}^n(Y))$, *we have*

$$\begin{aligned} \text{(A.4)} \qquad & \mathcal{H}_{n-m,\hat{F}^m(X)}(u) \circ \Psi_{i_1,\ldots,i_{n-m}} \\ = \; & \mathcal{H}_{n-m,\hat{F}^m(Y)}(u) \circ \Psi_{i_1,\ldots,i_{n-m}} \in A^{\pm}(\hat{F}^m(X), \hat{F}^m(Y)). \end{aligned}$$

Proof. (A.2) is from Proposition A.1 1). (A.3) is a simple consequence of (A.2) and the fact $h_-(X, Y) \leq h_+(X, Y)$. Next, if $h_{\pm,n} = h_{\pm,0}$

and $u \in A^-(\hat{F}^n(X), \hat{F}^n(Y))$, then we have

$$\begin{aligned} h_{-,0}\hat{S}_{\hat{F}^n(Y)}(u) &= \hat{S}_{\hat{F}^n(X)}(u) = \hat{\mathcal{E}}^{n-m}_{\hat{F}^m(X)}(\mathcal{H}_{n-m,\hat{F}^m(X)}(u)) \\ &\geq h_{-,m}\hat{\mathcal{E}}^{n-m}_{\hat{F}^m(Y)}(\mathcal{H}_{n-m,\hat{F}^m(X)}(u)) \\ &\geq h_{-,m}\hat{\mathcal{E}}^{n-m}_{\hat{F}^m(Y)}(\mathcal{H}_{n-m,\hat{F}^m(Y)}(u)) \\ &= h_{-,m}\hat{S}_{\hat{F}^n(Y)}(u) \geq h_{-,0}\hat{S}_{\hat{F}^n(Y)}(u). \end{aligned}$$

Thus all the inequalities above are in fact equalities. By the uniqueness of the harmonic extension, we obtain the $(-)$-version of (A.4). $(+)$-version of (A.4) can be proved similarly. Q.E.D.

We next mention a convergence result on positive matrices (cf. [15] Proposition 3.5).

Lemma A.3. *Let B be a finite set. Suppose $A_1, \cdots, A_n, \cdots, A_\infty$ are positive matrices from $l(B)$ to itself and suppose there exists a subsequence $\{\sigma(n)\}_n$ such that*

$$\text{(A.5)} \qquad \lim_{n\to\infty} (A_{\sigma(n)})_{ij} = (A_\infty)_{ij} \qquad \textit{for all } \ i,j \in B.$$

Then, for each family of non-negative non-zero vectors $\{v_n\}_n \subset l(B)$,

$$\text{(A.6)} \qquad \lim_{n\to\infty} \frac{A_1 \circ \cdots \circ A_n v_n}{\|A_1 \circ \cdots \circ A_n v_n\|}$$

exists and it is a positive vector.

Proof. For $p, r \geq 0$, define $T_{p,r} = A_{p+1} \circ \cdots \circ A_{p+r}$. By (A.5) and Theorem 3.6 in [17], $\lim_{r\to\infty}(T_{p,r})_{ij}/(\sum_{s\in B}(T_{p,r})_{is})$ exists for each $i,j \in B$, $p \geq 0$ and it is independent of i and p (in [17], such a property is called strongly ergodic). Using Lemma 3.3 in [17], we obtain the result. Q.E.D.

We now give a key lemma (cf. [15] Lemma 3.6, [16] Section 5.3).

Lemma A.4. *Let $X \in \mathcal{Q}_{irr}$. Then, $\hat{F}(X) = X$ if and only if*

$$\text{(A.7)} \qquad h(\hat{F}^n(X), \hat{F}^{n+1}(X)) = h(X, \hat{F}(X)) \qquad \textit{for all } \ n \in \mathbf{N}.$$

Proof. We will assume (A.7) and prove $\hat{F}(X) = X$ (since the other direction is clear). First, note that there exist $j \in I_F$ and a subsequence $\{s(n)\}_n$ so that we can take $u_{\pm,n} \in A^{\pm}(\hat{F}^{s(n)}(X), \hat{F}^{s(n)+1}(X))$ with

$u_{\pm,n}(a_j) = \min_x u_{\pm,n}(x) = 0$ (a_j is a fixed point of Ψ_j). Then by Lemma A.2,

$$\text{(A.8)} \quad \begin{aligned} u_{\pm,n,m} &:= \mathcal{H}_{s(n)-m,\hat{F}^m(X)}(u_{\pm,n}) \circ \Psi_j^{s(n)-m} \\ &= \mathcal{H}_{s(n)-m,\hat{F}^{m+1}(X)}(u_{\pm,n}) \circ \Psi_j^{s(n)-m} \in A^{\pm}(\hat{F}^m(X), \hat{F}^{m+1}(X)), \end{aligned}$$

for all $m \leq s(n)$ where Ψ_j^n is a n-th iteration of Ψ_j. By (A.1), we have

$$\text{(A.9)} \quad \begin{aligned} u_{\pm,n,m} &= A_{j,\hat{F}^m(X)} \circ \cdots \circ A_{j,\hat{F}^{s(n)-1}(X)}(u_{\pm,n}) \\ &= A_{j,\hat{F}^{m+1}(X)} \circ \cdots \circ A_{j,\hat{F}^{s(n)}(X)}(u_{\pm,n}). \end{aligned}$$

Now choose N_0 large enough so that $\hat{F}^m(X) \in \mathrm{Int}(\mathcal{Q}_M)$ for all $m \geq N_0$ (we use Assumption 2.3 1) here). Using Lemma 3.1, we see that $\{(A_{j,\hat{F}^m(X)})|_B\}_{m \geq N_0}$ are positive matrices from $l(B)$ to itself where $B := \hat{V}_0 \setminus \{a_j\}$. Since all the elements of the matrices are less than 1 (due to Lemma 3.1) and $\sharp B < \infty$, we see that (A.5) holds. We can thus apply Lemma A.3 and obtain that

$$\lim_{n \to \infty} \frac{(u_{\pm,n,m})|_B}{\|(u_{\pm,n,m})|_B\|}$$

exists for $m \geq N_0$. By (A.9), this limit is independent of $m \geq N_0$, we thus denote it as $u_{\pm}$.

We now regard $u_{\pm}$ as a function on $\hat{V}_0$. Then, by the choice of $u_{\pm,n}$, $u_{\pm}(a_j) = 0$ so that $u_{\pm}$ is non-constant. By (A.8), we see that $u_{\pm} \in A^{\pm}(\hat{F}^m(X), \hat{F}^{m+1}(X))$ for all $m \geq N_0$. Thus, $\hat{S}_{\hat{F}^m(X)}(u_{\pm}) = h_{\pm}(\hat{F}^m(X), \hat{F}^{m+1}(X))\hat{S}_{\hat{F}^{m+1}(X)}(u_{\pm})$. On the other hand, by (A.7), $h_{+}(\hat{F}^m(X), \hat{F}^{m+1}(X))/h_{-}(\hat{F}^m(X), \hat{F}^{m+1}(X)) \geq 1$ is independent of $m \geq 0$ which we denote by β. Then, we obtain

$$\frac{\hat{S}_{\hat{F}^{m+1}(X)}(u_+)}{\hat{S}_{\hat{F}^{m+1}(X)}(u_-)} = \beta^{-1} \frac{\hat{S}_{\hat{F}^m(X)}(u_+)}{\hat{S}_{\hat{F}^m(X)}(u_-)} = \cdots = \beta^{-(m+1-N_0)} \frac{\hat{S}_{\hat{F}^{N_0}(X)}(u_+)}{\hat{S}_{\hat{F}^{N_0}(X)}(u_-)},$$

for all $m \geq N_0$. If $\beta > 1$, it contradicts to Proposition A.1 2). So, $\beta = 1$ which means $h(X, \hat{F}(X)) = 0$ (by taking $m = 0$). Thus, $\hat{F}(X) = cX$ for some $c > 0$. Using Remark 2.4 1), we have $c = 1$. Q.E.D.

Proof of Proposition 4.1. First, since $c_1 Q_0 \leq \hat{F}^n(M) \leq c_2 Q_0$ for all $n \in \mathbf{N}$, there exists a subsequence (which could depend on M) $\{\sigma(n)\}_n$ and $Q_M \in \mathcal{Q}_{irr}$ such that

$$\text{(A.10)} \qquad \lim_{n \to \infty} \hat{F}^{\sigma(n)}(M) = Q_M.$$

On the other hand, by Lemma A.2, the following limit exists,

$$h_\pm := \lim_{n\to\infty} h_\pm(\hat{F}^n(M), \hat{F}^{n+1}(M)) \in (0,\infty).$$

Thus,

$$\begin{aligned} & h(\hat{F}^m(Q_M), \hat{F}^{m+1}(Q_M)) \\ = & \lim_{n\to\infty} h(\hat{F}^{m+\sigma(n)}(M), \hat{F}^{m+1+\sigma(n)}(M)) = \log \frac{h_+}{h_-}, \end{aligned}$$

for all $m \in \mathbf{N}$. By Lemma A.4, this implies $\hat{F}(Q_M) = Q_M$. In particular, $Q_M \in \mathrm{Int}(\mathcal{Q}_M)$ due to Assumption 2.3 1). Using Lemma A.2 again, $\lim_{n\to\infty} h_\pm(Q_M, \hat{F}^n(M))$ exists and the limit is 1 due to (A.10). By Proposition A.1 2), this implies (4.1). Q.E.D.

ACKNOWLEDGEMENTS. The author is grateful to Dr. V. Metz and Dr. R. Peirone for giving him information of the paper [15].

References

[1] M.T. Barlow, *Diffusions on fractals*, Lectures in Probability Theory and Statistics (Lect. Notes Math. **1690**), Springer, 1998.

[2] M.T. Barlow and R.F. Bass, *Stability of parabolic Harnack inequalities,* Preprint 2002.

[3] A. Grigor'yan and A. Telcs, *Sub-Gaussian estimates of heat kernels on infinite graphs,* Duke Math. J., **109** (2001), 451–510.

[4] A. Grigor'yan and A. Telcs, *Harnack inequalities and sub-Gaussian estimates for random walks,* Math. Annalen, **324** (2002), 521–556.

[5] B.M. Hambly and T. Kumagai, *Heat kernel estimates for symmetric random walks on a class of fractal graphs and stability under rough isometries,* Proc. of Symposia in Pure Math. (AMS), to appear.

[6] B.M. Hambly and T. Kumagai, *Transition density estimates for diffusions on p.c.f. fractals,* Proc. London Math. Soc., **78** (1999), 431–458.

[7] B.M. Hambly and V. Metz, *The homogenization problem for the Vicsek set,* Stochastic Process. Appl., **76** (1998), 167–190.

[8] V.V. Jikov, S.M. Kozlov and O.A. Oleĭnik, *Homogenization of differential operators and integral functionals,* Springer, 1994.

[9] J. Kigami, *Analysis on Fractals,* Cambridge Univ. Press, 2001.

[10] T. Kumagai and S. Kusuoka, *Homogenization on nested fractals,* Probab. Theory Relat. Fields, **104** (1996), 375–398.

[11] S. Kusuoka, *Diffusion processes on nested fractals,* Statistical Mechanics and Fractals (Lect. Notes Math. **1567**), Springer, 1993.

[12] T. Lindstrøm, *Brownian motion on nested fractals,* Mem. Amer. Math. Soc. 420 **83**, 1990.

[13] V. Metz, *Renormalization contracts on nested fractals,* J. Reine Angew. Math., **480** (1996), 161–175.
[14] R.D. Nussbaum, *Hilbert's projective metric and iterated nonlinear maps,* Mem. Amer. Math. Soc. 391 **75**, 1988.
[15] R. Peirone, *Convergence and uniqueness problems for Dirichlet forms on fractals,* Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8), **3** (2000), 431–460.
[16] C. Sabot, *Existence and uniqueness of diffusions on finitely ramified self-similar fractals,* Ann. Sci. École Norm. Sup. (4), **30** (1997), 605–673.
[17] E. Seneta, *Non-negative matrices and Markov chains,* Springer, 1981.

Research Institute for Mathematical Sciences
Kyoto University, Kyoto 606-8502, Japan
Email: kumagai@kurims.kyoto-u.ac.jp

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 209–232

Representation of Martingales with Jumps and Applications to Mathematical Finance

Hiroshi Kunita

Dedicated to Professor Kiyosi Itô on his 88-th birthday

Abstract.

We study representations of martingales with jumps based on the filtration generated by a Lévy process. Two types of representation theorem are obtained. The first formula is valid for any martingale and written as the sum of the stochastic integral based on the Brownian motion and that based on the compensated Poisson random measure. See (0.1). The second formula is valid only for a process which is a martingale for any equivalent martingale measure. See (0.2). The latter representation formula is then applied to a problem in mathematical finance. The upper hedging strategy and the lower hedging strategy of a contingent claim is obtained through the representation kernel.

§0. Introduction

It is a well known fact that any martingale with respect to the filtration generated by a Brownian motion can be represented as Itô's stochastic integral based on the Brownian motion. On the other hand, martingales with respect to the filtration generated by a Lévy process are not always represented by Itô's stochastic integrals based on the Lévy process, even the latter is a martingale. What is known is that any square integrable martingale with respect to the filtration generated by a Lévy process is represented by stochastic integrals based on the Brownian motion and the compensated Poisson random measure.

In the first half of this paper, we recall these representation theorems following Kunita-Watanabe [6] (Section 1). Let $(\mathcal{F}_t)$ be the filtration generated by a m-dimensional Lévy process. Then every (local)martingale $M(t)$ with respect to the filtration $(\mathcal{F}_t)$ is represented

Received April 7, 2003.

by

$$(0.1)\quad M(t) = M(0) + \sum_{i=1}^{m} \int_0^t \phi_i(s)dW^i(s) + \int_0^t \int_{\mathbf{R}^m} \psi(s,z)\tilde{N}(dsdz),$$

where $W(t) = (W^1(t), ..., W^m(t))$ is a standard Brownian motion and $\tilde{N}(dsdz)$ is the compensated Poisson random measure, which appear in the Lévy-Itô decomposition of the Lévy process. The pair $((\phi_1(s), ..., \phi_m(s)), \psi(s,z))$ is a predictable process with parameter z satisfying certain integrability conditions (See Theorem 1.1). We are particularly interested in the exponential representation of positive martingales (Theorem 2.1). We apply it to the study of Radon Nikodym density of equivalent probability measure and extend Girsanov's theorem to jump processes (Theorem 2.3).

In the second half of the paper, we apply these representation theorems to some problems in mathematical finance. Suppose that we are given a stochastic process $\xi_t = (\xi_t^1, ..., \xi_t^d)$ (e.g., a price process or its return process in mathematical finance) governed by a Lévy process. If the process ξ_t has jumps, there are infinitely many equivalent probability measures with respect to which ξ_t is a localmartingale (called equivalent martingale measures). Now suppose that $M(t)$ is a localmartingale for any equivalent martingale measure. We will show that under some conditions for ξ_t, $M(t)$ is represented by a stochastic integral based on ξ_t, i.e., it is written as

$$(0.2)\qquad M(t) = M(0) + \sum_{i=1}^{d} \int_0^t \varphi_i(s)d\xi_s^i.$$

The difference of these two representations (0.1) and (0.2) are big. We show further that an adapted process $X(t)$ is a supermartingale for any equivalent martingale measure if and only if it admits the unique Doob-Meyer decomposition (not depending on each equivalent martingale measure) and the localmartingle part $M(t)$ is represented as (0.2). See Theorem 3.4 in Section 3.

At the end of Section 3, we apply the above representation theorem to determine the upper hedging price and the lower hedging price of a given contingent claim (Theorem 3.5).

Finally, we mention that there are several works on determing the upper or the lower hedging prices of contingent claims in the case where the price processes have jumps. See e.g. Kabanov-Stricker [3] and references therein. In these works, more general price processes are studied in an abstract manner.

§1. Representation of localmartingales

Let T be a positive number and let $Z(t), t \in [0,T]$ be an m-dimensional Lévy process such that $Z(0) = 0$. Then it admits the Lévy-Itô decomposition:

$$(1.1) \qquad Z(t) = \sigma W(t) + bt + \int_{(0,t]} \int_{|z|>1} zN(dsdz) + \int_{(0,t]} \int_{|z|\le 1} z\{N(dsdz) - \hat{N}(dsdz)\},$$

where σ is an $m \times m$ matrix, $W(t) = (W^1(t), ..., W^m(t))$ is an m-dimensional standard Brownian motion and $N(dsdz)$ is a Poisson counting measure on $[0,T] \times \mathbf{R}^m$ with intensity measure $\hat{N}(dsdz) = ds\nu(dz)$, which is independent of $W(t)$. In the following, we denote

$$(1.2) \qquad \tilde{N}(dsdz) = N(dsdz) - \hat{N}(dsdz).$$

Let $(\mathcal{F}_t), t \in [0,T]$ be the filtration generated by the Brownian motion $W(t)$ and the Poisson random measure $N(dtdz)$. Then both $W(t)$ and $\int_{|z|\le 1} z\tilde{N}(dsdz)$ are martingales adapted to the filtration. Let $M(t), t \in [0,T]$ be an $(\mathcal{F}_t)$-adapted cadlag (right continuous with the left hand limits) process. It is called a localmartingale if there exists a non-decreasing sequence of stopping times $\tau_n, n = 1, 2, ...$ with values in $[0,T]$ such that $P(\tau_n < T) \to 0$ as $n \to \infty$ and the stopped process $M(t \wedge \tau_n)$ is a martingale for any n. In particular if we can choose the sequence such that the stopped process $M(t \wedge \tau_n)$ is a square integrable martingale for any n, $M(t)$ is called a locally square integrable martingale. Any continuous localmartingale is a locally square integrable martingale, but it is not always the case for a localmartingale with jumps. An $(\mathcal{F}_t)$ adapted cadlag process $X(t)$ is called a semimartingale if it is written as a sum of a localmartingale and a process of bounded variation. In particular if the corresponding process of bounded variation is locally integrable, $X(t)$ is called a special semimartingale. A special semimartingale is decomposed uniquely to the sum of a localmartingale and a predictable process of bounded variation.

We denote by Φ the set of all m dimensional predictable processes $\phi(t) = (\phi_1(t), ..., \phi_m(t))$ such that $\int_0^T |\phi(s)|^2 dt < \infty$ a.s. Then the stochastic integral based on the m-dimensional Brownian motion $W(t) = (W^1(t), ..., W^m(t))$ is well defined for $\phi \in \Phi$. We use the notation:

$$(1.3) \qquad \int_0^t (\phi(s), dW(s)) = \sum_{i=1}^m \int_0^t \phi_i(s) dW^i(s)$$

It is a continuous locally square integrable martingale.

Let $\mathcal{P}$ be the predictable σ-algebra on $[0,T]\times\Omega$ and let $\mathcal{B}$ be the Borel algebra on $\mathbf{R}^m$. A functional $\psi(s,z,\omega),(s,z,\omega)\in[0,T]\times\mathbf{R}^m\times\Omega$ is called a predictable process if it is $\mathcal{P}\times\mathcal{B}$-measurable.

We will recall the definition of the stochastic integral of the predictable process $\psi(s,z)$ based on the compensated Poisson random measure $\tilde{N}(dsdz)$ following Kunita-Watanabe [6]. Note first that if $E_1,...,E_n$ are disjoint Borel subsets of $[0,T]\times\mathbf{R}^m$ such that $\hat{N}(E_1)<\infty,...,\hat{N}(E_n)<\infty$, then $\tilde{N}(E_1),...,\tilde{N}(E_n)$ are independent random variables with mean 0 and variance $\hat{N}(E_1),...,\hat{N}(E_n)$, respectively. Now, let $\psi(t,z)$ be a step process of the form $\sum_{i,j}a_{ij}1_{(t_i,t_{i+1}]}(t)1_{F_{ij}}(z)$, where $0=t_0<\cdots<t_N=T$ and for each i $F_{i1},...,F_{in}$ are disjoint subsets of $\mathbf{R}^m$ satisfying $\nu(F_{ij})<\infty, j=1,...,n$ and a_{ij} are bounded $\mathcal{F}_{t_i}$-adapted random variables. We define the stochastic integral of ψ based on $\tilde{N}$ by

$$\int_0^T\int_{\mathbf{R}^m}\psi(t,z)\tilde{N}(dtdz)=\sum_{i,j}a_{ij}\tilde{N}((t_i,t_{i+1}]\times F_{ij}). \tag{1.4}$$

Then we have

(1.5)

$$E\left[\left(\int_0^T\int_{\mathbf{R}^m}\psi(t,z)\tilde{N}(dtdz)\right)^2\right]=\sum_i E\left[(\sum_j a_{ij}\hat{N}((t_i,t_{i+1}]\times F_{ij}))^2\right]$$
$$+\sum_{i<k}E\left[\{\sum_j a_{kj}\hat{N}((t_k,t_{k+1}]\times F_{kj})\}\{\sum_j a_{ij}\hat{N}((t_i,t_{i+1}]\times F_{ij})\}\right].$$

Since $\{a_{ij},j=1,2,...\}$ and $\tilde{N}((t_i,t_{i+1}]\times F_{ij}), j=1,2,...$ are independent and the latters are of mean 0, the first term of the left hand side is computed as

$$\sum_{ij}E\left[a_{ij}^2\tilde{N}((t_i,t_{i+1}]\times F_{ij})^2\right]=\sum_{ij}E\left[a_{ij}^2\hat{N}((t_i,t_{i+1}]\times F_{ij})\right] \tag{1.6}$$

The last term of (1.5) is 0, since $E[\sum_j a_{kj}\tilde{N}((t_j,t_{j+1}]\times F_{kj})|\mathcal{F}_{t_{i+1}}]=0$. Therefore we have

(1.7)

$$E\left[\left(\int_0^T\int_{\mathbf{R}^m}\psi(t,z)\tilde{N}(dtdz)\right)^2\right]=E\left[\int_0^T\int_{\mathbf{R}^m}|\psi(t,z)|^2\hat{N}(dtdz)\right].$$

Now suppose that $\psi(t,z)$ is a predictable process satisfying the condition $E[\int_0^T\int_{\mathbf{R}^m}|\psi(t,z)|^2 dt\nu(dz)] < \infty$. Then we can choose a sequence $\{\psi_n(t,z)\}$ of step processes such that $E[\int_0^T\int_{\mathbf{R}^m}|\psi(t,z)-\psi_n(t,z)|^2 dt\nu(dz)] \to 0$, as $n \to \infty$. Denote the stochastic integral of ψ_n (formula (1.4)) by M_n. Then M_n converges in L^2. We denote the limit by $M = \int_0^T\int_{\mathbf{R}^m}\psi(t,z)\tilde{N}(dtdz)$. Then it satisfies (1.7) again.

The stochastic integral $\int_0^T 1_{(0,t]}(s)\psi(s,z)\tilde{N}(dsdz)$ is denoted by $\int_0^t \psi(s,z)\tilde{N}(dsdz)$. It is a cadlag process with time t and in fact is a square integrable martingale. This fact can be shown directly in the case where $\psi(t,x)$ is a step process defined above. Then the martingale property is extended to any ψ such that (1.7) is finite.

We denote by $\Psi_2(\hat{N})$ $(\Psi_1(\hat{N}))$ the set of all predictable processes $\psi(t,z)$ which are square integrable (resp. integrable) with respect to the measure $\hat{N}(dtdz)$ a.s. Then we can define the stochastic integral $\int_0^t\int_{\mathbf{R}^m}\psi(s,z)\tilde{N}(dsdz)$ for $\psi \in \Psi_2(\hat{N})$ as a locally square integrable martingale. For $\psi \in \Psi_1(\hat{N})$, we define the stochastic integral by

$$(1.8)\quad \int_0^t\int_{\mathbf{R}^m}\psi(s,z)\tilde{N}(dsdz) := \int_0^t\int_{\mathbf{R}^m}\psi(s,z)N(dsdz) - \int_0^t\int_{\mathbf{R}^m}\psi(s,z)\hat{N}(dsdz).$$

It is a localmartingale. For a general $\psi(t,z)$, we set $\psi_1(t,z) = \psi(t,z)1_{\{|\psi|>1\}}(t,z), \psi_2(t,z) = \psi(t,z)1_{\{\{|\psi|\le 1\}}(t,z)$, and we denote by $\Psi_{1,2}(\hat{N})$ the set of all predictable process $\psi(t,z)$ such that $\psi_1 \in \Psi_1(\hat{N})$ and $\psi_2 \in \Psi_2(\hat{N})$. Then, for any $\psi \in \Psi_{1,2}(\hat{N})$, the stochastic integral is defined as the sum of stochastic integrals of ψ_1 and ψ_2. It is a localmartingale.

The following notations will be used

(1.9)

$$N_t(\psi) = \int_0^t\int_{\mathbf{R}^m}\psi(s,z)N(dsdz), \quad \tilde{N}_t(\psi) = \int_0^t\int_{\mathbf{R}^m}\psi(s,z)\tilde{N}(dsdz),$$

$$\hat{N}_t(\psi) = \int_0^t\int_{\mathbf{R}^m}\psi(s,z)\hat{N}(dsdz),$$

if these are well defined.

Now, we give a representation theorem of localmartingales.

Theorem 1.1. *([6], Example at p.227 and Proposition 5.2) Let $M(t)$ be a localmartingale. Then there exist $\phi(s) \in \Phi$, $\psi(s,z) \in \Psi_{1,2}(\hat{N})$,* *and $M(t)$ is represented by*

$$(1.10) \quad M(t) = M(0) + \int_0^t (\phi(s), dW(s)) + \int_0^t \int_{\mathbf{R}^m} \psi(s,z)\tilde{N}(dsdz).$$

The representation kernel $(\phi(s), \psi(s,z))$ is uniquely determined from $M(t)$, i.e., if $M(t)$ is represented by (1.10) with another $(\phi'(s), \psi'(s,z))$, then we have $\phi(s) = \phi'(s)$ a.e. $\lambda \otimes P$ and $\psi(s,z) = \psi'(s,z)$ a.e. $\hat{N} \otimes P$, where λ is the Lebesgue measure on $[0,T]$.

Proof. In the paper [6], the above theorem is proved for square integrable martingale by using the theory of additive functionals of Markov processes. Here we give a direct and simpler proof by applying Itô [2]. For simplicity we prove the theorem in the case $m = 1$ only. Let $\mathbf{Z} = (Z(t))$ be a one dimensional Lévy process and let (1.1) be the Lévy-Itô decomposition. We introduce a random measure $M(E)$ on $[0,T] \times \mathbf{R}$ by

$$(1.11) \qquad M(E) = \int_{E(0)} dW(t) + \int_{E-E(0)} \frac{z}{1+|z|}\tilde{N}(dtdz),$$

where $E(0) = \{(t,0);\ (t,0) \in E\}$. Then we have $E[M(E_1)M(E_2)] = \mu(E_1 \cap E_2)$, where

$$\mu(E) = |E(0)| + \int_{E-E(0)} \left(\frac{z}{1+|z|}\right)^2 dt\nu(dz).$$

For each positive integer p, we define the multiple Wiener integral by

$$(1.12) \qquad I_p(f) = \int \cdots \int f(\xi_1, ..., \xi_p) dM(\xi_1) \cdots dM(\xi_p).$$

Let $\mathbf{H_Z}$ be the L^2 space over $(\Omega, \mathcal{F}_T, P)$ and let $\mathbf{H_Z}^{(p)}$ be the closed linear manifold of $\{I_p(f); f \in L_p^2\}$, where L_p^2 is the L^2 space on $\mathbf{R}^p$ with the product measure of μ. Then it is shown in [2] that one has the direct sum expansion: $\mathbf{H_Z} = \sum_{p \geq 0} \oplus \mathbf{H_Z}^{(p)}$. Note that each $I_p(f)$ is written as the sum of the following terms

(1.13)

$$\int \cdot \int_{0 \leq t_1 < \cdots < t_p \leq T, (z_1, \ldots, z_p) \in \mathbf{R}^p} f((t_1, z_1), ..., (t_p, z_p)) dM(t_1 z_1) \cdots dM(t_p z_p)$$

$$= \int_0^T \int_{\mathbf{R}} \varphi(t_p, z_p) dM(t_p z_p),$$

where

$$\varphi(t_p, z_p) = \int \cdot \int_{\Lambda(t_p, z_p)} f((t_1, z_1), ..., (t_p, z_p)) dM(t_1 z_1) \cdots dM(t_{p-1} z_{p-1})$$

and $\Lambda(t_p, z_p) = \{0 < t_1 < \cdots < t_{p-1} < t_p, (z_1, ..., z_{p-1}, z_p) \in \mathbf{R}^{p-1}\}$. Setting $\phi(t) = \varphi(t, 0)$ and $\psi(t, z) = \varphi(t, z)\frac{1+|z|}{z}$ ($|z| > 0$), we find that the above is written as

$$(1.14) \qquad \int_0^T \phi(s) dW(s) + \int_0^T \int_{\mathbf{R}} \psi(s, z) \tilde{N}(dsdz).$$

Therefore any element of ${\mathbf{H_Z}}^{(p)}$ and hence any element X of $\mathbf{H_Z}$ with mean 0 is written as the above. Now taking the conditional expectation of (1.14), we obtain the representation (1.10) for square integrable martingale $M(t) = E[X|\mathcal{F}_t]$.

The extension to locally square integrable martingales will be obvious. The extension to localmartingales will be discussed after Theorem 2.1 in the next section.

Let $M(t)$ and $N(t)$ be two locally square integrable martingales such that $M(0) = N(0) = 0$. Then by the Doob-Meyer decomposition of a supermartingale, there exist adapted continuous increasing processes $\langle M \rangle_t$, $\langle N \rangle_t$ and an adapted continuous process of bounded variations $\langle M, N \rangle_t$ such that $\langle M \rangle_0 = \langle N \rangle_0 = \langle M, N \rangle_0 = 0$ and $M(t)^2 - \langle M \rangle_t$, $N(t)^2 - \langle N \rangle_t$ and $M(t)N(t) - \langle M, N \rangle_t$ are localmartingales. Such bracket processes are uniquely determined. Note that $\langle M, M \rangle_t = \langle M \rangle_t$ by the definition. If $M(t)$ is represented by (1.10) with $M(0) = 0$ and $N(t)$ is represented with the kernel $(\tilde{\phi}, \tilde{\psi})$, then we have the formula

$$(1.15) \quad \langle M, N \rangle_t = \int_0^t (\phi(s), \tilde{\phi}(s)) ds + \int_0^t \int_{\mathbf{R}^m} \psi(s, z) \tilde{\psi}(s, z) \hat{N}(dsdz).$$

We can define the quadratic co-variation of two semimartingales $X(t)$ and $Y(t)$ by

$$(1.16) \qquad [X, Y]_t = \exists \lim_{|\Delta| \to 0} \sum_{k=1}^{n} (X(t_k) - X(t_{k-1}))(Y(t_k) - Y(t_{k-1})),$$

where Δ are partitions of the time interval $[0, t]$ such that $0 = t_0 < t_1 < \cdots t_n = t$ and $|\Delta| = \max_{1 \le k \le n} |t_k - t_{k-1}|$. We set $[X]_t = [X, X]_t$. If $M(t)$ and $N(t)$ are continuous localmartingales, it is known that the bracket process $\langle M, N \rangle_t$ and the quadratic co-variation coincides, i.e., $[M, N]_t = \langle M, N \rangle_t$. However if both $M(t), N(t)$ have jumps, the bracket process is not equal to the quadratic variation. In the case where the representation kernels of $M(t)$ and $N(t)$ are (ϕ, ψ) and $(\tilde{\phi}, \tilde{\psi})$, respectively, we have

$$(1.17) \quad [M,N]_t = \int_0^t (\phi(s), \tilde{\phi}(s))ds + \int_0^t \int_{\mathbf{R}^m} \psi(s,z)\tilde{\psi}(s,z)N(dsdz).$$

Two locally square integrable martingales $M(t)$ and $N(t)$ are called *orthogonal* if $M(t)N(t)$ is a localmartingale or equivalently, the bracket process $\langle M, N\rangle_t$ is identically 0. By the formula (1.15) we see that the continuous local martingale $\sum_{i=1}^m \int_0^t \phi_i(s)dW^i(s)$ and the discontinuous one $\int_0^t \int_{\mathbf{R}^m} \psi(s,z)\tilde{N}(dsdz)$ are orthogonal.

Suppose that $M(t)$ and $N(t)$ are not orthogonal. There exists a predictable process $\varphi(t)$ such that $\langle M, N\rangle_t = \int_0^t \varphi(s)d\langle N\rangle_s$. We define new locally square integrable martingales by $M^1(t) = \int_0^t \varphi(s)dN(s)$ and $M^2(t) = M(t) - M^1(t)$. Then $M^1(t)$ and $M^2(t)$ are orthogonal each other because of the equality

$$\langle M^1, M^2\rangle_t = \langle M^1, M\rangle_t - \langle M^1, M^1\rangle_t = \int_0^t \varphi(s)^2 d\langle N\rangle_s - \int_0^t \varphi(s)^2 d\langle N\rangle_s = 0.$$

The locally square integrable martingale $M^1(t)$ is called the *orthogonal projection* of $M(t)$ to $N(t)$. In the case where $M(t)$ and $N(t)$ are represented with kernels (ϕ, ψ) and $(\tilde{\phi}, \tilde{\psi})$, respectively, the kernel φ of the orthogonal projection is given by

$$(1.18) \qquad \varphi(t) = \frac{(\phi(t), \tilde{\phi}(t)) + \int_{\mathbf{R}^m} \psi(t,z)\tilde{\psi}(t,z)\nu(dz)}{|\tilde{\phi}(t)|^2 + \int_{\mathbf{R}^m} \tilde{\psi}(t,z)^2\nu(dz)}.$$

We denote by $\mathcal{M}_{loc}^2$ (resp. $\mathcal{M}_{loc}^1$) the set of all locally square integrable martingales (resp. localmartingales) $M(t)$ with $M(0) = 0$. It is a vector space. A sequence $\{M_k(t), k = 1, 2, ..\}$ of $\mathcal{M}_{loc}^2$ is said to converge to $M(t)$ if there exists a nondecreasing sequence of stopping times $\tau_n, n = 1, ..$ such that $P(\tau_n < T) \to 0$ as $n \to \infty$ and each sequence of stopped processes $\{M_k^{\tau_n}(t) := M_k(t \wedge \tau_n), k = 1, 2, ...\}$ converges to $M^{\tau_n}(t)$ in L^2. Then $\mathcal{M}_{loc}^2$ is a complete space by this topology.

For a given $M(t) \in \mathcal{M}_{loc}^2$, we set

$$(1.19) \qquad \mathcal{L}(M) = \left\{ \int_0^t \varphi(s)dM(s);\ \varphi(s) \in \Phi(\langle M\rangle) \right\},$$

where $\Phi(\langle M\rangle)$ is the set of all predictable processes φ such that $\int_0^T |\varphi(t)|^2 d\langle M\rangle_t < \infty$ a.s. It is a subset of $\mathcal{M}^2_{loc}$. Let $\mathcal{N}$ be a subset of $\mathcal{M}^2_{loc}$. It is called a subspace of $\mathcal{M}^2_{loc}$ if it is a closed vector space including $\mathcal{L}(N)$ whenever $N \in \mathcal{N}$.

Given a subset $\mathcal{N}$ of $\mathcal{M}^2_{loc}$, we denote by $\mathcal{L}(N)$ the smallest closed subspace containing the set $\mathcal{N}$. We denote by $\mathcal{N}^\perp$ the set of all $M(t) \in \mathcal{M}^2_{loc}$ which is orthogonal to any $N \in \mathcal{N}$. Then $\mathcal{N}^\perp$ is a closed subspace of $\mathcal{M}^2_{loc}$. Further, if $\mathcal{N}$ is a closed subspace of $\mathcal{M}^2_{loc}$, every $M(t) \in \mathcal{M}^2_{loc}$ is decomposed uniquely to the sum of $M^1(t) \in \mathcal{N}$ and $M^2(t) \in \mathcal{N}^\perp$. We have thus the orthogonal decomposition

$$\mathcal{M}^2_{loc} = \mathcal{N} \oplus \mathcal{N}^\perp. \tag{1.20}$$

§2. Exponential representation of positive martingales and extension of Girsanov's theorem

We shall consider the exponential representation of a positive localmartingale. Here a localmartingale α_t is called positive if $\alpha_t > 0$ holds for all $t \in [0,T]$ a.s. For a predictable process $g(t,z)$, we set $g_1 = g1_{|g|>1}$ and $g_2 = g1_{|g|\le 1}$ as before. $g(s,z)$ is said to belong to $\Psi_{e,2}(\hat{N})$ if $e^{g_1(t,z)} - 1 \in \Psi_1(\hat{N})$ and $g_2 \in \Psi_2(\hat{N})$. Then it holds that $g \in \Psi_{e,2}(\hat{N})$ if and only if $\psi \equiv e^g - 1 \in \Psi_{1,2}(\hat{N})$.

Theorem 2.1. *(c.f. [6], Theorem 6.1) Let α_t be a positive localmartingale such that $\alpha_0 = 1$. Then there exists a pair of predictable process $f(t) = (f_1(t), ..., f_m(t))$ of Φ and $g(s,z)$ of $\Psi_{e,2}$ such that the localmartingale α_t is represented by*

(2.1)
$$\alpha_t = \exp\left\{\left(\int_0^t (f(s), dW(s)) - \frac{1}{2}\int_0^t |f(s)|^2 ds\right) + \left(N_t(g_1) - \hat{N}_t(e^{g_1} - 1)\right) + \left(\tilde{N}_t(g_2) - \hat{N}_t(e^{g_2} - 1 - g_2)\right)\right\}.$$

Further, the pair (f,g) is uniquely determined from α_t.

Conversely let $(f(t), g(t,z))$ be a pair of predictable processes belonging to Φ and $\Psi_{e,2}(\hat{N})$, respectively. Define α_t by (2.1). Then it is a positive localmartingale.

The above α_t is characterized as the solution of the following Itô's stochastic differential equation starting from 1 at time 0:

$$d\alpha_t = \alpha_{t-}(f(t), dW(t)) + \alpha_{t-}\int_{\mathbf{R}^m} \left(e^{g(t,z)} - 1\right)\tilde{N}(dtdz). \tag{2.2}$$

In fact, apply Itô's formula ([6], Theorem 5.1) to the function $F(x) = e^x$ and $X(t) = \log \alpha_t$, where α_t is given by (2.1). Note the obvious formula $e^g - 1 = (e^{g_1} - 1) + (e^{g_2} - 1)$. Then we find that α_t satisfies the above SDE. It is determined by two integrands $f(t)$ and $g(t,z)$. We denote the positive localmartingale by $\alpha_t = \alpha_t(f,g)$.

The above theorem is proved in [6] in the case where α_t is a multiplicative functional of a Markov process. We give here a direct and simpler proof.

Lemma 2.2. *(cf [6], Lemma 6.1.) Let* $f(t)=(f_1(t),...,f_m(t)), g(t,z)$, $h(t,z)$ *be predictable processes such that* $f \in \Phi$, h *is bounded belonging to* $\Psi_2(\hat{N})$, $gh = 0$ *and* $A(t)$ *is a right continuous predictable process of bounded variation. Set*

$$\beta_t = \exp\left\{\int_0^t (f(s), dW(s)) + N_t(g) + \tilde{N}_t(h) - A(t)\right\}. \tag{2.3}$$

Then β_t *is a localmartingale if and only if the following two conditions are satisfied.*

$$e^g - 1 \in \Psi_1(\hat{N}), \tag{2.4}$$

$$A(t) = \frac{1}{2}\int_0^t |f(s)|^2 ds + \hat{N}_t(e^g - 1) + \hat{N}_t(e^h - 1 - h). \tag{2.5}$$

Proof. By Itô's formula, we have

$$\begin{aligned}\beta_t - 1 = \int_0^t \beta_{s-}(f(s), dW(s)) + \int_0^t \int_{\mathbf{R}^m} \beta_{s-}(e^g - 1)dN \\ + \int_0^t \int_{\mathbf{R}^m} \beta_{s-}(e^h - 1)d\tilde{N} \\ + \frac{1}{2}\int_0^t \beta_{s-}|f(s)|^2 ds + \int_0^t \int_{\mathbf{R}^m} \beta_{s-}(e^h - 1 - h)d\hat{N} - \int_0^t \beta_{s-}dA(s).\end{aligned}$$

Therefore if (2.4) and (2.5) are satisfied, then

$$\begin{aligned}\beta_t - 1 = \int_0^t \beta_{s-}(f(s), dW(s)) + \int_0^t \int_{\mathbf{R}^m} \beta_{s-}(e^g - 1)d\tilde{N} \\ + \int_0^t \int_{\mathbf{R}^m} \beta_{s-}(e^h - 1)d\tilde{N}.\end{aligned}$$

Therefore β_t is a localmartingale.

Conversely suppose that β_t is a localmartingale. We want to prove (2.4). Set $g^+ = \max\{g, 0\}$ and $g^- = \max\{-g, 0\}$. Then $g = g^+ - g^-$. We shall prove first $e^{-g^-} - 1 \in \Psi_1(\hat{N})$. It holds by Itô's formula

$$e^{-N_t(g^-)} - 1 = -\int_0^t (1 - e^{-g^-})e^{-N_{s-}(g^-)}dN.$$

Since $-N_t(g^-) \leq 0$, the expectation of the above is finite and is equal to $-E[\int_0^t (1 - e^{-g^-})e^{-N_{s-}(g^-)}\hat{N}(dsdz)]$. Therefore, $(1 - e^{-g^-})e^{-N_{s-}(g^-)} \in \Phi_1(\hat{N})$ and this implies $(1 - e^{-g^-}) \in \Psi_1(\hat{N})$. Next, we have by Itô's formula,

$$\begin{aligned}\int_0^t \frac{d\beta_s}{\beta_{s-}} &= \int_0^t (f(s), dW(s)) + N_t(e^g - 1) + \tilde{N}_t(e^h - 1)\\ &+\frac{1}{2}\int_0^t |f(s)|^2 ds + \hat{N}_t(e^h - 1 - h) - A(t).\end{aligned}$$

The left hand side is a localmartingale. All terms except $N_t(e^g - 1)$ of the right hand side are locally integrable. Further we have $N_t(e^g - 1) = N_t(e^{g^+} - 1) + N_t(e^{-g^-} - 1)$ and the last term is locally integrable. Then $N_t(e^{g+} - 1)$ should be locally integrable, which shows that $\int_0^t (e^{g^+} - 1)d\hat{N}$ is also locally integrable, proving that $e^{g^+} - 1 \in \Psi_1(\hat{N})$. We have thus proved (2.4).

Now since (2.4) holds, the bounded variation part of $\beta_t - 1$ can be written as

$$\frac{1}{2}\int_0^t \beta_{s-}|f(s)|^2 ds + \int_0^t \beta_{s-}(e^g - 1)d\hat{N} + \int_0^t \beta_{s-}(e^h - 1 - h))d\hat{N} - \int_0^t \beta_{s-}dA(s).$$

It should be 0 since β_t is a localmartingale. This implies (2.5).

Proof of Theorem 2.1. Suppose that α_t is a positive localmartingale. Set $X(t) = \log \alpha_t$. It is a semimartingale. Consider

$$P^n(t) = \sum_{s \leq t, 1 \leq |\Delta X(s)| \leq n} \Delta X(s), \quad \text{(finite sum)}.$$

It is a locally integrable process of bounded variation. There exists a continuous process of boundecd variation $C^n(t)$ such that $M^n(t) = P^n(t) - C^n(t)$ is a locally square integrable martingale by Doob-Meyer decomposition. Then there exists $\psi_n \in \Psi_2(\hat{N})$ such that $M^n(t) = \tilde{N}(\psi_n)$ by Theorem 1.1 (for locally square integrable martingales). Jump parts of $P^n(t)$ and $N_t(\psi_n)$ coincide. Therefore we have $P^n(t) = N_t(\psi_n)$. It holds

$\psi_m = \psi_n 1_{|\psi_n|\le m}$ a.e. $\hat{N} \times P$ for any $m < n$. Then there exists ψ such that $\psi_n = \psi 1_{|\psi|\le n}$ and we have

$$N_t(\psi) = \sum_{s\le t, 1\le|\Delta X(s)|<\infty} \Delta X(s), \quad \text{(finite sum)}.$$

Now set $Y(t) = X(t) - N_t(\psi)$. It is a semimartingale such that $|\Delta Y(s)| \le 1$. Therefore it is a special semimartingale. Then it is decomposed uniquely to the sum of a martingale $M(t)$ and a predictable process of bounded variation, denoted by $B(t)$. Further $M(t)$ is locally square integrable so that it is written as $M(t) = \int_0^t (f(s), dW(s)) + \tilde{N}_t(\eta)$, where $f \in \Phi$ and $\eta \in \Psi_2(\hat{N})$. It holds $\psi\eta = 0$ since $N_t(\psi)$ and $Y(t)$ do not have common jumps. Then we get the decomposition:

$$\alpha_t = \exp\left\{\int_0^t (f(s), dW(s)) + N_t(\psi) + \tilde{N}_t(\eta) + B(t)\right\}. \tag{2.6}$$

Since α_t is a localmartingale, we have $e^{\psi} - 1 \in \Psi(\hat{N})$ and

$$-B(t) = \frac{1}{2}\int_0^t |f(s)|^2 ds + \hat{N}_t(e^{\psi} - 1) + \hat{N}_t(e^{\eta} - 1 - \eta)$$

by the previous lemma. Now set $g = \psi + \eta$. Then we have $g_1 = \psi$ and $g_2 = \eta$. Therefore we get the formula (2.1). The proof is complete.

Proof of Theorem 1.1 (continued). Let $M(t)$ be a martingale. We set $M^+ = M(T)\vee 0$ and $M^- = (-M)\vee 0$ and define $M_1(t) = E[M^+|\mathcal{F}(t)]$ and $M_2(t) = E[M^-|\mathcal{F}(t)]$. Then both are nonnegative martingales and $M(t) = M_1(t) - M_2(t)$. We consider positive martingales $M_{i,\epsilon}(t) = M_i(t) + \epsilon$ $(\epsilon > 0)$. These are represented by $M_{i,\epsilon}(t) = M_{i,\epsilon}(0)\alpha_t^i$, where $\alpha_t^i = \alpha_t(f_i', g_i')$ are exponential martingales. These satisfy SDE (2.2). Now set $\phi_i(t) = \alpha_{t-}^i f_i'(t)$ and $\psi_i(t,z) = \alpha_{t-}^i(e^{g_i'(t,z)} - 1)$. Then, since $\sup_t \alpha_t < \infty$ a.s., $\phi_i \in \Phi$ and $\psi_i \in \Psi_{1,2}(\hat{N})$. Further, we get the representation (1.10) for $M_{i,\epsilon}(t), i = 1, 2$. Thus we get the representation (1.10) where $\phi \in \Phi$ and $\psi \in \Psi_{1,2}(\hat{N})$.

Let α_t be a positive martingale with mean 1. We can define a probability measure Q by the formula

$$Q(A) = \int_A \alpha_T dP, \quad A \in \mathcal{F}. \tag{2.7}$$

Then $((\mathcal{F}_t), Q)$ and $((\mathcal{F}_t), P)$ are equivalent (mutually absolutely continuous). Conversely let $((\mathcal{F}_t), Q)$ be a probability measure equivalent to $((\mathcal{F}_t), P)$. Let α_t be the Radon-Nikodym density of $(\mathcal{F}_t, Q)$ with respect to $(\mathcal{F}_t, P)$. Then the stochastic process $\{\alpha_t, t \in [0,T]\}$ is a positive martingale with respect to P. Therefore it can be represented as $\alpha_t = \alpha_t(f,g)$.

A localmartingale with respect to $((\mathcal{F}_t), Q)$ is called a Q-localmartingale. The following is an extension of Girsanov's theorem.

Theorem 2.3. *(c.f. [6], Theorem 6.2) With respect to $((\mathcal{F}_t), Q)$, we have*
1) $W^f(t) := W(t) - \int_0^t f(s)ds$ is a standard Brownian motion.
2) The compensator of N is $\hat{N}^g(dsdz) = e^{g(s,z)}ds\nu(dz)$, that is $\tilde{N}^g(dsdz) := N(dsdz) - \hat{N}^g(dsdz)$ is a martingale measure. Further if ψ belongs to $\Psi_{1,2}(\hat{N}^g)$, the stochastic integral

$$\tilde{N}_t^g(\psi) := \int_0^t \int_{\mathbf{R}^m} \psi(s,z)\tilde{N}^g(dsdz) \tag{2.8}$$

is well defined as a Q-localmartingale.

3) Let $X(t)$ be a Q-localmartingale. Then there exists a pair of predictable processes $(\phi(t), \psi(t,z))$ belonging to Φ and $\Psi_{1,2}(\hat{N}^g)$, respectively and $X(t)$ is represented by

$$X(t) = X(0) + \int_0^t (\phi(s), dW^f(s)) + \tilde{N}_t^g(\psi). \tag{2.9}$$

Remark. 1) $N(dtdz)$ is no longer a Poisson random measure with respect to Q unless g is a deterministic function.
2) Set $\mathcal{F}_t(f,g) = \sigma(W_s^f, \tilde{N}^g(dsdz); s \leq t)$. Then it holds $\mathcal{F}_t(f,g) \subset \mathcal{F}_t$. The equality does not hold in general. The representation (2.9) is valid for localmartingale with respect to the filtration $(\mathcal{F}_t)$ but it is not clear if we have the similar representations for localmartingales with respect to the filtration $(\mathcal{F}_t(f,g))$.

Proof. The assertions (1) and (2) are shown in [6] in the case where α_t is a multiplicative functional of a Markov process. We give here an alternative proof. We first show that $W^f(t)$ is a Q-localmartingale. Set $X(t) = W^f(t)$. Then $X(t)$ is a Q-localmartingale if and only if the product $X(t)\alpha_t$ is a P-localmartingale. Note the equality

$$X(t)\alpha_t = \int_0^t X(s-)d\alpha_s + \int_0^t \alpha_{s-}dX(s) + [X,\alpha]_t. \tag{2.10}$$

The first term of the right hand side is a P-localmartingale. Since $[X,\alpha]_t = \langle X,\alpha\rangle_t = \int_0^t \alpha_{s-}f(s)ds$, we have $\int_0^t \alpha_{s-}dX(s) + [X,\alpha_t] = \int_0^t \alpha_{s-}dW(s)$, which is also a P-localmartingale. Therefore $X(t)\alpha_t$ is a

P-localmartingale or equivalently $X(t)$ is a continuous Q-localmartingale. It holds $[X]_t = [W]_t = t$, since the quadratic variation of $\int_0^t f(s)ds$ is 0. Hence $X(t) = W^f(t)$ is a Brownian motion with respect to Q.

We will next prove (2). Suppose first that $\psi(t,z)$ is bounded and $\int_0^T |\psi| d\hat{N} < \infty$ is satisfied. Then it holds valid $\int_0^T |\psi| e^g d\hat{N} < \infty$, since $g \in \Psi_{e,2}(\hat{N})$. Then $X(t) := \tilde{N}_t^g(\psi)$ is decomposed as $X(t) = \tilde{N}_t(\psi) - \int_0^t \int \psi(e^g - 1)d\hat{N}$. It holds (2.10) again. We have

$$\int_0^t \alpha_{s-} dX(s) + [X,\alpha]_t = \int_0^t \alpha_{s-}\psi d\tilde{N}$$
$$- \int_0^t \int_{\mathbf{R}^m} \alpha_{s-}\psi(e^g - 1)d\hat{N} + \int_0^t \int_{\mathbf{R}^m} \alpha_{s-}\psi(e^g - 1)dN,$$

which is a P-localmartingale. Consequently $X(t)\alpha_t$ is again a P-localmartingale, proving that $X(t) = \tilde{N}_t^g(\psi)$ is a Q-localmartingale. It can be extended to any $\psi \in \Psi_{1,2}(\hat{N}^g)$.

We will prove (3). Suppose first that $X(t)$ is a Q-localmartingale such that its jumps are bounded. Then $M(t) := X(t)\alpha_t$ is a P-localmartingale. Since α_t^{-1} is a P-semimartingale, the product $X(t) = M(t)\alpha_t^{-1}$ is a P-semimartingale. Note that jumps of $X(t)$ are bounded. Then $X(t)$ is a P special semimartingale. Then it is decomposed uniquely as $X(t) - X(0) = N(t) + A(t)$, where $N(t)$ is a P locally square integrable martingale and $A(t)$ is a right continuous predictable process of bounded variation. Now, $N(t)$ is represented by $\int \phi dW + \int \psi d\tilde{N}$, where ψ is a bounded predictable process. Then we can rewrite $X(t)$ as

$$\text{(2.11)}\quad X(t) = X(0) + \int_0^t (\phi(s), dW^f(s)) + N_t^g(\psi)$$
$$+ \left\{ \int_0^t (\phi(s), f(s))ds + \int_0^t \int_{\mathbf{R}^m} \psi(e^g - 1)d\hat{N} + A(t) \right\}.$$

The first and the second integrals of the right hand side are both Q-localmartingales. The last term $\{\cdots\}$ is a right continuous predictable process of bounded variation, which should be 0, since $X(t)$ is a Q-martingale. Therefore we get the representation of $X(t)$.

The representation can be extended to any Q locally square integrable martingale. Finally, the representation is valid for any Q-localmartingale. This can be verified through getting the exponential representation of positive Q-localmartingale similarly as in Theorem 2.1.

It is an interesting problem to find a condition for (f, g) which ensures that the localmartingale $\alpha_t(f, g)$ is a martingale. We give here a sufficient condition.

Theorem 2.4. *Suppose that $f \in \Phi$ and $g \in \Psi_{e.2}(\hat{N})$ satisfy*

$$(2.12)\quad E\Big[\exp\Big\{\int_0^T \Big((1+\epsilon)|f|^2 + \int_{g^+>\delta} e^{2(1+\epsilon)g^+}d\nu + 2(1+\epsilon)e^{2(1+\epsilon)\delta}\int_{|g|\le\delta} g^2 d\nu\Big)ds\Big\}\Big] < \infty,$$

for some $\epsilon > 0$ and $\delta > 0$, where $g^+ = \max(g, 0)$. Then $\alpha_t(f, g)$ is a martingale.

In particular, $\alpha_t(f, g)$ is a martingale if 1) $\int_0^T |f(s)|^2 ds$ is bounded a.s. and 2) $g^+, \hat{N}(g^+ > 1)$ and $\int_0^T \int_{|g|\le 1} |g|^2 \hat{N}(dsdz)$ are bounded a.s.

Proof. Let $\tau_n, n = 1, 2, \ldots$ be an increasing sequence of stopping times such that $P(\tau_n < T) \to 0$ as $n \to \infty$ and each stopped process $\alpha_{t\wedge\tau_n}$ is a martingale with mean 1. We want to prove that the above sequence of random variables (t is fixed) is uniformly integrable. If this property is verified, the limit process α_t is also a martingale. For this purpose it is sufficient to prove that $\sup_n E[\alpha_{t\wedge\tau_n}^p] < \infty$ holds for some $p > 1$. By a direct computation we can show that $\sup_n E[\alpha_{t\wedge\tau_n}^{1+\epsilon}] < \infty$, under the condition (2.12). Details are omitted.

§3. Processes with jumps and equivalent martingale measures

Let $\sigma(t) = (\sigma_j^i(t))$ be a $d \times m$ matrix valued predictable process, $b(t) = (b^i(t))$ be a d-vector predictable process and $v(t, z) = (v^i(t, z))$ be a d-vector predictable process continuous in $z \in \mathbf{R}^m$, which satisfy the integrability condition

$$(3.1)\quad \int_0^T |\sigma(t)|^2 + |b(t)|dt < \infty, \quad \int_0^T \int_{|z|\le 1} |v(s, z)|^2 ds\nu(dz) < \infty.$$

We shall consider a d-dimensional stochastic process with jumps defined by

$$(3.2)\quad \xi_t = \int_0^t \sigma(t)dW(t) + \int_0^t b(t)dt + \int_0^t \int_{|z|\le 1} v(t, z)\tilde{N}(dtdz) + \int_0^t \int_{|z|>1} v(t, z)N(dtdz),$$

where $W(t)$ is a m-dimensional standard Brownian motion and $N(dtdz)$ is a Poisson counting measure on $[0,T]\times\mathbf{R}^m$. To make the problem simple, we assume $d=m$ in this paper.

An equivalent probability measure Q such that ξ_t is a d-vector localmartingale with respect to $((\mathcal{F}_t),Q)$ is called an *equivalent martingale measure.* We denote by Γ the set of all equivalent martingale measures and by $\tilde{\Gamma}$ the set of all (f,g) such that $\alpha_t(f,g)dP\in\Gamma$. We shall characterize all equivalent martingale measures of a given process ξ_t by means of the pair (f,g).

Theorem 3.1. *Let $((\mathcal{F}_t),Q)$ be an equivalent probability measure and let $\alpha_t(f,g)$ be the density such that $dQ=\alpha_t(f,g)dP$, where $f\in\Phi$ and $g\in\Psi_{e,2}(\hat{N})$. Then the stochastic process ξ_t defined by (3.2) is a Q-localmartingale if and only if $v(s,z)(e^{g(s,z)}-1_{\{|z|\le 1\}})\in\Psi_1(\hat{N})$ and*

$$b(s)+\sigma(s)f(s)+\int_{\mathbf{R}^m}v(s,z)(e^{g(s,z)}-1_{\{|z|\le 1\}})\nu(dz)=0, \tag{3.3}$$

a.e. $\lambda\otimes P$, where λ is the Lebesgue measure.

Proof. In vector notation, we have by (2.2) and (3.2),

$$\begin{aligned}\xi_t\alpha_t &= \int_0^t \xi_{s-}d\alpha_s+\int_0^t\alpha_{s-}d\xi_s+[\xi,\alpha]_t\\ &= \text{a localmartingale}+\int_0^t\alpha_{s-}b(s)ds+\int_0^t\alpha_{s-}\sigma(s)f(s)ds\\ &\quad+\int_0^t\int_{\mathbf{R}^m}\alpha_{s-}v(s,z)(e^{g(s,z)}-1_{\{|z|\le 1\}})N(dsdz).\end{aligned}$$

If it is a d-vector localmartingale, the integrand with respect to $N(dsdz)$ should be integrable with respect to $\hat{N}(dsdz)$ and the equality

$$\alpha_{s-}b(s)+\alpha_{s-}\sigma(s)f(s)+\alpha_{s-}\int_{\mathbf{R}^m}v(s,z)(e^{g(s,z)}-1_{\{|z|\le 1\}})\nu(dz)=0, \tag{3.4}$$

holds a.e. (Theorem 1.1). Then we have (3.3), since $\inf_s\alpha_{s-}>0$ a.s. The converse will be shown similarly. The proof is complete.

An equivalent martingale measure $Q^0=\alpha_t(f^0,g^0)dP$ is said to be standard if ξ_t is a locally square integrable martingale with respect to Q^0. We will show the existence of such an equivalent martingale measure.

Lemma 3.2. *Assume that $\sigma(t)$ is invertible and $\sigma(t)^{-1}$ and $b(t)$ are bounded a.e. $\lambda\otimes P$. Then there exists a standard equivalent martingale measure. Further, for any given pair of $\phi\in\Phi$ and $\psi\in\Psi_{1,2}(\hat{N})$, there* *exists a standard equivalent martingale measure $Q^0 = \alpha_t(f^0, g^0)dP$ such that*

$$M^{Q^0}(t) := \int_0^t (\phi(s), dW^{f^0}(s)) + \tilde{N}_t^{g^0}(\psi) \tag{3.5}$$

is well defined as a locally square integrable martingale with respect to Q^0.

Proof. We will show that there exists a predictable pair $(f^0(s), g^0(s,z))$ of $\tilde{\Gamma}$ satisfying

$$\int_0^T \int_{\mathbf{R}^m} \{|v(s,z)|^2 + |\psi(s,z)|^2\} e^{g^0(s,z)} ds\nu(dz) < \infty. \tag{3.6}$$

For each $s \in [0,T]$, set $E(s) = E(s,\omega) = \{z : |\psi(s,z)| > 1\} \cup \{|z| > 1\}$. Then $\int_0^T \nu(E(s))ds < \infty$. Take first a nonpositive predictable process $g'(s,z)$ supported by $E(s)$, $|g'(s,z)| > 1$ on $E(s)$ and

$$\int_{E(s)} \{1 + |v(s,z)| + |v(s,z)|^2 + |\psi(s,z)|^2\} e^{g'(s,z)} \nu(dz)$$

is bounded in (s,ω) a.e. Take next a bounded predictable process $g''(s,z)$ supported by $E(s)^c$, $|g''(s,z)| < 1$ and $\int_{E(s)^c} (1 + |v(s,z)|)|g''(s,z)|\nu(dz)$ is bounded in (s,ω) a.e. Define $g^0 = g' 1_E + g'' 1_{E^c}$. Then $g^0 \in \Psi_{e,2}(\hat{N})$ and

$$\int_0^T \int_{E(s)} (|v|^2 + |\psi|^2) e^{g^0} ds\nu(dz) < \infty, \quad a.s.$$

Since $\int_0^T \int_{|z|\le 1} |v|^2 ds\nu(dz) < \infty$ and $\int_0^T \int_{\mathbf{R}^m} |\psi_2|^2 ds\nu(dz) < \infty$ for $\psi_2 = \psi 1_{|\psi|\le 1}$, we have $\int_0^T \int_{E(s)^c} (|v|^2 + |\psi|^2) e^{g^0} ds\nu(dz) < \infty$, *a.s.* Therefore (3.6) is satisfied.

The process $a(s) = \int_{\mathbf{R}^m} v(s,z)(e^{g^0(s,z)} - 1_{\{|z|\le 1\}})\nu(dz)$ is well defined since

(3.7)
$$\begin{aligned}
|a(s)| &\le \int_{E(s)} |v(s,z)(e^{g'(s,z)} - 1_{|z|\le 1})|\nu(dz) \\
&\qquad + \int_{E(s)^c} |v(s,z)(e^{g''} - 1_{|z|\le 1})|\nu(dz) \\
&\le \int_{E(s)} |v| e^{g'} \nu(dz) + \int_{E(s)^c} |v||g''|\nu(dz) \quad \text{bounded in } (s,\omega) \ a.e.
\end{aligned}$$

Then we can define $f^0(s)$ by $b(s) + \sigma(s)f^0(s) + a(s) = 0$. The pair (f^0, g^0) satisfies (3.3). Further, it satisfies conditions (1),(2) of Theorem 2.4. Indeed, we took g^0 so that it satisfies (2). By the estimation (3.7), $|a(s)|$ is bounded a.s. Since $|b(s)|$ is bounded and $\sigma(s)\sigma(s)^T$ is uniformly positive definite, $|f^0(s)|$ is also bounded a.s. Thus $f^0(s)$ satisfies (1) of the theorem. Then $\alpha_t(f^0, g^0)$ is a martingale.

Let $Q^0 = \alpha_T(f^0, g^0)dP$. We will show that ξ_t is a locally square integrable martingale with respect to Q^0. Observe (3.2) and (3.3). Then $\xi_t = (\xi_t^1, ..., \xi_t^d)$ is written as

$$\xi_t = \int_0^t \sigma(s)dW^{f^0}(s) + \int_0^t \int_{\mathbf{R}^m} v(s,z)\tilde{N}^{g^0}(dsdz).$$

The bracket process with respect to Q^0 is given by a $d \times d$ matrix

$$\left(\langle \xi^i, \xi^j \rangle_t^{Q^0}\right) = \int_0^t \sigma(s)\sigma(s)^T ds + \int_0^t \int_{\mathbf{R}^m} v(s,z)v(s,z)^T \hat{N}^{g^0}(dsdz).$$

It is finite a.s. This proves that ξ_t is a locally square integrable martingale.

Finally, $M^{Q^0}(t)$ of (3.5) is well defined as a locally square integrable martingale, because $\psi \in \Psi_2(\hat{N}^{g^0})$ by (3.6). The proof is complete.

We will fix the equivalent martingale measure Q^0 of Lemma 3.2. Set $\hat{f} = f - f^0$, $\hat{g} = g - g^0$. Then $\alpha_t(f,g)$ of Theorem 3.1 is decomposed to the product of two exponential semimartingales;

$$\alpha_t(f,g) = \alpha_t(f^0, g^0)\alpha_t^0(\hat{f}, \hat{g}), \tag{3.8}$$

(3.9)
$$\alpha_t^0(\hat{f},\hat{g}) = \exp\left\{\int_0^t (\hat{f}(s), dW^{f^0}(s)) - \frac{1}{2}\int_0^t |\hat{f}(s)|^2 ds \right.$$
$$\left. + N_t(\hat{g}_1) - \hat{N}_t^{g^0}(e^{\hat{g}_1} - 1) + \tilde{N}_t^{g^0}(\hat{g}_2) - \hat{N}_t^{g^0}(e^{\hat{g}_2} - 1 - \hat{g}_2)\right\}.$$

Since $dQ = \alpha_T(f,g)dP$ and $dQ^0 = \alpha_T(f^0, g^0)dP$, we have $dQ = \alpha_T^0(\hat{f}, \hat{g})dQ^0$. Hence Q is an equivalent martingale measure with respect to Q^0 and $\alpha_t^0(\hat{f}, \hat{g})$ is its density process. Then ξ_t, $\alpha_t^0(\hat{f}, \hat{g})$ and $\xi_t\alpha_t^0(\hat{f}, \hat{g})$ are all localmartingales with respect to Q^0.

Conversely if Q is an equivalent martingale measure with respect Q^0. The density process α_t^0 of Q with respect to Q^0 is represented by (3.9). We denote by $\hat{\Gamma}^0$ the set of all such density processes α_t^0 and by

$\hat{\Gamma}_2^0$ is the set of $\alpha_t^0 \in \hat{\Gamma}^0$ such that these are all locally square integrable martingales with respect to Q^0.

Let $\mathcal{M}_{loc}^2(Q^0)$ be the set of all locally square integrable martingales $M(t)$ with $M(0) = 0$ with respect to Q^0. Then the d-vector process ξ_t belongs to $\mathcal{M}_{loc}^2(Q^0)$. Further, ξ_t and $\alpha_t^0 - 1$ are orthogonal with respect to Q^0, if α_t^0 is locally square integrable with respect to Q^0. We claim;

Lemma 3.3. *Assume that $\sigma(t)$ is invertible and $\sigma(t)^{-1}$ and $b(t)$ are bounded a.e. $\lambda \otimes P$. Let Q^0 be a standard equivalent martingale measure. Then, with respect to Q^0, we have the orthogonal decomposition of $\mathcal{M}_{loc}^2(Q^0)$.*

$$\mathcal{M}_{loc}^2(Q^0) = \mathcal{L}(\xi_t^1, ..., \xi_t^d) \oplus \mathcal{L}\left(\alpha_t^0 - 1; \alpha_t^0 \in \hat{\Gamma}_2^0\right). \tag{3.10}$$

Proof. Let $\mathcal{K} = \mathcal{L}(\xi_t^1, ..., \xi_t^d)^{\perp}$ and let M be any element of $\mathcal{K}$ represented by $M = \int(\phi, dW^{f^0}) + \tilde{N}^{g^0}(\psi)$. Since it is orthogonal to $\xi_t^1, ..., \xi_t^d$ with respect to Q^0,

$$\langle \xi^i, M\rangle_t^{Q_0} = \int_0^t (\sigma^i(s), \phi(s))ds + \int_0^t \left(\int v^i(s,z)\psi(s,z)e^{g^0}\nu(dz)\right) ds = 0,$$
$$i = 1, ..., d.$$

Therefore, setting $\phi(t)=(\phi^1(t), ..., \phi^d(t))$ and $v(t,z)=(v^1(t,z), ..., v^d(t,z))$, we get

$$\sigma(t)\phi(t) + \int_{\mathbf{R}^d} \psi(t,z)v(t,z)e^{g^0(t,z)}\nu(dz) = 0, \quad \forall t.$$

We will show that

$$\mathcal{H} := \left\{\int(\phi, dW^{f^0}) + \tilde{N}^{g^0}(\psi) \in \mathcal{K};\ \psi \text{ are bounded}\right\}$$

is dense in $\mathcal{K}$. Let $M = \int(\phi, dW^{f^0}) + \tilde{N}^{g^0}(\psi)$ be any element of $\mathcal{K}$. We define trancated functions by $\psi_n = (\psi \wedge n) \vee (-n)$. Next define d-vector functions by $\phi_n(t) = -\sigma(t)^{-1} \int \psi^n(t,z)v(t,z)e^{g^0(t,z)}\nu(dz)$. Then it holds

$$\sigma(t)\phi^n(t) + \int \psi^n(t,z)v(t,z)e^{g^0(t,z)}\nu(dz) = 0, \quad \forall t.$$

Therefore $M^n = \int(\phi^n, dW^{f^0}) + \tilde{N}^{g^0}(\psi^n)$ belongs to $\mathcal{K}$. Further, since $\int_0^T \int |\psi_n - \psi|^2 e^{g^0}\nu(dz)ds \to 0$ holds valid as $n \to \infty$, $\int_0^T |\phi^n - \phi|^2 ds \to 0$ as $n \to \infty$. Therefore the sequence $\{M^n\}$ converges to M with respect to the topology of $\mathcal{M}_{loc}^2(Q^0)$. We have thus shown that $\mathcal{H}$ is dense in $\mathcal{K}$.

Let $\mathcal{J}$ be the set of all $M \in \mathcal{K}$ which is bounded from the below. Then we have $\mathcal{L}(\mathcal{J}) = \mathcal{L}(\alpha_t^0 - 1; \alpha_t^0 \in \hat{\Gamma}_2^0\}$. Further it holds $\mathcal{L}(\mathcal{J}) \supset \mathcal{L}(\mathcal{H})$. Indeed, we have $\{M^\tau(t); M \in \mathcal{H}\} \subset \mathcal{L}(\mathcal{J})$, where τ are stopping times such that $M^\tau(t) := M(t \wedge \tau)$ are bounded localmartingales. We have thus proved

$$\mathcal{L}(\xi_t^1, \ldots, \xi_t^d)^\perp = \mathcal{K} = \mathcal{L}(\mathcal{H}) \subset \mathcal{L}(\mathcal{J}) = \mathcal{L}(\alpha_t^0 - 1; \alpha_t^0 \in \hat{\Gamma}_2^0).$$

The proof is complete.

We are now in a position of stating a main result of the paper.

Theorem 3.4. *Assume that $\sigma(t)$ is invertible and $\sigma(t)^{-1}$ and $b(t)$ are bounded a.e. $\lambda \otimes P$. If $X(t)$ is a supermartingale for any equivalent martingale measure Q, then it is represented by*

$$X(t) = X(0) + \int_0^t (\varphi(s), d\xi_s) - A(t). \tag{3.11}$$

Here, $A(t)$ is a predictable increasing process and $\varphi(s)$ is a predictable process such that $\sigma(s)\varphi(s) \in \Phi$ and $(\varphi(s), v(s,z)) \in \Psi_{1,2}(\hat{N}^g)$ for any $(f,g) \in \tilde{\Gamma}$.

If $X(t)$ is a localmartingale for any equivalent martingale measure, then it is represented by

$$X(t) = X(0) + \int_0^t (\varphi(s), d\xi_s). \tag{3.12}$$

Proof. For each $Q \in \Gamma$, the supermartingale $X(t)$ is decomposed as $X(0) + M^Q(t) - A^Q(t)$, where $M^Q(t)$ is a Q-localmartingale with $M^Q(0) = 0$ and $A^Q(t)$ is a natural (=predictable) increasing process, by Doob-Meyer decomposition. The Q-localmartingale $M^Q(t)$ is represented by $M^Q(t) = \int_0^t (\phi(s), dW^f(s)) + \tilde{N}_t^g(\psi)$. We will show that the kernel (ϕ, ψ) does not depend on the choice of Q. Let Q^* be another equivalent martingale measure. Then M^{Q^*} is represented by $M^{Q^*} = \int (\phi^*(s), dW^{f^*}(s)) + \tilde{N}^{g^*}(\psi^*)$. Since $M^Q(t) - A^Q(t) = M^{Q^*}(t) - A^{Q^*}(t)$, we have

$$\tilde{N}_t^{g^*}(\psi^*) - \tilde{N}_t^g(\psi) = \left(\int_0^t (\phi(s), dW^f(s)) - \int_0^t (\phi^*(s), dW^{f^*}(s)) \right) - \left(A^Q(t) - A^{Q^*}(t) \right).$$

The right hand side is a predictable process, so that it has no common jumps with the Poisson random measure $N(dtdz)$. So both sides of the above can not have jumps. This shows $\psi = \psi^*$ a.e. $\hat{N} \otimes P$. Hence the right hand side should be a predictable process of bounded variation. Therefore $\int_0^t (\phi(s), dW^f(s)) - \int_0^t (\phi^*(s), dW^{f^*}(s))$ is also a predictable process of bounded variation, which shows $\phi = \phi^*$ a.e. $\lambda \otimes P$. We have $\phi \in \Phi$ and $\psi \in \Psi_{1,2}(\hat{N}^g)$ for any $(f, g) \in \tilde{\Gamma}$.

We want to prove $A^Q = A^{Q^0}$ in the case where both Q and Q^0 are standard equivalent martingale measures such that $\psi \in \Psi_2(\hat{N}^g) \cap \Psi_2(\hat{N}^{g^0})$. Comparing two equations for $M^Q(t)$ and $M^{Q^0}(t)$, $M^{Q^0}(t)$ can be written as

$$\begin{aligned} M^{Q^0}(t) =& M^Q(t) + \int_0^t (\phi(s), \hat{f}(s)) ds \\ & + \int_0^t \int_{\mathbf{R}^m} \psi(s,z)(e^{\hat{g}(s,z)} - 1)e^{g^0(s,z)} ds\nu(dz), \end{aligned}$$

where $\hat{f} = f - f^0$ and $\hat{g} = g - g^0$, because $\psi, e^{\hat{g}} - 1 \in \Psi_2(\hat{N}^{g^0})$. Therefore we have

$$A^Q(t) = \int_0^t \left\{ (\phi, \hat{f}) + \int_{\mathbf{R}^m} \psi(e^{\hat{g}} - 1)e^{g^0} d\nu \right\} ds + A^{Q^0}(t).$$

We claim

$$(\phi, \hat{f}) + \int_{\mathbf{R}^m} \psi(e^{\hat{g}} - 1)e^{g^0} d\nu = 0, \tag{3.13}$$

a.e. $\lambda \otimes P$, in the case where $g \leq g^0$ or equivalently $\hat{g} \leq 0$. If it is not the case, then either the set

$$F = \left\{ (s, \omega); (\phi, \hat{f}) + \int_{\mathbf{R}^m} \psi(e^{\hat{g}} - 1)e^{g^0} d\nu > 0 \right\} \quad \text{or}$$

$$F' = \left\{ (s, \omega); (\phi, \hat{f}) + \int_{\mathbf{R}^m} \psi(e^{\hat{g}} - 1)e^{g^0} d\nu < 0 \right\}$$

is of positive measure with respect to $\lambda \otimes P$. Suppose that F is of positive measure. We define (f', g') by $f' = f^0 - n\hat{f} 1_F$ and $g' = g^0 + \log\{1 - n(e^{\hat{g}} - 1)1_F\}$. Then it holds $e^{\hat{g}'} - 1 = -n(e^{\hat{g}} - 1)1_F$, where $\hat{g}' = g' - g^0$. Set $\hat{f}' = f' - f^0$. Then $\alpha'_t := \alpha^0_t(\hat{f}', \hat{g}')$ is a positive localmartingale with respect to Q^0. Further $\xi_t \alpha'_t$ is a localmartingale with respect to Q^0. Indeed, equalities

$$\langle \xi^i_t, \alpha'_t \rangle^{Q^0}_t = -n \int_0^t \alpha'_{s-} 1_F \left\{ (\sigma^i, \hat{f}) + \int_{\mathbf{R}^m} v^i(e^{\hat{g}} - 1)e^{g^0} \nu(dz) \right\} ds = 0,$$
$$i = 1, \ldots, d$$

hold valid since ξ_t^i and $\alpha_t^0(\hat{f},\hat{g})$ are orthogonal with respect to Q^0. Let $\{\tau_k, k=1,2,...\}$ be an increasing sequence of stopping times such that $P(\tau_k < T) \to 0$ as $k \to \infty$ and each stopped process $\alpha'_{t\wedge\tau_k}$ is a Q^0-martingale. Define a sequence of probability measures Q'_k by $dQ'_k = \alpha'_{\tau_k} dQ^0$. Then each Q'_k is an equivalent martingale measure for the stopped process $\xi_{t\wedge\tau_k}$. Then the stopped process $X(t\wedge\tau_k)$ is a supermartingale with respect to Q'_k for each k. Its Doob-Meyer decomposition is represented by

$$X(t\wedge\tau_k) = \int_0^{t\wedge\tau_k} (\phi(s), dW^{f'}) + \tilde{N}^{g'}_{t\wedge\tau_k}(\psi) - A^{\alpha'}(t\wedge\tau_k), \quad k=1,2,...$$

where $A^{\alpha'}(t)$ is a suitable predictable increasing process. It satisfies

$$A^{\alpha'}(t) = -n\int_0^t 1_F \left\{(\phi,\hat{f}) + \int_{\mathbf{R}^m} \psi(e^{\hat{g}}-1)e^{g^0}d\nu\right\} ds + A^{Q^0}(t).$$

This makes a contradiction since the right hand side is negative for sufficiently large n. Therefore we get $A^Q(t) = A^{Q^0}(t)$.

Now if F' is of positive measure instead of the set F, interchange the role of Q^0 and Q in the above discussion. Then we get the same conclusion. Further in the case where $g \geq g^0$, we get the same equality (3.13) by interchanging the role of Q^0 and Q.

We have thus seen that $M^Q(t) = M^{Q^0}(t)$ holds for any standard equivalent martingale measure Q such that its density process α_t^0 with respect to Q^0 is a locally square integrable martingale and $g \leq g^0$ or $g \geq g^0$ is satisfied. Let $\tilde{\Gamma}_2^0$ be the set of all $\alpha_t^0(\hat{f},\hat{g}) \in \Gamma_2^0$ such that Q with $dQ = \alpha^0(\hat{f},\hat{g})dP^0$ is a standard equivalent martingale measure and $\psi \in \Psi_2(\hat{N}^g)$. Then $M^{Q^0}(t)$ is orthogonal to any element of

$$\mathcal{N} = \left\{\alpha_t^0(\hat{f},\hat{g}) - 1;\ \alpha^0(\hat{f},\hat{g}) \in \tilde{\Gamma}_2^0,\ \hat{g} \leq 0 \text{ or } \hat{g} \geq 0\right\},$$

with respect to Q^0. Observe that $\mathcal{L}(\mathcal{N}) = \mathcal{L}(\alpha_t^0 - 1; \alpha_t^0 \in \Gamma_2^0)$. Then we see that $M^{Q^0}(t)$ belongs to $\mathcal{L}(\xi_t^1, ..., \xi_t^d)$ by the decomposition formula (3.10). Then it is represented by $\int_0^t (\varphi(s), d\xi_s)$ with respect to Q^0. Setting $A(t) = A^{Q^0}(t)$, we get the decomposition formula (3.11). Further, representation (3.11) should hold valid for any Q of Γ. Then, comparing this with the representation of Theorem 2.3, $\sigma(s)\varphi(s) \in \Phi$ and $(v(s,z), \varphi(s)) \in \Psi_{1,2}(\hat{N}^g)$.

The second assertion of the theorem follows immediately from the above discussion by setting $A^{Q^0}(t) = A^Q(t) \equiv 0$.

Applications to mathematical finance

We consider a simple market model, where the return process is given by a stochastic process ξ_t of (3.2) and the interest rate $r(t)$ is identically 0. Let $\pi(t)$ be a predictable process called a *strategy* or *portfolio* and C_t be a right continuous predictable increasing process called a *cumulative consumption process.* A stochastic process $X(t) = X^{x,\pi,C}(t)$ defined by

$$X(t) = x + \int_0^t (\pi_s, d\xi_s) - C(t)$$

is called a *wealth process.* We introduce admissible classes for pairs of portofolios and consumptions. Let $x > 0$. We denote by $\mathcal{A}^+(x)$ the set of the pair (π, C) such that $X^{x,\pi,C}(t) \geq 0$ holds a.s. for any $0 \leq t \leq T$. We denote by $\mathcal{A}^-(-x)$ the set of the pair (π, C) such that $X^{-x,\pi,C}(t) \leq 0$ for any $0 \leq t \leq T$.

A *Europian contingent claim* Y is a nonnegative $\mathcal{F}_T$-measurable random variable. The contingent claim is not always attainable, since the model is not complete due to jumps of the return process. We shall study the upper and lower hedging price. The *upper hedging price* and *lower hedging price* of the contingent claim Y are defined respectively by

$$\begin{aligned} h_{up} &= \inf\{x \geq 0; \\ &\qquad \text{there exists } (\pi, C) \in \mathcal{A}^+(x) \text{ such that } X^{x,\pi,C}(T) \geq Y \ a.s.\} \\ h_{low} &= \sup\{x \geq 0; \\ &\qquad \text{there exists } (\pi, C) \in \mathcal{A}^-(-x) \text{ such that } X^{-x,\pi,C}(T) \geq -Y \ a.s.\} \end{aligned}$$

Theorem 3.5. *Assume that $\sigma(t)\sigma(t)^T$ is uniformly positive definite, $v(t,z)$ is greater than -1 and $v(t,z) \neq 0$ a.e. $\lambda \times \nu \times P$. Let Y be an Europian contingent claim. We have*

$$h_{up} = \sup_{Q \in \Gamma} E_Q[Y] =: h, \tag{3.14}$$

$$h_{low} = \inf_{Q \in \Gamma} E[Y] =: f. \tag{3.15}$$

If h is finite (resp. f is positive), there exists a pair $(\pi, C) \in \mathcal{A}^+(h)$ (resp. $(\pi', C') \in \mathcal{A}^-(-f)$) such that

$$X^{h,\pi,C}(t) = \operatorname{ess\,sup}_{Q \in \Gamma} E_Q[Y|\mathcal{F}_t], \tag{3.16}$$

$$-X^{-f,\pi',C'}(t) = \operatorname{ess\,inf}_{Q \in \Gamma} E_Q[Y|\mathcal{F}_t], \tag{3.17}$$

holds for any t. *In particular,* $X^{h,\pi,C}(T)=Y$ *and* $X^{-f,\pi',C'}(T)=-Y$, *a.s.*

Proof. We consider the upper hedging price only. Set $h=\sup_{Q\in\Gamma}E_Q[Y]$. We want to prove $h=h_{up}$. We first show $h\le h_{up}$. The inequality is obvious if $h_{up}=\infty$. If $h_{up}<\infty$, let x be an arbitrary element in the set $\{\cdots\}$ appearing in the defintion of h_{up}. Then there exists a pair (π,C) of $\mathcal{A}^+(x)$ such that $X^{x,\pi,C}(T)\ge Y$. Then $X^{x,\pi,C}(t)$ is a supermartingale for any $Q\in\Gamma$. Therefore, $E_Q[Y]\le E_Q[X^{x,\pi,C}(T)]\le x$ holds for any $Q\in\Gamma$. Then we have $h\le x$, so that we have $h\le h_{up}$.

In order to prove the reverse inequality $h\ge h_{up}$, it is sufficient to construct $(\pi,C)\in\mathcal{A}^+(h)$ such that $X^{h,\pi,C}(t)=X(t)$, where $X(t)\equiv \operatorname{ess\,sup}_{Q\in\Gamma}E_Q[Y|\mathcal{F}_t]$. It is known that the process $X(t)$ is a supermartingale for any Q. We shall apply Theorem 3.4 to the return process ξ_t. Then $X(t)$ admits the decomposition (3.11) by Theorem 3.4. This implies $X(t)=X^{h,\pi,C}(t)$, by setting $\varphi=\pi$ and $A(t)=C(t)$. It is clearly nonnegative a.s. for any $t\in(0,T]$. Therefore (π,C) belongs to $\mathcal{A}^+(h)$.

References

[1] H. Höllmer, Y. Kabanov, Optional decomposition and Lagrange multpliers, Finance and Stochastics, 2(1998), 69-81.

[2] K. Itô, Spectral type of the shift transformation of differential processes with stationary increments, Trans. Amer. Math. Soc. 81(1956), 253-263.

[3] Yu.M.Kabanov, Ch.Stricker, Hedging of contingent claims under transaction costs, 2002(preprint).

[4] I. Karatzas, Lectures on the mathematics of finance, CRM 8, AMS, 1997.

[5] I. Karatzas, S.E. Shreve, Methods of Mathematical Finance, Springer, 1998.

[6] H. Kunita, S. Watanabe, On square integrable martingales, Nagoya Math. J. 30(1967), 209-245.

[7] P.A. Meyer, Un cours sur integrales stochastiques, Seminaire Proba. X, Lecture Notes in Math. 511, 246-400, Springer, Berlin Heidelberg New York, 1976.

Department of Mathematical Science,
Nanzan University,
Seto, Aichi 489-0863,
Japan

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 233–246

Stochastic Newton Equation with Reflecting Boundary Condition

Shigeo Kusuoka

§1. Introduction

Let D be a bounded domain in $\mathbf{R}^d$ with a smooth boundary and $n(x)$, $x \in \partial D$, be an outer normal vector. Let $a^{ij} : \mathbf{R}^d \to \mathbf{R}$, $i, j = 1, \ldots d$, be smooth functions such that $a^{ij}(x) = a^{ji}(x)$, $x \in \mathbf{R}^d$. Also, let $b^i : \mathbf{R}^{2d} \to \mathbf{R}$, $i = 1, \ldots d$, be bounded measurable functions. We assume that there are positive constants C_0, C_1 such that

$$C_0|\xi|^2 \leq \sum_{i,j=1}^{d} a^{ij}(x)\xi_i\xi_j \leq C_1|\xi|^2, \qquad x, \xi \in \mathbf{R}^d.$$

Let L_0 be a second order linear differential operator in $\mathbf{R}^{2d}$ given by

$$L_0 = \sum_{i=1}^{d} v^i \frac{\partial}{\partial x^i} + \frac{1}{2} \sum_{i,j=1}^{d} a^{ij}(x) \frac{\partial^2}{\partial v^i \partial v^j} + \sum_{i=1}^{d} b^i(x,v) \frac{\partial}{\partial v^i}$$

Let $\tilde{W}^d = C([0,\infty); \mathbf{R}^d) \times D([0,\infty); \mathbf{R}^d)$. Now let $\Phi : \mathbf{R}^d \times \partial D \to \mathbf{R}^d$ be a smooth map satisfying the following .
(i) $\Phi(\cdot, x) : \mathbf{R}^d \to \mathbf{R}^d$ is linear for all $x \in \partial D$.
(ii) $\Phi(v,x) = v$ for any $x \in \partial D$ and $v \in T_x(\partial D)$, i.e., $\Phi(v,x) = v$ if $x \in \partial M$, $v \in \mathbf{R}^d$ and $v \cdot n(x) = 0$.
(iii) $\Phi(\Phi(v,x),x) = v$ for all $v \in \mathbf{R}^d$ and $x \in \partial D$.
(iv) $\Phi(n(x),x) \neq n(x)$ for any $x \in \partial D$.

The main theorem in the present paper is the following.

Theorem 1. *Let $(x_0, v_0) \in (\bar{D})^c \times \mathbf{R}^d$. Then there exists a unique probability measure μ over $\tilde{W}^d$ satisfying the following conditions.*
(1) $\mu(w(0) = (x_0, v_0)) = 1$.
(2) $\mu(w(t) \in D^c \times \mathbf{R}^d, t \in [0,\infty)) = 1$.

Received April 4, 2003.

(3) *For any* $f \in C_0^\infty((\bar{D})^c \times \mathbf{R}^d)$, $\{f(w(t)) - \int_0^t L_0 f(w(s))ds; t \geq 0\}$ *is a martingale under* $\mu(dw)$.
(4) $\mu(1_{\partial D}(x(t))(v(t) - \Phi(v(t-), x(t))) = 0$ *for all* $t \in [0, \infty)) = 1$.
Here $w(\cdot) = (x(\cdot), v(\cdot)) \in \tilde{W}^d$.

Now let us think of the following Stochastic Newton equation

$$\begin{aligned} dX_t^\lambda &= V_t^\lambda dt \\ dV_t^\lambda &= \sigma(X_t^\lambda)dB(t) + (b(X_t^\lambda, V_t^\lambda) - \lambda\nabla U(X_t^\lambda))dt \end{aligned}$$

$$X_0^\lambda = x_0, \qquad V_0^\lambda = v_0.$$

Here $B(t)$ is a d-dimensional Brownian motion, $\sigma \in C^\infty(\mathbf{R}^d; \mathbf{R}^d)$, $b : \mathbf{R}^{2d} \to \mathbf{R}^d$ is a bounded Lipschitz continuous function, and $U \in C_0^\infty(\mathbf{R}^d)$.

We assume the following also.
(A-1) There are positive constants C_0, C_1 such that

$$C_0|\xi|^2 \leq |\sigma(x)\xi|^2 \leq C_1|\xi|^2, \qquad x, \xi \in \mathbf{R}^d.$$

(A-2) Let $D = \{x \in \mathbf{R}^d;\ U(x) > 0\}$. Then there are $\varepsilon_0 > 0$, $U_0 \in C^\infty(\mathbf{R}^d; \mathbf{R})$ and a non-increasing C^1-function $\rho : \mathbf{R} \to \mathbf{R}$ satisfying the following.
(1) $x \in \partial D$, if and only if $U_0(x) = 0$ and $dis(x, \partial D) < \varepsilon_0$.
(2) $\nabla U_0(x) \neq 0$, $x \in \partial D$.
(3) $\rho(t) = 0$, $t \geq 0$, $\rho(t) > 0$, $t < 0$, and $U(x) = \rho(U_0(x))$ for $x \in \mathbf{R}^d$ with $dis(x, \partial D) < \varepsilon_0$.
(4) $\displaystyle\lim_{t \uparrow 0} \frac{\rho'(t)}{\rho(t)} = -\infty$.

Now let $\tilde{dis}$ be a metric function on $\tilde{W}^d$ given by

$$\begin{aligned} &\tilde{dis}(w_0, w_1) \\ &= \sum_{n=1}^\infty 2^{-n}(1 \wedge ((\max_{t\in[0,n]} |x_0(t) - x_1(t)|) + (\int_0^n |v_0(t) - v_1(t)|^n)^{1/n})), \end{aligned}$$

for $w_i(\cdot) = (x_i(\cdot), v_i(\cdot)) \in \tilde{W}^d$, $i = 0, 1$.

Then we will show the following.

Theorem 2. *Let* ν^λ, $\lambda \in [1, \infty)$, *be the probability law of* $(X_t^\lambda, V_t^\lambda)$, $t \in [0, \infty)$, *on* $\tilde{W}_0$, *and* μ *be the probability measure given in Theorem 1 in the case when* $\Phi(v, x) = v - 2(v \cdot n(x))n(x)$, $v \in \mathbf{R}^d$, $x \in \partial D$. *Then* ν^λ *conveges to* μ *weakly as* $\lambda \to \infty$ *as probability measures on* $(\tilde{W}_0, \tilde{dis})$.

§2. Basic lemmas

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,\infty)}, P)$ be a filtered probability space, and $B(t) = (B^1(t), \dots, B^d(t))$ be a d-dimensional Brownian motion. Let $B^0(t) = t$, $t \in [0,\infty)$. Let $\sigma_i : \mathbf{R}^N \to \mathbf{R}^N$, $i = 0, 1, \dots, d$, be Lipschitz continuous functions, and let $X : [0,\infty) \times \mathbf{R}^N \times \Omega \to \mathbf{R}^N$ be the solution to the following SDE

$$X(t,x) = x + \sum_{i=0}^{d} \int_0^t \sigma_i(X(s,x))dB^i(s), \qquad t \geq 0,\ x \in \mathbf{R}^N.$$

We may assume that $X(t,x)$ is continuous in (t,x) (cf. Kunita [2]).

Then we have the following.

Lemma 3. *For any $T > 0$ and $p_0, p_1, \dots, p_m \in (1,\infty)$, $m \geq 1$, with $\sum_{k=0}^m p_k^{-1} = 1$, there is a constant $C > 0$ such that*

$$E[\int_{\mathbf{R}^N} \prod_{k=0}^{m} |f_k(X(t_k,x))|dx] \leq C \prod_{k=0}^{m} \| f_k \|_{L^{p_k}(\mathbf{R}^N,dx)}$$

for all $0 = t_0 < t_1 < \dots < t_m \leq T$, and $f_k \in C_0^\infty(\mathbf{R}^N)$, $k = 0, 1, \dots, m$.

Proof. From the assumption, there is a $C_0 > 0$ such that

$$|\sigma_i(x) - \sigma_i(y)| \leq C_0|x-y|, \qquad x, y \in \mathbf{R}^N.$$

Let $\varphi \in C_0^\infty(\mathbf{R}^N)$ such that $\int_{\mathbf{R}^N} \varphi(x)dx = 1$. Let $\varphi_n(x) = n^N\varphi(nx)$, $x \in \mathbf{R}^N$, for $n \geq 1$, and let $\sigma_i^{(n)} = \varphi_n * \sigma_i$, $i = 0, \dots, d$. Then $\sigma_i^{(n)} \in C^\infty(\mathbf{R}^N; \mathbf{R}^N)$. Let

$$W_{i,k}^{(n),j}(x) = \frac{\partial}{\partial x^k}\sigma_i^{(n),j}(x),$$
$$x \in \mathbf{R}^N,\ j,k = 1\dots,N,\ i = 0,1,\dots,d,\ n \geq 1.$$

Then we see that $|W_{i,k}^{(n),j}(x)| \leq C_0$, $x \in \mathbf{R}^N$. Let $X^{(n)} : [0,\infty) \times \mathbf{R}^N \times \Omega \to \mathbf{R}^N$ be the solution to the following SDE

$$X^{(n)}(t,x) = x + \sum_{i=0}^{d} \int_0^t \sigma_i^{(n)}(X^{(n)}(s,x))dB^i(s), \qquad t \geq 0,\ x \in \mathbf{R}^N.$$

Then we may think that $X^{(n)}(t,\cdot) : \mathbf{R}^N \to \mathbf{R}^N$ is a diffeomorphism with probability one. Let $J_k^{(n),j}(t,x) = \frac{\partial}{\partial x^k} X^{(n),j}(t,x)$. Let $W_i^n(x) =$

$(W_{i,k}^{(n),j}(x))_{k,j=1,\dots,N}$ and $J^{(n)}(t,x) = (J_k^{(n),j}(t,x))_{k,j=1,\dots,N}$. Then the $N \times N$-matrix valued process $J^{(n)}(t,x)$ satisfies the following SDE

$$J^{(n)}(t,x) = I_N + \sum_{i=0}^{d} \int_0^t W_i^{(n)}(X^{(n)}(s,x))J^{(n)}(s,x)dB_i(s).$$

Also, we see that

$$\begin{aligned} &J^{(n)}(t,x)^{-1} \\ &= I_N - \sum_{i=0}^{d} \int_0^t J^{(n)}(s,x)^{-1}W_i^{(n)}(X^{(n)}(s,x))dB_i(s) \\ &\quad + \frac{1}{2}\sum_{i=1}^{d} \int_0^t J^{(n)}(s,x)^{-1}W_i^{(n)}(X^{(n)}(s,x))^2 ds. \end{aligned}$$

Then we see that

$$C_T = \sup\{E[\det J^{(n)}(t,x)^{-p_0+1}];\ t \in [0,T], x \in \mathbf{R}^N, n \geq 1\} < \infty.$$

So we have

$$\begin{aligned} &E[\int_{\mathbf{R}^N} \prod_{k=0}^{n} |f_k(X^{(n)}(t_k,x))|dx] \\ &\leq E[\int_{\mathbf{R}^N} |f_0(x)|_0^p (\prod_{k=1}^{m} \det J^{(n)}(t_k,x)^{-p_0/p_k})dx]^{1/p_0} \\ &\quad \times \prod_{k=1}^{m} E[\int_{\mathbf{R}^N} |f_k(X^{(n)}(t_k,x))|^{p_k} \det J^{(n)}(t_k,x)dx]^{1/p_k} \\ &\leq C_T(\int_{\mathbf{R}^N} |f(x)|_0^p dx)^{1/p_0} \prod_{k=1}^{m}(\int_{\mathbf{R}^N} |f_k(x))|^{p_k}dx)^{1/p_k} \end{aligned}$$

Letting $n \to \infty$, we have our assertion. ∎

Now let D be a bounded domain in $\mathbf{R}^N$ and $F^j : \mathbf{R}^N \to \mathbf{R}$, $j = 1, 2$, be C^2 functions satisfying the following assumptions (F1),(F2), furthermore.

(F1) For $x \in D$ and $i = 1, \dots, d$,

$$\sum_{j=1}^{N} \sigma_i^j(x) \frac{\partial}{\partial x^j} F^1(x) = 0.$$

(F2) $\inf\{\det(\nabla F^i(x) \cdot \nabla F^j(x))_{i,j=1,2};\ x \in D\} > 0$.

Then we have the following

Lemma 4. *For a.e.x,*

$$P(X(t,x) \in D, F(X(t,x)) = 0 \text{ for some } t > 0) = 0.$$

Here $F = (F^1, F^2) : \mathbf{R}^N \to \mathbf{R}^2$.

Proof. Let

$$\tau(s,x) = \inf\{t \geq s; X(t,x) \in D^c\} \wedge (s+1), \qquad x \in \mathbf{R}^N, s > 0.$$

Also, let

$$p(x,s) = P(F(X(t,x)) = 0 \text{ for some } t \in [s, \tau(s,x))), \qquad x \in \mathbf{R}^N, s > 0.$$

Then we see that

$$P(X(t,x) \in D, F(X(t,x)) = 0 \text{ for some } t > 0) \leq \sum_{r \in \mathbf{Q}_+} p(x,r),$$

where $\mathbf{Q}_+$ is the set of positive rational numbers. Let $V(m) = \{x \in \mathbf{R}^N;\ |x| \leq m\}$, $m \geq 1$. Let us define random variables $Z_{T,m}$, $T > 0$, $m \geq 1$, and constant C_1 by

$$Z_{T,m} = \sup\{|t-s|^{-1/3}|X(t,x) - X(s,x)|;\ 0 \leq s < t \leq T,\ x \in V(m)\},$$

and

$$C_1 = \sup\{|\sigma_0(x)||\nabla F^1(x)| + \frac{1}{2}\sum_{i=1}^{d} |\nabla^2 F^1(x)||\sigma_i(x)|^2 + |\nabla F^2(x)|;\ x \in \bar{D}\}.$$

Then we see that $P(Z_{T,m} < \infty) = 1$ (cf. Kunita[2]). By the assumtion (F1), we see that

$$F^1(X(t,x)) = F^1(x) + \int_0^t (\sigma_0(X(s,x))\nabla F^1(X(s,x))$$

$$+ \sum_{i=1}^{d} \frac{1}{2}\nabla^2 F^1(X(s,x))(\sigma_i(X(s,x)), \sigma_i(X(s,x)))ds.$$

So we see that

$$|F^1(X(t,x)) - F^1(X(s,x))| \leq C_1|t-s|, \qquad t \in [s, \tau(s,x)), s \geq 0, x \in \mathbf{R}^N,$$

and

$$|F^2(X(t,x)) - F^2(X(s,x))| \leq C_1 Z_{T,m}|t-s|^{1/3} \qquad t,s \in [0,T], x \in V(m).$$

Also, by the assumption (F2), we see that there is a constant $C_2 > 0$ such that

$$\int_D 1_A(F(x))dx \le C_2|A|$$

for any Borel set A in $\mathbf{R}^2$, where $|A|$ denotes the area of A.

Let $\Delta_{\ell,n,k} = [-C_1 n^{-1}, C_1 n^{-1}] \times [-\ell C_1 n^{-1/3}, \ell C_1 n^{-1/3}]$, $\ell, n \ge 1$, $k = 1, \dots, n$. Then we have for any $\ell \ge 1$,

$$\begin{aligned}
&\int_{V(m)} dx P(F(X(t,x)) = 0 \text{ for some } t \in [s, \tau(s,x)), Z_{s+1,m} \le \ell) \\
\le &\sum_{k=1}^{n} \int_{V(m)} dx\ P(X(s,x) \in D, X(s+(k-1)/n, x) \in D, \\
&\qquad\qquad\qquad\qquad F(X(s+(k-1)/n, x)) \in \Delta_{\ell,n,k}) \\
= &\sum_{k=1}^{n} E[\int_{\mathbf{R}^N} dx 1_{V(m)}(x) 1_D(X(s,x)) \\
&\qquad 1_D(X(s+(k-1)/n, x)) 1_{\Delta_{\ell,n,k}}(F(X(s+(k-1)/n, x)))] \\
\le &C \sum_{k=1}^{n} |V(m)|^{1/10} |D|^{1/10} (\int_D 1_{\Delta_{\ell,n,k}}(F(x))dx)^{4/5} \\
\le &C C_2 n |V(m)|^{1/10} |D|^{1/10} (4C_1^2 \ell n^{-4/3})^{4/5}.
\end{aligned}$$

Here C is the constant in Lemma 3 for $T = s+1$, $p_0 = p_1 = 10$ and $p_3 = 5/4$. Since $n \ge 1$ is arbitrary, we see that

$$\int_{V(m)} dx P(F(X(t,x)) = 0 \text{ for some } t \in [s,, \tau(s,x)), Z_{s+1,m} \le \ell) = 0, \qquad \ell \ge 1.$$

This implies that $\int_{\mathbf{R}^N} p(x,s) = 0$, $s > 0$.

Therefore we have our assertion. ∎

Corollary 5. *Suppose moreover that* $x_0 \in (\bar{D})^c$, σ_i, $i = 0, \dots, d$, *are smooth around* x_0 *and that* $\dim Lie[\frac{\partial}{\partial t} - V_0, V_1, \dots, V_d](0, x_0) = N+1$. *Here*

$$V_i(x) = \sum_{j=1}^{d} \sigma_i^j(x) \frac{\partial}{\partial x^j}, \qquad i = 1, \dots, d,$$

and

$$V_0(x) = \sum_{j=1}^{d} (\sigma_0^j(x) - \frac{1}{2} \sum_{i=1}^{d} \sum_{k=1}^{N} \sigma_i^k(x) \frac{\partial \sigma_i^j}{\partial x^k}(x)) \frac{\partial}{\partial x^j}.$$

Then

$$P(X(t, x_0) \in D, F(X(t, x_0)) = 0 \textit{ for some } t > 0) = 0.$$

Proof. Let U be an open neighborhood of x_0 such that σ_i, $i = 0, \ldots, d$, are smooth around $\bar{U}$ and that $\bar{U} \cap \bar{D} = \emptyset$. Let $\tau = \inf\{t > 0; X(t, x_0) \in U^c\}$. Then we see that

$$\begin{aligned}
&P(X(t, x_0) \in D, F(X(t, x_0)) = 0 \text{ for some } t > 0) \\
\le &\sum_{n=1}^{\infty} P(X(t, x_0) \in D, F(X(t, x_0)) = 0 \text{ for some } t > \frac{1}{n}, \tau > \frac{1}{n}) \\
\le &\sum_{n=1}^{\infty} \int_U P(X(\frac{1}{n}, x_0) \in dx, \tau > \frac{1}{n}) P(X(t, x) \in D, F(X(t, x)) = 0 \\
&\hspace{20em} \text{for some } t > 0).
\end{aligned}$$

However, by [3], we see that $P(X(\frac{1}{n}, x_0) \in dx, \tau > \frac{1}{n})$ is absolutely continuous. So by Lemma 4, we have our assertion. ∎

§3. Proof of Theorem 1

Since the proof is similar, we prove Theorem 1 in the case that $D = \{x = (x^1, \ldots, x^d) \in \mathbf{R}^d;\ x^1 < 0\} \subset \mathbf{R}^d$, and $\Phi(v, x) = (-v^1, v^2, \ldots, v^d)$ for $v = (v^1, v^2, \ldots, v^d)$ and $x \in \partial D$. In general, if we take a double cover of D^c and change the coordinate functions, we can apply a similar proof. Let $a^{ij} : \mathbf{R}^d \to \mathbf{R}$, $i, j = 1, \ldots d$, be bounded Lipschitz continuous function such that $a^{ij}(x) = a^{ji}(x)$, $x \in \mathbf{R}^d$ and that there are positive constants C_0, C_1 such that

$$C_0 |\xi|^2 \le \sum_{i,j}^{d} a^{ij}(x) \xi_i \xi_j \le C_1 |\xi|^2, \qquad x, \xi \in \mathbf{R}^d.$$

Let $b : \mathbf{R}^{2d} \to \mathbf{R}^d$ be a bounded measurable function.

Let L_0 be a second order linear differential operator in $\mathbf{R}^{2d}$ given by

$$L_0 = \sum_{i=1}^{d} v^i \frac{\partial}{\partial x^i} + \frac{1}{2} \sum_{i,j=1}^{d} a^{ij}(x) \frac{\partial^2}{\partial v^i \partial v^j} + \sum_{i=1}^{d} b^i(x, v) \frac{\partial}{\partial v^i}$$

Then Theorem 1 is somehow equivalent to the following Theorem. So we prove this Theorem.

Theorem 6. *Let* $(x_0, v_0) \in (\bar{D})^c \times \mathbf{R}^d$, *and suppose that* a^{ij}, $i, j = 1, \ldots, d$, *are smooth around* x_0. *Then there exists a unique probability measure* μ *over* $\tilde{W}^d$ *satisfying the following conditions.*
(1) $\mu(w(0) = (x_0, v_0)) = 1$.
(2) $\mu(w(t) \in D^c \times \mathbf{R}^d, t \in [0, \infty)) = 1$.
(3) *For any* $f \in C_0^\infty((\bar{D})^c \times \mathbf{R}^d)$, $\{f(w(t)) - \int_0^t L_0 f(w(s))ds; t \geq 0\}$ *is a martingale under* $\mu(dw)$.
(4) $\mu(1_{\{0\}}(x^1(t))(v^1(t) + v^1(t-)) = 0,\ t \in [0, \infty)) = 1$ *and*

$$\mu(v^i(t) \textit{ is continuous in } t \in [0, \infty),\ i = 2, \ldots, d) = 1.$$

Proof. Let $\tilde{a}^{ij} : \mathbf{R}^d \to \mathbf{R}$, $i, j = 1, \ldots d$, be given by

$$\tilde{a}^{ij}(x) = a^{ij}(|x^1|, x^2, \ldots, x^d), \qquad x = (x^1, x^2, \ldots, x^d) \in \mathbf{R}^d.$$

Let $\tilde{b}^i : \mathbf{R}^{2d} \to \mathbf{R}$, $i = 1, \ldots d$, be given by

$$\tilde{b}^1(x) = sgn(x^1) b^1(|x^1|, x^2, \ldots, x^d),$$

and

$$\tilde{b}^i(x) = b^i(|x^1|, x^2, \ldots, x^d),\ i = 2, \ldots, d$$

for $x = (x^1, x^2, \ldots, x^d) \in \mathbf{R}^d$. Let $\tilde{L}_0$ be second order linear differential operators in $\mathbf{R}^{2d}$ given by

$$\tilde{L}_0 = \sum_{i=1}^d v^i \frac{\partial}{\partial x^i} + \frac{1}{2} \sum_{i,j=1}^d \tilde{a}^{ij}(x) \frac{\partial^2}{\partial v^i \partial v^j} + \sum_{i=1}^d \tilde{b}^i(x, v) \frac{\partial}{\partial v^i}.$$

Then by transformation of drift (cf. Ikeda-Watanabe[1]), we see that there is a unique probability measure ν on $C([0, \infty); \mathbf{R}^{2d})$ such that $\nu(w(0) = (x_0, v_0)) = 1$ and that $\{f(w(t)) - \int_0^t \tilde{L}_0 f(w(s))ds; t \geq 0\}$ is a martingale under $\nu(dw)$ for any $f \in C_0^\infty(\mathbf{R}^{2d})$.

Let $\tilde{\xi}(w) = \inf\{t > 0;\ x^1(t) = 0, v^1(t-) = 0\}$. Then by Corollary 5 and Girsanov's transformation, we see that $\nu(\tilde{\xi}(w) = \infty) = 1$. Let

$$X(t, w) = (|x^1(t)|, x^2(t), \ldots, x^d(t)), \qquad t \in [0, \infty),$$

and

$$V(t, w) = \frac{d^+}{dt} X(t, w), \qquad t \in [0, \infty).$$

Let μ is the probability law of $(X(\cdot, w), V(\cdot, w))$ under ν. Then we see that μ satisfies the conditions (1)-(4). So we see the existence.

Now let us prove the uniqueness. Let μ be a probability measure as in Theorem. Let $\xi(w) = \inf\{t > 0;\ x^1(t) = 0, v^1(t-) = 0\}$. Also, let us define stopping times $\tau_k : \tilde{W}_0 \to [0,\infty]$, $k = 0,1,2,\ldots$, inductively by $\tau_0(w) = 0$ and

$$\tau_{k+1}(w) = \inf\{t > \tau_k(w);\ x^1(t) = 0\}, \qquad w \in \tilde{W}^d,\ k = 0,1,\ldots.$$

Then we see from the assumption (4) that if $\tau_k(w) < \xi(w)$, then $\tau_k(w) < \tau_{k+1}(w)$ for μ-a.s.w. Also, it is easy to see that $\xi(w) \le \sup_k \tau_k(w)$, $w \in \tilde{W}^d$.

For any $\varepsilon > 0$ and $k = 0,1,2,\ldots$, let

$$\sigma_k^0(w) = \inf\{t > \tau_k(w);\ x^1(t) > \varepsilon\},$$

and

$$\sigma_k^1(w) = \inf\{t > \sigma_k^0(w);\ x^1(t) < \varepsilon/2\}, \qquad w \in \tilde{W}^d,\ k = 0,1,\ldots.$$

Then we see from the assumption (3) that

$$f(x(t\wedge\sigma_k^1), v(t\wedge\sigma_k^1)) - f(x(t\wedge\sigma_k^0), v(t\wedge\sigma_k^0)) - \int_{t\wedge\sigma_k^0}^{t\wedge\sigma_k^1} L_0 f(x(s), v(s))ds$$

is a bounded continuous martingale for any $f \in C_0^\infty(\mathbf{R}^{2d})$.

Now let

$$\tilde{X}(t,w) = \begin{cases} x(t), & t \in [\tau_k(w), \tau_{k+1}(w)), \text{ if } k \text{ is even}, \\ (-x^1(t), x^2(t), \ldots, x^d(t)), & t \in [\tau_k(w), \tau_{k+1}(w)), \text{ if } k \text{ is odd}, \end{cases}$$

$$\tilde{V}(t,w) = \begin{cases} v(t), & t \in [\tau_k(w), \tau_{k+1}(w)), \text{ if } k \text{ is even}, \\ (-v^1(t), v^2(t), \ldots, v^d(t)), & t \in [\tau_k(w), \tau_{k+1}(w)), \text{ if } k \text{ is odd}. \end{cases}$$

Then we can see that $(\tilde{X}(t\wedge\xi), \tilde{V}(t\wedge\xi))$ is continuous in t for μ-a.s.w. Also, we see that

$$f(\tilde{X}(t\wedge\sigma_k^1), \tilde{V}(t\wedge\sigma_k^1)) - f(\tilde{X}(t\wedge\sigma_k^0), \tilde{V}(t\wedge\sigma_k^0)) - \int_{t\wedge\sigma_k^0}^{t\wedge\sigma_k^1} \tilde{L}_0 f(\tilde{X}(s), \tilde{V}(s))ds$$

is a continuous martingale for any $f \in C_0^\infty(\mathbf{R}^{2d})$.

Therefore we see that

$$f(\tilde{X}(t\wedge\tau_{k+1}), \tilde{V}(t\wedge\tau_{k+1})) - f(\tilde{X}(t\wedge\tau_k), \tilde{V}(t\wedge\tau_k)) - \int_{t\wedge\tau_k}^{t\wedge\tau_{k+1}} \tilde{L}_0 f(\tilde{X}(s), \tilde{V}(s))ds$$

is a continuous martingale for any $f \in C_0^\infty(\mathbf{R}^{2d})$. So we can conclude that

$$f(\tilde{X}(t \wedge \xi), \tilde{V}(t \wedge \xi)) - \int_0^{t\wedge\xi} \tilde{L}_0 f(\tilde{X}(s), \tilde{V}(s))ds$$

is a continuous martingale for any $f \in C_0^\infty(\mathbf{R}^{2d})$.

Therefore we see that the probability law of $(\tilde{X}(\cdot \wedge \xi), \tilde{V}(\cdot \wedge \xi))$ under μ is the same of $w(\cdot \wedge \tilde{\xi})$ under ν, by the argument of shift of drift and the fact that a strong solution of stochastic differential equation with Lipschitz continuous coefficients is unique. So we see that $\mu(\xi(w) = \infty) = 1$. Since we see that

$$x(t) = (|\tilde{X}^1(t)|, \tilde{X}^2(t), \ldots, \tilde{X}^d(t)), \qquad t \in [0, \xi),$$

and

$$v(t) = (\frac{d^+}{dt}|\tilde{X}^1(t)|, \tilde{V}^2(t), \ldots, \tilde{V}^d(t)), \qquad t \in [0, \xi),$$

we see the uniqueness.

This completes the proof.

§4. Proof of Theorem 2

We will make some preparations to prove Theorem 2.

Proposition 7. *Let $T > 0$. Let A_0 be the set of $w \in D([0,T); \mathbf{R})$ for which $w(0) = 0$, $w(T-) \leq 1$, and $w(t)$ is non-decreasing in t. Then A_0 is compact in $L^p((0,T), dt)$, $p \in (1, \infty)$, and its cluster points are in $D([0,T); \mathbf{R})$.*

Proof. Suppose that $w_n \in A_0$, $n = 1, 2, \ldots$. Then we see that $w_n(t) \in [0,1]$, $t \in [0,T)$, $n \geq 1$. So taking subsequence if necessary, we may assume that $\{w_n(r)\}_{n=1}^\infty$ is convergent for any $r \in [0,T) \cap \mathbf{Q}$. Let $\tilde{w}(r) = \lim_{n\to\infty} w_n(r)$, $r \in \mathbf{Q}$, and let $w(t) = \lim_{r \downarrow t} \tilde{w}(r)$, $t \in [0,T)$, and $w(T)$ be arbitrary such that $\sup_{t\in[0,T)} w(t) \leq w(T) \leq 1$. Then we see that $w \in D([0,T); \mathbf{R})$ and w is non-decreasing, and that $w_n(t) \to w(t)$, $t \in [0,T)$, if t is a continuous point of w. So we see that $w_n \to w$, $n \to \infty$, in $L^p((0,T), dt)$.

This completes the proof. ∎

We have the following as an easy consequence of Proposition 7.

Corollary 8. *Let $T > 0$. Let A be the set of $w \in D([0,T); \mathbf{R}^d)$ for which $w(0) = 0$ and the total variation of w is less than 1. Then A is compact in $L^p((0,T); \mathbf{R}^d, dt)$, $p \in (1, \infty)$, and its cluster points are in $D([0,T); \mathbf{R}^d)$.*

Now let us prove Theorem 2. Let

$$H_t^\lambda = \lambda U(X_t^\lambda) + \frac{1}{2}|V_t^\lambda|^2, \qquad t \geq 0.$$

Then we have

$$H_t^\lambda = \frac{1}{2}|v_0|^2 + \int_0^t V_s^\lambda \cdot \sigma(X_s^\lambda)dB_s + \int_0^t V_s^\lambda \cdot b(X_s^\lambda, V_s^\lambda)ds + \frac{1}{2}\int_0^t trace(\sigma(X_s^\lambda)^*\sigma(X_s^\lambda))ds.$$

So we see that for any $p \in [2, \infty)$ there is a constant C independent of λ such that

$$E[\sup_{t\in[0,T]} (H_t^\lambda)^p] \leq C(|v_0|^{2p} + 1 + E[\int_0^T |V_t^\lambda|^p dt])$$
$$\leq C(|v_0|^{2p} + 1 + 2^{p/2}TE[\sup_{t\in[0,T]} (H_t^\lambda)^p]^{1/2}).$$

So we see that

$$\sup_{\lambda>0} E[\sup_{t\in[0,T]} (H_t^\lambda)^p] < \infty, \qquad p \in [1, \infty). \tag{1}$$

Therefore we see that

$$\sup_{\lambda>0} E[\sup_{t\in[0,T]} |V_t^\lambda|^p] < \infty, \qquad p \in [1, \infty).$$

So we see that $\{H_t^\lambda\}_{t\in[0,\infty)}$, and $\{X_t^\lambda\}_{t\in[0,\infty)}$, $\lambda \geq 0$, are tight in C. Moreover, we see that

$$E[\sup_{t\in[0,T]} U(X_t^\lambda)^p] \to 0, \quad \lambda \to \infty, \qquad p \in [1, \infty). \tag{2}$$

Let us take an $\varepsilon \in (0, \varepsilon_0)$ such that

$$C_0 = \sup\{|\nabla U_0(x)|^{-1};\ dis(x, \partial D) \leq \varepsilon\} < \infty.$$

Let $\varphi \in C_0^\infty(\mathbf{R}^d)$, such that $0 \leq \varphi \leq 1$, $\varphi(x) = 1$, if $dis(x, \partial D) < \varepsilon/3$, and $\varphi(x) = 0$, if $dis(x; \partial D) > \varepsilon/2$. Let $D_0 = \{x \in D;\ dis(x, \partial D) > \varepsilon/4\}$, and let $\tau = \tau^\lambda = \inf\{t > 0;\ X_t^\lambda \in D_0\}$. Then we see by Equation (2) that

$$P(\tau^\lambda < T) \to 0, \quad \lambda \to \infty,$$

for any $T > 0$. Let A_t^λ, $t \geq 0$ be a non-decreasing continuous process given by

$$A_t^\lambda = -\lambda \int_0^{t\wedge\tau^\lambda} \varphi(X_s^\lambda)\rho'(U_0(X_s^\lambda))|\nabla U_0(X_s^\lambda)|^2 ds, \qquad t \geq 0.$$

Note that $A_0^\lambda = 0$. Since we have

$$\begin{aligned}
&\varphi(X_{t\wedge\tau^\lambda}^\lambda)(\nabla U_0(X_{t\wedge\tau^\lambda}^\lambda)\cdot V_{t\wedge\tau^\lambda}^\lambda) - \varphi(X_0^\lambda)(\nabla U_0(X_0^\lambda)\cdot V_0^\lambda)\\
&= A_t^\lambda + \int_0^{t\wedge\tau^\lambda} \varphi(X_s^\lambda)\nabla^2 U_0(X_s^\lambda)(V_s^\lambda, V_s^\lambda)ds\\
&\quad + \int_0^{t\wedge\tau^\lambda} \varphi(X_s^\lambda)(\nabla U_0(X_s^\lambda)\cdot b(X_s^\lambda, V_s^\lambda))ds\\
&\quad + \int_0^{t\wedge\tau^\lambda} \varphi(X_s^\lambda)(\nabla U_0(X_s^\lambda))^*\sigma(X_s^\lambda)dB_s\\
&\quad + \int_0^{t\wedge\tau^\lambda} (\nabla\varphi(X_s^\lambda)\cdot V_s^\lambda)(\nabla U_0(X_s^\lambda)\cdot V_s^\lambda)ds,
\end{aligned}$$

we see that

$$\sup_{\lambda>0} E[(A_T^\lambda)^p] < \infty, \qquad p \in [1,\infty).$$

Since we have

$$\int_0^{T\wedge\tau^\lambda} \lambda U(X_t^\lambda)dt = \int_0^{T\wedge\tau^\lambda} \frac{\rho(U_0(X_t^\lambda))}{|\rho'(U_0(X_t^\lambda))|}|\nabla U_0(X_t^\lambda)|^{-2}dA_t^\lambda,$$

we see that

$$\begin{aligned}
&P(\int_0^{T\wedge\tau^\lambda} \lambda U(X_t^\lambda)dt > \delta)\\
&\leq P(\sup_{t\in[0,T]} U(X_t^\lambda) > \eta) + P(C_0^2 A_T^\lambda \sup_{\rho^{-1}(\eta)\leq s<0} \frac{\rho(s)}{|\rho'(s)|} > \delta)
\end{aligned}$$

for any $\delta, \eta > 0$. So we see that

$$P(\int_0^{T\wedge\tau^\lambda} |H_t^\lambda - \frac{1}{2}|V_t^\lambda|^2|dt > \delta) \to 0, \quad \lambda\to\infty \tag{3}$$

for any $\delta > 0$.

Also, we see that

$$V_{t\wedge\tau^\lambda}^\lambda = v_0 + V_t^{\lambda,0} + V_t^{\lambda,1},$$

where

$$V_t^{\lambda,0} = + \int_0^{t\wedge\tau^\lambda} |\nabla U_0(X_s^\lambda))|^{-2} \nabla U_0(X_s^\lambda) dA_s^\lambda,$$

and

$$V_t^{\lambda,1} = \int_0^{t\wedge\tau^\lambda} \sigma(X_s^\lambda) dB_s + \int_0^{t\wedge\tau} b(X_s^\lambda, V_s^\lambda) ds.$$

So we see that the total variation of $V_t^{\lambda,0}$, $t \in [0,T]$, is dominated by $C_0 A_T^\lambda$. Also, $\{V_t^{\lambda,0}\}_{t\in[0,\infty)}$ is tight in C.

Then by Corollary 8 it is easy to see that $\{V_t^\lambda\}_{t\in[0,T)}$ is tight in $L^p((0,T);\mathbf{R}^d)$ and its limit process is in $D([0,T);\mathbf{R}^d)$ with probability one for any $T>0$ and $p \in (1,\infty)$.

Let $F \in C^\infty(\mathbf{R}^d \times \mathbf{R}^d; \mathbf{R}^d)$ be given by

$$F(x,v) = \varphi(x)(v - |\nabla U_0(x)|^{-2}(\nabla U_0(x)\cdot v)\nabla U_0(x)), \qquad (x,v) \in \mathbf{R}^d \times \mathbf{R}^d.$$

Then by Itô's lemma it is easy to see that $\{F(X_t^\lambda, V_t^\lambda)\}_{t\in[0,\infty)}$, $\lambda \in (0,\infty)$, is tight in C, and that $\{f(X_t^\lambda, V_t^\lambda) - \int_0^t L_0 f(X_s^\lambda, V_s^\lambda) ds\}$ is a continuous martingale for any $\lambda \in (0,\infty)$ and $f \in C_0^\infty((\bar{D})^c \times \mathbf{R}^d)$.

So we see that there are stochastic processes $\{(X_t, V_t)\}_{t\in[0,\infty)}$ and $\{H_t\}_{t\in[0,\infty)}$ and a subsequence $\{\lambda_n\}_{n=1}^\infty$, $\lambda_n \to \infty$, $n \to \infty$, such that $\{((X_t^{\lambda_n}, V_t^{\lambda_n}), H_t^{\lambda_n})\}_{t\in[0,\infty)}$ converges in law to $\{((X_t, V_t), H_t)\}_{t\in[0,\infty)}$ in $\tilde{W}^d \times C$ with respect the metric function $\tilde{dis} + dis_C$.

Then we see that $\{f(X_t, V_t) - \int_0^t L_0 f(X_s, V_s) ds\}_{t\in[0,\infty)}$ is a continuous martingale for any $f \in C_0^\infty((\bar{D})^c \times \mathbf{R}^d)$, and that $\{F(X_t, V_t)\}_{t\in[0,\infty)}$ is a continuous process. Also, we see by Equation (3) that

$$\int_0^T |H_t - \frac{1}{2}|V_t|^2| dt = 0 \qquad a.s.$$

for any $T>0$. So we see that $\{|V_t|^2\}_{t\in[0,\infty)}$ is a continuous process. Therefore we have

$$P(1_{\partial D}(X_t)(V_t - V_{t-} - 2(n(X_t)\cdot V_{t-})n(X_t)) = 0,\ t \in [0,\infty)) = 1.$$

So we see that the probability law of $\{(X_t, V_t)\}_{t\in[0,\infty)}$ in $\tilde{W}$ is μ in Theorem 1.

This complets the proof of Theorem 2

References

[1] Ikeda, S., and S. Watanabe, Stochstic Differential Equations and Diffusion processes, 2nd Edition, North-Holland/Kodansha, 1989.

[2] Kunita, H., Stochastic flows and stochastic differential equations, Cambridge University Press, 1990.

[3] Kusuoka, S., and D.W. Stroock, Applications of Malliavin Calculus II, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 32(1985), 1-76.

Graduate School of Mathematical Sciences
The University of Tokyo
Komaba 3-8-1, Meguro-ku
Tokyo 153-8914, Japan

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 247–255

Cubic Schrödinger: The Petit Canonical Ensemble

Henry P. McKean

§1. Introduction

This report describes some aspects of the Gibbsian petit canonical ensemble for the cubic Schrödinger equation in the space of functions of period 1, say. §2–5 (defocussing case) represent joint work with K. Vaninsky[1]. §6 is a brief report on the much more difficult focussing case. The original hope, that the petit ensemble might provide a picture of the typical solution, is far from being achieved.

1.1. Preliminaries[2]

The mechanical state is a pair QP of nice functions of period 1, moving according to the defocussing flow:

$$\begin{aligned} \frac{\partial Q}{\partial t} &= -\frac{\partial^2 P}{\partial x^2} + (Q^2+P^2)\,P &= \frac{\partial H_3}{\partial P} \\ \frac{\partial P}{\partial t} &= +\frac{\partial^2 Q}{\partial x^2} - (Q^2+P^2)\,Q &= -\frac{\partial H_3}{\partial Q} \end{aligned}$$

This is a Hamiltonian system, relative to the classical bracket in function space, with Hamiltonian

$$H_3 = \frac{1}{2}\int_0^1 \left[(Q')^2 + (P')^2\right] + \frac{1}{4}\int_0^1 (Q^2+P^2)\,.$$

It is integrable in the full technical sense of the word, having an infinite series of (commuting) constants of motion $H_1 = \frac{1}{2}\int_0^1 (Q^2+P^2)$, $H_2 = \int_0^1 Q'P$, H_3, and so on. The flow is integrated with the help of the Dirac equation

$$M' = \left[\begin{pmatrix} Q & P \\ P & -Q \end{pmatrix} + \frac{\lambda}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right] M$$

Received November 4, 2002.

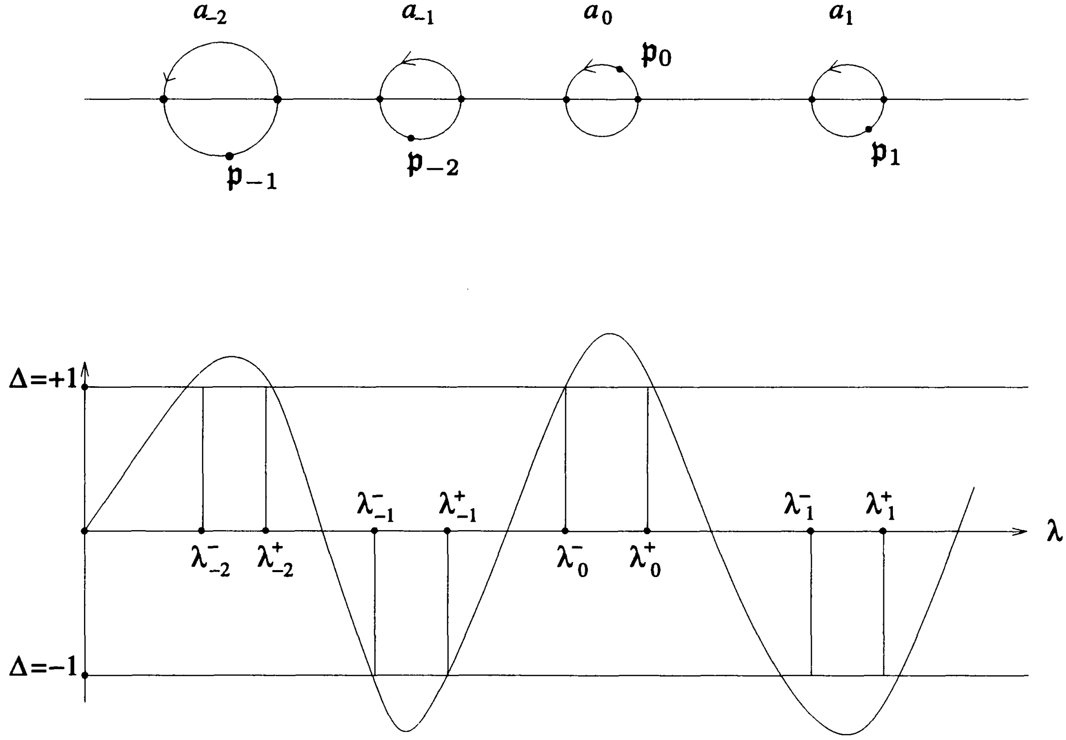
[1] McKean-Vaninsky [1997]

[2] Manakov et al. [1984] and/or McKean-Vaninsky [1997]

for the 2×2 monodromy matrix $M = [m_{ij} : 1 < i, j \leq 2]$ with $M(x = 0) = I$. Introduce the "discriminant" $\Delta(\lambda) = \frac{1}{2}\, sp\, M(x = 1)$ and the associated "Dirac curve" $\mathfrak{M}$ with points $\mathfrak{p} = \left[\lambda,\ \sqrt{\Delta^2(\lambda) - 1}\right]$. The latter is a double cover of the complex plane where λ lives, ramified over the roots

$$\ldots \lambda_{-1}^- \leq \lambda_{-1}^+ < \lambda_{-1}^- \leq \lambda_{-1}^+ < \lambda_0^- \leq \lambda_0^+ < \lambda_1^- \leq \lambda_1^+ < \ldots, \lambda_n^\pm \simeq 2\pi n \ etc.$$

of $\Delta(\lambda) = \pm 1$ indicated in the figure. These comprise the periodic/anti-periodic spectrum of the Dirac equation and may be interpreted as a



complete list of constants of motion, commuting among themselves and with the prior constants, H_1, H_2, H_3, etc. The cycles $a_n : n \in \mathbb{Z}$ seen in the upper part of the figure are the "real ovals" of $\mathfrak{M}$ covering the "gaps" $[\lambda_n^-, \lambda_n^+]$, these being all open for QP in general position, as is mostly assumed below. QP is encoded into a divisor $\mathfrak{P} = [\mathfrak{p}_n : n \in \mathbb{Z}]$ of $\mathfrak{M}$ having 1 point on each real oval: the numbers $\lambda(\mathfrak{p}_n) \equiv \mu_n \in [\lambda_n^-, \lambda_n^+]$ are the roots of $m_{12}(\mu) = 0$[3] and the radical $\sqrt{\Delta^2 - 1}\,(\mathfrak{p}_n)$ is declared to be $\frac{1}{2}\,(m_{11} - m_{12})\,(\mu_n)$[4]. The map $QP \to \mathfrak{P}$ is $1:1$ or to the product

[3] $m_{12}(\lambda)$ looks much like $-\sin(\lambda/2)$.

[4] $\det M(1) = 1$ so $m_{11}\, m_{12} = 1$ if $m_{12} = 0$ and $m_{11} + m_{12} = 2\Delta$ always, whence this possibility.

of all the ovals. The next actor in the play is the "Abel map" of the divisor into the (real) Jacobi variety Jac of $\mathfrak{M}$, determined as follows. DFK = the "differentials of the first kind" of $\mathfrak{M}$ are of the form $\omega = \phi_n(\lambda)\, d\lambda/\sqrt{\Delta^2(\lambda)-1}$ with certain entire functions ϕ, and a basis may be chosen so that $a_i(\omega_j) = 2\pi$ or 0 according as $i = j$ or not.[5] $\mathfrak{P}$ is now mapped to Jac via the "angles" $\theta_n = \sum_{k\in\mathbb{Z}} \int_{\mathfrak{o}_k}^{\mathfrak{p}_k} \omega_n$ construed *mod* 2π,[6] *i.e.*

$$\mathfrak{P} \to \Theta = [\theta_n : n \in \mathbb{Z}] \in (\mathbb{R}/2\pi\,\mathbb{Z})^\infty = \text{ Jac},$$

and this map likewise is 1 : 1 and onto. Now you have the composite map $QP \to$ divisor $\to$ Jac, the point of the whole exercise being that the (complicated) flow of QP is converted thereby into (simple) straight-line motion at constant speed in Jac which may be mapped back to the original (mechanical) variables with the help of a Riemann-like "theta" function. In this way, the flow is "integrated".

§2. Petit Ensemble at Levels 1 & 3

Level 1 is a warm-up for "level 3" to be described below. Introduce the "level 1 actions" $I_n = \frac{1}{4\pi}\, a_n \left(ch^{-1}\Delta,\, d\lambda\right)$[7] and note the trace formula $H_1 = \frac{1}{2}\int \left(Q^2 + P^2\right) = \sum_{\mathbb{Z}} I_n$. The petit ensemble[8]

$$\begin{aligned} e^{-H_1} d^\infty Q\, d^\infty P &= \frac{e^{-\frac{1}{2}\int_0^1 Q^2}}{(2\pi/0_+)^{\infty/2}} d^\infty\, Q \times \frac{e^{-\frac{1}{2}\int_0^1 P^2}}{(2\pi/0_+)^{\infty/2}}\, d^\infty\, P \\ &= \prod_{\mathbb{Z}} e^{-I_n}\, dI_n \times \prod_{\mathbb{Z}} d\,\theta_n/2\pi \;\; : \end{aligned}$$

is descriptive of 2 independent copies of white noise; line 2 comes from the trace formula plus the formal identification of the volume elements $d^\infty Q\, d^\infty P$ & $d^\infty I\, d^\infty\, \theta/2\pi$ prompted by the fact that actions & angles are canonically conjugate and together form a full coordinate system in QP-space. Naturally, line 2 requires proof as does the invariance of the ensemble under the flow, for which see McKean-Vaninsky [1997].

---

5) I should say differentials of the third kind as they have simple poles at the 2 points of $\mathfrak{M}$ covering ∞, but as they play the role of classical DFK, I keep the name. $\phi_n(\lambda)$ looks much like $m_{12}(\lambda)$ divided by $\lambda - \mu$, *i.e.* with 1 root left out.

6) $\mathfrak{o}_k = \left[\lambda_k^-, 0\right]$, some such choice being necessary for the convergence of the sum.

7) The name will be justified in §4.

8) Here and below, I will be free and easy with possibly infinite norming constants.

Level 3. The petit ensemble at "level 3":

$$e^{-H_3}\, d^\infty Q\, d^\infty P = \frac{e^{-\frac{1}{2}\int_0^1 (Q')^2}}{(2\pi 0_+)^{\infty/2}}\, d^\infty Q\, \frac{e^{-\frac{1}{2}\int_0^1 (P')^2}}{(2\pi 0_+)^{\infty/2}}\, d^\infty P \times e^{-\frac{1}{4}\int_0^1 (Q^2+P^2)^2}$$

is descriptive of 2 independent "circular" Brownian motions[9] coupled by the third factor; it is invariant under the flow as for level 1. To describe it in action/angle language requires a revision: DFK at level 3 is as before (level 1) but with a new basis $\omega'_n : n \in \mathbb{Z}$ normalized as in $a_i(\lambda^2 \omega'_j) = 2\pi$ or 0 according as $i = j$ or not. The level 3 actions are $I'_n = \frac{1}{q\pi}\, a_n(\lambda^2\, ch^{-1}\, \Delta\, d\lambda)$ and you have the trace formula $H_3 = \sum_{\mathbb{Z}} I'_n$, whence

$$\begin{aligned} e^{-H_3} d^\infty Q\, d^\infty P &= \prod_{\mathbb{Z}} e^{-I'_n} \times \left[d^\infty Q d^\infty P = d^\infty I\; d^\infty \frac{d\theta}{2\pi} \text{ at level } 1 \right] \\ &= \prod_{\mathbb{Z}} e^{-I'_n}\, dI'_n \prod_{\mathbb{Z}} d\,\frac{\theta_n}{2\pi} \times \det \frac{\partial I}{\partial I'}\,, \end{aligned}$$

in which the third (Jacobian) factor is still to be understood. The level 3 actions are canonically paired to the level 3 angles[10] $\theta'_n = \sum_{r\in\mathbb{Z}} \int_{\mathfrak{o}_k}^{\mathfrak{p}_k} \omega'_n$, so

$$\begin{aligned} \det \frac{\partial I}{\partial I'} &= \det \frac{\partial \theta'}{\partial \theta} \\ &= \frac{\det\, [\omega'_i / d\lambda(\mathfrak{p}_j)]}{\det\, [\omega_i / d\lambda(\mathfrak{p}_j)]} \\ &\times \int \frac{\det \prod_{i>j} (\mu_i - \mu_j)}{\prod_{\mathbb{Z}} \sqrt{\Delta^2 - 1(\mathfrak{p}_n)}}\, d^\infty \mu \\ &\text{divided by} \\ &\int \prod_{\mathbb{Z}} \mu^2 \times \text{ the same "volume element".} \end{aligned}$$

This rather fanciful expression comes from level 2 in case all but N gaps are closed and making $N \uparrow \infty$ with an (unpardonable) disregard of normalizing factors. Now the "volume element" seen in line 3 is

---

[9] CBM is standard Brownian motion, conditioned to end where it began, with this common displacement distributed over $\mathbb{R}$ by flat Lebesgue measure. The coupling holds down the total mass so that normalization is possible.

[10] These must be construed, not $mod\, 2\pi$, but relative to another, pretty complicated lattice of periods.

nothing but an un-normalized expression of the flat (level 1) volume element $d^\infty\theta/2\pi$ on Jac, written out in the language of the divisor; also $m_{12}(\lambda) = \frac{1}{2}(\mu_0 - \lambda)\prod_{\mathbb{Z}}(2\pi n)^{-1}(\mu_n - \lambda)$ precisely; and so it is an educated guess that, after proper normalization, the Jacobian det $\partial I/\partial I'$ ought to be the reciprocal of $N = \int_{Jac} m_{12}^2(0)\, d^\infty\,\theta/2\pi$.

This is correct as far as it goes[11], but what does N really look like? It is a function of actions alone, so the level 1 angles are still independent of them, with the same flat distribution as before. There are 10 integrals of products of 2 entries of $M(1)$, and I know 9 relations among them involving the constants of motion Δ and $\Delta^\bullet$, but the value of N is not revealed by these. Too bad! Crude estimates of N can be had but do not help to describe how the actions couple. I leave the subject in this unsatisfactory state.

§3. Some Tricks

I record here 3 amusing examples of averaging over Jac with respect to $d^\infty\theta/2\pi$, but first a general principle. Think of the (still to be normalized) expression

$$d^\infty\frac{\theta}{2\pi} = \prod_{i>j}(\mu_i - \mu_j)\, d^\infty\mu \text{ divided by } \prod_{\mathbb{Z}}\sqrt{\Delta^2 - 1}\,(\mathfrak{p}_n)$$

encountered in §3. The top, considered as a function of μ_n, say, is proportional to $m_{12}^\bullet(\mu_n)$, so you have the "splitting rule at $n \in \mathbb{Z}$":

$$d^\infty\frac{\theta}{2\pi} = \frac{m_{12}^\bullet(\mu_n)}{\sqrt{\Delta^2 - 1}\,(\mathfrak{p}_n)} \quad \text{on the oval } a_n$$

$\times$ a volume element on the product of all the other ovals.

This principle is now applied in 3 ways:

Example 1. $m_{12}(\lambda)$ looks like $-\sin(\lambda/2)$ and $\Delta(\lambda)$ like $\cos(\lambda/2)$, so you may expect $2\Delta^\bullet - m_{12}$ to be of "degree 1 lower" than m_{12} and that Lagrange interpolation would apply. This is correct:

$$2\Delta^\bullet(\lambda) - m_{12}(\lambda) = {}^{12)}\sum_{\mathbb{Z}}\frac{2\Delta^\bullet(\mu_n)}{m_{12}^\bullet(\mu_n)}\frac{m_{12}(\lambda)}{\lambda - \mu_n}\ .$$

---

[11] McKean-Vaninsky [1997].

[12] $m_{12}(\mu_n) = 0$ of course.

Now average over Jac, exchange sum and average, and split the volume at $n \in \mathbb{Z}$ to produce

$$2\Delta^\bullet(\lambda) - \int_{\mathrm{Jac}} m_{12}(\lambda)\,\frac{d^\infty\theta}{2\pi} = \sum_{\mathbb{Z}} 2 \int\limits_{\substack{\times a_k \\ k\neq n}} \frac{m_{12}(\lambda)}{\lambda-\mu_n} \int_{a_n} \frac{\Delta^\bullet d\mu_n}{\sqrt{\Delta^2-1}} \overset{13)}{=} 0,$$

i.e. $2\Delta^\bullet =$ average m_{12}.

Example 2. The numerator ϕ_n of $\omega_n \in$ DFK at level 1 looks like m_{12} with 1 root factored out, so it, too, should be capable of interpolation:

$$\phi_n(\lambda) = \sum_{i\in\mathbb{Z}} \frac{\phi(\mu_i)}{m_{12}^\bullet(\mu_i)}\,\frac{m_{12}^\bullet(\lambda)}{\lambda-\mu_i}\ .$$

But this object has nothing to do with angles, so an average over Jac does it no harm, and proceeding as in ex. 1, you find

$$\begin{aligned}\phi_n(\lambda) &= \sum_{i\in\mathbb{Z}} \int\limits_{\substack{\times a_j \\ j\neq i}} \frac{m_{12}(\lambda)}{\lambda-\mu_i} \int\limits_{a_i} \frac{\phi_n(\mu_i)\,d\mu_i}{\sqrt{\Delta^2-1}\,(\mathfrak{p}_i)} \\ &= \int\limits_{\substack{\times a_j \\ j\neq n}} \frac{m_{12}(\lambda)}{\lambda-\mu_n} \times 2\pi \\ &= \int\limits_{\times a_j = \text{ Jac.}} \frac{m_{12}(\lambda)}{m_{12}^\bullet(\mu_n)(\lambda-\mu_n)}\, d^\infty\,\frac{\theta}{2\pi}\end{aligned}$$

divided by

$$\frac{1}{2\pi}\int\limits_{a_n} \frac{d\mu}{\sqrt{\Delta^2-1}}\,,$$

i.e.

$$\omega_n = \quad \text{average} \quad \frac{m_{12}(\lambda)}{m_{12}^\bullet(\mu_n)(\lambda-\mu_n)} \quad \text{normalized to have mass } 2\pi \text{ on } a_n\ .$$

This seems to be a new way of writing DFK.

13) $\Delta^\bullet/\sqrt{\Delta^2-1} = d\,ch^{-1}\,\Delta$.

Example 3 identifies $I_n = \frac{1}{4\pi}\, a_n(ch^{-1}\,\Delta\, d\lambda)$ with a true mechanical action, as promised at start of §3.[14] The physical actions are $A_n = (2\pi)^{-1}a_n(PdQ) : n \in \mathbb{Z}$. To implement their evaluation, take the flow $e^{t\mathbb{X}}$ with Hamiltonian I_n which carries $\mathfrak{p}_n$ once about its private cycle a_n in time 2π, leaving the rest of the divisor fixed, and equate A_n with

$$\frac{1}{2\pi}\int_0^{2\pi} e^{t\mathbb{X}}\left[\int_0^1 P(x)\,\mathbb{X}\,Q(x)\,dx\right] dt.$$

Now A_n has nothing to do with angles, so you can average over Jac, exchange this average with the time-average, and use the invariance of the flat volume under the present flow and the flow of translation produced by $H_2 = \int_0^1 Q'P$ to reduce the previous display to $\int_{\mathrm{Jac}} P(0)\,\mathbb{X}\,Q(0)\,d^\infty\theta/2\pi$. Here,

$$\mathbb{X}\,Q(0) = \frac{1}{4\pi}\int_{a_n} (1/2)\,(m_{12}+m_{21})\,\frac{d\lambda}{\sqrt{\Delta^2-1}}\,,$$

$m_{12} - m_{21}$ is invariant under the "phase flow" $Q^\bullet = P$ and $P^\bullet = -Q$ produced by $H_1 = \frac{1}{2}\int_0^1 (Q^2+P^2)$, and the average of $P(0)$ under *this* flow is 0, permitting a further reduction to

$$A_n = \frac{1}{4\pi}\int_{a_n}\frac{d\lambda}{\sqrt{\Delta^2-1}}\int_{\mathrm{Jac}} P(0)\,[m_{12}(\lambda) - 2\Delta^\bullet(\lambda)]\,d^\infty\,\frac{\theta}{2\pi}\,.$$

The trace formula $P(0) = {}^{15)}\frac{1}{2}\sum(\mu_i - \lambda_i)$ and the interpolation of $2\Delta^\bullet - m_{12}$ from example 1 are now inserted under the average, sums and avarage are exchanged, and the volume element $d^\infty\theta/2\pi$ is split at

[14] The level 3 actions have also a mechanical interpretation, but I do not go into it here.

[15] $\lambda_n^\bullet : n \in \mathbb{Z}$ are the roots of $\Delta^\bullet(\lambda) = 0$.

$j \in \mathbb{Z}$, with the result that

$$
\begin{aligned}
A_n &= \frac{1}{4\pi}\int_{a_n} \frac{d\lambda}{\sqrt{\Delta^2-1}} \sum_{i\in\mathbb{Z}} \sum_{j\in\mathbb{Z}} \int_{\times a_k,\, k\neq j} \frac{m_{12}}{\lambda-\mu_j} \int_{a_j} (\lambda_i^\bullet - \mu_i)\, d\, ch^{-1}\,\Delta(\mathfrak{p}_j) \\
&= \frac{1}{4\pi}\int_{a_n} \frac{d\lambda}{\sqrt{\Delta^2-1}} \sum_{i\in\mathbb{Z}} \int_{\times a_k,\, k\neq i} \frac{m_{12}}{\lambda-\mu_i} \int_{a_n} ch^{-1}\,\Delta\, d\mu_i \\
&= \frac{1}{2\pi}\int_{a_n} \frac{d\lambda}{\sqrt{\Delta^2-1}} \sum_{i\in\mathbb{Z}} \left\{
\begin{array}{l}
\displaystyle\int_{\times a_k = \text{ Jac.}} \frac{m_{12}}{m_{12}^\bullet(\mu_i)(\lambda-\mu_i)}\, d^\infty \frac{\theta}{2\pi} \\
\text{divided by } \displaystyle\frac{1}{2\pi}\int_{a_i} \frac{d\mu_i}{\sqrt{\Delta^2-1}} \\
\text{multiplied by } \displaystyle\frac{1}{4\pi}\int_{a_i} ch^{-1}\,\Delta\, d\mu_i
\end{array}\right. \\
&= \frac{1}{2\pi}\int_{a_n} \frac{d\lambda}{\sqrt{\Delta^2-1}} \sum_{i\in\mathbb{Z}} \phi_i\, I_i \\
&= I_n,
\end{aligned}
$$

as advertised.

§4. Thermodynamic Limit

Now let Q & P have period L and take the large volume limit $L \uparrow \infty$. What happens to the petit ensemble $e^{-H_3}\, d^\infty Q\, d^\infty P$? The answer is nice and simple. Let ψ be the ground state of $-\frac{1}{2}\,\Delta + \frac{1}{4}\,r^4$ in $\mathbb{R}^2$. Then the mechanical variables $[Q(x), P(x)] : x \in \mathbb{R}$ tend (in law) to the stationary diffusion with infinitessimal operator $\frac{1}{2}\,\Delta + (\text{grad}\ \ell n\psi) \bullet \text{grad}$. This is even easy to prove.

§5. Focussing Case

This is much harder. The Hamiltonian is changed to $\frac{1}{2}\int \left[(Q')^2 + (P')^2\right]$ *minus* $\frac{1}{4}\int \left(Q^2 + P^2\right)$ and the associated petit ensemble has total mass $+\infty$. This prompted Lebowitz-Rose-Speer [1989] to introduce the micro-canonical ensemble obtained by conditioning upon the value N of the constant of motion $H_1 = \frac{1}{2}\int (Q^2 + P^2)$.[16] Their interest was in the thermodynamic limit: with fixed "density" D, "particle number" $N = DL$, and $L \uparrow \infty$, they found by numerical simulation, that the temperature dependent ensemble $e^{-H_3/T}\, d^\infty Q\, d^\infty P$ favors "solitons" / "radiation" at low/high temperatures, *i.e.* some kind of phase change takes place. Chorin [private communication] used a more sophisticated simulation of the Brownian motion and found the opposite: no phase change. This made me curious and, subsequently[17], I claimed to prove that the thermodynamical limit does not exist, explaining (as I thought) the discrepancy just described. But alas, all the big boys were wrong: in fact, my student B. Rider[18] proved that, at any values of temperature and density, the whole ensemble collapses onto $Q \equiv 0$ & $P \equiv 0$. A pity.

Bibliography

LEBOWITZ, J., H. A. ROSE, and E. R. SPEER: Statistical mechanics of the nonlinear Schrödinger equation(2). Mean field approximation. *J. Stat. Phys.* **54** (1989), 17–56.

MANKOV, S.V., S. NOVIKOV, L. P. PITAEVSKII, and V. E. ZAKAROV: *Theory of Solitons. The Inverse Scattering Method.* Contemporary Soviet Math., Consultants Bureau, New York-London, 1984.

MCKEAN, H. P. : Statistical mechanics of nonlinear wave equations (4): cubic Schrödinger. *Comm. Math. Phys.* **168** (1995), 479–491; erratum: **173** (1995) 675.

———, H. P. : A Martin boundary connected with the ∞-volume limit of the focussing cubic Schrödinger equation. *Itô's Stochastic Calculus and Probability Theory*, ed. Ikeda *et al.*, pp. 251–260, Springer-Verlag, Tokyo, 1995.

MCKEAN, H. P. and K. VANINSKY: Action-angle variables for the cubic Schrödinger equation. *CPAM* **50** (1997), 489–562.

——— : Cubic Schrödinger: the petit canonical ensemble in action-angle variables. *CPAM* **50** (1997), 593–622.

RIDER, B. : On the ∞-volume limit of the focussing cubic Schrödinger equation. *CPAM* **55** (2002), 1231–1248.

CIMS, 251 Mercer Street, New York, NY 10012

---

[16] The total mass is now finite, so the ensemble can be normalized. For a probabilistic proof of the existence of the flow and for the invariance of the micro-canonical ensemble under it, see McKean [1995].

[17] McKean [1995]

[18] Rider [2002]

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 257–278

Risk-sensitive Portfolio Optimization with Full and Partial Information

Hideo Nagai

Abstract.

We discuss an application of risk-sensitive control to portfolio optimization problems for a general factor model, which is considered a variation of Merton's intertemporal capital asset pricing model ([18]). In the model the instantaneous mean returns as well as volatilities of the security prices are affected by economic factors and the security prices. The economic factors are assumed to satisfy stocahstic differential equations whose coefficients depend on the security prices as well as themselves. In such general incomplete market models under Markovian setting we consider constructing optimal strategies for risk-sensitive portfolio optimization problems on a finite time horizon. We study the Bellman equations of parabolic type corresponding to the optimization problems. Through analysis of the Bellman equations we construct optimal strategies from the solution of the equation. We further discuss the problem with partial information. We shall obtain a necessary condition for optimality using backward stochastic partial differential equations.

§1. Introduction

Let us consider a market model with $m+1$ securities $(S_t^0, S_t) := (S_t^0, S_t^1, ..., S_t^m)^*$ and n factors $X_t = (X_t^1, X_t^2, ..., X_t^n)^*$. Here S^* stands for transposed matrix of S. We assume that the set of securities includes one bond, whose price is defined by the ordinary differential equation:

$$dS^0(t) = r(X_t, S_t)S^0(t)dt, \quad S^0(0) = s^0, \tag{1.1}$$

where $r(x, s)$ is a nonnegative function on R^{n+m}. The other secutity prices S_t^i, $i = 1, 2, ..., m$, and the factors X_t are assumed to satsfy the following stochastic differential equations:

Received February 28, 2003.

$$dS^i(t) = S^i(t)\{g^i(X_t, S_t)dt + \textstyle\sum_{k=1}^{n+m} \sigma^i_k(X_t, S_t)dW^k_t\},$$
$$S^i(0) = s^i, \quad i = 1, ..., m \tag{1.2}$$

and

$$dX_t = b(X_t, S_t)dt + \lambda(X_t, S_t)dW_t,$$
$$X(0) = x \in R^n, \tag{1.3}$$

where $W_t = (W^k_t)_{k=1,..,(n+m)}$ is an $m+n$ dimensional standard Brownian motion process defined on a filtered probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$. Here σ and λ are respectively $m \times (m+n)$, $n \times (m+n)$ matrix valued functions. Set

$$\mathcal{G}_t = \sigma(S(u), X(u); u \leq t)$$

and let us denote investment strategy to i-th security $S^i(t)$ by $h^i(t)$, $(i = 0, 1, ..., m)$ representing portfolio proportion of the amount of the i-th security to the total wealth V_t that the investor possesses, which is defined as follows:

Definition 1.1. $(h^0(t), h(t)) \equiv (h^0(t), (h^1(t), h^2(t), ..., , h^m(t))^*)$ *is said to be an invetment strategy if the following conditions are satisfied*

i) *$h(t)$ is an R^m valued $\mathcal{G}_t$ progressively measurable stochastic process such that*

$$\sum_{i=1}^{m} h^i(t) + h^0(t) = 1$$

ii) *and that*

$$P(\int_0^T |h(s)|^2 ds < \infty) = 1.$$

The set of all investment strategies will be denoted by $\mathcal{H}(T)$. When $(h^0(t), h(t)^*)_{0\leq t\leq T} \in \mathcal{H}(T)$ we will often write $h \in \mathcal{H}(T)$ for simplicity. In what follows we always assume that

$$\sigma\sigma^* > 0. \tag{1.4}$$

For given $h \in \mathcal{H}(T)$ the wealth process $V_t = V_t(h)$ satisfies

$$\begin{aligned}
\frac{dV_t}{V_t} &= \textstyle\sum_{i=0}^{m} h^i(t)\frac{dS^i(t)}{S^i(t)} \\
&= h^0(t)r(X_t, S_t)dt + \textstyle\sum_{i=1}^{m} h^i(t)\{g^i(X_t, S_t)dt \\
&\qquad\qquad + \textstyle\sum_{k=1}^{m+n} \sigma^i_k(X_t, S_t)dW^k_t\} \\
V_0 &= v
\end{aligned}$$

under the assumption of the self-financing condition. Then, taking i) above into account it turns out to be the solution of

$$\frac{dV_t}{V_t} = r(X_t, S_t)dt + h(t)^*(g(X_t, S_t) - r(X_t, S_t)\mathbf{1})dt + h(t)^*\sigma(X_t, S_t)dW_t,$$

$$V_0 = v,$$

where $\mathbf{1} = (1, 1, ..., 1)^*$.

We first consider the following problem. For a given constant $\mu < 1$, $\mu \neq 0$ maximize the following risk-sensitized expected growth rate up to time horizon T:

$$J(v, x; h; T) = \frac{1}{\mu} \log E[e^{\mu \log V_T(h)}] = \frac{1}{\mu} \log E[V_T(h)^\mu], \tag{1.5}$$

where h ranges over the set $\mathcal{A}(T)$ of all admissible strategies defined later. The meaning of the maximization is well understood by looking at the asymptotics of the criterion as $\mu \to 0$:

$$\frac{1}{\mu} \log E[e^{\mu \log V_T(h)}] \sim E[\log V_T(h)] + \frac{\mu}{2} \mathrm{Var}[\log V_T(h)] + O(\mu^2).$$

Maximizing (1.5) is a risk-sensitive counterpart of the problem maximizing the expected growth rate of the investor's wealth. The case where $\mu < 0$ is called risk averse and $\mu > 0$ risk seeking. Concerning this problem we introduce the Bellman equation corresponding to the value function and we present the results constructing an optimal strategy from the solution to the equation through its analysis in section 2. Note that the problem maximizing the criterion $J(v, x; h; T)$ is equivalent to HARA utility maximization:

$$\sup_h E[\frac{1}{\mu} V_T(h)^\mu] = \sup_h E[\frac{1}{\mu} e^{\mu \log V_T(h)}], \quad \mu < 1.$$

The problems on infinite time horizon maximizing

$$\liminf_{T \to \infty} \frac{1}{\mu T} \log E[e^{\mu \log V_T(h)}] \tag{1.6}$$

have been considered by several authors e.g. in [8],[9],[10],[14], in the case of linear Gaussian factor models since the work by Fleming [7], or under more general setting in [5] with the assumption that randomness of security price processes and that of factor processes are independent. In [23] we have discussed the problem under rather general setting for so called Merton's ICAPM ([18]), and the results in section 2 of the present paper are its genaralization in the case of finite time horizon. Here by Merton's ICAPM, we mean the case that

$$r(x,s) = r_1(x), \quad g(x,s) = g_1(x), \quad b(x,s) = b_1(x),$$
$$\sigma(x,s) = \sigma_1(x), \quad \lambda(x,s) = \lambda_1(x).$$

If r_1, g_1 and b_1 are linear functions and σ_1 and λ_1 are constant matrices the models are said to be linear Gaussian. In that case the solutions of the Bellman equations are expressed explicitly as the quadratic functions of x whose coefficients are determined as the solutions of the matrix Riccati differential equations and linear differential equations.

We then consider the maximization problem with partial infromation. In the above investment strategies are defined as $\mathcal{G}_t$ progressively measurable processes. However, it is not always realistic since economic factors X_t are to be considered implicit and so it might be better to select our strategies without using all past informations of securities S_t and factors X_t. Our strategies may be well selected by using only informations of security prices. Rishel [24] has considered the problem on a finite time horizon in such a way in a particular case, namely for a linear Gaussian model of one factor and one risky and one riskless assets under the assumtion that randomness of the factor process and that of the risky asset are independent. We have also considered the problem for general linear Gaussian factor models [21] on a finite time horizon and, by solving two kinds of Riccati differential equations, constructed an optimal strategy. The results are extended to the case of infinite time horizon in [22] by studying asymptotics of the solutions of inhomogeneous (time dependent) Riccati differential equations as time horizon goes to infinity. In the present paper we shall consider the maximization problem in section 3 under more general setting, namely the case where coefficients of security prices are nonlinearly depend on economic factors. In that case we don't have explicit expression of the optimal strategies but study necessity of optimality. We introduce backward stochastic partial differential equations (BSPDEs), which are considered to be adjoint equations of the problems, and find the necessary condition for optimality by using the solutions of the BSPDEs under suitable conditions. Such necessary condition is a kind of maximum principle and it has been studied by A. Bensoussan for stochastic control problems for partially observed diffusion processes (cf. [1],[2], [26]).

§2. Full information case

Let us set

$$Y_t^i = \log S_t^i, \quad i = 0, 1, 2, ..., m,$$

$Y_t = (Y_t^1, Y_t^2, ..., Y_t^m)^*$ and $\mathbf{e}^Y = (e^{Y^1}, ..., e^{Y^m})^*$. Then

$$dY_t^0 = r(X_t, \mathbf{e}^{Y_t})dt$$

and

$$dY_t = F(X_t, Y_t)dt + \Sigma(X_t, Y_t)dW_t, \tag{2.1}$$

where

$$\begin{aligned} F^i(x, y) &= g^i(x, \mathbf{e}^y) - \tfrac{1}{2}(\sigma\sigma^*)^{ii}(x, \mathbf{e}^y), \\ \Sigma_k^i(x, y) &= \sigma_k^i(x, \mathbf{e}^y). \end{aligned}$$

In the same way, set

$$B(x, y) = b(x, \mathbf{e}^y), \quad \Lambda(x, y) = \lambda(x, \mathbf{e}^y).$$

Then the factor prosess is described as

$$dX_t = B(X_t, Y_t)dt + \Lambda(X_t, Y_t)dW_t \tag{2.2}$$

So, by setting $Z_t = (X_t, Y_t)^*$ and

$$\beta(z) = (B(x, y), F(x, y))^*, \quad \alpha(z) = (\Lambda(x, y), \Sigma(x, y))^*,$$

we have

$$dZ_t = \beta(Z_t)dt + \alpha(Z_t)dW_t \tag{2.3}$$

Furthermore, by setting $\tilde{g}(z) = g(x, \mathbf{e}^y)$, $\tilde{r}(z) = r(x, \mathbf{e}^y)$ for simplicity we have

$$\frac{dV_t}{V_t} = \tilde{r}(Z_t)dt + h_t^*(\tilde{g}(Z_t) - \tilde{r}(Z_t)\mathbf{1})dt + h_t^*\Sigma(Z_t)dW_t$$

and so,

$$\begin{aligned} V_t^\mu = v^\mu \exp\{ -\mu \int_0^t \eta(Z_s, h_s)ds + \mu \int_0^t h_s^*\Sigma(Z_s)dW_s \\ -\frac{\mu^2}{2} \int_0^t h_s^*\Sigma\Sigma^*(Z_s)h_s ds\}, \end{aligned}$$

where

$$\eta(z, h) = \frac{1-\mu}{2} h^*\Sigma\Sigma^*(z)h - \tilde{r}(z) - h^*(\tilde{g}(z) - \tilde{r}(z)\mathbf{1})$$

If a given investment strategy h satisfies

$$(2.4)\qquad E[e^{\mu\int_0^T h_s^*\Sigma^*(Z_s)dW_s-\frac{\mu^2}{2}\int_0^T h_s^*\Sigma\Sigma^*(Z_s)h_s ds}]=1,$$

then we can introduce a probability measure P^h given by

$$P^h(A)=E[e^{\mu\int_0^T h_s^*\Sigma(Z_s)dW_s-\frac{\mu^2}{2}\int_0^T h_s^*\Sigma\Sigma^*(Z_s)h_s ds};A]$$

for $A\in\mathcal{F}_T$, $T>0$. By the probability measure P^h our criterion $J(v,x;h;T)$ can be written as follows:

$$(2.5)\qquad J(v,x;h,T)=\log v+\frac{1}{\mu}\log E^h[e^{-\mu\int_0^T\eta(Z_s,h_s)ds}].$$

On the other hand, under the probability measure

$$\begin{aligned} W_t^h &= W_t-\langle W_\cdot,\mu\textstyle\int_0^\cdot h^*(s)\Sigma(Z_s)dW_s\rangle_t \\ &= W_t-\mu\textstyle\int_0^t\Sigma^*(Z_s)h(s)ds \end{aligned}$$

is a standard Brownian motion process, and therefore the factor process X_t satisfies the following stochastic differential equation

$$(2.6)\qquad \begin{aligned} dX_s &= (B(X_s,Y_s)+\mu\Lambda\Sigma^*(X_s,Y_s)h_s)ds+\Lambda(X_s,Y_s)dW_s^h \\ dY_s &= (F(X_s,Y_s)+\mu\Sigma\Sigma^*(X_s,Y_s)h_s)ds+\Sigma(X_s,Y_s)dW_s^h. \end{aligned}$$

And so,

$$(2.7)\qquad dZ_t=\beta_\mu(Z_t,h_t)dt+\alpha(Z_t)dW_t^h,$$

where

$$\beta_\mu(z,h)=\beta(z)+\mu\alpha\Sigma^*(z)h.$$

We regard (2.7) as a stochastic differential equation controlled by h and the criterion function is written by P^h as follows:

$$(2.8)\qquad J(v,x;h;T-t)=\log v+\frac{1}{\mu}\log E^h[e^{-\mu\int_0^{T-t}\eta(Z_s,h(s))ds}]$$

and the value function

$$(2.9)\qquad u(t,z)=\sup_{h\in\mathcal{A}(T-t)} J(v,z;h;T-t),\ 0\le t\le T.$$

Here we denote by $\mathcal{A}(T)$ the set of all investment strategies satisfying (2.4). Then, according to Bellman's dynamic programming principle, it should satisfy the following Bellman equation

$$\begin{aligned} &\frac{\partial u}{\partial t} + \sup_{h\in R^m} L^h u = 0, \\ &u(T,z) = \log v, \end{aligned} \tag{2.10}$$

where L^h is defined by

$$L^h u(t,z) = \frac{1}{2}\mathrm{tr}(\alpha\alpha^*(z)D^2u) + \beta_\mu(z,h)Du + \frac{\mu}{2}(Du)^*\alpha\alpha^*(z)Du - \eta(z,h).$$

Note that $\sup_{h\in R^m} L^h u$ can be written as

$$\begin{aligned} \sup_{h\in R^m} L^h u(t,z) &= \frac{1}{2}\mathrm{tr}(\alpha\alpha^*(z)D^2u) + \beta(z)^*Du + \frac{\mu}{2}(Du)^*\alpha\alpha^*Du + \tilde{r} \\ &\quad + \sup_h\{\mu h^*\Sigma\alpha^*Du + h^*(\tilde{g} - \tilde{r}\mathbf{1}) - \tfrac{1-\mu}{2}h^*\Sigma\Sigma^*h\} \\ &= \tfrac{1}{2}\mathrm{tr}(\alpha\alpha^*(z)D^2u) + \beta(z)^*Du + \tfrac{\mu}{1-\mu}(\tilde{g} - \tilde{r}\mathbf{1})^*(\Sigma\Sigma^*)^{-1}\Sigma\alpha^*Du \\ &\quad + \tfrac{\mu}{2}(Du)^*\alpha(I + \tfrac{\mu}{1-\mu}\Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma)\alpha^*Du \\ &\quad + \tfrac{1}{2(1-\mu)}(\tilde{g} - \tilde{r}\mathbf{1})^*(\Sigma\Sigma^*)^{-1}(\tilde{g} - \tilde{r}\mathbf{1}) + \tilde{r} \end{aligned}$$

Therefore our Bellman equation (2.10) is written as follows:

$$\begin{aligned} &\tfrac{\partial u}{\partial t} + \tfrac{1}{2}\mathrm{tr}(\alpha\alpha^*D^2u) + \hat{\beta}_\mu^*Du + (Du)^*\alpha N^{-1}\alpha^*Du + U(z) = 0, \\ &u(T,z) = \log v, \end{aligned} \tag{2.11}$$

where

$$\begin{aligned} \hat{\beta}_\mu(z) &= \beta(z) + \tfrac{\mu}{1-\mu}\alpha\Sigma^*(\Sigma\Sigma^*)^{-1}(\tilde{g} - \tilde{r}\mathbf{1}) \\ N^{-1}(z) &= \tfrac{\mu}{2}(I + \tfrac{\mu}{1-\mu}\Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma(z)) \\ U(z) &= \tfrac{1}{2(1-\mu)}(\tilde{g} - \tilde{r}\mathbf{1})^*(\Sigma\Sigma^*)^{-1}(\tilde{g} - \tilde{r}\mathbf{1}) + \tilde{r}(z). \end{aligned} \tag{2.12}$$

As for (2.11) we note that if $\mu < 0$, then

$$\frac{\mu}{2(1-\mu)} I \leq N^{-1} \leq \frac{\mu}{2} I$$

and therefore we have

$$\frac{\mu}{2(1-\mu)} \alpha\alpha^* \leq \alpha N^{-1}\alpha^* \leq \frac{\mu}{2}\alpha\alpha^* < 0.$$

On the other hand if $0 < \mu < 1$, then

$$\frac{\mu}{2} I \leq N^{-1} \leq \frac{\mu}{2(1-\mu)} I$$

and therefore we have

$$0 < \frac{\mu}{2}\alpha\alpha^* \leq \alpha N^{-1}\alpha^* \leq \frac{\mu}{2(1-\mu)}\alpha\alpha^*.$$

In what follows we assume that

$$\text{(2.13)} \qquad \begin{aligned} &B, F, \Lambda, \Sigma \text{ are locally Lipshitz and that} \\ &\tfrac{1}{2} \parallel \Sigma\Sigma^* \parallel + \tfrac{1}{2} \parallel \Lambda\Lambda^* \parallel + \beta^* z \leq c(1+|z|^2), \end{aligned}$$

then we have a solution (X_t, Y_t) of (2.1) and (2.2), and so setting

$$S_t^i = e^{Y_t^i}, \quad i = 1, 2, ..., m, \quad S_t^0 = e^{Y_t^0} = e^{\log s^0 + \int_0^t r(X_s, e^{Y_s}) ds}$$

we have a market model (S_t^0, S_t) satisfying (1.2) and (1.3). Then we have the following theorem.

Theorem 2.1. *Let $u \in C^{1,2}([0,T) \times R^N)$ be a solution of (2.11). Define*

$$\hat{h}_t = \hat{h}(t, Z_t)$$
$$\hat{h}(t,z) = \tfrac{1}{1-\mu}(\Sigma\Sigma^*)^{-1}(\tilde{g} - \tilde{r}\mathbf{1} + \mu\Sigma\alpha^* Du)(t,z),$$

where Z_t is the solution of (2.3), then, under the assumption that

(2.14)

$$E[e^{-\int_0^T (2N^{-1}\alpha^* Du + 2\mu K)^* dW_s - \frac{1}{2}\int_0^T (2N^{-1}\alpha^* Du + 2\mu K)^*(2N^{-1}\alpha^* Du + 2\mu K) ds}] = 1,$$

with

$$K = \frac{1}{2(1-\mu)}\Sigma^*(\Sigma\Sigma^*)^{-1}(\tilde{g} - \tilde{r}\mathbf{1})$$

$\hat{h}_t \in \mathcal{A}_T$ *is an optimal strategy for the portfolio optimization problem of maximizing the criterion (1.5).*

The proof of this theorem is similar to that of Proposition 2.1 in [23] and we omit it here.

We then consider equation (2.11). Such kinds of equations have been studied in [20], or [3] in relation to risk-sensitive control problems under more general settings in the case of $\mu < 0$ and in [4] in the case where $\mu > 0$. Here we consider the case where $\mu < 0$ and obtain the following result along the line [3], Theorem 5.1 with refinement on estimate (2.16). It is a generalization of Theorem 2.1 in [23].

Theorem 2.2. *i) If, in addition to (2.13), $\mu < 0$ and*

$$\nu_r|\xi|^2 \leq \xi^*\alpha\alpha^*(z)\xi \leq \nu_r'|\xi|^2, \quad r = |z|, \quad \nu_r,\ \nu_r' > 0, \tag{2.15}$$

then we have a solution of (2.11) such that

$$u,\ \tfrac{\partial u}{\partial t},\ D_k u,\ D_{kj}u \in L^p(0,T; L^p_{loc}(R^{n+m})),$$

$$\tfrac{\partial^2 u}{\partial t^2},\ \tfrac{\partial D_k u}{\partial t},\ \tfrac{\partial D_{kj} u}{\partial t}, D_{kjl}u \in L^p(0,T; L^p_{loc}(R^{n+m})),$$

$$u \geq \log v,\ \ \tfrac{\partial u}{\partial t} \leq 0.$$

$$1 < \forall p < \infty$$

Furthermore we have the estimate

$$\begin{aligned} |\nabla u|^2(t,z) - \frac{c_0}{\nu_r}\frac{\partial u}{\partial t}(t,z) &\leq c_r(|\nabla Q|^2_{2r} + |Q|^2_{2r} \\ &+|\nabla(\alpha\alpha^*)|^2_{2r} + |\nabla\beta_\mu|_{2r} + |\beta_\mu|^2_{2r} \\ &+ |U|_{2r} + |\nabla U|^2_{2r} + 1), \qquad z \in B_r,\ \ t \in [0,T) \end{aligned} \tag{2.16}$$

where

$$Q = \alpha N^{-1}\alpha^*,\ \ c_0 = \tfrac{4(1+c)(1-\mu)}{-\mu},\ \ c > 0$$

$$|\cdot|_{2r} = \|\cdot\|_{L^\infty(B_{2r})}$$

and c_r is a positive constant depending on n, r, ν_r, ν'_r and c.

ii) If, in addition to the above conditions,

$$\inf_{|z|\geq r} U(z),\ r^2\frac{1}{\nu'_r}\inf_{|z|\geq r} U(z),\ r\inf_{|z|\geq r}\frac{U(z)}{|\beta_\mu(z)|} \to \infty$$

as $r \to \infty$, then the above solution u satisfies

$$\inf_{|z|\geq r, t\in(0,T)} u(z,t) \to \infty, \quad as\ \ r \to \infty.$$

Moreover, there exists at most one such solution in $L^\infty(0,T;W^{1,\infty}_{loc}(R^{n+m}))$

Remark. If

$$\frac{1}{\nu_r},\ \nu'_r \leq M(1+r^m), \quad \exists m > 0, \tag{2.17}$$

then we have

$$c_r \leq M'(1+r^{m'}),\ \exists m'$$

in estimete (2.16). In particular, if $m = 0$, then c_r can be taken independent of r.

Corollary 2.1. *Condition (2.14) is valid if*

$$\begin{aligned} &c_1|\xi|^2 \leq \xi^*\alpha\alpha^*(x)\xi \leq c_2|\xi|^2,\ \ c_1, c_2 > 0 \\ &B,\ F,\ \Lambda,\ \Sigma\ \ \textit{are globally Lipshitz.} \end{aligned} \tag{2.18}$$

The proofs of Theorem 2.2 and Corollary 2.1 are similar to those of Theorem 2.1 and Proposition 2.1 (ii) in [23] and we omit them here. Instead, we illustrate an example.

Example (Generalized linear Gaussian factor model)
Let us consider the case where B, $\tilde{g}$, and $\tilde{r}$ are all linear functions of z and Λ and Σ are constant matrices, namely

$$\beta(z) = \begin{pmatrix} B(x,y) \\ F(x,y) \end{pmatrix} = \begin{pmatrix} B_1x + B_2y + b \\ A_1x + A_2y + a - \frac{1}{2}\widehat{(\Sigma\Sigma^*)} \end{pmatrix}$$

$$\tilde{g}(x,y) = A_1x + A_2y + a, \quad \tilde{r}(x,y) = R_1^*x + R_2^*y + r,$$

where $\widehat{(\Sigma\Sigma^*)} = ((\Sigma\Sigma^*)^{ii}) \in R^m$. Then

$$\begin{aligned} \hat{\beta}_\mu(z) &= \beta(z) + \tfrac{\mu}{1-\mu}\begin{pmatrix} \Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}(\tilde{g} - r\mathbf{1}) \\ \tilde{g} - r\mathbf{1} \end{pmatrix} \\ &= K_1 z + L_1 \end{aligned}$$

where

$K_1 =$

$$\begin{pmatrix} B_1 + \tfrac{\mu}{1-\mu}\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}(A_1 - \mathbf{1}R_1^*) & B_2 + \tfrac{\mu}{1-\mu}\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}(A_2 - \mathbf{1}R_2^*) \\ \tfrac{1}{1-\mu}A_1 - \tfrac{\mu}{1-\mu}\mathbf{1}R_1^* & \tfrac{1}{1-\mu}A_2 - \tfrac{\mu}{1-\mu}\mathbf{1}R_2^* \end{pmatrix}$$

$$L_1 = \begin{pmatrix} b + \tfrac{\mu}{1-\mu}\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}(a - r\mathbf{1}) \\ a - \tfrac{1}{2}\widehat{(\Sigma\Sigma^*)} + \tfrac{\mu}{1-\mu}(a - r\mathbf{1}) \end{pmatrix}.$$

Furthermore

$$\alpha N^{-1}\alpha^* = \frac{\mu}{2}\begin{pmatrix} \Lambda(I + \tfrac{\mu}{1-\mu}\Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma)\Lambda^* & \tfrac{1}{1-\mu}\Lambda\Sigma^* \\ \tfrac{1}{1-\mu}\Sigma\Lambda^* & \tfrac{1}{1-\mu}\Sigma\Sigma^* \end{pmatrix} \equiv \frac{1}{2}K_0,$$

$$U(z) = \frac{1}{2}z^*K_2 z + L_2 z + r + \frac{1}{2(1-\mu)}(a - r\mathbf{1})^*(\Sigma\Sigma^*)^{-1}(a - r\mathbf{1}),$$

where

$K_2 =$

$$\frac{1}{1-\mu}\begin{pmatrix} (A_1 - \mathbf{1}R_1^*)^*(\Sigma\Sigma^*)^{-1}(A_1 - \mathbf{1}R_1^*) & (A_1 - \mathbf{1}R_1^*)^*(\Sigma\Sigma^*)^{-1}(A_2 - \mathbf{1}R_2^*) \\ (A_2 - \mathbf{1}R_2^*)^*(\Sigma\Sigma^*)^{-1}(A_1 - \mathbf{1}R_1^*) & (A_2 - \mathbf{1}R_2^*)^*(\Sigma\Sigma^*)^{-1}(A_2 - \mathbf{1}R_2^*) \end{pmatrix}$$

and

$$L_2 = \frac{1}{1-\mu}\begin{pmatrix} (A_1 - \mathbf{1}R_1^*)^*(\Sigma\Sigma^*)^{-1}(a - r\mathbf{1}) \\ (A_2 - \mathbf{1}R_2^*)^*(\Sigma\Sigma^*)^{-1}(a - r\mathbf{1}) \end{pmatrix}.$$

In this case the solution to (2.11) has an explicit form such that

$$u(t,z) = \frac{1}{2}z^*P(t)z + q(t)^*z + k(t),$$

provided that equation (2.19) below has a solution. Here $P(t)$, $q(t)$ and $k(t)$ are the solutions to the following ordinary differential equations:

(2.19) $\dot{P}(t) + K_1^*P(t) + P(t)K_1 + P(t)K_0P(t) + K_2 = 0, \quad P(T) = 0,$

(2.20)

$$\dot{q}(t) + K_1 q(t) + P(t)L_1 + P(t)K_0 q(t) + \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} + L_2 = 0, \quad q(T) = 0$$

and

(2.21)

$$\dot{k}(t) + \frac{1}{2}\mathrm{tr}(\alpha\alpha^* P(t)) + L_1^* q(t) + r + \frac{1}{2(1-\mu)}(a - r\mathbf{1})^*(\Sigma\Sigma^*)^{-1}(a - r\mathbf{1}) = 0,$$

$$k(T) = \log v.$$

Note that if $\mu < 0$, then (2.19) has a unique solution and so do (2.20) and (2.21).

§3. Partial information case

Now we consider a partial information case. Namely, the case where portfolio strategies are selected by using only past information of security prices. In this case the economic factor process X_t is considered unobservable and so we cannot use the information about it to choose our strategies. Thus the factor process X_t defined before by (2.2) may be reformulated as the solution with the initial condition $X_0 = x_0$, where x_0 is a random variable having a disrtibution density $\pi(x)$ on R^n. We then introduce

$$\tilde{\mathcal{G}}_t = \sigma(S(u); u \le t)$$

and the admissible strategies are assumed to be $\tilde{\mathcal{G}}_t$ measurable. In this case we consider more specific one than the above, namely we assume (2.18) and that $\sigma(x, S) = \sigma(S)$.

Then we consider the problem maximising the criterion (1.5) by selecting portfolio stratgies which are $\tilde{\mathcal{G}}_t$ measurable.

Let us set

$$I(v; h; T) = E[e^{\mu \log V_T(h)}], \tag{3.1}$$

and reformulate the problem as the one of partially observable stochastic control. Recall that Y_t is a solution of

$$dY_t = F(X_t, Y_t)dt + \Sigma(Y_t)dW_t, \tag{2.1$'$}$$

in the present case and we regard it as the SDE defining the observation process. On the other hand, X_t defined by (2.2) with the initial condition $X_0 = x_0$ is regarded as a system process. System noise $\Lambda(X_t, Y_t)dW_t$ and observation noise $\Sigma(Y_t)dW_t$ are correlated in general. $\sigma(Y_u, ; u \le t) = \sigma(S(u); u \le t)$ holds since log is a strictly increasing function, so our problem is to minimize (or maximize) the criterion (3.1) while looking at the observation process Y_t and choosing a $\sigma(Y_u, ; u \le t) = \tilde{\mathcal{G}}_t$ measurable strategy $h(t)$. Though there is no control in SDE (2.2) defining the system process X_t the criterion $I(v;h;T)$ is defined as a functional of the strategy $h(t)$ measurable with respect to observation and the problem is the one of stochstic control with partial observation.

In what follows we consider the case where $F(x,y) = F(x)$, $\Sigma(y) = \Sigma \equiv$ constant, $B(x,y) = B(x)$, $\Lambda(x,y) = \Lambda(x)$, $\tilde{r}(x,y) = r(x)$ for simplicity. Similar arguments are possible for general case as long as Σ does not depend on x. Now let us introduce a new probability measure $\hat{P}$ on $(\Omega, \mathcal{F})$ defined by

$$\left.\frac{d\hat{P}}{dP}\right|_{\mathcal{F}_T} = \rho_T,$$

where

$$\begin{aligned}\rho_t = \exp\{&-\textstyle\int_0^t F(X_s)^*(\Sigma\Sigma^*)^{-1}\Sigma dW_s \\ &-\tfrac{1}{2}\textstyle\int_0^t F(X_s)^*(\Sigma\Sigma^*)^{-1}F(X_s)ds\}.\end{aligned} \tag{3.2}$$

We see that $\hat{P}$ is a probability measure since it can be seen by standard arguments (cf. [1]) that ρ_t is a martingale and $E[\rho_T] = 1$ under assumption (2.18). Moreover, according to Girsanov theorem,

$$\hat{W}_t = W_t + \int_0^t \Sigma^*(\Sigma\Sigma^*)^{-1}F(X_s)ds \tag{3.3}$$

turns out to be a standard Brownian motion process under the probability measure $\hat{P}$ and we have

$$dY_t = \Sigma d\hat{W}_t \tag{3.4}$$

$$dX_t = \{B(X_t) - \Lambda(X_t)\Sigma^*(\Sigma\Sigma^*)^{-1}F(X_t)\}dt + \Lambda(X_t)d\hat{W}_t. \tag{3.5}$$

We rewrite our criterion $I(v;h;T)$ by new probability measure $\hat{P}$.

$$I(v;h;T) = v^{\mu}\hat{E}[\hat{E}[\exp\{-\mu\int_0^T \eta(X_s,h_s)ds\}\Psi_T|\tilde{\mathcal{G}}_t]], \tag{3.6}$$

where

$$\Psi_t = \exp\{\int_0^t Q(X_s,h_s)^* dY_s - \frac{1}{2}\int_0^t Q(X_s,h_s)^*(\Sigma\Sigma^*)Q(X_s,h_s)ds\}$$

and

$$\begin{aligned}Q(x,h)^* &= (\Sigma\Sigma^*)^{-1}F(x) + \mu h \\ &= (\Sigma\Sigma^*)^{-1}\{F(x) + \mu(\Sigma\Sigma^*)h\}.\end{aligned}$$

Set

$$(3.7)\qquad q^h(t)(\varphi(t)) = \hat{E}[\exp\{-\mu \int_0^t \eta(X_s, h_s))ds\}\Psi_t\varphi(t, X_t)|\mathcal{G}_t].$$

Then (3.6) reads

$$(3.8)\qquad I(v; h; T) = v^\mu \hat{E}[q^h(T)(1)]$$

Hence, if $\mu < 0$ (resp. $1 > \mu > 0$) our problem is reduced to minimize (resp. maximize) I of (3.8) when taking h over $\mathcal{H}(T)$. Let us set

$$(3.9)\qquad L\varphi = \frac{1}{2}(\Lambda\Lambda^*)^{ij}(x)D_{ij}\varphi + B(x)^i D_i\varphi.$$

Here and in what follows we utilize summation convention. Then, we can see that $q^h(t)$ satisfies a so called modified Zakai equation in a similar way to deducing Zakai equations as for conditional expectations of diffusion processes with respect to unnormalized conditional probabilities (cf. [2], [13], [21]). We actually have the following proposition.

Proposition 3.1. *Assume (2.18), then $q(t)(\varphi(t)) \equiv q^h(t)(\varphi(t))$ satisfies the following stochastic partial differential equation (SPDE):*
(3.10)

$$q(t)(\varphi(t)) = q(0)(\varphi(0)) + \int_0^t q(s)(\tfrac{\partial\varphi}{\partial t}(s,\cdot) + L\varphi(s,\cdot) + \mu h_s^*\Sigma\Lambda^*(\cdot)D\varphi(s,\cdot)$$

$$- \mu\eta_s(\cdot)\varphi(s,\cdot))ds + \int_0^t q(s)(\varphi(s,\cdot)Q(\cdot, h_s))dY_s$$

$$+ \int_0^t q(s)((D\varphi)^*(s,\cdot)\Lambda(\cdot)\Sigma^*(\Sigma\Sigma^*)^{-1})dY_s,$$

where $\eta_s(\cdot) = \eta(\cdot, h_s))$.

Let us introduce some notations and describe a strong form of stochstic partial differential equation (3.10). Set

$$L^0\varphi = \frac{1}{2}D_i(\Lambda\Lambda^*(x)^{ij}D_j\varphi),$$

$$\tilde{B}(x)^i = B(x)^i - \frac{1}{2}D_j(\Lambda\Lambda^*)^{ji}.$$

Then $L\varphi = L^0\varphi + \tilde{B}(x)^* D\varphi$ and its formal adjoint L^* is written as

$$L^*\varphi = L^0\varphi - D_i(\tilde{B}(x)^i\phi).$$

We set

$$G(h)\varphi = -D_i(\tilde{B}^i(\cdot)\varphi) - \mu h_s^i D_j((\Sigma\Lambda^*)^{ij}\varphi) - \mu\eta(\cdot, h_s)\varphi$$

and

$$M(h)_j\varphi = Q_j(\cdot, h_s)\varphi - D_i([\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}]^i_j\varphi)$$

We define

$$\begin{aligned}\mathcal{L}^2_Y(0,T;H^1(R^n)) &= \{v \in L^2(\Omega, \mathcal{F}, \hat{P}; L^2(0,T;H^1(R^n)),\\ &\qquad v(t) \in L^2(\Omega, \tilde{\mathcal{G}}_t, \hat{P}; H^1(R^n)) \text{ a.e. } t\}\end{aligned}$$

Then we consider the following stochastic partial differential equation which has a solution $q(t)$ such that $q_t e^{\delta\sqrt{1+|x|^2}} \in \mathcal{L}_Y([0,T];H^1)$.

$$dq_t = (L^0 q_t + G(h)q_t)dt + M(h)_j q_t dY^j_t. \tag{3.11}$$

Furthermore we assume

$$\Lambda,\ D\Lambda, B,\ DB,\ F, \quad \text{are bounded} \tag{3.12}$$

and the set of admissible strategies $\mathcal{A}_T$ is defined as the totality of $\tilde{\mathcal{G}}_t$ measurable strategy h satisfying the condition i) of definition 2.1 and $h_t \in \Gamma$, $\forall t$ for some convex compact $\Gamma \subset R^m$. Take a positive constant $\delta > 0$. Then we have the following theorem.

Proposition 3.2. *Let us assume (2.18), (3.12), and* $\pi e^{\delta\sqrt{1+|x|^2}} \in H^1$. *Then for each addmissible strategy h (3.11) has a unique solution* $q_t = q(t,x)$ *such that* $q_t e^{\delta\sqrt{1+|x|^2}} \in \mathcal{L}^2_Y(0,T;H^1(R^n)) \cap L^2(\Omega, \mathcal{F}, \hat{P}; C(0,T;L^2(R^n))$ *and that* $q_0 = \pi$. *Furthermore we have* $\int q(T,x)\psi(x)dx = q(T)(\psi)$ *for all bounded Borel function* ψ.

For the proof of this proposition we prepare the following lemma.

Lemma 3.1. *Under assumption (2.18)*

$$\Lambda(I_{n+m} - \Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma)\Lambda^* \geq c_1 I_n$$

Proof. Note that

$$(\xi_1^*, \xi_2^*)\begin{pmatrix}\Lambda\Lambda^* & \Lambda\Sigma^*\\ \Sigma\Lambda^* & \Sigma\Sigma^*\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix} \geq c_1|\xi|^2, \quad \forall\xi = \begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}. \tag{3.13}$$

under assumtion (2.18). Therefore, setting $\zeta = -(\Sigma\Sigma^*)^{-1}\Sigma\Lambda^*\xi_1$ for $\xi_1 \in R^n$, we see that

$$\begin{aligned}&\xi_1^*\Lambda(I_{n+m} - \Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma)\Lambda^*\xi_1\\ &= \xi_1^*\Lambda\Lambda^*\xi_1 + \xi_1^*\Lambda\Sigma^*\zeta + \zeta\Sigma\Lambda^*\xi_1 + \zeta\Sigma\Sigma^*\zeta\\ &\geq c_1(|\xi_1|^2 + |\zeta|^2) \geq c_1|\xi_1|^2\end{aligned}$$

□

Proof of Proposition 3.2. Set $\tilde{q}_t = q_t e^{\delta\sqrt{1+|x|^2}}$, $\delta > 0$ and $\nu(x) = \delta\sqrt{1+|x|^2}$. Then (3.11) can be written as

$$d\tilde{q}_t = (L^0\tilde{q}_t + \tilde{G}(h)\tilde{q}_t)dt + \tilde{M}(h)_j\tilde{q}_t dY_t^j, \tag{3.14}$$

where

$$\tilde{G}(h)\tilde{q} = G(h)\tilde{q} + (\Lambda\Lambda)^{ij}D_j\nu D_i\tilde{q}$$
$$+\{\tfrac{1}{2}D_i((\Lambda\Lambda)^{ij}D_j\nu)+\tfrac{1}{2}(\Lambda\Lambda^*)^{ij}D_i\nu D_j\nu-(\tilde{B}^iD_i\nu+\mu h^i(\Sigma\Lambda^*)^{ij}D_j\nu)\}\tilde{q},$$

$$\tilde{M}(h)_j\tilde{q} = \tilde{Q}_j\tilde{q} - D_i([\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}]_j^i\tilde{q})$$

and

$$\tilde{Q}_j = Q_j - D_i\nu[\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}]_j^i.$$

It suffices to check the coercivity condition for (3.14) because of general theory of stochastic partial differential equarions ([2],[13], [25]):

$$-2\langle L^0 q, q\rangle - 2\langle\tilde{G}(h)q, q\rangle + c_1 \| q \|_{L^2}^2 \geq c_0 \| q \|_{H^1}^2 + \langle\tilde{M}(h)_j q, (\Sigma\Sigma^*)^{jk}\tilde{M}(h)_k q\rangle$$

Indeed, setting $\hat{Q}_j = \tilde{Q}_j - D_i([\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}]_j^i)$ we see that

$$\begin{aligned}
&-2\langle L^0 q, q\rangle - \langle\tilde{M}(h)_j q, (\Sigma\Sigma^*)^{jk}\tilde{M}(h)_k q\rangle \\
&\quad= \textstyle\int (Dq)^*\Lambda\Lambda^* Dq dx - \int (Dq)^*\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma\Lambda^* Dq dx \\
&\qquad - \textstyle\int (\hat{Q}^*\Sigma\Sigma^*\hat{Q})q^2 dx - 2\int (Dq)^*\Lambda\Sigma^*\hat{Q} q dx \\
&\quad\geq \textstyle\int (Dq)^*\Lambda(I - \Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma)\Lambda^* Dq dx - \int \hat{Q}(\Sigma\Sigma^*)\hat{Q}q^2 dx \\
&\qquad - \epsilon \textstyle\int |Dq|^2 dx - \frac{1}{\epsilon}\int |\Lambda\Sigma^*\hat{Q}|^2 q^2 dx \\
&\quad\geq c_2 \textstyle\int |Dq|^2 dx - \epsilon\int |Dq|^2 dx - c_3\int |q|^2 dx
\end{aligned}$$

for some c_2, $c_3 > 0$ and sufficiently small $\epsilon > 0$ by using the above lemma. Since

$$|\langle\tilde{G}q, q\rangle| \leq \epsilon_1 \int |Dq|^2 dx + (\frac{c_4}{\epsilon_1} + c_5)\int |q|^2 dx$$

we can easily see the coercivity condition holds for (3.14).

□

Lemma 3.2. *Let us assume the assumptions of the above proposition and h, k be admissible strategies, then*

$$\lim_{\theta \to 0} \frac{I(v; h_\cdot + \theta k_\cdot; T) - I(v; h_\cdot; T)}{\theta} = v^\mu \hat{E}[< \zeta_T(k), 1 >], \tag{3.15}$$

where $\zeta_t = \zeta_t(k)$ is a solution of the following stochastic partial differential equation

(3.16)

$$d\zeta_t = (L^0\zeta_t + G(h)\zeta_t + k_t^i G_{h^i}(h) q_t)dt + (M(h)_j \zeta_t + k_t^i M_{h^i}(h)_j q_t) dY_t^j$$
$$\zeta_0 = 0,$$

where q_t is a solution to (3.11),

$$\begin{aligned} G_{h^i}(h)q &= -\mu D_j((\Sigma\Lambda^*)^{ij} q) - \mu \tfrac{\partial \eta}{\partial h^i} q \\ &= -\mu D_j((\Sigma\Lambda^*)^{ij} q) - \mu[(1-\mu)(\Sigma\Sigma^* h)^i - (g - r\mathbf{1})^i] \end{aligned}$$

and

$$M_{h^i}(h)_j q = \frac{\partial Q_j(\cdot, h)}{\partial h^i} q = \mu \delta_{ij} q.$$

Proof. Note that we can see that (3.14) has a unique solution such that $q_t e^{\delta\sqrt{1+|x|^2}} \in \mathcal{L}_Y^2(0,T; H^1(R^n)) \cap L^2(\Omega, \mathcal{F}, \hat{P}; C(0,T; L^2(R^n))$ in a similar way to the proof of the above proposition. Let us set

$$\bar{q}_\theta(t) = \frac{q_\theta(t) - q(t)}{\theta} - \zeta,$$

where $q_\theta(t)$ is the solution to :

(3.17)

$$dq_\theta(t) = \{L^0 q_\theta(t) + G(h + \theta k) q_\theta(t)\}dt + M(h + \theta k)_j q_\theta(t) dY_t^j$$

$$q_\theta(0) = \pi$$

We define in the same way as above

$$\tilde{q}_\theta(t) = q_\theta(t) e^{\nu(x)}, \quad \tilde{\zeta} = \zeta e^{\nu(x)}.$$

Then in a similar way to getting (3.14), we have stochastic partial differential equations for $\tilde{q}_\theta(t)$ and $\tilde{\zeta}$. We set

$$\tilde{\bar{q}}_\theta(t) = \frac{\tilde{q}_\theta(t) - \tilde{q}(t)}{\theta} - \tilde{\zeta}.$$

Then we can see that

$$\sup_{0 \le t \le T} E[\| \tilde{\bar{q}}_\theta(t) \|_{L^2}^2] \to 0$$

as $\theta \to 0$ by using the energy equality for $\tilde{\bar{q}}_\theta(t)$. Since

$$\frac{I(v;h_\cdot+\theta k_\cdot;T)-I(v;h_\cdot;T)}{\theta}-v^\mu\hat{E}[\langle\zeta_T,1\rangle]=v^\mu\hat{E}[\langle\tilde{\bar{q}}_\theta(T),e^{-\nu(x)}\rangle]$$

we obtain the present proposition.

□

Let us introduce the following backward stochastic partial differential equation.

$$(3.18)\qquad \begin{array}{rcl} -d\gamma_t &=& (L^0\gamma_t+\hat{G}(h)\gamma_t+\hat{M}(h)R_t)dt-R_t^*(\Sigma\Sigma^*)^{-1}dY_t \\ \gamma_T &=& 1 \end{array}$$

where

$$\hat{G}(h)\varphi=\tilde{B}^*D_\varphi+\mu h^*\Sigma\Lambda^*D\varphi-\mu\eta(\cdot,h)\varphi$$

$$\hat{M}(h)R=R^jQ_j(\cdot,h)+[\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}]^i_jD_iR^j.$$

Set

$$\check{\gamma}_t=e^{-\nu(x)}\gamma_t,\quad \check{R}_t=e^{-\nu(x)}R_t.$$

Then we have the following backward SPDE

$$(3.19)\qquad \begin{array}{rcl} -d\check{\gamma}_t &=& (L^0\check{\gamma}_t+\check{G}(h)\check{\gamma}_t+\check{M}(h)\check{R}_t)dt-\check{R}_t^*(\Sigma\Sigma^*)^{-1}dY_t \\ \check{\gamma}_T &=& e^{-\nu(x)}, \end{array}$$

where

$$\begin{aligned}\check{G}(h)\varphi &= \{(D\nu)^*\Lambda\Lambda^*+\tilde{B}^*+\mu h^*\Sigma\Lambda^*\}D\varphi \\ &+\{L^0\nu+\tfrac{1}{2}(D\nu)^*\Lambda\Lambda^*D\nu+(\tilde{B}^*+\mu h^*\Sigma\Lambda^*)D\nu-\mu\eta(\cdot,h)\}\varphi \\ &\equiv G_1^*D\varphi+G_2\varphi\end{aligned}$$

and

$$\begin{aligned}\check{M}(h)U &= \sum_{i,j}[\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}]^i_jD_iU^j \\ &+\textstyle\sum_j\{Q_j(\cdot,h)+\sum_i[\Lambda\Sigma^*(\Sigma\Sigma^*)^{-1}]^i_jD_i\nu\}U^j \\ &\equiv \textstyle\sum_{i,j}(M_1)^i_jD_iU^j+\sum_j(M_2)_jU^j.\end{aligned}$$

Let $\check{\gamma}_t$ be a solution to (3.19) with the terminal condition $\check{\gamma}_T = 0$ and set $(M_2')_j = (M_2)_j + D_j(G_1)^j$. We have by Itô's formula
(3.20)

$$
\begin{aligned}
&\hat{E}[\| \check{\gamma}_t \|_{L^2}^2] \\
&= \hat{E}[\int_t^T \{2\langle L^0\check{\gamma}_s + \check{G}(h)\check{\gamma}_s + \check{M}(h)\check{R}_s, \check{\gamma}_s\rangle - (\check{R}_s, (\Sigma\Sigma^*)^{-1}\check{R}_s)\}ds] \\
&= \hat{E}[\int_t^T \int\{-(D\check{\gamma}_s)^*\Lambda\Lambda^* D\check{\gamma}_s + G_1^* D(\check{\gamma}_s^2) + 2G_2\check{\gamma}_s^2 \\
&\quad + 2(M_2')_j\check{R}_s^j\check{\gamma}_s - 2\check{R}_s^j(M_1)_j^i D_i\check{\gamma}_s - \check{R}_s^*(\Sigma\Sigma^*)^{-1}\check{R}_s\}dx\, ds \\
&= \hat{E}[\int_t^T \int\{-(D\check{\gamma}_s)\Lambda(I - \Sigma^*(\Sigma\Sigma^*)^{-1}\Sigma)\Lambda^* D\check{\gamma}_s \\
&\quad - [(\Sigma\Sigma^*)^{-1}\check{R}_s + M_1^* D\check{\gamma}_s - (M_2')^*\gamma_s]^*(\Sigma\Sigma^*)[(\Sigma\Sigma^*)^{-1}\check{R}_s \\
&\qquad + M_1^* D\check{\gamma}_s - (M_2')^*\gamma_s] \\
&\quad + [M_2'\Sigma\Sigma^*(M_2')^* + 2G_2 - \sum_j D_j(G_1 + M_1\Sigma\Sigma^*(M_2')^*)^j]\check{\gamma}_s^2\}dx\, ds \\
&\le C\int_t^T \hat{E}[\| \check{\gamma}_s \|_{L^2}^2]ds
\end{aligned}
$$

for some constant $C > 0$. By using (3.20) we can obtain the following lemma.

Lemma 3.3. *Under the assumptions of Proposition 3.2 the solution (γ_t, R_t) to (3.18) such that $e^{-\delta\sqrt{1+|x|^2}}\gamma_t \in \mathcal{L}_Y^2(0,T;H^1(R^n)) \cap L^2(\Omega,\mathcal{F},\hat{P};C(0,T;L^2(R^n)))$ and $e^{-\delta\sqrt{1+|x|^2}}R^i \in \mathcal{L}_Y^2(0,T;H^1(R^n)) \cap L^2(\Omega,\mathcal{F},\hat{P};C(0,T;L^2(R^n)))$, $i = 1,2,...,m$ is unique.*

We can also see the existence of the solution to (3.18) in a similar way to Theorem 8.2.3 [2] through aproximation procedure, or directly thanks to Chapter 5, Theorem 2.2 in [16].

Lemma 3.4. *Under the assumptions of Proposition 3.2*

$$\hat{E}[< \zeta_T, 1 >] = \hat{E}[\int_0^T \{\langle\gamma_t, k_t^i G_{h^i}(h)q_t > + < R_t^j, k_t^i M_{h^i}(h)_j q_t >\}dt].$$

Proof. From (3.16) and (3.18) we obtain

$$
\begin{aligned}
d\langle\zeta_t,\gamma_t\rangle &= \{\langle k_t^i G_{h^i}(h)q_t, \gamma_t\rangle + \langle k_t^i M_{h^i}(h)_j q_t, R_t^j\rangle\}dt \\
&\quad + \{\langle M(h)_j\zeta_t,\gamma_t\rangle + \langle k_t^i M_{h^i}(h)_j q_t,\gamma_t\rangle + \langle M(h)_j\zeta_t,\gamma_t\rangle\}dY_t^j
\end{aligned}
$$

and we have the present lemma.

$\square$

Finally we have the following theorem

Theorem 3.1. *We assume the assumptions of Proposition 3.2. If h is optimal, then it satisfies*

(3.21)
$$(k-h_t)^*\{-(1-\mu)(\Sigma\Sigma^*)h_t < \gamma_t, q_t > + < \Sigma\Lambda^* D\gamma_t, q_t) > \\ + < (g - r\mathbf{1})\gamma_t, q_t > + < R_t, q_t >\} \le 0,$$

a.e. t a.s. $\forall k \in \Gamma$.

Proof. Let $h_t,\ k_t$ be admissible atrategies and h_t is an optimal one. Since Γ is convex $h + \theta(k-h) = (1-\theta)h + \theta k \in \Gamma$, for $h, k \in \Gamma$. Thus we have

$$I(v; h_.; T) \ge I(v; h_. + \theta(k_. - h_.); T), \quad 0 \le \forall\theta \le 1$$

if $\mu > 0$. Therefore, because of Lemma 3.2

$$\hat{E}[\langle \zeta_T(k-h), 1\rangle] \le 0,$$

which implies that

$$(3.22)\quad \hat{E}[\int_0^T \{\langle \gamma_t, (k_t^i - h_t^i) G_{h^i}(h) q_t\rangle + \langle R_t^j, (k_t^i - h_t^i) M_{h^i}(h)_j q_t\rangle\} dt] \le 0$$

for all admissible strategy k_t by Lemma 3.4. Set

$$(U_t)_i = \langle \gamma_t, G_{h^i}(h) q_t\rangle + \langle R_t^j, M_{h^i}(h)_j q_t\rangle.$$

For each $t_0 \in [0,T],\ \epsilon > 0,\ M > 0$ and $\tilde{G}_{t_0}$ measurable random variable k_{t_0}, we define

$$k_t = \begin{cases} k_{t_0} 1_{\{|U_t| \le M\}} + h_t 1_{\{|U_t| > M\}}, & t_0 \le t \le t_0 + \epsilon \\ h_t, & t \in [t_0, t_0+\epsilon]^c. \end{cases}$$

Then, through limiting procedure as $\epsilon \to 0$ and $M \to \infty$ after multilying (3.22) by $\frac{1}{\epsilon}$, we see that

$$\hat{E}[(k_{t_0} - h_{t_0})^i (U_{t_0})_i] \le 0$$

for each t_0 and $\tilde{G}_{t_0}$ measurable random variable k_{t_0}, which implies that

$$(k-h_t)^i \{\langle \gamma_t, G_{h^i}(h) q_t\rangle + \langle R_t^j, M_{h^i}(h)_j q\rangle\} \le 0 \quad \text{a.e. } t \text{ a.s.}$$

for each $k \in \Gamma$. Hence

(3.23)

$$\mu(k-h_t)^*\{-(1-\mu)(\Sigma\Sigma^*)h_t < \gamma_t, q_t > + < \Sigma\Lambda^* D\gamma_t, q_t) >$$

$$+ < (g - r\mathbf{1})\gamma_t, q_t > + < R_t, q_t >\} \leq 0.$$

Since $\mu > 0$ we have (3.21).

If $\mu < 0$ we obtain the converse inquality of (3.23) and we conclude the present theorem.

□

References

[1] A. Bensoussan, *Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions*, Stochastics 9 (1983) 169-222

[2] A. Bensoussan, *Stochastic Control of Partially Observable Systems*, Cambridge (1992)

[3] A. Bensoussan, J. Frehse and H. Nagai, *Some Results on Risk-sensitive with full observation*, Appl. Math. and its Optimization, vol.37 (1998)1-41

[4] A. Bensoussan and H. Nagai, *Condition for no breakdown and Bellman equations of risk-sensitive control*, Appl. Math. and its Optimization, vol.42 (2000)91-101

[5] T.R. Bielecki and S.R. Pliska, *Risk-Sensitive Dynamic Asset Management*, Appl. Math. Optim. vol. 39 (1999) 337-360

[6] T.R. Bielecki and S.R. Pliska, *Risk-Sensitive Intertemporal CAPM, With Application to Fixed Income Management*, preprint

[7] W.H. Fleming, *Optimal investment models and risk sensitive stochastic control*, Mathematical Finance, edited by M. Davis et al., Springer-Verlag (1995)75-88

[8] W.H. Fleming and S.J. Sheu, *Optimal Long Term Growth Rate of Expected Utility of Wealth*, Ann. Appl. Prob.9(1999)871-903

[9] W.H. Fleming and S.J. Sheu, *Risk-sensitive control and an optimal investment model*, Mathematical Finance, 10 (2000)197-213

[10] W.H. Fleming and S.J. Sheu, *Risk-sensitive control and an optimal investment model (II)*, Annals of Prob.,12(2002)730-767

[11] H. Kaise and H. Nagai, *Bellman-Isaacs equations of ergodic type related to risk-sensitive control and their singular limits*, Asymptotic Analysis, 16 (1998) 347-362

[12] H. Kaise and H. Nagai Ergodic type Bellman equations of risk-sensitive control with large parameters and their singular limits. Asymptotic Analysis, 20 (1999) 279-299

[13] N.V. Krylov and B.L. Rozovski, *On conditional distributions of diffusion processes*, Math USSR Izvestya 12 (1978) 336-356

[14] K. Kuroda and H. Nagai *Risk-sensitive portfolio optimization on infinite time horizon*, Stochastics and Stochastics Reports, 73 (2002) 309-331

[15] J. Ma and J. Yong, *On linear, degenerate backward stochastic partial differential equations*, Prob. Theory Rel. Fields, 113 (1999) 135-170

[16] J. Ma and J. Yong, *Forward-Backward Stochastic Differential equations and Their Applications*, Lect. Notes in Math. 1702, Springer(1999)

[17] R.C. Merton, *Optimum Consumption and Portfolio Rules in a Continuous-Time Model*, Journal of Economic Theory 3 (1971) 373-413

[18] R.C. Merton, *An Intertemporal Capital Asset Pricing Model*, Econometrica (1973) 867-887

[19] R.C. Merton, *Continuous Time Finance*, Blackwell, Malden (1990)

[20] H. Nagai, *Bellman equations of risk-sensitive control*, SIAM J. Control and Optimization. Vol. 34, No. 1 (1996) 74-101.

[21] H. Nagai, *Risk-senstive dynamic asset management with partial information*, Stochastics in finite and infinite dimension, a volume in honor of G. Kallianpur, Eds. Rajput et al., Birkhäuser (2000) 321-340

[22] H. Nagai and S. Peng, *Risk-sensitive dynamic portfolio optimization with partial information on infinite time horizon*, Annals of Applied Probability, vol.12, No.1(2002)173-195

[23] H. Nagai, *Optimal strategies for risk-sensitive portfolo optimization problems for general factor models*, to appear in SIAM J. Control and Optim.

[24] R. Rishel, *Optimal Portfolio Management with Partial Observations and Power Utility Function*, Stochastic Analysis, Control, Optimization and Applications, Birkhäuser (1999) 605-620

[25] B.L. Rozovski, *Stochastic Evolution Systems*, Kluwer Academic Publishers (1990)

[26] X.Y. Zhou, *On the necessary conditions of optimal controls for stochstic partial differential equations*, SIAM J. Control Optim. 31,(1993) 1462-1478

Department of Mathematical Science
Osaka University
Toyonaka, 560-8531
Japan

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 279–299

An Approximation for Exponential Hedging

Jun Sekine

Abstract.

An optimization problem in mathematical finance, called the exponential hedging problem is addressed. First, the relations between the problem and the backward stochastic differential equation (abbreviated to BSDE) having a quadratic growth term in the drift are reviewed. Next, the asymptotic analysis by Davis (2000) for the problem and the motivation of this paper are stated. Further, with some extensions, his analysis is reinterpreted by using the asymptotic expansion of the BSDE with respect to a small parameter, which suggests an alternative approach to the analysis, and the result on an approximated optimizer is obtained.

§1. Introduction

In [7], Rouge and El Karoui treated the following optimization problem of mathematical finance. For a fixed $T > 0$, let $S := (S_t)_{t\in[0,T]}$, $S_t := \left(S_t^1, \dots S_t^n\right)'$ be the price process of n-risky assets defined by the stochastic differential equation:

$$\begin{aligned} dS_t &= \operatorname{diag}(S_t)\left(\sigma_t dw_t + \mu_t dt\right), \quad S_0 \in \mathbf{R}^n_+, \\ \operatorname{diag}(S_t) &:= \begin{pmatrix} S_t^1 & 0 & \cdots & 0 \\ 0 & S_t^2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & S_t^n \end{pmatrix} \end{aligned}$$

on the probability space $(\Omega, \mathcal{F}, P)$ with a $d(\geq n)$-dimensional Brownian motion $w := (w_t)_{t\in[0,T]}$ on it and the augmented Brownian filtration $(\mathcal{F}_t)_{t\in[0,T]}$. Here, σ is an $n \times d$-matrix-valued left-continuous adapted process such that $\sigma\sigma' \in L^\infty([0,T]\times\Omega, \mathbf{R}^{n\times n})$ and that $\sigma_t\sigma_t'$ is invertible

Received March 3, 2003.
Revised September 1, 2003.
The author is grateful to the anonymous referees for their careful reading, helpful suggestions, and pointing out errors in the original version.

for all $t \in [0,T]$ ($(\cdot)'$ denotes the transpose of a matrix or a vector), μ is an n-dimensional predictable process, and $\lambda := \sigma' (\sigma\sigma')^{-1} (\mu - r\mathbf{1})$ is an element of $L^\infty([0,T] \times \Omega, \mathbf{R}^d)$, where $r(> 0)$ is the constant interest rate and $\mathbf{1} := (1, \ldots, 1)' \in \mathbf{R}^n$. On the other hand, let $F \in L^\infty(\Omega, \mathcal{F}_T)$ be the payoff of a derivative security maturing at time T and consider a seller of the derivative security, who trades the assets continuously in self-financing way on the time-interval $[0,T]$ to control the terminal wealth. The value process of the self-financing portfolio is given by

$$dX_t^{x,\pi} = \pi_t' (\mathrm{diag}(S_t))^{-1} dS_t + (X_t^{x,\pi} - \pi_t'\mathbf{1})\, r dt, \quad X_0^{x,\pi} = x,$$

or equivalently,

$$X_t^{x,\pi} := e^{rt} \left\{ x + \int_0^t \pi_u' \sigma_u (dw_t + \lambda_t dt) \right\}$$

where x is the initial capital and an n-dimensional predictable process π is the asset holding strategy. To optimize the terminal wealth $-F + X_T^{x,\pi}$ of the seller, the utility maximization problem (called the exponential hedging problem in this paper, following Delbaen et. al; 2002, [2])

$$\text{(P)} \qquad V(x) := \sup_{\pi \in \mathcal{A}} E\left[U_\gamma\left(-F + X_T^{x,\pi}\right)\right]$$

with respect to the exponential utility function:

$$U_\gamma(x) := -\frac{e^{-\gamma x}}{\gamma}, \quad (\gamma > 0)$$

over an appropriately chosen space $\mathcal{A}$ of admissible strategies is considered.

The importance of this problem is, from a viewpoint of mathematical finance, that it relates to the pricing and hedging problems of derivative securities in incomplete markets: the quantity called utility indifference price,

$$(1) \qquad p(x, F) := \inf_{y \in \mathbf{R}} \left\{ V(x + y) \geq \sup_{\pi \in \mathcal{A}} E\left[U_\gamma\left(X_T^{x,\pi}\right)\right] \right\},$$

is proposed as a coherent price of the derivative security in Davis (2000), [1] and [7], and the optimizer of the problem (P) is focused and studied to control (hedge) the "risk" of the seller in [1], [2] and [7].

Duality argument is well established for utility maximization (cf., Karatzas and Shreve; 1998, [5], for example) and is often used to attack this problem, as follows. For example, let us employ the space

$$\mathcal{A}_2 := \left\{ \pi \in \mathcal{L}_T^{2,n};\ \pi_t \in C \text{ for all } t \in [0,T],\ E\left[\int_0^T |\pi_t|^2 dt \right] < \infty \right\},$$

as the set of admissible strategies, $\mathcal{A}$, where $C \subset \mathbf{R}^n$ is a fixed closed convex cone and $\mathcal{L}_T^{2,n}$ is the totality of the n-dimensional predictable processes π on the time-interval $[0,T]$ such that $\int_0^T |\pi_t|^2 dt < \infty$, a.s. For $f, x \in \mathbf{R}$, and $y > 0$, denote

$$u_\gamma(x;y,f) := U_\gamma(-f+x) - yx \quad \text{and} \quad I_\gamma(y) := \left(U_\gamma'\right)^{-1}(y) = -\frac{1}{\gamma}\log(y)$$

to see the relation

$$\sup_{x\in\mathbf{R}} u_\gamma(x;y,f) = u_\gamma\left(f + I_\gamma(y); y, f\right) = -y\left(f - \frac{1+\log y}{\gamma}\right).$$

Define

$$Z^\nu := \mathcal{E}\left(-(\lambda-\nu)\cdot w\right) \quad \text{and} \quad \widetilde{Z}_t^\nu := e^{-rt} Z_t^\nu,$$

where ν is an element of

$$\mathcal{D} := \left\{ \nu \in \mathcal{L}_T^{2,d};\ \text{bounded and } \ \nu_t \in \widehat{(\sigma_t' C)} \text{ for all } t \in [0,T] \right\}$$

and $\widehat{(\sigma_t' C)}$ is the notation for the negative polar cone of $\sigma_t' C$, i.e.,

$$\widehat{(\sigma_t' C)}(\omega) := \left\{ y \in \mathbf{R}^d;\ xy \le 0 \text{ for all } x \in \sigma_t'(\omega) C \right\}.$$

For $\pi \in \mathcal{A} := \mathcal{A}_2$ and $\nu \in \mathcal{D}$, observe that

$$e^{-rt} X_t^{x,\pi} \le x + \int_0^t \pi_u' \sigma_u \left\{ dw_u + (\lambda_u - \nu_u) du \right\}$$

and that $Z^\nu \int \pi' \sigma \left\{ dw + (\lambda-\nu) du \right\}$ is a martingale since $E\left[\sup_{t\in[0,T]} |Z_t^\nu|^2\right] < \infty$ and since

$$E\left[\sup_{t\in[0,T]} \left| \int_0^t \pi_u' \sigma_u \left\{ dw_u + (\lambda_u - \nu_u) du \right\} \right|^2 \right] \le C_1 E\left[\int_0^T |\pi_u|^2 du \right] < \infty$$

from Doob's inequality and the boundedness assumptions of $\sigma\sigma'$, λ and ν. Therefore, the relation

$$E\left[\widetilde{Z}_T^\nu X_T^{x,\pi} \right] \le x$$

follows. Based on the relation, for $\pi \in \mathcal{A}$ and $x \in \mathbf{R}, y > 0$, we observe the inequalities

$$
\begin{aligned}
(2) \qquad & E\left[U_\gamma\left(-F + X_T^{x,\pi}\right)\right] - yx \\
\leq\ & \inf_{\nu \in \mathcal{D}} E\left[U_\gamma\left(-F + X_T^{x,\pi}\right) - y\widetilde{Z}_T^\nu X_T^{x,\pi}\right] \\
\leq\ & \inf_{\nu \in \mathcal{D}} \sup_{\pi \in \mathcal{A}} E\left[u_\gamma\left(X_T^{x,\pi}; y\widetilde{Z}_T^\nu, F\right)\right] \\
\leq\ & \inf_{\nu \in \mathcal{D}} E\left[u_\gamma\left(F + I_\gamma\left(y\widetilde{Z}_T^\nu\right); y\widetilde{Z}_T^\nu, F\right)\right].
\end{aligned}
$$

The minimization problem

$$
\text{(D)} \qquad \widehat{V}(y) := \inf_{\nu \in \mathcal{D}} E\left[u_\gamma\left(F + I_\gamma\left(y\widetilde{Z}_T^\nu\right); y\widetilde{Z}_T^\nu, F\right)\right]
$$

is called the dual problem of the primal problem (P), and the inequality

$$
(3) \qquad V(x) \leq \inf_{y>0}\left(\widehat{V}(y) + yx\right)
$$

is deduced from (2). Indeed, the equality can be established in (3) (i.e., there is no "duality-gap") and the following expression is obtained.

Theorem 1. *(Theorem 2.1 of Rouge and El Karoui, [7]) For $\mathcal{A} := \mathcal{A}_2$, it holds that*

$$
(4) \qquad V(x) = U_\gamma\left(e^{rT}x - \frac{1}{\gamma}\sup_{\nu \in \mathcal{D}}\left\{E^\nu[\gamma F] - H\left(P^\nu|P\right)\right\}\right),
$$

where $E^\nu[\cdot]$ denotes the expectation with respect to the probability measure P^ν on $(\Omega, \mathcal{F}_T)$ defined by

$$
\left.\frac{dP^\nu}{dP}\right|_{\mathcal{F}_t} := Z_t^\nu
$$

and

$$
H(Q|P) := \begin{cases} E\left[\frac{dQ}{dP}\log\frac{dQ}{dP}\right] & \text{if } Q \ll P, \\ +\infty & \text{otherwise} \end{cases}
$$

is the relative entropy of Q with respect to P.

Remark 1. The duality relations similar to (4) have been obtained for more general semimartingale S and for other choices of the set of admissible strategies $\mathcal{A}$ by Delbaen et. al. in [2]. Also, the work by Kabanov and Stricker (2002), [4], should be referred.

For the computations of the value $V(x)$ and the optimizer, one can solve the BSDE for the value process of the dual problem, described as follows.

Theorem 2. *(Theorem 4.1-2 of Rouge and El Karoui, [7]) Denote* $Z^{\nu}_{t,T} := Z^{\nu}_T / Z^{\nu}_t$, $\widetilde{Z}^{\nu}_{t,T} := \widetilde{Z}^{\nu}_T / \widetilde{Z}^{\nu}_t$, *and* $\tau := T - t$ *for* $0 \le t \le T$. *Let*

$$
\begin{aligned}
&\operatorname*{essinf}_{\nu \in \mathcal{D}} E\left[u_\gamma\left(F + I_\gamma\left(y\widetilde{Z}^{\nu}_{t,T}\right); y\widetilde{Z}^{\nu}_{t,T}, F\right) \mid \mathcal{F}_t\right] \\
= \quad & \frac{ye^{-r\tau}}{\gamma}\left\{-\operatorname*{esssup}_{\nu \in \mathcal{D}} E^{\nu}\left[\gamma F - \log Z^{\nu}_{t,T} | \mathcal{F}_t\right] + (1 + \log y - r\tau)\right\} \\
=: \quad & \frac{ye^{-r\tau}}{\gamma}\left\{-Y_t + (1 + \log y - r\tau)\right\}.
\end{aligned}
$$

There exists $\Xi \in \mathbf{H}^{2,d}_T := \left\{f \in \mathcal{L}^{2,d}_T;\ E\left[\int_0^T |f_t|^2 dt\right] < \infty\right\}$ *such that* (Y, Ξ) *satisfies*

$$
\begin{aligned}
dY_t &= f(t, \Xi_t)\, dt + \Xi'_t dw_t, \quad Y_T = \gamma F, \\
\textit{where } f(t,\xi) &:= \lambda'_t \Pi_{\sigma'_t C}(\xi + \lambda_t) - \frac{1}{2}\left|\xi - \Pi_{\sigma'_t C}(\xi + \lambda_t)\right|^2 .
\end{aligned}
\tag{5}
$$

and $\Pi_{\sigma'_t(\omega)C} : \mathbf{R}^d \ni x \mapsto \Pi_{\sigma'_t(\omega)C}\, x \in \sigma'_t(\omega)C (\subset \mathbf{R}^d)$ *is the projection operator onto the closed convex cone* $\sigma'_t(\omega)C$.

In particular, $\pi^* \in \mathcal{A}_2$ *satisfying*

$$
\sigma'_t \pi^*_t := \frac{e^{-rT}}{\gamma} \Pi_{\sigma'_t C}(\Xi_t + \lambda_t) \quad \textit{for all } t \in [0, T]
\tag{6}
$$

is an optimizer of the primal problem (P) with $\mathcal{A} := \mathcal{A}_2$, *and* $\nu^* := (I - \Pi_{\sigma' C})(\Xi + \lambda)$ *attains the infimum of the dual problem (D). Further,*

$$
V(x) = U_\gamma\left(e^{rT}x - \frac{Y_0}{\gamma}\right)
\tag{7}
$$

holds.

Remark 2. The existence and the uniqueness of the solution (Y, Ξ) of the quadratic BSDE (5) in the space $\mathbf{H}^{\infty}_T \times \mathbf{H}^{2,d}_T$, where $\mathbf{H}^{\infty}_T := \{f \in L^{\infty}([0,T] \times \Omega);\ \text{predictable}\}$ is ensured by the work of Kobylanski (2000), [6]. Further, utilizing the dynamic programming principle and the comparison theorems between linear BSDEs and between quadratic BSDEs in [6], the above theorem is established.

On the other hand, if the model has a Markovian structure, one can solve a dynamic programming equation to compute the value, which is suggested in Delbaen et al (2002), [2], and is employed and studied in Davis (2000), [1]. In particular, in [1], a special but a typical situation is addressed, which can be stated as follows in our setting.

(i) Let $d = n = 2$. σ is the following constant matrix

(8)
$$\sigma := \begin{pmatrix} \sigma_1 & 0 \\ \sigma_2\sqrt{1-\epsilon^2} & \sigma_2\epsilon \end{pmatrix}$$

with $\sigma_1, \sigma_2 > 0$, $\epsilon \in [-1,1]$. $\mu = (\mu_1, \mu_2)'$ is also a constant vector. Further, $\epsilon \neq 0$, $\epsilon \ll 1$ is assumed, i.e., two assets S^1 and S^2 are *closely correlated*:

$$\frac{d\langle S^1, S^2\rangle}{\sqrt{d\langle S^1\rangle d\langle S^2\rangle}} = \sqrt{1-\epsilon^2} \approx 1.$$

(ii) $F := h(S_T^1)$ with continuous, piecewise linear $h : \mathbf{R}_+ \mapsto \mathbf{R}$ bounded from above.

(iii) The constraint of the asset-holding strategy π is given by $\pi_t \in C := \{0\} \times \mathbf{R}$: only S^2 is *tradable*, and the derivative security is written on the *untradable* asset S^1.

Recall, in the situation, that the expressions

$$\sigma' C = \{k d_\epsilon;\ k \in \mathbf{R}\}, \quad \mathcal{D} = \left\{\eta d_\epsilon^\perp;\ \eta \in \mathcal{L}_T^{2,1}, \text{bounded}\right\},$$

and

$$\lambda^\epsilon := \sigma^{-1}(\mu - r\mathbf{1}) = \frac{1}{\epsilon\sigma_1\sigma_2}\begin{pmatrix} \epsilon\sigma_2 & 0 \\ -\sigma_2\sqrt{1-\epsilon^2} & \sigma_1 \end{pmatrix}\begin{pmatrix} \mu_1 - r \\ \mu_2 - r \end{pmatrix}$$

hold, where we denote

$$d_\epsilon := \left(\sqrt{1-\epsilon^2}, \epsilon\right)' \quad \text{and} \quad d_\epsilon^\perp := \left(\epsilon, -\sqrt{1-\epsilon^2}\right)'.$$

The dual problem is now, rewritten as

$$\begin{aligned} &\inf_{\nu\in\mathcal{D}} E\left[\left(-y\widetilde{Z}_T^\nu\right)\left\{h\left(S_T^1\right) - \frac{1}{\gamma}\left(1 + \log y + \log \widetilde{Z}_T^\nu\right)\right\}\right] \\ = \ & \frac{ye^{-rT}}{\gamma}\left\{-\sup_{\nu\in\mathcal{D}} E^\nu\left[\gamma h\left(S_T^1\right) - \log Z_T^\nu\right] + (1 + \log y - rT)\right\}. \end{aligned}$$

Since

$$\log Z_T^\nu = -\int_0^T (\lambda^\epsilon - \nu_t)' dw_t^\nu + \frac{1}{2}\int_0^T |\lambda^\epsilon - \nu_t|^2 dt,$$

where $w^\nu := (w_1^\nu, w_2^\nu)'$, $w_t^\nu := w_t + \int_0^t (\lambda^\epsilon - \nu_u)du$ is a 2-dimensional P^ν-Brownian motion, it is equivalent to solve the following:

$$\sup_{\nu\in\mathcal{D}} E^\nu\left[\gamma h\left(S_T^1\right) - \frac{1}{2}\int_0^T |\lambda^\epsilon - \nu_t|^2 dt\right],$$

in which the process S^1 has the dynamics:

$$\begin{aligned} dS_t^1 &= S_t^1 \left[\sigma_1 dw_1^\nu(t) + \left\{ \mu_1 - \sigma_1 (\overline{\lambda}^\epsilon - \nu_t^1) \right\} dt \right] \\ &= S_t^1 \left\{ \sigma_1 dw_1^\nu(t) + (r - \epsilon \sigma_1 \eta_t)\, dt \right\}, \end{aligned}$$

where we denote $\nu := \eta d_\epsilon^\perp$ with some bounded predictable η. For the value function

$$v^\epsilon(t,y) := \operatorname*{esssup}_{\nu \in \mathcal{D}} E^\nu \left[\gamma h\left(S_T^1\right) - \frac{1}{2} \int_t^T |\lambda^\epsilon - \nu_t|^2 \, dt \;\middle|\; S_t^1 = y \right],$$

a dynamic-programming equation is derived and the existence of its smooth solution is checked in the setting of [1]. Moreover, the following expressions are obtained.

Theorem 3. *(Theorem 6.1, 6.4 and 7.3 of Davis,* [1]*)*
1. *An optimal strategy of the problem (P) is given by*

$$\begin{aligned} \pi_t^* &= \frac{e^{-rT}}{\gamma} (\sigma')^{-1} \Pi_{\sigma' C} \left\{ \begin{pmatrix} -\partial_x v^\epsilon(t, S_t^1) S_t^1 \sigma_1 \\ 0 \end{pmatrix} + \lambda_t^\epsilon \right\} \\ &= \begin{pmatrix} 0 \\ \frac{e^{-rT}}{\gamma} \left\{ \frac{\mu_2 - r}{\sigma_2^2} - \sqrt{1-\epsilon^2} \frac{\sigma_1}{\sigma_2} \partial_x v^\epsilon(t, S_t^1) S_t^1 \right\} \end{pmatrix}. \end{aligned}$$

2. *For the utility indifference price defined by (1),*

$$p(x,F) = \frac{e^{-rT}}{\gamma} \left\{ v^\epsilon(0, S_0^1) + \frac{T}{2} \left(\frac{\mu_2 - r}{\sigma_2} \right)^2 \right\}$$

holds for any $x \in \mathbf{R}$.
3. *As $\epsilon \downarrow 0$, the value function has the expansion*

$$\begin{aligned} v^\epsilon(t,y) &= \gamma E\left[h(A_T) | A_t = y\right] - \frac{T-t}{2} \left(\frac{\mu_2 - r}{\sigma_2} \right)^2 \\ &\quad + \epsilon^2 \frac{\gamma^2}{2} \operatorname{Var}\left[h(A_T) | A_t = y\right] + O(\epsilon^4), \end{aligned} \tag{9}$$

where $\operatorname{Var}[*|\cdot] := E[(*)^2|\cdot] - (E[*|\cdot])^2$, $O(\epsilon^4)$ *depends on the value* (t,y), *and the process A is defined by*

$$dA_t = A_t \left[\sigma_1 dw_1(t) + \left\{ \mu_1 - \sqrt{1-\epsilon^2} \frac{\sigma_1(\mu_2 - r)}{\sigma_2} \right\} dt \right], \quad A_0 = S_0^1.$$

In particular, we are interested in the expansion (9). From a practical viewpoint, it is an effective and useful expansion: it gives nice approximations of the value of the problem (P) and the utility indifference price. By using the relation (7),

$$\log V^{\epsilon}(x) - \log U_{\gamma}\Bigg(e^{rT}x - \gamma E\left[h(A_T)\right] - \frac{T}{2}\left(\frac{\mu_2 - r}{\sigma_2}\right)^2 \\ -\epsilon^2\frac{\gamma^2}{2}\mathrm{Var}\left[h(A_T)\right]\Bigg) = O(\epsilon^4)$$

is observed, where we denote the value by $V^{\epsilon}(x)$ emphasizing ϵ, and

$$p(x,F) = e^{-rT}\left\{E\left[h(A_T)\right] + \epsilon^2\frac{\gamma}{2}\mathrm{Var}\left[h(A_T)\right]\right\} + O(\epsilon^4)$$

holds for any $x \in \mathbf{R}$. Also, both quantities $E\left[h(A_T)|A_t = y\right]$ and $\mathrm{Var}\left[h(A_T)|A_t = y\right]$ are fairly "computable". In [1], it is derived from a clever observation, however, the reason why the second term has $O(\epsilon^2)$ and the error term has $O(\epsilon^4)$ seems to be obscure. To see its intrinsic reason is one of our motivations.

Further, we are interested in the approximation of the optimal strategy (optimizer), which is not mentioned in [1]. It looks natural to deduce the strategy $\check{\pi} := \left(\check{\pi}^1, \check{\pi}^2\right)'$ defined by $\check{\pi}^1 \equiv 0$ and

$$\begin{aligned}\check{\pi}^2_t \;&:=\; \frac{e^{-rT}}{\gamma}\Bigg[\frac{\mu_2 - r}{\sigma_2^2} - \sqrt{1-\epsilon^2}\frac{\sigma_1}{\sigma_2}S^1_t \\ &\qquad \times\partial_y\left(\gamma E\left[h(A_T)|A_t = y\right] + \epsilon^2\frac{\gamma^2}{2}\mathrm{Var}\left[h(A_T)|A_t = y\right]\right)\Bigg|_{y=S^1_t}\Bigg]\end{aligned}$$

and expect the approximation such that

$$\log V^{\epsilon}(x) - \log E\left[U_{\gamma}\left(-F + X_T^{x,\check{\pi}}\right)\right] = O(\epsilon^4), \tag{10}$$

for example.

In the next section, using the BSDE in Theorem 2 and its asymptotic expansion with respect to ϵ (precisely saying, with respect to ϵ', cf., the BSDE (14)), we reconstruct the expansions (9-10), which yields an alternative approach to the above analysis. The main contribution of this paper is Theorem 4 in the next section, extensions of (9-10) under Assumption 1. It is also suggested that the i-th derivatives $\left(\partial^i_{\epsilon'}Y^{0,\epsilon}, \partial^i_{\epsilon'}\Xi^{0,\epsilon}\right) \equiv 0$ for odd numbers $i = 1, 3, 5, \cdots$ (cf., Remark 4).

§2. An approximated optimizer

In this section, the probability space is assumed to be the product of Wiener spaces: $(\Omega, \mathcal{F}, P) := \prod_{i=1}^{2}(\Omega_i, \mathcal{F}^i, P_i)$, where $\Omega_i := C_0([0,T], \mathbf{R})$, $\mathcal{F}^i := \mathcal{B}(\Omega_i)$ and P_i is the Wiener measure, the law of the i-th canonical Brownian motion $w^i := \left(w^i_t\right)_{t\in[0,T]}$. The filtration $(\mathcal{F}_t)_{t\in[0,T]} := (\mathcal{F}^1_t \times \mathcal{F}^2_t)_{t\in[0,T]}$ is the augmented natural filtration. Sometimes a random variable X on $(\Omega_1, \mathcal{F}^1, P_1)$ is identified with $X \circ j_1$ on $(\Omega, \mathcal{F}, P)$, where $j_1 : \Omega \ni \omega := (\omega_1, \omega_2) \mapsto \omega_1 \in \Omega_1$ is the projection onto the first probability space.

We now impose the following conditions.

(i)' The volatility matrix of the process S is given by (8). On the other hand, $\mu = (\mu_1, \mu_2)'$ is a bounded $\mathcal{F}^1_t$-predictable process, i.e., $\mu : [0,T] \times \Omega_1 \ni (t, \omega_1) \mapsto \mu(t, \omega_1) \in \mathbf{R}^2$ is measurable with respect to the predictable σ-algebra on $[0,T] \times \Omega_1$.

(ii)' $F(\omega_1) = h(S^1(\omega_1))$ with a bounded measurable function h on $C([0,T], \mathbf{R}_+)$.

(iii) The constraint of the strategy π is given by $\pi_t \in C := \{0\} \times \mathbf{R}$.

Remark 3. The condition (i)' is considered as an extension of the constant μ case employed in (i) in the previous section, though Assumption 1.1 will be added later. On the other hand, the conditoin (ii)' does not include the conditon (ii) in the previous section.

Further, we consider the problem (P) over the extended space: $\mathcal{A} := \mathcal{A}_1$, where

$$\mathcal{A}_1 := \left\{ \pi \in \mathcal{L}^{2,2}_T;\ \pi_t \in C \text{ for } \forall t \in [0,T], \right.$$
$$\left. E\left[\left(\int_0^T |\pi_t|^2 dt\right)^{q/2}\right] < \infty \text{ for } \exists q > 1 \right\}$$

and construct an approximated optimizer in $\mathcal{A}_1$, not in $\mathcal{A}_2$. We first remark the following.

Proposition 1. *Let π^* be the process defined by the formula (6) and by the solution $(Y, \Xi) \in \mathbf{H}^\infty_T \times \mathbf{H}^{2,2}_T$ of the BSDE (5). It is also an optimizer of the problem (P) with $\mathcal{A} := \mathcal{A}_1$.*

Proof. We first observe that $E\left[\widetilde{Z}^\nu_T X^{x,\pi}_T\right] \le x$ for all $(\pi, \nu) \in \mathcal{A}_1 \times \mathcal{D}$ and $x \in \mathbf{R}$. For the purpose, since

$$e^{-rt} X^{x,\pi}_t = x + \int_0^t \pi'_u \sigma_u dw^\nu_u$$

holds, to show the martingale property of the process $Z^\nu \int \pi' \sigma dw^\nu$ is sufficient, which can be verified by checking

$$
\begin{aligned}
& E\left[\sup_{t\in[0,T]}\left|Z_t^\nu \int_0^t \pi_u' \sigma_u dw_t^\nu\right|\right] \\
\leq\ & E\left[\sup_{t\in[0,T]}\left|Z_t^\nu \int_0^t \pi_u' \sigma_u dw_t\right|\right] + E\left[\sup_{t\in[0,T]}\left|Z_t^\nu \int_0^t \pi_u' \sigma_u (\lambda_u - \nu_u) du\right|\right] \\
\leq\ & E\left[\sup_{t\in[0,T]} (Z_t^\nu)^p\right]^{1/p} \left\{ E\left[\sup_{t\in[0,T]}\left|\int_0^t \pi_u' \sigma_u dw_u\right|^q\right]^{1/q} \right. \\
& \qquad\qquad \left. + E\left[\sup_{t\in[0,T]}\left|\int_0^t \pi_u' \sigma_u (\lambda_u - \nu_u) du\right|^q\right]^{1/q} \right\} \\
\leq\ & C_1 E\left[\langle Z^\nu\rangle_T^{p/2}\right]^{1/p} \\
& \times \left\{ C_2 E\left[\left(\int_0^T |\pi_u|^2 du\right)^{q/2}\right]^{1/q} + C_3 E\left[\left(\int_0^T |\pi_u| du\right)^q\right]^{1/q} \right\} \\
\leq\ & C_4 E\left[\langle Z^\nu\rangle_T^{p/2}\right]^{1/p} E\left[\left(\int_0^T |\pi_u|^2 du\right)^{q/2}\right]^{1/q} < \infty
\end{aligned}
$$

for $p, q > 1$ satisfying $1/p + 1/q = 1$ by using the Hölder inequality and the Burkholder-Davis-Gundy inequality. In particular, the inequalities (2) and

$$
E\left[U_\gamma\left(-F + X_T^{x,\pi}\right)\right] - yx \leq \widehat{V}(y) \tag{11}
$$

are deduced for any $\pi \in \mathcal{A}_1$ $x \in \mathbf{R}$ and $y > 0$.

Next, note that the pair (π^*, ν^*) defined by (6) and the formula $\nu_t^* := \left(I - \Pi_{\sigma_t' C}\right)\left(\Xi_t^\epsilon + \lambda_t^\epsilon\right)$ satisfies the relation

$$
\left\{
\begin{array}{l}
F + I_\gamma\left(\mathcal{Y}(x) \widetilde{Z}_T^{\nu^*}\right) = X_T^{x,\pi^*} \\
\text{with} \quad \mathcal{Y}(x) := \exp\left(Y_0 + rT - \gamma e^{rT} x\right).
\end{array}
\right. \tag{12}
$$

In fact, the solution of the BSDE satisfies

$$
\begin{aligned}
F &= \frac{1}{\gamma}\left\{Y_0 + \int_0^T f(t,\Xi_t)\,dt + \int_0^T (\Xi_t)'\,dw_t\right\} \\
&= \frac{Y_0}{\gamma} + \int_0^T \left\{e^{rT}(\pi_t^*)'\sigma_t\lambda_t - \frac{1}{2\gamma}|\lambda_t - \nu_t^*|^2\right\}dt \\
&\quad + \int_0^T \left\{e^{rT}\sigma_t'\pi_t^* - \frac{1}{\gamma}(\lambda_t - \nu_t^*)\right\}' dw_t \\
&= \frac{Y_0}{\gamma} + X_T^{0,\pi^*} - I_\gamma\left(Z_T^{\nu^*}\right),
\end{aligned}
$$

which is equivalent to (12). Using (12) and Theorem 2, we observe that

$$
\begin{aligned}
E\left[\widetilde{Z}_T^{\nu^*}X_T^{x,\pi^*}\right] &= E\left[\widetilde{Z}_T^{\nu^*}\left(F + I_\gamma\left(\mathcal{Y}(x)\widetilde{Z}_T^{\nu^*}\right)\right)\right] \\
&= \frac{e^{-rT}}{\gamma}\left\{Y_0 - \log\left(\mathcal{Y}(x)e^{-rT}\right)\right\} = x.
\end{aligned} \tag{13}
$$

Finally, replacing y by $\mathcal{Y}(x)$ in (11) and using Theorem 2 and (12-3), we deduce that

$$
\begin{aligned}
& E\left[U_\gamma\left(-F + X_T^{x,\pi}\right)\right] - \mathcal{Y}(x)x \\
\le\ & E\left[u_\gamma\left(F + I_\gamma\left(\mathcal{Y}(x)\widetilde{Z}_T^{\nu^*}\right); \mathcal{Y}(x)\widetilde{Z}_T^{\nu^*}, F\right)\right] \\
=\ & E\left[U_\gamma\left(-F + X_T^{x,\pi^*}\right)\right] - \mathcal{Y}(x)x
\end{aligned}
$$

for all $\pi \in \mathcal{A}_1$, which implies the optimality of π^*. ∎

The BSDE (5) for the optimizer is now rewritten as, in the situation of this section,

$$
\begin{aligned}
dY_t^\epsilon &= f(t,\Xi_t^\epsilon,\epsilon)\,dt + (\Xi_t^\epsilon)'\,dw_t, \quad Y_T^\epsilon = \gamma F, \\
\text{where}\quad f(t,\xi,\epsilon) &:= \frac{1}{2}\left\{\overline{\lambda}_t^2 - \left(\xi, d_\epsilon^\perp\right)^2\right\} + \overline{\lambda}_t\left(\xi, d_\epsilon\right),
\end{aligned}
$$

$(\cdot,\cdot)$ denotes the standard inner-product in $\mathbf{R}^2$ and

$$
\overline{\lambda} := (\lambda^\epsilon, d_\epsilon) = \frac{\mu_2 - r}{\sigma_2}.
$$

Denote

$$
\overline{w}_t^\epsilon = (\overline{w}_1^\epsilon(t), \overline{w}_2^\epsilon(t))' := w_t + \left(\int_0^t \overline{\lambda}_u du\right) d_\epsilon
$$

to reexpress the solution $(Y^\epsilon, \Xi^\epsilon) = (Y^{\epsilon,\epsilon}, \Xi^{\epsilon,\epsilon})$ by using the BSDE:

$$
\begin{aligned}
(14) \qquad dY_t^{\epsilon',\epsilon} &= g\left(t, \Xi_t^{\epsilon',\epsilon}, \epsilon'\right) dt + \left(\Xi_t^{\epsilon',\epsilon}\right)' d\overline{w}_t^\epsilon, \quad Y_T^{\epsilon',\epsilon} = \gamma F, \\
\text{where} \quad g(t,\xi,\epsilon') &:= \frac{1}{2}\left\{\overline{\lambda}_t^2 - (\xi, d_{\epsilon'}^\perp)^2\right\}.
\end{aligned}
$$

We consider the asymptotic expansion of $\left(Y^{\epsilon',\epsilon}, \Xi^{\epsilon',\epsilon}\right)$ with respect to ϵ' at 0. Let $\left(\partial_{\epsilon'}^0 Y^{0,\epsilon}, \partial_{\epsilon'}^0 \Xi^{0,\epsilon}\right) := \left(Y^{0,\epsilon}, \Xi^{0,\epsilon}\right)$ and introduce the BSDEs:

$$
(15) \quad \begin{cases} d\left(\partial_{\epsilon'}^i Y_t^{0,\epsilon}\right) = g_i\left(t, \left(\partial_{\epsilon'}^j \Xi_t^{0,\epsilon}\right)_{j=0,\ldots,i}, 0\right) dt + \left(\partial_{\epsilon'}^i \Xi_t^{0,\epsilon}\right)' d\overline{w}_t^\epsilon, \\ \partial_{\epsilon'}^i Y_T^{0,\epsilon} = 0, \end{cases}
$$

using the functions g_i defined inductively

$$
\begin{aligned}
g_0\left(t, \xi^0, \epsilon'\right) &:= g\left(t, \xi^0, \epsilon'\right) \\
\text{and} \quad g_i\left(t, (\xi^j)_{j=0,\ldots,i}, \epsilon'\right) &:= \sum_{j=0}^{i-1} \left(\partial_{\xi^j} g_{i-1}\left(t, (\xi^k)_{k=0,\ldots,i-1}, \epsilon'\right), \xi^{j+1}\right) \\
&\quad + \partial_{\epsilon'} g_{i-1}\left(t, (\xi^k)_{k=0,\ldots,i-1}, \epsilon'\right).
\end{aligned}
$$

Formally, it is expected that $\left(\partial_{\epsilon'}^i Y^{0,\epsilon}, \partial_{\epsilon'}^i \Xi^{0,\epsilon}\right)$ is the i-th derivative of $\left(Y^{\epsilon',\epsilon}, \Xi^{\epsilon',\epsilon}\right)$ with respect to ϵ' at $\epsilon' = 0$, although we have not been able to show the property. The standard results on the differentiation of the solution of BSDE with respect to a parameter (cf., El Karoui et. al ; 1997, [3], for example) cannot be applied to our quadratic BSDE (14).

Define the probability measure $\overline{P}^\epsilon$ on $(\Omega, \mathcal{F}_T)$ by

$$
\begin{aligned}
\left.\frac{d\overline{P}^\epsilon}{dP}\right|_{\mathcal{F}_t} &:= \mathcal{E}_t\left(-\int \overline{\lambda} d_\epsilon' dw\right) =: \overline{Z}_t^\epsilon \\
&= \mathcal{E}_t\left(-\sqrt{1-\epsilon^2}\int \overline{\lambda} dw_1\right) \mathcal{E}_t\left(-\epsilon \int \overline{\lambda} dw_2\right) =: \overline{Z}_1^\epsilon(t)\overline{Z}_2^\epsilon(t)
\end{aligned}
$$

and the space $\mathbf{H}_T^{2,2,\epsilon} := \left\{f \in \mathcal{L}_T^{2,2};\ \int_0^T |f_t|^2 dt \in L^1(\overline{P}^\epsilon)\right\}$ to obtain the expressions for the solution of (15) for $i = 0, 1, 2, 3$, as follows.

Lemma 1. 1. *The solution* $(Y^{0,\epsilon}, \Xi^{0,\epsilon})$ *in the space* $\mathbf{H}_T^\infty \times \mathbf{H}_T^{2,2}$ *has the expressions:*

$$\begin{aligned} Y_t^{0,\epsilon} &= \overline{E}^\epsilon \left[\gamma F - \frac{1}{2}\int_t^T \overline{\lambda}_u^2 du \,\Big|\, \mathcal{F}_t \right], \\ Y_0^{0,\epsilon} + \int_0^t \Xi_1^{0,\epsilon}(u) d\overline{w}_1^\epsilon(u) &= \overline{E}^\epsilon \left[\gamma F - \frac{1}{2}\int_0^T \overline{\lambda}_u^2 du \,\Big|\, \mathcal{F}_t \right], \end{aligned}$$

and $\Xi_2^{0,\epsilon}(t) = 0$ *for all* $t \in [0,T]$, *where* $\overline{E}^\epsilon[\cdot]$ *denotes the expectation with respect to the probability measure* $\overline{P}^\epsilon$.
2. $\left(\partial_{\epsilon'}^i Y^{0,\epsilon}, \partial_{\epsilon'}^i \Xi^{0,\epsilon}\right) \equiv 0$ *for* $i = 1, 3$.
3. *A solution of (15) with* $i = 2$ *exists in* $\mathbf{H}_T^\infty \times \mathbf{H}_T^{2,2,\epsilon}$ *and is given by*

$$\begin{aligned} \partial_{\epsilon'}^2 Y_t^{0,\epsilon} &= \overline{E}^\epsilon \left[\int_t^T \left(\Xi_1^{0,\epsilon}(u) \right)^2 du \,\Big|\, \mathcal{F}_t \right] \\ &= \overline{\mathrm{Var}}^\epsilon \left[\gamma F - \frac{1}{2}\int_t^T \overline{\lambda}_u^2 du \,\Big|\, \mathcal{F}_t \right], \\ \partial_{\epsilon'}^2 Y_0^{0,\epsilon} + \int_0^t \partial_{\epsilon'}^2 \Xi_1^{0,\epsilon}(u) d\overline{w}_1^\epsilon(u) &= \overline{E}^\epsilon \left[\int_0^T \left(\Xi_1^{0,\epsilon}(u) \right)^2 du \,\Big|\, \mathcal{F}_t \right], \end{aligned}$$

and $\partial_{\epsilon'}^2 \Xi_2^{0,\epsilon}(t) = 0$ *for all* $t \in [0,T]$, *where we denote* $\overline{\mathrm{Var}}^\epsilon[\cdot|\mathcal{F}_t] := \overline{E}^\epsilon[(\cdot)^2|\mathcal{F}_t] - \left(\overline{E}^\epsilon[\cdot|\mathcal{F}_t]\right)^2$.

Proof. 1. Suppose $\Xi_2^{0,\epsilon} \equiv 0$, then

$$dY_t^{0,\epsilon} = \frac{1}{2}\overline{\lambda}_t^2 dt + \Xi_1^{0,\epsilon}(t) d\overline{w}_1^\epsilon(t), \quad Y_T^{0,\epsilon} = \gamma F$$

is observed. **1** is now a consequence of the standard result of linear BSDE (cf., El Karoui et. al, [3]) and the result on the uniqueness of the quadratic BSDE studied in Kobylanski (2000), [6].
2-3. Observe that

$$\begin{aligned} d_\epsilon^\perp &= \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \epsilon \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{\epsilon^2}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{\epsilon^3}{3!}\begin{pmatrix} 0 \\ 0 \end{pmatrix} + O(\epsilon^4) \\ &=: d_0^\perp + \sum_{i=1}^3 \frac{\epsilon^i}{i!}\partial_{\epsilon'}^i d_0^\perp + O(\epsilon^4), \end{aligned}$$

where $O(\epsilon^4) \in \mathbf{R}^2$ is a vector with the norm $|O(\epsilon^4)| \sim \epsilon^4$.

(i) Noting that

$$g_1\left(t,(\xi^j)_{j=0,1},0\right) = -\left(\xi^0, d_0^\perp\right)\left\{\left(\xi^1, d_0^\perp\right) + \left(\xi^0, \partial_{\epsilon'} d_0^\perp\right)\right\}$$

and that $\Xi_2^0 \equiv 0$, we can deduce

$$d\left(\partial_{\epsilon'} Y_t^{0,\epsilon}\right) = \partial_{\epsilon'} \Xi_t^{0,\epsilon} d\overline{w}_t^\epsilon, \quad \partial_{\epsilon'} Y_T^{0,\epsilon} \equiv 0$$

and $\left(\partial_{\epsilon'} Y^{0,\epsilon}, \partial_{\epsilon'} \Xi^{0,\epsilon}\right) \equiv 0$.
(ii) Observing that

$$\begin{aligned} & g_2\left(t,(\xi^j)_{j=0,1,2},0\right) \\ = \; & -\left(\xi^1, d_0^\perp\right)\left\{\left(\xi^1, d_0^\perp\right) + \left(\xi^0, \partial_{\epsilon'} d_0^\perp\right)\right\} \\ & -\left(\xi^0, d_0^\perp\right)\left\{\left(\xi^2, d_0^\perp\right) + \left(\xi^1, \partial_{\epsilon'} d_0^\perp\right)\right\} \\ & -\left(\xi^0, \partial_{\epsilon'} d_0^\perp\right)\left\{\left(\xi^1, d_0^\perp\right) + \left(\xi^0, \partial_{\epsilon'} d_0^\perp\right)\right\} \\ & -\left(\xi^0, d_0^\perp\right)\left\{\left(\xi^1, \partial_{\epsilon'} d_0^\perp\right) + \left(\xi^0, \partial_{\epsilon'}^2 d_0^\perp\right)\right\}, \end{aligned}$$

we rewrite the BSDE for $\left(\partial_{\epsilon'}^2 Y^{0,\epsilon}, \partial_{\epsilon'}^2 \Xi^{0,\epsilon}\right)$ as

$$d\left(\partial_{\epsilon'}^2 Y_t^{0,\epsilon}\right) = -\left(\Xi_1^{0,\epsilon}(t)\right)^2 dt + \left(\partial_{\epsilon'}^2 \Xi_t^{0,\epsilon}\right)' d\overline{w}_t^\epsilon, \quad \partial_{\epsilon'}^2 Y_T^{0,\epsilon} \equiv 0$$

since $\Xi_2^{0,\epsilon} \equiv 0$ and $\partial_{\epsilon'} \Xi^{0,\epsilon} \equiv 0$. Define $\overline{P}^\epsilon$-martingales M, N by the formulas

$$M_t := \int_0^t \Xi_1^{0,\epsilon}(u) d\overline{w}_1^\epsilon(u) := -Y_0^{0,\epsilon} + \overline{E}^\epsilon\left[\gamma F - \frac{1}{2}\int_0^T \overline{\lambda}_u^2 du \,\middle|\, \mathcal{F}_t\right]$$

and $N_t := \overline{E}^\epsilon\left[\langle M\rangle_T | \mathcal{F}_t\right]$ for $t \in [0,T]$, respectively. Note that M is bounded and that N is $\overline{P}^\epsilon$-square integrable:

$$\overline{E}^\epsilon\left[N_t^2\right] \le \overline{E}^\epsilon\left[\langle M\rangle_t^2\right] = \overline{E}^\epsilon\left[\left(M_t^2 - M_0^2 - 2\int_0^t M_u dM_u\right)^2\right] < \infty.$$

The martingale representation theorem implies that $H_t := E[\overline{Z}_1^\epsilon(T) N_T | \mathcal{F}_t]$ $= N_0 + \int_0^t \phi_u dw_1(u)$ holds for all $t \in [0,T]$ and for some $\mathcal{F}_t^1$-predictable ϕ such that $\int_0^T \phi_u^2 du < \infty$. Therefore,

$$N_t - N_0 = \int_0^t d\left(\frac{\overline{Z}_1^\epsilon(u) N_u}{\overline{Z}_1^\epsilon(u)}\right) = \int_0^t \frac{\phi_u + \sqrt{1-\epsilon^2} H_u \overline{\lambda}_u}{\overline{Z}_1^\epsilon(u)} d\overline{w}_1^\epsilon(u)$$

is observed from the Itô-formula. The solution is now constructed by setting

$$\partial_{\epsilon'}^2 Y_t^{0,\epsilon} := N_t - \int_0^t \left(\Xi_1^{0,\epsilon}(u)\right)^2 du, \quad \partial_{\epsilon'}^2 \Xi_1^{0,\epsilon}(t) := \frac{\phi_t + \sqrt{1-\epsilon^2} H_t \overline{\lambda}_t}{\overline{Z}_1^{\epsilon}(t)}$$

and $\partial_{\epsilon'}^2 \Xi_2^{0,\epsilon} \equiv 0$.

(iii) For $\left(\xi^j\right)_{j=0,1,2,3}$ such that $\xi_2^0 = \xi_2^2 = 0$ and $\xi^1 = 0$, we can check that

$$g_3\left(t, (\xi^j)_{j=0,1,2,3}, 0\right) = 0,$$

so the equation

$$d\left(\partial_{\epsilon'}^3 Y_t^{0,\epsilon}\right) = \partial_{\epsilon'}^3 \Xi_t^{0,\epsilon} d\overline{w}_t^{\epsilon}, \quad \partial_{\epsilon'}^3 Y_T^{0,\epsilon} \equiv 0$$

and $\left(\partial_{\epsilon'}^3 Y^{0,\epsilon}, \partial_{\epsilon'}^3 \Xi^{0,\epsilon}\right) \equiv 0$ are deduced. ∎

We are now in the position to state our last theorem, an extension of Theorem 3.3. We require the conditions:

Assumption 1. 1. *The process μ_2 is bounded and deterministic.* **2.** *There is a kernel ∂F, finite measures $\partial F(\omega_1, \cdot)$ on $\mathcal{B}([0,T])$ for each $\omega_1 \in \Omega_1$, satisfying*

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon}\{F(\omega_1 + \epsilon\phi) - F(\omega_1)\} = \int_0^T \phi(t) \partial F(\omega_1, dt)$$

for all $\phi \in C^1([0,T])$ and the Clark formula:

$$F = E[F] + \int_0^T E[\partial F(\cdot, (t,T]) | \mathcal{F}_t] dw_1(t).$$

(For sufficient conditions on F and ∂F to ensure the formula, cf., Appendix E of [5], *for example). Moreover, $\partial F(\cdot, (\cdot, T]) \in L^\infty(\Omega_1 \times [0,T])$ holds.*

For $n = 1, 2, \ldots$, define

$$\overline{Y}^{\epsilon,n} := \sum_{i=0}^{n} \partial_{\epsilon'}^{2i} Y_t^{0,\epsilon} \frac{\epsilon^{2i}}{(2i)!} \quad \text{and} \quad \overline{\Xi}^{\epsilon,n} := \sum_{i=0}^{n} \partial_{\epsilon'}^{2i} \Xi_t^{0,\epsilon} \frac{\epsilon^{2i}}{(2i)!}$$

and introduce the approximated strategy $\overline{\pi}^{\epsilon,n} := (\overline{\pi}_t^{\epsilon,n})_{t \in [0,T]}$ by the formula

$$(16)\ \overline{\pi}_t^{\epsilon,n} := \frac{e^{-rT}}{\gamma} (\sigma')^{-1} \Pi_{\sigma' C}\left(\overline{\Xi}_t^{\epsilon,n} + \lambda^{\epsilon}(t)\right) = \begin{pmatrix} 0 \\ \frac{e^{-rT}}{\gamma}\left\{\sigma_2^{-2}(\mu_2(t) - r) + \sqrt{1-\epsilon^2}\sigma_2^{-1} \overline{\Xi}_1^{\epsilon,n}(t)\right\} \end{pmatrix}.$$

Note that $\overline{\pi}^{\epsilon,1} \in \mathcal{A} := \mathcal{A}_1$ since $E\left[\int_0^T \left|\Xi_t^{0,\epsilon}\right|^2 dt\right] < \infty$ and since

$$\begin{aligned}
& E\left[\left(\int_0^T \left|\partial_{\epsilon'}^2 \Xi_t^{0,\epsilon}\right|^2 dt\right)^{q/2}\right] \\
= \; & \overline{E}^{\epsilon}\left[\left(\overline{Z}_T^{\epsilon}\right)^{-1}\left(\int_0^T \left|\partial_{\epsilon'}^2 \Xi_t^{0,\epsilon}\right|^2 dt\right)^{q/2}\right] \\
\leq \; & \overline{E}^{\epsilon}\left[\left(\overline{Z}_T^{\epsilon}\right)^{-\frac{2}{2-q}}\right]^{\frac{2-q}{2}} \overline{E}^{\epsilon}\left[\int_0^T \left|\partial_{\epsilon'}^2 \Xi_t^{0,\epsilon}\right|^2 dt\right]^{q/2} < \infty
\end{aligned}$$

for $0 < q < 2$. We obtain the following.

Theorem 4. *Under Assumption 1, the relations*

$$\begin{aligned}
\left\|Y^{\epsilon} - \overline{Y}^{\epsilon,1}\right\|_{L^{\infty}([0,T]\times\Omega)} &= O(\epsilon^4) \\
\textit{and} \quad \log V^{\epsilon}(x) - \log E\left[U_{\gamma}\left(-F + X_T^{x,\overline{\pi}^{\epsilon,1}}\right)\right] &= O(\epsilon^4)
\end{aligned}$$

hold as $\epsilon \downarrow 0$.

Proof. Denote $\Lambda^{\epsilon}(t) := (\int_0^t \overline{\lambda}_u du)d_{\epsilon}$ and define the Wiener functional and the kernel:

$$G(\omega_1) := \gamma F(\omega_1 - \Lambda_1^{\epsilon}) - \frac{1}{2}\int_0^T \overline{\lambda}_u^2 du, \quad \partial G(\omega_1, dt) := \gamma \partial F(\omega_1 - \Lambda_1^{\epsilon}, dt).$$

First, we observe the following:

$$\begin{aligned}
\Xi_1^{0,\epsilon}(t,\omega_1) &= E[\partial G(\cdot,(t,T])|\mathcal{F}_t](\omega_1 + \Lambda_1^{\epsilon}), \\
\partial_{\epsilon'}^2 \Xi_1^{0,\epsilon}(t,\omega_1) &= 2\{E[\partial G(\cdot,(t,T])G|\mathcal{F}_t] \\
&\quad - E[\partial G(\cdot,(t,T])|\mathcal{F}_t]\, E[G|\mathcal{F}_t]\}(\omega_1 + \Lambda_1^{\epsilon}).
\end{aligned}$$

In fact, the first expression is a consequence from the Clark formula,

$$G(\omega_1) = E[G] + \int_0^T E[\partial G(\cdot,(t,T])|\mathcal{F}_t](\omega_1) dw_1(t,\omega_1),$$

the Cameron-Martin formula, $\overline{P}^{\epsilon}(\cdot) = P(\cdot + \Lambda^{\epsilon})$, and the relation $w_t(\omega + \Lambda^{\epsilon}) = \overline{w}_t^{\epsilon}(\omega)$,

$$\begin{aligned}
G(\omega_1 + \Lambda_1^{\epsilon}) &= \overline{E}^{\epsilon}[G(\omega_1 + \Lambda_1^{\epsilon})] \\
&\quad + \int_0^T E[\partial G(\cdot,(t,T])|\mathcal{F}_t](\omega_1 + \Lambda_1^{\epsilon}) d\overline{w}_1^{\epsilon}(t,\omega_1).
\end{aligned}$$

The second expression is deduced from the relation

$$\begin{aligned}
&\int_0^T \left(\Xi_1^{0,\epsilon}(t)\right)^2 dt \\
=\;& \left(\int_0^T \Xi_1^{0,\epsilon}(t)d\overline{w}_1^\epsilon(t)\right)^2 - 2\int_0^T \left(\int_0^t \Xi_1^{0,\epsilon}(u)d\overline{w}_1^\epsilon(u)\right)\Xi_1^{0,\epsilon}(t)d\overline{w}_1^\epsilon(t) \\
=\;& \left(G - \overline{E}^\epsilon[G]\right)^2 - 2\int_0^T \{E[G|\mathcal{F}_t](\cdot + \Lambda^\epsilon) - E[G]\}\,\Xi_1^{0,\epsilon}(t)d\overline{w}_1^\epsilon(t) \\
=\;& G^2 - \left(\overline{E}^\epsilon[G]\right)^2 - 2\int_0^T \left(E[G|\mathcal{F}_t]E[\partial G(\cdot,(t,T])|\mathcal{F}_t]\right)(\cdot + \Lambda_1^\epsilon)d\overline{w}_1^\epsilon(t),
\end{aligned}$$

the Clark formula, and the chain rule for differentiation. In particular, it holds that $\Xi_1^{0,\epsilon}, \partial_{\epsilon'}^2\Xi_1^{0,\epsilon} \in \mathbf{H}_T^\infty$. Therefore, in the BSDE for $\left(\overline{Y}^{\epsilon,n}, \overline{\Xi}^{n,\epsilon}\right)$:

$$\begin{aligned}
(17)\qquad d\overline{Y}_t^{\epsilon,n} &= \left\{g\left(t, \overline{\Xi}_t^{\epsilon,n}, \epsilon\right) + R_t^{\epsilon,n}\right\}dt + \overline{\Xi}_t^{\epsilon,n}d\overline{w}_t^\epsilon, \quad \overline{Y}_T^{\epsilon,n} = \gamma F, \\
\text{where}\quad R_t^{\epsilon,n} &:= \sum_{i=0}^{2n+1} g_i\left(t, \left(\partial_{\epsilon'}^j\Xi_t^0\right)_{j=0,\dots,i}, 0\right)\frac{\epsilon^i}{i!} - g\left(t, \overline{\Xi}_t^{\epsilon,n}, \epsilon\right),
\end{aligned}$$

$\|R^{\epsilon,1}\|_{L^\infty([0,T],\Omega)} = O(\epsilon^4)$ is satisfied because of the boundedness of λ^ϵ, $\partial_{\epsilon'}^i d_0^\perp$, and $\partial_{\epsilon'}^i\Xi^{0,\epsilon}$ $(i = 0,\dots,3)$.

Next, we introduce the linear BSDE for $(\Delta Y^{\epsilon,n}, \Delta\Xi^{\epsilon,n}) := \left(Y^\epsilon - \overline{Y}^{\epsilon,n}, \Xi^\epsilon - \overline{\Xi}^{\epsilon,n}\right)$, described as

$$\begin{cases}
d\Delta Y_t^{\epsilon,n} = \left\{-\frac{1}{2}\left(\Xi_t^\epsilon + \overline{\Xi}_t^{\epsilon,n}, d_\epsilon^\perp\right)\left(\Delta\Xi_t^{\epsilon,n}, d_\epsilon^\perp\right) - R_t^{\epsilon,n}\right\}dt + \Delta\Xi_t^{\epsilon,n}d\overline{w}_t^\epsilon, \\
\Delta Y_T^{\epsilon,n} \equiv 0
\end{cases}$$

to observe the expression:

$$(18)\qquad -\Gamma_s\Delta Y_s^{\epsilon,n} = -\Gamma_t\Delta Y_t^{\epsilon,n} - \int_s^t \Gamma_u R_u^{\epsilon,n}du + M_t - M_s$$

for $0 \le s \le t \le T$, where $\Gamma := (\Gamma_t)_{t\in[0,T]}$ is the solution of the SDE:

$$d\Gamma_t = \Gamma_t\left\{\frac{1}{2}\left(\Xi_t^\epsilon + \overline{\Xi}_t^{\epsilon,n}, d_\epsilon^\perp\right)\left(d_\epsilon^\perp\right)' d\overline{w}_t^\epsilon\right\}, \quad \Gamma_0 = 1$$

and $M := (M_t)_{t\in[0,T]}$ is the $\overline{P}^\epsilon$-local-martingale defined by

$$M_t := \int_0^t \Gamma_u\left\{\Delta\Xi_u^{\epsilon,n} + \frac{1}{2}\Delta Y_u^{\epsilon,n}\left(\Xi_u^\epsilon + \overline{\Xi}_u^{\epsilon,n}, d_\epsilon^\perp\right)d_\epsilon^\perp\right\}' d\overline{w}_u^\epsilon.$$

Let $n = 1$. For a sequence of increasing stopping times $(\tau_m)_{m\in\mathbf{N}}$, which localizes the local martingale M, we deduce the relation

$$\Gamma_{t\wedge\tau_m}\left|\Delta Y^{\epsilon,1}_{t\wedge\tau_m}\right| \le \overline{E}^{\epsilon}\left[\Gamma_{T\wedge\tau_m}\left|\Delta Y^{\epsilon,1}_{T\wedge\tau_m}\right| + \epsilon^4 C_1\int_{t\wedge\tau_m}^{T\wedge\tau_m}\Gamma_u du \,\Big|\, \mathcal{F}_{t\wedge\tau_m}\right]$$

with some constant $C_1 > 0$ from (18). The first term of the right-hand-side is

$$\le \overline{E}^{\epsilon}\left[\Gamma_{T\wedge\tau_m} \mid \mathcal{F}_{t\wedge\tau_m}\right]\left\|\Delta Y^{\epsilon,1}_{T\wedge\tau_m}\right\|_{L^\infty(\Omega)} \le \Gamma_{t\wedge\tau_m}\left\|\Delta Y^{\epsilon,1}_{T\wedge\tau_m}\right\|_{L^\infty(\Omega)} \to 0$$

as $m\to\infty$ by using the optional stopping theorem, and the second term of the right-hand-side is

$$= \epsilon^4 C_1\overline{E}^{\epsilon}\left[\int_{t\wedge\tau_m}^{T\wedge\tau_m}\Gamma_u du \,\Big|\, \mathcal{F}_{t\wedge\tau_m}\right] \to \epsilon^4 C_1\overline{E}^{\epsilon}\left[\int_t^T \Gamma_u du \,\Big|\, \mathcal{F}_t\right] \le \epsilon^4 C_1 T\Gamma_t$$

as $m\to\infty$ for a continuous version of $\overline{E}^{\epsilon}[\int_\cdot^T \Gamma_u du|\mathcal{F}_\cdot]$ by using the monotone convergence theorem. Therefore, $\|\Delta Y^{\epsilon,1}\|_{L^\infty([0,T]\times\Omega)} = O(\epsilon^4)$ follows.

Finally, define the process $\overline{\nu}^{\epsilon,n} := (\overline{\nu}^{\epsilon,n})_{t\in[0,T]}$ by

$$\overline{\nu}^{\epsilon,n}_t := (I - \Pi_{\sigma' C})\left(\overline{\Xi}^{\epsilon,n}_t + \lambda^{\epsilon}_t\right) \tag{19}$$

to deduce

$$\begin{aligned}\gamma F &= \overline{Y}^{\epsilon,n}_0 + \int_0^T \left(e^{rT}\gamma\sigma'\overline{\pi}^{\epsilon,n}_t - \lambda^{\epsilon}_t + \overline{\nu}^{\epsilon,n}_t\right)' d\overline{w}^{\epsilon}_t, \\ &\quad + \int_0^T\left(\frac{|\lambda^{\epsilon}_t|^2 - |\overline{\nu}^{\epsilon,n}_t|^2}{2} + R^{\epsilon,n}_t\right)dt\end{aligned}$$

from (16-7) and (19). Therefore, for $x\in\mathbf{R}$, we obtain that

$$\begin{aligned}F + I_\gamma\left(\overline{\mathcal{Y}}^{\epsilon,n}(x) Z^{\overline{\nu}^{\epsilon,n}}_T\right) &= X^{x,\overline{\pi}^{\epsilon,n}}_T + \int_0^T R^{\epsilon,n}_t dt, \\ \text{where}\quad \overline{\mathcal{Y}}^{\epsilon,n}(x) &= \exp\left(\overline{Y}^{\epsilon,n}_0 - \gamma e^{rT}x\right),\end{aligned}$$

which implies

$$\begin{aligned}
& \log E\left[U_\gamma\left(-F + X_T^{x,\overline{\pi}^{\epsilon,1}}\right)\right] \\
= & \log E\left[U_\gamma\left(I_\gamma\left(\overline{\mathcal{Y}}^{\epsilon,1}(x)Z_T^{\overline{\nu}^{\epsilon,1}}\right) - \int_0^T R_t^{\epsilon,1}dt\right)\right] \\
= & -\frac{1}{\gamma}\overline{\mathcal{Y}}^{\epsilon,1}(x) + O(\epsilon^4) \\
= & \log U_\gamma\left(e^{rT}x - \frac{\overline{Y}_0^{\epsilon,1}}{\gamma}\right) + O(\epsilon^4) \\
= & \log U_\gamma\left(e^{rT}x - \frac{Y_0^{\epsilon}}{\gamma}\right) + O(\epsilon^4) \quad \text{as } \epsilon \downarrow 0. \blacksquare
\end{aligned}$$

Remark 4. For the higher order terms, the following is observed, for example:

$$\begin{aligned}
\partial_\epsilon^4 Y_t^{0,\epsilon} &= -12\overline{E}^\epsilon\left[\int_t^T \Xi_1^{0,\epsilon}(u)\partial_{\epsilon'}^2\Xi_1^{0,\epsilon}(u)du \;\middle|\; \mathcal{F}_t\right], \\
\partial_{\epsilon'}^4 Y_0^{0,\epsilon} - \int_0^t \partial_{\epsilon'}^4 \Xi_1^{0,\epsilon}(u)d\overline{w}_1^\epsilon(u) & \\
&= -12\overline{E}^\epsilon\left[\int_0^T \Xi_1^{0,\epsilon}(u)\partial_{\epsilon'}^2\Xi_1^{0,\epsilon}(u)du \;\middle|\; \mathcal{F}_t\right],
\end{aligned}$$

$\partial_{\epsilon'}^4\Xi_2^{0,\epsilon}(t) = 0$, and $\partial_{\epsilon'}^5\Xi^{0,\epsilon}(t) = 0$ for all $t \in [0,T]$. So, if we assume $\partial_{\epsilon'}^4\Xi^{0,\epsilon} \in \mathbf{H}_T^\infty$, then

$$\begin{aligned}
\left\|Y^\epsilon - \overline{Y}^{\epsilon,2}\right\|_{L^\infty([0,T]\times\Omega)} &= O(\epsilon^6) \\
\text{and} \quad \log V^\epsilon(x) - \log E\left[U_\gamma\left(-F + X_T^{x,\overline{\pi}^{\epsilon,2}}\right)\right] &= O(\epsilon^6)
\end{aligned}$$

are deduced as $\epsilon \downarrow 0$.

Example: European put option case. Let μ_1, μ_2 be constant and set $F = (k - S_T^1)^+$ $(k > 0)$. Then, Assumption 1 is satisfied, and we have that

$$
\begin{aligned}
Y_t^{0,\epsilon} &= \gamma\left\{k\Phi(-c_t^-) - e^{\eta^\epsilon(T-t)}S_t^1\Phi(-c_t^+)\right\} - \frac{\overline{\lambda}^2(T-t)}{2}, \\
\partial_{\epsilon'}^2 Y_t^{0,\epsilon} &= \gamma^2\left\{k^2\Phi(-c_t^-) - 2ke^{\eta^\epsilon(T-t)}S_t^1\Phi(-c_t^+)\right. \\
&\qquad \left. + e^{(2\eta^\epsilon+\sigma_1^2)(T-t)}(S_t^1)^2\Phi(-c_t^{++})\right\}, \\
\Xi_1^{0,\epsilon}(t) &= -\gamma\sigma_1 e^{\eta^\epsilon(T-t)}\Phi(-c_t^+)S_t^1, \\
\partial_{\epsilon'}^2 \Xi_1^{0,\epsilon}(t) &= 2\gamma^2\sigma_1\left\{-ke^{\eta^\epsilon(T-t)}S_t^1\Phi(-c_t^+)\right. \\
&\qquad \left. + e^{(2\eta^\epsilon+\sigma_1^2)(T-t)}(S_t^1)^2\Phi(-c_t^{++})\right\} \\
&\quad + 2\gamma\sigma_1 e^{\eta^\epsilon(T-t)}\Phi(-c_t^+)S_t^1 \\
&\quad \times\left[\gamma\left\{k\Phi(-c_t^-) - e^{\eta^\epsilon(T-t)}S_t^1\Phi(-c_t^+)\right\} - \frac{\overline{\lambda}^2(T-t)}{2}\right],
\end{aligned}
$$

where

$$
\begin{aligned}
\eta^\epsilon &:= \mu_1 - \sqrt{1-\epsilon^2}\sigma_1\sigma_2^{-1}(\mu_2 - r), \\
\Phi(d) &:= \int_{-\infty}^{d}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx, \\
c_t^- &:= \frac{1}{\sigma_1\sqrt{T-t}}\left\{\log\left(\frac{S_t^1}{k}\right) + \left(\eta^\epsilon - \frac{\sigma_1^2}{2}\right)(T-t)\right\}, \\
c_t^+ &:= c_t^- + \sigma_1\sqrt{T-t}, \quad \text{and} \quad c_t^{++} := c_t^- + 2\sigma_1\sqrt{T-t}.
\end{aligned}
$$

References

[1] Davis, M.H.A. (2000). Optimal Hedging with Basis Risk, preprint, Imperial College, London, URL: `http://www.ma.ic.ac.uk/~mdavis/`

[2] Delbaen, F., P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer, and C. Stricker (2002). Exponential Hedging and Entropic Penalties, *Mathematical Finance*, **12**(1), 99–124.

[3] El Karoui, N., S. Peng, and M.C. Quenez (1997). Backward Stochastic Differential Equations in Finance, *Mathematical Finance*, **7**(1), 1–71.

[4] Kabanov, Y. and C. Stricker (2002). On the Optimal Portfolio for the Exponential Utility Maximization: Remarks to the Six-Author, *Mathematical Finance*, **12**(1), 125–134.

[5] Karatzas, I. and S. Shreve (1998). "Methods of Mathematical Finance," Springer.

[6] KOBYLANSKI, M. (2000). Backward Stochastic Differential Equations and Partial Differential Equations with Quadratic Growth, *Annals of Probability*, **28**(2), 558–602.

[7] ROUGE, R. AND N. EL KAROUI (2000). Pricing via Utility Maximization and Entropy, *Mathematical Finance.* **10** (2), 259–276.

Graduate School of Engineering Science,
Osaka University,
Toyonaka, Osaka 560-8531,
Japan.

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 301–317

Orlicz Norm Equivalence for the Ornstein-Uhlenbeck Operator

Ichiro Shigekawa

Dedicated to Professor Kiyosi Itô on the occasion of his 88th birthday

Abstract.

The Meyer equivalence on an abstract Wiener space states that the L^p-norm of square root of the Ornstein-Uhlenbeck operator is equivalent to L^p-norm of the Malliavin derivative. We prove the equivalence in the framework of Orlicz space. We also discuss the logarithmic Sobolev inequality in L^p setting and higher order logarithmic Sobolev inequality.

§1. Introduction

Let (B, H, μ) be an abstract Wiener space: B is a separable real Banach space, H is a separable real Hilbert space which is embedded densely and continuously in B and μ is a Gaussian measure with

$$\int_B \exp\{\sqrt{-1}\,{}_{B^*}\langle l, x\rangle_B\}\mu(dx) = \exp\Big\{-\frac{1}{2}|l|_{H^*}^2\Big\}, \quad l \in B^* \hookrightarrow H^*.$$

On an abstract Wiener space, the Ornstein-Uhlenbeck semigroup is defined as

$$T_t f(x) = \int_B f(e^{-t}x + \sqrt{1-e^{-2t}}y)\mu(dy). \tag{1}$$

The generator of the semigroup $\{T_t\}$ is called the Ornstein-Uhlenbeck operator and we denote it by L. Then the following Meyer equivalence is well-known: for any $1 < p < \infty$, there exists positive constants C_1 and C_2 such that

$$C_1\{\|Df\|_p + \|f\|_p\} \le \|\sqrt{1-L}f\|_p \le C_2\{\|Df\|_p + \|f\|_p\}. \tag{2}$$

Received March 20, 2003.
Partially supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Scientific Research (A) 14204008, 2002.

Here D is the Malliavin H-derivation and $\| \ \|_p$ is the L^p-norm. The constants C_1 and C_2 depend only on p.

In this paper we show that similar inequalities hold in the framework of Orlicz space, i.e., the above inequalities hold for the Orlicz norm in place of L^p-norm. Typical example we are in mind is $L^p \log^\beta L$. As an application, we discuss the logarithmic Sobolev inequality in L^p setting and higher order logarithmic Sobolev inequality.

§2. Orlicz space

In this section, we review the Orlicz space (see, e.g., [1] or [8] for details). First we need the notion of Young function. Young function is a function Φ defined as

$$\Phi(x) = \int_0^x \phi(t)dt, \quad x \geq 0 \tag{3}$$

where ϕ is a non-negative, right continuous, non-decreasing function. If, in addition, ϕ satisfies $\phi(0) = 0$, $\phi(t) > 0$ for $t > 0$, $\phi(\infty) = \infty$, then Φ is called a nice Young function or N-function. Define ψ by

$$\psi(u) = \inf\{t\,;\, \phi(t) > u\}.$$

ψ is right continuous and non-decreasing. The function Ψ defined by

$$\Psi(y) = \int_0^y \psi(u)du, \quad y \geq 0$$

is called a complementary function. The following properties are fundamental.

$$xy \leq \Phi(x) + \Psi(y), \tag{4}$$

$$x\phi(x) = \Phi(x) + \Psi(\phi(x)). \tag{5}$$

(4) is called the Young inequality.

The Orlicz space associated with Φ is defined as follows. Let (M, m) be a measure space and Φ be a nice Young function. Define a norm $\| \ \|_\Phi$ by

$$\|f\|_\Phi := \inf\{\lambda > 0\,;\, \int_M \Phi(|f|/\lambda)dm \leq 1\}. \tag{6}$$

$L^\Phi(m)$ is the set of all measurable functions f which satisfy $\|f\|_\Phi < \infty$. We call $L^\Phi(m)$ an Orlicz space. It is a Banach space with the norm $\| \ \|_\Phi$. If Φ satisfies the Δ_2 condition, i.e., there exists a constant C such that $\Phi(2x) \le C\Phi(x)$, then the dual space is identified with $L^{\Psi}(m)$, Ψ being the complementary function of Φ.

We introduce some classes of functions.

Definition 2.1. For non-negative constant α, we define a set of functions $L(\alpha)$, $U(\alpha)$ as follows:

(i) $\phi \in L(\alpha) \overset{\text{def}}{\Longleftrightarrow} \alpha\phi(t) \le t\phi'(t), \quad \forall t > 0.$

(ii) $\phi \in U(\alpha) \overset{\text{def}}{\Longleftrightarrow} t\phi'(t) \le \alpha\phi(t), \quad \forall t > 0.$

The following inequality for semimartingales is important in our later argument.

Let (Z_t) $(t \in [0,\infty])$ be a non-negative submartingale. We assume that (Z_t) is right continuous and has left hand limits. By the Doob-Meyer decomposition theorem, (Z_t) can be decomposed as

$$Z_t = M_t + A_t$$

where (M_t) is a martingale and (A_t) is an increasing process. We assume that (A_t) is continuous and $A_0 = 0$. If $\Phi \in U(\alpha)$, then the following inequality holds (see [4, Theorem VI.99]):

$$E[\Phi(A_\infty)] \le E[\Phi(\alpha Z_\infty)]. \tag{7}$$

Further, a generalization of the Doob's inequality also holds. It is stated as follows (see [4, Chapter VI, Section 3]). We assume that $\Phi \in L(\alpha)$ for an $\alpha > 1$. Then, setting $Z_t^* := \sup_{s \le t} Z_s$, it holds that

$$E[\Phi(Z_\infty^*)] \le E[\Phi(\alpha Z_\infty)]. \tag{8}$$

From this inequality, we can have the following maximal ergodic inequality.

$$\int_B \Phi(\sup_{t\ge 0} |T_t f(x)|)\mu(dx) \le \int_B \Phi(\alpha |f(x)|)\mu(dx). \tag{9}$$

Here $\{T_t\}$ is the Ornstein-Uhlenbeck semigroup on an abstract Wiener space (B, H, μ).

§3. Littlewood-Paley inequality

Let (B, H, μ) be an abstract Wiener space and K be a separable Hilbert space. $\{T_t\}$ is the Ornstein-Uhlenbeck semigroup on $L^p(E, \mu; K)$ defined by (1). For $\alpha > 0$, set

$$T_t^{(\alpha)} = e^{-\alpha t} T_t.$$

Then the generator of $\{T_t^{(\alpha)}\}$ is $L-\alpha$. We further define a semigroup $\{Q_t^{(\alpha)}\}$ by subordination as follows:

$$Q_t^{(\alpha)} = \int_0^\infty T_s^{(\alpha)} \lambda_t(ds) = \int_0^\infty e^{-\alpha s} T_s \lambda_t(ds).$$

Here λ_t is a probability measure on $[0,\infty)$ whose Laplace transform is given by

$$\int_0^\infty e^{-\gamma s} \lambda_t(ds) = e^{-\sqrt{\gamma} t}.$$

When $\alpha = 0$, $Q_t^{(0)}$ is simply denoted by Q_t and called the Cauchy semigroup. For $F \in L^\Phi(B,\mu;K)$, it holds that

$$\|Q_t^{(\alpha)} F\|_\Phi \le e^{-\sqrt{\alpha} t} \|F\|_\Phi$$

and $\{Q_t^{(\alpha)}\}$ is a strongly continuous semigroup on L^Φ. The generator will be denoted by $-\sqrt{\alpha - L}$.

We denote by $\mathcal{P}(K)$ a set of all functions $f\colon B \to K$ which can be expressed as

$$f(x) = \sum_i p_i(\langle l_1, x\rangle), \dots, \langle l_n, x\rangle) k_i$$

where p_i is a polynomial on $\mathbb{R}^n$ and $k_1, \dots, k_n \in K$, $l_1, \dots, l_n \in B^*$.

For $f \in \mathcal{P}(K)$, define

$$\begin{aligned} g^{\rightarrow} f(x,t) &= |\partial_t Q_t^{(\alpha)}(x,f)|_K, \\ g^{\uparrow} f(x,t) &= |D Q_t^{(\alpha)}(x,f)|_{\mathrm{HS}}, \\ g f(x,t) &= \sqrt{g^{\rightarrow} f(x,t)^2 + g^{\uparrow} f(x,t)^2}. \end{aligned}$$

Here $Q_a^{(\alpha)}(x,f) = Q_a^{(\alpha)} f(x)$ and the norm $|\ |_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm. $g^{\rightarrow} f$, $g^{\uparrow} f$, gf all depend on α but we fix it throughout the argument and suppress it for simplicity. We further define

$$\begin{aligned} G^{\rightarrow} f(x) &= \left\{ \int_0^\infty t g^{\rightarrow} f(x,t)^2 dt \right\}^{1/2}, \\ G^{\uparrow} f(x) &= \left\{ \int_0^\infty t g^{\uparrow} f(x,t)^2 dt \right\}^{1/2}, \\ G f(x) &= \left\{ \int_0^\infty t g f(x,t)^2 dt \right\}^{1/2}. \end{aligned}$$

We call them Littlewood-Paley G-functions.

Our aim in this section is to prove the following theorem.

Theorem 3.1. *Assume that $\Phi \in L(\alpha) \cap U(\beta)$ for constants $1 < \alpha < \beta$. Further assume that ϕ is either convex or concave. Then we have*

$$\|\Phi(Gf)\|_1 \lesssim \|\Phi(|f|)\|_1, \tag{10}$$

$$\|\Phi(|f|)\|_1 \lesssim \|\Phi(G^{\rightarrow}f)\|_1. \tag{11}$$

In the above theorem, $A \lesssim B$ stands for $A \leq CB$ for a positive constant C that is independent of f. We use this convention in the sequel without mentioning.

We give a probabilistic proof. To do this, we take the Ornstein-Uhlenbeck process (X_t) on B, i.e., the diffusion process generated by L. We also take a process (B_t) on $\mathbb{R}$ generated by $\frac{d^2}{da^2}$. We assume that the initial distribution of (X_t) is the stationary measure μ so that the process becomes stationary. We denote the starting point of the Brownian motion (B_t) by N. E_N stands for the integration with respect to this measure. Later we let $N \to \infty$.

Now, for $f \in \mathcal{P}(K)$, set $u(x,a) = Q_a^{(\alpha)}(x,f)$. Then $u(x,a)$ satisfies

$$\begin{cases} u(x,0) = f(x) \\ L_x u(x,a) + \partial_a^2 u(x,a) - \alpha u(x,a) = 0. \end{cases} \tag{12}$$

Define a stopping time τ by

$$\tau = \inf\{t > 0 \mid B_t = 0\}.$$

Then we can think of $u(X_t, B_t)$ for $t \leq \tau$. Set

$$\begin{aligned} M_t &= Q_{B_{t\wedge\tau}}(X_{t\wedge\tau}, f) - \alpha \int_0^{t\wedge\tau} Q_{B_s}(X_s, f)ds \\ &= u(X_{t\wedge\tau}, B_{t\wedge\tau}) - \alpha \int_0^{t\wedge\tau} Q_{B_s}(X_s, f)ds. \end{aligned}$$

Then (M_t) is a martingale with $M_0 = Q_{B_0}f(X_0)$. The quadratic variation is given as

$$\langle M\rangle_t = 2\int_0^{t\wedge\tau} gf^2(X_s, B_s)ds. \tag{13}$$

Therefore we have

$$\begin{aligned} d|u|^2 &= 2(u, dM) + 2\alpha|u|^2 dt + \langle dM, dM\rangle \\ &= 2(u, dM) + (2\alpha|u|^2 + 2gf^2)dt \end{aligned} \tag{14}$$

Now set

$$Z_t = |u(X_{t\wedge T}, B_{t\wedge T})|^2.$$

(Z_t) is a non-negative submartingale. To compute $\Phi(\sqrt{Z_t})$, we approximate it as follows. Take any $\varepsilon > 0$ and set $F(x) = \Phi(\sqrt{x+\varepsilon})$. Recall that ϕ is either convex or concave. We divide into tow cases.

(i) ϕ is concave.

We need the following proposition.

Proposition 3.2. *Assume $\Phi \in L(\alpha)$ $(\alpha > 1)$. Then it holds that, for u, $v \geq 0$,*

$$\Phi(v) \leq \frac{1}{\alpha-1}\Big(\frac{1}{2}\phi'(u)v^2 + \Phi(u)\Big). \tag{15}$$

Proof. From the assumption, $\alpha\Phi(x) \leq x\phi(x)$ holds. Since ϕ is concave, $\Phi(x) \geq \frac{1}{2}x\phi(x)$ which leads to $\alpha \leq 2$. Hence (15) clearly holds when $u \geq v$.

If $v \geq u$, we have

$$|\{(x,y)\,;\, 0 \leq x \leq u,\, \phi(x) \leq y \leq \phi(u)\}| \leq \frac{1}{2}u\phi(u)$$
$$|\{(x,y)\,;\, 0 \leq x \leq u,\, \phi(u) \leq y \leq \phi(v)\}| \leq u(\phi(v) - \phi(u)) \leq u\phi'(u)(v-u)$$
$$|\{(x,y)\,;\, u \leq x \leq v,\, \phi(x) \leq y \leq \phi(v)\}| \leq \frac{1}{2}(v-u)^2\phi'(u).$$

These are easily obtained by observing the graph.

Summing up three terms of the left-hand side and $\Phi(v)$, we have $v\phi(v)$. Therefore

$$\frac{1}{2}u\phi(u) + u\phi'(u)(v-u) + \frac{1}{2}(v-u)^2\phi'(u) + \Phi(v) \geq v\phi(v) \geq \alpha\Phi(v).$$

Hence we have

$$\begin{aligned}(\alpha-1)\Phi(v) &\leq \frac{1}{2}u\phi(u) + \phi'(u)(v-u)\Big(u + \frac{1}{2}v - \frac{1}{2}u\Big)\\ &= \frac{1}{2}u\phi(u) + \frac{1}{2}\phi'(u)(v^2 - u^2)\\ &\leq \Phi(u) + \frac{1}{2}\phi'(u)v^2\end{aligned}$$

which is the desired result. Q.E.D.

The derivatives of $F(x) = \Phi(\sqrt{x+\varepsilon})$ are

$$F'(x) = \Phi'(\sqrt{x+\varepsilon})\frac{1}{2\sqrt{x+\varepsilon}},$$

$$F''(x) = \Phi''(\sqrt{x+\varepsilon})\frac{1}{4(x+\varepsilon)} + \Phi'(\sqrt{x+\varepsilon})\frac{1}{2}\Big(-\frac{1}{2}\Big)\frac{1}{\sqrt{x+\varepsilon}^3}.$$

By the Itô formula,

$$\begin{aligned}
d\Phi(\sqrt{Z_t+\varepsilon}) &= \frac{\Phi'(\sqrt{Z_t+\varepsilon})}{2\sqrt{Z_t+\varepsilon}}dZ_t \\
&\quad + \frac{1}{2}\left\{\frac{\Phi''(\sqrt{Z_t+\varepsilon})}{4(Z_t+\varepsilon)} - \frac{1}{4}\frac{\Phi'(\sqrt{Z_t+\varepsilon})}{\sqrt{Z_t+\varepsilon}^3}\right\}\langle dZ, dZ\rangle \\
&= \frac{\phi(\sqrt{|u|^2+\varepsilon})}{2\sqrt{|u|^2+\varepsilon}}\{2(u,dM) + 2(\alpha|u|^2+gf^2)dt\} \\
&\quad + \frac{1}{2}\left\{\frac{\phi'(\sqrt{|u|^2+\varepsilon})}{4(|u|^2+\varepsilon)} - \frac{1}{4}\frac{\phi(\sqrt{|u|^2+\varepsilon})}{\sqrt{|u|^2+\varepsilon}^3}\right\}\langle dZ, dZ\rangle \\
&= \frac{\phi(\sqrt{|u|^2+\varepsilon})}{\sqrt{|u|^2+\varepsilon}}(u,dM) + \frac{\phi(\sqrt{|u|^2+\varepsilon})}{\sqrt{|u|^2+\varepsilon}}(\alpha|u|^2+gf^2)dt \\
&\quad + \frac{1}{8}\frac{1}{|u|^2+\varepsilon}\left\{\phi'(\sqrt{|u|^2+\varepsilon}) - \frac{\phi(\sqrt{|u|^2+\varepsilon})}{\sqrt{|u|^2+\varepsilon}}\right\}\langle dZ, dZ\rangle.
\end{aligned}$$

Now we note $\langle dZ, dZ\rangle \le 4|u|^2\langle dM, dM\rangle = 8|u|^2 gf^2 dt$. Further $\phi'(t) \le \phi(t)/t$ since ϕ is concave. We therefore have

$$\begin{aligned}
&\frac{\phi(\sqrt{|u|^2+\varepsilon})}{\sqrt{|u|^2+\varepsilon}}gf^2dt + \frac{1}{8}\frac{1}{|u|^2+\varepsilon}\left\{\phi'(\sqrt{|u|^2+\varepsilon}) - \frac{\phi(\sqrt{|u|^2+\varepsilon})}{\sqrt{|u|^2+\varepsilon}}\right\}\langle dZ, dZ\rangle \\
&\ge \frac{\phi(\sqrt{|u|^2+\varepsilon})}{\sqrt{|u|^2+\varepsilon}}gf^2dt + \left\{\phi'(\sqrt{|u|^2+\varepsilon}) - \frac{\phi(\sqrt{|u|^2+\varepsilon})}{\sqrt{|u|^2+\varepsilon}}\right\}gf^2dt \\
&= \phi'(\sqrt{|u|^2+\varepsilon})gf^2dt.
\end{aligned}$$

Integrating from 0 to τ and taking expectation, we have

$$\|\Phi(\sqrt{|f|^2+\varepsilon})\|_1 \ge E_N\left[\int_0^\tau \phi'(\sqrt{|u|^2+\varepsilon})gf^2dt\right]. \tag{16}$$

We will give an estimate from below of the right-hand side. We note that $f^*(x) := \sup_{t\ge 0}|T_tf(x)| \ge \sup_{a\ge 0}|u(x,a)|$.

$$\begin{aligned}
E_N\left[\int_0^\tau \phi'(\sqrt{|u|^2+\varepsilon})gf^2dt\right] &= \left\|\int_0^\infty \phi'(\sqrt{|u|^2+\varepsilon})gf^2(\,\cdot\,,a)(a\wedge N)da\right\|_1 \\
&\ge \left\|\int_0^\infty \phi'(\sqrt{f^{*2}+\varepsilon})gf^2(\,\cdot\,,a)(a\wedge N)da\right\|_1.
\end{aligned}$$

($\because$ ϕ' is non-increasing)

Combining this with (16) and letting $N \to \infty$

$$\begin{aligned}\|\Phi(\sqrt{|f|^2+\varepsilon})\|_1 &\geq \left\|\int_0^\infty \phi'(\sqrt{f^{*2}+\varepsilon})gf^2(\cdot\,,a)ada\right\|_1\\ &= \|\phi'(\sqrt{f^{*2}+\varepsilon})Gf^2\|_1\end{aligned}$$

Now we use the inequality $\Phi(v) \leq \frac{1}{\alpha-1}(\frac{1}{2}\phi'(u)v^2 + \Phi(u))$ in Proposition 3.2 and get

$$\begin{aligned}\|\Phi(Gf)\|_1 &\lesssim \|\phi'(\sqrt{f^{*2}+\varepsilon})Gf^2\|_1 + \|\Phi(\sqrt{f^{*2}+\varepsilon})\|_1\\ &\leq \|\Phi(\sqrt{|f|^2+\varepsilon})\|_1 + \|\Phi(\sqrt{f^{*2}+\varepsilon})\|_1.\end{aligned}$$

Letting $\varepsilon \to 0$ and using the maximal ergodic inequality (9), we have

$$\|\Phi(Gf)\|_1 \lesssim \|\Phi(|f|)\|_1 + \|\Phi(f^*)\|_1 \lesssim \|\Phi(|f|)\|_1.$$

This completes the proof in the case that ϕ is concave.

<u>(ii) ϕ is convex.</u>

Set $\tilde{\Phi}(x) = \Phi(\sqrt{x})$. Then $\tilde{\Phi}$ is convex. In fact, by differentiating, we have

$$\frac{d}{dx}\Phi(\sqrt{x}) = \Phi'(\sqrt{x})\frac{1}{2\sqrt{x}} = \frac{\phi(\sqrt{x})}{2\sqrt{x}}.$$

The function is increasing since ϕ is convex and so the convexity of $\tilde{\Phi}$ follows. Further $\tilde{\Phi} \in U(\alpha/2)$ since

$$\frac{x\tilde{\Phi}'(x)}{\tilde{\Phi}(x)} = \frac{x\Phi'(\sqrt{x})}{2\sqrt{x}\Phi(\sqrt{x})} = \frac{\sqrt{x}\Phi'(\sqrt{x})}{2\Phi(\sqrt{x})}.$$

The submartingale $Z_t = |u(X_{t\wedge\tau}, B_{t\wedge\tau})|^2$ is decomposed as a sum of a martingale and an increasing process as in (14). By using (7), we get

$$\begin{aligned}(17)\quad E_N[\tilde{\Phi}(\int_0^\tau gf(X_s,B_s)^2ds)] &\lesssim E_N[\tilde{\Phi}(Z_\infty)] = E_N[\tilde{\Phi}(|f(X_\tau)|^2)]\\ &= E_N[\Phi(|f(X_\tau)|)] = \|\Phi(|f|)\|_1\end{aligned}$$

Now we introduce H-functions as follows.

$$H^{\rightarrow}f(x) = \left\{\int_0^\infty tQ_tg^{\rightarrow}f(x,t)^2dt\right\}^{1/2},$$

$$H^\uparrow f(x) = \left\{ \int_0^\infty tQ_t g^\uparrow f(x,t)^2 dt \right\}^{1/2},$$
$$Hf(x) = \left\{ \int_0^\infty tQ_t gf(x,t)^2 dt \right\}^{1/2}.$$

Then we have

$$\begin{aligned}
\|\Phi(Hf)\|_1 &= \|\tilde{\Phi}(Hf^2)\|_1 \\
&= \lim_{N\to\infty} \left\| \tilde{\Phi}\left(\int_0^\infty Q_a gf(\,\cdot\,,a)^2 (a \wedge N) da \right) \right\|_1 \\
&= \lim_{N\to\infty} \int_B \tilde{\Phi}\left(E_N \left[\int_0^\tau gf^2(X_t, B_t) dt \middle| X_\tau = x \right] \right) \mu(dx) \\
&\le \lim_{N\to\infty} \int_B E_N \left[\tilde{\Phi}\left(\int_0^\tau gf^2(X_t, B_t) dt \right) \middle| X_\tau = x \right] \mu(dx) \\
&\le \lim_{N\to\infty} E_N \left[\tilde{\Phi}\left(\int_0^\tau gf^2(X_t, B_t) dt \right) \right] \\
&\lesssim \|\Phi(|f|)\|_1. \quad (\because\ (17))
\end{aligned}$$

It is well-known that Gf is dominated by Hf (see [7]) and so (10) follows. (11) can be shown by a standard duality argument. This completes the proof of Theorem 3.1.

Using this theorem, the Meyer equivalence in Orlicz space, which is of our main interest, follows easily. In fact, the same proof as in L^p setting works (see e.g., [9]).

Theorem 3.3. *Assume that $\Phi \in L(\alpha) \cap U(\beta)$ for $1 < \alpha < \beta$ and that ϕ is either convex or concave. Then there exist positive constants C_1 and C_2 such that*

$$C_1\{\|Df\|_\Phi + \|f\|_\Phi\} \le \|\sqrt{1-L}f\|_\Phi \le C_2\{\|Df\|_\Phi + \|f\|_\Phi\}. \tag{18}$$

§4. Examples

We give some example of nice Young functions that satisfy the condition of Theorem 3.3. For indicies $p > 1$, $\beta \in \mathbb{R}$, $k \ge 1$, we set

$$\phi_{p,\beta,k}(x) = x^{p-1} \log^{p\beta}(k+x), \quad x \ge 0 \tag{19}$$

and define

$$\Phi_{p,\beta,k}(x) = \int_0^x \phi_{p,\beta,k}(y) dy. \tag{20}$$

We regards this as a Young function. The function does not satisfy the condition of Young function in general since β might be negative. We see when it is a Young function. To avoid complexity, we simply denote ϕ and Φ in place of $\phi_{p,\beta,k}$ and $\Phi_{p,\beta,k}$, respectively. Differentiating ϕ, we have

$$\begin{aligned}\phi'(x) &= (p-1)x^{p-2}\log^{p\beta}(k+x) + p\beta x^{p-1}\{\log^{p\beta-1}(k+x)\}\frac{1}{k+x} \\ &= x^{p-2}\log^{p\beta}(k+x)\left\{p-1+p\beta\frac{x}{(k+x)\log(k+x)}\right\}.\end{aligned}$$

We look for the condition so that ϕ' is positive. To do this, set

$$f(x) = \frac{x}{(k+x)\log(k+x)}. \tag{21}$$

If $k=1$, f takes its maximum 1 at $x=0$. If $k>1$, f takes its maximum at $x=\alpha$ where α is the solution of $k\log(k+x)-x=0$. We can see that $f(\alpha) \leq \frac{1}{1+\log k}$. Therefore, in all cases of k, it holds that

$$0 \leq \frac{x}{(k+x)\log(k+x)} \leq \frac{1}{1+\log k}. \tag{22}$$

Now it is easy to see that Φ is a nice Young function if $p\left(1+\frac{\beta}{1+\log k}\right) \geq 1$. Further we easily have the following proposition.

Proposition 4.1. *ϕ satisfies following inequalities:*

$$(p-1)\phi(x) \leq x\phi'(x) \leq \left(p-1+\frac{p\beta}{1+\log k}\right)\phi(x), \quad \textit{for } \beta \geq 0, \tag{23}$$

$$\left(p-1+\frac{p\beta}{1+\log k}\right)\phi(x) \leq x\phi'(x) \leq (p-1)\phi(x), \quad \textit{for } \beta < 0. \tag{24}$$

Similar inequalities hold for Φ. To see this, we need the following proposition.

Proposition 4.2. *For positive constant α, it holds that*

(i) *if $\phi \in L(\alpha)$, then $\Phi \in L(\alpha+1)$,*
(ii) *if $\phi \in U(\alpha)$, then $\Phi \in U(\alpha+1)$.*

Proof. Suppose $\phi \in L(\alpha)$, i.e., $\alpha\phi(t) \leq t\phi'(t)$. By integrating both hands, we have

$$\alpha\Phi(x) \leq \int_0^x t\phi'(t)dt. \tag{25}$$

On the other hand, since $(t\phi(t))' = \phi(t) + t\phi'(t)$, we have

$$x\phi(x) = \int_0^x \{\phi(t) + t\phi'(t)\}dt$$

and hence

$$x\phi(x) - \Phi(x) = \int_0^x t\phi'(t)dt.$$

This combined with (25) leads

$$x\phi(x) \geq (\alpha + 1)\Phi(x).$$

(ii) can be shown similarly. Q.E.D.

Now the following proposition easily follows.

Proposition 4.3. *The following inequalities hold:*

$$p\Phi(x) \leq x\Phi'(x) \leq p\Big(1 + \frac{\beta}{1 + \log k}\Big)\Phi(x), \quad (\beta \geq 0), \tag{26}$$

$$p\Big(1 + \frac{\beta}{1 + \log k}\Big)\Phi(x) \leq x\Phi'(x) \leq p\Phi(x), \quad (\beta < 0). \tag{27}$$

Lastly, we will see the asymptotic behavior of the complementary function Ψ. We use the notation $f \sim g$ when $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$ holds.

Proposition 4.4. *Assume $p > 1$ and let q be the conjugate exponent of p: $\frac{1}{p} + \frac{1}{q} = 1$. Then it holds that*

$$p\Phi(x) \sim x^p \log^{p\beta} x, \tag{28}$$

$$(q-1)^{q\beta} q\Psi(x) \sim x^q \log^{-q\beta} x. \tag{29}$$

Proof. By the l'Hôpital theorem, we have

$$\begin{aligned}
\lim_{x\to\infty} \frac{p\Phi(x)}{x^p \log^{p\beta} x} &= \lim_{x\to\infty} \frac{p\phi(x)}{px^{p-1}\log^{p\beta} x + x^p p\beta(\log^{p\beta} x)/x} \\
&= \lim_{x\to\infty} \frac{px^{p-1}\log^{p\beta}(k+x)}{px^{p-1}\log^{p\beta} x + p\beta x^{p-1}\log^{p\beta-1} x} \\
&= \lim_{x\to\infty} \frac{\log^{p\beta}(k+x)}{\log^{p\beta} x + \beta \log^{p\beta-1} x} \\
&= 1
\end{aligned}$$

which shows (28).

As for Ψ, we have, by the l'Hôpital theorem,

$$
\begin{aligned}
&\lim_{x\to\infty} \frac{q\Psi(x)}{x^q \log^{-q\beta} x} \\
&= \lim_{x\to\infty} \frac{q\Psi(\phi(x))}{\phi(x)^q \log^{-q\beta} \phi(x)} \\
&= \lim_{x\to\infty} \frac{q\Psi'(\phi(x))\phi'(x)}{q\phi(x)^{q-1}\phi'(x)\log^{-q\beta}\phi(x) + \phi(x)^q(-q\beta)\{\log^{-q\beta-1}\phi(x)\}\phi'(x)/\phi(x)} \\
&= \lim_{x\to\infty} \frac{x}{\phi(x)^{q-1}\log^{-q\beta}\phi(x) - \beta\phi(x)^{q-1}\log^{-q\beta-1}\phi(x)} \\
&= \lim_{x\to\infty} \frac{x}{\phi(x)^{q-1}\log^{-q\beta}\phi(x)(1-\beta/\log\phi(x))} \\
&= \lim_{x\to\infty} \frac{x}{\phi(x)^{q-1}\log^{-q\beta}\phi(x)} \\
&= \lim_{x\to\infty} \frac{x}{\{x^{p-1}\log^{p\beta}(k+x)\}^{q-1}\log^{-q\beta}(x^{p-1}\log^{p\beta}(k+x))} \\
&= \lim_{x\to\infty} \frac{x}{x^{(p-1)(q-1)}\log^{p\beta(q-1)}(k+x)\{(p-1)\log x + p\beta\log\log(k+x)\}^{-q\beta}} \\
&= \lim_{x\to\infty} \frac{\{(p-1)\log x + p\beta\log\log(k+x)\}^{q\beta}}{\log^{p\beta(q-1)}(k+x)} \qquad (\because (p-1)(q-1)=1) \\
&= \lim_{x\to\infty} \left\{\frac{\{(p-1)\log x + p\beta\log\log(k+x)}{\log(k+x)}\right\}^{q\beta} \qquad (\because q = p(q-1)) \\
&= (p-1)^{q\beta}
\end{aligned}
$$

which shows (29). Q.E.D.

We denotes the Orlicz space L^Φ associated with Φ by $L^p\log^{p\beta} L$. We do not specify k since it does not affect the asymptotic behavior at infinity. Since the Wiener measure is finite, $L^p\log^{p\beta} L$ is independent of k. The above theorem means that the dual space of $L^p\log^{p\beta} L$ is $L^q\log^{-q\beta} L$.

§5. Logarithmic Sobolev inequality

The logarithmic Sobolev inequality in L^p setting was discussed by D. Bakry-P. A. Meyer [3] and higher order Logarithmic Sobolev inequality was discussed by G. F. Feissner [5] and R. A. Adams [2]. They all used the interpolation theorem. Here we take a different approach.

The following logarithmic Sobolev inequality holds for the Ornstein-Uhlenbeck process.

$$E\left[f^2 \log(f^2/\|f\|_2^2)\right] \leq 2E\left[|Df|^2\right].$$

Here $E[\,]$ stands for the integration with respect to μ. Hereafter we use this notation. Recall that $\int_B (Df, Dg)_{H^*} d\mu$ is the Dirichlet form associated with the Ornstein-Uhlenbeck process. We remark that the following argument works for the diffusion Dirichlet form satisfying the logarithmic Sobolev inequality if we assume the Dirichlet form is of the gradient type.

We introduce a new Young function. Set

$$\theta(x) = \{x^2 \log(e + x^2)\}^{(p-2)/4} \log^{p\beta/4}(k + x^2 \log(e + x)) \tag{30}$$

and define

$$\Theta(x) = \int_0^x \theta(y)dy. \tag{31}$$

Then we have the following proposition.

Proposition 5.1. *For sufficient large k if necessary, there exists a positive constant K such that*

$$\begin{aligned} x^p \log^{p/2}(e + x^2) \log^{p\beta/2}(k + x^2 \log(e + x^2)) \\ \leq K\Theta(x)^2 \log(e + \Theta(x)^2). \end{aligned} \tag{32}$$

Proof. We divide the proof into two cases.
(a) $\beta \geq 0$, $k = 1$.

Let us see the asymptotic behavior as $x \to 0$.

$$\text{LHS} \sim x^p \cdot x^{(p\beta/2)2} = x^{p(1+\beta)}.$$

On the other hand,

$$\theta(x) \sim x^{(p-2)/2+(p\beta)/2} = x^{p(\beta+1)/2-1}.$$

and hence

$$\Theta(x) \sim \frac{2}{p(\beta+1)} x^{p(\beta+1)/2}$$
$$\Theta(x)^2 \sim \frac{4}{p^2(\beta+1)^2} x^{p(\beta+1)}.$$

Thus both hands have the same asymptotic behavior.

As $x \to \infty$,

$$\text{LHS} \sim x^p 2^{p/2}(\log^{p/2} x) 2^{p\beta/2} \log^{p\beta/2} x = 2^{p(\beta+1)/2} x^p \log^{p(\beta+1)/2} x.$$

On the other hand,

$$\begin{aligned}
\theta(x) &\sim x^{(p-2)/2} 2^{(p-2)/4} (\log^{(p-2)/4} x) 2^{p\beta/4} \log^{p\beta/4} x \\
\Theta(x) &\sim (2/p) 2^{(p+p\beta-2)/4} x^{p/2} \log^{(p+p\beta-2)/4} x \\
\Theta(x)^2 \log(e+\Theta(x)^2) &\sim p^{-2} 2^{(p+p\beta+2)/2} x^p (\log^{(p+p\beta-2)/2} x) p \log x \\
&= p^{-1} 2^{(p+p\beta+2)/2} x^p \log^{(p+p\beta)/2} x.
\end{aligned}$$

Hence they have the same asymptotic behavior.
(b) $\beta < 0$ and large k.

The asymptotic behavior at $x = \infty$ can be obtained similarly as in the case $\beta \geq 0$.

As $x \to 0$, LHS $\sim x^p$ is clear. Further we have

$$\begin{aligned}
\theta(x) &\sim x^{(p-2)/2} \\
\Theta(x) &\sim \frac{2}{p} x^{p/2} \\
\Theta(x)^2 \log(e+\Theta(x)^2) &\sim \frac{4}{p^2} x^p
\end{aligned}$$

Thus we have the desired result. Q.E.D.

We recall the following fact. Let U and V be a non-negative functions on a measure space (M, m). Assume that

$$\begin{aligned}
&\int_M U\phi(U)dm < \infty, \\
&\int_M U\phi(U)dm \leq \int_M V\phi(U)dm + C.
\end{aligned}$$

Then it follows that

$$\int_M \Phi(U)dm \leq \int_M \Phi(V)dm + C. \tag{33}$$

For the proof, see [4, Lemma VI.98]. Now we have the following theorem. In the sequel, we denote by $\Phi_{p,\beta}$ in place of $\Phi_{p,\beta,k}$ because the index k is not essential.

Proposition 5.2. *For $p > 2$, $\beta \in \mathbb{R}$, there exists a positive constant C such that*

(34)
$$E[\Phi_{p,(\beta+1)/2}(|f|)] \le CE[\Phi_{p,(1+\beta)/2-(1/p)}(|f|)] + CE[\Phi_{p,\beta/2}(|Df|)].$$

Proof. Set $g = \sqrt{\Theta(|f|)^2 + e}$. Then

$$Dg = \frac{2\Theta(|f|)\Theta'(|f|)D|f|}{2\sqrt{\Theta(|f|)^2 + e}}$$

and hence $|Dg| \le \theta(|f|)|Df|$. Now, by using the logarithmic Sobolev inequality

$$E\left[g^2 \log(g^2/\|g\|_2^2)\right] \le 2E\left[|Dg|^2\right],$$

we have

$$\begin{aligned}
&E\left[\{\Theta(|f|)^2 + e\}\log(e + \Theta(|f|)^2)\right] \\
&\le E\left[\Theta(|f|)^2 + e\right]\log E\left[e + \Theta(|f|)^2\right] \\
&\quad + 2E\left[|Df|^2\{|f|^2\log(e+|f|^2)\}^{(p-2)/2}\log^{p\beta/2}(k + |f|^2\log(e+|f|^2))\right],
\end{aligned}$$

and

$$\begin{aligned}
&E\left[\Theta(|f|)^2\log(e + \Theta(|f|)^2)\right] \\
&\le E\left[\Theta(|f|)^2\right]\log E\left[e + \Theta(|f|)^2\right] \\
&\quad + 2E\left[|Df|^2\{|f|^2\log(e+|f|^2)\}^{(p-2)/2}\log^{p\beta/2}(k + |f|^2\log(e+|f|^2))\right].
\end{aligned}$$

We set

$$\begin{aligned}
\phi(x) &= \phi_{p/2,\beta,k}(x) = x^{(p/2)-1}\log^{p\beta/2}(k+x), \\
U &= |f|^2\log(e+|f|^2).
\end{aligned}$$

Then

$$\begin{aligned}
U\phi(U) &= |f|^2\log(e+|f|^2) \\
&\qquad \times \{|f|^2\log(e+|f|^2)\}^{(p/2)-1}\log^{p\beta/2}(k + |f|^2\log(e+|f|^2)) \\
&= |f|^p\log^{p/2}(e+|f|^2)\log^{p\beta/2}(k+|f|^2\log(e+|f|^2)) \\
&\le K\Theta(|f|)^2\log(e+\Theta(|f|)^2). \quad (\because\ (32))
\end{aligned}$$

Combining this with the previous result, we have

$$K^{-1}E[U\phi(U)] \le E[e + \Theta(|f|)^2]\log E[e + \Theta(|f|)^2] + 2E[|Df|^2\phi(U)].$$

Now, by (33), it follows that

$$E[\Phi(U)] \le KE[e+\Theta(|f|)^2]\log E[e+\Theta(|f|)^2] + 2KE[\Phi(|Df|^2)].$$

Here Φ is the integral of ϕ. Since $\Phi = \Phi_{p/2,\beta}$,

$$\begin{aligned}\Phi(x^2) &\le c_1 x^2\phi(x^2)\\ &\le c_1 x^2 (x^2)^{(p/2)-1}\log^{p\beta/2}(k+x^2)\\ &\le c_1 x^p \log^{p\beta/2}(k+x^2)\\ &\le c_2\Phi_{p,\beta/2}(x).\end{aligned}$$

Further

$$\begin{aligned}\Phi(x^2\log(e+x^2)) &\ge c_3 x^2\log(e+x^2)\phi(x^2\log(e+x^2))\\ &= c_3x^2\log(e+x^2)\{x^2\log(e+x^2)\}^{(p/2)-1}\\ &\quad\times \log^{p\beta/2}(k+x^2\log(e+x^2))\\ &= c_3x^p\log^{p/2}(e+x^2)\log^{p\beta/2}(k+x^2\log(e+x^2))\\ &\ge c_4x^p\log^{p/2}(e+x)\log^{p\beta/2}(k+x)\\ &\ge c_5x^p\log^{p(1+\beta)/2}(k+x)\\ &\ge c_6\Phi_{p,(\beta+1)/2}(x)\end{aligned}$$

and

$$\begin{aligned}\Theta(x)^2 &\le x^2\theta(x)^2\\ &\le x^2\{x^2\log(e+x^2)\}^{(p-2)/2}\log^{p\beta/2}(k+x^2\log(e+x))\\ &\le c_7x^p\log^{(p-2)/2}(e+x)\log^{p\beta/2}(k+x)\\ &\le c_7x^p\log^{(p+p\beta-2)/2}(k+x)\\ &\le c_8\Phi_{p,(1+\beta)/2-(1/p)}(x).\end{aligned}$$

Thus we have eventually obtained

$$E[\Phi_{p,(\beta+1)/2}(|f|)] \le CE[\Phi_{p,(1+\beta)/2-(1/p)}(|f|)] + CE[\Phi_{p,\beta/2}(|Df|)].$$

This completes the proof. Q.E.D.

If $p=2$ and $\beta \ge 0$, the above proof works as well in this case. We only state the result.

Proposition 5.3. *For $p=2$, $\beta\ge 0$, there exists a positive constant C such that*

$$(35)\qquad E[\Phi_{2,(\beta+1)/2}(|f|)] \le CE[\Phi_{2,\beta/2}(|f|)] + CE[\Phi_{p,\beta/2}(|Df|)].$$

In Section 3, we showed that the right hand side of (34) is equivalent to $E\left[\Phi_{p,\beta/2}(\sqrt{1-L}f)\right]$. Therefore we easily get the following theorem.

Theorem 5.4. *For $p>1$, $\beta\geq 0$, the following map is continuous:*

$$\sqrt{1-L}^{\,-1}: L^p\log^{p\beta}L \to L^p\log^{p(\beta+1/2)}L. \tag{36}$$

Recall that the dual space of $L^p\log^{p\beta}$ is $L^q\log^{-q\beta}L$ (see Proposition 4.4). Hence, when $1<p<2$, the above equation (36) is shown by the duality. By iterating the map $\sqrt{1-L}^{\,-1}$, we can have the continuity of $(1-L)^{-1}$ from $L^p\log^{p\beta}L$ to $L^p\log^{p(\beta+1)}L$.

References

[1] R. A. Adams, "*Sobolev spaces*," Academic press, New York, 1975.

[2] R. A. Adams, General logarithmic Sobolev inequalities and Orlicz imbeddings, *J. Funct. Anal.*, **34** (1979), 292–303.

[3] D. Bakry and P. A. Meyer, Sur les inégalités de Sobolev logarithmiques, I, II, Lecture Notes in Math., **920** (1982), 138–145, 146–150, Springer-Verlag, Berlin.

[4] C. Dellacherie and P.-A. Meyer, "*Probabilities and potential. B. Theory of martingales*," North-Holland Publishing Co., Amsterdam, 1982.

[5] G. F. Feissner, Hypercontractive semigroups and Sobolev's inequality, *Trans. Amer. Math. Soc.*, **210** (1975), 51–62.

[6] P. A. Meyer, Notes sur les processus d'Ornstein-Uhlenbeck, *Séminaire de Prob.* XVI, ed. par J. Azema et M. Yor, Lecture Notes in Math., vol. 920 (1982), 95–133, Springer-Verlag, Berlin-Heidelberg-New York.

[7] P. A. Meyer, Quelques results analytiques sur le semigroupe d'Ornstein-Uhlenbeck en dimension infinie, Theory and application of random fields, Proceedings of IFIP-WG 7/1 Working conf. at Bangalore, ed. by G. Kallianpur, Lecture Notes in Cont. and Inform. Sci., vol. 49 (1983), 201–214, Springer-Verlag, Berlin-Heidelberg-New York.

[8] M. M. Rao and Z. D. Ren, "*Theory of Orlicz spaces*," Marcel Dekker, New York, 1991.

[9] I. Shigekawa, Littlewood-Paley inequality for a diffusion satisfying the logarithmic Sobolev inequality and for the Brownian motion on a Riemannian manifold with boundary, *Osaka J. Math.*, **39** (2002), 897–930.

[10] I. Shigekawa and N. Yoshida, Littlewood-Paley-Stein inequality for a symmetric diffusion, *J. Math. Soc. Japan*, **44** (1992), 251–280.

Department of Mathematics
Graduate School of Science
Kyoto University
Kyoto, 606-8502
Japan

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 319–331

Some Comments about Itô's Construction Procedure

Daniel W. Stroock

For K. Itô on his 88th birthday

Abstract.

This article reviews Itô's procedure for constructing the Markov process generated by variable coefficient Lévy–Khinchine operators. In particular, it examines conditions under which Itô's procedure succeeds but more analytic procedures appear to fail.

§0 Introduction

In his famous memoir [1], Itô dealt with the construction of Markov processes corresponding to variable coefficient Lévy–Khinchine operators. His method rests on the ability to represent of the action of Lévy–Khinchine operator L with diffusion coefficient $x \rightsquigarrow a(x)$, drift coefficient $x \rightsquigarrow b(x)$, and Lévy measure $x \rightsquigarrow M(x,\,\cdot\,)$ on a $\varphi \in C^2_{\mathrm{c}}(\mathbb{R}^n;\mathbb{R})$ in the form

$$
\begin{aligned}
L\varphi(x) = \tfrac{1}{2}\sum_{i,j=1}^{n}\big(\sigma(x)\sigma(x)^\top\big)_{ij}\partial_i\partial_j\varphi(x) + \sum_{i=1}^{n} c(x)_i\partial_i\varphi(x) \\
+ \int_{\mathbb{R}^n\setminus\{0\}}\Big(\varphi\big(x+F(x,y)\big) - \varphi(x) \\
- \mathbf{1}_{[0,1]}\big(|y|\big)\big(F(x,y),\mathrm{grad}_x\varphi\big)_{\mathbb{R}^n}\Big)\,M(dy),
\end{aligned}
\tag{0.1}
$$

for appropriate functions $\sigma : \mathbb{R}^n \longrightarrow \mathrm{Hom}(\mathbb{R}^n;\mathbb{R}^n)$, $c : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $F : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$, and Lévy measure M. In order for his method to have a chance of working, these functions must be at least (Borel) measurable, and, in practice, they must be much better than that. Indeed, apart from refinements (cf. [6]), which are important but of restricted

Received January 7, 2003.
Research contained in this article was partially supported by NSF grant DMS #6890047.

applicability, what one needs is that σ and c be uniformly Lipschitz continuous and that F satisfy conditions of the form:

$$
\begin{gathered}
\lim_{r\searrow 0}\sup_{x\in\mathbb{R}^n}\frac{1}{1+|x|^2}\int_{0<|y|\le r}|F(x,y)|^2\,\frac{dy}{|y|^{n+1}}=0\\
\sup_{x\in\mathbb{R}^n}\frac{1}{1+|x|^2}\int_{0<|y|\le R}|F(x,y)|^2\,\frac{dy}{|y|^{n+1}}<\infty\\
\sup_{x_2\neq x_1}\frac{1}{|x_2-x_1|^2}\int_{0<|y|\le R}\left|F(x_2,y)-F(x_1,y)\right|^2\frac{dy}{|y|^{n+1}}<\infty
\end{gathered}
\tag{0.2}
$$

for each $R\in(0,\infty)$. Under these conditions it is possible to carry out (cf. §3.1 and §4.1 in [3]) Itô's procedure for constructing the Markov process corresponding to $x\rightsquigarrow\big(a(x),b(x),M(x.\,\cdot\,)\big)$ by transforming the paths of the Lévy process whose continuous part is standard Brownian motion and whose Lévy part is the symmetric Cauchy process whose Lévy measure is $M_0(dy)=\mathbf{1}_{\mathbb{R}^n\setminus\{0\}}(y)\frac{dy}{|y|^{n+1}}$.

Assuming that $x\rightsquigarrow a(x)$, $x\rightsquigarrow b(x)$, and $x\rightsquigarrow M(x,\,\cdot\,)$ are measurable, it is always possible (cf. §3.2, and especially Theorem 3.2.5, in [3]) to make measurable choices of $x\rightsquigarrow\sigma$ and $(x,y)\rightsquigarrow F(x,y)$ so that (0.1) holds. In addition, it is well-known (cf. §3.2.1 in [3]) that the non-negative definite, symmetric square root of $x\rightsquigarrow a(x)$ will be uniformly Lipschitz if either $x\rightsquigarrow a(x)$ is uniformly Lipschitz and uniformly positive definite or a and its second derivatives a are uniformly bounded. On the other hand, it is much less clear what smoothness properties of $x\rightsquigarrow M(x,\,\cdot\,)$ will guarantee that F can be chosen so that (0.2) holds. Because it is the Lévy term which poses the greatest challenge to traditional analytic techniques, it may be of interest to investigate how successful Itô's theory is with it, and that is what we will be doing here.

§1 Basic Result

In this section we will show how to construct an F satisfying (0.2) when $x\rightsquigarrow M(x,\,\cdot\,)$ can be expressed in the form

$$
M(x,\Gamma)=\omega_{n-1}\int_{\mathbb{S}^{n-1}}\left(\int_{(0,\infty)}\mathbf{1}_\Gamma(r\omega)\beta(x,\omega,r)\,dr\right)\mu(d\omega)
\tag{1.1}
$$

for $\Gamma\in\mathcal{B}_{\mathbb{R}^n\setminus\{0\}}$, where ω_{n-1} is the area of the unit sphere $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$, $\mu\in\mathbf{M}_1(\mathbb{S}^{n-1})$ (i.e., μ is a Borel probability measure on $\mathbb{S}^{n-1}$), and $\beta:\mathbb{R}^n\times\mathbb{S}^{n-1}\times(0,\infty)\longrightarrow(0,\infty)$ is a measurable function with the properties that

$$\int_{(0,\infty)} \beta(x,\omega,r)\,dr = \infty \quad \text{for all } (x,\omega) \in \mathbb{R}^n \times \mathbb{S}^{n-1}, \tag{1.2}$$

$$\inf_{(x,\omega)} \int_{[1,\infty)} \beta(x,\omega,s)\,ds > 0 \;\&\; \sup_{(x,\omega)} \int_{(0,\infty)} \frac{r^2\beta(x,\omega,r)}{1+r^2}\,dr < \infty, \tag{1.3}$$

$$\lim_{R\searrow 0} \sup_{(x,\omega)} \int_{(0,R]} r^2\beta(x,\omega,r)\,dr = 0 = \lim_{R\to\infty} \sup_{(x,\omega)} \int_{[R,\infty)} \beta(x,\omega,s)\,ds, \tag{1.4}$$

and, for each $(\omega,r) \in \mathbb{S}^{n-1} \times (0,\infty)$, $\beta(\,\cdot\,,\omega,r)$ has a continuous derivative which satisfies

$$\sup_{(x,\omega)} \int_{(0,R]} \frac{\left(\int_{[r,\infty)} \left|\mathrm{grad}_x \beta(\,\cdot\,,\omega,s)\right| ds\right)^2}{\beta(x,\omega,r)}\,dr < \infty \tag{1.5}$$

for each $R \in (0,\infty)$.

The construction of F in this case can be carried out as follows. First, one determines $\rho : \mathbb{R}^n \times \mathbb{S}^{n-1} \times (0,\infty) \longrightarrow (0,\infty)$ so that

$$\int_{[\rho(x,\pm 1,r),\infty)} \beta(x,\pm 1,s)\,ds = \frac{2\mu(\{\pm 1\})}{r} \quad \text{when } n = 1$$

and

$$\int_{[\rho(x,\omega,r),\infty)} \beta(x,\omega,s)\,ds = \frac{1}{r} \qquad \text{when } n \geq 2.$$

Second, $f : \mathbb{S}^{n-1} \longrightarrow \mathbb{S}^{n-1}$ is chosen so that $f(\omega) = \omega$ when $n = 1$ and, when $n \geq 2$, f is a measurable map with the property that $f_*\lambda_{\mathbb{S}^{n-1}} = \omega_{n-1}\mu$, where $f_*\lambda_{\mathbb{S}^{n-1}}$ denotes the pushforward under f of the standard surface measure $\lambda_{\mathbb{S}^{n-1}}$ on $\mathbb{S}^{n-1}$. (The existence of such an f is assured by Theorem 3.2.5 in [3].) Finally, one takes $F(x, r\omega) = \rho\big(x, f(\omega), r\big) f(\omega)$.

To see that this F does the job, begin by observing that, by construction,

$$\int_{\mathbb{R}^n\setminus\{0\}} \varphi\circ F(x,y)\,\frac{dy}{|y|^{n+1}} = \int_{\mathbb{S}^{n-1}}\left(\int_{(0,\infty)} \varphi\circ F(x,r\omega)\frac{dr}{r^2}\right)\lambda_{\mathbb{S}^{n-1}}(d\omega)$$

$$= \int_{\mathbb{S}^{n-1}}\left(\int_{(0,\infty)} \varphi\big(rf(\omega)\big)\beta\big(x,f(\omega),r\big)\,dr\right)\lambda_{\mathbb{S}^{n-1}}(d\omega)$$

$$= \int_{\mathbb{R}^n\setminus\{0\}} \varphi(y)\,M(x,dy)$$

for any $\varphi\in C\big(\mathbb{R}^n\setminus\{0\};[0,\infty)\big)$. That is,

$$\int_{\mathbb{R}^n\setminus\{0\}} \varphi\circ F(x,y)\,M_0(dy) = \int_{\mathbb{R}^n\setminus\{0\}} \varphi(y)\,M(x,dy) \tag{1.6}$$

for $\varphi\in C\big(\mathbb{R}^n\setminus\{0\};[0,\infty)\big)$. Thus, if $\psi\in C^\infty\big(B_{\mathbb{R}^n}(0,1);\mathbb{R}^n\big)$ satisfies $|\psi(y)-y|\le C|y|^2$ for some $C<\infty$ and we adopt

$$K_M\varphi(x) = \int_{\mathbb{R}^n\setminus\{0\}}\Big(\varphi(x+y)-\varphi(x)-\big(\psi(y),\mathrm{grad}_x\varphi\big)_{\mathbb{R}^n}\Big)\,M(x,dy) \tag{1.7}$$

as the operator associated with $x\rightsquigarrow M(x,\,\cdot\,)$, then $K_M\varphi(x)$ is equal to

$$\sum_{i=1}^n c_i(x)\partial_i\varphi(x) + \int_{\mathbb{R}^n\setminus\{0\}}\Big(\varphi\big(x+F(x,y)\big)-\varphi(x)$$
$$- \mathbf{1}_{[0,1]}(|y|)\big(F(x,y),\mathrm{grad}_x\varphi\big)_{\mathbb{R}^n}\Big)\,M_0(dy)$$

where $c:\mathbb{R}^n\longrightarrow\mathbb{R}^n$ is given by

$$c(x) = \int_{\mathbb{R}^n\setminus\{0\}}\Big(\mathbf{1}_{[0,1]}(|y|)F(x,y)-\psi\big(F(x,y)\big)\Big)\,M_0(dy). \tag{1.8}$$

Next, we need to check that F satisfies (0.2) and c is uniformly Lipschitz. To this end, first observe that, by (1.6),

$$\int_{0<|y|\le r} |F(x,y)|^2\frac{dy}{|y|^{n+1}} = \int_{\{y:0<|y|\le R(r)\}} |y|^2\,M(x,dy)$$
$$\le \sup_\omega \int_{\{0<|y|\le R(r)\}} r^2\beta(x,\omega,r)\,dr,$$

where $R(r)\equiv \sup_{(x,\omega)}\rho(x,\omega,r)$. Since, by the second part of (1.4), $R(r)<\infty$ for all $r\in(0,\infty)$, we know that the second line of (0.2) holds.

At the same time, from second part of (1.3), we know that $R(r) \searrow 0$ as $r \searrow 0$, and so the first line of (0.2) also holds.

Turning to the last line of (0.2), observe that

$$\partial_x F(\,\cdot\,, r\omega) = \frac{\int_{[\rho(x,f(\omega),r),\infty)} \partial_x \beta(\,\cdot\,, f(\omega), s)\, ds}{\beta(x, f(\omega), \rho(x, f(\omega), r)} f(\omega),$$

and therefore, by (1.6), that

$$\begin{aligned}
&\int_{0<|y|\le R} \big|F(x_2,y) - F(x_1,y)\big|^2 \frac{dy}{|y|^{n+1}} \\
&\qquad \le |x_2 - x_1|^2 \sup_{(x,\omega)} \int_{(0,R(r)]} \frac{\left(\int_{[r,\infty)} |\mathrm{grad}_x \beta(\,\cdot\,,\omega,s)|\, ds\right)^2}{\beta(x,\omega,r)}\, dr.
\end{aligned}$$

Hence the third line of (0.2) follows from (1.5).

Finally, we must check that the c in (1.8) is uniformly Lipschitz continuous. But, since

$$c(x) = \int_{0<|y|\le 1} \Big(\psi\big(F(x,y)\big) - F(x,y)\Big)\, M_0(dy) + \int_{|y|\ge 1} \psi\big(F(x,y)\big)\, M_0(dy),$$

we can use the first part of (1.3) and the same line of reasoning as above to see that there is an $r \in (0,\infty)$ and a $C < \infty$ for which

$$\begin{aligned}
|\partial_x c| &\le C \sup_\omega \int_{(0,r]} \rho(x,\omega,s) \big|\partial_x \rho(\,\cdot\,,\omega,s)\big| \frac{ds}{s^2} \\
&\le C \sup_\omega \int_{(0,R]} s \left(\int_{[s,\infty)} |\partial_x \beta(\,\cdot\,,\omega,\sigma)|\, d\sigma \right) ds \\
&\le C \sup_\omega \sqrt{\int_{(0,R]} s^2 \beta(x,\omega,s)\, ds} \sqrt{\int_{(0,R]} \frac{\left(\int_{[s,\infty)} |\partial_x \beta(\,\cdot\,,\omega,\sigma)|\, d\sigma\right)^2}{\beta(x,\omega,s)}\, ds}.
\end{aligned}$$

Hence, by the second part of (1.3) and (1.5), it is clear that c is uniformly Lipschitz.

By the results in §3.1 of [3], we can now say that when M is given by (1.1) with a β satisfying (1.2)–(1.5), then Itô's construction leads to a Markov process which corresponds to the operator K_M in (1.7) in the sense that, starting at each $x \in \mathbb{R}^n$, the process solves the martingale problem (cf. §3 below) for K_M on $C_c^2(\mathbb{R}^n; \mathbb{R})$.

An Example: In order to demonstrate that Itô's theory can handle situations which defy more analytic methodology, consider the case when

$$\beta(x,\omega,r) = \alpha(x,\omega)r^{-1-\lambda(x,\omega)},$$

where $\alpha : \mathbb{R}^n \times \mathbb{S}^{n-1} \longrightarrow [\alpha_1,\alpha_2]$ and $\lambda : \mathbb{R}^n \times \mathbb{S}^{n-1} \longrightarrow [\lambda_1,\lambda_2]$ are measurable functions, $0 < \alpha_1 \le \alpha_2 < \infty$, and $0 < \lambda_1 \le \lambda_2 < 2$. Assuming that $\alpha(\,\cdot\,,\omega)$ and $\lambda(\,\cdot\,,\omega)$ are continuously differentiable for each ω and that $(x,\omega) \rightsquigarrow \mathrm{grad}_x\alpha(\,\cdot\,,\omega)$ and $(x,\omega) \rightsquigarrow \mathrm{grad}_x\lambda(\,\cdot\,,\omega)$ are bounded, one can easily verify that β's of this sort satisfy (1.2)–(1.5). The reason why traditional analytic approaches would have difficulties with an operator K_M of the form in (1.7) when M is given by (1.1) with these β's is that, unless λ is independent of x, K_M will have no principal part. For this reason, perturbative techniques, like those on which standard pseudodifferential arguments (cf. [2]) depend, do not apply.

Remark: It is reasonable to ask whether there is any advantage to be gained by considering reference Lévy measures other than $M_0(dy) = \mathbf{1}_{\mathbb{R}^n\setminus\{0\}}\frac{dy}{|y|^{n+1}}$. However, at least so far as the considerations in this and the next sections[1)] are concerned, the answer seems to be *no*. Indeed, without any change in the proof, one can show that Itô's procedure works when M_0 in (0.2) is replaced by any Lévy measure M and the conditions there are replaced by

$$\begin{aligned}
&\lim_{N\to\infty}\sup_{x\in\mathbb{R}^n}\frac{1}{1+|x|^2}\int_{\Delta_N}|F(x,y)|^2\,M(dy) = 0\\
&\sup_{x\in\mathbb{R}^n}\frac{1}{1+|x|^2}\int_{\Gamma_\epsilon}|F(x,y)|^2\,M(dy) < \infty\\
&\sup_{x_2\neq x_1}\frac{1}{|x_2-x_1|^2}\int_{\Gamma_\epsilon}\left|F(x_2,y)-F(x_1,y)\right|^2 M(dy) < \infty,
\end{aligned}\tag{1.9}$$

where, for each $N \ge 1$, $0 \in \Delta_N \in \mathcal{B}_{\mathbb{R}^n}$ satisfies $M(\mathbb{R}^n \setminus \Delta_N) < \infty$, and, for each $\epsilon > 0$, $0 \in \Gamma_\epsilon \in \mathcal{B}_{\mathbb{R}^n}$ satisfies $M(\mathbb{R}^n \setminus \Gamma_\epsilon) < \epsilon$. However, because one can always find a measurable $f : \mathbb{R}^n \setminus \{0\} \longrightarrow \mathbb{R}^n$ such that $M = f_*M_0$, one can easily check that if M and F satisfies $F(x,\,\cdot\,)_*M = M(x,\,\cdot\,)$ and the conditions in (1.9), then $(x,y) \rightsquigarrow \tilde{F}(x,y) \equiv F\big(x,f(y)\big)$ will satisfy $\tilde{F}(x,\,\cdot\,)_*M_0 = M(x,\,\cdot\,)$ and (1.9) with $\tilde{F}$ and M_0 replacing F and M and $f^{-1}(\Delta_N)$ and $f^{-1}(\Gamma_\epsilon)$ replacing Δ_N and Γ_ϵ.

1) See the concluding Remark in §3 for a consideration in which there is an advantage to allowing more general reference Lévy measures.

§2 Some Extensions

It is important to note that there are situations in which it is impossible to construct an F which satisfies (1.9) for any choice of Lévy measure M, even though $x \rightsquigarrow M(x, \cdot)$ is as smooth as one could hope. For example, consider the seemingly trivial case in which $n = 1$ and $M(x, dy) = \alpha(x)\delta_1(dy)$, where $\alpha : \mathbb{R} \longrightarrow [1, 2]$ is smooth and δ_1 is the unit point mass at 1. Clearly, if M is a Lévy measure and $F : \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \longrightarrow \mathbb{R}$ satisfies $F(x, \cdot)_* M = M(x, \cdot)$, then, for each x, $F(x, \cdot) \in \{0, 1\}$ M-almost everywhere and $M\big(\Gamma(x)\big) = \alpha(x)$ when $\Gamma(x) \equiv \{y : F(x, y) = 1\}$. Thus, for each $\epsilon > 0$,

$$\begin{aligned}\int_{\Gamma_\epsilon} \Big(F(x_1, y) - F(x_0, y)\Big)^2 M(dy) &\geq \alpha_\epsilon(x_1) + \alpha_\epsilon(x_0) - 2\alpha_\epsilon(x_1) \wedge \alpha_\epsilon(x_0)\\ &= |\alpha_\epsilon(x_1) - \alpha_\epsilon(x_0)|,\end{aligned}$$

where $\alpha_\epsilon(x) \equiv M\big(\Gamma(x) \cap \Gamma_\epsilon\big) \nearrow \alpha(x)$ uniformly as $\epsilon \searrow 0$. In particular, the only way that the third line of (1.9) could hold is that α_ϵ be constant for each $\epsilon > 0$, which means that α itself would have to be constant. Of course, one can object that this example is a little ridiculous since it is easy to carry out Itô's construction whenever $x \rightsquigarrow M(x, \mathbb{R}^n)$ is bounded, even if the third line of (1.9) fails. On the other hand, one can overcome this objection by considering $M(x, \cdot) = \sum_{m=0}^\infty \alpha_m(x)\delta_{3^{-m}}$ where each $\alpha_m \in C_\mathrm{b}^\infty\big(\mathbb{R}; (0, \infty)\big)$ satisfies $\|\alpha_m\|_{C_\mathrm{b}^1(\mathbb{R};\mathbb{R})} \leq C8^m$. Proceeding as before, we know that $M\big(\Gamma_m(x)\big) = \alpha_m(x)$ and, M-almost everywhere, $F(x, \cdot) = \sum_{m=0}^\infty 3^{-m}\mathbf{1}_{\Gamma_m(x)}$, where $\Gamma_m(x) \equiv \{y : F(x, y) = 3^{-m}\}$. Hence, if $M_\epsilon(dy) = \mathbf{1}_{\Gamma_\epsilon}(y)M(dy)$, then

$$\begin{aligned}&\int_{\Gamma_\epsilon} \big|F(x_1, y) - F(x_0, y)\big|^2 M(dy)\\ &= \sum_{m=0}^\infty 9^{-m}\Big(M_\epsilon\big(\Gamma_m(x_1)\big) + M_\epsilon\big(\Gamma_m(x_1)\big) - 2M_\epsilon\big(\Gamma_m(x_0) \cap \Gamma_m(x_1)\big)\Big)\\ &\quad - 2\sum_{m=0}^\infty 3^{-m}\sum_{n>m} 3^{-n}\Big(M_\epsilon\big(\Gamma_m(x_0) \cap \Gamma_n(x_1)\big) + M_\epsilon\big(\Gamma_n(x_0) \cap \Gamma_m(x_1)\big)\Big)\\ &\geq \sum_{m=0}^\infty 9^{-m}\Big(M_\epsilon\big(\Gamma_m(x_1)\big) + M_\epsilon\big(\Gamma_m(x_1)\big) - 2M_\epsilon\big(\Gamma_m(x_0) \cap \Gamma_m(x_1)\big)\Big)\\ &\quad - \frac{2}{3}\sum_{m=0}^\infty 9^{-m}\Big(M_\epsilon\big(\Gamma_m(x_0) \cap \Gamma_m(x_1)\complement\big) + M_\epsilon\big(\Gamma_m(x_1) \cap \Gamma_m(x_0)\complement\big)\Big)\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{3}\sum_{m=0}^{\infty} 9^{-m}\Big(M_\epsilon\big(\Gamma_m(x_0)\cup\Gamma_m(x_1)\big) - M_\epsilon\big(\Gamma_m(x_0)\cap\Gamma_m(x_1)\big)\Big)\\
&\ge \frac{1}{3}\sum_{m=0}^{\infty} 9^{-m}\big|M_\epsilon\big(\Gamma_m(x_1)\big) - M_\epsilon\big(\Gamma_m(x_1)\big)\big|.
\end{aligned}
$$

Hence, by the same argument as the one just used, the third line of (1.9) can hold only if each of the α_m's is constant.

In view of the preceding example, it is interesting to note that the problems encountered there disappear if the measure has a sufficiently strong absolutely continuous part. To be more precise, let $\beta : \mathbb{R}^n \times \mathbb{S}^{n-1} \longrightarrow (0,\infty)$ be a function which satisfies the conditions in (1.2)–(1.4), and suppose that $(x,\omega) \rightsquigarrow \mu(x,\omega,\,\cdot\,)$ is a measurable map from $\mathbb{R}^n \times \mathbb{S}^{n-1}$ into measures on $(0,\infty)$ such that

$$
\begin{gathered}
\mu(x,\omega,dr) = \beta(x,\omega,r)\,dr + \nu(x,\omega,dr) \quad \text{where}\\
\sup_{(x,\omega)} \int_{(0,\infty)} \frac{r^2}{1+r^2}\,\nu(x,\omega,dr) < \infty \\
\lim_{R\searrow 0}\sup_{(x,\omega)}\int_{(0,R]} r^2\,\nu(x,\omega,dr) = 0 = \lim_{R\to\infty}\sup_{(x,\omega)} \nu\big(x,\omega,[R,\infty)\big).
\end{gathered}
\tag{2.1}
$$

Further, assume that $x \rightsquigarrow \beta(x,\omega,r)$ and $x \rightsquigarrow \nu\big(x,\omega,[r,\infty)\big)$ are continuously differentiable for each $(x,r) \in \mathbb{R}^n \times (0,\infty)$. Finally, choose $\eta \in C^\infty_{\mathrm{c}}\big((0,\infty);[0,\infty)\big)$ with total integral 1, set

$$
\beta_\epsilon(x,\omega,r) = \beta(x,\omega,r) + \int_{(0,\infty)} \eta_\epsilon(s-r)\,\nu(x,\omega,ds) \quad \text{for } \epsilon \in (0,1],
$$

where $\eta_\epsilon(s) \equiv \epsilon^{-1}\eta\left(\frac{s}{\epsilon}\right)$, and assume that, for each $R \in (0,\infty)$,

$$
\sup_{\substack{(x,\omega)\\ \epsilon\in(0,1]}} \int_{(0,R]} \frac{\left(\int_{[r,\infty)} \big|\mathrm{grad}_x\beta_\epsilon(\,\cdot\,,\omega,s)\big|\,ds\right)^2}{\beta(x,\omega,r)}\,dr < \infty.
\tag{2.2}
$$

Next, given a probability measure μ on $\mathbb{S}^{n-1}$, define $\rho_\epsilon : \mathbb{R}^n \times \mathbb{S}^{n-1} \longrightarrow (0,\infty)$ and $F_\epsilon : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ relative to β_ϵ by the prescription used in §1 (cf. the discussion preceding (1.6)). By the arguments used in §1, we know that, when F and β are replaced throughout by F_ϵ and β_ϵ, then everything in (0.2) as well as the Lipschitz continuity of the associated c_ϵ in (1.8) can be controlled in terms of quantity in (1.5). But clearly the quantity in (1.5) is dominated uniformly for $\epsilon \in (0,1]$ by the quantity in (2.2). At the same time, if

$$M(x,\Gamma) = \int_{\mathbb{S}^{n-1}} \left(\int_{(0,\infty)} \mathbf{1}_\Gamma(r\omega)\,\mu(x,\omega,dr) \right) \mu(d\omega), \tag{2.3}$$

then $F_\epsilon \longrightarrow F$ where $F(x,\,\cdot\,)_* M = M(x,\,\cdot\,)$. Hence, when $x \rightsquigarrow M(x,\,\cdot\,)$ is given by (2.3) for any $\mu \in \mathbf{M}_1(\mathbb{S}^{n-1})$ and a $(x,\omega) \in \mathbb{R}^n \times \mathbb{S}^{n-1} \longmapsto \mu(x,\omega,\,\cdot\,) \in \mathbf{M}_1\big((0,\infty)\big)$ which satisfies (2.1) and (2.2), then a choice of F satisfying (0.2) is available.

§3 Uniqueness for the Martingale Problem

Suppose that (cf. (1.7))

$$L\varphi(x) = \frac{1}{2}\sum_{i,j=1}^{n} a(x)_{ij}\partial_i\partial_j\varphi(x) + \sum_{i=1}^{n} b(x)_i + K_M\varphi(x) \tag{3.1}$$

for $\varphi \in C_c^2(\mathbb{R}^n;\mathbb{R})$, where $x \rightsquigarrow a(x)$ and $x \rightsquigarrow b(x)$ are continuous mappings into, respectively, non-negative definite, symmetric $n\times n$-matrices and $\mathbb{R}^n$ and $x \rightsquigarrow M(x,\,\cdot\,)$ takes its values in Lévy measures and satisfies

$$\sup_{|x|\le R} \int_{\mathbb{R}^n\setminus\{0\}} \frac{|y|^2}{1+|y|^2}\, M(x,dy) < \infty \quad \text{for all } R \in (0,\infty).$$

Let $D\big([0,\infty);\mathbb{R}^n\big)$ be the space of right continuous paths $p:[0,\infty) \longrightarrow \mathbb{R}^n$ which possess a left limit $p(t-)$ at each $t \in (0,\infty)$, and use $\mathcal{B}_t$ to denote the σ-algebra over $D\big([0,\infty);\mathbb{R}^n\big)$ generated by $p \rightsquigarrow p(\tau)$ for $\tau \in [0,t]$. We will say that $\mathbb{P} \in \mathbf{M}_1\big(D\big([0,\infty);\mathbb{R}^n\big)\big)$ *solves the martingale problem for* L if

$$\left(\varphi\big(p(t)\big) - \int_0^t L\varphi\big(p(\tau)\big)\,d\tau, \mathcal{B}_t, \mathbb{P} \right) \quad \text{is a martingale}$$

for all $\varphi \in C_c^2(\mathbb{R}^n;\mathbb{R})$. If, in addition, $\mathbb{P}\big(p(0)=x\big) = 1$, then we will say that $\mathbb{P}$ *solves the martingale problem for* L *starting from* x.

In §3.1.5 of [3], it is shown that when L admits a representation of the form in (0.1) with uniformly Lipschitz continuous σ and b and an $x \rightsquigarrow F(x,\,\cdot\,)$ satisfying (0.2), Itô's construction leads to a solution to the martingale problem for L starting from x. On the other hand, there are lots of other ways in which one might go about constructing solutions to this martingale problem. (In fact, even if one restricts ones attention to Itô's method, there is are lots of choices of σ and F, and each one gives rise to a different construction.) Thus, it is of some importance to determine conditions which guarantee that there is only one solution to the martingale problem for a given L starting from a given x.

Under the condition that $M = 0$, the problem of determining when uniqueness holds for the martingale problem was studied systematically in Chapter 6 of [4]. The methods used there are of two types. Methods of the first type work by duality and yield (cf. Theorem 6.3.2 in *op cit*) uniqueness for solutions to the martingale problem as a consequence of existence of solutions to the evolution equation

$$\partial_t u = Lu \quad \text{with } u(0, \cdot\,) = \varphi \tag{3.2}$$

for sufficiently many φ's. This duality method is quite powerful and leads to the most refined results obtained in [4]. For example, when $M = 0$ and a and b have two bounded, continuous derivatives, it is shown in §3.2 of [4] that (3.2) admits classical solutions for $\varphi \in C^2_c(\mathbb{R}^n;\mathbb{R})$, and this is more than enough to check uniqueness for the associated martingale problem. In §4.2 of [3], this sort of reasoning is extended to situations with $M \neq 0$, when the quantities σ and b entering (0.1) have two bounded, continuous derivatives and $x \rightsquigarrow F(x, \cdot\,)$ has continuous second derivatives which satisfies appropriate (cf. $\mathbf{(H2)}^2$ in *op cit*) mean square bounds. For example, these conditions are often met by F's of the sort constructed in §1. Unfortunately, they seem unlikely to hold for situations requiring the extension introduced in §2.

The second method introduced in [4] is more directly dependent on Itô's theory. Namely, when $M = 0$, it is shown there (cf. Theorem 5.3.2 in *op cit*) that any solution to the martingale problem can be realized as the solution of an Itô stochastic integral equation. Thus, when $M = 0$ and σ and b are Lipschitz continuous, uniqueness for the martingale problem comes quite easily as a consequence of Itô's theory. (This is the result which was refined in [6].) In this concluding section, we will examine possible extensions of this line of reasoning to the case when $M \neq 0$.

Suppose that $\mathbb{P}$ solves the martingale problem for L starting from x. Using the techniques developed in §1 of [5], one can make an Itô decompostion of the paths p into their "continuous" and "discontinuous parts" parts. More precisely, given $p \in D([0,\infty);\mathbb{R}^n)$, a $\Gamma \in \mathcal{B}_{[0,\infty)} \times \mathcal{B}_{\mathbb{R}^n}$ with $\big([0,\infty) \times \{0\}\big) \cap \overline{\Gamma} = \emptyset$, define $\nu(\Gamma;p)$ to be the number of $\tau \in (0,\infty)$ such that $\big(\tau, p(\tau) - p(\tau-)\big) \in \Gamma$. One can then show that there exists a measurable map $p \in D\big([0,\infty);\mathbb{R}^n\big) \longmapsto p_1 \in D\big([0,\infty);\mathbb{R}^n\big)$ such that (cf. (1.7))

$$(3.4)\quad p_1(t) = \lim_{r\searrow 0} \iint\limits_{[0,t]\times B_{\mathbb{R}^n}(0,r)\complement} \Big(y\nu\big(d\tau\times dy;p\big) - \psi(y)d\tau\times M\big(p(\tau),dy\big)\Big),$$

uniformly for t's in compacts, in $\mathbb{P}$-probability. Moreover, if $p_0 = p - p_1$, then

(a) $p_0 \in C\big([0,\infty);\mathbb{R}^n\big)$ $\mathbb{P}$-almost surely,

(b) for each $\varphi \in C^2_c(\mathbb{R}^n;\mathbb{R})$,

$$\left(\varphi\big(p_0(t)\big) - \int_0^t L_0\varphi\big(p(\tau)\big)\,d\tau, \mathcal{B}_t, \mathbb{P}\right)$$

is a martingale, where

$$L_0\varphi(x) = \frac{1}{2}\sum_{i,j=1}^n a(x)_{ij}\partial_i\partial_j\varphi(x) + \sum_{i=1}^n b(x)_i\partial_i\varphi(x),$$

(c) for each $\varphi \in C^2_c(\mathbb{R}^n;\mathbb{R})$,

$$\left(\varphi\big(p_1(t)\big) - \int_0^t K_M\varphi\big(p(\tau)\big)\,d\tau, \mathcal{B}_t, \mathbb{P}\right)$$

is a martingale.

In spite of the obvious ambiguity in this decomposition, we will call p_0 and p_1 the *continuous part* and the *discontinuous part* of p.

Given a measurable $x \rightsquigarrow \sigma(x)$ satisfying $a(x) = \sigma(x)\sigma(x)^\top$, one can start from **(b)** above and, by mimicking the procedure in Theorem 5.3.2 of [4], produce a Brownian motion β such that

$$(3.5)\qquad p_0(t) = x + \int_0^t \sigma\big(p(\tau)\big)\,d\beta(\tau) + \int_0^t b\big(p(\tau)\big)\,d\tau, \quad t\in[0,\infty).$$

There are technical difficulties which arise when σ becomes degenerate, and these necessitate the introduction of a larger probability space, one which is big enough to support a full blown Brownian motion. However, as is explained in the theorem just cited, the resolution of such difficulties is well understood. On the other hand, it is not so clear how to treat the analogous difficulties for the discontinuous part p_1. Specifically, given a measurable $F : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ and a Lévy measure[2)] M for

---

[2)] For reasons which will be explained below in the concluding Remark, it is best to allow general reference Lévy measures here rather than always taking M_0.

which $F(x,\,\cdot\,)_*M = M(x,\,\cdot\,)$, it is not clear in general how one can use F to produce a Lévy process from which p_1 can be re-constructed via Itô's procedure. Nonetheless, when F is *non-degenerate* in the sense that, for each $x \in \mathbb{R}^n$, $F(x,\,\cdot\,)$ is one-to-one from $\mathbb{R}^n \setminus \{0\}$ onto itself, then one can construct such a Lévy process. Namely, take $F^{-1}(x,\,\cdot\,)$ to be the inverse[3] of $F(x,\,\cdot\,)$, and set

$$q_r(t,p) = \iint\limits_{[0,t]\times B_{\mathbb{R}^n}(0,r)\complement} \Phi^{-1}(\tau,y;p)\Big(\nu(d\tau\times dy;p) - d\tau\times M\big(p(\tau),dy\big)\Big),$$

where $\Phi^{-1}(\tau,y;p) \equiv F^{-1}\big(p(\tau-),y\big)$. Then one can show that there exists a $\{\mathcal{B}_t : t \geq 0\}$-progressively measurable $p \in D\big([0,\infty);\mathbb{R}^n\big) \longrightarrow q(\,\cdot\,,p) \in D\big([0,\infty);\mathbb{R}^n\big)$ such that, as $r \searrow 0$, $q_r(\,\cdot\,,p) \longrightarrow q(\,\cdot\,,p)$ uniformly on compacts in $\mathbb{P}$-probability. Moreover, the $\mathbb{P}$-distribution of $p \rightsquigarrow q(\,\cdot\,,p)$ is that of the Lévy process corresponding to M in the sense that, for each $\xi \in \mathbb{R}^n$,

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}\Big[e^{\sqrt{-1}(\xi,q(1,p))_{\mathbb{R}^n}}\Big]\\&= \exp\left[\int_{\mathbb{R}^n\setminus\{0\}}\Big(e^{\sqrt{-1}(\xi,y)_{\mathbb{R}^n}} - 1 - \mathbf{1}_{[0,1]}(|y|)(\xi,y)_{\mathbb{R}^n}\Big)M(dy)\right].\end{aligned}$$

In addition, it should be clear that, $\mathbb{P}$-almost surely, $\nu\big(\,\cdot\,;q(\,\cdot\,,p)\big) = \Phi^{-1}(\,\cdot\,;p)_*\nu(\,\cdot\,;p)$. In particular, if $\Phi(\tau,y;q) \equiv F\big(q(\tau-),y\big)$, then

$$\nu(\,\cdot\,;p) = \Phi\big(\,\cdot\,;q(\,\cdot\,,p)\big)_*\nu\big(\,\cdot\,;q(p)\big) \quad \mathbb{P}\text{-almost surely,}$$

and so (cf. (3.4)) $p_1(t)$ is equal to

$$\begin{aligned}\lim_{r\searrow 0} \iint\limits_{[0,t]\times B_{\mathbb{R}^n}(0,r)} &\Big(F\big(p(\tau-),y\big)\nu\big(d\tau\times dy;q(\,\cdot\,;p)\big)\\&\qquad - \psi(y)\big(F(p(\tau),y)\big)_{\mathbb{R}^n}\,d\tau\times M(dy)\Big).\end{aligned}$$

Hence, after putting this together with (3.5), the path p can be recovered via Itô's procedure from a Lévy process for which

$$\xi \rightsquigarrow \exp\left[-\frac{|\xi|^2}{2} + \int_{\mathbb{R}^n\setminus\{0\}}\Big(e^{\sqrt{-1}(\xi,y)_{\mathbb{R}^n}} - 1 - \mathbf{1}_{[0,1]}(|y|)(\xi,y)_{\mathbb{R}^n}\Big)\,M(dy)\right]$$

[3] By a famous theorem due to C. Kuratowski, this inverse will be Borel measurable with respect to (x,y).

is the characteristic function of the distribution at time 1.

These considerations yield the following uniqueness theorem.

Theorem. *Suppose that $\sigma : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ and $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ are uniformly Lipschitz continuous functions, M is a Lévy measure, and the $F : \mathbb{R}^n \times \mathbb{R}^n \setminus \{0\} \longrightarrow \mathbb{R}^n \setminus \{0\}$ is a measurable function which satisfies the conditions in (1.9). Further, assume that, for each $x \in \mathbb{R}^n$, $F(x, \cdot\,)$ is one-to-one from $\mathbb{R}^n \setminus \{0\}$ onto itself. Then, for each starting point, Itô's construction yields the one and only solution to the martingale problem for operator L described in (3.1) when $a(x) = \sigma(x)\sigma(x)^\top$ and $M(x, \cdot\,) = F(x, \cdot\,)_* M$ for all $x \in \mathbb{R}^n$.*

Remark: As distinguished from our earlier results, there is an advantage to allowing reference Lévy measures other than M_0 when applying the preceding theorem. For instance, suppose that $(x, \omega, r) \rightsquigarrow \mu(x, \omega, r)$ satisfies the conditions in (2.1) and (2.2), and let $x \rightsquigarrow M(x, \cdot\,)$ be given by (2.3) for some $\mu \in \mathbf{M}_1(\mathbb{S}^{n-1})$. Then the function F which was constructed so that $M(x, \cdot\,) = F(x, \cdot\,)_* M_0$ need not be one-to-one and onto because the map $f : \mathbb{S}^{n-1} \longrightarrow \mathbb{S}^{n-1}$ may fail to be. On the other hand, if we take $M(dy) = \frac{1}{|y|^{n+1}} \mu(dy)$, then the construction given in §1 does not require the use of f and leads to an F which is one-to-one and onto and satisfies $M(x, \cdot\,) = F(x, \cdot\,)_* M$ as well as the conditions in (1.9).

References

[1] Itô, K., "On stochastic differential equations", Mem. Amer. Math. Soc. #4, A.M.S., Providence, RI, 1951.

[2] Jacob, N. & Schilling, R., Lévy–Type processes and pseudodifferential operators, in "Lévy Processes, Theory and Applications", (Barndorff –Nielsen, O., Mikosch, T., and Resnik, S., eds.), Birkhäuser, 2001.

[3] Stroock, D., "Markov Procsses from Itô's Perspective", Annals of Math. Studies #155, Princeton Univ. Press, 2003.

[4] Stroock, D. & Varadhan, S.R.S., "Multidimensional Diffusion Processes", Grundlehren #233, Springer–Verlag, Heidelberg & NYC, 1979.

[5] Stroock, D., Diffusion processes associated with Lévy processes, Z. Wahr. verw. Geb. 32 (1975), 209–244.

[6] Watanabe, S. & Yamada, T., On the uniqueness of solutions of stochastic differential equations, J. Math. Kyoto Univ. #11 (1971), 155–167.

M.I.T., 2-272
Cambridge, MA, 02139-4307 USA
dws@math.mit.edu

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 333–350

Criticality of Generalized Schrödinger Operators and Differentiability of Spectral Functions

Masayoshi Takeda and Kaneharu Tsuchida

Abstract.

Let μ be a positive Radon measure in the Kato class. We consider the spectral bound $C(\lambda) = -\inf \sigma(\mathcal{H}^{\lambda\mu})$ $(\lambda \in \mathbb{R}^1)$ of a generalized Schrödinger operator $\mathcal{H}^{\lambda\mu} = -\frac{1}{2}\Delta - \lambda\mu$ on $\mathbb{R}^d$, and show that the spectral bound is differentiable if $d \leq 4$ and μ is Green-tight.

§1. Introduction

Let $(\mathbf{D}, H^1(\mathbb{R}^d))$ be the classical Dirichlet integral and μ a positive Radon measure in the Kato class. For a Schrödinger operator $\mathcal{H}^{\lambda\mu} = -\frac{1}{2}\Delta - \lambda\mu$, $\lambda \in \mathbb{R}^1$, define *the spectral function* $C(\lambda)$ by

$$\begin{aligned} C(\lambda) &= -\inf\{\theta : \theta \in \sigma(\mathcal{H}^{\lambda\mu})\} \\ &= -\inf\left\{\frac{1}{2}\mathbf{D}(u,u) - \lambda\int_{\mathbb{R}^d}\tilde{u}^2 d\mu : u \in H^1(\mathbb{R}^d),\ \int_{\mathbb{R}^d} u^2 dx = 1\right\}, \end{aligned}$$

where $\sigma(\mathcal{H}^{\lambda\mu})$ is the set of the spectrum of $\mathcal{H}^{\lambda\mu}$ and $\tilde{u}$ is a quasi-continuous version of u. In this paper, we study the differentiability of the function $C(\lambda)$.

When the potential μ is a function in a certain Kato class, Arendt and Batty [3] proved that the spectral function is differentiable at $\lambda = 0$ and its derivative equals to zero ([3, Corollary 2.10]). Using a large deviation principle for additive functionals of the Brownian motion, Wu [27] obtained a necessary and sufficient condition for the spectral function being differentiable at 0. In [24] one of the authors extended Wu's result to measures which may be singular with respect to the Lebesgue measure. Furthermore, one of the authors showed that if $d \leq 2$ and the measure μ is Green-tight (in notation, $\mu \in \mathcal{K}_d^\infty$), the spectral function is differentiable on $\mathbb{R}^1$. Here the class $\mathcal{K}_d^\infty$ was introduced in Zhao [29](see

Received February 12, 2003.

Definition 2.1 (II) below). A main objective of this paper is to extend the results in [24] as follows:

Theorem 1.1. *If $d \leq 4$ and $\mu \in \mathcal{K}_d^\infty$, then the spectral function $C(\lambda)$ is differentiable for all $\lambda \in \mathbb{R}^1$.*

Define $\lambda^+ = \inf\{\lambda > 0 : C(\lambda) > 0\}$. We then see that $\lambda^+ = 0$ for $d \leq 2$ and $\lambda^+ > 0$ for $d \geq 3$ and the proof of Theorem 1.1 is reduced to the proof of the differentiability of $C(\lambda)$ at $\lambda = \lambda^+$. In [24], the differentiablity at $\lambda = 0$ is derived from the fact that for $d \leq 2$ the Brownian motion is a Harris recurrent process with infinite invariant measure, the Lebesgue measure. We will extend this method for $d = 3,\ 4$ by applying the criticality theory of Schrödinger operators.

We first extend the criticality theory to the generalized Schrödinger operator $\mathcal{H}^\mu$; we show in Corollary 3.5 below that if $d \geq 3$, then the operator $\mathcal{H}^{\lambda^+\mu}$ is *critical*, that is, $\mathcal{H}^{\lambda^+\mu}$ does not admit the minimal positive Green function but admits a positive continuous $\mathcal{H}^{\lambda^+\mu}$-harmonic function. This harmonic function is called a *ground state*, which is uniquely determined up to constant multiplication. Moreover, if $d = 3,\ 4$, $\mathcal{H}^{\lambda^+\mu}$ is *null critical*, that is, the ground state does not belong to L^2. In fact, denoting by h the ground state, we prove in section 5 that $h(x)$ is equivalent to the Green function $G(0, x)$ of the Brownian motion on a neighbourhood of the infinity; there exist positive constants c, C such that

$$\frac{c}{|x|^{d-2}} \leq h(x) \leq \frac{C}{|x|^{d-2}}, \quad |x| > 1. \tag{1}$$

The criticality and the null criticality are regarded as extended notions of recurrence and null recurrence respectively. Using these facts, we see that if $d = 3,\ 4$, the h-transformed process generated by the Markov semigroup

$$P_t^{\lambda^+\mu,h} f(x) = \frac{1}{h(x)} \exp(-t\mathcal{H}^{\lambda^+\mu})(hf)(x)$$

becomes a Harris recurrent Markov process with infinite invariant measure h^2dx. Furthermore, through the h-transformation a functional inequality for the *critical Schrödinger form* is derived (Theorem 4.4) ; the inequality is an extenstion of Oshima's inequality ([11]) which holds for the Dirichlet forms generated by symmetric Harris recurrent Markov processes. We now obtain Theorem 1.1 by applying the argument in [24] to the transformed process. This is a key idea of the proof of Theorem 1.1. The equation (1) tells us that if $d \geq 5$, $\mathcal{H}^{\lambda^+\mu}$ becomes *positive critical*, that is, the ground state belongs to L^2. Thus we can not use our method and have not known yet whether $C(\lambda)$ is differentable or not.

The criticality of Schrödinger operators is studied by many people (M. Murata, Y. Pinchover, R. Pinsky,...). In particular, the equation (1) was shown by Murata [10] for classical Schrödinger operators on $\mathbb{R}^d$ and extended by Pinchover [12] to second order elliptic operators in a domain of $\mathbb{R}^d$.

Our motivation lies in the proof of the large deviation principle for continuous additive functional A_t^μ in the Revuz correspondence with μ. The function $C(\lambda)$ is regarded as a *logarithmic moment generating function* of the additive functional A^μ (see [21]), and the differentiability of logarithmic moment generating functions play a crucial role in the Gärtner-Ellis Theorem (see [7]). In fact, using Theorem 1.1, we can show the large deviation principle for additive functional A_t^μ associated with $\mu \in \mathcal{K}_d^\infty$.

§2. Preliminaries

Let $\mathbb{W} = (P_x, B_t)$ be a Brownian motion on $\mathbb{R}^d$ ($d \geq 3$). Let $p(t,x,y)$ be the transition density function of $\mathbb{W}$ and $G(x,y)$ its Green function, $G(x,y) = C(d)|x-y|^{2-d}$, where $C(d) = (2\pi)^{-1}\Gamma(\frac{d}{2}-1)$. For a measure μ, the 0-potential of μ is defined by $G\mu(x) = \int_{\mathbb{R}^d} G(x,y)\mu(dy)$. Let P_t be the semigroup of $\mathbb{W}$, $P_t f(x) = \int_{\mathbb{R}^d} p(t,x,y)f(y)dy = E_x[f(B_t)]$. The Dirichlet form of $\mathbb{W}$ is given by $(1/2\mathbf{D}, H^1(\mathbb{R}^d))$ where $\mathbf{D}$ denotes the classical Dirichlet integral and $H^1(\mathbb{R}^d)$ is the Sobolev space of order 1 ([8, Example 4.4.1]). Let $(1/2\mathbf{D}, H_e^1(\mathbb{R}^d))$ denote the extended Dirichlet form of $(1/2\mathbf{D}, H^1(\mathbb{R}^d))$ ([8, p.36]). Note that $H_e^1(\mathbb{R}^d)$ is a Hilbert space with inner product $\mathbf{D}$ because $\mathbb{W}$ is transient ([8, Theorem 1.5.3]). Let $G_\alpha(x,y)$ be the α-resolvent kernel of $\mathbb{W}$.

Throughout this paper, the Lebesgue measure is denoted by m and $m(dx)$ is abbriviated to dx. For $r > 0$, we denote by $B(r)$ an open ball with radius R centered at the origin. We use $c, C, ..., etc$ as positive constants which may be different at different occurrences. We now define classes of measures which play an important role in this paper.

Definition 2.1. (I) *A positive Radon measure μ on $\mathbb{R}^d$ is said to be in the Kato class ($\mu \in \mathcal{K}_d$ in notation), if*

$$\lim_{a\to 0}\sup_{x\in\mathbb{R}^d}\int_{|x-y|\leq a} G(x,y)\mu(dy) = 0. \tag{2}$$

(II) *A measure μ is in $\mathcal{K}_d^\infty$ if μ is in $\mathcal{K}_d$ and satisfies*

$$\lim_{R\to\infty}\sup_{x\in\mathbb{R}^d}\int_{|y|>R}G(x,y)\mu(dy)=0. \tag{3}$$

For $\mu\in\mathcal{K}_d$, define a symmetric bilinear form $\mathcal{E}^\mu$ by

$$\mathcal{E}^\mu(u,u)=\frac{1}{2}\mathbf{D}(u,u)-\int_{\mathbb{R}^d}\widetilde{u}^2d\mu,\quad u\in H^1(\mathbb{R}^d), \tag{4}$$

where $\widetilde{u}$ is a quasi continuous version of u ([8, Theorem 2.1.3]). In the sequel, we always assume that every function $u\in H_e^1(\mathbb{R}^d)$ is represented by its quasi continuous version. Since $\mu\in\mathcal{K}_d$ charges no set of zero capacity by [2, Theorem 3.3], the form $\mathcal{E}^\mu$ is well defined. We see from [2, Theorem 4.1] that $(\mathcal{E}^\mu,H^1(\mathbb{R}^d))$ becomes a lower semi-bounded closed symmetric form. We call $(\mathcal{E}^\mu,H^1(\mathbb{R}^d))$ a *Schrödinger form*. Denote by $\mathcal{H}^\mu$ the self-adjoint operator generated by $(\mathcal{E}^\mu,H^1(\mathbb{R}^d))$: $\mathcal{E}^\mu(u,v)=(\mathcal{H}^\mu u,v)$. Let P_t^μ be the L^2-semigroup generated by $\mathcal{H}^\mu$: $P_t^\mu=\exp(-t\mathcal{H}^\mu)$. We see from [2, Theorem 6.3(iv)] that P_t^μ admits a symmetric integral kernel $p^\mu(t,x,y)$ which is jointly continuous on $(0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d$.

For $\mu\in\mathcal{K}_d$, A_t^μ denotes a positive continuous additive functional which is in the *Revuz correspondence* with μ: for any positive Borel function f and γ-excessive function h,

$$<h\mu,f>=\lim_{t\to 0}\frac{1}{t}E_{hm}\left[\int_0^t f(B_s)dA_s^\mu\right], \tag{5}$$

([8, p.188]). By the Feynman-Kac formula, the semigroup P_t^μ is written as

$$P_t^\mu f(x)=E_x[\exp(A_t^\mu)f(B_t)]. \tag{6}$$

§3. Criticality and ground state

Definition 3.1. *A real-valued function h is said to be harmonic on a domain D with respect to $\mathcal{H}^\mu$ if for any relatively compact open set $G\subset\bar{G}\subset D$,*

$$h(x)=E_x[\exp(A_{\tau_G}^\mu)h(B_{\tau_G})],\quad x\in G, \tag{7}$$

where τ_G is the first exit time from G, $\tau_G=\inf\{t>0:B_t\notin G\}$.

We formally write a $\mathcal{H}^\mu$-harmonic function h as $\mathcal{H}^\mu h=0$. An operator $\mathcal{H}^\mu$ is said to be *subcritical* if $\mathcal{H}^\mu$ possesses the minimal positive

Green function $G^\mu(x,y)$, that is,

$$G^\mu(x,y) = \int_0^\infty p^\mu(t,x,y)dt < \infty, \quad x \neq y.$$

The operator $\mathcal{H}^\mu$ is said to be *critical* if $G^\mu(x,y) = \infty$ and a positive continuous $\mathcal{H}^\mu$-harmonic function exists. If the operator $\mathcal{H}^\mu$ is neither subcritical nor critical, it is said to be *supercritical* (see [13, p.145]).

The *spectral function* $C(\lambda)$ is defined by the bottom of the spectrum of $\mathcal{H}^{\lambda\mu}$: for $\mu \in \mathcal{K}_d^\infty$,

$$C(\lambda) = -\inf\left\{\mathcal{E}^{\lambda\mu}(u,u)\ ;\ u \in H^1(\mathbb{R}^d), \int_{\mathbb{R}^d} u^2 dx = 1\right\}. \tag{8}$$

Define

$$\lambda^+ \quad = \quad \inf\{\lambda > 0 : C(\lambda) > 0\}.$$

We then see that $C(\lambda) = 0$ for $\lambda \leq \lambda^+$ ([23]).

Lemma 3.1. *For $\mu \in \mathcal{K}_d^\infty$, there exists a positive continuous function such that $\mathcal{H}^{\lambda^+\mu}h = 0$.*

Proof. Let λ_n be the bottom of spectrum of $\mathcal{H}^{\lambda^+\mu}$ for the Dirichlet problem on $B(n)$. Since $0 = -C(\lambda^+) < \lambda_{n+1} < \lambda_n$, $\mathcal{H}^{\lambda^+\mu}$ is subcritical on $B(n)$. Let G^n denotes the Green operator of $\mathcal{H}^{\lambda^+\mu}$ on $B(n)$. We define a function h_n by $h_n(x) = c_n G^{n+1} I_{A_n}(x)$, where I_{A_n} is the indicator function of $A_n(= B(n+1) \setminus B(n))$ and c_n is the normalized constant, $c_n = (G^{n+1}I_{A_n}(0))^{-1}$. Then h_n is a harmonic function on $B(m)$, $m < n$. Indeed, for $x \in B(m)$

$$\begin{aligned} E_x[\exp(\lambda^+ A^\mu_{\tau_m})h_n(B_{\tau_m})] &= c_n E_x[\exp(\lambda^+ A^\mu_{\tau_m})G^{n+1}I_{A_n}(B_{\tau_m})] \\ = c_n E_x\left[\exp(\lambda^+ A^\mu_{\tau_m}) E_{B_{\tau_m}}\left[\int_0^{\tau_{n+1}} \exp(\lambda^+ A^\mu_t) I_{A_n}(B_t)dt\right]\right], \end{aligned}$$

where $\tau_m = \inf\{t > 0 : B_t \notin B(m)\}$. By the strong Markov property, the right hand side is equal to

$$\begin{aligned} & c_n E_x\left[\int_0^{\tau_{n+1}\circ\theta_{\tau_m}} \exp(\lambda^+(A^\mu_{\tau_m} + A^\mu_t \circ \theta_{\tau_m}) I_{A_n}(B_{t+\tau_m})dt\right] \\ = \quad & c_n E_x\left[\int_{\tau_m}^{\tau_{n+1}\circ\theta_{\tau_m}+\tau_m} \exp(\lambda^+ A^\mu_t) I_{A_n}(B_t)dt\right]. \end{aligned}$$

Noting that $\tau_{n+1}\circ\theta_{\tau_m} + \tau_m = \tau_{n+1}$ and $\int_0^{\tau_m} \exp(\lambda^+ A^\mu_t) I_{A_n}(B_t)dt = 0$, we see that the last term is equal to $h_n(x)$. Therefore h_n satisfies (7) for $G = B(m)$.

Now by [4, Corollary 7.8], $\{h_n\}$ is uniformly bounded and equicontinuous on $B(1)$, so we can choose a subsequence of $\{h_n\}$ which converges uniformly on $B(1)$. We denote the subsequence by $\{h_n^{(1)}\}$. Next take a subsequence $\{h_n^{(2)}\}$ of $\{h_n^{(1)}\}$ so that it converges uniformly on $B(2)$. By the same procedure, we take a subsequence $\{h_n^{(m+1)}\}$ of $\{h_n^{(m)}\}$ so that it converges uniformly on $B(m+1)$. Then the function, $h(x) = \lim_{n\to\infty} h_n^{(n)}(x)$, is a desired one. Q.E.D.

Lemma 3.2. *The following statements are equivalent:*

(i) $\inf\left\{\frac{1}{2}\mathbf{D}(u,u) : u \in H^1(\mathbb{R}^d),\ \int_{\mathbb{R}^d} u^2 d\mu = 1\right\} < 1;$

(ii) $\inf\left\{\mathcal{E}^\mu(u,u) : u \in H^1(\mathbb{R}^d),\ \int_{\mathbb{R}^d} u^2 dx = 1\right\} < 0.$

Proof. Assume (i). Then there exists a $\varphi_0 \in C_0^\infty(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d} \varphi_0^2 d\mu = 1$ and $1/2\mathbf{D}(\varphi_0,\varphi_0) < 1$. Letting $u_0 = \varphi_0/\sqrt{\int_{\mathbb{R}^d}\varphi_0^2 dx}$, we have $\mathcal{E}^\mu(u_0,u_0) < 0$.
(ii) $\Longrightarrow$ (i) follows similarly. Q.E.D.

Remark 3.3. We see from [25, Lemma 3.5] that if

$$\inf\left\{\frac{1}{2}\mathbf{D}(u,u) : \int_{\mathbb{R}^d} u^2 d\mu = 1\right\} \le 1,$$

then

$$\inf\left\{\mathcal{E}^\mu(u,u) : \int_{\mathbb{R}^d} u^2 dx = 1\right\} \le 0.$$

However, the converse does not hold in general. Indeed, let $\mu = \sigma_R$, the surface measure of the sphere $\partial B(R)$. Then if $R < \frac{d-2}{2}$, the first infimum is greater than 1, while the second infimum is equal to 0 ([25]).

Lemma 3.4. *Let $\mu \in \mathcal{K}_d^\infty$. Then the number λ^+ is characterized as a unique positive number such that*

$$\inf\left\{\frac{1}{2}\mathbf{D}(u,u) : \lambda^+ \int_{\mathbb{R}^d} u^2 d\mu = 1\right\} = 1. \tag{9}$$

Proof. Define

$$F(\lambda) = \inf\left\{\frac{1}{2}\mathbf{D}(u,u) : \lambda \int_{\mathbb{R}^d} u^2(x)\mu(dx) = 1\right\},$$

Note that $F(\lambda) = F(1)/\lambda$. Then $F(1)$ is nothing but the bottom of spectrum of the time changed process by the additive functional A_t^μ([22, Lemma 3.1]). We see by [23, Lemma 3.1] that 1-resolvent R_1^μ of the time changed process satisfies $R_1^\mu 1 \in C_\infty(\mathbb{R}^d)$. Hence it follows from [17, Corollary 3.2] and [23, Corollary 2.2] that $F(1) > 0$. Consequently we see that $\lambda^0 = F(1)$ is a unique positive constant such that $F(\lambda^0) = 1$. Lemma 3.2 leads us that $\lambda^0 = \lambda^+$. Q.E.D.

Corollary 3.5. *For $\mu \in \mathcal{K}_d^\infty$, the operator $\mathcal{H}^{\lambda^+\mu}$ is critical.*

Proof. Let $F(\lambda)$ be the function in the proof of Lemma 3.4. Then it is known in [25, Theorem 3.9] that the operator $\mathcal{H}^{\lambda\mu}$ is subcritical if and only if $F(\lambda) > 1$. Hence by Lemma 3.1 and Lemma 3.4, $\mathcal{H}^{\lambda^+\mu}$ is critical. Q.E.D.

Lemma 3.6. *A positive $\mathcal{H}^{\lambda^+\mu}$-harmonic function h satisfies $P_t^{\lambda^+\mu}h(x) \leq h(x)$.*

Proof. Let $x \in B(m)$. By Definition 3.1, h satisfies

$$h(x) = E_x[\exp(\lambda^+ A_{\tau_n}^\mu)h(B_{\tau_n})]$$

for any $n > m$. Here τ_n is the first exit time from $B(n)$. It follows from the Markov property that

$$\begin{aligned}
&E_x[\exp(\lambda^+ A_t^\mu)h(B_t); t < \tau_m] \\
&= E_x[\exp(\lambda^+ A_t^\mu)\exp(\lambda^+ A_{\tau_n}^\mu \circ \theta_t)h(B_{\tau_n} \circ \theta_t); t < \tau_m] \\
&= E_x[\exp(\lambda^+ A_{\tau_n}^\mu)h(B_{\tau_n}); t < \tau_m] \leq h(x).
\end{aligned}$$

Hence we have

$$P_t^{\lambda^+\mu}h(x) = \lim_{m\to\infty} E_x[\exp(\lambda^+ A_t^\mu)h(B_t); t < \tau_m] \leq h(x).$$

Q.E.D.

Let P_t be a positive semigroup with integral kernel $p(t,x,y)$. A positive function h is called P_t-*excessive* if h satisfies $P_t h(x) \leq h(x)$. For a P_t-excessive function $h(x)$, the h-transformed semigroup P_t^h is defined by

$$P_t^h f(x) = \int_{\mathbb{R}^d} \frac{1}{h(x)} p(t,x,y)h(y)f(y)dy, \quad t > 0,\ x, y \in \mathbb{R}^d. \tag{10}$$

Then P_t^h becomes a Markovian semigroup.

Let h be the function defined in Lemma 3.1. We see from Lemma 3.6 that the h-transformed semigroup $P_t^{\lambda^+\mu,h}$ generates a h^2m-symmetric Markov process $\mathbb{W}^{\lambda^+\mu,h} = (P_x^{\lambda^+\mu,h}, X_t)$. Note that $\mathbb{W}^{\lambda^+\mu,h}$ is recurrent because of the criticality of $\mathcal{H}^{\lambda^+\mu}$.

Lemma 3.7. *Finely continuous $P_t^{\lambda^+\mu}$-excessive functions are unique up to constant multiplication.*

Proof. We follow the argument in [13, Theorem 4.3.4]. Let h, h' be finely continuous $P_t^{\lambda^+\mu}$-excessive functions. Since

$$E_x\left[\exp(\lambda^+ A_t^\mu)h(B_t)\left(\frac{h'}{h}\right)(B_t)\right] \leq h \cdot \frac{h'}{h}(x),$$

we have

$$E_x^{\lambda^+\mu,h}\left[\frac{h'}{h}(X_t)\right] \leq \frac{h'}{h}(x),$$

where $E_x^{\lambda^+\mu,h}$ is the expectation of h-transformed process $\mathbb{W}^{\lambda^+\mu,h}$. For $y \in \mathbb{R}^d$ and $\epsilon > 0$, we put $U_\epsilon(y) = \{z : |h(z) - h(y)| < \epsilon\}$. Since $U_\epsilon(y)$ is finely open, $\sigma_{U_\epsilon(y)} < \infty$, $P_x^{\lambda^+\mu,h}$-a.s [8, Problem 4.6.3]. Replacing t by σ_ϵ, we have

$$E_x^{\lambda^+\mu,h}\left[\frac{h'}{h}(X_{\sigma_\epsilon})\right] \leq \frac{h'}{h}(x). \tag{11}$$

Note that the left hand side of (11) converges to $\frac{h'}{h}(y)$ as $\epsilon \to 0$. We then have

$$\begin{aligned}\frac{h'}{h}(y) &= E_x^{\lambda^+\mu,h}\left[\liminf_{\epsilon\to 0}\frac{h'}{h}(X_{\sigma_\epsilon})\right] \leq \liminf_{\epsilon\to 0} E_x^{\lambda^+\mu,h}\left[\frac{h'}{h}(X_{\sigma_\epsilon})\right] \\ &\leq \frac{h'}{h}(x).\end{aligned}$$

Since x and y are arbitrary, h'/h is a constant function. Q.E.D.

Now we give known facts on the Kato class.

Theorem 3.8 ([20]). *Let $\mu \in \mathcal{K}_d$. Then for any $u \in H^1(\mathbb{R}^d)$*

$$\int_{\mathbb{R}^d} u^2(x)\mu(dx) \leq \|G_\alpha\mu\|_\infty\left(\mathbf{D}(u,u) + \alpha\int_{\mathbb{R}^d} u^2(x)dx\right). \tag{12}$$

It is known from [1] (also see [28]) that $\mu \in \mathcal{K}_d$ if and only if

$$\lim_{\alpha\to\infty}\|G_\alpha\mu\|_\infty = 0. \tag{13}$$

Therefore we see that for any ϵ there exists a constant $M(\epsilon)$ such that for any $u \in H^1(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} u^2(x)\mu(dx) \leq \epsilon\mathbf{D}(u,u) + M(\epsilon)\int_{\mathbb{R}^d} u^2(x)dx. \tag{14}$$

For a measure μ, let $\mu_R(\cdot) = \mu(\cdot \cap B(R))$ and $\mu_{R^c} = \mu(\cdot \cap B(R)^c)$.

Lemma 3.9. *If $\mu \in \mathcal{K}_d^\infty$, then the embedding of $H_e^1(\mathbb{R}^d)$ to $L^2(\mu)$ is compact.*

Proof. Let $\{u_n\}$ be a sequence in $H_e^1(\mathbb{R}^d)$ such that

$$u_n \to u_0 \in H_e^1(\mathbb{R}^d), \ \mathbf{D}\text{-weakly.}$$

Rellich's theorem says that for any compact set $K \subset \mathbb{R}^d$

$$u_n I_K \to u_0 I_K \quad L^2(m)\text{-strongly.} \tag{15}$$

Now, for $\varphi \in C_0^\infty(\mathbb{R}^d)$ with $\varphi = 1$ on $B(R)$

$$\begin{aligned}\int_{\mathbb{R}^d} |u_n - u_0|^2 \mu_R(dx) &= \int_{\mathbb{R}^d} |u_n\varphi - u_0\varphi|^2 \mu_R(dx) \\ &\le \epsilon \mathbf{D}(u_n\varphi - u_0\varphi, u_n\varphi - u_0\varphi) + M(\epsilon) \int_{\mathbb{R}^d} |u_n\varphi - u_0\varphi|^2 dx\end{aligned}$$

by (14), and the second term converges to 0 as $n \to \infty$ by (15). Since

$$\sup_n \mathbf{D}(u_n\varphi - u_0\varphi, u_n\varphi - u_0\varphi) < \infty$$

by the principle of uniform boundedness and ϵ is arbitrary, u_n converges to u_0 in $L^2(\mu_R)$. Moreover, since by Theorem 3.8,

$$\begin{aligned}\int_{\mathbb{R}^d} |u_n - u_0|^2 \mu(dx) &= \int_{\mathbb{R}^d} |u_n - u_0|^2 \mu_R(dx) + \int_{\mathbb{R}^d} |u_n - u_0|^2 \mu_{R^c}(dx) \\ &\le \int_{\mathbb{R}^d} |u_n - u_0|^2 \mu_R(dx) + \|G\mu_{R^c}\|_\infty \mathbf{D}(u_n - u_0, u_n - u_0),\end{aligned}$$

$$\limsup_{n\to\infty} \int_{\mathbb{R}^d} |u_n - u_0|^2 \mu(dx) \le \|G\mu_{R^c}\|_\infty \sup_n \mathbf{D}(u_n - u_0, u_n - u_0).$$

Hence according to the definition of $\mathcal{K}_d^\infty$ the right hand side converges to 0 by letting R to ∞. Therefore $\{u_n\}$ is an $L^2(\mu)$-convergent sequence. Q.E.D.

Assume that $\mathcal{H}^\mu$ is subcritical or critical. Let h be a positive $\mathcal{H}^\mu$-harmonic function. We denote by $\mathcal{D}_e(\mathcal{E}^\mu)$ the family of m-measurable function u on $\mathbb{R}^d$ such that $|u| < \infty$ m-a.e. and there exists an $\mathcal{E}^\mu$-Cauchy sequence $\{u_n\}$ of functions in $H^1(\mathbb{R}^d)$ such that $\lim_{n\to\infty} u_n = u$ m-a.e. We call $\{u_n\}$ as above an *approximating sequence* for $u \in \mathcal{D}_e(\mathcal{E}^\mu)$.

Note that the Dirichlet form $(\mathcal{E}^{\mu,h}, \mathcal{D}(\mathcal{E}^{\mu,h}))$ associated with the Markov semigroup $P_t^{\mu,h}$ is given by

$$\begin{aligned} \mathcal{E}^{\mu,h}(u,v) &= \mathcal{E}^{\mu}(hu,hv) \\ \mathcal{D}(\mathcal{E}^{\mu,h}) &= \left\{u \in L^2(\mathbb{R}^d; h^2dx) :\ hu \in \mathcal{D}(\mathcal{E}^{\mu})\right\}. \end{aligned}$$

Then we see that $u \in \mathcal{D}_e(\mathcal{E}^{\mu})$ if and only if $u/h \in \mathcal{D}_e(\mathcal{E}^{\mu,h})$, where $\mathcal{D}_e(\mathcal{E}^{\mu,h})$ is the entended Dirichlet space of $(\mathcal{E}^{\mu,h}, \mathcal{D}(\mathcal{E}^{\mu,h}))$. Consequently, the Schrödinger form $\mathcal{E}^{\mu}$ can be well extended to $\mathcal{D}_e(\mathcal{E}^{\mu})$ as a symmetric form: for $u \in \mathcal{D}_e(\mathcal{E}^{\mu})$ and its approximating sequence $\{u_n\}$

$$\mathcal{E}^{\mu}(u,u) = \lim_{n\to\infty} \mathcal{E}^{\mu}(u_n,u_n), \qquad u \in \mathcal{D}_e(\mathcal{E}^{\mu}) \tag{16}$$

(see [8, p.35]). We call $(\mathcal{E}^{\mu}, \mathcal{D}_e(\mathcal{E}^{\mu}))$ the *extended Schrödinger form.* We see from [18, Definition 1.6] that a function u belongs to $\mathcal{D}_e(\mathcal{E}^{\mu})$ if there exists a sequence $\{u_n\}$ of functions in $H^1(\mathbb{R}^d)$ such that $\lim_{n\to\infty} u_n = u$ m-a.e. and

$$\sup_n \mathcal{E}^{\mu}(u_n,u_n) < \infty.$$

If $(\mathcal{E}^{\mu}, H^1(\mathbb{R}^d))$ is a subcritical Schrödinger form, that is, the associated operator $\mathcal{H}^{\mu}$ be subcritical, then $(\mathcal{E}^{\mu}, \mathcal{D}_e(\mathcal{E}^{\mu}))$ becomes a Hilbert space by [8, Lemma 1.5.5]. In particular, a positive $\mathcal{H}^{\mu}$-harmonic function h does not belong to $\mathcal{D}_e(\mathcal{E}^{\mu})$. If $(\mathcal{E}^{\mu}, H^1(\mathbb{R}^d))$ is a critical Schrödinger form, that is, the associated operator $\mathcal{H}^{\mu}$ be critical, its ground state h belongs to $\mathcal{D}_e(\mathcal{E}^{\mu})$ on account of [8, Theotem 1.6.3]. Noting that for $\mu \in \mathcal{K}_d^{\infty}$

$$\mathcal{E}^{\mu}(u,u) \le (1/2 + \|G\mu\|_{\infty})\mathbf{D}(u,u)$$

by Theorem 3.8, we see that $\mathcal{D}_e(\mathcal{E}^{\mu})$ includes $H_e^1(\mathbb{R}^d)$.

For $w \ge 0 \in C_0(\mathbb{R}^d)$ define $\nu = \lambda^+\mu - w\cdot m$. We then see that $\mathcal{H}^{\nu}$ is subcritical. Let $G^{\nu}(x,y)$ be the Green function of $\mathcal{H}^{\nu}$ and G^{ν} the Green operator,

$$G^{\nu}f(x) = \int_{\mathbb{R}^d} G^{\nu}(x,y)f(y)dy. \tag{17}$$

By [26, Theorem 3.1], the Green function $G^{\nu}(x,y)$ is equivalent to $G(x,y)$: there exist positive constants c, C such that

$$cG(x,y) \le G^{\nu}(x,y) \le CG(x,y) \quad \text{for } x \ne y. \tag{18}$$

Lemma 3.10. *For a positive function $\varphi \in C_0(\mathbb{R}^d)$, $G^{\nu}\varphi$ belongs to $\mathcal{D}_e(\mathcal{E}^{\nu})$*

Proof. Let G^ν_β be the β-resolvent associated with $\mathcal{H}^\nu$. Then $G^\nu_\beta\varphi$ belongs to $H^1(\mathbb{R}^d)$ and $G^\nu_\beta\varphi \to G^\nu\varphi$ as $\beta \to 0$. Moreover,

$$\mathcal{E}^\nu(G^\nu_\beta\varphi, G^\nu_\beta\varphi) \le \mathcal{E}^\nu_\beta(G^\nu_\beta\varphi, G^\nu_\beta\varphi) = (\varphi, G^\nu_\beta\varphi) \le (\varphi, G^\nu\varphi)$$

and the right hand side is not greater than $C(\varphi, G\varphi) < \infty$ by (18). Q.E.D.

The next theorem is first obtained by Murata [10, Theorem 2.2] when the potential μ is absolutely continuous with respect to the Lebesgue measure.

Theorem 3.11. *For $w \in C_0(\mathbb{R}^d)$ with $w \ge 0$, $w \not\equiv 0$, let $\nu = \lambda^+\mu - w\cdot m$. The positive continuous $\mathcal{H}^{\lambda^+\mu}$-harmonic function h satisfies*

$$h(x) = \int_{\mathbb{R}^d} G^\nu(x,y)h(y)w(y)dy. \tag{19}$$

Proof. Note that by Lemma 3.9 there exists a function $u_0 \in H^1_e(\mathbb{R}^d)$ such that u_0 attains the infimum:

$$\inf\left\{\frac{1}{2}\mathbf{D}(u,u) : u \in H^1_e(\mathbb{R}^d),\ \lambda^+\int_{\mathbb{R}^d} u^2 d\mu = 1\right\} = 1.$$

The function u_0 then satisfies the following equation:

$$\frac{1}{2}\mathbf{D}(u_0, f) = \lambda^+\int_{\mathbb{R}^d} u_0 f d\mu \quad \text{for all } f \in H^1_e(\mathbb{R}^d),$$

and thus by the definition of ν

$$\mathcal{E}^\nu(u_0, f) = \int_{\mathbb{R}^d} u_0 f w dx \quad \text{for all } f \in H^1_e(\mathbb{R}^d).$$

On account of the definition of the extended Schrödinger form, we see that the equation above is extended to any $f \in \mathcal{D}_e(\mathcal{E}^\nu)$. Since $G^\nu\varphi \in \mathcal{D}_e(\mathcal{E}^\nu)$ for any $\varphi \in C_0(\mathbb{R}^d)$ by Lemma 3.10, we obtain, by substituting $G^\nu\varphi$ for f

$$\int_{\mathbb{R}^d} u_0(x)\varphi(x)dx = \int_{\mathbb{R}^d} u_0(x)w(x)G^\nu\varphi(x)dx = \int_{\mathbb{R}^d} G^\nu(u_0 w)(x)\varphi(x)dx,$$

thus

$$u_0(x) = \int_{\mathbb{R}^d} G^\nu(x,y)u_0(y)w(y)dy, \quad m\text{-a.e.}$$

Let

$$v(x) = E_x\left[\int_0^\infty \exp(A_t^\nu)u_0(B_t)w(B_t)dt\right],$$

Then the function $v(x)$ equals to $u_0(x)$ m-a.e. and satisfies

$$v(x) = \int_{\mathbb{R}^d} G^\nu(x,y)v(y)w(y)dy, \ m\text{-a.e.}$$

Moreover, $v(x)$ is a finely continuous $P_t^{\lambda^+\mu}$-excessive function. Indeed,

$$\tag{20} \begin{aligned} v(B_s) &= E_{B_s}\left[\textstyle\int_0^\infty \exp(A_t^\nu)u_0(B_t)w(B_t)dt\right] \\ &= E_x\left[\textstyle\int_0^\infty \exp(A_t^\nu\circ\theta_s)u_0(B_{t+s})w(B_{t+s})dt\Big|\mathcal{F}_s\right] \\ &= \exp(-A_s^\nu)E_x\left[\textstyle\int_0^\infty \exp(A_t^\nu)u_0(B_t)w(B_t)dt\Big|\mathcal{F}_s\right] \\ &\quad - \exp(-A_s^\nu)\textstyle\int_0^s \exp(A_t^\nu)u_0(B_t)w(B_t)dt. \end{aligned}$$

and the first term of the last equality is right continuous because of the right continuity of $\mathcal{F}_s$. Hence v is finely continuous ([10,Theorem A.2.7]), and thus $v(x) = u_0(x)$ q.e. Consequently

$$\tag{21} v(x) = E_x\left[\int_0^\infty \exp(A_t^\nu)v(B_t)w(B_t)dt\right] \text{ for any } x.$$

Let $M_t = E_x[\int_0^\infty \exp(A_t^\nu)v(B_t)w(B_t)dt|\mathcal{F}_s]$. Then according to (20) and (21)

$$\begin{aligned} &\exp(A_t^{\lambda^+\mu})v(B_t) = \exp(\int_0^t w(B_u)du)(\exp(A_t^\nu)v(B_t)) \\ = \ & v(B_0) + \int_0^t \exp(\int_0^s w(B_u)du)dM_s - \int_0^t \exp(A_s^{\lambda^+\mu})v(B_s)w(B_s)ds \\ &+ \int_0^t \exp(A_s^\nu)v(B_s)\exp(\int_0^s w(B_u)du)w(B_s)ds \\ = \ & v(B_0) + \int_0^t \exp(\int_0^s w(B_u)du)dM_s, \end{aligned}$$

which implies that

$$E_x[\exp(A_t^{\lambda^+\mu})v(B_t)] \le v(x).$$

Hence $h(x) = cv(x)$ by Lemma 3.7, and thus for all x

$$h(x) = \int_{\mathbb{R}^d} G^\nu(x, y)h(y)w(y)dy. \tag{22}$$

Q.E.D.

§4. An extension of Oshima's inequality

In this section, we extend Oshima's inequality in [11] to critical Schrödinger forms. The inequality plays a crucial role for the proof of the differentiability of $C(\lambda)$.

Lemma 4.1. *Let h be a positive continuous $\mathcal{H}^{\lambda^+\mu}$-harmonic function. Then the h-transformed semigroup $P_t^{\lambda^+\mu,h}$ of $P_t^{\lambda^+\mu}$ has the strong Feller property.*

Proof. Following the argument in [6, Corollary 5.2.7], we can prove this lemma. Q.E.D.

Proposition 4.2. *For the ground state h, the h-trasformed process $\mathbb{W}^{\lambda^+\mu,h} = (P_x^{\lambda^+\mu,h}, X_t)$ is Harris recurrent, that is, for a non-negative function f,*

$$\int_0^\infty f(X_t)dt = \infty, \quad P_x^{\lambda^+\mu,h}\text{-}a.s. \tag{23}$$

whenever $m(\{x : f(x) > 0\}) > 0$.

Proof. Since $P_t^{\lambda^+\mu,h}$ generates the h^2m-symmetric recurrent Markov process, we see from [8, Theorem 4.6.6] that

$$P_x[\sigma_A \circ \theta_n < \infty, \ \forall n \geq 0] = 1 \ \text{ for q.e. } x \in \mathbb{R}^d, \tag{24}$$

where $A = \{x : f(x) > 0\}$. Moreover, since the Markov process $\mathbb{W}^{\lambda^+\mu,h}$ has transition density with respect to h^2m, (24) holds for all $x \in \mathbb{R}^d$ by [8, Problem 4.6.3]. Hence according to [16, Chapter X, Proposition (3.11)], we have the equation (23). Q.E.D.

Theorem 4.3. *For the form $\mathcal{E}^{\lambda^+\mu}$ and its ground state h, there exist a positive function $g \in L^1(h^2m)$ and a function $\psi \in C_0(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \psi h^2 dx = 1$ such that for $u \in \mathcal{D}(\mathcal{E}^{\lambda^+\mu,h})$*

$$\int_{\mathbb{R}^d} |u(x) - h(x)L(\frac{u}{h})|g(x)h(x)dx \leq C\mathcal{E}^{\lambda^+\mu}(u, u)^{1/2}, \tag{25}$$

where C is a positive constant and

$$L(u)=\int_{\mathbb{R}^d} u\psi h^2 dx.$$

Proof. We can apply Oshima's inequality to the Dirichlet form $(\mathcal{E}^{\lambda^+\mu,h},\mathcal{D}(\mathcal{E}^{\lambda^+\mu,h}))$ satisfying the Harris recurrence condition: there exist a positive function $g\in L^1(h^2m)$ and a function $\psi\in C_0(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d}\psi h^2dx=1$ such that for any $u\in\mathcal{D}(\mathcal{E}^{\lambda^+\mu,h})$

$$\int_{\mathbb{R}^d}|u(x)-L(u)|g(x)h^2(x)dx\le C\mathcal{E}^{\lambda^+\mu,h}(u,u)^{1/2}, \tag{26}$$

where

$$L(u)=\int_{\mathbb{R}^d} u\psi h^2 dx.$$

Therefore substituting v/h for u in (26) and noting the relation

$$\mathcal{E}^{\lambda^+\mu,h}(v,v)=\mathcal{E}^{\lambda^+\mu}(hv,hv),$$

we obtain the equality (25). Q.E.D.

§5. Differentiability of spectral function

Lemma 5.1 ([24, Lemma 4.3]). *Let $\mu\in\mathcal{K}_d^\infty$. Then for any $\lambda>\lambda^+$, the negative spectrum of $\sigma(\mathcal{E}^{\lambda\mu})$ consists of isolated eigenvalues with finite multiplicities.*

Let $\mathcal{H}^\mu$ be critical and h its groung state. Then we call $\mathcal{H}^\mu$ *null critical* if the function h does not belong to $L^2(m)$,

Theorem 5.2. *Let $\mu\in\mathcal{K}_d^\infty$. If $\mathcal{H}^{\lambda^+\mu}$ is null critical, then its spectral function $C(\lambda)$ is diffentiable.*

Proof. Note that by Lemma 5.1, for $\lambda>\lambda^+$, $-C(\lambda)$ is the principal eigenvalue of Schrödinger operator $\mathcal{H}^{\lambda\mu}=-\frac{1}{2}\Delta-\lambda\mu$. By analytic perturbation theory [9, Chapter VII], we can see that $C(\lambda)$ is differentiable on $\lambda>\lambda^+$. Hence we only need to prove the differentiability of $C(\lambda)$ at $\lambda=\lambda^+$. Since $C(\lambda)$ is convex, it is enough to prove that there exists a sequence $\{\lambda_n\}$ such that $\lambda_n\downarrow\lambda^+$ and $dC(\lambda_n)/d\lambda\downarrow 0$. By the perturbation theory [9, p.405, Chapter VII (4.44)], we see

$$\frac{dC(\lambda)}{d\lambda}=\int_{\mathbb{R}^d}u_\lambda^2 d\mu, \tag{27}$$

where u_λ is the L^2-normalized eigenfunction corresponding to $-C(\lambda)$, that is,

$$C(\lambda) = \lambda \int_{\mathbb{R}^d} u_\lambda^2 d\mu - \frac{1}{2}\mathbf{D}(u_\lambda, u_\lambda). \tag{28}$$

Using (14) and taking $\epsilon > 0$ so small that $\lambda_n \epsilon < 1/2$, we have

$$\mathbf{D}(u_{\lambda_n}, u_{\lambda_n}) \leq \frac{-C(\lambda_n) + \lambda_n M(\epsilon)}{1/2 - \lambda_n \epsilon}.$$

Noting that $C(\lambda_n) \to 0$ as $n \to \infty$, we see

$$\sup_n \mathbf{D}(u_{\lambda_n}, u_{\lambda_n}) < \infty. \tag{29}$$

Since

$$\mathcal{E}^{\lambda^+\mu}(u_{\lambda_n}, u_{\lambda_n}) - \mathcal{E}^{\lambda_n\mu}(u_{\lambda_n}, u_{\lambda_n}) \leq (\lambda_n - \lambda^+)\|G\mu\|_\infty \mathbf{D}(u_{\lambda_n}, u_{\lambda_n})$$

the right hand side converges to 0 as $n \to \infty$ by (29). Thus we obtain

$$\lim_{n\to\infty} \mathcal{E}^{\lambda^+\mu}(u_{\lambda_n}, u_{\lambda_n}) = 0. \tag{30}$$

For the ground state h of $\mathcal{H}^{\lambda^+\mu}$ let $\mathcal{H}^{\lambda^+\mu,h}$ be the h-transformed operator. For $\psi \in C_0(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \psi h^2 dx = 1$, let $L(u) = \int_{\mathbb{R}^d} u(x)\psi(x)h^2(x)dx$. Then we have

$$\left|L(\frac{u_{\lambda_n}}{h})\right| \leq \sqrt{\int_{\mathbb{R}^d} u_{\lambda_n}^2 dx}\sqrt{\int_{\mathbb{R}^d} \psi^2(x)h^2(x)dx} < \infty.$$

Hence we can choose a sequence $\{\lambda_n\}$ tending to λ^+ such that $L(u_{\lambda_n}/h)$ converges to a certain constant C. Noting by Thorem 4.3,

$$\begin{aligned}
&\int_{\mathbb{R}^d} |u_{\lambda_n} - Ch| g h dx \\
\leq\ & \int_{\mathbb{R}^d} |u_{\lambda_n} - hL(\frac{u_{\lambda_n}}{h})| g h dx + \int_{\mathbb{R}^d} |hL(\frac{u_{\lambda_n}}{h}) - Ch| g h dx \\
\leq\ & C\mathcal{E}^{\lambda^+\mu}(u_{\lambda_n}, u_{\lambda_n})^{1/2} + \int_{\mathbb{R}^d} |L(\frac{u_{\lambda_n}}{h}) - C| g h^2 dx \to 0,
\end{aligned}$$

we see $u_{\lambda_n} \to Ch$ a.e. by choosing a subsequence if necessary. Since

$$1 = \liminf_{n\to\infty} \int_{\mathbb{R}^d} u_{\lambda_n}^2 dx \geq \int_{\mathbb{R}^d} \liminf_{n\to\infty} u_{\lambda_n}^2 dx = C^2 \int_{\mathbb{R}^d} h^2 dx,$$

the constant C must be equal to 0 on account of the null criticality. Since $C(\lambda_n)$ is an eigenvalue for $-\mathcal{H}^{\lambda_n\mu}$, $u_{\lambda_n} = e^{-C(\lambda_n)t}P_t^{\lambda_n\mu}u_{\lambda_n}$. Thus we have by [2, Theorem 6.1 (iii)]

$$\|u_{\lambda_n}\|_\infty \le e^{-C(\lambda_n)t}\|P_t^{\lambda_n\mu}\|_{2,\infty} \le \|P_t^{\lambda_1\mu}\|_{2,\infty} < \infty.$$

Also we can assume that $u_{\lambda_n} \to 0$ $q.e.$ as $k \to \infty$ by choosing a subsequence. Therefore we have

$$\begin{aligned} & \limsup_{n\to\infty} \int_{\mathbb{R}^d} u_{\lambda_n}^2 d\mu \\ \le\ & \limsup_{n\to\infty} \int_{\mathbb{R}^d} u_{\lambda_n}^2 d\mu_R + \limsup_{n\to\infty} \|G\mu_{R^c}\|_\infty \mathbf{D}(u_{\lambda_n}, u_{\lambda_n}) \\ \le\ & \|G\mu_{R^c}\|_\infty M. \end{aligned}$$

By letting R to ∞, we complete the proof. Q.E.D.

Finally we consider the situation in Theorem 5.2. By Theorem 3.11 we have

$$c\int_K G^\nu(x,y)w(y)dy \le h(x) \le C\int_K G^\nu(x,y)w(y)dy,$$

where K is the support of w. Let $B(R) \supset K$. Applying the Harnack inequality to $G^\nu(x,\cdot)$, $x \in B(R)^c$, we see that

$$cG^\nu(x,0) \le h(x) \le CG^\nu(x,0) \quad \text{on } x \in B(R)^c.$$

We see from the equation (18) that the ground state h satisfies

$$cG(x,0) \le h(x) \le CG(x,0), \text{ on } x \in B(R)^c. \tag{31}$$

Hence we see that if $d \le 4$, h is not in L^2, that is, $\mathcal{H}^{\lambda^+\mu}$ is null critical. Therefore combining [24, Theorem 4.3] and Theorem 5.2, we obtain Theorem 1.1.

References

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

[2] Albeverio, S., Blanchard, P., Ma, Z.M.: Feynman-Kac semigroups in terms of signed smooth measures, in "Random Partial Differential Equations" ed. U. Hornung et al., Birkhäuser, 1-31, (1991).

[3] Arendt, W., Batty, C.J.K.: The spectral function and principal eigenvalues for Schrödinger operators, Potential Anal. **7**, 415-436, (1997).

[4] Boukricha, A., Hansen, W., Hueber, H.: Continuous solutions of the generalized Schrödinger equation and perturbation of harmonic spaces, Expo. Math. **5**, 97-135, (1987).

[5] Brasche, J.F., Exner, P., Kuperin, Y., Seba, P.: Schrödinger operators with singular interactions, J. Math. Anal. Appl. **184**, 112-139, (1994).

[6] Davies, E.B., Heat Kernals and Spectral Theory, Cambridge Univ. Press, Cambridge, U.K., (1989).

[7] Dembo, A., Zeitouni, O.: Large deviation techniques and applications, Second edition, Applications of Mathematics **38**, Springer-Verlag, New York, (1998).

[8] Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes, Walter de Gruyter, Berlin, (1994).

[9] Kato, T.: Perturbation Theory for Linear Operators, Springer-Verlag, Berlin-Heidelberg-New York, (1984).

[10] Murata, M.: Structure of positive solutions to $(-\Delta + V)u = 0$ in R^n, Duke Math. J. **53**, 869–943, (1986).

[11] Oshima, Y.: Potential of recurrent symmetric Markov processes and its associated Dirichlet spaces, in Funtional Analysis in Markov Processes, ed. M. Fukushima, Lecture Notes in Math. **923**, Springer-Verlag, Berlin-Heidelberg-New York, 260-275, (1982).

[12] Pinchover, Y.: Criticality and ground states for scond-order elliptic equations, J. Differential Equations, **80**, 237-250, (1989).

[13] Pinsky, R.G.: Positive Harmonic Functions and Diffusion, Cambridge Studies in Advanced Mathematics **45**, Cambridge University Press, (1995).

[14] Reed, M., Simon, B.: Methods of Modern Mathematical Physics, Volume I, Functional Analysis, Academic Press, (1972).

[15] Reed, M., Simon, B.: Methods of Modern Mathematical Physics, Volume IV, Analysis of Operators, Academic Press, (1979).

[16] Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edition, Springer, New York, (1998).

[17] Sato, S.: An inequality for the spectral radius of Markov processes, Kodai Math. J. **8**, 5-13, (1985).

[18] Silverstein, M. L.: Symmetric Markov processes, Lecture Notes in Math ematics, **426**, Springer-Verlag, Berlin-New York, (1974).

[19] Simon, B.: Schrödinger semigroups, Bull. Amer. Math. Soc. **7**, 447-536, (1982).

[20] Stollmann, P., Voigt, J.: Perturbation of Dirichlet forms by measures, Potential Anal. **5**, 109-138, (1996).

[21] Takeda, M.: Asymptotic properties of generalized Feynman-Kac functionals, Potential Analysis **9**, 261-291, (1998).

[22] Takeda, M.: Exponential decay of lifetimes and a theorem of Kac on total occupation times, Potential Analysis **11**, 235-247, (1999).

[23] Takeda, M.: L^p-independence of the spectral radius of symmetric Markov semigroups, Stochastic Processes, Physics and Geometry: New Interplays. **II**: A Volume in Honor of Sergio Albeverio, Edited by Fritz Gesztesy, et.al, (2000).

[24] Takeda, M.: Large deviation principle for additive functionals of Brownian motion corresponding to Kato measures, To appear in Potential Analysis.

[25] Takeda, M.: Conditional gaugeability and subcriticality of generalized Schrödinger operators, J. Funct. Anal. **191**, 343-376, (2002).

[26] Takeda, M., Uemura, T.: Subcriticality and Gaugeability for Symmetric α-Stable Processes, To appear in Forum Math.

[27] Wu,L.: Exponential convergence in probability for emprical means of Brownian motion and random walks, J. Theoretical Probab. **12**, 661-673, (1999).

[28] Zhao, Z.: A probabilistic principle and generalized Schrödinger perturbation. J. Funct. Anal. **101**, 162-176, (1991).

[29] Zhao, Z.: Subcriticality and gaugeability of the Schrödinger operator, Trans. Amer. Math. Soc. **334**, 75-96, (1992).

Masayoshi Takeda
Mathematical Institute
Tohoku University
Aoba, Sendai, 980-8578
Japan
Kaneharu Tsuchida
Mathematical Institute
Tohoku University
Aoba, Sendai, 980-8578
Japan

Advanced Studies in Pure Mathematics 41, 2004
Stochastic Analysis and Related Topics
pp. 351–373

On Spectra of Noises Associated with Harris Flows

Jon Warren and Shinzo Watanabe

Dedicated to Professor Kiyosi Itô on his 88th birthday

Abstract.

A Harris flow is a stochastic flow on the real line given by SDE (2.1) below. We study the *noise* generated by Harris flows, particularly *spectra* of the noise. Our aim is to understand what lies beyond the finite order terms in the chaos expansion (the Wiener-Itô expansion) for *nonstrong* solutions of SDE (2.1).

§1. Definitions and main results

The notion of noises in continuous time (i.e., the case of time $t \in \mathbf{R}$) has been introduced by Tsirelson (cf. [T 1], [T 2], [T 5]):

Definition 1.1. *A* noise $\mathbf{N} = [\{\mathcal{F}_{s,t}\}_{s \le t}, \{T_h\}_{h \in \mathbf{R}}]$ *is a two parameter family of sub σ-fields $\mathcal{F}_{s,t}$, $s \le t$, of events defined on a probability space $(\Omega, \mathcal{F}, P)$ which is stationary in time and possesses the following property:*

$$\mathcal{F}_{s,u} = \mathcal{F}_{s,t} \otimes \mathcal{F}_{t,u}, \quad s \le t \le u, \tag{1.1}$$

that is, $\mathcal{F}_{s,t}$ and $\mathcal{F}_{t,u}$ are independent and generate $\mathcal{F}_{s,u}$, for every $s \le t \le u$. By the stationarity in time, we mean the existence of a measurable flow $\{T_h\}$, i.e., a measurable one-parameter group of automorphisms, on $(\Omega, \mathcal{F}_{-\infty,\infty} := \bigvee_{s \le t} \mathcal{F}_{s,t})$, in which $\mathcal{F}_{s,t}$ is sent to $\mathcal{F}_{s+h,t+h}$ by T_h.

In this article, it is always assumed that the probability space is complete and separable and that a sub σ-field contains all P-null sets.

In the discrete time case (i.e., the case of time $n \in \mathbf{Z}$), a noise can be defined similarly but it is essentially equivalent to giving an i.i.d. random sequence. In the continuous time case, noises generated by increments of a Wiener process (of finite or countably infinite dimension), a stationary Poisson point process, or an independent pair of them, are typical

Received March 31, 2003.

examples which we call *white*, *linearizable* or *classical noises*. There are many non-classical noises, however. Every noise $\mathbf{N} = \{\mathcal{F}_{s,t}\}$ contains a unique maximal (i.e., the largest) classical subnoise which is denoted by $\mathbf{N}^{lin} = \{\mathcal{F}_{s,t}^{lin}\}$.

A *Harris flow* (as will be defined precisely in Def.1.3 below) is a stochastic flow on the real line $\mathbf{R}$ determined uniquely by giving a real positive definite function $b(x)$ such that $b(0) = 1$, (cf. [H]). Note that $b(x) = b(-x)$. We assume that either $b(x) = \mathbf{1}_{\{0\}}(x)$ or $b(x)$ is continuous, $\mathcal{C}^2$ on $\mathbf{R} \setminus \{0\}$ and *strictly positive-definite* in the sense that the matrix $\{b(x_i - x_j)\}$ is strictly positive-definite for any choice of finite different points $\{x_i\}$ in $\mathbf{R}$. The Harris flow in the discontinuous case of $b(x) = \mathbf{1}_{\{0\}}(x)$ is known as the *Arratia flow* ([A]).

Here is a formal definition of stochastic flows on the real line: Let $\mathcal{T}$ be the set of all non-decreasing right-continuous functions $\varphi : x \in \mathbf{R} \mapsto \varphi(x) \in \mathbf{R}$ with the metric defined by $\rho(\varphi, \psi) = \sum_{n=1}^{\infty} 2^{-n} \left(\rho_n(\varphi, \psi) \wedge 1\right)$ where

$$\rho_n(\varphi, \psi) = \inf\{\ \varepsilon > 0 \mid \varphi(x - \varepsilon) - \varepsilon \le \psi(x) \le \varphi(x + \varepsilon) + \varepsilon \quad \text{for all } x \in [-n, n]\}.$$

Then $\mathcal{T}$ is a Polish space: The composite $(\varphi, \psi) \in \mathcal{T} \times \mathcal{T} \mapsto \psi \circ \varphi \in \mathcal{T}$, defined by $\psi \circ \varphi(x) = \psi(\varphi(x))$, and the evaluation map $\mathcal{T} \times \mathbf{R} \ni (\varphi, x) \mapsto \varphi(x) \in \mathbf{R}$ are all Borel measurable even though they are generally not continuous.

Definition 1.2. *By a* stochastic flow *on* $\mathbf{R}$, *we mean a family* $\mathbf{X} = \{X_{s,t};\ s \le t\ \}$ *of* $\mathcal{T}$*-valued random variables* $X_{s,t}$ *having the following properties:*

(1) *(Flow property),* $X_{s,u} = X_{t,u} \circ X_{s,t}$ *and* $X_{t,t} = \mathrm{id}$, *a.s. for every* $s \le t \le u$,
(2) *(Independence property), for any sequence* $t_0 \le t_1 \le \cdots \le t_n$, $\mathcal{T}$*-valued random variables* X_{t_{k-1},t_k}, $k = 1, \cdots, n$, *are independent,*
(3) *(Stationarity), for any* $h > 0$, $X_{s,t} \stackrel{d}{=} X_{s+h,t+h}$,
(4) *(Stochastic continuity),* $X_{0,h} \to \mathrm{id}$ *in probability as* $h \downarrow 0$.

Given a stochastic flow $\mathbf{X} = \{X_{s,t}\}$, it generates a noise $\mathbf{N}^X = \left[\{\mathcal{F}_{s,t}^X\}, \{T_h\}\right]$ by letting $\mathcal{F}_{s,t}^X$ to be the σ-field generated by $\mathcal{T}$-valued random variables $X_{u,v}$, $s \le u \le v \le t$, and $\{T_h\}$ to be a unique one-parameter family of automorphisms on $(\Omega, \mathcal{F}_{-\infty,\infty}^X)$ such that $(T_h)_*(X_{u,v}(x)) = X_{u+h,v+h}(x)$, $u \le v$, $x \in \mathbf{R}$.

Now we give a formal definition of Harris flows. Generally, for a given filtration $\mathbf{F} = \{\mathcal{F}_t\}_{t \ge 0}$, we denote by $\mathcal{M}_2(\mathbf{F})$ the space of all locally square-integrable $\mathbf{F}$-martingales $M = (M_t)_{t\geq 0}$ with $M_0 = 0$ and by $\mathcal{M}_2^c(\mathbf{F})$ the subspace formed of all continuous elements in $\mathcal{M}_2(\mathbf{F})$.

Definition 1.3. *The* Harris flow $\mathbf{X} = \{X_{s,t}\}$ associated with the correlation function $b(x)$ *is a stochastic flow on* $\mathbf{R}$ *such that, for every* $x \in \mathbf{R}$*, if we define the process* $M(x) = (M_t(x))_{t\geq 0}$ *by setting* $M_t(x) = X_{0,t}(x) - x$ *and the filtration* $\mathbf{F}^X = \{\mathcal{F}_t^X\}$ *by setting* $\mathcal{F}_t^X = \mathcal{F}_{0,t}^X$*, then* $M(x) \in \mathcal{M}_2^c(\mathbf{F}^X)$ *and, for every* $x, y \in \mathbf{R}$*, we have*

$$\langle M(x), M(y)\rangle_t = \int_0^t b\left(X_{0,s}(x) - X_{0,s}(y)\right) ds. \tag{1.2}$$

The law of a Harris flow is uniquely determined under our assumption on functions $b(x)$. The existence of Harris flows has been established in [H] (cf. also [LR 1]). A Harris flow is equivalently given by a stochastic differential equation (SDE) (2.1) in Section 2.

Let $\mathbf{X} = \{X_{s,t}\}$ be a Harris flow associated with the function $b(x)$ and $\mathbf{N}^X$ be the noise generated by it. *Suppose that* $b(x)$ *is continuous.* Then we can construct a centered Gaussian system $\mathbf{W} = \{W(t,x); t \in \mathbf{R}, x \in \mathbf{R}\}$ contained in $L_2(\mathcal{F}_{-\infty,\infty}^X)$ such that $(T_h)_*[W(t,x) - W(s,x)] = W(t+h,x) - W(s+h,x)$, $s \leq t$, $x \in \mathbf{R}$ and, if we set $w_t(x) = W(t,x) - W(0,x)$, then $w(x) = (w_t(x))_{t\geq 0} \in \mathcal{M}_2^c(\mathbf{F}^X)$ and, for every $x, y \in \mathbf{R}$, we have $\langle w(x), w(y)\rangle_t = tb(x-y)$. Indeed, $W(t,x) - W(s,x)$ is the L_2-limit of $M_\Delta^x(s,t)$ as $|\Delta| \to 0$. Here, for a sequence of times $\Delta : s = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$ and $x \in \mathbf{R}$, $M_\Delta^x(s,t) = \sum_{k=1}^n (X_{t_{k-1},t_k}(x) - x)$ and $|\Delta| = \max_k |t_k - t_{k-1}|$. $\mathbf{W}$ defines a Gaussian white noise $\mathbf{N}^W = \left[\{\mathcal{F}_{s,t}^W\}, \{T_h\}\right]$ where $\mathcal{F}_{s,t}^W = \sigma[W(v,x) - W(u,x); s \leq u \leq v \leq t, x \in \mathbf{R}]$. It is obvious that $\mathbf{N}^W$ is a subnoise of $\mathbf{N}^X$.

Theorem 1.1. *Suppose that the function* $b(x)$ *is continuous. Then, it holds that* $[\mathbf{N}^X]^{lin} = \mathbf{N}^W$*. Furthermore,* $\mathbf{N}^X = \mathbf{N}^W$ *holds, that is, the noise* $\mathbf{N}^X$ *generated by the Harris flow* $\mathbf{X}$ *is classical, if and only if*

$$\int_{0+}^1 (1 - b(x))^{-1} dx = \infty. \tag{1.3}$$

Hence, the noise $\mathbf{N}^X$ is nonclassical if and only if

$$\int_{0+}^1 (1 - b(x))^{-1} dx < \infty. \tag{1.4}$$

In the case of the Arratia flow, it generates a nonclassical noise: Tsirelson [T 3] (cf. also [LR 2]) showed that this noise is *black* in the sense that $(\mathcal{F}_{s,t}^X)^{lin} = \{\emptyset, \Omega\}$ for every $s \leq t$.

Tsirelson ([T 2], [T 5]) introduced the notion of *spectral measures* for noises which is an invariant under the isomorphism of noises and which can measure the degree of non-linearity (or sensitivity in the discrete-time approximation) of noises. Let $\mathcal{C}$ be the space formed of all compact sets in $\mathbf{R}$ endowed with the Hausdorff distance and $\mathcal{C}^f$ be its subclass formed of all finite sets: $\mathcal{C}^f = \{ S \in \mathcal{C} \mid |S| < \infty \}$. Here, $|S|$ denotes the number of elements in S.

Definition 1.4. *Let* $\mathbf{N} = [\{\mathcal{F}_{s,t}\}, \{T_h\}]$ *be a noise. To every* $\Phi \in L_2(\mathcal{F}_{-\infty,\infty})$*, there corresponds a unique finite Borel measure* μ_Φ *on* $\mathcal{C}$ *such that*

$$(1.5)\qquad \mu_\Phi(\ \{S \in \mathcal{C} \mid S \subset J\}\ ) = E\left[E(\Phi|\mathcal{F}(J))^2\right]$$

for every elementary set $J \subset \mathbf{R}$*. Here, by an elementary set* J*, we mean a finite union* $J = \bigcup_k [t_k, t_{k+1}]$ *of non-overlapping intervals and we set* $\mathcal{F}(J) = \bigvee_k \mathcal{F}_{t_k,t_{k+1}}$*.* μ_Φ *is called the* spectral measure *of the noise* $\mathbf{N}$ *associated with* $\Phi \in L_2(\mathcal{F}_{-\infty,\infty})$*.*

When $\Phi \in L_2(\mathcal{F}_{s,t})$, we have $\mu_\Phi(\mathcal{C} \setminus \mathcal{C}_{[s,t]}) = 0$ where $\mathcal{C}_{[s,t]} = \{S \in \mathcal{C} \mid S \subset [s,t]\}$, so that μ_Φ is a measure on $\mathcal{C}_{[s,t]}$. The following is an important characterization of classical noises due to Tsirelson: *a noise is classical if and only if* $\mu_\Phi(\mathcal{C} \setminus \mathcal{C}^f) = 0$ *for every* $\Phi \in L_2(\mathcal{F}_{-\infty,\infty})$.

Set $L_2^{us}(\mathcal{F}_{s,t}) = \{ \Phi \in L_2(\mathcal{F}_{s,t}) \mid ||\Phi||_2 = 1 \}$; the unit sphere in $L_2(\mathcal{F}_{s,t})$. If $\Phi \in L_2^{us}(\mathcal{F}_{-\infty,\infty})$, then μ_Φ is a Borel probability on $\mathcal{C}$ so that we can speak of a $\mathcal{C}$-value random variable with the distribution μ_Φ. We denote it by S_Φ and call it the *spectral set* of the noise associated with Φ.

We wish to describe the spectral set S_Φ for the noise $\mathbf{N}^X$ generated by a Harris flow $\mathbf{X}$ when $\Phi = X_{0,1}(0) \in L_2^{us}(\mathcal{F}_{0,1}^X)$. The random set S_Φ in this case is denoted by S_X. We would also obtain some information on S_Φ for general Φ. We consider naturally the case when the noise is nonclassical so that we assume (1.4). Furthermore, we assume that

$$(1.6)\quad b(x) \quad \textit{is non-increasing in } (0,\infty) \textit{ and satisfies } \lim_{x\to\infty} b(x) = 0.$$

Functions $b(x) = \exp(-c|x|^\alpha)$ for $c > 0$ and $0 < \alpha < 1$ are typical examples. Also, $b(x) = \mathbf{1}_{\{0\}}(x)$ (the case of the Arratia flow) is another typical example.

For $S \in \mathcal{C}$, let S^{acc} be the the set of all accumulation points of S, so that $S^{acc} \neq \emptyset$ if and only if $S \notin \mathcal{C}^f$.

Theorem 1.2. *Let* $\mathbf{X}$ *be the Harris flow associated with the function* $b(x)$ *which satisfies (1.4) and (1.6) and let* S_X *be the spectral set* S_Φ *of* *the noise* $\mathbf{N}^X$ *for* $\Phi = X_{0,1}(0)$. *Then the random set* S_X^{acc} *has the same law as the random set* $\widetilde{S}$ *in* $[0,1]$ *defined by*

$$\widetilde{S} = \{ \ t \mid 0 \le t \le \tau, \ \widehat{\xi}^+(\tau - t) = 0 \ \} \tag{1.7}$$

where $\widehat{\xi}^+ = \{\widehat{\xi}^+(t)\}_{t\ge 0}$ *is the reflecting diffusion process on* $[0,\infty)$ *with the generator*

$$\widehat{L} = \frac{d}{dx}(1 - b(x))\frac{d}{dx} \tag{1.8}$$

and the initial distribution $\mu(dx) := -db(x)$. *Here,* τ *is a* $[0,1]$*-valued and uniformly distributed random variable independent of* $\widehat{\xi}^+$.

In particular, we have

$$P(S_X^{acc} \ne \emptyset) = P(|S_X| = \infty) = P(\widetilde{S} \ne \emptyset) = P\left\{\exists t \in [0,\tau]; \widehat{\xi}^+(t) = 0\right\}$$

and this probability is also equal to $E\left[\int_0^1 (1 - b(\xi^+(t)))dt\right]$ where $\xi^+ = \{\xi^+(t)\}_{t\ge 0}$ is the reflecting diffusion process on $[0,\infty)$ with the generator

$$L = (1 - b(x))\frac{d^2}{dx^2} \tag{1.9}$$

which starts at 0. Still another expression of this probability is given by the expectation $\frac{1}{2}E\left[A^{-1}(1)\right]$, where $A(t)$ is an additive functional of the one-dimensional Wiener process $\beta(t)$ with $\beta(0) = 0$, defined by

$$A(t) = \frac{1}{2}\int_0^t (1 - b(\beta(s)))^{-1}ds, \tag{1.10}$$

and $t \to A^{-1}(t)$ is the inverse function of $t \to A(t)$.

In the case of the Arratia flow, $S_X^{acc} = S_X$ and it is a perfect set, a.s.. It is described as a zero points set of a (double speed) reflecting Brownian motion starting at 0 as in the theorem. This recovers a result of Tsirelson ([T 4]) who obtained it by an approximation by coalescing random walks.

In the following, we consider the class of Harris flows associated with the correlation functions $b(x)$ which satisfy (1.4), (1.6) and, for some $0 \le \alpha < 1$,

$$1 - b(x) \asymp |x|^\alpha \quad \text{as} \quad x \to 0. \tag{1.11}$$

Again, functions $b(x) = \exp(-c|x|^\alpha)$ for $c > 0$ and $0 < \alpha < 1$ are typical examples. Note also that the function $b(x) = \mathbf{1}_{\{0\}}(x)$ (the case of the Arratia flow) is a typical example of the case when $\alpha = 0$. From Theorem 1.2, we can obtain the following: Denoting by $\dim(S)$ the Hausdorff dimension of a subset S in $\mathbf{R}$,

Corollary 1.1. $\dim(S_X^{acc}) = \frac{1-\alpha}{2-\alpha}$ a.s., *under the condition that it is not empty.*

Theorem 1.3. *Let* $\gamma = \inf\{ \beta \mid \dim(S_\Phi) \leq \beta,$ a.s. *for any* $\Phi \in L_2^{us}(\mathcal{F}_{-\infty,\infty}^X) \}$. *Then*

$$\gamma = \frac{1-\alpha}{2-\alpha}.$$

The proof of these theorems will be given in the subsequent sections by appealing to two main tools: *joinings of Harris flows* and certain *duality relations* between the reflecting (absorbing) L-diffusion and the absorbing (resp. reflecting) $\widehat{L}$-diffusion.

§2. The joining of Harris flows: The proof of Th. 1.1.

Suppose that the correlation function $b(x)$ of a Harris flow $\mathbf{X}$ is continuous. Let $H(\subset \mathbf{C}_b(\mathbf{R} \to \mathbf{R}))$ be the (real) reproducing kernel Hilbert space associated with $b(x)$ so that, defining $f_x \in H$ by $f_x(y) = b(y-x)$, linear combinations $\sum c_i f_{x_i}$ are dense in H and $(f_x, f_y)_H = b(x-y)$. The Gaussian system $\mathbf{W}$ introduced in Section 1 can be given equivalently by a Gaussian system $\{W(t,f); t \in \mathbf{R}, f \in H\}$ contained in $L^2(\mathcal{F}_{-\infty,\infty}^X)$ such that $(T_h)_*[W(t,f) - W(s,f)] = W(t+h,f) - W(s+h,f)$, $s \leq t$, $f \in H$ and, if we set $w_t(f) = W(t,f) - W(0,f)$, then $w(f) = (w_t(f))_{t \geq 0} \in \mathcal{M}_2^c(\mathbf{F}^X)$ and, for every $f, g \in H$, we have $\langle w(f), w(g)\rangle_t = t(f,g)_H$. Indeed, we set $W(t,f) = \sum_i c_i W(t, x_i)$ when $f = \sum c_i f_{x_i}$ and extend this to general $f \in H$ by routine arguments.

We define an Itô-type stochastic integral $\int_0^t \psi_s \cdot W(ds, \varphi_s)$ for $\mathbf{F}^X$-predictable processes φ and ψ satisfying that $\int_0^t |\psi_s|^2 ds < \infty$, a.s., by

$$\int_0^t \psi_s \cdot W(ds, \varphi_s) = \sum_k \int_0^t \psi_s \cdot e_k(\varphi_s) db_k(s),$$

where $\{e_k\}$ is an orthonormal basis (ONB) in H and $b_k(t) = W(t, e_k)$, so that $\{b_k(t)\}$ is an independent family of one-dimensional Wiener processes. As is easily seen, the definition is independent of a particular choice of ONB. Note that $\sum_k e_k(\varphi_s) e_k(\varphi_s') = b(\varphi_s - \varphi_s')$, so that, in particular, $\sum_k |e_k(\varphi_s)|^2 \equiv 1$. Now, (1.2) is equivalently given in the form of SDE for $X_t := X_{0,t}(x)$:

$$X_t = x + \int_0^t W(ds, X_s) = x + \sum_k \int_0^t e_k(X_s) db_k(s). \tag{2.1}$$

Since $\sum_k |e_k(x) - e_k(y)|^2 = 2(1 - b(x-y))$, the condition (1.3) implies the pathwise uniqueness of solutions for SDE (2.1) (cf. [IW], p.182). Hence, if the function b satisfies the condition (1.3), then X_t is a unique strong solution to SDE (2.1) so that $X_{0,t}(x)$ is $\mathcal{F}^W_{0,t}$-measurable for every x. By the stationarity, we see that $X_{s,t}(x)$ is $\mathcal{F}^W_{s,t}$-measurable for every x and $s \le t$. Therefore, $\mathbf{N}^X = \mathbf{N}^W$ holds. Thus, the *if part* of Th. 1.1 is proved.

To prove the *only if part*, we first remark the following martingale representation theorem for Harris flows.

Proposition 2.1. *Suppose the correlation function $b(x)$ of the Harris flow is continuous. Then, $M \in \mathcal{M}_2(\mathbf{F}^X)$ if and only if there exists a sequence $\varphi_k = (\varphi_k(t))$, $k = 1, 2, \dots$, of $\mathbf{F}^X$-predictable processes satisfying that $\sum_k \int_0^t \varphi_k^2(s)ds < \infty$, a.s., for each $t > 0$, and*

$$M(t) = \sum_k \int_0^t \varphi_k(s)db_k(s).$$

In particular, it holds that $\mathcal{M}_2(\mathbf{F}^X) = \mathcal{M}_2^c(\mathbf{F}^X)$.

Proof. Given distinct $x_1, x_2, \dots x_n \in \mathbf{R}$, any $\mathbf{R}^n$-valued process $(X_t^1, X_t^2, \dots X_t^n)$ of which each component X_t^k solves the SDE (2.1) starting from x_k and these components satisfy the coalescing property, has the same law as the n-point motion of the Harris flow $(X_{0,t}(x_1), X_{0,t}(x_2) \dots, X_{0,t}(x_n))$. From this uniqueness in law, it follows by the usual methods that any $M \in \mathcal{M}_2(\mathbf{F}^X)$ that is measurable with respect to this n-point motion is continuous and has the desired representation as a stochastic integral. The result can then be extended to an arbitrary $M \in \mathcal{M}_2(\mathbf{F}^X)$ using the fact that the set of representable martingales is closed in this space. ∎

From this proposition, we can easily deduce that $[\mathbf{N}^X]^{lin} = \mathbf{N}^W$, see also Lemma 6a5 of [T 5]. Indeed, if $\mathbf{N}^W$ is smaller than $[\mathbf{N}^X]^{lin}$, then there should exist some martingale in $\mathcal{M}_2(\mathbf{F}^X)$ which cannot be given by a sum of stochastic integrals by b_k. Hence, in order to prove the *only if part*, it is sufficient to show that *(1.4) implies that $\mathbf{N}^W$ is strictly smaller than $\mathbf{N}^X$*. For this, we introduce the following notion.

Definition 2.1. *By a* joining of a Harris flow, *we mean a pair $(\mathbf{X} = \{X_{s,t}\}, \mathbf{X}' = \{X'_{s,t}\})$ of copies of the Harris flow defined on a same probability space such that the joint process $\Xi = \{\Xi_{s,t} = (X_{s,t}, X'_{s,t}); s \le t\}$ has the independence property (2) in Def.1.2. Given $0 \le \rho \le 1$, it is* *called a* ρ-joining *if it satisfies further the following:* $\mathbf{X}$ *and* $\mathbf{X}'$ *are stationarily correlated in the sense that the joint process* Ξ *has the stationarity property (3) of Def.1.2 and, if filtrations* $\mathbf{F}^X = \{\mathcal{F}_t^X\}$, $\mathbf{F}^{X'} = \{\mathcal{F}_t^{X'}\}$ *and martingales* $M(x) = (M_t(x))$, $M'(x) = (M_t'(x))$ *are defined similarly as in Def.1.3 for* $\mathbf{X}$ *and* $\mathbf{X}'$, *respectively, then* $\mathbf{F}^X$ *and* $\mathbf{F}^{X'}$ *are jointly immersed, i.e.,* $\mathcal{M}_2(\mathbf{F}^X) \cup \mathcal{M}_2(\mathbf{F}^{X'}) \subset \mathcal{M}_2(\mathbf{F}^X \bigvee \mathbf{F}^{X'})$, *and, for every* $x, y \in \mathbf{R}$,

$$\langle M(x), M'(y)\rangle_t = \int_0^t \rho \cdot b\left(X_{0,s}(x) - X'_{0,s}(y)\right) ds, \tag{2.2}$$

$b(x)$ *being the correlation function of the Harris flow.*

It is obvious that, for a ρ-joining, the corresponding Gaussian noises $\mathbf{W}$ and $\mathbf{W}'$ are jointly Gaussian and ρ-correlated.

Lemma 2.1. *For* $0 \le \rho < 1$, *a* ρ*-joining exists and is unique in law. If, in particular,* $\rho = 0$, *then it is a pair of independent copies.*

This lemma can be deduced from the fact that the following differential operator Λ with variables $x = (x_1, \cdots, x_n) \in \mathbf{R}^n$ and $x' = (x_1', \cdots, x_m') \in \mathbf{R}^m$ is non degenerate at all such points $(x, x') \in \mathbf{R}^n \times \mathbf{R}^m$ as all coordinates in x are different and also all coordinates in x' are different:

$$\begin{aligned}\Lambda &= \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n b(x_i - x_j)\frac{\partial^2}{\partial x_i \partial x_j} + \frac{1}{2}\sum_{k=1}^m\sum_{l=1}^m b(x_k' - x_l')\frac{\partial^2}{\partial x_k' \partial x_l'} \\ &+ \rho\sum_{i=1}^n\sum_{k=1}^m b(x_i - x_k')\frac{\partial^2}{\partial x_i \partial x_k'}.\end{aligned}$$

Note that, for a ρ-joining $(\mathbf{X}, \mathbf{X}')$, the process

$$[0,\infty) \ni t \mapsto (X_{0,t}(x_1), \cdots, X_{0,t}(x_n), X'_{0,t}(x_1'), \cdots, X'_{0,t}(x_m'))$$

is a solution to the Λ-martingale problem.

We now assume (1.4) and prove that $\mathbf{N}^W$ is strictly smaller than $\mathbf{N}^X$. Take ρ-joinings $(\mathbf{X}^{(\rho)}, \mathbf{X}'^{(\rho)})$ for $\rho \in [0, 1)$. By (2.2), the process $\xi^{(\rho)}(t) = X_{0,t}^{(\rho)}(0) - X_{0,t}'^{(\rho)}(0)$ is a Feller diffusion on $\mathbf{R}$ with the canonical scale $s(x) = x$ and the speed measure $m(dx) = (1 - \rho \cdot b(x))^{-1}dx$ which starts from the origin at time 0, (cf. [IM] for a general theory of Feller diffusions). As $\rho \nearrow 1$, the processes $\xi^{(\rho)}(t)$ converge to the Feller diffusion $\xi(t)$ with the canonical scale $s(x) = x$ and the speed measure $m(dx) = (1 - b(x))^{-1}dx$ which starts from the origin 0 at time 0. As is well-known, $\xi(t) = \beta(A^{-1}(t))$ for a one-dimensional Wiener process $\beta(t)$ and $A(t)$ is defined by (1.10). Then we have

$$\lim_{\rho \nearrow 1} E\left[|\xi^{(\rho)}(t)|^2\right] = E\left[|\xi(t)|^2\right] = \frac{1}{2}E[A^{-1}(t)] > 0$$

for $t > 0$. Suppose $\mathbf{N}^X \subset \mathbf{N}^W$ be true. Then $X_{0,t}^{(\rho)}(0) := \Phi \in L_2(\mathcal{F}_{0,t}^W)$ and $E\left[X_{0,t}^{\prime(\rho)}(0)|\mathbf{W}\right] = P_{-\log\rho}\Phi$ where $(P_s)_{s\geq 0}$ is the *Ornstein-Uhlenbeck semigroup* acing on $L_2(\mathcal{F}_{0,t}^W)$. Hence $E\left[|\xi^{(\rho)}(t)|^2\right] = 2\left(||\Phi||_2^2 - (\Phi, P_{-\log\rho}\Phi)_2\right)$. By the L^2-continuity of the Ornstein-Uhlenbeck semigroup, we have

$$\lim_{\rho \nearrow 1} E\left[|\xi^{(\rho)}(t)|^2\right] = \lim_{\rho \nearrow 1} 2\left(||\Phi||_2^2 - (\Phi, P_{-\log\rho}\Phi)_2\right) = 0.$$

Thus we have a contradiction and hence we cannot have $\mathbf{N}^X \subset \mathbf{N}^W$. This proves the *only if part* of Th.1.1 so that its proof now is completed.

In the following, we assume that (1.4) holds so that the noise generated by the Harris flow is nonclassical. In this case, 1-joinings are not unique. We specify two of them as the 1^+*-joining* and the 1^-*-joining.*

Definition 2.2. *The* 1^+*-joining* $(\mathbf{X}, \mathbf{X}')$ *is the identity joining: i.e.,* $\mathbf{X} = \mathbf{X}'$. *The* 1^-*-joining is the limit in law of the* ρ*-joinings* $(\mathbf{X}^{(\rho)}, \mathbf{X}'^{(\rho)})$ *as* $\rho \nearrow 1$. *It is such that* $[0,\infty) \ni t \mapsto X_{0,t}(x) - X'_{0,t}(y)$, *for fixed* $x, y \in \mathbf{R}$, *is the Feller diffusion on* $\mathbf{R}$ *with the canonical scale* $s(x) = x$ *and the speed measure* $m(dx) = (1 - b(x))^{-1}dx$ *which starts at* $x - y$ *at time* 0.

For $\rho \in [0, 1)$, let $(\mathbf{X}, \mathbf{X}')$ be a ρ-joining with corresponding ρ-correlated Gaussian processes $\mathbf{W}$ and $\mathbf{W}'$. It is easy to see that the joint law $\Pi(d\mathcal{X}d\mathcal{X}'d\mathcal{W}d\mathcal{W}')$ of $(\mathbf{X}, \mathbf{X}', \mathbf{W}, \mathbf{W}')$ is given by

$$P(\mathbf{X} \in d\mathcal{X}|\mathbf{W} = \mathcal{W})P(\mathbf{X}' \in d\mathcal{X}'|\mathbf{W}' = \mathcal{W}')P(\mathbf{W} \in d\mathcal{W}, \mathbf{W}' \in d\mathcal{W}').$$

From this, we deduce that

$$\begin{aligned} E[\Phi \cdot \pi_*(\Psi)] &= E\big[E[\Phi|\mathbf{W}] \cdot E[\pi_*(\Psi)|\mathbf{W}']\big] \\ &= E\big[E[\Phi|\mathbf{W}] \cdot E\big(E[\pi_*(\Psi)|\mathbf{W}']|\mathbf{W}\big)\big] \\ &= E\big[E[\Phi|\mathbf{W}] \cdot E\big[\pi_*(E(\Psi|\mathbf{W}))|\mathbf{W}\big]\big] \\ &= E\big[E[\Phi|\mathbf{W}] \cdot P_{-\log\rho}(E(\Psi|\mathbf{W}))\big] \end{aligned}$$

whenever $\Phi, \Psi \in L_2(\mathcal{F}_{-\infty,\infty}^X)$. Here, π_* is the unique isomorphism $\pi_* : L_0(\mathcal{F}_{-\infty,\infty}^X) \to L_0(\mathcal{F}_{-\infty,\infty}^{X'})$ such that $\pi_*(X_{s,t}(x)) = X'_{s,t}(x)$ for every s, t and x, and (P_s) is the Ornstein-Uhlenbeck semigroup acting on $L_2(\mathcal{F}^W_{-\infty,\infty})$. By the L^2-continuity of the Ornstein-Uhlenbeck semigroup, the above expectation converges to $E\big[E[\Phi|\mathbf{W}] \cdot E[\Psi|\mathbf{W}]\big]$ as $\rho \nearrow 1$. This proves existence of the 1^--joining as the limit of ρ-joinings. Moreover for a 1^--joining $(\mathbf{X}, \mathbf{X}')$ the corresponding Gaussian systems $\mathbf{W}$ and $\mathbf{W}'$ are equal and $\mathbf{X}$ and $\mathbf{X}'$ are conditionally independent given this common Gaussian process.

Remark 2.1. *For the Arratia flow, its ρ-joining for $\rho \in [0,1)$ is independent of ρ and coincides with 0-joining, that is, a pair of independent copies of the Arratia flow. Hence, its 1^--joining is also a pair of independent copies of the Arratia flow.*

Let $F = \bigcup_{k=1}^n [t_{2k-2}, t_{2k-1}]$ be an elementary set in $\mathbf{R}$ defined for a sequence $t_0 < t_1 < \cdots < t_{2n-2} < t_{2n-1}$ of times. We would introduce the notion of (ρ, F)-*joining* $(\mathbf{X}, \mathbf{X}')$ of the Harris flow when $\rho \in [0,1)$, which is roughly the ρ-joining on F and the identity joining outside F. To be more precise, set $t_{-1} = -\infty$ and $t_{2n} = \infty$ by convention. Take a ρ-joining $(\mathbf{Y}, \mathbf{Y}')$ and a 1^+-joining $(\mathbf{Z}, \mathbf{Z}')$ which are mutually independent. Define $\mathbf{X} = [\{X_{s,t}\}_{s \le t}]$ as follows: First, set $X_{s,t} = Y_{s,t}$ if $t_{2k-2} \le s \le t \le t_{2k-1}$, $k = 1, \cdots, n$ and $X_{s,t} = Z_{s,t}$ if $t_{2k-1} \le s \le t \le t_{2k}$, $k = 0, \cdots, n$. Then, define $X_{s,t}$ for general $s \le t$, by

$$X_{s,t} = X_{t_l,t} \circ X_{t_{l-1},t_l} \circ \cdots \circ X_{t_k,t_{k+1}} \circ X_{s,t_k}$$

when $t_{k-1} < s \le t_k \le t_l \le t < t_{l+1}$, $0 \le k \le l \le 2n-1$. Define $\mathbf{X}' = [\{X'_{s,t}\}_{s \le t}]$ similarly from $\mathbf{Y}'$ and $\mathbf{Z}'$. Then $(\mathbf{X}, \mathbf{X}')$ defines a joining of the Harris flow in which, however, $\mathbf{X}$ and $\mathbf{X}'$ are not stationarily correlated.

Definition 2.3. *The pair $(\mathbf{X}, \mathbf{X}')$ defined above is called the (ρ, F)-*joining *of the Harris flow.*

Next, take mutually independent 1^--joining $(\mathbf{Y}, \mathbf{Y}')$ and 1^+-joining $(\mathbf{Z}, \mathbf{Z}')$ and construct the pair $(\mathbf{X}, \mathbf{X}')$ in the same way.

Definition 2.4. *The pair $(\mathbf{X}, \mathbf{X}')$ defined above is called the $(1^-, F)$-*joining *of the Harris flow*

We turn now to the notion of the spectral measure μ_Φ associated with some $\Phi \in L_2(\mathcal{F}^X_{-\infty,\infty})$ as defined in Def.1.4. This notion is intimately related to chaos expansions. The spectral measure of a random variable $\Phi \in L_2(\mathcal{F}^W_{-\infty,\infty})$, measurable with respect to $\mathbf{W}$, can be expressed by expanding Φ as a sum of multiple Wiener-Itô integrals with respect to the Brownian motions b_k. To be more precise, $\Phi = \sum_{m=1}^\infty I_m$

where I_0 is a constant and I_m, for $m = 1, 2. \cdots$, is given by an iterated Itô stochastic integral

$$I_m = \sum_{(k_1,\cdots,k_m)} \int \cdots \int_{-\infty<t_m<\cdots<t_1<\infty} f_\Phi^{(k_1,\cdots,k_m)}(t_1,\cdots \\ \cdots, t_m) db_{k_m}(t_m) \cdots db_{k_1}(t_1).$$

μ_Φ is supported on $\mathcal{C}^f = \{S \in \mathcal{C} : |S| < \infty\}$ and

$$\mu_\Phi(\mathcal{C}^f) = E(\Phi^2) = \sum_{m=0}^{\infty} E(|I_m|^2)$$

$$= \sum_{m=0}^{\infty} \sum_{(k_1,\cdots,k_m)} \int \cdots \int_{-\infty<t_m<\cdots<t_1<\infty} |f_\Phi^{(k_1,\cdots,k_m)}(t_1,\cdots \\ \cdots, t_m)|^2 dt_m \cdots dt_1 < \infty.$$

The restriction of μ_Φ to $\{S \in \mathcal{C} : |S| = m\}$ is determined (denoting $S = \{t_m, \cdots, t_1\}$, $-\infty < t_m < \cdots < t_1 < \infty$) by

$$\mu_\Phi(dS; |S| = m) = |f_\Phi^{(k_1,\cdots,k_m)}(t_1, \cdots, t_m)|^2 dt_m \cdots dt_1.$$

In particular, $\mu_\Phi(|S| = m) = E(|I_m|^2)$.

For a general $\Phi \in L_2(\mathcal{F}^X_{-\infty,\infty})$, the chaos expansion of $E[\Phi|\mathbf{W}]$ given by $E[\Phi|\mathbf{W}] = \sum_{m=0}^{\infty} I_m$, yields in the same fashion the restriction of μ_Φ to $\mathcal{C}^f$ and in particular

$$E\left[E[\Phi|\mathbf{W}]^2\right] = \mu_\Phi(\mathcal{C}^f).$$

If $(\mathbf{X}, \mathbf{X}')$ is a ρ-joining for $\rho \in [0, 1)$ and $\Phi' = \pi_*(\Phi)$ as above, we have

$$E\big(E[\Phi'|\mathbf{W}']|\mathbf{W}\big) = P_{-log\rho}\big(E[\Phi|\mathbf{W}]\big) = \sum_{m=0}^{\infty} \rho^m I_m.$$

As was remarked above, the relation $E(\Phi\Phi') = E\big(E[\Phi|\mathbf{W}]E[\Phi'|\mathbf{W}']\big)$ holds. Hence,

$$E\big(\Phi\Phi'\big) = \sum_{m=0}^{\infty} \rho^m E(|I_m|^2) = \int_{\mathcal{C}} \rho^{|S|} \mu_\Phi(dS). \tag{2.3}$$

In the same way, we deduce for a 1^--joining $(\mathbf{X}, \mathbf{X}')$,

$$E\big(\Phi\Phi'\big) = \mu_\Phi(\mathcal{C}^f). \tag{2.4}$$

Example 2.1. *Consider the case $\Phi = g(X_{0,1}(x))$ for a bounded continuous function g on $\mathbf{R}$. Note that $E(\Phi^2) = \int_{\mathbf{R}} p(1, x-y)g(y)^2 dy$ where*

$$p(t,x) = \frac{1}{\sqrt{2\pi t}} \exp\{-\frac{x^2}{2t}\}, \quad t > 0, \ x \in \mathbf{R}.$$

The chaos expansion of $E[\Phi|\mathbf{W}]$ was obtained explicitly by Veretennikov and Krylov (cf. [VK]): By setting

$$T_t f(x) = \int_{\mathbf{R}} p(t, x-y) f(y) dy \quad \text{and} \quad Q_t^k f(x) = e_k(x) \frac{\partial}{\partial x} T_t f(x),$$

we have

$$g(X_{0,1}(x)) = \sum_{m=0}^{n} I_m + R_n, \quad I_0 = T_1 g(x) = E[\Phi],$$

where I_m, $m = 1, \dots, n$, and R_n are given by the following iterated Itô stochastic integrals:

$$\begin{aligned} I_m \quad = \quad & \sum_{(k_1, k_2, \cdots, k_m)} \int \cdots \int_{0 < t_m < t_{m-1} < \cdots < t_2 < t_1 < 1} \Big[T_{t_m} Q^{k_m}_{t_{m-1} - t_m} \cdots \\ & \cdots \ Q^{k_2}_{t_1 - t_2} Q^{k_1}_{1 - t_1} g(x) \Big] \, db_{k_m}(t_m) db_{k_{m-1}}(t_{m-1}) \cdots db_{k_2}(t_2) db_{k_1}(t_1), \end{aligned}$$

$$\begin{aligned} R_n \quad = \quad & \sum_{(k_1, k_2, \cdots, k_n, k_{n+1})} \int \cdots \\ & \cdots \int_{0 < t_{n+1} < t_n < \cdots < t_2 < t_1 < 1} \Big[Q^{k_{n+1}}_{t_n - t_{n+1}} Q^{k_n}_{t_{n-1} - t_n} \cdots \\ & \cdots Q^{k_2}_{t_1 - t_2} Q^{k_1}_{1 - t_1} g(X_{0, t_{n+1}}(x)) \Big] \, db_{k_{n+1}}(t_{n+1}) db_{k_n}(t_n) \cdots \\ & \qquad \cdots db_{k_2}(t_2) db_{k_1}(t_1). \end{aligned}$$

From this, we obtain that

$$E[\Phi|\mathbf{W}] = \sum_{m=0}^{\infty} I_m.$$

The following is a key lemma for the proof of Theorem 1.2 which records various generalizations of the identities (2.3) and (2.4). As above, we denote by S_X the spectral set S_Φ when $\Phi = X_{0,1}(0)$ which is a $\mathcal{C}_{[0,1]}$-valued random variable.

Lemma 2.2. *(i) If* $(\mathbf{X},\mathbf{X}')$ *is a* (ρ,F)*-joining of the Harris flow for* $\rho\in[0,1)$*, then,*

$$E\left[\rho^{|S_X\cap F|}\right]=E\left[X_{0,1}(0)X'_{0,1}(0)\right], \tag{2.5}$$

equivalently,

$$E\left[1-\rho^{|S_X\cap F|}\right]=\frac{1}{2}E\left[|X_{0,1}(0)-X'_{0,1}(0)|^2\right]. \tag{2.6}$$

(ii) If $(\mathbf{X},\mathbf{X}')$ *is a* $(1^-,F)$*-joining of the Harris flow, then,*

$$P(\ |S_X\cap F|<\infty\ )=E\left[X_{0,1}(0)X'_{0,1}(0)\right], \tag{2.7}$$

equivalently,

$$P(\ |S_X\cap F|=\infty\ )=\frac{1}{2}E\left[|X_{0,1}(0)-X'_{0,1}(0)|^2\right]. \tag{2.8}$$

(iii) More generally, let $(\mathbf{X},\mathbf{X}')$ *be a* (ρ,F)*-joining for* $0\le\rho<1$ *(a* $(1^-,F)$*-joining) and* $\Phi\in L_2^{us}(\mathcal{F}^X_{-\infty,\infty})$*. There is a unique isomorphism* $\pi_*:L_0(\mathcal{F}^X_{-\infty,\infty})\to L_0(\mathcal{F}^{X'}_{-\infty,\infty})$ *such that* $\pi_*(X_{s,t}(x))=X'_{s,t}(x)$ *for every* s,t *and* x*. Set* $\Phi'=\pi_*(\Phi)$*. Then we have*

$$E\left[\rho^{|S_\Phi\cap F|}\right]\ \left(\text{ resp. } P(\ |S_\Phi\cap F|<\infty\ )\ \right)=E\left[\Phi\Phi'\right], \tag{2.9}$$

equivalently,

$$E\left[1-\rho^{|S_\Phi\cap F|}\right]\ \left(\text{ resp. } P(\ |S_\Phi\cap F|=\infty\ )\ \right)=\frac{1}{2}E\left[|\Phi-\Phi'|^2\right] \tag{2.10}$$

Proof. In the case when $\Phi\in L_2^{us}(\mathcal{F}^X_{s,t})$ and $F=[s,t]$, (2.9) is nothing but (2.3) and (2.4). From this, we can deduce (2.9) in the general case of an elementary set $F=\bigcup_{k=1}^n[t_{2k-2},t_{2k-1}]$, $t_{-1}=-\infty<t_0<\cdots<t_{2n-1}<t_{2n}=\infty$, by considering the following L^2-space factorization:

$$L_2(\mathcal{F}^X_{-\infty,\infty})=\bigotimes_{k=0}^{2n}L_2(\mathcal{F}^X_{t_{k-1},t_k}).$$

We omit the details. ∎

§3. Duality relations for L- and $\widehat{L}$-diffusions in the time reversal: The proof of Th. 1.2.

Let $\{\xi^+(t),P_x\}$ and $\{\widehat{\xi}^+(t),\widehat{P}_x\}$ be the reflecting L- and $\widehat{L}$-diffusion processes on $[0,\infty)$ introduced in Section 1. The associated Markovian semigroups of operators acting on the space $\mathbf{B}([0,\infty))$ of real bounded Borel functions are defined by

$$(3.1) \qquad T_t^+ f(x) = E_x[f(\xi^+(t))] \quad \text{and} \quad \widehat{T}_t^+ f(x) = \widehat{E}_x[f(\widehat{\xi}^+(t))].$$

Define also the semigroups for *absorbing processes* by

$$(3.2) \quad T_t^- f(x) = E_x[f(\xi^+(t \wedge \sigma_0))] \quad \text{and} \quad \widehat{T}_t^- f(x) = \widehat{E}_x[f(\widehat{\xi}^+(t \wedge \widehat{\sigma}_0))],$$

where σ_0 and $\widehat{\sigma}_0$ are the first hitting time to 0 of $\xi^+(t)$ and $\widehat{\xi}^+(t)$, respectively. Introduce, further, the semigroups for *processes with the extinction at hitting to* 0 by

(3.3)

$$T_t^0 f(x) = E_x\left[f(\xi^+(t)) \cdot \mathbf{1}_{[t<\sigma_0]}\right] \text{ and } \widehat{T}_t^0 f(x) = \widehat{E}_x\left[f(\widehat{\xi}^+(t)) \cdot \mathbf{1}_{[t<\widehat{\sigma}_0]}\right].$$

T_t^- and $\widehat{T}_t^-$ are Markovian semigroups and T_t^0 and $\widehat{T}_t^0$ are sub-Markovian semigroups. Note also that T_t^+ , $\widehat{T}_t^+$, T_t^0 and $\widehat{T}_t^0$ have the strong Feller property but T_t^- and $\widehat{T}_t^-$ have the Feller property only. It holds that

$$(3.4) \qquad T_t^0 f = T_t^-(\mathbf{1}_{(0,\infty)} \cdot f) \quad \text{and} \quad \widehat{T}_t^0 f = \widehat{T}_t^-(\mathbf{1}_{(0,\infty)} \cdot f).$$

We have the following duality relations which form another key lemma in the proof of Th.1.2:

Lemma 3.1. *For $x, y \in [0,\infty)$ and $t > 0$,*

$$(3.5) \quad T_t^+ \mathbf{1}_{[0,y]}(x) = \widehat{T}_t^0 \mathbf{1}_{[x,\infty)}(y) \quad \text{and} \quad T_t^- \mathbf{1}_{[0,y]}(x) = \widehat{T}_t^+ \mathbf{1}_{[x,\infty)}(y).$$

More generally, for $x, y \in [0,\infty)$ and $0 \le t_0 < t_1 < \ldots < t_{2n-1} < t_{2n} < t_{2n+1}$,

$$\begin{aligned}(3.6) \qquad & T^+_{t_1-t_0} T^-_{t_2-t_1} T^+_{t_3-t_2} \cdots T^+_{t_{2n-1}-t_{2n-2}} T^-_{t_{2n}-t_{2n-1}} \mathbf{1}_{[0,y]}(x) \\ = \; & \widehat{T}^+_{t_{2n}-t_{2n-1}} \widehat{T}^0_{t_{2n-1}-t_{2n-2}} \cdots \widehat{T}^0_{t_3-t_2} \widehat{T}^+_{t_2-t_1} \widehat{T}^0_{t_1-t_0} \mathbf{1}_{[x,\infty)}(y),\end{aligned}$$

and

$$\begin{aligned}(3.7) \quad & T^+_{t_1-t_0} T^-_{t_2-t_1} T^+_{t_3-t_2} \cdots T^+_{t_{2n-1}-t_{2n-2}} T^-_{t_{2n}-t_{2n-1}} T^+_{t_{2n+1}-t_{2n}} \mathbf{1}_{[0,y]}(x) \\ = \; & \widehat{T}^0_{t_{2n+1}-t_{2n}} \widehat{T}^+_{t_{2n}-t_{2n-1}} \widehat{T}^0_{t_{2n-1}-t_{2n-2}} \cdots \widehat{T}^0_{t_3-t_2} \widehat{T}^+_{t_2-t_1} \widehat{T}^0_{t_1-t_0} \mathbf{1}_{[x,\infty)}(y).\end{aligned}$$

Admitting this lemma for a moment, we now proceed to prove Th. 1.2.

Proof of Th. 1.2. Let $F = [t_0, t_1] \cup [t_2, t_3] \ldots \cup [t_{2n-2}, t_{2n-1}]$ be an elementary set in $[0,1]$ and $(\mathbf{X}, \mathbf{X}')$ be a $(1^-, F)$-coupling of the Harris flow. Set $\xi(t) = X_{0,t}(0) - X'_{0,t}(0)$. Then $|\xi(t)|$ is a time-inhomogeneous diffusion process which behaves as a reflecting L-diffusion when $t \in F$ and as an absorbing L-diffusion (i.e., L-diffusion with 0 as a trap) when $t \in [0,1] \setminus F$. It is known that $P(S_X \ni t) = 0$ for every $t \in [0,1]$ (cf. [T 2]). Then (2.8), combined with this remark, yields that

$$P(|S_X \cap F| = \infty) = P(S_X^{acc} \cap F \neq \emptyset) = \frac{1}{2} E[|\xi(1)|^2].$$

By applying the Itô formula for $\xi(t)$ on each interval $[t_k, t_{k+1}]$, we have

$$\frac{1}{2} E[|\xi(1)|^2] = \int_0^1 E[(1-b)(\xi(t))]dt = 1 - \int_0^1 E[b(\xi(t))]dt,$$

and hence,

$$P(S_X^{acc} \cap F = \emptyset) = \int_0^1 E[b(\xi(t))]dt. \tag{3.8}$$

On the other hand,

$$\begin{aligned} &E[b(\xi(t))] \\ &= \begin{cases} T^+_{t_1-t_0} T^-_{t_2-t_1} \cdots T^-_{t_{2k}-t_{2k-1}} T^+_{t-t_{2k}} b(0), & \text{if } t_{2k} \le t < t_{2k+1} \\ T^+_{t_1-t_0} T^-_{t_2-t_1} \cdots T^+_{t_{2k-1}-t_{2k-2}} T^-_{t-t_{2k-1}} b(0), & \text{if } t_{2k-1} \le t < t_{2k} \end{cases}. \end{aligned}$$

Noting $b(x) = \int_{[0,\infty)} \mathbf{1}_{[0,y]}(x)\mu(dy)$, we have by Lemma 3.1 the following:

$$\begin{aligned} &E[b(\xi(t))] \\ &= \begin{cases} \int_0^\infty \mu(dy) (\widehat{T}^0_{t-t_{2k}} \widehat{T}^+_{t_{2k}-t_{2k-1}} \cdots \widehat{T}^+_{t_2-t_1} \widehat{T}^0_{t_1-t_0} \mathbf{1}_{[0,\infty)})(y), \\ \qquad\qquad\qquad\qquad\qquad \text{if } t_{2k} \le t < t_{2k+1} \\ \int_0^\infty \mu(dy) (\widehat{T}^+_{t-t_{2k-1}} \widehat{T}^0_{t_{2k-1}-t_{2k-2}} \cdots \widehat{T}^+_{t_2-t_1} \widehat{T}^0_{t_1-t_0} \mathbf{1}_{[0,\infty)})(y), \\ \qquad\qquad\qquad\qquad\qquad \text{if } t_{2k-1} \le t < t_{2k} \end{cases}. \end{aligned}$$

If the random set $\widetilde{S}$ is defined by (1.7), it is not difficult to deduce, from the last expression of $E[b(\xi(t))]$, that $\int_0^1 E[b(\xi(t))]dt$ coincides with $P(\widetilde{S} \cap F = \emptyset)$. Then $P(\widetilde{S} \cap F = \emptyset) = P(S_X^{acc} \cap F = \emptyset)$ by (3.8). Since this holds for every elementary set F, we can conclude that $S_X^{acc} \stackrel{d}{=} \widetilde{S}$. ∎

Proof of Lemma 3.1. First, we prove (3.5). For $\lambda > 0$, let U_λ^+ and $\hat{U}_\lambda^0$ be the resolvent operators associated with the semigroups T_t^+

and $\hat{T}_t^0$ respectively. Let f be continuous and compactly supported in $(0,\infty)$. Then $u = U_\lambda^+ f$ solves Poisson's equation

$$Lu - \lambda u = -f,$$

with the boundary conditions $u'(0+) = u(\infty) = 0$. Define functions g and v via

$$g(y) = \int_0^y \frac{f(x)}{a(x)} dx \quad \text{and} \quad v(y) = \int_0^y \frac{u(x)}{a(x)} dx,$$

where $a(x) = (1 - b(x))$. Dividing Poisson's equation through by $a(x)$ and integrating, we obtain

$$\hat{L}v - \lambda v = -g.$$

Moreover v and g are bounded and $v(0) = 0$. Thus we must have $v = \hat{U}_\lambda^0 g$. Letting f approach a delta function we may write the relationship between u and v as:

$$\frac{1}{a(z)} \hat{U}_\lambda^0 \mathbf{1}_{[z,\infty)}(y) = \int_0^y \frac{u_\lambda^+(x,z)}{a(x)} dx,$$

where u_λ^+ is the continuous version of the resolvent density corresponding to U_λ^+. Recalling the symmetry relation,

$$\frac{1}{a(x)} u_\lambda^+(x,z) a(z) = u_\lambda^+(z,x),$$

we obtain

$$\hat{U}_\lambda^0 \mathbf{1}_{[z,\infty)}(y) = \hat{U}_\lambda^+ \mathbf{1}_{[0,y]}(z),$$

from which the first equality of (3.5) follows by uniqueness of Laplace transforms. The second equality may be proved by a similar method.

(3.6) and (3.7) can be proved by applying (3.5) successively: For example,

$$\begin{aligned} T^+_{t_1-t_0} \cdot T^-_{t_2-t_1}\mathbf{1}_{[0,y]}(x) &= \int_{[0,\infty)} T^+_{t_1-t_0}(x,du)T^-_{t_2-t_1}\mathbf{1}_{[0,y]}(u) \\ &= \int_{[0,\infty)} T^+_{t_1-t_0}(x,du)\widehat{T}^+_{t_2-t_1}\mathbf{1}_{[u,\infty)}(y) \\ &= \iint_{0\le u\le v<\infty} T^+_{t_1-t_0}(x,du)\widehat{T}^+_{t_2-t_1}(y,dv) \\ &= \int_{[0,\infty)} \widehat{T}^+_{t_2-t_1}(y,dv)T^+_{t_1-t_0}\mathbf{1}_{[0,v]}(x) \\ &= \int_{[0,\infty)} \widehat{T}^+_{t_2-t_1}(y,dv)\widehat{T}^0_{t_1-t_0}\mathbf{1}_{[x,\infty)}(v) \\ &= \widehat{T}^+_{t_2-t_1} \cdot \widehat{T}^0_{t_1-t_0}\mathbf{1}_{[x,\infty)}(y). \end{aligned}$$

This proves a particular case of (3.6). In the same way, the general case can be proved easily by induction. ■

Remark 3.1. *We remark that an alternative proof of (3.5) is possible by means of the time reversal of stochastic flows on the half line. A stochastic flow on the half line $[0,\infty)$ is defined similarly by replacing the whole line $\mathbf{R}$ by $[0,\infty)$ in Def.1.2. A key idea in the proof is to construct a stochastic flow $\mathbf{X} = (X_{s,t})$ on $[0,\infty)$ whose one-point motion $t \mapsto X_{0,t}(x), x \in \mathbf{R}$, is given by the absorbing L-diffusion $\xi^-(t)$, i.e., the diffusion with the semigroup T_t^-, and then show that its* time reversed flow $\widehat{\mathbf{X}} = (\widehat{X}_{s,t})$, *defined by $\widehat{X}_{s,t} = (X_{-t,-s})^{-1}$, has the one-point motion given by the reflecting $\widehat{L}$-diffusion $\widehat{\xi}^+(t)$, i.e., the diffusion with the semigroup $\widehat{T}_t^+$. Here, for a right-continuous and non-decreasing $\varphi : [0,\infty) \to [0,\infty)$ such that $\lim_{x\nearrow\infty}\varphi(x) = \infty$, φ^{-1} is the right-continuous inverse of φ: $\varphi^{-1}(x) = \inf\{y|\varphi(y) > x\}$. This is connected to the fact that L and $\widehat{L}$, when written in Hörmander form, differ only in the sign of the drift term. The corresponding fact in the case of stochastic flows of homeomorphisms is well-known (cf. [K] p.131, [IW] p.265).*

§4. Proof of Th. 1.3.

Consider a Harris flow $\mathbf{X}$ satisfying (1.4), (1.6) and (1.11).

Proof of Cor. 1.1. It is sufficient to show that the set of zeros of $\widehat{L}$-diffusion $\widehat{\xi}(t)$ has the Hausdorff dimension $(1-\alpha)/(2-\alpha)$, $\widehat{P}_0$-almosy surely. The set of zeros of $\widehat{\xi}(t)$ is the range of the inverse local time $l^{-1}(t)$ at 0 of $\widehat{\xi}(t)$, which is a subordinator with exponent $\Psi(\lambda)=g_\lambda(0,0)^{-1}$:

$$E\left(e^{-\lambda l^{-1}(t)}\right)=e^{-t\Psi(\lambda)}=e^{-t/g_\lambda(0,0)}.$$

Here, $g_\lambda(x,y)$ is the Green function (resolvent density) with respect to the speed measure dx of reflecting $\widehat{L}$-diffusion where $\widehat{L}=\frac{d}{dx}(1-b(x))\frac{d}{dx}$. If we introduce the scale $\xi=\int_0^x(1-b(y))^{-1}dy$ as the coordinate of $[0,\infty)$, then $\widehat{L}=(1-\tilde{b}(\xi))^{-1}\frac{d^2}{d^2\xi}$ where $\tilde{b}(\xi)=b(x(\xi))$, so that the speed measure in the new coordinate is given by $d\tilde{m}(\xi)=a(\xi)d\xi$ with $a(\xi)=1-\tilde{b}(\xi)$. It is easy to deduce from (1.11) that $a(\xi)\asymp\xi^{\alpha/(1-\alpha)}$ as $\xi\to 0$. Let $\tilde{g}_\lambda(\xi,\eta)$ be the Green function for $\widehat{L}$-diffusion with respect to the speed measure so that $\tilde{g}_\lambda(0,0)=g_\lambda(0,0)$. By Th.2.3 in p.243 of [KW], we have

$$\Psi(\lambda)=\tilde{g}_\lambda(0,0)^{-1}\asymp\lambda^{1/(2+\frac{\alpha}{1-\alpha})}=\lambda^{\frac{1-\alpha}{2-\alpha}}\quad\text{as }\lambda\to\infty.$$

Then we can conclude that the range of the subordinator $l^{-1}(t)$ has the Hausdorff dimension $\frac{1-\alpha}{2-\alpha}$ almost surely, by a result of Blumenthal and Getoor (cf. [B], p. 94, Th. 16). ■

Now we proceed to prove Th. 1.3. We need several lemmas.

Lemma 4.1. *(i) Let $\Phi_1,\Phi_2\in L_2^{us}(\mathcal{F}^X_{-\infty,\infty})$ and consider their linear combination $\Phi=\alpha\Phi_1+\beta\Phi_2\in L_2^{us}(\mathcal{F}^X_{-\infty,\infty})$. If $A\in\mathcal{B}(\mathcal{C})$ satisfies $P(S_{\Phi_1}\in A)=P(S_{\Phi_2}\in A)=1$, then it holds that $P(S_\Phi\in A)=1$.*
(ii) Let $\Phi_n\in L_2^{us}(\mathcal{F}^X_{-\infty,\infty})$, $n=1,2,\dots$, constitute a dense family in $L_2^{us}(\mathcal{F}^X_{-\infty,\infty})$. If $A\in\mathcal{B}(\mathcal{C})$ satisfies $P(S_{\Phi_n}\in A)=1$ for all n, then it holds that $P(S_\Phi\in A)=1$ for all $\Phi\in L_2^{us}(\mathcal{F}^X_{-\infty,\infty})$.

Proof. According to Theorem 3d12 of [T 5], every $A\in\mathcal{B}(\mathcal{C})$ is associated with a closed subspace $\mathcal{H}_A$ of $L_2(\mathcal{F}^X_{-\infty,\infty})$ such that the spectral measure μ_Φ of any Φ satisfies

$$||P_A\Phi||^2=\mu_\Phi(A),$$

where P_A denotes the orthogonal projection onto $\mathcal{H}_A$. Both parts of this lemma are immediate consequences. ■

Lemma 4.2. *Let $t_1<t_2<t_3$ and $\Phi=\Phi_1\Phi_2\in L_2^{us}(\mathcal{F}^X_{t_1,t_3})$ such that $\Phi_1\in L_2^{us}(\mathcal{F}^X_{t_1,t_2})$ and $\Phi_2\in L_2^{us}(\mathcal{F}^X_{t_2,t_3})$. Then,*

$$S_\Phi\cap[t_1,t_2]\stackrel{d}{=}S_{\Phi_1},\qquad S_\Phi\cap[t_2,t_3]\stackrel{d}{=}S_{\Phi_2}.$$

Furthermore, $S_\Phi\cap[t_1,t_2]$ and $S_\Phi\cap[t_2,t_3]$ are mutually independent.

The proof is easy and omitted.

Lemma 4.3. *Let S be a $C_{[0,1]}$-valued random variable and assume, for $0 < \beta < 1$ and $K > 0$, that*

$$P(S \cap [t, t+\epsilon] \neq \emptyset) \leq K\epsilon^{\beta} \quad \text{for all } 0 < \epsilon < 1 \quad \text{and} \quad t \in [0,1].$$

Then, $P(\dim S \leq 1 - \beta) = 1$.

Proof. For every $a > 1 - \beta$, we have

$$\begin{aligned} & E\left(\sum_{k=1}^{n} \mathbf{1}_{\{S \cap [\frac{k-1}{n}, \frac{k}{n}] \neq \emptyset\}} \cdot \left(\frac{1}{n}\right)^{a}\right) \\ = & \sum_{k=1}^{n} P\left(S \cap \left[\frac{k-1}{n}, \frac{k}{n}\right] \neq \emptyset\right) \cdot \left(\frac{1}{n}\right)^{a} \\ \leq & \; nK\left(\frac{1}{n}\right)^{\beta} \cdot \left(\frac{1}{n}\right)^{a} = K \cdot n^{1-(\beta+a)} \to 0 \quad \text{as} \quad n \to \infty. \end{aligned}$$

Hence, there exists a subsequence $n_\nu \to \infty$ such that, almost surely,

$$\sum_{k=1}^{n_\nu} \mathbf{1}_{\{S \cap [\frac{k-1}{n_\nu}, \frac{k}{n_\nu}] \neq \emptyset\}} \cdot \left(\frac{1}{n_\nu}\right)^{a} \to 0 \quad \text{as} \quad \nu \to \infty.$$

Let $\mathcal{C}_\nu$ be the collection of intervals $E_k = [\frac{k-1}{n_\nu}, \frac{k}{n_\nu}]$, $k = 1, \ldots, n_\nu$, which have nonempty intersections with the set S. Then $\mathcal{C}_\nu$ is a covering of S and

$$\sum_{E_k \in \mathcal{C}_\nu} (\text{diam } E_k)^{a} \to 0 \quad \text{a.s., as} \quad \nu \to \infty.$$

Hence, $\dim S \leq a$, a.s., implying that $\dim S \leq 1 - \beta$, a.s. ∎

Proof of Th. 1.3. It is sufficient to show that

$$\dim S_\Phi \leq \frac{1-\alpha}{2-\alpha} \quad \text{a.s.} \tag{4.1}$$

for $\Phi \in L_2^{us}(\mathcal{F}_{0,1}^X)$. Indeed, if (4.1) is true for $\Phi \in L_2^{us}(\mathcal{F}_{0,1}^X)$, then by the stationarity of the flow, it is also true for $\Phi \in L_2^{us}(\mathcal{F}_{n,n+1}^X)$. By Lemma 4.2, (4.1) is true for a finite product of such Φ's. Since linear combinations of such products are dense in $L_2(\mathcal{F}_{-\infty,\infty}^X)$, we can conclude by Lemma 4.1 that (4.1) is true for any $\Phi \in L_2^{us}(\mathcal{F}_{-\infty,\infty}^X)$.

First, we consider the case when $\Phi \in L_2^{us}(\mathcal{F}_{0,1}^X)$ is given by

$$\Phi = f(X_{0,1}(x_1), \ldots, X_{0,1}(x_n)), \quad x_1, \ldots, x_n \in \mathbf{R},$$

and a function f is uniformly Lipschitz-continuous on $\mathbf{R}^n$.

Let $F = [t, t+\epsilon]$, $0 \leq t < t+\epsilon \leq 1$, and let $(\mathbf{X}, \mathbf{X}')$ be a $(1^-, F)$-joining. Then we know by Lemma 2.2 that $2P(S_X^{acc} \cap F \neq \emptyset) = E(|X_{0,1}(0) - X'_{0,1}(0)|^2)$ and similarly, we have $2P(S_\Phi^{acc} \cap F \neq \emptyset) = E(|\Phi - \Phi'|^2)$ where $\Phi' = f(X'_{0,1}(x_1), \ldots, X'_{0,1}(x_n))$. Therefore, noting that $E(|X_{0,1}(x) - X'_{0,1}(x)|^2)$ is independent of x, we have

$$
\begin{aligned}
P(S_\Phi^{acc} \cap F \neq \emptyset) &= \frac{1}{2} E(|\Phi - \Phi'|^2) \\
&\leq KE(|X_{0,1}(0) - X'_{0,1}(0)|^2) = 2KP(S_X^{acc} \cap F \neq \emptyset) \qquad (4.2)
\end{aligned}
$$

where a constant K depends on n and the Lipschitz constant of f.

Let $\{\widehat{\xi}^+(t), \widehat{P}_\xi\}$ be the reflecting $\widehat{L}$-diffusion on $[0, \infty)$. As in the proof of Cor.1.1, *take a canonical scale* ξ *as the coordinate* so that $\widehat{L} = \frac{d^2}{a(\xi) d\xi^2}$ and we have $a(\xi) \asymp \xi^{\alpha/(1-\alpha)}$ as $\xi \to 0$ and $a(\xi) \to 1$ as $\xi \to \infty$. Let $\mu(d\xi) = da(\xi)$. By what we have shown above,

$$
\begin{aligned}
& P(S_X^{acc} \cap [t, t+\epsilon] \neq \emptyset) = P(\widetilde{S} \cap [t, t+\epsilon] \neq \emptyset) \\
= & \int_0^1 \widehat{P}_\mu \left(\widehat{\xi}^+(u-s) = 0 \text{ for some } s \in [0, u] \cap [t, t+\epsilon] \right) du \\
= & \int_t^1 \widehat{P}_\mu \left(\widehat{\xi}^+(\theta) = 0 \text{ for some } \theta \in [(u-t-\epsilon)_+, u-t] \right) du \\
= & O(\epsilon) + \int_t^1 \widehat{P}_\mu \left(\widehat{\xi}^+(\theta) = 0 \text{ for some } \theta \in [u-t, u-t+\epsilon] \right) du.
\end{aligned}
$$

We would show
(4.3)

$$
I(t) := \int_t^1 \widehat{P}_\mu \left(\widehat{\xi}^+(\theta) = 0 \text{ for some } \theta \in [u-t, u-t+\epsilon] \right) du = O\left(\epsilon^{\frac{1}{2-\alpha}} \right)
$$

as $\epsilon \to 0$ uniformly in $t \in [0, 1]$. If we can show this, then

$$
P(S_X^{acc} \cap [t, t+\epsilon] \neq \emptyset) = O(\epsilon^{1/(2-\alpha)})
$$

as $\epsilon \to 0$ uniformly in $t \in [0, 1]$ and, combining this with (4.2), we see that $P(S_\Phi^{acc} \cap [t, t+\epsilon] \neq \emptyset) = O(\epsilon^{1/(2-\alpha)})$, so that, by Lemma 4.3, we can conclude that the estimate (4.1) holds for Φ because $1 - 1/(2-\alpha) = (1-\alpha)/(2-\alpha)$.

To obtain (4.3), we estimate

$$
\begin{aligned}
I(t) &\le \int_0^1 \widehat{P}_\mu\left(\widehat{\xi}^+(\theta)=0 \text{ for some } \theta\in[u,u+\epsilon]\right)du \\
&= \int_0^1 \widehat{E}_\mu\left(\widehat{P}_{\widehat{\xi}^+(u)}[\widehat{\sigma}_0\le\epsilon]\right)du \le e\int_0^1 e^{-u}\widehat{E}_\mu\left(\widehat{P}_{\widehat{\xi}^+(u)}[\widehat{\sigma}_0\le\epsilon]\right)du \\
&\le e\int_0^\infty e^{-u}\widehat{E}_\mu\left(\widehat{P}_{\widehat{\xi}^+(u)}[\widehat{\sigma}_0\le\epsilon]\right)du \\
&= e\int_{[0,\infty)}\mu(d\xi)\int_{[0,\infty)}\tilde{g}_1(\xi,\eta)\widehat{P}_\eta[\widehat{\sigma}_0\le\epsilon]a(\eta)d\eta,
\end{aligned}
$$

where $\widehat{\sigma}_0$ is the first hitting time of $\widehat{\xi}^+(t)$ to 0. Since the resolvent density $\tilde{g}_1(\xi,\eta)$ is bounded, we have, for some $C>0$,

$$
I(t)\le C\int_{[0,\infty)}\widehat{P}_\eta[\widehat{\sigma}_0\le\epsilon]a(\eta)d\eta.
$$

The process $\widehat{\xi}^+(t)$ under $\widehat{P}_\eta$, $\eta>0$, and in the coordinate ξ, is obtained from a one-dimensional Brownian motion $B(t)$ with $B(0)=0$ by

$$
\widehat{\xi}^+(t)=|\eta+B(A^{-1}(t))| \quad \text{where} \quad A(t)=\int_0^t a(|\eta+B(s)|)ds.
$$

Hence,

$$
\widehat{P}_\eta(\widehat{\sigma}_0\le\epsilon)=P\left(\int_0^{\sigma_0}a(|\eta+B(s)|)ds\le\epsilon\right) \qquad \text{where} \quad \sigma_0=\min\{s|\eta+B(s)=0\}.
$$

and, noting $a(\xi)\ge K^{-1}\cdot\xi^{\alpha/(1-\alpha)}\wedge 1$ for some $K>0$,

$$
\widehat{P}_\eta(\widehat{\sigma}_0\le\epsilon)\le P\left(\int_0^{\sigma_0}\left(|\eta+B(s)|^{\alpha/(1-\alpha)}\wedge 1\right)ds\le K\epsilon\right).
$$

The scaling property of $B(t)$ combined with an easy inequality $(\epsilon a)\wedge 1\ge\epsilon(a\wedge 1)$ for $a>0$ and $1\ge\epsilon>0$ yields that the RHS is dominated by $\phi(\epsilon^{-(1-\alpha)/(2-\alpha)}\eta)$, where

$$
\phi(\eta)=P\left(\int_0^{\sigma_0}\left(|\eta+B(s)|^{\alpha/(1-\alpha)}\wedge 1\right)ds\le K\right).
$$

Then,

$$\begin{aligned} I(t) &\le C \int_{[0,\infty)} \phi(\epsilon^{-(1-\alpha)/(2-\alpha)}\eta)a(\eta)d\eta \\ &\le K' \int_{[0,\infty)} \phi(\epsilon^{-(1-\alpha)/(2-\alpha)}\eta)\eta^{\alpha/(1-\alpha)}d\eta \\ &= K'\epsilon^{1/(2-\alpha)} \int_{[0,\infty)} \phi(\eta)\eta^{\alpha/(1-\alpha)}d\eta \end{aligned}$$

and we have otained (4.3).

In the same way, we have the estimate (4.1) for $\Phi = f(X_{s,t}(x_1), \dots, X_{s,t}(x_n))$, $x_1, \dots, x_n \in \mathbf{R}$, $s < t$, where f is uniformly Lipschitz continuous on $\mathbf{R}^n$. Then, by Lemma 4.2, we have the estimate (4.1) for $\Phi = \Phi_1\Phi_2\cdots\Phi_m \in L_2^{us}(\mathcal{F}_{0,1}^X)$ if $t_0 = 0 < t_1 < t_2 < \cdots < t_m = 1$, and $\Phi_k \in \mathrm{ub}[L^2(\mathcal{F}_{t_{k-1},t_k}^X)]$, $k = 1, 2, \dots, m$, is given in the form $\Phi_k = f_k\left(X_{t_{k-1},t_k}(x_1^{(k)}), \dots, X_{t_{k-1},t_k}(x_{n_k}^{(k)})\right)$, $x_1^{(k)}, \dots, x_{n_k}^{(k)} \in \mathbf{R}$, where f_k is uniformly Lipschitz continuous on $\mathbf{R}^{n_k}$. By Lemma 4.1 (i), the estimate (4.1) still holds for a finite linear combination of such functionals and this class of functionals is dense in $L_2^{us}(\mathcal{F}_{0,1}^X)$. ■

References

[A] R. A. Arratia, Coalescing Brownian motions, and voter model on Z, Unpublished manuscript, Univ. of Southern California.

[B] J. Bertoin, *Lévy Processes*, Cambridge University Press, 1996

[H] T. E. Harris, Coalescing and noncoalescing stochastic flows in $\mathbf{R}_1$, Stochastic Processes and their Applications, **17**(1984), 187-210

[IW] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, Second Edition, North-Holland/Kodansha, Amsterdam/Tokyo, 1988

[IM] K. Itô and H. P. McKean, Jr., *Diffusion Processes and their Sample Paths*, Springer, Berlin, 1965, Second Printing 1974, in *Classics in Mathematics*, 1996

[K] H. Kunita, *Stochastic flows and stochastic differential equations,* Cambridge University Press, Cambridge, 1990

[KW] S. Kotani and S. Watanabe, Krein's spectral theory of strings and generalized diffusion processes, *in Functional Analysis in Markov Processes, ed. by M. Fukushima*, **LNM 923**, Springer(1982), 235-259

[LR 1] Y. Le Jan and O. Raimond, Flow, coalescence and noise, PR/0203221(2002)

[LR 2] Y. Le Jan and O. Raimond, The noise of Brownian sticky flow is black, PR/0212269(2003)

[T 1] B. Tsirelson, Within and beyond the reach of Brownian innovation, "Documenta Mathematica", Extra Volume ICM 1998, III, 311-320

[T 2] B. Tsirelson, Unitary Brownian motions are linearizable, PR/9806112(1998)

[T 3] B. Tsirelson, Brownian coalescence as a black noise. I, preprint 1998

[T 4] B. Tsirelson, Fourier-Walsh coefficients for a coalescing flow (discrete time), PR/9903068(1999)

[T 5] B. Tsirelson, Scaling Limit, Noise, Stability, In: Saint-Flour Summer School for Probability 2002, Lect. Notes in Math. to appear.

[VK] A. Ju. Veretennikov and N. V. Krylov, On explicit formulas for solutions of stochastic equations, Math. USSR Sbornik, **29**(1976), 239-256

Jon Warren
Department of Statistics, University of Warwick
Coventry, CV4 7AL, United Kingdom

Shinzo Watanabe
527-10, Chaya-cho, Higashiyama-ku
Kyoto, 605-0931, Japan