The Crux of the Non-Gradient Method

Julian Kern

May 09, 2023

1 Generator Magic (Heuristics)

Consider a Markov Process X on state space S with generator A (and domain $D(A) \subseteq \mathcal{C}_b(S)$). Then, for any $f \in D(A)$,

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (Af)(X_s) \, \mathrm{d}s \tag{1}$$

is a mean-zero martingale with quadratic variation

$$[M^f]_t := \int_0^t \Big(A(f^2) - 2fAf \Big)(X_s) \, \mathrm{d}s.$$

Conversely, under some technical conditions, if a process X satisfies (1) for all $f \in D(A)$, then X is a strong Markov Process with generator A. That has the following consequence:

Theorem 1.1 (Handwaving [EthKur86])**.** *Consider Markov Processes X^N with generators A^N. Suppose that*

i) The sequence $(X^N)_N$ is tight (so that we can extract a convergent subsequence).

ii) The generators $A^N \to A$ converge in a suitable way.

iii) Many technical conditions...

Then, $X^N \to X$ in distribution to the Markov Process X with generator A.

Furthermore, if the **quadratic variation vanishes** in the limit, the **limiting process is deterministic.** In this case, the above formulation is nothing else than the weak formulation of PDEs of the type

$$\partial_t \rho_t = A\rho_t,$$

where the test functions $f = f_G : L^2(\mathbb{T}) \to \mathbb{R}$ usually take the form

$$f_G(\rho) := \langle \rho, G \rangle := \int_{\mathbb{T}} \rho(u)G(u) \, \mathrm{d}u$$

for some $G : \mathbb{T} \to \mathbb{R}$ smooth enough. In the case of the heat equation ($A = \Delta$), the action of the generator is

$$\Delta f_G := f_{\Delta G}$$

and the weak formulation becomes

$$0 = \langle \rho_t, G \rangle - \langle \rho_0, G \rangle - \int_0^t \langle \rho_s, \Delta G \rangle \, \mathrm{d}s.$$

Note that usually the test functions are *time dependent* (to prove uniqueness of the solution). That would add a term $\int_0^t (\partial_s f_s)(\rho_s) \, \mathrm{d}s$ or $\int_0^t \langle \rho_s, \partial_s G_s \rangle \, \mathrm{d}s$ to the formulations. In our case, this additional term does not make any difference, so I don't mention it for simplicity!

2 The Symmetric Simple Exclusion Process (SSEP)

Everything will be defined on the torus for simplicity: $\mathbb{T}_N := \mathbb{Z}/(N\mathbb{Z})$ and $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. The exclusion processes take values in the space of all particle configurations $\Omega_N := \{0,1\}^{\mathbb{T}_N}$.

We then consider the Markov process $\eta := (\eta_t^N)_{t \geq 0}$ with values in Ω_N, started in μ_N, with generator

$$\begin{aligned}\mathcal{L}^N f(\eta) &:= \frac{1}{2} \sum_{x \in \mathbb{T}_N} \Big(\eta_x (1 - \eta_{x+1}) \cdot \big(f(\eta^{x,x+1}) - f(\eta) \big) + \eta_x (1 - \eta_{x-1}) \cdot \big(f(\eta^{x-1,x}) - f(\eta) \big) \Big) \\ &= \frac{1}{2} \sum_{x \in \mathbb{T}_N} \Big(f(\eta^{x,x+1}) - f(\eta) \Big).\end{aligned}$$

We expect the limit ρ to be deterministic. (One can easily show that the quadratic variation vanishes.) This leads us to the *ansatz* $\rho_t(x/N) \approx \rho_t^N(x) := \mathbb{E}[\eta_t^N(x)]$. Hence, we want to compute

$$\begin{aligned}\partial_t \rho_t^N(x) = \mathbb{E}[\mathcal{L}^N \eta_t(x)] &= \mathbb{E}\left[\frac{1}{2} \cdot \big(\eta_t(x+1) + \eta_t(x-1) - 2\eta_t(x) \big) \right] \\ &= \frac{1}{2} \cdot \Delta^N \rho_t^N(x),\end{aligned}$$

where Δ^N is the discrete Laplacian on $\mathbb{T}_N$. Since, at least heuristically via parabolic scaling,

$$N^2 \Delta^N \longrightarrow \Delta,$$

we might expect that the sped-up process $(\eta_{N^2 t}^N)_{t \geq 0}$ converges to the solution of the heat equation

$$\partial_t \rho_t = \frac{1}{2} \Delta \rho_t.$$

More rigorously: We are interested in test functions

$$\langle \rho_t, G \rangle = \int_{\mathbb{T}} \rho_t(u) G(u) \, \mathrm{d}u \approx \sum_{x \in \mathbb{T}_N} \rho_{N^2 t}^N(x) G\left(\frac{x}{N}\right) \approx \langle \pi_t^N, G \rangle,$$

where

$$\pi_t^N := \frac{1}{N} \sum_{x \in \mathbb{T}_N} \eta_{N^2 t}(x) \delta_{x/N}$$

is the rescaled empirical measure process. Then, π^N is a Markov process with generator (using the same symbol)

$$\begin{aligned} N^2 \mathcal{L}^N \langle \pi^N, G \rangle &= N^2 \sum_{x \in \mathbb{T}_N} (\mathcal{L}^N \eta_x) \cdot G\left(\frac{x}{N}\right) \\ &= \frac{N^2}{2} \sum_{x \in \mathbb{T}_N} \Delta^N \eta^N(x) \cdot G\left(\frac{x}{N}\right) \\ &= \left\langle \pi^N, \frac{1}{2}(N^2 \Delta^N) G \right\rangle . \end{aligned}$$

Note that we performed *two summations by parts* to shift the Laplacian from η to the test function G.

We conclude that

$$M_t^{G,N} := \langle \pi_t^N, G \rangle - \langle \pi_0^N, G \rangle - \int_0^t \left\langle \pi_s^N, \frac{1}{2} \Delta_N G \right\rangle \, \mathrm{d}s$$

is a mean-zero martingale. If $\pi^N \longrightarrow \pi$ in distribution, then these quantities converge to

$$\langle \pi_t, G \rangle - \langle \pi_0, G \rangle - \int_0^t \left\langle \pi_s, \frac{1}{2} \Delta G \right\rangle \, \mathrm{d}s.$$

Furthermore, one shows that $M^{G,N}$ converges to 0 and that, necessarily, $\pi(\,\mathrm{d}x) = \rho(x)\,\mathrm{d}x$ for some $\rho \in L^1(\mathbb{T})$.

This gives that ρ satisfies

$$\partial_t \rho_t = \frac{1}{2} \Delta \rho$$

weakly.

The above covers everything that was discussed during the talk on May 09. The rest will be covered during the first session of the minicourse on May 11. It is not necessary to read it beforehand and is meant as an overview of the motivation for the rest of the minicourse.

3 Gradient Models

One way of writing the generator from the SSEP is

$$\mathcal{L}^N \eta_x = J_{x-1,x} - J_{x,x+1},$$

where $J_{x,x+1}$ is the *current* of particles flowing from x to $x+1$, given by

$$\begin{aligned} J_{x,x+1} &= (\text{flux from } x \text{ to } x+1) - (\text{flux from } x+1 \text{ to } x) \\ &= \eta_x(1-\eta_{x+1}) - \eta_{x+1}(1-\eta_x) \\ &= \eta_x - \eta_{x+1} = -\nabla^N \eta_x. \end{aligned}$$

This decomposition into currents is often possible. This allows us to do a summation by parts via

$$\begin{aligned} N^2\mathcal{L}^N\langle \pi^N, G\rangle &= N^2 \sum_{x\in\mathbb{T}_N} (\mathcal{L}^N \eta_x)\cdot G\left(\frac{x}{N}\right) \\ &= N^2 \sum_{x\in\mathbb{T}_N} \big(J_{x-1,x} - J_{x,x+1}\big)\cdot G\left(\frac{x}{N}\right) \\ &= \sum_{x\in\mathbb{T}_N} N J_{x,x+1}\cdot N\nabla^N G\left(\frac{x}{N}\right), \end{aligned}$$

where $\nabla^N G\left(\frac{x}{N}\right) = G\left(\frac{x+1}{N}\right) - G\left(\frac{x}{N}\right)$ is the discrete derivative.

To proceed as before, we need to perform a **second summation by parts**. In analogy to the SSEP for which $J_{x,x+1} = -\nabla^N \eta_x$, we assume the *gradient condition* for the current:

$$J_{x,x+1} = h(\eta_x) - h(\eta_{x+1}).$$

Then, we get

$$\begin{aligned} N^2\mathcal{L}^N\langle \pi, G\rangle &= \sum_{x\in\mathbb{T}_N} N\Big(h(\eta_x) - h(\eta_{x+1})\Big)\cdot N\nabla^N G\left(\frac{x}{N}\right) \\ &= \sum_{x\in\mathbb{T}_N} h(\eta_x)\cdot N^2\Delta^N G\left(\frac{x}{N}\right) \end{aligned}$$

and we might expect the limiting equation to be

$$\partial_t \rho_t = \Delta\big(h(\rho_t)\big).$$

Closing the Equation: There is one technical issue to overcome. In the present form, we cannot use the weak convergence of π^N as the above is not (necessarily) a continuous functional of π^N. The usual approach is to replace η_x by the average

$$\eta_x^{\varepsilon N} := \frac{1}{2\varepsilon N}\sum_{|y|\le \varepsilon N} \eta_{x+y}$$

over a macroscopic box of size ε. The advantage is that we may write

$$\eta_x^{\varepsilon N} = \langle \pi^N, \iota_\varepsilon^x\rangle$$

as a continuous functional of π, where $\iota_\varepsilon^x := \frac{1}{2\varepsilon}\mathbb{1}_{[x-\varepsilon,x+\varepsilon]}$. This implies

$$N^2\mathcal{L}^N\langle\pi^N,G\rangle \approx \sum_{x\in\mathbb{T}_N} h\big(\langle\pi^N,\iota_\varepsilon^x\rangle\big)\cdot N^2\Delta^N G\left(\frac{x}{N}\right)$$

which converges in N to

$$\int_{\mathbb{T}} h\left(\frac{1}{2\varepsilon}\int_{u-\varepsilon}^{u+\varepsilon}\rho(v)\ \mathrm{d}v\right)\cdot\Delta G(u)\ \mathrm{d}u \underset{\varepsilon\downarrow 0}{\longrightarrow} \int_{\mathbb{T}} h(\rho(u))\cdot\Delta G(u)\ \mathrm{d}u.$$

Although this part of the proof is not trivial and the so-called *Replacement Lemmas* usually take up a big part of the overall proof, these will not be the main point of this talk. So from now on, we will more or less ignore this technicality and simply assume that everything works as just mentioned.

4 Examples Non-Gradient Models

Let us modify the Simple Symmetric Exclusion Process by changing the jumping rate from one site to another by a factor that may depend on the surrounding sites:

$$\mathcal{G}^N f(\eta) := \sum_{x\in\mathbb{T}_N}\sum_{z=\pm1} c(\eta;x)\cdot\eta_x(1-\eta_{x+z})\cdot\big(f(\eta^{x,z})-f(\eta)\big),$$

where $c(\eta;x) = c(\tau_x\eta)$ is a translationally invariant function depending on a finite number of values of η around x. This model is considered in [VarYau97].

As before, we can define the current

$$J_{x,x+1} = c(\eta;x)\cdot\eta_x(1-\eta_{x+1}) - c(\eta;x+1)\eta_{x+1}(1-\eta_x)$$

yielding the current representation

$$\mathcal{G}^N\eta_x = J_{x-1,x} - J_{x,x+1}$$

and thus

$$N^2\mathcal{G}^N\langle\pi^N,G\rangle = \sum_{x\in\mathbb{T}_N} NJ_{x,x+1}\cdot N\nabla^N G\left(\frac{x}{N}\right).$$

The problem is that J **does not satisfy the gradient condition** as soon as c is not constant!

Heuristically, we might expect that

$$\begin{aligned} NJ_{x,x+1} &= Nc(\eta;x)\big(\eta_x-\eta_{x+1}\big) + N\eta_x(1-\eta_{x+1})\big(c(\eta;x)-c(\eta;x+1)\big) \\ &\approx -c(\rho(x))\nabla\rho(x) + \rho(x)(1-\rho(x))\nabla c(\rho(x)). \end{aligned}$$

In other words, we expect to see a limiting equation of the form

$$\partial_t\rho_t(x) = \nabla\cdot\big(D(\rho_t(x))\nabla\rho_t(x)\big)$$

for some diffusion coefficient D.

The problem is that we do not have any estimates that would allow us to deduce this type of convergence of NJ. This is the object of the *Non-Gradient Method*: In what sense can NJ be replaced by a gradient (rescaled by a diffusion coefficient)?

A second example can be obtained by introducing a second species. In this context, a site can be in one of three states: 0 (no particle), 1 (particle of type +) or −1 (particle of type −). Then, the generator of the simply symmetric multi-species exclusion process becomes

$$\mathcal{H}^N f(\eta) := \sum_{x \in \mathbb{T}_N} \sum_{z=\pm 1} |\eta_x| \big(1 - |\eta_{x+z}|\big) \cdot \big(f(\eta^{x,y}) - f(\eta)\big),$$

where $|\eta_x| = 1$ iff there is a particle at site x. The process is uniquely characterised by the pair $(|\eta|, \eta^+)$, where $|\eta|$ tracks where the particles are (it is a standard SSEP!) and η^+ tracks only the particles of type +. In particular, that means that $|\eta|$ can be described by the gradient methods. Let us concentrate on η^+. One easily checks that

$$\mathcal{H}^N \eta_x^+ = J^+_{x-1,x} - J^+_{x,x+1},$$

where the current of positive particles is given by

$$J^+_{x,x+1} = \eta^+_x \big(1 - |\eta_{x+1}|\big) - \big(1 - |\eta_x|\big)\eta^+_{x+1}$$

which is not in gradient form. For details on this particular model, see [Éri22] for an introduction and [Qua92] for a full treatment.

References

[EthKur86] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley series in probability and mathematical statistics. Wiley, 1986.

[KipLan98] Claude Kipnis and Claudio Landim. *Scaling Limits of Interacting Particle Systems* -. Berlin Heidelberg: Springer Science and Business Media, 1998. ISBN: 978-3-540-64913-7.

[Qua92] Jeremy Quastel. "Diffusion of color in the simple exclusion process". In: *Communications on Pure and Applied Mathematics* 45.6 (1992), pp. 623–679. DOI: https://doi.org/10.1002/cpa.3160450602. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.3160450602. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160450602.

[Éri22] "Varadhan's method for non gradient diffusive hydrodynamics". Mini-Course. 2022. URL: http://chercheurs.lille.inria.fr/cerignou/contenu/NG.pdf.

[VarYau97] S. R. S. Varadhan and Horng-Tzer Yau. "Diffusive limit of lattice gas with mixing conditions". In: *Asian Journal of Mathematics* 1 (1997), pp. 623–678.