MVA25 - Mathematical methods for neurosciences

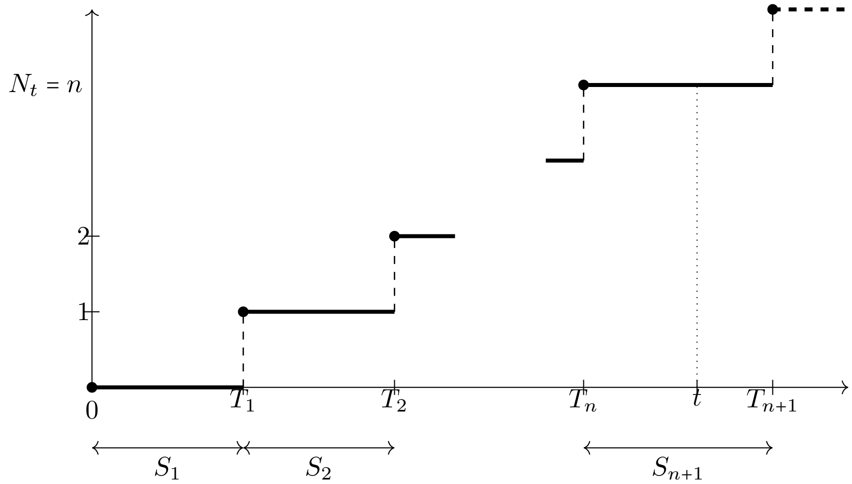
Quentin Cormier
September 11, 2026

Contents

1 Motivations

We would like to have mathematical models to study either one neuron, or few hundreds of neurons interacting with each other (mesoscopic population) or to model a macroscopic population (more than few thousands). Crucially, we need stochastic models to account for the variability of experimental results: the same repeated experiment typically gives different spike trains.

In practice, what we can measure is the spiking times of the neurons. We do not have access in many experimental settings to the full information such as the membrane voltage. For us, one neuron will be modeled by a spike train. In other words, we care about the trajectory of the spikes and their statistical properties. We care much less about the precise description of the membrane voltage which leads to these spikes. Therefore for us, the information about one neuron can be summarized by a counting process (N_t) which gives the spiking times of the neuron $T_1, T_2, \cdots$.



The Poisson process

The most basic example is the Poisson process with constant activity. In that case, the times between two successive events (*interspikes* in neurosciences) are i.i.d. with exponential law $\mathcal{E}(\lambda)$:

$$(T_{i+1} - T_i)_{i\in\mathbb{N}} \text{ i.i.d. with law } \mathcal{E}(\lambda).$$

The exponential law $\mathcal{E}(\lambda)$ has density $\lambda e^{-\lambda s}$ on $\mathbb{R}_+$. It has a very special property, the absence of memory.

Exercice 1.1. *Let S be a random variable of law $\mathcal{E}(\lambda)$. Show that*

$$\mathbb{P}(S > t + h \mid S > t) = \mathbb{P}(S > h).$$

In fact, the exponential laws are the only distributions which have this memoryless property. We shall return to this example later, and generalize it so that the intensity can depend on time (inhomogeneous Poisson process). However, the absence of memory/structure of the spike trains makes the model too limited to reproduce the statistics of real neurons.

The stochastic Integrate and fire model

The typical example we would like to study is the following stochastic variant of the Integrate and Fire model. In this model, (V_t) is the membrane potential and (N_t) is the number of spikes of the neuron up to time t. Between two spikes, the dynamics of (V_t) is governed by a deterministic ODE:

$$\dot{V}_t = b(V_t) + a_t.$$

The drift b is typically of the form $b(x) = C_1 - C_2 x$ and a_t is a time dependent current to account for the activity of other neurons or the external activity related to the experiment. For now, assume that (a_t) is deterministic as well. Neurons then spike randomly at rate $f(V_t)$, which means that:

$$\mathbb{P}(N_{t+\mathrm{d}t} - N_t > 0 \mid \text{all the past up to time } t) \approx f(V_t)\mathrm{d}t.$$

When a neuron spike, its potential is instantaneously reset to a resting value V_{rest}. For simplicity, we take $V_{rest} = 0$. One way to describe the dynamics is to write that:

$$V_t = v + \int_0^t (b(V_s) + a_s)\mathrm{d}s - \int_0^t V_{s-}\mathrm{d}N_s,$$

where N_t is the counting process: N_t is the number of spikes up to time t. Here V_{s-} is the potential immediately before time s. At a spike, the last integral subtracts this value and resets the potential to zero. Therefore, (V_t) is simply a deterministic transformation of (N_t): if we know the spiking times, we know the potential (V_t) between consecutive jumps. All the information is stored in (N_t). The first objective of the course is to explain how to define properly this object (N_t), and how to simulate trajectories of (N_t). We then study the mathematical property of (N_t, V_t): the Markov property, the renewal property due to the reset mechanism and finally the long time behavior. We shall see also some links with PDE.

2 Stochastic processes, stopping times

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We recall that $\mathcal{F}$ is a σ-field (or σ-algebra, "tribu" in French).

Definition 2.1. *A family of random variables $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F})$ is called a stochastic process.*

Given such a stochastic process $X = (X_t)_{t\geq 0}$, we let

$$\mathcal{F}_t^X = \sigma(X_s, s \in [0,t]),$$

the σ-field generated by $(X_s)_{0\leq s\leq t}$. The family $(\mathcal{F}_t^X)$ is the internal history of X.

Heuristic 2.2. *If a random variable Y is $\mathcal{F}_t^X$-measurable, then there exists a deterministic function Φ such that $Y = \Phi((X_s), s \leq t)$. In other words, Y depends causally on $(X_s), s \leq t$.*

Let $T \in \bar{\mathbb{R}}_+$ a random variable on $(\Omega, \mathcal{F})$.

Heuristic 2.3. *We say T is a stopping time if T models the arrival time of a random phenomenon which depends causally on (X_t). In other words, T is a stopping time if we can answer the following question, for all $t \geq 0$: "has the phenomenon already occurred at time t"?*

Definition 2.4. *T is a $(\mathcal{F}_t^X)$ stopping time if for all $t \geq 0$, $\{T \leq t\} \in \mathcal{F}_t^X$.*

We now generalize slightly theses two definitions.

Definition 2.5. *A sequence $(\mathcal{F}_t)_{t\geq 0}$ of σ-fields is called a filtration if for all $t, h \geq 0$ $\mathcal{F}_t \subset \mathcal{F}$ (it is a sub σ-field of $\mathcal{F}$) and $\mathcal{F}_t \subset \mathcal{F}_{t+h}$ (it is increasing). We also let $\mathcal{F}_\infty$ be the smallest σ-field which contains all the $\mathcal{F}_t$.*

Example 2.6. *Let $Z_t = (X_t, Y_t)$ for two processes X and Y. Then $\mathcal{F}_t^X \subset \mathcal{F}_t^Z$. $\mathcal{F}_t^X$ corresponds to the observation of $(X_s)_{0\leq s\leq t}$, while $\mathcal{F}_t^Z$ corresponds to the observation of both trajectories $((X_s, Y_s))_{0\leq s\leq t}$.*

If a process (X_t) is such that $\mathcal{F}_t^X \subset \mathcal{F}_t$ then we say that $(\mathcal{F}_t)$ is an history of (X_t); $(\mathcal{F}_t^X)$ is the smallest history of (X_t).

Definition 2.7. *A random variable $T : \Omega \to \bar{\mathbb{R}}_+$ is a $(\mathcal{F}_t)$-stopping time if*

$$\forall t \geq 0, \quad \{T \leq t\} \in \mathcal{F}_t.$$

Finally, we let:

$$\mathcal{F}_T := \{A \in \mathcal{F}_\infty \mid A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\}.$$

3 Martingales

Assume we know (X_t) up to the time t, i.e. we "know" $\mathcal{F}_t^X$. Let Y be random variable correlated to (X_t). We would like to find the best approximation of Y, knowing $\mathcal{F}_t^X$. In other words, we seek for $\hat{Y}_t$ a $\mathcal{F}_t^X$-measurable random variable such that

$$\forall Z \ \mathcal{F}_t^X - \text{measurable}, \quad \mathbb{E}(Y - \hat{Y}_t)^2 \leq \mathbb{E}(Y - Z)^2.$$

The answer to this problem is the **conditional expectation**:

$$\hat{Y}_t = \mathbb{E}[Y \mid \mathcal{F}_t^X].$$

In addition, we have by the tower property:

$$\mathbb{E}[\hat{Y}_t \mid \mathcal{F}_s^X] = \mathbb{E}[\mathbb{E}[Y \mid \mathcal{F}_t^X] \mid \mathcal{F}_s^X] = \mathbb{E}[Y \mid \mathcal{F}_s^X] = \hat{Y}_s.$$

We say, according to the following definition, that $(\hat{Y}_t)$ is a martingale.

Definition 3.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $(\mathcal{F}_t)$ a filtration. A $(\mathbb{P}, \mathcal{F}_t)$-martingale on $[0, T]$ is a $\mathbb{R}$-valued process (X_t) such that:*

- *X_t is $\mathcal{F}_t$-adapted.*
- *$\mathbb{E}|X_t| < \infty$, for all $t \in [0, T]$.*
- *For all $0 \leq s \leq t \leq T$, $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$, almost surely.*

Remark 3.2. *If for all $s \leq t \leq T$, $\mathbb{E}[X_t \mid \mathcal{F}_s] \leq X_s$, we say X is a super-martingale. If $\mathbb{E}[X_t \mid \mathcal{F}_s] \geq X_s$, we say X is a sub-martingale.*

Exercice 3.3. *Let (X_t) be a process with independent increment, that is*

$$X_t - X_s \text{ is independent of } \mathcal{F}_s^X, \quad \forall 0 \leq s \leq t.$$

Assume that $\mathbb{E}|X_t| < \infty$ and $\mathbb{E}X_t = 0$. Show that (X_t) is a $(\mathbb{P}, \mathcal{F}_t^X)$-martingale.

Exercice 3.4. *Assume in addition that $\mathbb{E}X_t^2 < \infty$. Show that*

1. *$X_t^2 - \mathbb{E}X_t^2$ is a $(\mathbb{P}, \mathcal{F}_t^X)$-martingale.*
2. *X_t^2 is a $(\mathbb{P}, \mathcal{F}_t^X)$ sub-martingale.*

Definition 3.5. *Let (X_t) be a stochastic process adapted to $(\mathcal{F}_t)$. Let (S_n) be a sequence of increasing stopping times, such that $\lim_{n\uparrow\infty} S_n = \infty$. If for all n, $(X_{t\wedge S_n})$ is a martingale, we say (X_t) is a local martingale. We say (X_t) is localized by the sequence (S_n).*

Exercice 3.6. *We recall Fatou's conditional lemma. Let $X_n \geq 0$ be a sequence of random variables and $\mathcal{G} \subset \mathcal{F}$. Then*

$$\liminf_{n\to\infty} \mathbb{E}[X_n \mid \mathcal{G}] \geq \mathbb{E}[\liminf_{n\to\infty} X_n \mid \mathcal{G}].$$

Let (X_t) be a $(\mathbb{P},\mathcal{F}_t)$-local martingale, such that $X_t \geq 0$ and $\mathbb{E}|X_t| < \infty$, for all t. Show that X_t is a $(\mathbb{P},\mathcal{F}_t)$ super-martingale.

Exercice 3.7. *Let $(\Omega,\mathcal{F},(\mathcal{F}_t),P)$ be a filtered probability space. Assume (M_t) is a $(\mathbb{P},\mathcal{F}_t)$ martingale. Let $(\mathcal{G}_t)$ be a filtration with less information than $(\mathcal{F}_t)$, that is: $\mathcal{G}_t \subset \mathcal{F}_t$ for all t. Assume that M_t is adapted to $(\mathcal{G}_t)$ (a possible choice is $\mathcal{G}_t = \mathcal{F}_t^M$...) Show that M_t is also a $(\mathbb{P},\mathcal{G}_t)$-martingale.*

4 Predictable processes

Definition 4.1. *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space. Let $\mathbb{F} = (\mathcal{F}_t)$ be a filtration. We define $\mathcal{P}(\mathbb{F})$ the σ-field generated by the rectangles of the form:*

$$(s,t] \times A, \quad A \in \mathcal{F}_s, 0 \leq s \leq t.$$

We say $\mathcal{P}(\mathbb{F})$ is the predictable σ-field on $(0,\infty)\times\Omega$.

A process (X_t) such that X_0 is $\mathcal{F}_0$-measurable and such that $(t,\omega) \to X_t(\omega)$ is $\mathcal{P}(\mathbb{F})$-measurable is said to be $(\mathcal{F}_t)$ predictable.

In practice, we will mostly use:

Theorem 4.2. *A process (X_t) adapted to $(\mathcal{F}_t)$ and with left-continuous trajectories is $\mathcal{F}_t$-predictable.*

Example 4.3. *Take (N_t) a Poisson process. The trajectories of (N_t) are cadlag (continue à droite, limité à gauche). Then N_{t-} is caglad (continue à gauche, limité à droite) and so $\mathcal{F}_t^N$-predictable.*

5 Integration with respect to a martingale with finite variation

Let (M_t) be a $(\mathbb{P},\mathcal{F}_t)$ martingale. Assume that (M_t) has almost surely finite variation on $[0,t]$, for all $t \geq 0$. That means that

$$\int_0^t |\mathrm{d}M_s| = \sup\left\{\sum_{i=1}^p |M_{t_i} - M_{t_{i-1}}|\right\} < \infty,$$

where the supremum is taken on all the subdivisions $0 = t_0 < t_1 < \cdots < t_p = t$ of $[0,t]$. We assume that

$$\mathbb{E}\int_0^t |\mathrm{d}M_s| < \infty.$$

Theorem 5.1. *Let (C_t) be a $(\mathcal{F}_t)$-predictable process such that $\mathbb{E}\int_0^1 |C_s||\mathrm{d}M_s| < \infty$. Then $t \mapsto \int_0^t C_s \mathrm{d}M_s$ is a $(\mathbb{P}, \mathcal{F}_t)$ martingale on $[0,1]$.*

We shall not prove the result here. But to explain this result, consider the predictable process

$$C_u = \mathbb{1}_A \mathbb{1}_{(s,t]}(u), \quad A \in \mathcal{F}_s.$$

Then

$$\tilde{M}_\theta := \int_0^\theta C_u \mathrm{d}M_u = \mathbb{1}_A (M_{t\wedge\theta} - M_{s\wedge\theta}) = \begin{cases} 0 & \text{if } \theta \le s \\ \mathbb{1}_A (M_\theta - M_s) & \text{if } \theta \in [s,t] \\ \mathbb{1}_A (M_t - M_s) & \text{if } \theta > t \end{cases}.$$

Therefore, for all $u \le \theta$,

$$\mathbb{E}[\tilde{M}_\theta \mid \mathcal{F}_u] = \mathbb{E}[\mathbb{1}_A (M_{t\wedge\theta} - M_{s\wedge\theta}) \mid \mathcal{F}_u] = \tilde{M}_u.$$

Indeed, consider first the case $\theta \le s$. Then $\theta \ge s, u \le s$. Then $\theta, u \ge s$. And use each time the martingale property of M. In fact (M_t) is $(\mathbb{P}, \mathcal{F}_t)$ martingale is equivalent to

$$\mathbb{E}\left[\int_0^\infty C_u \mathrm{d}M_u\right] = 0,$$

for all predictable process of the form $C_u = \mathbb{1}_A \mathbb{1}_{(s,t]}(u)$, where $A \in \mathcal{F}_s$.

6 Point processes

A point processes is a sequence of non-negative and increasing random variables (T_n):

$$T_0 = 0, \quad \text{and} \quad [T_n < \infty \implies T_{n+1} > T_n].$$

Let $T_\infty := \lim_{n\uparrow\infty} T_n$. We define the counting process:

$$N_t = \begin{cases} n & \text{if } t \in [T_n, T_{n+1}) \\ +\infty & \text{if } t \ge T_\infty. \end{cases}.$$

By abuse of notation, we say that (N_t) is the point process as well. There is a one to one relation between (N_t) and the sequence (T_n). On every interval $[0,t]$ with $t < T_\infty$, the trajectories of (N_t) are cadlag and have finite variation. For a predictable process (C_t), we define

$$\int_0^t C_s \mathrm{d}N_s = \sum_{\substack{n \ge 1 \\ T_n \le t}} C_{T_n}, \qquad t < T_\infty.$$

For $C \ge 0$, the same sum is defined for all $t \ge 0$, possibly with value $+\infty$. The theorem below will rule out explosion when N admits an intensity. Now, the fundamental definition is the following:

Definition 6.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space, $(\mathcal{F}_t)$ a filtration and (N_t) a point process adapted to $(\mathcal{F}_t)$. Let (λ_t) be a non-negative $(\mathcal{F}_t)$ predictable process such that*

$$a.s, \quad \forall t \ge 0, \quad \int_0^t \lambda_s \mathrm{d}s < \infty.$$

If for all non-negative $\mathcal{F}_t$-predictable process (C_t) the equality

$$\mathbb{E}\left[\int_0^\infty C_s \mathrm{d}N_s\right] = \mathbb{E}\left[\int_0^\infty C_s \lambda_s \mathrm{d}s\right]$$

holds, we say that N_t admits the $(\mathbb{P}, \mathcal{F}_t)$ intensity λ_t.

Theorem 6.2. *Let (N_t) a point process which admits the $(\mathbb{P}, \mathcal{F}_t)$ intensity λ_t. Then:*

1. *N_t is not explosive, that is $N_t < \infty$ for all $t > 0$, with probability 1.*
2. *$M_t = N_t - \int_0^t \lambda_s \mathrm{d}s$ is a local martingale.*
3. *Let (X_t) be $(\mathcal{F}_t)$-predictable with $\mathbb{E}\int_0^t |X_s|\lambda_s \mathrm{d}s < \infty$ for all t. Then $\int_0^t X_s \mathrm{d}M_s$ is a martingale.*
4. *Let (X_t) be $(\mathcal{F}_t)$-predictable with a.s. $\int_0^t |X_s|\lambda_s \mathrm{d}s < \infty$ for all t. Then $\int_0^t X_s \mathrm{d}M_s$ is a local martingale.*

Proof. First point. Let $S_n = \inf\{t \geq 0 : \int_0^t \lambda_s \mathrm{d}s \geq n\}$. It is a $(\mathcal{F}_t)$ stopping-time (easy exercise!). Let $C_t = \mathbb{1}_{t \leq S_n}$, (C_t) is a predictable process (exercice!). So

$$\mathbb{E}N_{S_n} = \mathbb{E}\int_0^{S_n} \lambda_s \mathrm{d}s \leq n < \infty.$$

Therefore, by Markov's inequality:

$$\mathbb{P}(N_{S_n} > A) \leq \frac{n}{A}.$$

So $\mathbb{P}(N_{S_n} = \infty) = 0$. In addition, $S_n \uparrow \infty$ so

$$\mathbb{P}(\exists t : N_t = \infty) \leq \mathbb{P}(\exists n : N_{S_n} = \infty) \leq \sum_n \mathbb{P}(N_{S_n} = \infty) = 0.$$

To conclude, $\mathbb{P}(\forall t \geq 0,\ N_t < \infty) = 1$.

Second point. We choose $C_s = \mathbb{1}_A \mathbb{1}_{a \leq s \leq b \wedge T_n}$. So we have:

$$\mathbb{E}\mathbb{1}_A (N_{b\wedge T_n} - N_{a \wedge T_n}) = \mathbb{E}\mathbb{1}_A \int_{a\wedge T_n}^{b \wedge T_n} \lambda_u \mathrm{d}u.$$

So:

$$\mathbb{E}\left[\mathbb{1}_A\left(N_{b\wedge T_n} - \int_0^{b\wedge T_n} \lambda_u \mathrm{d}u\right)\right] = \mathbb{E}\left[\mathbb{1}_A\left(N_{a\wedge T_n} - \int_0^{a\wedge T_n} \lambda_u \mathrm{d}u\right)\right],$$

therefore $N_{b\wedge T_n} - \int_0^{b\wedge T_n} \lambda_u \mathrm{d}u$ is a $(\mathbb{P}, \mathcal{F}_t)$ martingale. As (N_t) is not explosive, $T_n \uparrow \infty$ a.s. and so M_t is a local martingale (localized by the T_n).

Third point. Take $C_s = \mathbb{1}_{a \leq s \leq b} \mathbb{1}_A X_s$. We find that

$$\mathbb{E}\mathbb{1}_A \int_a^b X_s \mathrm{d}N_s = \mathbb{E}\mathbb{1}_A \int_a^b X_s \lambda_s \mathrm{d}s.$$

Equivalently,

$$\mathbb{E}\left[\int_0^b X_s \mathrm{d}N_s - \int_0^b X_s \lambda_s \mathrm{d}s \mid \mathcal{F}_a\right] = \int_0^a X_s \mathrm{d}N_s - \int_0^a X_s \lambda_s \mathrm{d}s.$$

So $\int_0^t X_s \mathrm{d}M_s$ is a martingale. The last point is left as an exercise. □

7 The Poisson process

In this section, we assume that $\lambda(t)$ is deterministic. Consider (N_t) a $(\mathcal{F}_t)$ point process with stochastic (in fact deterministic!) intensity $\lambda(t)$. We note, as N is constant between its jumps that:

Lemma 7.1. *Let $g \in C_b(\mathbb{R}+)$ be a bounded and continuous function. Then*

$$g(N_t) = g(0) + \int_0^t [g(N_s) - g(N_{s-})]\mathrm{d}N_s.$$

We apply this result with $g(r) = e^{iur}$. We find:

$$e^{iuN_t} = 1 + \int_0^t [(e^{iu} - 1)e^{iuN_{s-}}]\mathrm{d}N_s.$$

Therefore, we find that

$$m_t := e^{iuN_t} - \int_0^t [(e^{iu} - 1)e^{iuN_{s-}}]\lambda_s \mathrm{d}s = 1 + \int_0^t [(e^{iu} - 1)e^{iuN_{s-}}]\mathrm{d}M_s,$$

where $M_t = N_t - \int_0^t \lambda(s)\mathrm{d}s$. The r.h.s. is a martingale, so for all $A \in \mathcal{F}_s$, $\mathbb{E}\mathbb{1}_A(m_t - m_s) = 0$, provided that $t \geq s$. Therefore:

$$\mathbb{E}\mathbb{1}_A\left[e^{iuN_t} - e^{iuN_s} - \int_s^t [(e^{iu} - 1)e^{iuN_{v-}}]\lambda_v \mathrm{d}v\right] = 0$$

As the jumps times are of Lebesgue measure zero, we can remove the N_{v-} in the integral. We divide by e^{iuN_s} (or multiply by its bounded inverse e^{-iuN_s}) and find that:

$$\mathbb{E}\mathbb{1}_A(e^{iu(N_t - N_s)} - 1) = \mathbb{E}\mathbb{1}_A \int_s^t [(e^{iu} - 1)e^{iu(N_v - N_s)}]\lambda_v \mathrm{d}v.$$

We let

$$H(t, s, A) = \mathbb{E}\mathbb{1}_A e^{iu(N_t - N_s)}.$$

So

$$H(t, s, A) = \mathbb{P}(A) + \int_s^t (e^{iu} - 1)H(v, s, A)\lambda(v)\mathrm{d}v.$$

That is

$$\frac{d}{dt}H(t, s, A) = (e^{iu} - 1)\lambda(t)H(t, s, A).$$

Therefore,

$$H(t, s, A) = \mathbb{P}(A)\exp\left((e^{iu} - 1)\int_s^t \lambda(v)\mathrm{d}v\right).$$

In particular

$$\mathbb{E}[e^{iu(N_t - N_s)}] = \exp\left((e^{iu} - 1)\int_s^t \lambda(v)\mathrm{d}v\right).$$

We recognize the Fourier transform of the Poisson distribution of parameter $\int_s^t \lambda(v)\mathrm{d}v$. Altogether, we have proved that: $N_t - N_s$ is independent of $\mathcal{F}_s$ (independent increment) and

$$\mathbb{P}(N_t - N_s = k) = \frac{1}{k!}\left(\int_s^t \lambda(v)\mathrm{d}v\right)^k \exp\left(-\int_s^t \lambda(v)\mathrm{d}v\right).$$

Exercice 7.2. *Show directly from the definition of (N_t) that*

$$\mathbb{P}(N_t - N_s = 0) = \exp\left(-\int_s^t \lambda(v)\mathrm{d}v\right).$$

Hint: use $C_u = \mathbb{1}_{u \le T_1 \wedge t}$.

Exercice 7.3. *What is*

$$\mathbb{E}\int_0^t N_s \mathrm{d}N_s?$$

Exercice 7.4. *Let τ be a random variable with law $\mathcal{E}(a)$, $a > 0$. Let $Z_t = \mathbb{1}_{\{\tau \le t\}}$. Show that:*

$$M_t = Z_t - a(t \wedge \tau).$$

is a $(\mathcal{F}_t^Z)$-martingale.

8 Girsanov transformation

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{Q})$ that supports a standard Poisson process N of rate 1. Denote by $(\tau_n)_{n\ge1}$ the jumps of N. We shall see how to transform this standard Poisson process to a point process of stochastic intensity by changing the probability measure. This gives a weak construction on any fixed finite interval $[0, T]$.

Theorem 8.1. *Let (μ_t) be a non-negative $(\mathcal{F}_t)$-predictable process such that*

$$\forall t \ge 0, \mathbb{Q}(\mathrm{d}\omega) a.s., \quad \int_0^t \mu_s \mathrm{d}s < \infty.$$

Let

$$L_t := \prod_{\tau_i \le t} \mu_{T_i} \exp\left(t - \int_0^t \mu_s \mathrm{d}s\right).$$

Then (L_t) is a $(\mathbb{Q}, \mathcal{F}_t)$-local martingale.

Proof. The idea is that L_t solves the following equation, where $M_t := N_t - t$:

$$L_t = 1 + \int_0^t L_{s-}(\mu_s - 1)\mathrm{d}M_s.$$

As M_t is $(\mathbb{Q}, \mathcal{F}_t)$ martingale, L_t is a local martingale by Theorem 6.2. □

Remark 8.2. *If in addition μ_s is bounded by a deterministic constant C for all t, it holds that*

$$L_t \le C^{N_t} e^t.$$

So:

$$\mathbb{E}_{\mathbb{Q}} \int_0^t L_{s-}|\mu_s - 1|\mathrm{d}s \le t e^t (C+1)\mathbb{E}_{\mathbb{Q}} C^{N_t}.$$

It holds that

$$\mathbb{E}_{\mathbb{Q}} C^{N_t} = e^{t(C-1)}.$$

Therefore,

$$\mathbb{E}_{\mathbb{Q}} \int_0^t L_{s-}|\mu_s - 1|\mathrm{d}s < \infty, \quad \forall t \ge 0.$$

Using Theorem 6.2, we deduce that L_t is true $(\mathbb{Q}, \mathcal{F}_t)$-martingale.

Theorem 8.3 (Girsanov's theorem)**.** *Fix a deterministic horizon $T > 0$. Let (μ_t) be a non-negative $(\mathcal{F}_t)$-predictable process such that $(L_t)_{0\le t\le T}$ is a true martingale. Define $\mathbb{P}$ on $\mathcal{F}_T$ by*

$$\mathrm{d}\mathbb{P} := L_T \mathrm{d}\mathbb{Q}.$$

Then $(N_t)_{0\le t\le T}$ has the $(\mathbb{P}, \mathcal{F}_t)$-intensity μ_t on $[0,T]$.

Proof. Let $(C_t)_{0\le t\le T}$ be a non-negative $\mathcal{F}_t$-predictable process. We have to prove that

$$\mathbb{E}_{\mathbb{P}} \int_0^T C_s \mathrm{d}N_s = \mathbb{E}_{\mathbb{P}} \int_0^T C_s \mu_s \mathrm{d}s.$$

Equivalently, we have to show that

$$\mathbb{E}_{\mathbb{Q}} L_T \int_0^T C_s \mathrm{d}N_s = \mathbb{E}_{\mathbb{Q}} L_T \int_0^T C_s \mu_s \mathrm{d}s.$$

The left hand side is equal to:

$$\mathbb{E}_{\mathbb{Q}} \int_0^T L_s C_s \mathrm{d}N_s$$

Similar equation can be written for the right hand side. Therefore, we have to prove that:

$$\mathbb{E}_{\mathbb{Q}} \int_0^T L_s C_s \mathrm{d}N_s = \mathbb{E}_{\mathbb{Q}} \int_0^T L_s C_s \mu_s \mathrm{d}s.$$

Finally, as $L_s \mathrm{d}N_s = L_{s-}\mu_s \mathrm{d}N_s$, we have to prove that

$$\mathbb{E}_{\mathbb{Q}} \int_0^T L_{s-} C_s \mu_s \mathrm{d}N_s = \mathbb{E}_{\mathbb{Q}} \int_0^T L_s C_s \mu_s \mathrm{d}s.$$

This is true. Why? By definition of N_t being a $(\mathbb{Q}, \mathcal{F}_t)$-point process of stochastic intensity 1! □

Remark 8.4 (Choice of the time horizon)**.** *For $0 \le t \le T$ and $A \in \mathcal{F}_t$, the martingale property gives*

$$\mathbb{P}(A) = \mathbb{E}_{\mathbb{Q}}[L_T \mathbb{1}_A] = \mathbb{E}_{\mathbb{Q}}[\mathbb{E}_{\mathbb{Q}}[L_T \mid \mathcal{F}_t]\mathbb{1}_A] = \mathbb{E}_{\mathbb{Q}}[L_t \mathbb{1}_A].$$

Thus the law of observations up to t is unchanged if we choose a larger horizon on which L is a true martingale. All applications of Girsanov below are understood on a fixed finite horizon. For the neuron model, the Poisson measure construction in Section 10 will provide trajectories for all $t \ge 0$ directly, without a measure extension argument.

Exercice 8.5. *We used that:*

$$\mathbb{E} \int_0^T L_T C_s \mathrm{d}N_s = \mathbb{E} \int_0^T L_s C_s \mathrm{d}N_s = \mathbb{E} \int_0^T L_{s-} \mu_s C_s \mathrm{d}N_s.$$

Justify this equality.
Hint: write down the definition of the integral with respect to the jump times of N_t.

9 Spiking neurons

Fix a deterministic horizon $T > 0$. In this section, the neuron model is considered on $[0,T]$; all observation times in the statements below lie in this interval. Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space and $(\mathcal{F}_t)$ a filtration. We now consider (N_t) a $(\mathbb{P}, \mathcal{F}_t)$ point process with predictable stochastic intensity $f(V_{t-})$, where:

$$\mathrm{d}V_t = b(V_t)\mathrm{d}t - V_{t-}\mathrm{d}N_t.$$

The idea is that: between the spikes, the dynamics is deterministic and given by the ODE $\dot{V}_t = b(V_t)$; spikes occur with intensity $f(V_{t-})$ and when a spike occurs, the potential is reset to zero (resting potential).

Assumption 9.1. *The function $b : \mathbb{R}_+ \to \mathbb{R}$ is Lipschitz with $b(0) \geq 0$. In addition, $f \in C^1(\mathbb{R}_+; \mathbb{R}_+)$, with $\|f\|_\infty + \|f'\|_\infty < \infty$.*

Remark 9.2. *It is not clear apriori on how to construct this object, as the stochastic intensity of N_t is a function of V_{t-}, and (V_t) itself depends on (N_t). We shall see two different constructions in a moment, one using Girsanov theorem.*

We denote by $\varphi_t(v)$ the solution of the ODE:

$$\frac{d}{dt}\varphi_t(v) = b(\varphi_t(v)); \quad \varphi_0(v) = v.$$

Proposition 9.3. *Assume that $V_0 = v$. Let τ_1 be the time of the first spike, with $\tau_1 = \infty$ if there is no spike on $[0,T]$. For $0 \leq t \leq T$, it holds that:*

$$\mathbb{P}(\tau_1 > t) = \exp\left(-\int_0^t f(\varphi_s(v))\mathrm{d}s\right).$$

Proof. Let $C_u = \mathbb{1}_{u \leq \tau_1}\mathbb{1}_{u \leq t}$. We have

$$\mathbb{E}\int_0^{t\wedge\tau_1} \mathrm{d}N_u = \mathbb{E}\int_0^{t\wedge\tau_1} f(V_u)\mathrm{d}u.$$

Therefore:

$$\mathbb{P}(\tau_1 \leq t) = \mathbb{E}\int_0^{t\wedge\tau_1} f(\varphi_u(v))\mathrm{d}u.$$

or equivalently:

$$1 - \mathbb{P}(\tau_1 > t) = \mathbb{E}\int_0^t \mathbb{1}_{\tau_1 > u} f(\varphi_u(v))\mathrm{d}u.$$

So

$$\frac{d}{dt}\mathbb{P}(\tau_1 > t) = -\mathbb{P}(\tau_1 > t) f(\varphi_t(v)).$$

This gives the stated formula. □

Representation using Girsanov's theorem

Recall under $\mathbb{Q}$, N_t is a Poisson process of rate 1. Let for all $t \geq s$:

$$V^v_{t,s} = v + \int_s^t b(V^v_{u,s})\mathrm{d}u - \int_s^t V^v_{u-,s}\mathrm{d}N_u.$$

This equation has a path-wise unique solution, because b is Lipschitz. Let also:

$$L^v_{t,s} = \prod_{s<\tau_i\leq t} f(V^v_{\tau_i-,s}) \exp\Big((t-s) - \int_s^t f(V^v_{u,s})\mathrm{d}u\Big).$$

To simplify the notations, we write:

$$L^v_t := L^v_{t,0}, \quad V^v_t := V^v_{t,0}.$$

As f is bounded, $(L^v_t)_{t\geq 0}$ is a $(\mathbb{Q}, \mathcal{F}_t)$-martingale. For each initial potential v, define $\mathbb{P}^v$ on $\mathcal{F}_T$ by

$$\mathrm{d}\mathbb{P}^v = L^v_T \mathrm{d}\mathbb{Q}.$$

Then $(N_t)_{0\leq t\leq T}$ has the $(\mathbb{P}^v, \mathcal{F}_t)$ stochastic intensity $f(V^v_{t-})$ on $[0,T]$. The measure $\mathbb{P}^v$ depends on v: each initial potential has its own change of measure. For $0 \leq t \leq T$, its restriction to $\mathcal{F}_t$ has density L^v_t with respect to $\mathbb{Q}$.

Exercice 9.4. *Using this construction, compute $\mathbb{P}^v(\tau_1 > t)$ for $0 \leq t \leq T$.*

Markov semi-group

We now prove the following Markov property.

Proposition 9.5. *Let $g : \mathbb{R}_+ \to \mathbb{R}_+$ be a non-negative measurable function. For $0 \leq t \leq T$, let*

$$\psi_t(v) := \mathbb{E}_{\mathbb{P}^v}[g(V^v_t)] = \mathbb{E}_{\mathbb{Q}}[L^v_t g(V^v_t)].$$

For $s, t \geq 0$ with $s + t \leq T$, time homogeneity under the reference measure $\mathbb{Q}$ gives

$$\mathbb{E}_{\mathbb{Q}}[L^v_{s+t,s} g(V^v_{s+t,s})] = \psi_t(v).$$

For the same s, t, the Markov property is

$$\mathbb{E}_{\mathbb{P}^v}[g(V^v_{t+s}) \mid \mathcal{F}_s] = \psi_t(V^v_s) \quad \mathbb{P}^v\text{-}a.s.$$

Proof. Under $\mathbb{Q}$, $(N_{s+u} - N_s)_{u\geq 0}$ is a standard Poisson process independent of $\mathcal{F}_s$. Since the coefficients do not depend on time, $(L^v_{s+t,s}, V^v_{s+t,s})$ has the same law under $\mathbb{Q}$ as (L^v_t, V^v_t). This proves the time homogeneity identity. Take $A \in \mathcal{F}_s$. We have

$$\begin{aligned}\mathbb{E}_{\mathbb{P}^v}[g(V^v_{t+s})\mathbb{1}_A] &= \mathbb{E}_{\mathbb{Q}}\left[L^v_{t+s} g(V^v_{t+s})\mathbb{1}_A\right] \\ &= \mathbb{E}_{\mathbb{Q}}\Big[L^v_s L^{V^v_s}_{t+s,s} g(V^{V^v_s}_{t+s,s})\mathbb{1}_A\Big].\end{aligned}$$

We used the fact that (stochastic flow property)

$$a.s, \quad V^v_{t+s,0} = V^{V^v_s}_{t+s,s},$$

and

$$L_{t+s}^{v} = L_s^v L_{t+s,s}^{V_s^v}.$$

We now take the conditional expectation by $\mathcal{F}_s$. We obtain:

$$\mathbb{E}_{\mathbb{P}^v}[g(V_{t+s}^v)\mathbb{1}_A] = \mathbb{E}_{\mathbb{Q}}\left[\mathbb{1}_A L_s^v \mathbb{E}_{\mathbb{Q}}\left[L_{t+s,s}^{V_s^v} g(V_{t+s,s}^{V_s^v}) \mid \mathcal{F}_s\right]\right].$$

Using this independence and the time homogeneity identity with the $\mathcal{F}_s$-measurable starting potential V_s^v, we have:

$$\mathbb{E}_{\mathbb{Q}}\left[L_{t+s,s}^{V_s^v} g(V_{t+s,s}^{V_s^v}) \mid \mathcal{F}_s\right] = \psi_t(V_s^v).$$

Altogether, we have

$$\mathbb{E}_{\mathbb{P}^v}[g(V_{t+s}^v)\mathbb{1}_A] = \mathbb{E}_{\mathbb{Q}}[L_s^v \mathbb{1}_A \psi_t(V_s^v)] = \mathbb{E}_{\mathbb{P}^v}[\mathbb{1}_A \psi_t(V_s^v)].$$

As this is true for all $A \in \mathcal{F}_s$ (and that V_s^v is $\mathcal{F}_s$-measurable), the result is proved. □

Exercice 9.6. *Let $0 \le s \le t_1 \le t_2 \le T$. For two non-negative test functions g_1 and g_2, define*

$$F(v') := \mathbb{E}_{\mathbb{P}^{v'}}[g_1(V_{t_1-s}^{v'}) g_2(V_{t_2-s}^{v'})].$$

Show that

$$\mathbb{E}_{\mathbb{P}^v}[g_1(V_{t_1}^v) g_2(V_{t_2}^v) \mid \mathcal{F}_s] = F(V_s^v) \quad \mathbb{P}^v\text{-a.s.}$$

Clearly this generalizes to any number of test functions $g_1, \cdots, g_k$.

The exercise above shows that $(V_t^v)_{0\le t\le T}$ has the Markov property under $\mathbb{P}^v$. Given g a non-negative measurable function and $0 \le t \le T$, let

$$S_t g(v) := \mathbb{E}_{\mathbb{P}^v}[g(V_t^v)] = \mathbb{E}_{\mathbb{Q}}[L_t^v g(V_t^v)] = \psi_t(v).$$

Proposition 9.7. *S_t satisfies the semi-group identity: for all $s, t \ge 0$ with $s + t \le T$,*

$$S_{t+s} g = S_t(S_s g).$$

Proof. For every v, the Markov property gives

$$\begin{aligned} S_{t+s} g(v) &= \mathbb{E}_{\mathbb{P}^v}\left[\mathbb{E}_{\mathbb{P}^v}[g(V_{t+s}^v) \mid \mathcal{F}_t]\right] \\ &= \mathbb{E}_{\mathbb{P}^v}[(S_s g)(V_t^v)] = S_t(S_s g)(v). \end{aligned}$$

□

Since the deterministic horizon T can be chosen arbitrarily large and the finite-horizon laws are consistent, S_t is defined for every $t \ge 0$. To verify the semi-group identity for given s, t, simply choose $T \ge s + t$. This defines the transition operators without constructing a probability measure on an infinite time interval. The generator of this semi-group is defined as follows, provided that the limit exists:

$$\mathcal{L}g(v) = \lim_{\epsilon \downarrow 0} \frac{\mathbb{E}_{\mathbb{P}^v} g(V_\epsilon^v) - g(v)}{\epsilon}.$$

Proposition 9.8. *For all $g \in C^1(\mathbb{R})$ with compact support. It holds that:*

$$\mathcal{L}g(v) = g'(v)b(v) + [g(0) - g(v)]f(v).$$

Proof. Let $M_t = N_t - \int_0^t f(V_{u-}^v)\mathrm{d}u$. At a spike, the change in $g(V^v)$ is $g(0) - g(V_{u-}^v)$. The change of variables formula therefore gives

$$g(V_t^v) = g(v) + \int_0^t g'(V_u^v)b(V_u^v)\mathrm{d}u + \int_0^t [g(0) - g(V_{u-}^v)]\mathrm{d}N_u.$$

Substituting $\mathrm{d}N_u = \mathrm{d}M_u + f(V_{u-}^v)\mathrm{d}u$, we obtain

$$\begin{aligned} g(V_t^v) = g(v) &+ \int_0^t \big[g'(V_u^v)b(V_u^v) + (g(0) - g(V_u^v))f(V_u^v)\big]\mathrm{d}u \\ &+ \int_0^t [g(0) - g(V_{u-}^v)]\mathrm{d}M_u. \end{aligned}$$

The last integrand is predictable and bounded, and f is bounded, so the last term is a $(\mathbb{P}^v, \mathcal{F}_t)$-martingale on $[0,T]$. Taking expectations, subtracting $g(v)$, dividing by t and letting $t \downarrow 0$ gives the stated generator: the drift integrand is bounded and continuous as a function of the potential, and $V_t^v \to v$ almost surely.

We used that V_u^v and V_{u-}^v differ only at spike times, which form a locally finite set. Thus they can be interchanged under an integral against $\mathrm{d}u$. This does not hold against $\mathrm{d}N_u$ or $\mathrm{d}M_u$, or in a product evaluated at spike times. In particular, using $g(0) - g(V_u^v)$ at a spike would give zero instead of the actual change in $g(V^v)$. □

Strong Markov property

Proposition 9.9. *Let $g : \mathbb{R}_+ \to \mathbb{R}_+$ be a non-negative measurable function. Let $t \in [0,T]$ and let τ be an $(\mathcal{F}_t)$ stopping time with values in $[0, T-t]$.*

$$\psi_t(v) := \mathbb{E}_{\mathbb{P}^v}[g(V_t^v)].$$

It holds that

$$\mathbb{E}_{\mathbb{P}^v}[g(V_{t+\tau}^v) \mid \mathcal{F}_\tau] = \psi_t(V_\tau^v) \quad \mathbb{P}^v\text{-a.s.}$$

Proof. This is the same proof as before, but this time we use that under $\mathbb{Q}$, $(N_{u+\tau} - N_\tau)_{u\geq 0}$ is a standard Poisson process independent of $\mathcal{F}_\tau$. This claim is admitted (but standard). □

Volterra integral equation

For $0 \leq t \leq T$, we let:

$$r^v(t) = \mathbb{E}_{\mathbb{P}^v}[f(V_t^v)] = S_t f(v),$$

be the spiking rate of the neuron.

Exercice 9.10. *Fix $0 \leq t < T$. Show that, as $\epsilon \downarrow 0$ with $t + \epsilon \leq T$:*

$$\mathbb{P}^v(N_{t+\epsilon} - N_t > 0) = \epsilon \mathbb{E}_{\mathbb{P}^v} f(V_t^v) + o(\epsilon).$$

Let τ_1 be the first jumping time of N_t. We let

$$K^v(t) = -\frac{d}{dt}\mathbb{P}^v(\tau_1 > t) = f(\varphi_t(v)) \exp\Big(-\int_0^t f(\varphi_u(v))\mathrm{d}u\Big).$$

Proposition 9.11. $(r^v(t))_{0\le t\le T}$ *solves the following Volterra integral equation on* $[0,T]$*:*

$$r^v(t) = K^v(t) + \int_0^t r^0(t-s)K^v(s)\mathrm{d}s.$$

Proof. First,

$$r^v(t) = \mathbb{E}_{\mathbb{P}^v}[f(V_t^v)] = \mathbb{E}_{\mathbb{P}^v}[f(V_t^v)\mathbb{1}_{\tau_1>t}] + \mathbb{E}_{\mathbb{P}^v}[f(V_t^v)\mathbb{1}_{\tau_1\le t}] =: A + B.$$

The first term is

$$A = K^v(t).$$

For the second term, set $\rho = \tau_1 \wedge t$, so that $\rho \le t \le T$. The strong Markov property at ρ gives the conditional rate $r^{V_\rho^v}(t-\rho)$ at time t: the remaining duration $t-\rho$ is known at time ρ. On $\{\tau_1 \le t\} \in \mathcal{F}_\rho$, we have $V_\rho^v = 0$; the rate after this reset is $r^0(u) = \mathbb{E}_{\mathbb{P}^0}[f(V_u^0)]$. Therefore

$$\begin{aligned} B &= \mathbb{E}_{\mathbb{P}^v}[\mathbb{1}_{\tau_1\le t}\mathbb{E}_{\mathbb{P}^v}[f(V_t^v) \mid \mathcal{F}_\rho]] \\ &= \mathbb{E}_{\mathbb{P}^v}[\mathbb{1}_{\tau_1\le t} r^0(t-\tau_1)]. \\ &= \int_0^t K^v(s) r^0(t-s)\mathrm{d}s. \end{aligned}$$

Altogether, the result is proved. □

Exercice 9.12 (Modelling two interacting neurons)**.** *We wish to construct a stochastic model for two neurons on* $[0,T]$*, with initial potentials* $(v_1,v_2)\in\mathbb{R}_+^2$ *and coupling strengths* $\alpha,\beta\ge 0$*. Denote their potentials by* V^1, V^2 *and their spike counts by* N^1, N^2*, with* $N_0^1 = N_0^2 = 0$*. The model should satisfy the following requirements.*

Between spikes, each potential follows the ODE $\dot V_t^i = b(V_t^i)$*. Neuron* i *should spike with predictable intensity* $f(V_{t-}^i)$ *relative to the joint history of both neurons. When neuron 1 spikes alone, its potential is reset to zero and the potential of neuron 2 increases by* β*. When neuron 2 spikes alone, its potential is reset to zero and the potential of neuron 1 increases by* α*.*

1. *Propose a construction of the spike counts and potentials using Poisson measures on* $\mathbb{R}_+\times\mathbb{R}_+$*. Specify the dependence between the driving measures and write the corresponding stochastic equations. Can simultaneous spikes occur in your model?*

2. *Describe how to to simulate trajectories "exactly" on* $[0,T]$*.*

3. *Give a weak construction of this model on* $[0,T]$ *using Girsanov's theorem.*

4. *Explain why* (V_t^1, V_t^2) *is a Markov process and compute its generator.*

5. *Let* $g:\mathbb{R}_+^2\to\mathbb{R}$ *be a smooth compactly supported test function and define* $u(t,v_1,v_2) = \mathbb{E}_{v_1,v_2}[g(V_t^1,V_t^2)]$*. Assuming that* u *is continuously differentiable in time and in the potentials, derive the PDE satisfied by* u*.*

10 Poisson measures

Let $(S,\mathcal{S})$ be a separable metric space; where $\mathcal{S}$ is the Borel σ-algebra. In the applications, we will use mostly $S=\mathbb{R}_+^2$, or $\mathbb{R}_+\times\mathbb{R}^d$.

Definition 10.1. *A point measure on S is measure of the form:*

$$\mu = \sum_{i \in I} \delta_{x_i},$$

where $x_i \in S$, and I is a finite or countable set. We denote by $M_p(S)$ the set of all punctual measures as above. We equip $M_p(S)$ with the σ-algebra $\mathcal{M}$ generated by the functions $\mu \mapsto \mu(A), A \in \mathcal{S}$ and $\mu \in M_p(S)$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space.

Definition 10.2. *Let ν be a σ-finite measure on $(S, \mathcal{S})$. A (random) Poisson measure Π on S with intensity ν is a random variable on $(M_p(S), \mathcal{M})$ such that:*

1. *For all $A \in \mathcal{S}$, $\Pi(A) \overset{(d)}{=} Poisson(\nu(A))$.*
2. *For all disjoints $A_1, \cdots, A_n$, the random variables $\Pi(A_1), \cdots, \Pi(A_n)$ are independent.*

The intensity measure gives the expected number of points in each measurable set A:

$$\mathbb{E}[\Pi(A)] = \nu(A).$$

For a set of small measure $\nu(A) = \epsilon$, the Poisson distribution gives

$$\mathbb{P}(\Pi(A) = 1) = \epsilon e^{-\epsilon} = \epsilon + O(\epsilon^2), \qquad \epsilon \downarrow 0.$$

The motivation for this definition is the following. Take $A \in \mathcal{S}$ and cover it by small disjoints balls of ν-volume ϵ. Then

$$\Pi(A) \approx \sum \text{independent bernoulli r.v.} \approx \sum_{i=1}^{\nu(A)/\epsilon} Ber(\epsilon) = Bin(\frac{\nu(A)}{\epsilon}, \epsilon) \to_{\epsilon \downarrow 0} \text{Poisson}(\nu(A)).$$

.

Theorem 10.3. *Let ν be a σ-finite measure on $(S, \mathcal{S})$. There exists a Poisson measure with intensity ν.*

To prove this result, we start with a lemma.

Lemma 10.4 (Rarefaction)**.** *Let X be a random variable with distribution $Poisson(\lambda)$. Let I be a discrete set. Let $\xi_1, \xi_2, \cdots$ a collection of i.i.d I-valued random variables, also independent of X. Denote by*

$$p_i = \mathbb{P}(\xi_1 = i), \quad i \in I.$$

Then the random variables

$$X_i = \sum_{j=1}^{X} \mathbb{1}_{\xi_j = i}$$

are independent with law $Poisson(p_i \lambda)$

Exercice 10.5. *Prove this lemma. Hint: compute the Fourier transform.*

We also have:

Lemma 10.6. *It holds that for all $\lambda, \lambda' \geq 0$*

$$Poisson(\lambda) * Poisson(\lambda') = Poisson(\lambda + \lambda').$$

That is if $X \sim Poisson(\lambda)$ and $Y \sim Poisson(\lambda')$ are independent then $X + Y \sim Poisson(\lambda + \lambda')$.

Exercice 10.7. *Prove this result.*

Proof of Theorem 10.3. We first assume that $\nu(S) < \infty$. Let $\bar{\nu} = \frac{\nu}{\nu(S)}$. Let N a random variable with law $Poisson(\nu(S))$. Let (X_j) be a collection of i.i.d. random variable (also independent of N) of law $\bar{\nu}$. Set:

$$\Pi = \sum_{j=1}^{N} \delta_{X_j}.$$

Let $A_1, \cdots, A_{n-1} \in \mathcal{S}$ be disjoint sets. Let $A_n = S \smallsetminus \cup_i A_i$ such that $A_1, \cdots A_n$ is a partition of S. We also set:

$$\xi_j = \sum_{i=1}^{n} i \mathbb{1}_{\{X_j \in A_i\}}.$$

We have:

$$\Pi(A_i) = \sum_{j=1}^{N} \mathbb{1}_{A_i}(X_j) = \sum_{j=1}^{N} \mathbb{1}_{\xi_j = i}.$$

By the rarefaction lemma, we deduce that the r.v. $\Pi(A_i)$ are independent and

$$\Pi(A_i) \sim Poisson(\nu(S) \frac{\nu(A_i)}{\nu(S)}).$$

Therefore, Π is a Poisson measure of intensity ν.

Assume now that $\nu(S) = \infty$, and take S_n a partition of S with $\nu(S_n) < \infty$. Consider Π_n a family of independent Poisson measures on S_n, with intensity $\nu(\cdot \cap S_n)$. Set:

$$\Pi = \sum_{n \geq 1} \Pi_n.$$

It holds that Π is a Poisson measure with intensity ν. Indeed:

1. almost surely, $\Pi_n(S \smallsetminus S_n) = 0$, because it has law $Poisson(0)$.
2. Take $A_1, \cdots A_k$ two by two disjoint, $A_i \in \mathcal{S}$. Then $\Pi_n(A_i), n$ are independent; of law $Poisson(\nu(A_i \cap S_n))$. Therefore:

$$\Pi(A_i) = \sum_{n \geq 0} \Pi_n(A_i) \sim Poisson(\nu(A_i)).$$

3. As all the random variables $\Pi_n(A_i), n, i$ are independent, we deduce that the random variables

$$\Pi(A_i) = \sum_n \Pi_n(A_i), \quad i \in 1 \cdots k$$

are independent.

This end the proof. □

Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be a filtration on $(\Omega, \mathcal{F})$.

Definition 10.8. *Let ν be a σ-finite measure on $(S, \mathcal{S})$. We say that $\Pi(\mathrm{d}s, \mathrm{d}z)$ is an $(\mathcal{F}_t)$ adapted Poisson measure of intensity $\mathrm{d}s\nu(\mathrm{d}z)$ if Π is a Poisson measure on $\mathbb{R}_+ \times S$ with intensity $\mathrm{d}s\nu(\mathrm{d}z)$ and if for all $t, s \geq 0$:*

- *$\Pi([0,t] \times A)$ is $\mathcal{F}_t$ measurable for all $A \in \mathcal{S}$.*
- *$\Pi((t, t+s] \times A)$ is independent of $\mathcal{F}_t$ for all $A \in \mathcal{S}$.*

The restriction of Π to $[0,t] \times S$ is itself a Poisson measure, and can be written of the form

$$\Pi_{|[0,t]} = \sum_n \delta_{(s_n, z_n)}.$$

Recall that $\mathcal{P}(\mathbb{F})$ denotes the predictable σ-field on $(0,\infty) \times \Omega$ associated to $\mathbb{F} = (\mathcal{F}_t)$. Let $H : \mathbb{R}_+ \times \Omega \times S \to [0, \infty]$ be $\mathcal{P}(\mathbb{F}) \times \mathcal{S}$ measurable function. We denote by

$$\int_0^t \int_S H_s(\omega, z)\Pi(\mathrm{d}s, \mathrm{d}z) := \sum_n H_{s_n}(\omega, z_n),$$

provided that $\Pi_{|[0,t]} = \sum_n \delta_{(s_n, z_n)}$. We have (proof is admitted).

Theorem 10.9 (Master formula)**.** *It holds that:*

1. *$\int_0^t \int_S H_s(\omega, z)\Pi(\mathrm{d}s, \mathrm{d}z)$ and $\int_0^t \int_S \nu(\mathrm{d}z) H_s(\omega, z)\mathrm{d}s$ are both $\mathcal{F}_t$-random variable and have the same expectation.*
2. *If moreover $\mathbb{E}\int_0^t \int_S \nu(\mathrm{d}z) H_s(\omega, z)\mathrm{d}s < \infty$ for all t, then*

$$M_t = \int_0^t \int_S H_s(\omega, z)\Pi(\mathrm{d}s, \mathrm{d}z) - \int_0^t \int_S \nu(\mathrm{d}z) H_s(\omega, z)\mathrm{d}s,$$

is a $(\mathbb{P}, \mathcal{F}_t)$ martingale.

In most of the applications, we choose:

$$H_s(\omega, z) = X_s(\omega)\mathbb{1}_{\{z \leq Y_s(\omega)\}},$$

where $(X_s), (Y_s)$ are predictable processes with respect to $(\mathcal{F}_t)$. This provides a direct and efficient way to construct sophisticated point processes with stochastic intensity.

Application to the model of spiking neurons

Fix $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ a filtered probability space, and Π a $(\mathcal{F}_t)$ adapted Poisson measure of intensity $\mathrm{d}s\mathrm{d}z$ on $\mathbb{R}_+^2$. Consider the following equation

$$V_t = v + \int_0^t b(V_s)\mathrm{d}s - \int_0^t \int_{\mathbb{R}_+} V_{s-}\mathbb{1}_{\{z \leq f(V_{s-})\}}\Pi(\mathrm{d}s, \mathrm{d}z).$$

This equation can be solved jump after jump for all $t \geq 0$: since f is bounded, only finitely many points of Π can produce a spike on each bounded time interval. In addition, let

$$N_t = \int_0^t \int_{\mathbb{R}_+} \mathbb{1}_{\{z \leq f(V_{s-})\}}\Pi(\mathrm{d}s, \mathrm{d}z).$$

Then N_t is a point processes of stochastic intensity $f(V_{t-})$. Indeed,

$$N_t - \int_0^t f(V_{s-})\mathrm{d}s$$

is a $(\mathbb{P}, \mathcal{F}_t)$-martingale. This construction defines the neuron for all $t \geq 0$. For initial potential v, its law on each $[0,T]$ is the law of V^v under $\mathbb{P}^v$ in the Girsanov construction: both follow the same deterministic flow between spikes and have the same conditional law of the next spike time. The finite-horizon Markov properties therefore apply to this process; the strong Markov property at an almost surely finite stopping time follows by localization. We use this construction for the long-time study below.

11 Long time behavior via coupling arguments

The previous section provides a representation of stochastic processes using an integral with respect to a Poisson measure. This representation is very useful to do coupling arguments, where two processes are driven by the same Poisson measure. We illustrate this coupling argument by studying the long time behavior of (V_t). Let V_t^x be the solution of:

$$V_t^x = x + \int_0^t b(V_s^x)\mathrm{d}s - \int_0^t \int_{\mathbb{R}_+} V_{s-}^x \mathbb{1}_{\{z \leq f(V_{s-}^x)\}} \Pi(\mathrm{d}s, \mathrm{d}z).$$

Given $x, y \in \mathbb{R}_+$ two initial conditions, we denote by

$$\tau_c = \inf\{t \geq 0 : V_s^x = V_s^y \ \ \forall s \geq t\},$$

with the convention that $\inf = \infty$. Importantly here, V^x and V^y are driven by the same Poisson measure. We denote by $\varphi_t(x)$ the solution of the ODE:

$$\frac{d}{dt}\varphi_t(x) = b(\varphi_t(x)); \quad \varphi_0(x) = x.$$

We assume that f is bounded by $f_{\max}$ and there exists $T, \lambda > 0$ such that:

$$\forall x, y \in \mathbb{R}_+, \quad \int_0^T f(\varphi_s(x)) \wedge f(\varphi_s(y))\mathrm{d}s \geq \lambda.$$

Proposition 11.1. *Set $\mu = e^{-T f_{\max}}\lambda \in (0,1)$. For all initial potentials $x, y \in \mathbb{R}_+$,*

$$\mathbb{P}(\tau_c > kT) \leq (1-\mu)^k, \qquad k \in \mathbb{N}.$$

In particular,

$$\mathbb{E}e^{\eta\tau_c} < \infty \qquad \textit{for } 0 < \eta < -\frac{1}{T}\log(1-\mu).$$

Proof. Once the two potentials coincide, pathwise uniqueness and the common Poisson measure ensure that they stay together. Thus

$$\{\tau_c > kT\} = \{V_{kT}^x \neq V_{kT}^y\} \in \mathcal{F}_{kT}.$$

Fix $k \in \mathbb{N}$. Conditional on $\mathcal{F}_{kT}$, define

$$q_s^{(k)} := f(\varphi_s(V_{kT}^x)) \wedge f(\varphi_s(V_{kT}^y)), \qquad s \geq 0.$$

Let $(kT + S_k, Z_k)$ be the first point of Π after kT in the strip of heights $[0, f_{\max}]$. Consider the event

$$E_k := \{S_k \le T,\ Z_k \le q^{(k)}_{S_k}\}.$$

Before this first point, neither neuron has spiked, so both follow their deterministic flows. On E_k, this point triggers a spike in both neurons and resets both potentials to zero. Hence E_k is sufficient for coupling by $(k+1)T$; coupling may also occur after earlier points of the strip have been rejected.

The Poisson measure after kT is independent of $\mathcal{F}_{kT}$. Conditional on this past, S_k has exponential distribution with rate $f_{\max}$, and Z_k is independent of S_k and uniform on $[0, f_{\max}]$. Therefore

$$\begin{aligned}\mathbb{P}(E_k \mid \mathcal{F}_{kT}) &= \int_0^T f_{\max} e^{-f_{\max} s} \frac{q^{(k)}_s}{f_{\max}} \mathrm{d}s \\ &\ge e^{-T f_{\max}} \int_0^T q^{(k)}_s \mathrm{d}s \ge e^{-T f_{\max}} \lambda = \mu.\end{aligned}$$

Here $0 < \mu < 1$, since $0 < \lambda \le T f_{\max}$ and $T f_{\max} e^{-T f_{\max}} \le e^{-1}$.

On $\{\tau_c > kT\}$, coupling fails throughout the next block only if E_k fails. Consequently,

$$\begin{aligned}\mathbb{P}(\tau_c > (k+1)T \mid \mathcal{F}_{kT}) &\le \mathbb{1}_{\{\tau_c > kT\}} \mathbb{P}(E_k^c \mid \mathcal{F}_{kT}) \\ &\le (1-\mu) \mathbb{1}_{\{\tau_c > kT\}}.\end{aligned}$$

Taking expectations and iterating gives

$$\mathbb{P}(\tau_c > (k+1)T) \le (1-\mu)\mathbb{P}(\tau_c > kT), \qquad \mathbb{P}(\tau_c > kT) \le (1-\mu)^k.$$

In particular, $\tau_c < \infty$ almost surely. For $\eta > 0$, Tonelli's theorem and the bound on each interval $[kT, (k+1)T)$ give

$$\begin{aligned}\mathbb{E} e^{\eta \tau_c} &= 1 + \eta \int_0^\infty e^{\eta t} \mathbb{P}(\tau_c > t) \mathrm{d}t \\ &\le 1 + (e^{\eta T} - 1) \sum_{k=0}^\infty \left[e^{\eta T}(1-\mu)\right]^k.\end{aligned}$$

The series converges when $e^{\eta T}(1-\mu) < 1$, which is precisely the stated condition on η. □

As a corollary, we have:

Proposition 11.2. *There is a constant $c > 0$ such that:*

$$\forall t \ge 0, \quad d_{TV}(\mathcal{L}(V^x_t), \mathcal{L}(V^y_t)) \le C e^{-ct}.$$

Proof. We recall that the total variaton distance satisfies:

$$d_{TV}(\mu, \nu) = \sup_g \int g(x)(\nu - \mu)(dx),$$

where the supremum is taken over measurable functions g with $|g(x)| \le 1$. Let g be such a function. Let $\mu_t^x = \mathcal{L}(V_t^x), \mu_t^y = \mathcal{L}(V_t^y)$. We have:

$$\begin{aligned}\int g(x)(\mu_t^x - \mu_t^y)(\mathrm{d}x) &= \mathbb{E}[g(V_t^x) - g(V_t^y)] \\ &\le 2\mathbb{P}(V_t^x \ne V_t^y) \le 2\mathbb{P}(\tau_c > t) \\ &= 2\mathbb{P}(e^{\eta\tau_c} > e^{\eta t}) \le 2e^{-\eta t}\mathbb{E}e^{\eta\tau_c}.\end{aligned}$$

This gives the result. □

Proposition 11.3. *Assume in addition that $b(0) > 0$. The process (V_t) has a unique invariant distribution which is given by*

$$\nu^\infty(\mathrm{d}x) = \frac{\gamma}{b(x)} \exp\left(-\int_0^x \frac{f(y)}{b(y)}\mathrm{d}y\right) \mathbb{1}_{[0,\sigma)}(x)\mathrm{d}x,$$

where γ is a normalizing factor (so that $\int \nu^\infty = 1$) and $\sigma = \lim_{t\to\infty} \varphi_t(0)$. The normalizing factor satisfies:

$$\gamma = \int_{\mathbb{R}_+} f(x)\nu^\infty(\mathrm{d}x).$$

Proof. We construct the invariant distribution from the occupation of a spike cycle. Write $\mathbb{P}_0$ and $\mathbb{E}_0$ for probability and expectation when $V_0 = 0$, and let $\tau_1 = \inf\{t > 0 : N_t \ge 1\}$ be the first spike time. Before this spike, $V_a = \varphi_a(0)$. Thus

$$H(a) := \mathbb{P}_0(\tau_1 > a) = \exp\left(-\int_0^a f(\varphi_s(0))\mathrm{d}s\right), \qquad m := \mathbb{E}_0\tau_1 = \int_0^\infty H(a)\mathrm{d}a.$$

The uniform assumption of this section, applied with $x = y = \varphi_{kT}(0)$, gives $H(kT) \le e^{-k\lambda}$. Consequently, τ_1 is almost surely finite and $0 < m \le T/(1 - e^{-\lambda}) < \infty$.

Each spike resets the potential to zero. By the strong Markov property, successive cycles between spikes are independent and have the same law. This suggests defining a candidate invariant probability of a Borel set $B \subset \mathbb{R}_+$ as the expected time spent in B during one cycle, divided by the expected duration of a cycle:

$$\nu^\infty(B) = \frac{1}{m}\mathbb{E}_0 \int_0^{\tau_1} \mathbb{1}_B(V_a)\mathrm{d}a = \frac{1}{m}\int_0^\infty \mathbb{1}_B(\varphi_a(0))H(a)\mathrm{d}a. \tag{11.1}$$

The second equality follows from Tonelli's theorem: the cycle still lasts at age a with probability $H(a)$. In particular, this defines a probability measure because its total mass is $m/m = 1$.

We now verify invariance, keeping $t \ge 0$ fixed. Recall that the transition semi-group $(S_t)_{t\ge0}$ acts on a bounded measurable function g by

$$S_t g(v) := \mathbb{E}_v[g(V_t)],$$

where $\mathbb{E}_v$ denotes expectation for the neuron started at $V_0 = v$. Thus $S_t g$ is a deterministic function of the starting potential: it gives the expected value of g after a time interval of length t. Write $\nu^\infty(h) = \int h\mathrm{d}\nu^\infty$ for a bounded measurable function h. To prove invariance, we must show that $\nu^\infty(S_t g) = \nu^\infty(g)$ for every such g and every $t \ge 0$.

First, for each deterministic $s \ge 0$, the Markov property gives

$$S_t g(V_s) = \mathbb{E}_0[g(V_{s+t}) \mid \mathcal{F}_s] \quad \text{a.s.}$$

Since τ_1 is a stopping time, the event $\{s < \tau_1\}$ belongs to $\mathcal{F}_s$. Multiplying by its indicator and taking expectations therefore yields

$$\begin{aligned}\mathbb{E}_0[\mathbb{1}_{\{s<\tau_1\}} S_t g(V_s)] &= \mathbb{E}_0[\mathbb{1}_{\{s<\tau_1\}} \mathbb{E}_0[g(V_{s+t}) \mid \mathcal{F}_s]] \\ &= \mathbb{E}_0[\mathbb{1}_{\{s<\tau_1\}} g(V_{s+t})].\end{aligned}$$

We can integrate this equality over $s \in [0,\infty)$. Fubini's theorem applies because g and $S_t g$ are bounded by $\|g\|_\infty$ in absolute value, and $\mathbb{E}_0\tau_1 = m < \infty$. Applying the occupation formula to $h = S_t g$, we obtain

$$\begin{aligned}m\nu^\infty(S_t g) &= \mathbb{E}_0 \int_0^{\tau_1} S_t g(V_s)\mathrm{d}s \\ &= \int_0^\infty \mathbb{E}_0[\mathbb{1}_{\{s<\tau_1\}} S_t g(V_s)]\mathrm{d}s \\ &= \int_0^\infty \mathbb{E}_0[\mathbb{1}_{\{s<\tau_1\}} g(V_{s+t})]\mathrm{d}s = \mathbb{E}_0 \int_0^{\tau_1} g(V_{s+t})\mathrm{d}s.\end{aligned}$$

Next, on each trajectory, the change of variables $u = s + t$ and the additivity of integrals give

$$\begin{aligned}\int_0^{\tau_1} g(V_{s+t})\mathrm{d}s &= \int_t^{\tau_1+t} g(V_u)\mathrm{d}u \\ &= \int_0^{\tau_1+t} g(V_u)\mathrm{d}u - \int_0^t g(V_u)\mathrm{d}u \\ &= \int_0^{\tau_1} g(V_u)\mathrm{d}u + \int_{\tau_1}^{\tau_1+t} g(V_u)\mathrm{d}u - \int_0^t g(V_u)\mathrm{d}u.\end{aligned}$$

This identity holds whether $\tau_1 < t$ or $\tau_1 \geq t$: we have simply split the integral over $[0, \tau_1 + t]$ at τ_1. Taking expectations and using the occupation formula with $h = g$ for the first term gives

$$m\nu^\infty(S_t g) = m\nu^\infty(g) + \mathbb{E}_0 \int_{\tau_1}^{\tau_1+t} g(V_u)\mathrm{d}u - \mathbb{E}_0 \int_0^t g(V_u)\mathrm{d}u.$$

Finally, at τ_1 the potential resets to zero. By the strong Markov property, its evolution during the next t units of time has the same law as a fresh trajectory started at zero. More explicitly, conditional Fubini and the strong Markov property give

$$\begin{aligned}\mathbb{E}_0\left[\int_{\tau_1}^{\tau_1+t} g(V_u)\mathrm{d}u \,\middle|\, \mathcal{F}_{\tau_1}\right] &= \int_0^t \mathbb{E}_0[g(V_{\tau_1+r}) \mid \mathcal{F}_{\tau_1}]\mathrm{d}r \\ &= \int_0^t S_r g(V_{\tau_1})\mathrm{d}r = \int_0^t S_r g(0)\mathrm{d}r \\ &= \mathbb{E}_0 \int_0^t g(V_r)\mathrm{d}r.\end{aligned}$$

Taking expectations shows that the last two terms in the preceding expression for $m\nu^\infty(S_t g)$ cancel. Hence $\nu^\infty(S_t g) = \nu^\infty(g)$, since $m > 0$. This proves invariance. Uniqueness follows from the preceding coupling estimate.

Since $b(0) > 0$, the map $a \mapsto \varphi_a(0)$ is strictly increasing from 0 towards σ, and $b(x) > 0$ for $0 \leq x < \sigma$. The change of variables $x = \varphi_a(0)$ gives

$$\mathrm{d}a = \frac{\mathrm{d}x}{b(x)}, \qquad H(a) = \exp\left(-\int_0^x \frac{f(y)}{b(y)}\mathrm{d}y\right).$$

This recovers the density in the proposition, with $\gamma = 1/m$. Moreover, $H'(a) = -f(\varphi_a(0))H(a)$ and $H(\infty) = 0$, so

$$\int_{\mathbb{R}_+} f(x)\nu^\infty(\mathrm{d}x) = \frac{1}{m}\int_0^\infty f(\varphi_a(0))H(a)\mathrm{d}a = \frac{1}{m} = \gamma.$$

□

Remark 11.4. *If instead $b(0) = 0$, uniqueness of the ODE gives $\varphi_a(0) = 0$. Under the assumptions of this section, formula* (11.1) *remains valid and gives $\nu^\infty = \delta_0$, rather than a density.*

12 Hawkes process

Let (N_t) be a point process with $N_0 = 0$ and predictable stochastic intensity

$$\lambda_t = \mu + \int_{(0,t)} \phi(t-s)\mathrm{d}N_s, \qquad t > 0, \qquad \lambda_0 = \mu,$$

where $\mu \geq 0$ and ϕ is a non-negative continuous kernel. We assume that there are no events before time zero. The open endpoint at t excludes the current spike: the intensity uses only the spikes strictly before t. Consider the exponential kernel

$$\phi(s) = \alpha e^{-\beta s}, \qquad s \geq 0, \qquad \alpha \geq 0, \quad \beta > 0.$$

Define the cadlag state

$$Y_t := \mu + \int_{(0,t]} \alpha e^{-\beta(t-s)}\mathrm{d}N_s.$$

Then $Y_0 = \mu$ and $\lambda_t = Y_{t-}$ for $t > 0$, while

$$\mathrm{d}Y_t = -\beta(Y_t - \mu)\mathrm{d}t + \alpha \mathrm{d}N_t.$$

Between spikes, Y_t decays towards μ; at each spike, it increases by α. Thus Y_t includes the current spike, whereas λ_t is its value just before that spike.

Exercice 12.1. *Show that $(Y_t, N_t)_{t\geq 0}$ is a Markov process and that its generator is*

$$\mathcal{A}g(y,n) = -\beta(y-\mu)\partial_y g(y,n) + y[g(y+\alpha, n+1) - g(y,n)],$$

for compactly supported test functions $g : \mathbb{R}_+ \times \mathbb{N} \to \mathbb{R}$ that are continuously differentiable in y. Hint: identify the deterministic flow, the jump rate and the state immediately after a spike.

13 Filtering

Fix a deterministic observation horizon $T > 0$. Throughout this section, all observation times satisfy $0 \leq t \leq T$. We start with the following theoretical result. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{Q})$ be a probability space. Let $\mathbb{P}$ be a probability measure on $\mathcal{F}_T$, and denote by $\mathbb{P}_t, \mathbb{Q}_t$ the restrictions of $\mathbb{P}$ and $\mathbb{Q}$ to $\mathcal{F}_t$, for $0 \leq t \leq T$. We assume that $\mathbb{P}_t \ll \mathbb{Q}_t$ and denote by

$$L_t = \frac{\mathrm{d}\mathbb{P}_t}{\mathrm{d}\mathbb{Q}_t}.$$

Let Z_t be a real valued bounded process adapted to $\mathcal{F}_t$. Let $(\mathcal{G}_t)$ be a filtration such that $\mathcal{G}_t \subset \mathcal{F}_t$.

Proposition 13.1. *For $0 \le t \le T$, it holds that $\mathbb{P}(\mathrm{d}\omega)$ a.s.:*

$$\mathbb{E}_{\mathbb{P}}[Z_t \mid \mathcal{G}_t] = \frac{\mathbb{E}_{\mathbb{Q}}[Z_t L_t \mid \mathcal{G}_t]}{\mathbb{E}_{\mathbb{Q}}[L_t \mid \mathcal{G}_t]}.$$

Proof. Take $A \in \mathcal{G}_t$. We have by definition of condition expectation:

$$\mathbb{E}_{\mathbb{P}}[\mathbb{1}_A \mathbb{E}_{\mathbb{P}}[Z_t \mid \mathcal{G}_t]] = \mathbb{E}_{\mathbb{P}}[\mathbb{1}_A Z_t].$$

As $\mathrm{d}\mathbb{P}_t = L_t \mathrm{d}\mathbb{Q}_t$, we deduce that

$$\mathbb{E}_{\mathbb{Q}}[\mathbb{1}_A L_t \mathbb{E}_{\mathbb{P}}[Z_t \mid \mathcal{G}_t]] = \mathbb{E}_{\mathbb{Q}}[\mathbb{1}_A Z_t L_t].$$

Taking the conditional expectation with respect to $\mathcal{G}_t$, we deduce that:

$$\mathbb{E}_{\mathbb{Q}}[\mathbb{1}_A \mathbb{E}_{\mathbb{Q}}[L_t \mid \mathcal{G}_t] \mathbb{E}_{\mathbb{P}}[Z_t \mid \mathcal{G}_t]] = \mathbb{E}_{\mathbb{Q}}[\mathbb{1}_A \mathbb{E}_{\mathbb{Q}}[Z_t L_t \mid \mathcal{G}_t]].$$

As this is true for all A:

$$\mathbb{E}_{\mathbb{Q}}[L_t \mid \mathcal{G}_t] \mathbb{E}_{\mathbb{P}}[Z_t \mid \mathcal{G}_t] = \mathbb{E}_{\mathbb{Q}}[Z_t L_t \mid \mathcal{G}_t].$$

Let now C_t the set:

$$C_t = \{\omega \in \Omega : \mathbb{E}_{\mathbb{Q}}[L_t \mid \mathcal{G}_t] = 0\}.$$

Then $C_t \in \mathcal{G}_t$ and so:

$$\mathbb{P}(C_t) = \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_{C_t} L_t) = \mathbb{E}_{\mathbb{Q}}[\mathbb{1}_{C_t} \mathbb{E}_{\mathbb{Q}}[L_t \mid \mathcal{G}_t]] = 0.$$

□

We apply this formula to the estimation of a hidden parameter. Let Θ be a Borel subset of $\mathbb{R}^d$, equipped with its Borel σ-field $\mathcal{B}(\Theta)$. We start with a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{Q})$ such that:

1. N is a standard Poisson process of rate 1 relative to $(\mathcal{F}_t)$: its future increments after t are independent of $\mathcal{F}_t$.
2. Λ is a Θ-valued, $\mathcal{F}_0$-measurable random variable of law m_0, independent of the whole process N under $\mathbb{Q}$.
3. The observation filtration is $\mathcal{G}_t \coloneqq \mathcal{F}_t^N = \sigma(N_u, 0 \le u \le t)$, and we write $\mathbb{F}^N = (\mathcal{F}_t^N)$.

We assume that the non-negative family

$$(t, \omega, \lambda) \longmapsto \mu_t(\lambda, \omega)$$

is measurable for $\mathcal{P}(\mathbb{F}^N) \otimes \mathcal{B}(\Theta)$ on $(0, T] \times \Omega \times \Theta$. Thus, for each fixed parameter λ, the intensity is a predictable function of the observed spike history. Any dependence on the unknown parameter is through the argument λ; the random variable Λ is substituted only afterwards. For every $\lambda \in \Theta$, assume that $\int_0^T \mu_s(\lambda)\mathrm{d}s < \infty$ $\mathbb{Q}$-almost surely. Define

$$L_t(\lambda) = \prod_{\tau_i \le t} \mu_{\tau_i}(\lambda) \exp\left(t - \int_0^t \mu_s(\lambda)\mathrm{d}s\right),$$

where τ_i are the spike times of N, and assume that $(L_t(\lambda))_{0 \le t \le T}$ is a true $(\mathbb{Q}, \mathcal{F}_t)$-martingale for every λ. This holds, for example, if each process $\mu(\lambda)$ is bounded by a deterministic constant $C(\lambda) < \infty$; a bound uniform in λ is not required. The joint measurability assumption ensures that, for each t, $(\omega, \lambda) \mapsto L_t(\lambda, \omega)$ is $\mathcal{F}_t^N \otimes \mathcal{B}(\Theta)$-measurable.

Lemma 13.2. *The process $(L_t(\Lambda))_{0\le t\le T}$ is a $(\mathbb{Q},\mathcal{F}_t)$-martingale.*

Proof. Since Λ is $\mathcal{F}_0$-measurable, the joint predictability assumption implies that $\mu_t(\Lambda)$ is $(\mathcal{F}_t)$-predictable. Independence of Λ and N gives

$$\mathbb{Q}\left(\int_0^T \mu_s(\Lambda)\mathrm{d}s = \infty\right) = \int_\Theta \mathbb{Q}\left(\int_0^T \mu_s(\lambda)\mathrm{d}s = \infty\right) m_0(\mathrm{d}\lambda) = 0.$$

The density process $L(\Lambda)$ is therefore a non-negative local martingale by the result preceding Girsanov's theorem. Moreover, for each $t \le T$, independence and Tonelli's theorem yield

$$\mathbb{E}_{\mathbb{Q}}[L_t(\Lambda)] = \int_\Theta \mathbb{E}_{\mathbb{Q}}[L_t(\lambda)] m_0(\mathrm{d}\lambda) = 1.$$

A non-negative local martingale is a supermartingale. Since its expectation is constant, the conditional supermartingale inequality must be an equality, proving that $L(\Lambda)$ is a true martingale. □

We then define $\mathbb{P}$ on $\mathcal{F}_T$ by:

$$\mathrm{d}\mathbb{P} = L_T(\Lambda)\mathrm{d}\mathbb{Q}.$$

Under $\mathbb{P}$, $(N_t)_{0\le t\le T}$ is a point process of stochastic intensity $\mu_t(\Lambda)$ on $[0,T]$. The prior law of Λ remains m_0, since $\mathbb{E}_{\mathbb{Q}}[L_T(\Lambda) \mid \Lambda] = 1$ by the same independence argument.

For a fixed $t \le T$, conditioning on $\mathcal{F}_t^N$ fixes the observed spike history. Under $\mathbb{Q}$, the independent parameter Λ still has law m_0. Joint measurability therefore gives, for every bounded measurable $\phi : \Theta \to \mathbb{R}$,

$$\mathbb{E}_{\mathbb{Q}}[\phi(\Lambda)L_t(\Lambda) \mid \mathcal{F}_t^N] = \int_\Theta \phi(\lambda)L_t(\lambda)m_0(\mathrm{d}\lambda).$$

Applying the Bayes formula above, with this identity for ϕ and for the constant function 1, yields

$$\mathbb{E}_{\mathbb{P}}[\phi(\Lambda) \mid \mathcal{F}_t^N] = \frac{\int_\Theta \phi(\lambda)L_t(\lambda)m_0(\mathrm{d}\lambda)}{\int_\Theta L_t(\lambda)m_0(\mathrm{d}\lambda)}, \qquad \mathbb{P}\text{-a.s.}$$

The denominator is finite and strictly positive $\mathbb{P}$-almost surely, as in the proof of the Bayes formula.

Application: estimation of the rate of a Poisson process

Let Λ a random variable of law m_0. This is the prior distribution. Assume that, conditionally on Λ, $(N_t)_{0\le t\le T}$ is a Poisson process of rate Λ on $[0,T]$. We observe $\mathcal{F}_t^N$. We want to compute the posterior distribution $M_t = \mathcal{L}(\Lambda \mid \mathcal{F}_t^N)$. By the result above; it holds that:

$$\begin{aligned} M_t(\mathrm{d}\lambda) &\propto L_t(\lambda)m_0(\mathrm{d}\lambda) \\ &\propto \lambda^{N_t}\exp(-\lambda t)m_0(\mathrm{d}\lambda). \end{aligned}$$

Example 13.3. *Assume that $m_0 = \Gamma(\alpha,\beta)$, that is:*

$$m_0(\mathrm{d}\lambda) = \frac{\lambda^{\alpha-1}e^{-\beta\lambda}\mathrm{d}\lambda}{Z_{\alpha,\beta}}.$$

Then

$$M_t = \Gamma(\alpha + N_t, \beta + t).$$

Therefore, the mean of M_t is $(\alpha + N_t)/(\beta + t)$. And the variance is $(\alpha + N_t)/(\beta + t)^2 = \frac{\text{Mean of } M_t}{\beta + t}$.

Application: the disruption problem

Let τ be a random variable of law $\mathbb{E}(\theta)$, $\theta > 0$. Let $0 < a < b$. Let $(N_t)_{0\le t\le T}$ be a point process with the following stochastic intensity on $[0,T]$:

$$a + (b-a)\mathbb{1}_{t\geq\tau}.$$

We assume we only have access to $\mathcal{F}_t^N$. We want to estimate τ, given the observations $\mathcal{F}_t^N$. In other word, we want to compute

$$\Lambda_t = \mathbb{P}(\tau > t \mid \mathcal{F}_t^N).$$

Let $M_t = \mathcal{L}(\tau \mid \mathcal{F}_t^N)$. It holds that

$$M_t(\mathrm{d}\lambda) \propto \left[a^{N_t}e^{-at}\mathbb{1}_{\{\lambda\geq t\}} + a^{N_\lambda}b^{N_t-N_\lambda}e^{-a\lambda}e^{-b(t-\lambda)}\mathbb{1}_{\{\lambda<t\}}\right]\theta e^{-\theta\lambda}\mathrm{d}\lambda.$$

Therefore, we find that:

$$\begin{aligned}
\Lambda_t &= \frac{\int_t^\infty a^{N_t}e^{-at}\theta e^{-\theta\lambda}\mathrm{d}\lambda}{\int_t^\infty a^{N_t}e^{-at}\theta e^{-\theta\lambda}\mathrm{d}\lambda + \int_0^t a^{N_\lambda}b^{N_t-N_\lambda}e^{-a\lambda}e^{-b(t-\lambda)}\theta e^{-\theta\lambda}\mathrm{d}\lambda}\\
&= \frac{a^{N_t}e^{-at}e^{-\theta t}}{a^{N_t}e^{-at}e^{-\theta t} + \int_0^t a^{N_\lambda}b^{N_t-N_\lambda}e^{-a\lambda}e^{-b(t-\lambda)}\theta e^{-\theta\lambda}\mathrm{d}\lambda}\\
&= \frac{1}{1 + \left(\frac{b}{a}\right)^{N_t} e^{-(b-a)t}e^{\theta t}\int_0^t \left(\frac{a}{b}\right)^{N_\lambda} e^{(b-a)\lambda}\theta e^{-\theta\lambda}\mathrm{d}\lambda}.
\end{aligned}$$

We note that between two successive jumps, Λ solves an ODE. Indeed,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Lambda_t &= -\Lambda_t^2\left(\frac{b}{a}\right)^{N_t} e^{-(b-a)t}e^{\theta t}\left[(\theta-(b-a))\int_0^t\left(\frac{a}{b}\right)^{N_\lambda} e^{(b-a)\lambda}\theta e^{-\theta\lambda}\mathrm{d}\lambda + \left(\frac{a}{b}\right)^{N_t} e^{(b-a)t}\theta e^{-\theta t}\right]\\
&= -\Lambda_t^2\left[\left(\frac{1}{\Lambda_t}-1\right)(\theta-(b-a))+\theta\right]\\
&= -\Lambda_t(\theta-(b-a)) - \Lambda_t^2(b-a)\\
&= -\theta\Lambda_t + \Lambda_t(1-\Lambda_t)(b-a).
\end{aligned}$$

Let $Z_t = 1-\Lambda_t = \mathbb{P}(t\geq\tau \mid \mathcal{F}_t^N)$. Between the jumps, it therefore solves:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}Z_t = +\theta(1-Z_t) - Z_t(1-Z_t)(b-a)}.$$

At a spike time t, we denote the values before the spike by Λ_{t-}, Z_{t-} and the values after the spike by Λ_t, Z_t (cadlag convention). We have

$$\Lambda_t = \frac{1}{1+\frac{b}{a}(\frac{1}{\Lambda_{t-}}-1)} = \frac{a\Lambda_{t-}}{a\Lambda_{t-}+b(1-\Lambda_{t-})}.$$

Therefore, $Z_t = 1-\Lambda_t$ satisfies:

$$Z_t = \frac{bZ_{t-}}{a+(b-a)Z_{t-}}.$$

In other words,

$$\boxed{Z_t - Z_{t-} = \frac{(b-a)Z_{t-}(1-Z_{t-})}{a+(b-a)Z_{t-}}.}$$

Application: estimation of a parameter

Exercice 13.4. *Let Λ be a random variable of distribution m_0. Let (V_t) be the membrane potential of a neuron, whose dynamics is given by*

$$V_t = \int_0^t b(V_s)\mathrm{d}s - \int_0^t V_{s-}\mathrm{d}N_s,$$

and (N_t) is point process of stochastic intensity $\Lambda f(V_{t-})$. Assume we only observe the spiking trains. What is

$$\mathcal{L}(\Lambda \mid \mathcal{F}_t^N)?$$

14 Doeblin's theorem

We use the same total variation convention as in Section 11, with values in $[0,2]$:

$$d_{TV}(\nu,\mu) = 2\inf_{X,Y}\mathbb{P}(X \neq Y),$$

where (X,Y) are any random variables defined on the same probability space with law $\mathcal{L}(X) = \nu$ and $\mathcal{L}(Y) = \mu$. Equivalently,

$$d_{TV}(\nu,\mu) = \int |\nu - \mu|(\mathrm{d}x).$$

Theorem 14.1. *Let (X_n) be a Markov chain with transition kernel $P(x,\mathrm{d}y)$:*

$$a.s., \forall n \geq 0, \quad \mathbb{P}(X_{n+1} \in \mathrm{d}y \mid X_n) = P(X_n, \mathrm{d}y).$$

Assume that the transition kernel satisfies the following Doeblin's condition: there exist a probability measure ν and $\alpha \in (0,1]$ such that

$$P(x,\mathrm{d}y) \geq \alpha\nu(\mathrm{d}y) \qquad \textit{for every } x.$$

Then, for all initial states x,y and all $n \geq 1$,

$$d_{TV}(\mathcal{L}(X_n^x), \mathcal{L}(X_n^y)) \leq 2(1-\alpha)^n.$$