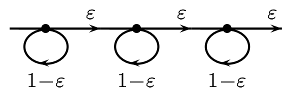
7 Some non-smooth stochastic flows: stickiness

7a Poisson process revisited

Here is a very simple discrete model that leads to the Poisson process in the scaling limit ($\varepsilon \to 0$):

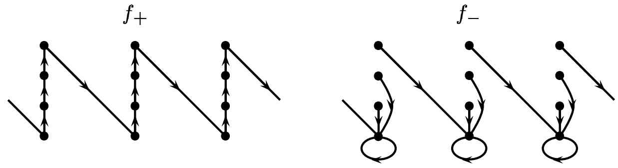


At every instant $k\varepsilon$, the process either jumps by 1 (with probability ε) or stays (with probability $1-\varepsilon$). Jumps at different instants are independent, of course. The process may be treated as a stochastic flow,

$$f_0, f_1 : \mathbb{Z} \to \mathbb{Z}\,, \quad f_0(x) = x, \quad f_1(x) = x+1\,,$$

but f_0, f_1 are not at all equiprobable; rather, f_1 is of probability ε, and f_0 is of probability $1-\varepsilon$.

Here is another model; it leads to the same Poisson process in the scaling limit, but uses equiprobable maps f_+, f_-:



$$\begin{gathered} f_+, f_- : \mathbb{Z} \times \{0, \varepsilon, \dots, m\varepsilon\} \to \mathbb{Z} \times \{0, \varepsilon, \dots, m\varepsilon\}\,; \\ f_+(k, l\varepsilon) = (k, l\varepsilon + \varepsilon) \quad \text{if } l < m\,, \\ f_+(k, m\varepsilon) = (k+1, 0)\,; \\ f_-(k, l\varepsilon) = (k, 0) \quad \text{if } l < m\,, \\ f_-(k, m\varepsilon) = (k+1, 0)\,. \end{gathered}$$

The time pitch is ε (as before), and (similarly to 2b), the scaling limit is taken for

$$\varepsilon = 2^{-(m+1)} \to 0\,, \quad m \to \infty\,.$$

The single-point motion may be thought of as a sequence of *excursions* above the 'ground level' $\mathbb{Z} \times \{0\}$. Each excursion consists of several $(0, 1, \dots, m)$ steps upwards and one step downwards — return to the ground level, either to the starting point (after $0, \dots, m-1$ steps upwards), or to the next column (after m steps upwards). The latter case may be called a success. Excursions are independent[1] and identically distributed. Each excursion is successful with probability 2^{-m}. The first success appears after a random number, $N-1$, of unsuccessful excursions; N is distributed geometrically, $\mathrm{G}(2^{-m})$,

N	1	2	3	...
probability	2^{-m}	$(1-2^{-m})2^{-m}$	$(1-2^{-m})^2 2^{-m}$	...

[1]The needed 'no memory' property of random signs seems evident, but is not so easy to formulate and prove.

therefore, $\mathbb{E}\,N = 2^m$. The duration of a single excursion has a truncated geometric distribution (rescaled by the time pitch ε),

duration	ε	2ε	$\dots$	$(m-1)\varepsilon$	$m\varepsilon$	$(m+1)\varepsilon$
probability	2^{-1}	2^{-2}	$\dots$	$2^{-(m-1)}$	2^{-m}	2^{-m}

Note that duration $(m+1)\varepsilon$ corresponds to success, others — to failure.

The total duration (until the first success) consists of durations of $N-1$ unsuccessful excursions, plus $(m+1)\varepsilon$ (the duration of the N-th, successful excursion). Conditionally, given $N = n$, the total duration is $(m+1)\varepsilon$ plus the sum of $n-1$ independent random variables, each having expectation[2]

$$\frac{\varepsilon \cdot 2^{-1} + 2\varepsilon \cdot 2^{-2} + \dots + m\varepsilon \cdot 2^{-m}}{2^{-1} + 2^{-2} + \dots + 2^{-m}} = 2\varepsilon\big(1 + o(1)\big)$$

and variance $\le 2\varepsilon^2$. Thus, the total duration T has expectation $(n-1)\cdot 2\varepsilon\big(1+o(1)\big)+(m+1)\varepsilon$ and variance $\le 2(n-1)\varepsilon^2$. These are conditional values, given $N = n$. Taking into account that $\varepsilon = 2^{-(m+1)}$ we see that the total duration T satisfies

$$\begin{aligned} \mathbb{E}\,\big(T \,\big|\, N \big) &= \frac{N-1}{2^m}\big(1+o(1)\big) + o(1) = \frac{N}{2^m} + o(1)\,, \\ \operatorname{Var}\big(T \,\big|\, N \big) &\le \frac{N-1}{2^{2m+1}} = o\Big(\frac{N}{2^m}\Big)\,. \end{aligned}$$

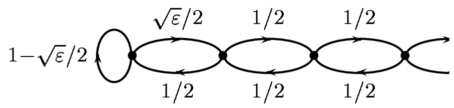
Therefore

$$\begin{aligned} \|T - 2^{-m}N\|^2 &= \mathbb{E}\,(T - 2^{-m}N)^2 = \mathbb{E}\Big(\mathbb{E}\big(\,(T-2^{-m}N)^2 \,\big|\, N\,\big)\Big) = \\ &= \mathbb{E}\Big(\big(\mathbb{E}\,\big(\,T - 2^{-m}N \,\big|\, N\,\big)\,\big)^2 + \operatorname{Var}\big(\,T - 2^{-m}N \,\big|\, N\,\big)\Big) = \\ &= \mathbb{E}\Big(\big(\underbrace{\mathbb{E}\,\big(\,T \,\big|\, N\,\big) - 2^{-m}N}_{o(1)}\big)^2 + \underbrace{\operatorname{Var}\big(\,T \,\big|\, N\,\big)}_{o(2^{-m}N)}\Big) = o\big(\mathbb{E}\,(2^{-m}N)\big) = o(1)\,, \end{aligned}$$

that is, $\|T - 2^{-m}N\| \to 0$ when $\varepsilon = 2^{-(m+1)} \to 0$. In the scaling limit, $2^{-m}N$ becomes an exponential $\operatorname{Exp}(1)$ random variable, and the same for T. This is why we get the Poisson process.

7b Stickiness in discrete time

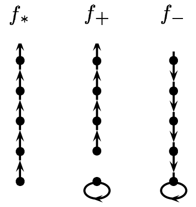
Here is an interesting hybrid of models considered in Sect. 6 and in 7a:



Being at a point $k\sqrt{\varepsilon} > 0$, the process jumps to $(k \pm 1)\sqrt{\varepsilon}$ with probabilities 50%, 50%. However, being at 0, it jumps to $\sqrt{\varepsilon}$ with probability $\frac{1}{2}\sqrt{\varepsilon}$, or stays at 0 (with probability

[2]Recall that geometric distribution $\mathrm{G}(p)$ has expectation $1/p$ and variance $(1-p)/p^2$.

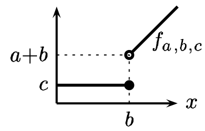
$1 - \frac{1}{2}\sqrt{\varepsilon})$. It is the one-point motion of a stochastic flow built out of three maps:



$$\xi_{k\varepsilon} = \begin{cases} f_* & \text{with probability } \frac{\sqrt{\varepsilon}}{2}, \\ f_+ & \text{with probability } \frac{1-\sqrt{\varepsilon}}{2}, \\ f_- & \text{with probability } \frac{1}{2}. \end{cases}$$

Note that $f_- \circ f_+ = f_- \circ f_* = \mathrm{id}$. Therefore, in every composition we may concentrate all f_- at the right end; for example, $f_+ \circ f_- \circ f_+ \circ f_* \circ f_+ \circ f_- = f_+ \circ f_* \circ f_+ \circ f_-$. Note also that $f_+ \circ f_* = f_* \circ f_*$. Therefore, every composition may be reduced to $f_*^k \circ f_+^l \circ f_-^m$; for example, $f_+ \circ f_- \circ f_+ \circ f_* \circ f_+ \circ f_- = f_+ \circ f_* \circ f_+ \circ f_- = f_* \circ f_* \circ f_+ \circ f_-$. The maps f_-, f_+, f_*, as well as their compositions $\xi_{s,t}$ (defined like (6a2)) belong to a three-parameter family of maps $f_{a,b,c} : [0,\infty) \to [0,\infty)$,

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



for $b \in [0,\infty)$, $0 \le c \le a + b$, as we'll see now.

7b1 Exercise.

(a) $$f_- = f_{-\sqrt{\varepsilon},\sqrt{\varepsilon},0}; \quad f_+ = f_{\sqrt{\varepsilon},0,0}; \quad f_* = f_{\sqrt{\varepsilon},0,\sqrt{\varepsilon}};$$

(b) $$f_-^n = f_{-n\sqrt{\varepsilon},n\sqrt{\varepsilon},0}; \quad f_+^n = f_{n\sqrt{\varepsilon},0,0}; \quad f_*^n = f_{n\sqrt{\varepsilon},0,n\sqrt{\varepsilon}};$$

(c) $$f_*^k \circ f_+^l \circ f_-^m = f_{(k+l-m)\sqrt{\varepsilon},m\sqrt{\varepsilon},k\sqrt{\varepsilon}};$$

(d) $$f_{a,b,c} = f_{c,0,c} \circ f_{a+b-c,0,0} \circ f_{-b,b,0};$$

(e) $$\begin{cases} f_{a_1,0,0} \circ f_{a_2,0,0} = f_{a_1+a_2,0,0}; \\ f_{-b_1,b_1,0} \circ f_{-b_2,b_2,0} = f_{-b_1-b_2,b_1+b_2,0}; \\ f_{c_1,0,c_1} \circ f_{c_2,0,c_2} = f_{c_1+c_2,0,c_1+c_2}; \end{cases}$$

(f) $$\begin{cases} f_{-a,a,0} \circ f_{a,0,0} = \mathrm{id}; \\ f_{-a,a,0} \circ f_{a,0,a} = \mathrm{id}; \\ \quad \text{moreover, } f_{-a,a,0} \circ f_{a,0,c} = \mathrm{id}; \\ f_{a,0,0} \circ f_{c,0,c} = f_{a,0,a} \circ f_{c,0,c} = f_{a+c,0,a+c} \text{ for } c > 0; \\ \quad \text{moreover, } f_{a_2,b_2,c_2} \circ f_{a_1,b_1,c_1} = f_{a_2,0,a_2} \circ f_{a_1,b_1,c_1} = f_{a_1+a_2,b_1,c_1+a_2} \text{ for } c_1 > b_2; \end{cases}$$

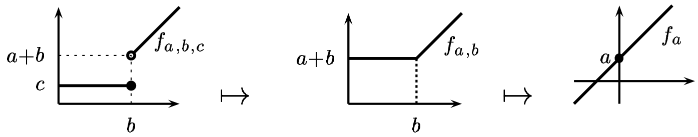
(g) $$f_{-b,b,0} \circ f_{a,0,0} = \begin{cases} f_{a-b,0,0} & \text{if } a \ge b, \\ f_{a-b,b-a,0} & \text{if } a \le b; \end{cases} \quad f_{-b,b,0} \circ f_{a,0,a} = \begin{cases} f_{a-b,0,a-b} & \text{if } a \ge b, \\ f_{a-b,b-a,0} & \text{if } a \le b; \end{cases}$$

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

Prove it.

The composition law 7b1(h) shows that c does not influence a, b; also, b does not influence a. The composition law for (a, b) in 7b1(h) is the same as in 6b3(h). We have three semigroups G_1, G_2, G_3 (one-dimensional, two-dimensional and three-dimensional, respectively), considered in 6a, 6b and 7a (here) respectively, and canonical homomorphisms $G_3 \to G_2 \to G_1$:

(7b2)

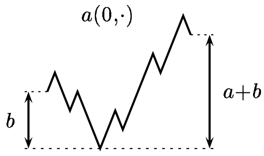


We have random maps $\xi_{s,t} = f_{a(s,t),b(s,t),c(s,t)}$. Note that $f_+ = f_{\sqrt{\varepsilon},0,0}$ and $f_* = f_{\sqrt{\varepsilon},0,\sqrt{\varepsilon}}$ differ only in the third parameter (c); their first two parameters (a, b) are the same as (a, b) of f_+ in 6b. Also, f_- here and in 6b conform in (a, b) parameters. Therefore, the joint distribution of $a(s, t)$ and $b(s, t)$ here is the same as in 6b, 6c. Moreover, corresponding two-dimensional random processes $t \mapsto \big(a(0, t), b(0, t)\big)$ are identically distributed.

Similarly (but simpler), $a(s, t)$ here (and in 6b) conforms to $a(s, t)$ of 6a, just the random walk.

Relation 6b4 still holds:

(7b3)



you see, a path of the process $a(0, \cdot)$ determines uniquely the corresponding path of the process $b(0, \cdot)$.

The third process $c(0, \cdot)$ is new, and is not uniquely determined by $a(0, \cdot)$ and $b(0, \cdot)$, since the distinction between f_+ and f_* comes into play.

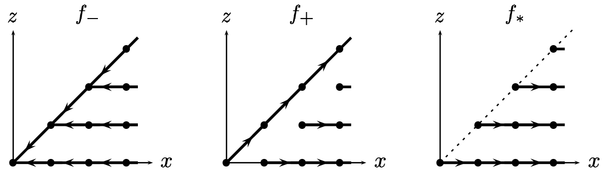
The process $k\varepsilon \mapsto \xi_{0,k\varepsilon}$ is a Markov chain in G_3; being at $f_{a,b,c}$, it jumps to

$$\begin{aligned} f_{-\sqrt{\varepsilon},\sqrt{\varepsilon},0} \circ f_{a,b,c} \quad &\text{with probability } \frac{1}{2}, \\ f_{\sqrt{\varepsilon},0,0} \circ f_{a,b,c} \quad &\text{with probability } \frac{1-\sqrt{\varepsilon}}{2}, \\ f_{\sqrt{\varepsilon},0,\sqrt{\varepsilon}} \circ f_{a,b,c} \quad &\text{with probability } \frac{\sqrt{\varepsilon}}{2}. \end{aligned}$$

The Markov chain is time-homogeneous. The three-dimensional process $k\varepsilon \mapsto \big(a(0, k\varepsilon), b(0, k\varepsilon), c(0, k\varepsilon)\big)$ also is a time-homogeneous Markov chain. Conditionally, given paths $a(0, \cdot)$ and $b(0, \cdot)$, the process $c(0, \cdot)$ is a time-inhomogeneous Markov chain. If $a(0, k\varepsilon) = a(0, k\varepsilon - \varepsilon) - \sqrt{\varepsilon}$, it means that $\xi_{k\varepsilon} = f_-$, therefore $c(0, k\varepsilon) = \max\big(0, c(0, k\varepsilon - \varepsilon) - \sqrt{\varepsilon}\big)$. If $a(0, k\varepsilon) = a(0, k\varepsilon - \varepsilon) + \sqrt{\varepsilon}$, it means that $\xi_{k\varepsilon}$ is either f_+ (with conditional probability $1 - \sqrt{\varepsilon}$) or f_* (with conditional probability $\sqrt{\varepsilon}$). These two cases give us

$$c(0, k\varepsilon) = \begin{cases} 0 & \text{if } c(0, k\varepsilon - \varepsilon) = 0 \text{ and } \xi_{k\varepsilon} = f_+, \\ c(0, k\varepsilon - \varepsilon) + \sqrt{\varepsilon} & \text{otherwise.} \end{cases}$$

Similarly to 6c, we introduce the process $X(t) = a(0,t) + b(0,t)$. However, instead of $Y(t) = b(0,t)$, we introduce now $Z(t) = X(t) - c(0,t)$. Note that $0 \le Z(t) \le X(t)$. The point $\big(X(k\varepsilon), Z(k\varepsilon)\big)$ is a function of $\xi(k\varepsilon)$ and $\big(X(k\varepsilon-\varepsilon), Z(k\varepsilon-\varepsilon)\big)$, and we can calculate the function:



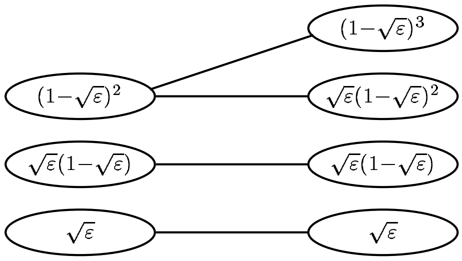
7b4 Exercise. Check the figures shown above.

7b5 Exercise.

$$\mathbb{P}\left(Z(n\varepsilon) = l\sqrt{\varepsilon} \,\middle|\, X(\varepsilon), \dots, X(n\varepsilon) \right) = \begin{cases} \sqrt{\varepsilon}(1-\sqrt{\varepsilon})^l & \text{for } 0 \le l\sqrt{\varepsilon} < X(n\varepsilon), \\ (1-\sqrt{\varepsilon})^l & \text{for } l\sqrt{\varepsilon} = X(n\varepsilon). \end{cases}$$

Prove it.

Hint: use 7b4;



The conditional probability depends only on $X(k\varepsilon)$, thus,

$$\text{(7b6)} \qquad \mathbb{P}\left(Z(n\varepsilon) = l\sqrt{\varepsilon} \,\middle|\, X(n\varepsilon) \right) = \mathbb{P}\left(Z(n\varepsilon) = l\sqrt{\varepsilon} \,\middle|\, X(\varepsilon), \dots, X(n\varepsilon) \right) .$$

Using the distribution of $X(n\varepsilon)$ found in 6c12, we get the joint distribution of $X(n\varepsilon)$ and $Z(n\varepsilon)$:

$$\mathbb{P}\left(X(n\varepsilon) = Z(n\varepsilon) = k\sqrt{\varepsilon} \right) = 2^{-n} \binom{n}{m} (1-\sqrt{\varepsilon})^k ;$$

$$\mathbb{P}\left(X(n\varepsilon) = k\sqrt{\varepsilon}, Z(n\varepsilon) = l\sqrt{\varepsilon} \right) = 2^{-n} \binom{n}{m} \sqrt{\varepsilon}(1-\sqrt{\varepsilon})^l \quad \text{for } 0 \le l < k ;$$

here m is either $\frac{n+k}{2}$ (for $n+k$ even), or $\frac{n+k+1}{2}$ (for $n+k$ odd). Moreover, (7b6) means that the random vector $\big(X(\varepsilon), \dots, X(n\varepsilon-\varepsilon)\big)$ and the random variable $Z(n\varepsilon)$ are conditionally independent, given $X(n\varepsilon)$ (think, why). On the other hand, $X(\varepsilon), \dots, X(n\varepsilon)$ determine $a(0,\varepsilon), \dots, a(0,n\varepsilon)$ uniquely; indeed, a increases (by $\sqrt{\varepsilon}$) if and only if X increases (by $\sqrt{\varepsilon}$). Also, $a(0,\varepsilon), \dots, a(0,n\varepsilon)$ determine $b(0,\varepsilon), \dots, b(0,n\varepsilon)$ uniquely (recall (7b3)). Thus, $b(0,n\varepsilon)$ is a function of $X(\varepsilon), \dots, X(n\varepsilon)$. Therefore, $b(0,n\varepsilon)$ and $Z(n\varepsilon)$ are conditionally independent, given $X(n\varepsilon)$. We get

$$\begin{aligned} &\mathbb{P}\left(b(0,n\varepsilon) = i\sqrt{\varepsilon},\, X(n\varepsilon) = k\sqrt{\varepsilon},\, Z(n\varepsilon) = l\sqrt{\varepsilon} \right) = \\ &= \mathbb{P}\left(Z(n\varepsilon) = l\sqrt{\varepsilon} \,\middle|\, b(0,n\varepsilon) = i\sqrt{\varepsilon},\, X(n\varepsilon) = k\sqrt{\varepsilon} \right) \cdot \mathbb{P}\left(b(0,n\varepsilon) = i\sqrt{\varepsilon},\, X(n\varepsilon) = k\sqrt{\varepsilon} \right) = \\ &\qquad = \underbrace{\mathbb{P}\left(Z(n\varepsilon) = l\sqrt{\varepsilon} \,\middle|\, X(n\varepsilon) = k\sqrt{\varepsilon} \right)}_{\text{see 7b5}} \cdot \underbrace{\mathbb{P}\left(b(0,n\varepsilon) = i\sqrt{\varepsilon},\, X(n\varepsilon) = k\sqrt{\varepsilon} \right)}_{\text{see the formula after 6c7}} ; \end{aligned}$$

the (3-dimensional) joint distribution of $a(0, n\varepsilon), b(0, n\varepsilon), c(0, n\varepsilon)$ is found.

7b7 Exercise. For every n, the joint distribution of $a(0, n\varepsilon)$, $b(0, n\varepsilon)$ and $c(0, n\varepsilon)$ is the same as the joint distribution of $a(0, n\varepsilon)$, $b(0, n\varepsilon)$ and $\max\big(0, a(0, n\varepsilon)+b(0, n\varepsilon)-\sqrt{\varepsilon}(G-1)\big)$, where G is a random variable independent of the random vector $\big(a(0, n\varepsilon), b(0, n\varepsilon)\big)$ and distributed geometrically, $\mathrm{G}(\sqrt{\varepsilon})$.

Prove it.

7c Scaling limit

The scaling limit (for $\varepsilon \to 0$) of the (3-dimensional) joint distribution of $a(s, t)$, $b(s, t)$ and $c(s, t)$ results from 7b7 and (6c9). It may be thought of as the joint distribution of $a(s, t)$, $b(s, t)$ and

(7c1) $$\max\big(0, a(s, t) + b(s, t) - \eta\big) = \big(a(s, t) + b(s, t) - \eta\big)^+ ,$$

where η is a random variable distributed exponentially, Exp(1), independent of the random vector $\big(a(s, t), b(s, t)\big)$, the vector being distributed (6c9). We have a probability measure μ_{t-s} on the 3-dimensional semigroup G_3 (recall (7b2)).

However, existence of such a 3-dimensional scaling limit is necessary but not sufficient for existence of scaling limit of our process. We need a limiting probability distribution on the infinite-dimensional space of functions $[0, \infty) \to G_3$, or equivalently, $[0, \infty) \to \mathbb{R}^3$, $t \mapsto \big(a(0, t), b(0, t), c(0, t)\big)$. No problem with $a(0, \cdot)$; this one is just the usual Brownian motion (recall (7b2)). No problem with $b(0, \cdot)$, too; this one is the 'cumulative minimum' of the Brownian motion, see (6b5) and the figure after (6c9). However, $c(0, \cdot)$ is a problem.

Postpone infinite dimension, and consider for now dimension 6:

$$g(0, s) = \big(a(0, s), b(0, s), c(0, s)\big) ,$$
$$g(0, t) = \big(a(0, t), b(0, t), c(0, t)\big) ;$$

here $t > s > 0$ are given. What about a scaling limit for these six?

We have a scaling limit of another 6-tuple,

$$g(0, s) = \big(a(0, s), b(0, s), c(0, s)\big) ,$$
$$g(s, t) = \big(a(s, t), b(s, t), c(s, t)\big)$$

(think, why). Also, $g(0, t)$ is a function of $g(0, s)$ and $g(s, t)$, see 7b1(h). However, the composition $G_3 \times G_3 \to G_3$ is discontinuous (on the surface $c_1 = b_2$). We have weak convergence of measures,

$$\mu_{\varepsilon,s} \to \mu_s , \quad \mu_{\varepsilon,t-s} \to \mu_{t-s} \quad \text{for } \varepsilon \to 0 ,$$

which means that

$$\int \varphi \, d\mu_{\varepsilon,s} \to \int \varphi \, d\mu_s$$

for every bounded continuous $\varphi : G_3 \to \mathbb{R}$. However, what happens if φ is discontinuous? Still convergence, provided that φ is continuous μ_s-almost everywhere.

We have

$$\mu_{\varepsilon,s} \otimes \mu_{\varepsilon,t-s} \to \mu_s \otimes \mu_{t-s} \quad \text{on } G_3 \times G_3 \,,$$

and the question is, whether the set of continuity, $\{c_1 \ne b_2\}$, contains $(\mu_s \otimes \mu_{t-s})$-almost all points, or not. The answer is positive, since these c_1 and b_2 are independent, and b_2 is nonatomic. The 6-dimensional scaling limit exists!

Similarly, for every n and every $0 < t_1 < \dots < t_n$, there exists a ($3n$-dimensional) scaling limit of $g(0,t_1),\dots,g(0,t_n)$, that is, of $a(0,t_1),b(0,t_1),c(0,t_1),\dots,a(0,t_n),b(0,t_n),c(0,t_n)$. Finite dimensions are done. What happens in the space of functions?

Recall 6b6: the Brownian motion satisfies the Hölder condition

$$\sup_{0\le s<t\le 1} \frac{|B(t)-B(s)|}{(t-s)^{1/3}} < \infty \quad \text{a.s.},$$

and the same for the random walk, uniformly in ε.[3] For a large C, the set of functions satisfying $\sup(\dots) \le C$, is of high probability, and a continuous function on such a set is close to a 'finite-dimensional' function.

The process $b(0,\cdot)$ satisfies the same Hölder condition, since it is the 'accumulated minimum' of the Brownian motion. Also, the process $c(0,\cdot)$ satisfies the condition (both in discrete time and in continuous time), since

$$\max_{|i-j|\le k} |c(0,i\varepsilon) - c(0,j\varepsilon)| \le \max_{|i-j|\le k} |a(0,i\varepsilon) - a(0,j\varepsilon)|$$

(think, why). It leads to a proposition similar to 6b6 for the sticky flow.

So, the scaling limit exists. The process $a(0,\cdot)$ is the usual Brownian motion; the process $b(0,\cdot)$ is a function of $a(0,\cdot)$.[4] What about $c(0,\cdot)$? Is it also a function of $a(0,\cdot)$? Of course, in discrete time the answer is negative (because a does not discern f_+ and f_*). So what? The additional randomness could disappear in the scaling limit. Does it, really?

Recalling (7b6), 7b7 and (7c1), it is easy to guess that the conditional distribution of $c(0,t)$, given the whole Brownian past $a(0,\cdot)|_{[0,t]}$ (rather than $a(0,t)+b(0,t)$), is still described by (7c1) (and therefore $c(0,\cdot)$ is not a function of $a(0,\cdot)$). Is the guess true? Conditioning by a Brownian path is rather subtle...[5]

We introduce a random variable η_ε independent of $a_\varepsilon(0,\cdot)$, distributed such that $(\eta_\varepsilon/\sqrt{\varepsilon})+1 \sim \mathrm{G}(\sqrt{\varepsilon})$, and a random variable η independent of $a(0,\cdot)$, distributed exponentially, $\mathrm{Exp}(1)$. We have

$$\eta_\varepsilon \xrightarrow[\varepsilon\to 0]{} \eta \quad \text{in distribution,}$$

that is, $\mathbb{E} f(\eta_\varepsilon) \to \mathbb{E} f(\eta)$ for all bounded continuous functions $f : \mathbb{R} \to \mathbb{R}$. Also,

$$a_\varepsilon(0,\cdot)|_{[0,1]} \xrightarrow[\varepsilon\to 0]{} a(0,\cdot)|_{[0,1]} \quad \text{in distribution,}$$

[3] In discrete time the denominator should be corrected: $(t-s+\varepsilon)^{1/3}$.

[4] I mean, a function of the whole past, not just the current value.

[5] Recall 4a for a clash between conditioning by Brownian path and conditioning by the path of the random walk.

that is, $\mathbb{E} f(a_\varepsilon(0,\cdot)|_{[0,1]}) \to \mathbb{E} f(a(0,\cdot)|_{[0,1]})$ for all bounded continuous functions $f : C[0,1] \to \mathbb{R}$. It follows that

$$\big(a_\varepsilon(0,\cdot)|_{[0,1]}, \eta_\varepsilon\big) \xrightarrow[\varepsilon\to 0]{} \big(a(0,\cdot)|_{[0,1]}, \eta\big) \quad \text{in distribution;}$$

this time, $f : C[0,1] \times \mathbb{R} \to \mathbb{R}$. That is clear when f is a factorizable function ($f = f_1 \otimes f_2$, $f_1 : C[0,1] \to \mathbb{R}$, $f_2 : \mathbb{R} \to \mathbb{R}$), or a linear combination of factorizable functions. Are they dense (uniformly)? Yes, if spaces are compact, but they are not... However, a compact set of high probability exists for every distribution, and moreover, for every converging (therefore, precompact) sequence of distributions.

Convergence in distribution is preserved by continuous maps. Thus,

$$\big(a_\varepsilon(0,\cdot)|_{[0,1]},\, (a_\varepsilon(0,1) + b_\varepsilon(0,1) - \eta_\varepsilon)^+\big) \xrightarrow[\varepsilon\to 0]{} \big(a(0,\cdot)|_{[0,1]},\, (a(0,1) + b(0,1) - \eta)^+\big)\,.$$

However,

$$\big(a_\varepsilon(0,\cdot)|_{[0,1]},\, (a_\varepsilon(0,1) + b_\varepsilon(0,1) - \eta_\varepsilon)^+\big) \sim \big(a_\varepsilon(0,\cdot)|_{[0,1]},\, c_\varepsilon(0,1)\big)$$

by (7b6) and 7b7. Also,

$$\big(a_\varepsilon(0,\cdot)|_{[0,1]},\, c_\varepsilon(0,1)\big) \xrightarrow[\varepsilon\to 0]{} \big(a(0,\cdot)|_{[0,1]},\, c(0,1)\big) \quad \text{in distribution}$$

(since a, b, c are the scaling limit of $a_\varepsilon, b_\varepsilon, c_\varepsilon$). Therefore

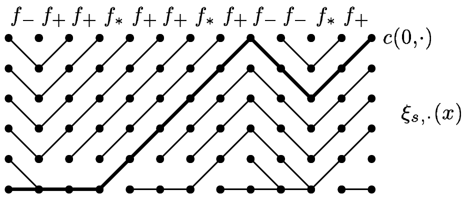
$$\big(a(0,\cdot)|_{[0,1]},\, c(0,1)\big) \sim \big(a(0,\cdot)|_{[0,1]},\, (a(0,1) + b(0,1) - \eta)^+\big)\,,$$

which was discovered by Jon Warren.[6] So, $b(0,\cdot)$ is a function of $a(0,\cdot)$, but $c(0,\cdot)$ is not.

7d When space becomes time...

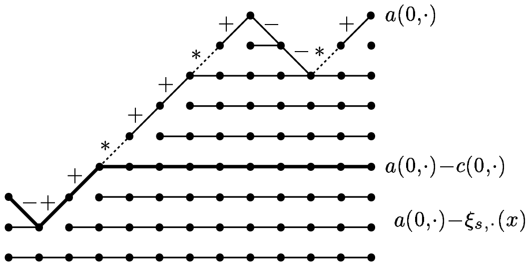
Let us look again at the sticky flow in discrete time:

(7d1)



For large x we have $a(0,\cdot) - \xi_{s,\cdot}(x) = \text{const}$. It is instructive to look at $a(0,\cdot) - \xi_{s,\cdot}(x)$ for all x:

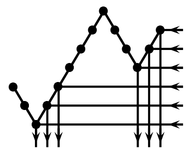
(7d2)



[6] Jonathan Warren, "Branching processes, the Ray-Knight theorem, and sticky Brownian motion", Lecture Notes in Mathematics **1655**, 1–15 (1997).

Given a path $\big(a(0,k\varepsilon)\big)_{k=0,\dots,n}$ of the driver, we consider

(7d3)
$$\sigma_n(l\sqrt{\varepsilon}) = \sqrt{\varepsilon}\cdot\max\{k\in\{0,\dots,n\} : a(0,k\varepsilon) = l\sqrt{\varepsilon}\}$$
$$\text{for } l\sqrt{\varepsilon} \in [\underbrace{\min_{[0,n\varepsilon]} a(0,\cdot)}_{=-b(0,n\varepsilon)}, a(0,n\varepsilon)]\,.$$



7d4 Exercise.

$$a(0,n\varepsilon) - c(0,n\varepsilon) = \min\big(a(0,n\varepsilon), \min\{x : \xi_{\sigma_n(x+\sqrt{\varepsilon})} = f_*\}\big)\,;$$

here x runs over $\sqrt{\varepsilon}\mathbb{Z}\cap[-b(0,n\varepsilon), a(0,n\varepsilon))$.

Prove it.

Hint: use 7b1 and/or look at (7d2), (7d3).

7d5 Exercise. Deduce 7b5 from 7d4.

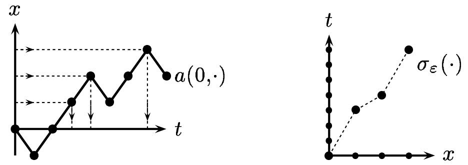
It is natural to treat such random functions as $a(0,\cdot)$ as random maps from the time to the space. Then, $\sigma_n(\cdot)$ is a random map from the space to the time! Nevertheless, we may treat it as a random process, which leads to interesting results. However, for a more elegant theory, we change the construction a little.

Having the random walk

$$a(0,k\varepsilon) = \sqrt{\varepsilon}\big(\tau(\varepsilon)+\dots+\tau(k\varepsilon)\big)\,,$$

we define

(7d6)
$$\sigma_\varepsilon(l\sqrt{\varepsilon}) = \varepsilon\min\{k : a(0,k\varepsilon) = l\sqrt{\varepsilon}\} \quad \text{for } l = 0,1,2,\dots$$

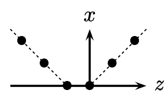


Increments $\sigma_\varepsilon(\sqrt{\varepsilon}) = \sigma_\varepsilon(\sqrt{\varepsilon}) - \sigma_\varepsilon(0)$, $\sigma_\varepsilon(2\sqrt{\varepsilon}) - \sigma_\varepsilon(\sqrt{\varepsilon})$, $\sigma_\varepsilon(3\sqrt{\varepsilon}) - \sigma_\varepsilon(2\sqrt{\varepsilon})$, ... are independent and identically distributed.[7] The distribution of $\sigma_\varepsilon(\sqrt{\varepsilon})$ is closely related to the distribution of $b(0,k\varepsilon) = \max_{i=0,\dots,k}\big(-a(0,i\varepsilon)\big) \sim \max_{i=0,\dots,k} a(0,i\varepsilon)$, namely,

$$\mathbb{P}\big(\,\sigma_\varepsilon(\sqrt{\varepsilon}) > k\varepsilon\,\big) = \mathbb{P}\big(\,b(0,k\varepsilon) = 0\,\big)$$

(think, why). The distribution of $b(0,k\varepsilon)$ was calculated in 6c12. It appeared (recall (6b10)), that

$$b(0,k\varepsilon) \sim \left| a(0,k\varepsilon) + \frac{\sqrt{\varepsilon}}{2}\right| - \frac{\sqrt{\varepsilon}}{2}\,,$$



thus $\mathbb{P}\big(\,b(0,k\varepsilon)=0\,\big) = \mathbb{P}\big(-\sqrt{\varepsilon} \le a(0,k\varepsilon) \le 0\,\big)$, that is,

$$\mathbb{P}\big(\,\sigma_\varepsilon(\sqrt{\varepsilon}) > k\varepsilon\,\big) = \begin{cases} 2^{-k}\binom{k}{k/2} & \text{for } k \text{ even,} \\ 2^{-k}\binom{k}{(k-1)/2} = 2^{-k}\binom{k}{(k+1)/2} & \text{for } k \text{ odd,} \end{cases}$$

[7] See Footnote 1 on page 63.

therefore

$$\mathbb{P}\left(\,\sigma_\varepsilon(\sqrt{\varepsilon}) = k\varepsilon\,\right) = \begin{cases} 0 & \text{for } k \text{ even,} \\ 2^{-(k-1)}\frac{1}{k+1}\binom{k-1}{(k-1)/2} & \text{for } k \text{ odd.} \end{cases}$$

(Of course, "0 for k even" is evident; think, why.)

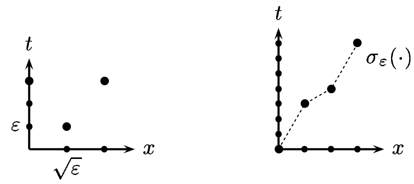
k	1	3	5	7	...
$\mathbb{P}\left(\,\sigma_\varepsilon(\sqrt{\varepsilon}) = k\varepsilon\,\right)$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{5}{128}$	...

Using the local limit theorem (or just the Stirling formula) we get

(7d7)
$$2^{-(k-1)}\binom{k-1}{(k-1)/2} \sim \sqrt{\frac{2}{\pi k}} \quad \text{for } k \to \infty\,;$$
$$\mathbb{P}\left(\,\sigma_\varepsilon(\sqrt{\varepsilon}) = t\,\right) \sim \sqrt{\frac{2}{\pi}}\left(\frac{\varepsilon}{t}\right)^{3/2} \quad \text{for } t \to \infty,\ t \in \varepsilon(2\mathbb{Z}+1)\,.$$

The graph of the function

(7d8)
$$\sqrt{\varepsilon}\mathbb{Z}_+ \ni x \mapsto \sigma_\varepsilon(x+\sqrt{\varepsilon}) - \sigma_\varepsilon(x)$$



is a random subset of $(\sqrt{\varepsilon}\mathbb{Z}_+) \times (\varepsilon\mathbb{Z}_+)$. Assigning measure 1 to each point of the subset we get a random measure Π_ε such that

$$\iint \varphi\,d\Pi_\varepsilon = \iint \varphi(x,t)\,d\Pi_\varepsilon(x,t) = \sum_{l=0}^{\infty} \varphi\big(l\sqrt{\varepsilon}, \sigma_\varepsilon(l\sqrt{\varepsilon}+\sqrt{\varepsilon}) - \sigma_\varepsilon(l\sqrt{\varepsilon})\big)\,;$$
$$\mathbb{E}\iint \varphi\,d\Pi_\varepsilon = \sum_{k,l\in\mathbb{Z}_+} \varphi(l\sqrt{\varepsilon}, k\varepsilon)\mathbb{P}\left(\,\sigma_\varepsilon(\sqrt{\varepsilon}) = k\varepsilon\,\right).$$

Note that

(7d9)
$$\sigma_\varepsilon(l\sqrt{\varepsilon}) = \iint_{[0,l\sqrt{\varepsilon})\times[0,\infty)} t\,d\Pi_\varepsilon\,.$$

Let $\varphi : [0,\infty)\times[0,\infty) \to \mathbb{R}$ be a Riemann integrable function, concentrated on $[0,x_{\max}]\times[t_{\min},t_{\max}]$ for some $t_{\min} > 0$, $t_{\max} < \infty$, $x_{\max} < \infty$, then

$$\mathbb{E}\iint \varphi\,d\Pi_\varepsilon \xrightarrow[\varepsilon\to 0]{} \frac{1}{\sqrt{2\pi}}\iint \frac{\varphi(x,t)}{t^{3/2}}\,dxdt\,.$$

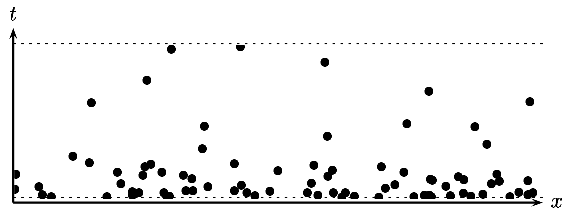
We may guess that the scaling limit of Π_ε is a Poisson process Π satisfying

$$\mathbb{E}\iint \varphi\,d\Pi = \frac{1}{\sqrt{2\pi}}\iint \frac{\varphi(x,t)}{t^{3/2}}\,dxdt\,.$$

True, events $A_{k,l} = \{\sigma_\varepsilon(l\sqrt{\varepsilon}+\sqrt{\varepsilon}) - \sigma_\varepsilon(l\sqrt{\varepsilon}) = k\varepsilon\}$ are not independent; rather,

$$\mathbb{P}\left(A_{k_1,l_1} \cap A_{k_2,l_2}\right) = \begin{cases} \mathbb{P}\left(A_{k_1,l_1}\right) & \text{if } k_1 = k_2,\ l_1 = l_2, \\ \mathbb{P}\left(A_{k_1,l_1}\right) \cdot \mathbb{P}\left(A_{k_2,l_2}\right) & \text{if } l_1 \ne l_2, \\ 0 & \text{if } k_1 \ne k_2,\ l_1 = l_2; \end{cases}$$

but we saw a similar situation in 2b, and it did not prevent the scaling limit from being a Poisson process. The same argument works here. Note also that the construction of the Poisson process, given in 2b, works over any measure space, not just $\mathbb{R}$ with Lebesgue measure. Here, we use it over the measure space $[0,\infty) \times [0,\infty)$ with the measure $\frac{1}{\sqrt{2\pi}} \frac{dxdt}{t^{3/2}}$.



The random variable $\iint \varphi \, d\Pi$ is well-defined for every measurable function φ such that $\iint |\varphi(x,t)| \frac{dxdt}{t^{3/2}} < \infty$. (If φ takes on two values $0, 1$ only, then the random variable has a Poisson distribution, of course.) Especially, the random process

$$\sigma(x) = \iint_{[0,x)\times(0,\infty)} t \, d\Pi$$

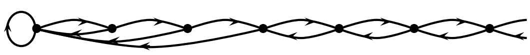
is the scaling limit of σ_ε; it is related to the Brownian motion $a(0,\cdot)$ by

$$\sigma(x) = \inf\{t \in [0,\infty) : a(0,t) = x\} \,.$$

The process $\sigma(\cdot)$ is an increasing process with stationary independent increments.[8]

7e Another discrete model

Two models converging to Poisson process were considered in 7a, but only one sticky model — in 7b. Here is the other:



The first m points form a trap. Being outside the trap, at $k\sqrt{\varepsilon}$ where $k > m$, the process jumps to $(k \pm 1)\sqrt{\varepsilon}$ with probabilities $50\%, 50\%$. Being inside the trap, at $k\sqrt{\varepsilon}$ where $k \in \{0, 1, \dots, m\}$, the process jumps either to 0 or to $(k+1)\sqrt{\varepsilon}$; probabilities are $50\%, 50\%$. That is the one-point motion of a stochastic flow built out of two maps:

$$g_+(x) = x + \sqrt{\varepsilon} \,,$$

$$g_-(x) = \begin{cases} x - \sqrt{\varepsilon} & \text{for } x > m\sqrt{\varepsilon}, \\ 0 & \text{for } x \le m\sqrt{\varepsilon}; \end{cases}$$

$$\xi_{k\varepsilon} = \begin{cases} g_+ & \text{if } \tau(k\varepsilon) = +1, \\ g_- & \text{if } \tau(k\varepsilon) = -1. \end{cases}$$

[8]Such processes are called 'subordinators'.

The space pitch is $\sqrt{\varepsilon}$, the time pitch is ε, and the scaling limit will be taken for $\varepsilon = 2^{-m} \to 0$, $m \to \infty$.

In terms of f_*, f_+, f_- of 7b we have

$$g_+ = f_* = f_{\sqrt{\varepsilon},0,\sqrt{\varepsilon}}\,; \quad g_- = f_+^{m-1} \circ f_-^m = f_{-\sqrt{\varepsilon},m\sqrt{\varepsilon},0}\,.$$

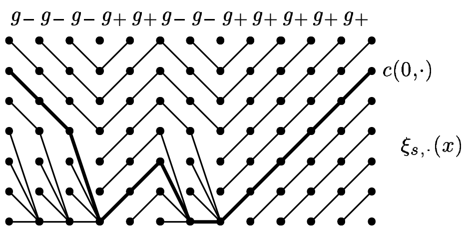
All maps belong to the 3-parameter semigroup G_3 introduced in 7b;

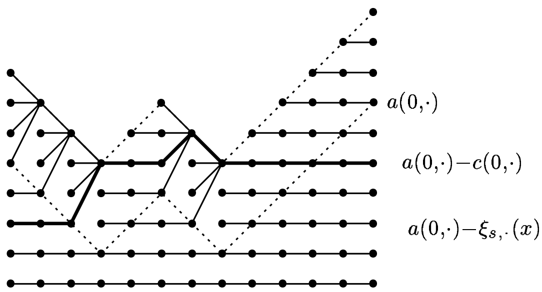
$$\xi_{s,t} = f_{a(s,t),b(s,t),c(s,t)}\,.$$

A path of the driver, $a(0,\cdot)$, is the same function of random signs $\tau(\cdot)$ as before, and it determines uniquely these random signs; therefore, $b(0,\cdot)$ and $c(0,\cdot)$ are functions of $a(0,\cdot)$.[9] Our old formula $b(0,t) = -\min_{[0,t]} a(0,\cdot)$ is too simple for the new model; however, we have

$$b(0,n\varepsilon) = m - 1 - \min_{k=0,\dots,n} a(0,k\varepsilon) \quad \text{if} \min_{k=0,\dots,n} a(0,k\varepsilon) < 0\,,$$
$$b(0,n\varepsilon) \in [0, m-1] \quad \text{if} \min_{k=0,\dots,n} a(0,k\varepsilon) = 0$$

(think, why). That is enough for the scaling limit of $b(0,\cdot)$; it is the same as before. In order to understand $c(0,\cdot)$ we use the approach of 7d. Here are modified versions of (7d1), (7d2):





We restrict ourselves to the case $\min_{[0,t]} a(0,\cdot) < 0$, since the case $\min(\dots) = 0$ disappears in the scaling limit. In order to escape the trap at $(k\varepsilon, l\sqrt{\varepsilon})$ we need $(m+1)$ positive steps of $a(0,\cdot)$ in succession, $a(0,(k-i)\varepsilon) = (l-i)\sqrt{\varepsilon}$ for $i = 0, 1, \dots, m$. In order to remain outside the trap until $n\varepsilon$, we need $\sigma_n\big((l-i)\sqrt{\varepsilon}\big) = (k-i)\varepsilon$ for $i = 1, \dots, m$.[10] Thus, instead of 7d4 we get

$$a(0,n\varepsilon) - c(0,n\varepsilon) = \min\{x : \sigma_n(x - \sqrt{\varepsilon}) - \sigma_n(x - m\sqrt{\varepsilon}) = (m-1)\varepsilon\}\,,$$

if such x exists in the set $\sqrt{\varepsilon}\mathbb{Z} \cap [\min_{[0,n\varepsilon]} a(0,\cdot) + m\sqrt{\varepsilon}, a(0,n\varepsilon)]$. Otherwise, $c(0,n\varepsilon) \in [-m\sqrt{\varepsilon}, 0]$.

We turn to σ_ε as defined by (7d6), and consider the random variable

$$X_\varepsilon = \sqrt{\varepsilon}\min\{l : \sigma_\varepsilon(l\sqrt{\varepsilon}) - \sigma_\varepsilon((l-m+1)\sqrt{\varepsilon}) = (m-1)\varepsilon\}\,. \tag{7e1}$$

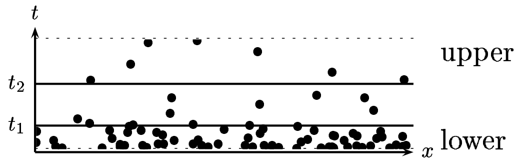
In terms of the random set (7d8), X locates the first occurrence of $m-1$ points, in succession, in the bottom row. The distribution of X_ε converges (in the scaling limit $\varepsilon = 2^{-m} \to 0$) to

[9] In discrete time, of course.

[10] It need not hold for $i = 0$; think, why.

the exponential distribution, $X \sim \mathrm{Exp}(1)$, by the same argument as in 7a. In order to check that models of 7b and 7d have the same scaling limit, we'll prove that X is independent of $a(0,\cdot)$ in the scaling limit.[11]

In fact, we'll prove a much stronger statement — weak dependence between 'upper' and 'lower' parts of the random subset of $[0,\infty)\times[0,\infty)$.[12] By the 'upper' part I mean its intersection with $[0,\infty)\times[t_2,\infty)$; by the 'lower' part — its intersection with $[0,\infty)\times[0,t_1)$; and $t_1 \ll t_2$.



7e2 Exercise.

$$|\operatorname{Corr}\big(\varphi(\min(\sigma_\varepsilon(\sqrt{\varepsilon}),t_1)),\ \psi(\max(t_2,\sigma_\varepsilon(\sqrt{\varepsilon})))\big)| \le \sqrt{\frac{p_1}{1-p_1}}\sqrt{\frac{p_2}{1-p_2}}$$

for every bounded functions $\varphi,\psi:\mathbb{R}\to\mathbb{R}$ such that the correlation coefficient is well-defined;[13] here $p_1 = \mathbb{P}\big(\sigma_\varepsilon(\sqrt{\varepsilon}) < t_1\big)$, $p_2 = \mathbb{P}\big(\sigma_\varepsilon(\sqrt{\varepsilon}) > t_2\big)$.

Prove it.[14]

Hint:

$$\big(\varphi(\min(\sigma_\varepsilon(\sqrt{\varepsilon}),t_1)) - \varphi(t_1)\big)\big(\psi(\max(t_2,\sigma_\varepsilon(\sqrt{\varepsilon}))) - \psi(t_2)\big) = 0\,;$$

$$\operatorname{Corr}(\dots) = -\frac{\mathbb{E}\,(\varphi(\min(\dots)) - \varphi(t_1))}{\sqrt{\operatorname{Var}(\varphi(\min(\dots)))}}\frac{\mathbb{E}\,(\varphi(\max(\dots)) - \varphi(t_2))}{\sqrt{\operatorname{Var}(\psi(\max(\dots)))}}\,;$$

find worst functions φ,ψ.

In other words, the so-called *maximal correlation coefficient* between two sub-σ-fields does not exceed (in fact, is equal to) $\sqrt{\frac{p_1}{1-p_1}}\sqrt{\frac{p_2}{1-p_2}}$; one sub-$\sigma$-field is generated by the random variable $Y_1 = \min(\sigma_\varepsilon(\sqrt{\varepsilon}),t_1))$, the other — by $Z_1 = \max(t_2,\sigma_\varepsilon(\sqrt{\varepsilon})))$.

Consider also $Y_2 = \min\big(\sigma_\varepsilon(2\sqrt{\varepsilon}) - \sigma_\varepsilon(\sqrt{\varepsilon}), t_1\big)$ and $Z_2 = \max\big(\sigma_\varepsilon(2\sqrt{\varepsilon}) - \sigma_\varepsilon(\sqrt{\varepsilon}), t_2\big)$. Clearly, pairs (Y_1,Z_1) and (Y_2,Z_2) are independent and identically distributed.

7e3 Exercise.

$$|\operatorname{Corr}\big(\varphi(Y_1,Y_2),\psi(Z_1,Z_2)\big)| \le \sqrt{\frac{p_1}{1-p_1}}\sqrt{\frac{p_2}{1-p_2}}\,;$$

here p_1,p_2 are the same as in 7e2, and $\varphi,\psi:\mathbb{R}^2\to\mathbb{R}$ are arbitrary bounded functions.[15]

Prove it.[16]

Hint: $\varphi(Y_1,Y_2) = \sum_{i,j}\varphi_i(Y_1)\varphi_j(Y_2)$, where $\varphi_0,\varphi_1,\dots$ are an orthonormal basis (in the corresponding L_2), and $\varphi_0(\cdot) = 1$. The same for $\psi(Z_1,Z_2)$.

---

[11] In spite of functional dependence of X_ε on $a_\varepsilon(0,\cdot)$ in discrete time.

[12] In the scaling limit these are independent, of course. However, in discrete time some dependence exists.

[13] $\operatorname{Corr}(Y,Z) = \frac{\operatorname{Cov}(Y,Z)}{\sqrt{\operatorname{Var}(Y)}\sqrt{\operatorname{Var}(Z)}}$ is well-defined, if $\operatorname{Var}(Y)\ne 0$, $\operatorname{Var}(Z)\ne 0$.

[14] You may also recall 4f5.

[15] Such that the correlation coefficient is well-defined.

[16] You may also recall 3b18.

That fairly general argument is called *tensorization* of the maximal correlation coefficient. It works equally well for three and more pairs. We see that the maximal correlation coefficient between the two parts ('lower' and 'upper') of our random set does not exceed $\sqrt{\frac{p_1}{1-p_1}}\sqrt{\frac{p_2}{1-p_2}}$.

For $\varepsilon \to 0$ we have (recall 7d7)

$$1 - p_1 = \mathbb{P}\left(\sigma_\varepsilon(\sqrt{\varepsilon}) \ge t_1\right) \sim \sqrt{\frac{2\varepsilon}{\pi t_1}},$$

$$p_2 = \mathbb{P}\left(\sigma_\varepsilon(\sqrt{\varepsilon}) > t_2\right) \sim \sqrt{\frac{2\varepsilon}{\pi t_2}},$$

$$\sqrt{\frac{p_1}{1-p_1}}\sqrt{\frac{p_2}{1-p_2}} \to \sqrt[4]{\frac{t_1}{t_2}}.$$

The scaling limit X of X_ε (defined by (7e1)) is therefore independent of the 'upper' part of the 2-dimensional Poisson process, for every t_2. It means that X is independent of the whole Poisson process. So, models of 7d and 7b give the same scaling limit.[17]

[17]... assuming existence of a scaling limit of the model of 7d. Its existence can be proven by combining arguments of 7e and 7c.