Natural OU-processes on Lie groups with Applications to simulated annealing

Josef Teichmann
(joint work with Fabrice Baudoin and Martin Hairer)

Vienna University of Technology,
Institute for Mathematical Methods in Economics

October 2007

We consider the following structures:

- G a finite dimensional Lie group, $\mathfrak{g}$ its Lie algebra.

We consider the following structures:

- G a finite dimensional Lie group, $\mathfrak{g}$ its Lie algebra.
- $d \geq 1$ and left-invariant vector fields $V_1, \ldots, V_d \in \mathfrak{g}$ such that

$$\langle V_1, \ldots, V_d \rangle_{\mathrm{LA}} = \mathfrak{g}.$$

We consider the following structures:

- G a finite dimensional Lie group, $\mathfrak{g}$ its Lie algebra.
- $d \geq 1$ and left-invariant vector fields $V_1, \ldots, V_d \in \mathfrak{g}$ such that

$$\langle V_1, \ldots, V_d \rangle_{\mathrm{LA}} = \mathfrak{g}.$$

- A right-invariant Haar measure μ on G.

We consider the following structures:

- G a finite dimensional Lie group, $\mathfrak{g}$ its Lie algebra.
- $d \geq 1$ and left-invariant vector fields $V_1, \ldots, V_d \in \mathfrak{g}$ such that

$$\langle V_1, \ldots, V_d \rangle_{\mathrm{LA}} = \mathfrak{g}.$$

- A right-invariant Haar measure μ on G.
- X the G-valued diffusion process

$$dX_t^x = \sum_{i=1}^{d} V_i(X_t^x) \circ dB_t^i$$
$$X_t^x = x \in G.$$

We consider the following structures:

- G a finite dimensional Lie group, $\mathfrak{g}$ its Lie algebra.
- $d \geq 1$ and left-invariant vector fields $V_1, \dots, V_d \in \mathfrak{g}$ such that

$$\langle V_1, \dots, V_d \rangle_{\text{LA}} = \mathfrak{g}.$$

- A right-invariant Haar measure μ on G.
- X the G-valued diffusion process

$$dX_t^x = \sum_{i=1}^{d} V_i(X_t^x) \circ dB_t^i$$

$$X_t^x = x \in G.$$

- $\mathcal{L} = \frac{1}{2} \sum_{i=1}^{d} V_i^2$ the generator of X and $P_t f(x) = E(f(X_t^x))$ its expectation functional on Borel functions.

We have the following facts:

- The generator $\mathcal{L}$ is symmetric on $C_c^\infty(G) \subset L^2(\mu)$ and admits a self-adjoint extension $\mathcal{L}$, which generates a symmetric, strongly continuous semigroup P on $L^2(\mu)$.

We have the following facts:

- The generator $\mathcal{L}$ is symmetric on $C_c^\infty(G) \subset L^2(\mu)$ and admits a self-adjoint extension $\mathcal{L}$, which generates a symmetric, strongly continuous semigroup P on $L^2(\mu)$.
- The carré du champ operator

$$\Gamma(f,g) = \mathcal{L}(fg) - f\mathcal{L}(g) - g\mathcal{L}(f)$$

defines a symmetric, positive bi-linear form

$$\mathcal{E}(f,g) = \int \Gamma(f,g)\mu = -2\int f\mathcal{L}(g)\mu$$

on test functions $f, g \in C_c^\infty(G)$.

We have the following facts:

- The generator $\mathcal{L}$ is symmetric on $C_c^\infty(G) \subset L^2(\mu)$ and admits a self-adjoint extension $\mathcal{L}$, which generates a symmetric, strongly continuous semigroup P on $L^2(\mu)$.
- The carré du champ operator

  $$\Gamma(f,g) = \mathcal{L}(fg) - f\mathcal{L}(g) - g\mathcal{L}(f)$$

  defines a symmetric, positive bi-linear form

  $$\mathcal{E}(f,g) = \int \Gamma(f,g)\mu = -2\int f\mathcal{L}(g)\mu$$

  on test functions $f, g \in C_c^\infty(G)$.
- The random variable X_t^e admits a smooth, positive density $p_t : G \to \mathbb{R}_{>0}$ with respect to μ.

- For smooth functions $U : G \to \mathbb{R}$ with
$$\int_G \exp(-U)\mu = Z^U < \infty,$$
we consider the vector field $\Gamma(U,.)$ and the operator
$$\mathcal{L}^U := \mathcal{L} - \frac{1}{2}\Gamma(U,.).$$
Then $\mu^U = \exp(-U)\mu$ is an invariant (finite) measure for $\mathcal{L}^U$, since
$$\int \mathcal{L}^U f \mu^U = \int \mathcal{L} f \exp(-U)\mu + \frac{1}{2}\int \Gamma(f, \exp(-U))\mu = 0$$
by symmetry of $\mathcal{L}$.

- For smooth functions $U : G \to \mathbb{R}$ with
$$\int_G \exp(-U)\mu = Z^U < \infty,$$
we consider the vector field $\Gamma(U,.)$ and the operator
$$\mathcal{L}^U := \mathcal{L} - \frac{1}{2}\Gamma(U,.).$$
Then $\mu^U = \exp(-U)\mu$ is an invariant (finite) measure for $\mathcal{L}^U$, since
$$\int \mathcal{L}^U f \mu^U = \int \mathcal{L} f \exp(-U)\mu + \frac{1}{2}\int \Gamma(f, \exp(-U))\mu = 0$$
by symmetry of $\mathcal{L}$.
- $\mathcal{L}^U$ has a spectral gap of size $a > 0$ if and only if
$$\int \Gamma(f,f)\mu^U \geq 2a\big(Z^U \int f^2 \mu^U - (\int f \mu^U)^2\big)$$
for compactly supported smooth functions f on G.

Proposition

Consider a smooth potential $U : G \to \mathbb{R}$ such that

$$\int \exp(-U)\,\mu < \infty \, .$$

Consider the Stratonovich SDE with values in G

$$dY_t^y = V_0(Y_t^y)dt + \sum_{i=1}^{d} V_i(Y_t^y) \circ dB_t^i, \quad Y_0^y = y \in G,$$

where $V_0 f = -\frac{1}{2}\Gamma(U, f)$ for test functions f. Then there exists a global strong solution for all initial values $y \in G$.

Corollary

We consider now the case

$$W_s(x) = -\log p(s,x),$$

with fixed $s > 0$. We write for short $\mathcal{L}^s = \mathcal{L}^{W_s}$.
By the previous Proposition there are strong solutions to the associated Stratonovich SDE on G,

$$dY_t^y = V_0(Y_t^y)dt + \sum_{i=1}^{d} V_i(Y_t^y) \circ dB_t^i, \quad Y_0^y = y \in G,$$

where $V_0 = \frac{1}{2}\sum_{i=1}^{d}(V_i \log p_s)V_i$ is the horizontal gradient of $-\log p_s$. We call this diffusion process an **OU-process** on G.

Proof

A direct proof by Lyapunov-function techniques would work if the following assertion holds true: there is a compact set K such that there is a constant $C > 0$ with

$$-\frac{1}{2}\Gamma(U,U) + \mathcal{L}U \leq CU$$

outside K. In case of W_s for some fixed $s > 0$ this is implied by the assertion that

$$\mathcal{L}p_s \geq 0$$

holds true outside a compact set $K \subset G$. By asymptotic expansions of the hypo-elliptic heat kernel we can prove the previous inequality for small times t on compact sets off the cut-locus, but not on non-compact sets.

Theorem

The following assertions are equivalent:

- The operator $\mathcal{L}^t$ has a spectral gap of size $a_t \geq 0$ for $t > 0$ and a positive H^1-function $a : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$.

Theorem

The following assertions are equivalent:

- The operator $\mathcal{L}^t$ has a spectral gap of size $a_t \geq 0$ for $t > 0$ and a positive H^1-function $a : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$.
- The local estimate

$$P_t(\Gamma(f,f))(e) \geq 2a_t((P_T f^2)(e) - ((P_t f)(e))^2)$$

holds true for all test functions $f : G \to \mathbb{R}$, $t > 0$ and a positive H^1-function $a : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$.

Theorem

If we know that

$$\Gamma(P_t f, P_t f)(e) \leq \phi(t) P_t(\Gamma(f,f))(e)$$

holds true for test functions $f : G \longrightarrow \mathbb{R}$, $t \geq 0$ and a positive locally integrable function $\phi : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{>0}$, then we can choose $t \mapsto a_t$ due to the following equation

$$a_t \int_0^t \phi(t-s) ds = \frac{1}{2}$$

for $t > 0$ and one of the two equivalent assertions holds true.

Proof

The equivalence of the first two statements is apparent. We fix a test function $f : G \to \mathbb{R}$ as well as $t > 0$ and consider

$$H(s) = P_s((P_{t-s}f)^2)$$

for $0 \leq s \leq t$. The derivative of this function equals

$$H'(s) = P_s(\Gamma(P_{t-s}f, P_{t-s}f))$$

and therefore – assuming the third statement – we obtain

$$H'(s) \leq \phi(t-s)P_t(\Gamma(f,f)).$$

Whence we can conclude

$$H(t) - H(0) \leq \int_0^t \phi(t-s)ds\, P_t(\Gamma(f,f)),$$

which is the second of the two equivalent assertions.

Let G be a nilpotent Lie group with d generators $e_1, \ldots, e_d$ and denote by X the canonical diffusion process on G, i.e.

$$dX_t = \sum_{i=1}^{d} X_t e_i \circ dB_t^i.$$

The Driver-Melcher inequality asserts that there is a constant K such that

$$\Gamma(P_t f, P_t f)(e) \leq K P_t(\Gamma(f, f))(e)$$

for all test functions $f : G \to \mathbb{R}$ and for all times $t \geq 0$ holds true. This allows to conclude that there is a spectral gap for $\mathcal{L}^t$. We can choose $a_t K t = \frac{1}{2}$, hence the OU-process on G (associated to $W_t = -\log p_t$) has a spectral gap of size $\frac{1}{2Kt}$.

Spectral gaps for general potentials

Fix $t > 0$ and let $U : G \to \mathbb{R}$ be a smooth potential U with

$$|U + \log p(t, .)| \leq K$$

for some constant $K > 0$. Assume furthermore that a Poincaré inequality holds for $\mathcal{L}^t$, i.e.

$$P_t(f^2)(e) \leq C P_t(\Gamma(f, f))(e)$$

for test function $f : G \to \mathbb{R}$ with $P_t f(e) = 0$ and some constant $C > 0$. Then one has $\exp(-U) \in L^1(\mu(dx))$ and the desired Poincare inequality

$$\int f^2 \exp(-U)\mu \leq C' \int \Gamma(f, f) \exp(-U)\mu$$

for test function $f : G \to \mathbb{R}$ with $\int f \exp(-U)\mu = 0$ and some constant $C' = C \exp(2K) > 0$. The second inequality leads to a spectral gap for $\mathcal{L}^U$ of size $\frac{1}{C'}$.

Let G be a nilpotent Lie group and M a homogeneous space with respect to the Lie group G (with right action) and projection $\pi : G \to M$. We assume a measure μ^M on M, which is invariant with respect to the action.
The left invariant vector fields V_i induce vector fields V_i^M on M by means of the action. Due to the invariance of the measure with respect to the action, the vector fields V_i^M are anti-symmetric operators on $L^2(\mu^M)$ and the generator

$$\mathcal{L}^M = \frac{1}{2} \sum_{i=1}^{d} (V_i^M)^2$$

is consequently symmetric on $L^2(\mu^M)$. In particular we have

$$(V_i^M f) \circ \pi = V_i(f \circ \pi)$$

for $i = 1, \ldots, d$.

The local Driver-Melcher inequality on G therefore translates to the same local inequality on M due to

$$P_t^M(f) \circ \pi = P_t(f \circ \pi)$$

for test functions $f : M \to \mathbb{R}$ and $t \geq 0$, hence we obtain the corresponding Driver-Melcher inequality on M.

Simulated Annealing on nil-manifolds

Let M be a compact nil-manifold, $U : M \to \mathbb{R}$ a smooth function together with a constant D and a point $x_0 \in M$ such that

$$|U(x) - d(x_0, x)^2| \leq D.$$

Then there constants $R, c > 0$ such that for $\varepsilon(t) = \frac{c}{\sqrt{\log(R+t)}}$ the process

$$dZ_t^z = V_0(Z_t^z)dt + \sum_{i=1}^{d} \varepsilon(t) V_i(Z_t^z) \circ dB_t^i, \quad Z_0^z = z \in G,$$

with $V_0 f = -\frac{1}{2}\Gamma(U, f)$ satisfies for a constant A

$$\mathrm{var}_{\varepsilon(t)}(f) \leq A(1+t)\mathcal{E}_{\varepsilon(t)}(f, f),$$

for test functions $f : M \to \mathbb{R}$ and $t \geq 0$.

Proof

We apply R. Leandré's beautiful result that

$$\lim_{t \to 0} t \log p(t, x_0, x) = -d(x_0, x)^2$$

uniformly on the compact manifold M. This result yields that we can choose a constant $\tilde{D}$ such that

$$|U(.) + \varepsilon^2 \log p(\varepsilon^2, x_0, .)| \leq D + \tilde{D}$$

for all $0 < \varepsilon < 1$. Hence we can start to collect results. The spectral gap for the operator $\mathcal{L}^\varepsilon$ has size

$$\frac{1}{K\varepsilon^2} \exp(-\frac{2(D + \tilde{D})}{\varepsilon^2})$$

for $\varepsilon < 1$.

Proof

Whence for $\varepsilon^2\mathcal{L}^\varepsilon$ the spectral gap has size

$$\frac{1}{K}\exp\left(-\frac{2(D+\tilde{D})}{\varepsilon^2}\right).$$

We choose $c^2 = 2(D+\tilde{D})$ and R so big that $\varepsilon(t) < 1$ for $t \geq 0$, and we conclude that

$$K\exp\left(\frac{2(D+\tilde{D})}{\varepsilon(t)^2}\right) \leq K(R+t) \leq A(1+t)$$

for some large constant A and for $t \geq 0$. □

The algorithm

Assume that U has a unique minimizer at $x_0 \in M$ then we obtain as a consequence of the previous Theorem that

$$Z_t^z \longrightarrow \delta_{x_0},$$

where we obtain precise estimates on the speed of convergence, too.

Important References

G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Ann. Sci. École Norm. Sup. (4)* **21**, no. 3, (1988), 307–331.

B. K. Driver and T. Melcher. Hypoelliptic heat kernel inequalities on the Heisenberg group. *J. Funct. Anal.* **221**, no. 2, (2005), 340–365.

R. Holley and D. Stroock. Simulated Annealing via Sobolev Inequalities. *Comm. in Math. Physics* **115**, (1988), 553–569.

R. Léandre. Développement asymptotique de la densité d'une diffusion dégénérée. *Forum Math.* **4**, no. 1, (1992), 45–75.

X.-M. Li. Stochastic flows on noncompact manifolds, 1992. Ph.D. thesis.