5 Stochastic flows

5a Flows of maps: definition and simple examples

5a1 Definition. A *representation* of a measurable semigroup[1] G on a measurable space Z is a measurable map $Z \times G \ni (z,g) \mapsto zg \in Z$ satisfying

$$z(g_1 g_2) = (zg_1)g_2 \quad \text{for all } z \in Z \text{ and } g_1, g_2 \in G\,.$$

Given an abstract stochastic flow[2] $(X_{s,t})_{s\le t}$ in a measurable semigroup G and a representation of G on Z, we get a *stochastic flow* (of maps), — a family of Z-valued random variables $zX_{s,t}$ for all $s \le t$ and $z \in Z$. Note that

$$zX_{r,t} = (zX_{r,s})X_{s,t}\,,$$

$X_{r,s}$ and $X_{s,t}$ being independent. Recall also that the distribution of $X_{s,t}$ depends on $t-s$ only. The Z-valued process $(zX_{0,t})_t$ is called the *one-point motion* (of the flow) starting at z. The Z^n-valued process $(z_1X_{0,t},\dots,z_nX_{0,t})_t$ is called the *n-point motion* (of the flow) starting at $(z_1,\dots,z_n \in Z^n$. Thus, the n-point motion consists of n (dependent!) one-point motions.

5a2 Example. Using the standard Brownian motion $(B_t)_t$ in $\mathbb{R}$ and the representation $(z,g) \mapsto z+g$ of $\mathbb{R}$ on $\mathbb{R}$ we get a flow of shifts $\mathbb{R} \to \mathbb{R}$,

$$(z,s,t) \mapsto z + B_t - B_s\,.$$

A one-point motion $t \mapsto z + B_t$ is the Brownian motion starting at z. Any two one-point motions (starting at different points) keep a constant distance.

5a3 Example. Every Brownian motion $(Y_t)_t$ in $\mathrm{SO}(n)$ leads (via the natural representation of $\mathrm{SO}(n)$ on $\mathbb{R}^n$) to a flow of rotations $\mathbb{R}^n \to \mathbb{R}^n$,

$$(z,s,t) \mapsto zY_{s,t} = zY_s^{-1}Y_t\,.$$

Having $Y_t = \mathrm{Texp}\big(i\int_0^t dX_s\big)$ we may write informally

$$(z,t,t+dt) \mapsto z\exp(i\,dX_t) = z + iz\,dX - \frac12 z(dX)^2\,.$$

Especially, the case of 3c3 gives

$$(z,t,t+dt) \mapsto z + iz\sigma_1\,dB_1(t) + iz\sigma_2\,dB_2(t) + iz\sigma_3\,dB_3(t) - z\,dt\,.$$

A one-point motion is a Brownian motion on a sphere. Any two one-point motions keep a constant distance.

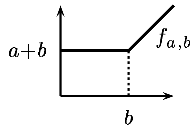
From Brownian motions in groups we turn now to abstract stochastic flows in semigroups. It is straightforward to generalize Definition 3a3 (morphisms) to abstract stochastic flows.

---

[1]Recall 1a5.
[2]Recall 1b6.

5a4 Example. Consider the measurable semigroup G_2 of maps $[0,\infty) \to [0,\infty)$ of the form

$$f_{a,b}(x) = a + \max(x, b) \,,$$



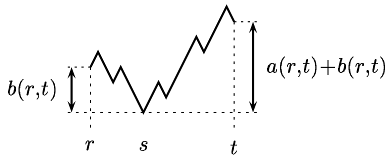
where $b \ge 0$, $a + b \ge 0$. It is indeed a semigroup:

$$f_{a_1,b_1} f_{a_2,b_2} = f_{a,b} \,, \qquad \begin{aligned} a &= a_1 + a_2 \,, \\ b &= \max(b_1, b_2 - a_1) \,. \end{aligned}$$

Alternatively one may define G_2 as the set $\{(a,b) : a,b \in \mathbb{R},\ b \ge 0,\ a+b \ge 0\}$ equipped with the composition law written above. Note that G_2 is not commutative. Note also a homomorphism $G_2 \to \mathbb{R}$, $f_{a,b} \mapsto a$.

Here is a morphism of the standard Brownian motion $(B_t)_t$ in $\mathbb{R}$ to an abstract stochastic flow in G_2;

$$a(r,t) = B_{r,t} \,, \quad b(r,t) = - \min_{s \in [r,t]} B_{r,s} \,,$$
$$a(r,t) + b(r,t) = \max_{s \in [r,t]} B_{s,t} \,.$$



In fact, the corresponding convolution semigroup $(\mu_t^{(2)})_t$ on G_2 can be calculated explicitly:

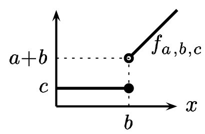
$$\frac{\mu_t^{(2)}(da\,db)}{da\,db} = \frac{2(a+2b)}{\sqrt{2\pi}\, t^{3/2}} \exp\Big(- \frac{(a+2b)^2}{2t} \Big) \quad \text{for } b \ge 0, a+b \ge 0 \,.$$

Using the natural representation of G_2 on $[0,\infty)$ we get a stochastic flow (of maps) on $[0,\infty)$. A one-point motion $t \mapsto a(0,t) + \max(b(0,t), x) = B_t + \max\big(x, \max_{[0,t]}(-B)\big)$ is the so-called *reflecting Brownian motion.* In fact, the one-point motion $t \mapsto a(0,t) + b(0,t) = \max_{s \in [0,t]} B_{s,t}$ starting at 0 is distributed like $(|B_{0,t}|)_t$.[3] Two one-point motions keep a constant distance until one of them reaches the origin. Generally, the distance decreases and ultimately vanishes forever, — the two particles *coalesce.*

5a5 Exercise. Find the support of the single-time distribution of a two-point motion. That is, the support of $(z_1 X_{0,t}, z_2 X_{0,t})$.

5a6 Example. Consider the measurable semigroup G_3 of maps $[0,\infty) \to [0,\infty)$ of the form

$$f_{a,b,c}(x) = \begin{cases} c & \text{for } 0 \le x \le b, \\ x + a & \text{for } x > b \end{cases}$$



where $b \ge 0$ and $0 \le c \le a + b$. It is indeed a semigroup:

$$f_{a_1,b_1,c_1} f_{a_2,b_2,c_2} = f_{a,b,c} \,, \qquad \begin{aligned} a &= a_1 + a_2 \,, \\ b &= \max(b_1, b_2 - a_1) \,, \end{aligned} \qquad c = \begin{cases} a_2 + c_1 & \text{if } c_1 > b_2, \\ c_2 & \text{otherwise} \end{cases}$$

[3]Moreover, the pair $\big(a(0,t) + b(0,t), b(0,t)\big)_t$ is distributed like the pair of $|B_t|$ and the so-called local time of B at 0. Note that $b(0,\cdot)$ increases (continuously) only when $a(0,\cdot) + b(0,\cdot) = 0$ (thus, on a closed set of Lebesgue measure zero).

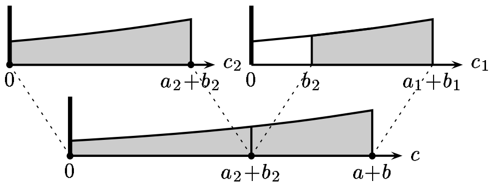
Clearly, $G_2 \subset G_3$, and the abstract stochastic flow in G_2 (introduced by 5a4) is also an abstract stochastic flow in G_3. However, its convolution semigroup $\big(\mu_t^{(2)}\big)_t$ is only a special case ($\lambda = 0$) of a family of convolution semigroups $\big(\mu_t^{(3,\lambda)}\big)_t$ on G_3; the parameter λ runs over $[0,\infty)$. Namely,

(5a7)
$$\begin{gathered}\mu_t^{(3,\lambda)} \text{ is the joint distribution of } a, b \text{ and } c = \max(0, a+b-\lambda\eta)\,, \\ \text{where } (a,b) \sim \mu_t^{(2)} \\ \text{and } \eta \sim \operatorname{Exp}(1) \text{ is independent of } (a,b)\,.\end{gathered}$$

It is indeed a convolution semigroup, due to a property of the composition in G_3: for every a_1, b_1, a_2, b_2,

$$\begin{gathered}\text{if } c_1 \sim \max(0, a_1+b_1-\lambda\eta) \quad \text{and} \quad c_2 \sim \max(0, a_2+b_2-\lambda\eta) \\ \text{then } c \sim \max(0, a+b-\lambda\eta)\,.\end{gathered}$$

(5a8)



Note that the measure $\mu_t^{(3,\lambda)}$ has an absolutely continuous part (its three-dimensional density can be written explicitly, using the two-dimensional density of $\mu_t^{(2)}$ and the one-dimensional exponential density of η) and a singular part concentrated on the plane $c = 0$; the singular part has a two-dimensional density (it can also be written explicitly).

Having no appropriate morphism,[4] we use Theorem 1c3 for constructing an abstract stochastic flow corresponding to $\big(\mu_t^{(3,\lambda)}\big)_t$. We identify G_3 with the subset $\{(a,b,c) : b \ge 0,\ 0 \le c \le a+b\}$ of $\mathbb{R}^3$ and equip G_3 with the topology induced from $\mathbb{R}^3$. Say,

$$\rho_0(f_{a_1,b_1,c_1}, f_{a_2,b_2,c_2}) = \max\big(|a_1-a_2|, |b_1-b_2|, |c_1-c_2|\big)\,.$$

It does not turn G_3 into a topological semigroup! The composition is discontinuous (when $c_1 = b_2$). However,

$$\rho_0(g_1g_2, g_1) \le \rho_0(g_2, e) \quad \text{and} \quad \rho_0(g_1g_2, g_2) \le \rho_0(g_1, e)$$

(check it). Also, the convolution semigroup is continuous in probability[5] (check it). And of course, each measure $\mu_t^{(3,\lambda)}$ is concentrated on a countable union of compact sets. Existence of the abstract stochastic flow is thus ensured. Sample continuity follows from 1e5.

Note the homomorphism $G_3 \to G_2$, $f_{a,b,c} \mapsto f_{a,b}$. It leads to a morphism of the flow constructed here (irrespective of λ) to the flow of 5a4.

The one-point motion (corresponding to the constructed flow) is the so-called *sticky Brownian motion.* It spends a positive time at the origin, but never sits there during a time interval. Two one-point motions keep a constant distance until one of them reaches the origin. Generally, the distance is non-monotone. But ultimately two particles coalesce.

[4] We will prove (in 5c) that it does not exist!

[5] Recall 1c2.

5a9 Exercise. Find the support of the single-time distribution of a two-point motion (once again). That is, the support of $(z_1 X_{0,t}, z_2 X_{0,t})$.

5b Flows of kernels: definition and simple examples[6]

Given a measurable space Z, we denote by $\mathcal{P}(Z)$ the set of all probability measures on Z. We turn $\mathcal{P}(Z)$ into a new measurable space, equipping it with the σ-field generated by functions $\mathcal{P}(Z) \ni \mu \mapsto \mu(A) \in \mathbb{R}$ where A runs over all measurable subsets of Z.[7]

We define a *kernel* on Z as a measurable map $K : Z \to \mathcal{P}(Z)$. Some convenient notation:

$$\begin{aligned} zK &\quad \text{rather than} \quad K(z)\,, \\ zKA &\quad \text{rather than} \quad K(z)(A)\,, \\ \mu f &\quad \text{rather than} \quad \textstyle\int f\,d\mu\,, \\ zKf &\quad \text{rather than} \quad \textstyle\int f\,dK(z)\,, \\ Kf &\quad \text{rather than} \quad z \mapsto \textstyle\int f\,dK(z)\,, \\ \mu K &\quad \text{rather than} \quad A \mapsto \textstyle\int K(z)(A)\,d\mu(z)\,; \end{aligned}$$

here K is a kernel, $z \in Z$, $A \subset Z$ a measurable set, $f : Z \to \mathbb{R}$ a bounded measurable function, and $\mu \in \mathcal{P}(Z)$. Given two kernels K_1, K_2 on Z, their composition, the kernel $K = K_1 K_2$ on Z is defined by

$$zK = (zK_1)K_2 \quad \text{for } z \in Z\,.$$

Note that $z(K_1K_2)f = (zK_1)(K_2 f)$.

A kernel on Z may be thought of as a map $\mathcal{P}(Z) \to \mathcal{P}(Z)$ (a very special map, of course), and the composition of these maps conforms to the composition of kernels.

5b1 Definition. A *kernel representation* of a measurable semigroup G on a measurable space Z is a family of kernels $(K_g)_{g \in G}$ on Z such that

(a) $K_{g_1 g_2} = K_{g_1} K_{g_2}$ for all $g_1, g_2 \in G$;

(b) the function $Z \times G \ni (z, g) \mapsto zK_g \in \mathcal{P}(Z)$ is measurable.

Note that (b) means just measurability of $(z, g) \mapsto zK_g A$ for every A. We may identify K_g with g writing zg rather than zK_g; then $z(g_1 g_2) = (zg_1)g_2$.

A kernel representation of G on Z may be thought of as a representation of G on $\mathcal{P}(Z)$ (a very special representation, of course).

5b2 Example. A kernel representation of the measurable semigroup G_2 (introduced by 5a4) on the measurable space $\mathbb{R}$:

$$x f_{a,b} = \begin{cases} \delta(x-a) & \text{for } x \in (-\infty, -b), \\ \frac12 \delta(-a-b) + \frac12 \delta(a+b) & \text{for } x \in [-b, b], \\ \delta(x+a) & \text{for } x \in (b, \infty); \end{cases}$$

here $\delta(z)$ is the atom at z (of mass 1).

[6] See also: Y. Le Jan, O. Raimond. *Flows, coalescence and noise.*

[7] Or equivalently, over an algebra that generates the σ-field.

Given a kernel representation of G on Z and an abstract stochastic flow $(X_{s,t})_{s\le t}$ in G, we get a stochastic flow (of maps) $\mu \mapsto \mu X_{s,t}$ on $\mathcal{P}(Z)$ of a special kind; it corresponds to a *stochastic flow of kernels,* — a family of $\mathcal{P}(Z)$-valued random variables $zX_{s,t}$ for all $s \le t$ and $z \in Z$. Still,

$$zX_{r,t} = (zX_{r,s})X_{s,t} \,,$$

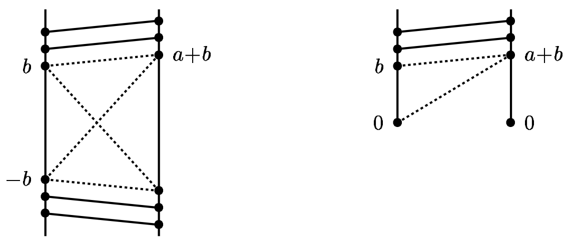
but now these are measures, not points.

5b3 Example. Combining the abstract stochastic flow in G_2 (and the morphism), introduced by 5a4, with the kernel representation introduced by 5b2 we get for $x > 0$ such a sample path in $\mathcal{P}(\mathbb{R})$:

$$xf_{a,b} = \begin{cases} \delta(x+a) & \text{for } b < x, \\ \frac12\delta(-a-b) + \frac12\delta(a+b) & \text{for } x \le b, \end{cases}$$
$$= \begin{cases} \delta(x+B_{0,t}) & \text{for } -\min_{s\in[0,t]} B_{0,s} < x, \\ \frac12\delta(-\max_{s\in[0,t]} B_{s,t}) + \frac12\delta(\max_{s\in[0,t]} B_{s,t}) & \text{for } x \le -\min_{s\in[0,t]} B_{0,s}]. \end{cases}$$

The atom at x moves according to $x + B_{0,t}$ until it reaches the origin. Then it splits in two symmetric atoms. Sometimes the two halves meet at the origin and separate again.

The map $\mathbb{R} \ni x \mapsto |x| \in [0,\infty)$ transforms this stochastic flow of kernels to the stochastic flow of maps on $[0,\infty)$ introduced by 5a4. Here we identify a stochastic flow of maps with the corresponding stochastic flow of degenerate (deterministic) kernels.



$x \mapsto xf_{a,b}$ the kernel $\qquad$ $x \mapsto f_{a,b}(x)$ the map

A particle moves according to the Brownian motion until reaching the origin. Here it decides: to cross the origin, or reflect. (However, it visits the origin a continuum of times.) Still, its law is Brownian.[8]

Particles are conditionally independent, given $(B_t)_t$. That is, their decisions at the origin (crossing or reflecting) are independent. Do not think that, introducing additional randomness, we could get a flow of maps that describes the motion of the continuum of particles. This is impossible. Indeed, the additional randomness should contain a continuum of independent random variables. And even if two particles start together, they can (and will) separate. Each point may contain a lot of particles…

[8]In fact, if the particle starts at the origin then its motion $(\alpha_t)_t$ is distributed like $(B_t)_t$, and B_t equals to $|\alpha_t|$ minus the local time of α at 0.

More formally, the one-point motion $(\alpha_t)_t$, conditioned on $(X_{s,t})_{s\le t}$, at a time t is distributed $\alpha_0 X_{0,t}$; in other words,

$$\mathbb{E}\big(f(\alpha_t) \,\big|\, X \big) = \alpha_0 X_{0,t} f \,;$$

accordingly, its unconditional distribution is

$$\mathbb{E} f(\alpha_t) = \mathbb{E}(\alpha_0 X_{0,t} f) = \mathbb{E} \int_Z f(z)(\alpha_0 X_{0,t})(dz)\,. \tag{5b4}$$

However, we did not really define the process $(\alpha_t)_t$.[9] Rather, we define its single-time distributions by (5b4), keeping $\mathbb{E}\big(f(\alpha_t) \,\big|\, X \big)$ an informal idea.

Similarly, we define the joint distributions (at time t) of n particles,

$$\mathbb{E} f\big(\alpha_1(t), \dots, \alpha_n(t)\big) = \mathbb{E} \int \cdots \int_{Z^n} f(z_1, \dots, z_n)\, (\alpha_1(0) X_{0,t})(dz_1) \dots (\alpha_n(0) X_{0,t})(dz_n)\,,$$

treating the particles as being conditionally independent given X. In other words, every kernel K on Z leads to a kernel $K^{\otimes n}$ on Z^n, namely,

$$K^{\otimes n}(z_1, \dots, z_n) = K(z_1) \times \cdots \times K(z_n)\,,$$

and so, the joint distribution (at time t) of n particles $\alpha_1(t), \dots, \alpha_n(t)$ is

$$\mathbb{E} K^{\otimes n}_{X_{0,t}}\big(\alpha_1(0), \dots, \alpha_n(0)\big)\,.$$

Moreover, every kernel K on Z leads to a kernel $K^{\otimes\infty}$ on $Z^\infty = Z \times Z \times \dots$, namely

$$K^{\otimes\infty}(z_1, z_2 \dots) = K(z_1) \times K(z_2) \times \dots$$

and we may introduce another kernel $P_{\infty,t}$ on Z^∞ by

$$P_{\infty,t}(z_1, z_2, \dots) = \mathbb{E} K^{\otimes\infty}_{X_{0,t}}\,.$$

5b5 Exercise. The family of kernels $(P_{\infty,t})_t$ is a one-parameter semigroup. That is,

$$P_{\infty,s} P_{\infty,t} = P_{\infty,s+t}\,.$$

Prove it.

5b6 Exercise. Each kernel $P_{\infty,t}$ satisfies

$$\int f(y_{k_1}, y_{k_2}, \dots)\, P_{\infty,t}(x_1, x_2, \dots; dy_1, dy_2, \dots) = \\ = \int f(y_1, y_2, \dots)\, P_{\infty,t}(x_{k_1}, x_{k_2}, \dots; dy_1, dy_2, \dots)$$

for all $k_1, k_2, \dots \in \{1, 2, 3, \dots\}$.

Prove it.

In fact, under some conditions kernels on Z^∞ are in a natural one-to-one correspondence with probability distributions on the space of kernels on Z.

[9]It can be done, see Le Jan and Raimond, the section 'Sampling the flow'.

5c Failure of the classical approach

By the classical approach I mean investigation of a given (abstract) stochastic flow by means of a morphism of a standard Brownian motion in $\mathbb{R}^m$ to the given flow. According to Sect. 4, the classical approach is successful for all Brownian motions in $\mathrm{SO}(n)$, thus, for all flows considered in 5a3. Now we will see that it fails for the flow introduced in 5a6 (sticky at the origin).

5c1 Theorem. No morphism exists from the standard Brownian motion in $\mathbb{R}^m$ (for whatever m) to the abstract stochastic flow introduced in 5a6 (for $\lambda \ne 0$).

The proof[10] is based on the idea of influence of a variable, useful both in computer science and probability theory. Given independent random variables $U_1, \dots, U_n$ and $V = f(U_1, \dots, U_n)$, the *influence* of U_n (on V) is, by definition,

$$\mathbb{E} \sqrt{\operatorname{Var} \big(V \,\big|\, U_1, \dots, U_{n-1} \big)} \; ;$$

of course, the conditional variance $\operatorname{Var}\big(V \,|\, \dots \big) = \mathbb{E}\big(V^2 \,|\, \dots \big) - \big(\mathbb{E}\big(V \,|\, \dots \big) \big)^2$. The influence of U_k is defined similarly. (Think, what does it mean when $U_1, \dots, U_n, V$ take on two values only.) The sum of squared influences,

$$H(f) = \sum_{k=1}^n \Big(\mathbb{E} \sqrt{\operatorname{Var} \big(V \,\big|\, U_1, \dots, U_{k-1}, U_{k+1}, \dots, U_n \big)} \Big)^2 ,$$

was used by Benjamini, Kalai and Schramm.[11] Note that $H(f) = \operatorname{Var}(V)$ for the linear case, $V = c_1 U_1 + \dots + c_n V_n$, but not in general.

Given a function $X \in L_2(\Omega, \mathcal{F}^B_{0,1}, P)$ of a Brownian path $B|_{[0,1]}$, we may define

$$H_n(X) = \sum_{k=1}^n \Big(\mathbb{E} \sqrt{\operatorname{Var} \big(X \,\big|\, \mathcal{F}^B_{0,\frac{k-1}{n}} \vee \mathcal{F}^B_{\frac{k}{n},1} \big)} \Big)^2 ,$$

where $\mathcal{F}^B_{s,t}$ is the σ-field generated by $B_{u,v}$ for all $(u,v) \subset (s,t)$.

5c2 Lemma. If $H_n(X) \to 0$ for $n \to \infty$, then $X = \text{const}$.

5c3 Exercise. If $\operatorname{Var}(X) \ne 0$ then there exist an interval $(s,t) \subset (0,1)$ and a random variable $Y \in L_\infty(\Omega, \mathcal{F}^B_{0,s} \vee \mathcal{F}^B_{t,1}, P)$ such that $\mathbb{E}\,(X B_{s,t} Y) \ne 0$.

Prove it.

Hint. First, replace $L_\infty(\Omega, \mathcal{F}^B_{0,s} \vee \mathcal{F}^B_{t,1}, P)$ with $L_2(\Omega, \mathcal{F}^B_{0,s} \vee \mathcal{F}^B_{t,1}, P)$. Second, $X = \sum_n \int \dots \int f_n(t_1, \dots, t_n) \, dB_{t_1} \dots dB_{t_n}$; consider the least n such that $\|f_n\| \ne 0$.

---

[10]Several proofs are known.

[11]They proved, roughly speaking, that a small $H(f)$ implies noise sensitivity of f. See Theorem 1.3 in: I. Benjamini, G. Kalai, O. Schramm (1999) *Noise sensitivity of Boolean functions and applications to percolation.* Inst. Hautes Études Sci. Publ. Math. no. 90, 5–43.

5c4 Exercise. Prove Lemma 5c2.

Hint. Use 5c3. Prove that

$$|\mathbb{E}(XB_{u,v}Y)| \le \sqrt{v-u}\,\|Y\|_\infty \sqrt{\operatorname{Var}\big(X \,\big|\, \mathcal{F}^B_{0,u} \vee \mathcal{F}^B_{v,1}\big)}$$

for every subinterval $(u,v) \subset (s,t)$.

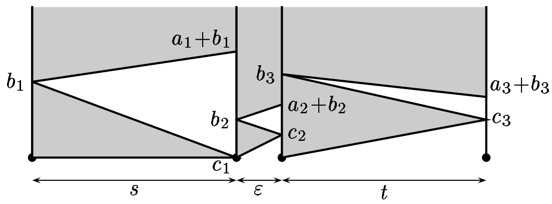
If $X \in L_2(\Omega, \mathcal{F}^B_{0,\infty}, P)$ and $H_n(X) \to 0$ then $X \in L_2(\Omega, \mathcal{F}^B_{1,\infty}, P)$. Similarly, for two independent Brownian motions B_1, B_2, if $X \in L_2(\Omega, \mathcal{F}^{B_1,B_2}_{0,1}, P)$ and

$$\sum_{k=1}^n \bigg(\mathbb{E} \sqrt{\operatorname{Var}\big(X \,\big|\, \mathcal{F}^{B_1,B_2}_{0,\frac{k-1}{n}} \vee \mathcal{F}^{B_2}_{\frac{k-1}{n},\frac{k}{n}} \vee \mathcal{F}^{B_1,B_2}_{\frac{k}{n},1}\big)} \bigg)^2 \xrightarrow[n\to\infty]{} 0\,,$$

then $X \in L_2(\Omega, \mathcal{F}^{B_2}_{0,1}, P)$.

5c5 Lemma. Let $(X_{s,t})_{s\le t}$ be the abstract stochastic flow introduced in 5a6 (for $\lambda \ne 0$), $X_{s,t} = f_{a(s,t),b(s,t),c(s,t)}$.[12] Let $\varphi : G_3 \to \mathbb{R}$ be a bounded Borel function. Then

$$\sum_{k=1}^n \bigg(\mathbb{E} \sqrt{\operatorname{Var}\big(\varphi(X_{0,1}) \,\big|\, \mathcal{F}^X_{0,\frac{k-1}{n}} \vee \mathcal{F}^a_{\frac{k-1}{n},\frac{k}{n}} \vee \mathcal{F}^X_{\frac{k}{n},1}\big)} \bigg)^2 \xrightarrow[n\to\infty]{} 0\,.$$



Proof. (Sketch.) Replacing $c\big(\frac{k-1}{n}, \frac{k}{n}\big)$ with 0 we do not change $X_{0,1}$ unless $b_3 < a_2 + b_2$. The latter event is of a small probability, namely, $O(1/\sqrt{n})$.[13] If the event does not happen, then the conditional variance vanishes. And if the event happens, then the conditional variance is small, namely, $O(1/\sqrt{n})$, since such is the conditional probability of $c_2 \ne 0$.[14] We get

$$\mathbb{E}\sqrt{\operatorname{Var}(\dots \mid \dots)} = O\Big(\frac{1}{\sqrt{n}} \cdot \sqrt{\frac{1}{\sqrt{n}}}\Big) = O(n^{-3/4}) = o(n^{-1/2})\,.$$

□

The theorem follows.

[12] Recall that a is a Brownian motion.
[13] Provided that k/n is bounded away from 0 and 1.
[14] Unless $a_2 + b_2$ is much greater than $1/\sqrt{n}$.