4 Example: Warren's "noise made by a Poisson snake"

This section is devoted to a noise discovered and investigated by J. Warren in a manuscript "The noise made by a Poisson snake" [12].

4a Three discrete semigroups: algebraic definition

A discrete semigroup (with unit; non-commutative, in general) may be defined by generators and relations.

Two generators f_+, f_- with two relations $f_+f_- = 1$, $f_-f_+ = 1$ generate a semigroup G_1^{discrete} that is in fact a group, just the cyclic group $\mathbb{Z}$. Indeed, every word reduces to some f_+^k or f_-^k (or 1).

Two generators f_+, f_- with a single relation $f_+f_- = 1$ generate a semigroup G_2^{discrete}. Every word reduces to some $f_-^k f_+^l$. The composition is

$$(f_-^{k_1} f_+^{l_1})(f_-^{k_2} f_+^{l_2}) = f_-^k f_+^l \,, \qquad \begin{aligned} k &= k_1 + \max(0, k_2 - l_1)\,, \\ l &= l_2 + \max(0, l_1 - k_2)\,. \end{aligned} \tag{4a1}$$

The canonical homomorphism $G_2^{\text{discrete}} \to G_1^{\text{discrete}}$ maps f_+ to f_+, f_- to f_-, and $f_-^k f_+^l$ into f_-^{k-l} (if $k > l$), or f_+^{l-k} (if $k < l$), or 1 (if $k = l$). Accordingly, the composition law (4a1) satisfies

$$l - k = (l_1 - k_1) + (l_2 - k_2)\,.$$

There is a more convenient pair of parameters, $a = l - k$, $b = k$; that is,[20]

$$\begin{gathered} f_{a,b} = f_-^b f_+^{a+b} \quad \text{for } a, b \in \mathbb{Z},\, b \ge 0,\, a + b \ge 0\,; \\ f_{a_1,b_1} f_{a_2,b_2} = f_{a,b}\,, \qquad \begin{aligned} a &= a_1 + a_2\,, \\ b &= \max(b_1, b_2 - a_1)\,. \end{aligned} \end{gathered} \tag{4a2}$$

The canonical homomorphism $G_2^{\text{discrete}} \to G_1^{\text{discrete}}$ maps $f_{a,b}$ to f_a, where $f_a \in G_1^{\text{discrete}}$ is f_+^a for $a > 0$, $f_-^{|a|}$ for $a < 0$, and 1 for $a = 0$.

Three generators f_-, f_+, f_* with three relations

$$f_+f_- = 1\,, \quad f_*f_- = 1\,, \quad f_*f_+ = f_*f_* \tag{4a3}$$

generate a semigroup G_3^{discrete}. Every word reduces to some $f_-^k f_+^l f_*^m$. The following homomorphism $G_3^{\text{discrete}} \to G_2^{\text{discrete}}$ will be called canonical: $f_- \mapsto f_-$, $f_+ \mapsto f_+$, $f_* \mapsto f_+$. We have $f_-^k f_+^l f_*^m \mapsto f_-^k f_+^{l+m}$, which suggests such a triple of parameters for G_3^{discrete}: $a = l + m - k$, $b = k$, $c = m$; that is,

$$\begin{gathered} f_{a,b,c} = f_-^b f_+^{a+b-c} f_*^c \quad \text{for } a, b, c \in \mathbb{Z},\ b \ge 0,\ 0 \le c \le a + b\,; \\ f_{a_1,b_1,c_1} f_{a_2,b_2,c_2} = f_{a,b,c}\,, \qquad \begin{aligned} a &= a_1 + a_2\,, \\ b &= \max(b_1, b_2 - a_1)\,, \end{aligned} \qquad c = \begin{cases} a_2 + c_1 & \text{if } c_1 > b_2, \\ c_2 & \text{otherwise.} \end{cases} \end{gathered} \tag{4a4}$$

The canonical homomorphism $G_3^{\text{discrete}} \to G_2^{\text{discrete}}$ is just $f_{a,b,c} \mapsto f_{a,b}$.

Note that G_1^{discrete} is commutative, but G_2^{discrete} and G_3^{discrete} are not.

[20]Parameters a, b of (4a2) and a, b, c of (4a4) are suggested by S. Watanabe.

4b The three discrete semigroups: representation

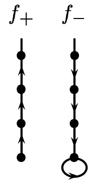
By a representation of a semigroup G on a set S we mean a map $G \times S \ni (g, s) \mapsto g(s) \in S$ such that

$$(g_1 g_2)(s) = g_2\big(g_1(s)\big) \quad \text{and} \quad 1(s) = s$$

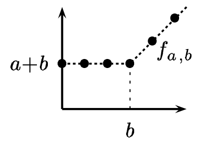
for all $g_1, g_2 \in G$, $s \in S$. The representation is called faithful, if

$$g_1 \ne g_2 \quad \Longrightarrow \quad \exists s \in S \ \ \big(g_1(s) \ne g_2(s)\big) \, .$$

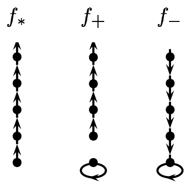
Every G has a faithful representation on itself, $S = G$, namely, the regular representation, $g(g_0) = g_0 g$. Fortunately, G_2^{discrete} and G_3^{discrete} have more economical faithful representations, on the set $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$. Namely, for G_2^{discrete},



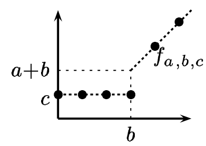
(4b1)
$$f_+(x) = x + 1 \, , \quad f_-(x) = \max(0, x-1) \, , \\ f_{a,b}(x) = a + \max(x, b) \, ,$$



$x \in \mathbb{Z}_+$. For G_3^{discrete},



(4b2)
$$f_*(x) = x + 1 \, , \quad f_-(x) = \max(0, x-1) \, ,$$
$$f_+(x) = \begin{cases} x+1 & \text{for } x > 0, \\ 0 & \text{for } x = 0; \end{cases}$$
$$f_{a,b,c}(x) = \begin{cases} c & \text{for } 0 \le x \le b, \\ x + a & \text{for } x > b. \end{cases}$$



4c Random walks and stochastic flows in discrete semigroups

4c1 Example. The standard random walk on $\mathbb{Z}$ may be described by G_1^{discrete}-valued random variables

(4c2)
$$\xi_{s,t} = \xi_{s,s+1} \xi_{s+1,s+2} \dots \xi_{t-1,t} \quad \text{for } s, t \in \mathbb{Z},\ s \le t \, ;$$
$$\xi_{t,t+1} \text{ are independent random variables } (t \in \mathbb{Z}) \, ;$$
$$\mathbb{P}\left(\xi_{t,t+1} = f_- \right) = \frac{1}{2} = \mathbb{P}\left(\xi_{t,t+1} = f_+ \right) \quad \text{for each } t \in \mathbb{Z} \, .$$

Note that $\xi_{r,s} \xi_{s,t} = \xi_{r,t}$ whenever $r \le s \le t$. Everyone knows that

(4c3)
$$\mathbb{P}\left(\xi_{0,t} = f_a \right) = \frac{1}{2^t} \binom{t}{\frac{t+a}{2}}$$

for $a = -t, -t+2, -t+4, \dots, t$.

In fact, 'the standard random walk' is the random process $t \mapsto \xi_{0,t}$. Taking into account that G_1^{discrete} is a group, $\xi_{s,t}$ may be thought of as an increment, $\xi_{s,t} = \xi_{0,s}^{-1} \xi_{s,t}$.

4c4 Example. Formulas (4c2) work equally well on G_2^{discrete}. Still, $\xi_{r,s}\xi_{s,t} = \xi_{r,t}$. However, G_2^{discrete} is not a group, and $\xi_{s,t}$ is not an increment; moreover, it is not a function of $\xi_{0,s}$ and $\xi_{0,t}$. Indeed, knowing a_1, b_1 and $a_1 + a_2$, $\max(b_1, b_2 - a_1)$ (recall (4a2)) we can find a_2 but not b_2. Thus, the two-parameter family $(\xi_{s,t})_{s\le t}$ of random variables is more than just a random walk. Let us call such a family an *abstract (stochastic) flow.* Why 'abstract'? Since G_2^{discrete} is an abstract semigroup rather than a semigroup of transformations (of some set). So, we have the standard abstract flow in G_2^{discrete}. In order to get a (usual, not abstract) stochastic flow, we have to choose a representation of G_2^{discrete}. Of course, the regular representation could be used, but the representation (4b1) is more useful. Introducing integer-valued random variables $a(s,t), b(s,t)$ by

$$\xi_{s,t} = f_{a(s,t),b(s,t)}$$

we express the stochastic flow as

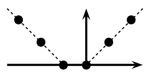
$$\xi_{s,t}(x) = a(s,t) + \max(x, b(s,t)) .$$

Fixing s and x we get a random process called a single-point motion of the flow. Namely, it is a reflecting random walk. Especially, for $s = 0$ and $x = 0$, the process

$$t \mapsto \xi_{0,t}(0) = a(0,t) + b(0,t)$$

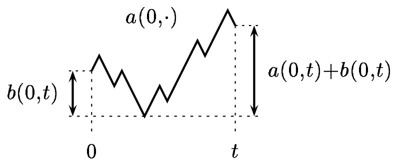
is a reflecting random walk. It is easy to see that two processes

$$t \mapsto \xi_{0,t}(0) = a(0,t) + b(0,t) ,$$
$$t \mapsto \left| a(0,t) + \frac{1}{2} \right| - \frac{1}{2}$$



are identically distributed. Also,

(4c5)
$$b(0,t) = - \min_{s=0,1,\dots,t} a(0,s) ,$$
$$a(0,t) + b(0,t) = \max_{s=0,1,\dots,t} a(s,t) ,$$



and $a(\cdot,\cdot)$ is the standard random walk on $G_1^{\text{discrete}} = \mathbb{Z}$. That is, the canonical homomorphism $G_2^{\text{discrete}} \to G_1^{\text{discrete}}$ transforms the standard flow on G_2^{discrete} into the standard flow (or random walk) on G_1^{discrete}. Using the reflection principle one gets

(4c6)
$$\mathbb{P}\left(\xi_{0,t} = f_{a,b}\right) = \frac{a+2b+1}{2^t} \frac{t!}{\left(\frac{t+a}{2} + b + 1\right)!\left(\frac{t-a}{2} - b\right)!} .$$

Note that a, b occur only in the combination $a + 2b$.

4c7 Example. On G_3^{discrete}, we have no 'standard' random walk or flow; rather, we introduce a one-parameter family of abstract stochastic flows,

(4c8)
$$\xi_{s,t} = \xi_{s,s+1}\xi_{s+1,s+2}\dots\xi_{t-1,t} \quad \text{for } s,t \in \mathbb{Z},\ s \le t ;$$
$$\xi_{t,t+1} \text{ are independent random variables } (t \in \mathbb{Z}) ;$$
$$\mathbb{P}\left(\xi_t = f_-\right) = \frac{1}{2}, \quad \mathbb{P}\left(\xi_t = f_+\right) = \frac{1-p}{2}, \quad \mathbb{P}\left(\xi_t = f_*\right) = \frac{p}{2} ;$$

$p \in (0,1)$ is the parameter. The canonical homomorphism $G_3^{\text{discrete}} \to G_2^{\text{discrete}}$ glues together f_+ and f_*, thus eliminating the parameter p and giving the standard abstract flow on G_2^{discrete}. Defining $a(\cdot,\cdot), b(\cdot,\cdot), c(\cdot,\cdot)$ by

$$\xi_{s,t} = f_{a(s,t),b(s,t),c(s,t)}$$

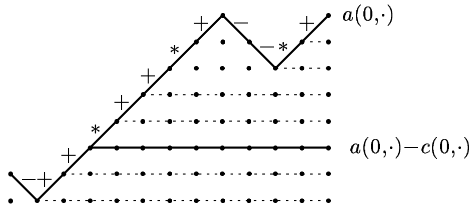
we see that the joint distribution of $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ is the same as before.

Representation (4b2) of G_3^{discrete} turns the abstract flow into a stochastic flow on $\mathbb{Z}_+$. Its single-point motion is a sticky random walk,

$$t \mapsto \xi_{0,t}(0) = c(0,t)\,.$$

In order to find the conditional distribution of $c(\cdot,\cdot)$ given $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ we observe that

(4c9)
$$a(0,t) - c(0,t) = \min\bigl(a(0,t), \min\{x : \xi_{\sigma(x),\sigma(x+1)} = f_*\}\bigr)$$
$$\text{where } \sigma(x) = \max\{s = 0,\dots,t : a(0,s) = x\}\,, \quad -b(0,t) \le x < a(0,t).$$



Therefore the conditional distribution of $c(0,t)$ is basically the truncated geometric distribution. More exactly, it is the (conditional) distribution of

(4c10)
$$\max\bigl(0, a(0,t) + b(0,t) - G + 1\bigr)\,, \qquad G \sim \operatorname{Geom}(p)\,;$$

here G is a random variable, independent of $a(\cdot,\cdot), b(\cdot,\cdot)$, such that $\mathbb{P}\bigl(\, G = g \,\bigr) = p(1-p)^{g-1}$ for $g = 1,2,\dots$ That is the discrete counterpart of a well-known result of J. Warren. So,

(4c11)
$$\mathbb{P}\bigl(\, \xi_{0,t} = f_{a,b,c} \,\bigr) = \frac{a+2b+1}{2^t} \frac{t!}{\Bigl(\frac{t+a}{2} + b + 1\Bigr)!\Bigl(\frac{t-a}{2} - b\Bigr)!} \cdot p(1-p)^{a+b-c}$$

for $c > 0$; for $c = 0$ the factor $p(1-p)^{a+b-c}$ turns into $(1-p)^{a+b}$, rather than $p(1-p)^{a+b}$, because of truncation.

4d Three continuous semigroups

The continuous counterpart of the discrete semigroup $G_1^{\text{discrete}} = \mathbb{Z}$ is the semigroup $G_1 = \mathbb{R} = \{f_a : a \in \mathbb{R}\}$, $f_{a_1} f_{a_2} = f_{a_1+a_2}$.

The continuous counterpart of the discrete semigroup $G_2^{\text{discrete}} = \{f_{a,b} : a,b \in \mathbb{Z}, b \ge 0, a+b \ge 0\}$ is the semigroup

(4d1)
$$G_2 = \{f_{a,b} : a,b \in \mathbb{R},\ b \ge 0,\ a+b \ge 0\}\,,$$
$$f_{a_1,b_1} f_{a_2,b_2} = f_{a,b}\,, \quad \begin{array}{l} a = a_1 + a_2\,, \\ b = \max(b_1, b_2 - a_1) \end{array}$$

(recall (4a2)). The canonical homomorphism $G_2 \to G_1$ maps $f_{a,b}$ to f_a.

The continuous counterpart of the discrete semigroup $G_3^{\text{discrete}} = \{f_{a,b,c} : a,b,c \in \mathbb{Z},\ b \ge 0,\ 0 \le c \le a+b\}$ is the semigroup

$$G_3 = \{f_{a,b,c} : a,b,c \in \mathbb{R},\ b \ge 0,\ 0 \le c \le a+b\}\,,$$

(4d2) $$f_{a_1,b_1,c_1} f_{a_2,b_2,c_2} = f_{a,b,c}\,, \qquad \begin{matrix} a = a_1 + a_2\,, \\ b = \max(b_1, b_2 - a_1)\,, \end{matrix} \qquad c = \begin{cases} a_2 + c_1 & \text{if } c_1 > b_2, \\ c_2 & \text{otherwise} \end{cases}$$

(recall (4a4)). The canonical homomorphism $G_3 \to G_2$ maps $f_{a,b,c}$ to $f_{a,b}$.

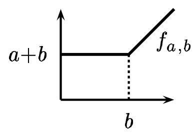
Note that G_1 is commutative but G_2, G_3 are not. Also, G_1 and G_2 are topological semigroups, but G_3 is not (since the composition is discontinuous at $c_1 = b_2$).

There are two one-parameter semigroups in G_2, $\{f_{a,0} : a \in [0,\infty)\}$ and $\{f_{-b,b} : b \in [0,\infty)\}$. They generate G_2 according to the relation $f_{b,0} f_{-b,b} = 1$; namely, $f_{a,b} = f_{-b,b} f_{a+b,0}$.

There are three one-parameter semigroups in G_3, $\{f_{a,0,0} : a \in [0,\infty)\}$, $\{f_{-b,b,0} : b \in [0,\infty)\}$ and $\{f_{c,0,c} : c \in [0,\infty)\}$. They generate G_3 according to relations $f_{b,0,0} f_{-b,b,0} = 1$, $f_{b,0,b} f_{-b,b,0} = 1$, and $f_{c,0,c} f_{a,0,0} = f_{c,0,c} f_{a,0,a}$ for $c > 0$; namely, $f_{a,b,c} = f_{-b,b,0} f_{a+b-c,0,0} f_{c,0,c}$.

Here is a faithful representation of G_2 on $[0,\infty)$ (recall (4b1)):

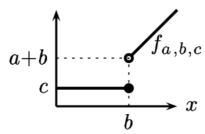
(4d3) $$f_{a,b}(x) = a + \max(x,b)\,,$$



$x \in [0,\infty)$.

Here is a faithful representation of G_3 on $[0,\infty)$ (recall (4b2)):

(4d4) $$f_{a,b,c}(x) = \begin{cases} c & \text{for } 0 \le x \le b, \\ x + a & \text{for } x > b. \end{cases}$$



All functions are increasing, but $f_{a,b}$ are continuous, while $f_{a,b,c}$ are not.

4e Convolution semigroups in these continuous semigroups

4e1 Example. Everyone knows that the binomial distribution (4c3) is asymptotically normal. That is, the distribution of $\sqrt{\varepsilon} a(0, t/\varepsilon)$ converges weakly (for $\varepsilon \to 0$) to the normal distribution $\mu_t^{(1)} = \mathrm{N}(0,t)$. These form a convolution semigroup, $\mu_s^{(1)} * \mu_t^{(1)} = \mu_{s+t}^{(1)}$.

Note however, that $a(s,t)$ and $\xi_{s,t}$ are defined (see (4c2)) only for integers s,t. We may extend them, in one way or another, to real s,t. Or alternatively, we may use coarse instants $t = \big(t[i]\big)_{i=1}^{\infty}$, $t[i] \in \frac{1}{i}\mathbb{Z}$, $t[i] \to t[\infty]$, introduced in 3b. For every coarse instant t, the distribution of $i^{-1/2} a(0, it[i])$ converges weakly (for $i \to \infty$) to $\mu_{t[\infty]}^{(1)} = \mathrm{N}(0, t[\infty])$.

4e2 Example. The two-dimensional distribution (4c6) on G_2^{discrete} has its asymptotics. Namely, the joint distribution of $i^{-1/2} a(0, it[i])$ and $i^{-1/2} b(0, it[i])$ converges weakly (for $i \to \infty$) to the measure $\mu_{t[\infty]}^{(2)}$ with such a density (on the relevant domain $b > 0$, $a + b > 0$; t means $t[\infty]$):

(4e3) $$\frac{\mu_t^{(2)}(da db)}{da db} = \frac{2(a+2b)}{\sqrt{2\pi}\, t^{3/2}} \exp\Big(-\frac{(a+2b)^2}{2t} \Big)\,.$$

Treating $\mu_t^{(2)}$ (for $t \in [0,\infty)$) as a measure on G_2, we get a convolution semigroup: $\mu_s^{(2)} * \mu_t^{(2)} = \mu_{s+t}^{(2)}$. Of course, the convolution is taken according to the composition (4d1).

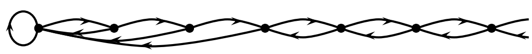
4e4 Example. What about the three-dimensional distribution (4c11) on G_3^{discrete}? It has a parameter p. In order to get a non-degenerate asymptotics, we let p depend on i, namely,

$$p = \frac{1}{\sqrt{i}} \to 0 \, ,$$

then the distribution of $i^{-1/2}G$, where $G \sim \operatorname{Geom}(p)$ (recall (4c10)), converges weakly to the exponential distribution $\operatorname{Exp}(1)$, and the joint distribution of $i^{-1/2}a(0,it[i])$, $i^{-1/2}b(0,it[i])$ and $i^{-1/2}c(0,it[i])$ converges weakly to a measure $\mu_{t[\infty]}^{(3)}$. The measure has an absolutely continuous part and a singular part (at $c=0$), and may be described (somewhat indirectly) as the joint distribution of three random variables a, b and $(a+b-\eta)^+$, where the pair (a,b) is distributed $\mu_t^{(2)}$ (see (4e3)), η is independent of (a,b), and $\eta \sim \operatorname{Exp}(1)$. Treating $\mu_t^{(3)}$ (for $t \in [0,\infty)$) as a measure on G_3, we get a convolution semigroup: $\mu_s^{(3)} * \mu_t^{(3)} = \mu_{s+t}^{(3)}$, the convolution being taken according to the composition (4d2). No need to check the relation 'by hand'; it follows from its discrete counterpart. The latter follows from the construction of 4c (you see, random variables $\xi_{0,1}, \xi_{1,2}, \dots, \xi_{s+t-1,s+t}$ are independent). It may seem that the limiting procedure does not work, since G_3 is not a topological semigroup; the composition (4d2) is discontinuous at $c_1 = b_2$. However, that is not an obstacle, since the equality $c_1 = b_2$ is of zero probability, as far as triples (a_1,b_1,c_1) and (a_2,b_2,c_2) are independent and distributed $\mu_s^{(3)}$, $\mu_t^{(3)}$ respectively $(s,t>0)$. The atom of c_1 at 0 does not matter, since b_2 is nonatomic. The composition is continuous almost everywhere!

4f Getting dyadic

Our flows in G_1^{discrete} and G_2^{discrete} are dyadic (two equiprobable possibilities on each step), which cannot be said about G_3^{discrete}; here, on each step, we have 3 possibilities f_-, f_+, f_* of probabilities $1/2, (1-p)/2, p/2$. Can a dyadic model produce the same asymptotic behavior? Yes, it can, at the expense of using $i \in \{1,4,16,64,\dots\}$ only (recall 3b6); and, of course, the dyadic model is more complicated.[21] Instead of the trap at 0, we design a trap near 0 as follows:



$$g_+ = f_* = f_{1,0,1} \, ; \quad g_- = f_-^m f_+^{m-1} = f_{-1,m,0} \, ;$$
$$\mathbb{P}\big(\, \xi_{t,t+1} = g_- \,\big) = \frac{1}{2} = \mathbb{P}\big(\, \xi_{t,t+1} = g_+ \,\big) \, .$$

The old (small) parameter p disappears, and a new (large) parameter m appears. We'll see that the two models are asymptotically equivalent, when $p = 2^{-m}$.

As before, we may denote

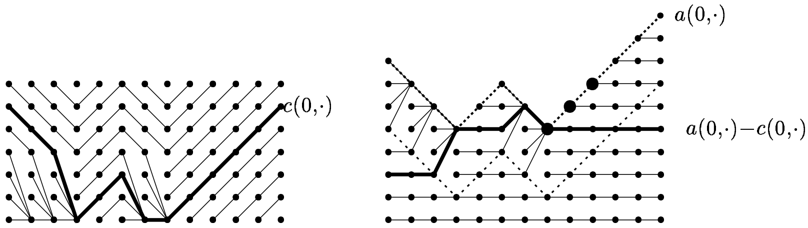
$$\xi_{s,t} = f_{a(s,t),b(s,t),c(s,t)} \, ;$$

[21] Maybe, a still more complicated construction can use all i; I do not know.

note however that only $a(s,t)$ is the same as before; $b(s,t)$, $c(s,t)$ and $\xi_{s,t}$ are modified. Formula (4c5) for $b(0,t)$ fails, but still,

(4f1) $$b(0,t) = - \min_{s=0,1,\dots,t} a(0,s) + O(m)\,,$$

which is asymptotically the same. Formula (4c9) for $c(0,t)$ also fails; instead,



(4f2) $$a(0,t) - c(0,t) = \min\{x : \sigma(x+m-1) - \sigma(x) = m-1\}\,,$$

if such x exists in the set $\mathbb{Z} \cap [\min_{[0,t]} a(0,\cdot), a(0,t)-m+1]$; otherwise, $c(0,t) = O(m)$. (Here σ is the same as in (4c9).)

The conditional distribution of $c(0,t)$, given the path $a(0,\cdot)$, is not at all geometric (unlike (4c10)), since now $c(0,t)$ is uniquely determined by $a(0,\cdot)$. However, according to (4f2), $c(0,t)$ is determined by small increments of the process $\sigma(\cdot)$. On the other hand, the large-scale structure of the path $a(0,\cdot)$ is correlated mostly with large increments of $\sigma(\cdot)$; small increments are numerous, but contribute a little to the sum. Using this argument, one can show that $c(0,t)$ is asymptotically independent of $a(0,t)$ (and $b(0,t)$, due to (4f1)).

The unconditional distribution of $c(0,t)$ can be found from (4f2), taking into account that increments $\sigma(x+1) - \sigma(x)$ are independent, and each increment is equal to 1 with probability 1/2. We have Bernoulli trials, and we wait for the first block of $m-1$ 'successes'. For large m, the waiting time is approximately exponential, with the mean 2^m.[22] Thus, $2^{-m}\big(a(0,t) - c(0,t) - \min_{[0,t]} a(0,\cdot)\big)$ is asymptotically Exp(1), truncated (at $c=0$) as in 4e.

Taking the limit $i = 2^{2m} \to \infty$ we get for $i^{-1/2}a(0,it[i])$, $i^{-1/2}b(0,it[i])$, $i^{-1/2}c(0,it[i])$ the limiting distribution $\mu^{(3)}_{t[\infty]}$, the same as in 4e.

4g Scaling limit

For any coarse instants s,t, the distribution $\mu^{(n)}_{s,t}[i]$ of $i^{-1/2}\xi^{(n)}_{is[i],it[i]}$ converges weakly (for $i \to \infty$) to the measure $\mu^{(n)}_{s,t}[\infty] = \mu^{(n)}_{t[\infty]-s[\infty]}$ on G_n, for our three models, $n = 1,2,3$. Of course, multiplication of ξ by $i^{-1/2}$ is understood as multiplication of $a(\cdot,\cdot)$, $b(\cdot,\cdot)$, $c(\cdot,\cdot)$ by $i^{-1/2}$, which is a homomorphic embedding of G_n^{discrete} into G_n.

Let r,s,t be coarse instants, $r \le s \le t$. Due to independence, the joint distribution $\mu^{(n)}_{r,s}[i] \otimes \mu^{(n)}_{s,t}[i]$ of random variables $i^{-1/2}\xi^{(n)}_{ir[i],is[i]}$ and $i^{-1/2}\xi^{(n)}_{is[i],it[i]}$ converges weakly to $\mu^{(n)}_{r,s}[\infty] \otimes \mu^{(n)}_{s,t}[\infty]$. However, we need the joint distribution of three random variables,

$$i^{-1/2}\xi^{(n)}_{ir[i],is[i]}\,, \quad i^{-1/2}\xi^{(n)}_{is[i],it[i]}\,, \quad i^{-1/2}\xi^{(n)}_{ir[i],it[i]}\,,$$

[22] Such a block appears, in the mean, after 2^{m-1} shorter blocks, of mean length ≈ 2 each.

the third being the product of the first and the second in the semigroup G_n. For $n = 1, 2$ weak convergence for the triple follows immediately from weak convergence for the pair, since the composition is continuous. For $n = 3$, discontinuity of the composition in G_3 does not invalidate the argument, since the composition is continuous almost everywhere w.r.t. the relevant measure (recall 4e).

Similarly, for every k and every coarse instants $t_1 \le \dots \le t_k$, the joint distribution of $k(k-1)/2$ random variables $i^{-1/2}\xi^{(n)}_{it_l[i],it_m[i]}$, $1 \le l < m \le k$, converges weakly (for $i \to \infty$). We choose a sequence $(t_k)_{k=1}^\infty$ of coarse instants such that the sequence of numbers $(t_k[\infty])_{k=1}^\infty$ is dense in $\mathbb{R}$, and use Lemma 2c10, getting a coarse probability space.

The Hölder condition, the same as in 2a3, holds for all three models. I mean Hölder continuity of $a(\cdot,\cdot)$, $b(\cdot,\cdot)$, $c(\cdot,\cdot)$. Indeed, $a(\cdot,\cdot)$ is the same as in 2a3; $b(\cdot,\cdot)$ is related to $a(\cdot,\cdot)$ via (4c5) or (4f1); and $c(\cdot,\cdot)$ satisfies (on any interval)

$$\max_{|s-t|\le x} |c(0,s) - c(0,t)| \le \max_{|s-t|\le x} |a(0,s) - a(0,t)| \,;$$

though, for the model of 4f, $O(m)$ must be added.

Thus, a joint σ-compactification is constructed for all three models (the third model — in two versions, (4c7) and 4f).

4h Noises

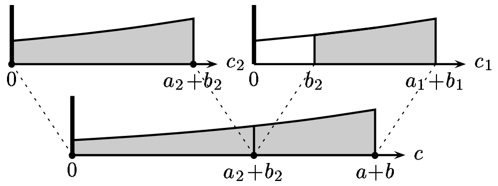
4h1 Example. The standard flow in G_1^{discrete}, rescaled by $i^{-1/2}$, gives us a coarse probability space, identical to that of 3b2. It is a dyadic coarse factorization. Its refinement is the Brownian continuous factorization (see 3d2). Equipped with the natural time shift, it is a noise (see 3e4).

4h2 Example. The standard flow in G_2^{discrete}, rescaled by $i^{-1/2}$, gives us another coarse probability space. It is also a dyadic coarse factorization (the proof is similar to the previous case). Its 'two-dimensional nature' is a delusion; the dyadic coarse factorization is identical to that of 4h1. The second dimension $b(\cdot,\cdot)$ reduces to the first dimension, $a(\cdot,\cdot)$, by (4c5).

4h3 Example. The flow in G_3, introduced in 4c7, rescaled by $i^{-1/2}$ with $p = i^{-1/2}$ (recall 4e4), gives us a coarse probability space. It is not a dyadic coarse factorization, since it is not dyadic. However, it satisfies a natural generalization of Definition 3b1 to non-dyadic case (the proof is as before). Its refinement is a continuous factorization, and (with natural time shift), a noise.

Once again, the second dimension, $b(\cdot,\cdot)$, reduces to the first dimension, $a(\cdot,\cdot)$. Indeed, the joint distribution of $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ is the same as in 4h2. What about the third dimension, $c(\cdot,\cdot)$?

The conditional distribution of $c(s,t)$, given $a(s,t)$ and $b(s,t)$, is basically truncated exponential. Namely, it is the distribution of $\big(a(s,t) + b(s,t) - \eta\big)^+$ where $\eta \sim \operatorname{Exp}(1)$; see 4e4. Moreover, for any $r < s < t$, the conditional distribution of $c(r,t)$ given $a(r,s), b(r,s)$ and $a(s,t), b(s,t)$, is still the distribution of $\big(a(r,t) + b(r,t) - \eta\big)^+$. In other words, $c(r,t)$ is conditionally independent of $a(r,s), b(r,s), a(s,t), b(s,t)$, given $a(r,t), b(r,t)$. That is a property of the composition (4a4); if $c_1 \sim \big(a_1 + b_1 - \eta_1\big)^+$ and $c_2 \sim \big(a_2 + b_2 - \eta_2\big)^+$ then $c \sim \big(a + b - \eta\big)^+$.



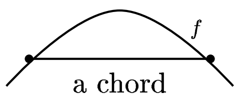
It follows by induction that the conditional distribution of $c(t_1, t_n)$, given all $a(t_i, t_j)$ and $b(t_i, t_j)$, is given by the same formula $\big(a(t_1, t_n) + b(t_1, t_n) - \eta\big)^+$, $\eta \sim \mathrm{Exp}(1)$, for every n and $t_1 < \cdots < t_n$. Therefore, the same holds for the conditional distribution of $c(s, t)$ given all $a(u, v)$ and $b(u, v)$ for u, v such that $s \le u \le v \le t$ (a result of J. Warren). We see that $c(\cdot, \cdot)$ is not a function of $a(\cdot, \cdot)$ (and $b(\cdot, \cdot)$).

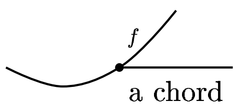
4h4 Example. Another flow in G_3^{discrete}, introduced in 4f, rescaled by $i^{-1/2}$ with $i = 2^{2m}$, gives us a dyadic coarse factorization. Its refinement is the same continuous factorization (and noise) as in 4h3.

4i The Poisson snake

Formula (4c9) suggests a description of the sticky flow in G_3^{discrete} by a combination of a simple random walk $a(\cdot, \cdot)$ and a random subset of the set of its 'chords'. A chord may be defined as an interval $[s, t]$, $s, t \in \mathbb{Z}$, $s < t$, such that $a(s, t) = 0$ and $a(s, u) > 0$ for all $u \in (s, t) \cap \mathbb{Z}$. Or equivalently, a chord is a horizontal straight segment on the plane that connects points $\big(s, a(0, s)\big)$ and $\big(t, a(0, t)\big)$ and goes below the graph of $a(0, \cdot)$. The random subset of chords is very simple: every chord belongs to the subset with probability p, independently of others. Note that $p = i^{-1/2}$ is equal to the vertical pitch (after rescaling $a(\cdot, \cdot)$ by $i^{-1/2}$). The scaling limit suggests itself: a Poisson random subset of the set of all chords of the Brownian sample path.

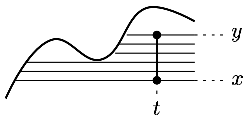
4i1 Definition. A *finite chord* of a continuous function $f : \mathbb{R} \to \mathbb{R}$ is a set of the form $[s, t] \times \{x\} \subset \mathbb{R}^2$ where $s < t$, $x = f(s)$ and $t = \inf\{u \in (s, \infty) : f(u) > x\}$. An *infinite chord* of f is a set of the form $[s, \infty) \times \{x\} \subset \mathbb{R}^2$ where $x = f(s)$ and $f(t) > x$ for all $t \in (s, \infty)$. A *chord* of f is either a finite chord of f, or an infinite chord of f.





If f decreases, it has no chords. Otherwise it has continuum of chords. The set of chords is, naturally, a standard Borel space, due to the one-one correspondence between a chord and its initial point $(s, x) \in \mathbb{R}^2$.

4i2 Lemma. For every continuous function $f : \mathbb{R} \to \mathbb{R}$ there exists one and only one σ-finite positive Borel measure on the space of all chords of f, such that the set of chords that intersect a vertical segment $\{t\} \times [x, y]$ is of measure $y - x$, whenever t, x, y are such that $\inf_{s \in (-\infty, t)} f(s) \le x < y \le f(t)$.



The proof is left to the reader. Hint: for every $\varepsilon > 0$, the set of chords longer than ε is elementary; on this set, the measure is locally finite.

The map $[s,t] \times \{x\} \mapsto s$ (also $[s,\infty) \times \{x\} \mapsto s$, of course) sends the measure on the set of chords into a measure on $\mathbb{R}$. If f is of locally finite variation, then the measure on $\mathbb{R}$ is just $(df)^+$, the positive part of the Lebesgue-Stieltjes measure. However, we need the opposite case: f is of infinite variation on every interval, and the measure is also infinite on every interval. Nevertheless, it is σ-finite. We'll denote it $(df)^+$ anyway.

The measure $(df)^+$ is concentrated on the set of points of 'local minimum from the right'. If f is a Brownian sample path then the set is of Lebesgue measure 0.

So, the set of all chords is a measure space, it carries a natural σ-finite (sometimes, finite) measure. The latter is the intensity measure of a unique Poisson random measure.[23] This way, (the distribution of) a random set of chords is well-defined.

Or equivalently, we may consider a Poisson random subset of $\mathbb{R}$, whose intensity measure is $(df)^+$.

However, it is not so easy, to substitute a Brownian sample path $B(\cdot)$ for $f(\cdot)$. In order to get a (Poisson) random variable, we may ask, how many random points belong to a given Borel set $A \subset \mathbb{R}$ such that $(dB)^+(A) < \infty$. Note that for any *interval* A, $(dB)^+(A) = \infty$ a.s. We cannot choose an appropriate A without knowing the path $B(\cdot)$. Countable dense subsets of $\mathbb{R}$ do not carry a natural (non-pathological) Borel structure.

In this aspect, chords are better than points. They are parametrized by three (or two) numbers, thus, they carry a natural Borel structure, irrespective of $B(\cdot)$. The random countable set of chords is not dense; rather, it accumulates toward short chords.

A point (t,x) belongs to a random chord of $B(\cdot)$ if and only if

$$x \in \sigma_t^{-1}(\Pi)\,, \quad \text{that is,} \quad \sigma_t(x) \in \Pi\,,$$
$$\text{where } \sigma_t(x) = \inf\{s \in (-\infty,t] : B(s) > x\} \text{ for } x \in (-\infty, B(t))$$

(recall (4c9)), and Π is the Poisson random subset of $\mathbb{R}$, whose intensity measure is $(dB)^+$. Do not confuse the inverse image $\sigma_t^{-1}(\Pi)$ with the image $B(\Pi)$. True, $B(\sigma_t(x)) = x$, but $\sigma_t(B(s)) \ne s$. Sets Π and $B(\Pi)$ are dense, but the set $\sigma_t^{-1}(\Pi)$ is locally finite. Moreover, $\sigma_t^{-1}(\Pi)$ is a Poisson random subset of $(-\infty, B(t)]$, its intensity being just 1.

The random countable dense set Π itself is bad; we have no measurable functions of it. However, the pair $\big(B(\cdot),\Pi\big)$ of the Brownian path and the set is good; we have measurable functions of the pair; in particular, measurable functions of the locally finite set $\sigma_t^{-1}(\Pi)$. Especially,

$$a(0,t) - c(0,t) = \min\big(a(0,t), \min\{x : \sigma_t(x) \in \Pi \cap (0,\infty)\}\big)\,.$$

4i3 Lemma. The σ-field $\mathcal{F}_{s,t}$ of the sticky noise is generated by Brownian increments $B(u) - B(s)$ for $u \in (s,t)$ and random sets $\sigma_u^{-1}\big(\Pi \cap (s,t)\big)$ for $u \in (s,t)$ (treated as random variables whose values are finite subsets of $\mathbb{R}$).

The proof is left to the reader.

[23] See for instance [5, XII.1.18].