

# The Fokker-Planck Equation

Pillon, Erik

July 4, 2018

## Abstract

Something

All the codes were written by myself and where not explicitly said all the graphics were produced in MATLAB or Julia v0.6.3.

We begin with a particular Fokker-Planck equation where the drift is the gradient of a given smooth potential  $V: \mathbb{R}^d \rightarrow \mathbb{R}$

$$\begin{cases} \frac{\partial}{\partial t} n(t, x) - \Delta n(t, x) - \operatorname{div}(n(t, x) \nabla V(x)) = 0, & t \geq 0, x \in \mathbb{R}^d, \\ n(t = 0, x) = n^0(x) \end{cases} \quad (1)$$

## 1 Elementary a priori estimates

The main properties of solutions of (1) are the nonnegativity principle, the ‘mass’ conservation and the existence of a non-zero steady state,

$$n^0 \geq 0 \longrightarrow n(t, x) \geq 0, \quad (2)$$

$$\int_{\mathbb{R}^d} n(t, x) dx = \int_{\mathbb{R}^d} n^0(x) dx, \quad \forall t \geq 0, \quad (3)$$

$$\mu e^{-V(x)} \text{ are steady states for all } \mu \in \mathbb{R}. \quad (4)$$

this last statement is because  $\Delta e^{-V} = \operatorname{div}(\nabla e^{-V}) = -\operatorname{div}(e^{-V} \nabla V)$

It is intuitive that the decay properties of this steady state should play a role, for example because its integrability is a desirable property. This relies on growth assumptions of  $V(x)$  at infinity. A potential  $V$  is called *confining* if  $V(x) \rightarrow +\infty$  as  $|x| \rightarrow \infty$ ; typically  $V(x) = \frac{\|x\|^2}{2}$  determines a Gaussian steady state.

A simple idea is to look directly for a priori estimates in  $L^2$ . After multiplication by  $n$  and integration by parts, one finds

$$\frac{1}{2} \frac{\partial}{\partial t} \int_{\mathbb{R}^d} n(t, x)^2 dx + \int_{\mathbb{R}^d} |\nabla n(t, x)|^2 dx = - \int_{\mathbb{R}^d} \nabla n(t, x) \nabla V(x) n(t, x) dx \quad (5)$$

$$= - \frac{1}{2} \int_{\mathbb{R}^d} \nabla n(t, x)^2 \nabla V(x) dx \quad (6)$$

and thus we arrive at

$$\frac{1}{2} \frac{\partial}{\partial t} \int_{\mathbb{R}^d} n(t, x)^2 dx + \int_{\mathbb{R}^d} |\nabla n(t, x)|^2 dx = \frac{1}{2} \int_{\mathbb{R}^d} n(t, x)^2 \Delta V(x) dx. \quad (7)$$

When  $\Delta V \leq -\nu < 0$ , with  $\nu$  a constant, one finds that

$$\int_{\mathbb{R}^d} n(t, x)^2 dx \leq e^{-\mu} \int_{\mathbb{R}^d} n^0(x)^2 dx$$

Note that, in this case, the steady state in (4) is not integrable and this calculation shows the interest of confining potentials.

## 2 Weak Solution

We define now a weak solution of (1) by testing against a function  $\Phi \in \mathcal{D}(\mathbb{R}^+ \times \mathbb{R}^d)$

$$\int_0^\infty \int_{\mathbb{R}^d} n(t, x) \left[ -\frac{\partial \Phi}{\partial t} - \Delta \Phi + \nabla V \cdot \nabla \Phi \right] dx dt = \int_{\mathbb{R}^d} n^0(x) \Phi(0, x) dx. \quad (8)$$

We see that this makes sense under different assumptions for  $V$ . Examples are

- assume  $\nabla V \in L_{loc}^2(\mathbb{R}^d)$  and  $n \in L^\infty((0, T); L^2(\mathbb{R}^d))$  for all  $T > 0$ ,
- assume  $\nabla V \in C(\mathbb{R}^d)$  and  $n \in L^\infty(\mathbb{R}^+; M^1(\mathbb{R}^d))$  (bounded measure), which is relevant in view of mass conservation.

Another option is to use, as we did for the relative entropy structure, the alternative formulation

$$\begin{cases} \frac{\partial}{\partial t} n(t, x) - \operatorname{div} \left( e^{-V(x)} \nabla (n(t, x) e^{V(x)}) \right) = 0, & t \geq 0, x \in \mathbb{R}^d, \\ n(t = 0, x) = n^0(x). \end{cases} \quad (9)$$

This leads to another possible weak formulation

$$\int_0^\infty \int_{\mathbb{R}^d} n(t, x) \left[ -\frac{\partial \Phi}{\partial t} - e^{V(x)} \operatorname{div} \left( e^{-V(x)} \nabla \Phi \right) \right] dx dt = \int_{\mathbb{R}^d} n^0(x) \Phi(0, x) dx,$$

which is not very different from (8).

When applying the Lax-Milgram theory [2] for steady states, the two formulations are however very different. The direct formulation leads to the use of the scalar product (definiteness should come from the zero order term for example)

$$((n, \Phi))_1 = \int_{\mathbb{R}^d} \nabla n(x) \cdot \nabla \Phi(x) dx + \int_{\mathbb{R}^d} n(x) \nabla \Phi(x) \cdot \nabla V(x) dx.$$

The modified formulation leads to the scalar product (we use the test function  $\Phi(x) = \varphi(x) e^{V(x)}$ )

$$((n, \varphi))_2 = \int_{\mathbb{R}^d} \nabla (n(x) e^{V(x)}) \cdot \nabla (\varphi(x) e^{V(x)}) e^{-V(x)} dx = \int_{\mathbb{R}^d} \nabla \tilde{n}(x) \cdot \nabla \Phi(x) e^{-V(x)} dx.$$

This is a self-adjoint scalar product in the Hilbert space  $H^1(e^{-V(x)} dx)$ , with definiteness directly related to the Poincaré inequality in the full space.

### 3 The Complete Fokker-Planck Equation

The complete Fokker-Planck equation (also called the Kolmogorov equation in the theory of Markov/diffusion processes) is given by

$$\begin{cases} \frac{\partial}{\partial t} n(t, x) - \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j} (B_{ij}(t, x) n(t, x)) + \operatorname{div}(n(t, x) U(t, x)) = 0, & t \geq 0, x \in \mathbb{R}^d, \\ n(t = 0, x) = n^0(x). \end{cases} \quad (10)$$

Here  $U: \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the velocity field and  $B: \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathcal{M}_+^{d \times d}$  is the nonnegative symmetric diffusion matrix. In other words, it is always assumed to satisfy, for some constant  $\nu \geq 0$ ,

$$\sum_{i,j=1}^d B_{i,j}(t, x) \xi_i \xi_j \geq \nu |\xi|^2, \quad \forall \xi \in \mathbb{R}^d.$$

The main properties are still the sign property

$$n^0 \geq 0 \Rightarrow n \geq 0,$$

and 'mass' conservation

$$\int_{\mathbb{R}^d} n(t, x) dx = \int_{\mathbb{R}^d} n^0(x) dx, \quad \forall t \geq 0.$$

We assume that there is a steady state  $N(x)$  to (1)

$$-\frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j} (B_{ij}(x) N(x)) + \operatorname{div}(N(x) U(x)) = 0, \quad \forall x \in \mathbb{R}^d, \quad N(x) > 0. \quad (11)$$

The existence of solutions depends heavily on the coefficients  $(B, U)$  and on the desired properties for  $N$  (usually  $N \in L^1 \cap L^\infty(\mathbb{R}^d)$ ), but a typical example is as before

$$\frac{1}{2} B = I \quad (\text{identity matrix}), \quad U(x) = -\nabla V(x), \quad N(x) = \mu e^{-V(x)}, \quad \mu \in \mathbb{R}.$$

More generally for a constant matrix  $B$ , when  $U = -B \nabla V(x)$ , we find the same steady states.

### 4 Stochastic Differential Equations (SDE)

The Fokker-Planck equation is intimately connected to Stochastic Differential Equations and we explain this in the following. We give here a very intuitive and non-rigorous construction based on numerical approximations. For a given, sufficiently smooth, vector field  $U \in \mathbb{R}^d$  and a matrix  $\sigma(t, x) \in M_{d,p}$ , one can build the solution of the Itô

Stochastic Differential Equation

$$\begin{cases} dX(t) = U(t, X(t)) dt + \sigma(t, X(t)) dW(t), \\ X(0) = X^0 \in \mathbb{R}^d \quad (\text{a random vector}) \end{cases} \quad (12)$$

with  $W(t) = (W^1(t), \dots, W^p(t))$ ,  $p$  independent Brownian motions.

The numerical construction is essentially based on the numerical approximations through subsequent finite increments. This defines the Euler scheme

$$X^{k+1} = X^k + \Delta t U(t^k, X^k) + \sqrt{\Delta t} \sigma(t^k, X^k) Y^k. \quad (13)$$

Because the random variable  $Y^k$  is independent of  $X^k$  and  $\mathbb{E}[Y^k] = 0$ , we have

$$\mathbb{E}[X^{k+1} - X^k] = \Delta t \mathbb{E}[U(t^k, X^k)], \quad (14)$$

and

$$\mathbb{E}[(X^{k+1} - X^k)^2] = \mathbb{E}[\Delta t^2 U^2(t^k, X^k) + 2\Delta t^{3/2} U(t^k, X^k) \sigma(t^k, X^k) Y^k + \Delta t |\sigma(t^k, X^k) Y^k|^2] \quad (15)$$

$$= \Delta t \mathbb{E}[\sigma^2(t^k, X^k)] \mathbb{E}[(Y^k)^2] + O(\Delta t^2) \quad (16)$$

$$= \Delta t \mathbb{E}[\sigma^2(t^k, X^k)] + O(\Delta t^2) \quad (17)$$

Except when  $\sigma \equiv 0$ , this scales like the Brownian motion. The type of approximation scheme used in (13) is very important, in opposition to the deterministic case. Changing this approximation also changes the limit. The Stratonovich convention is an example and the notion is denoted as

$$\begin{cases} dX(t) = U(t, X(t)) dt + \sigma(t, X(t)) \circ dW(t), \\ X(0) = X^0 \in \mathbb{R}^d \quad (\text{a random vector}) \end{cases} \quad (18)$$

The Stratonovich SDE corresponds to the semi-implicit approximation in the stochastic term

$$X^{k+1} = X^k + \Delta t U(t^k, X^k) + \sqrt{\Delta t} \sigma\left(t^k, \frac{X^k + X^{k+1}}{2}\right) Y^k. \quad (19)$$

It may be approximated as

$$X^{k+1} \approx X^k + \Delta t U(t^k, X^k) + \sqrt{\Delta t} \left[ \sigma(t^k, X^k) + \frac{1}{2} D\sigma(t^k, X^k)(X^{k+1} - X^k) \right] Y^k \quad (20)$$

$$\approx X^k + \Delta t U(t^k, X^k) + \frac{\Delta t}{2} D\sigma(t^k, X^k) \sigma(t^k, X^k) (Y^k)^2 + \sqrt{\Delta t} \sigma(t^k, X^k) Y^k + O(\Delta t^{3/2}). \quad (21)$$

This means that

$$\mathbb{E}[X^{k+1} - X^k] = \Delta t \mathbb{E}[U(t^k, X^k)] + \frac{\Delta t}{2} \mathbb{E}[D\sigma(t^k, X^k) \sigma(t^k, X^k)], \quad (22)$$

In other words, one cannot take the expectation as simply as for Itô SDE.

Because  $\mathbb{E}[(Y^k)^3] = 0$ , we also find that

$$\mathbb{E}[(X^{k+1} - X^k)^2] = \Delta t \mathbb{E}[\sigma^2(t^k, X^k)] + O(\Delta t^2).$$

It is remarkable that the Stratonovich semi-implicit scheme leads to the same first two expectations as the Itô formula with the drift  $U$  changed in  $U - \frac{1}{2}\sigma'\sigma$ ; in fact for the limit  $\Delta t \rightarrow 0$ , the two constructions are equivalent with this modification for the drift. This is to say that the limit of the term  $(Y^k)^2$  is 1 in probability.

## 5 The Itô Formula

An important tool in the theory of SDEs is the Itô formula that gives the equation satisfied by the random variable  $u(t, X(t))$ , when  $u \in C^2(\mathbb{R}^{d+1}; \mathbb{R})$ ,

$$du(t, X(t)) = \left[ \frac{\partial u(t, X(t))}{\partial t} + U(t, X(t)) \cdot Du(t, X(t)) + \frac{1}{2} B(t, X(t)) \cdot D^2 u(t, X(t)) \right] dt + \sigma(t, X(t)) \cdot Du(t, X(t)) \cdot dW(t) \quad (23)$$

with

$$B(t, x) = \frac{1}{2} \sigma(t, x) \cdot \sigma^t(t, x) \quad (\sigma^t \text{ denotes the transposed matrix}). \quad (24)$$

In other words, the chain rule does not apply to SDEs as it applies to ODEs ( $\sigma(t, x) = 0$ ) and it gives an additional term, namely the term  $B(t, X(t)) \cdot D^2 u(t, X(t))$ .

This formula can be derived from the approximation (13). Using Taylor's expansion in time, we find

$$u(t^{k+1}, X^{k+1}) = u(t^k, X^k) + \Delta t \frac{\partial u(t^k, X^{k+1})}{\partial t} + O(\Delta t^2) \quad (25)$$

$$= u(t^k, X^{k+1}) + \Delta t \frac{\partial u(t^k, X^k)}{\partial t} + O(\Delta t^{3/2}). \quad (26)$$

Therefore, using now Taylor's expansion in space, we find

$$\begin{aligned} u(t^{k+1}, X^{k+1}) &= u(t^k, X^k) + \Delta t \frac{\partial u(t^k, X^k)}{\partial t} + O(\Delta t^{3/2}) \\ &\quad + \left[ \Delta t U(t^k, X^k) + \sqrt{\Delta t} \sigma(t^k, X^k) Y^k \right] \cdot Du(t^k, X^k) \\ &\quad + \frac{1}{2} \Delta t \sigma^2(t^k, X^k) (Y^k)^2 D^2 u(t^k, X^k). \end{aligned}$$

As we already saw for the Stratonovich model, the limit of the term  $(Y^k)^2$  is 1 for the limit  $\Delta t \rightarrow 0$  and we see that this is the discrete version of (23). In the case of the Stratonovich convention, one can use the chain rule as in the deterministic case (and this is the advantage gained from the more complicated average)

$$du(t, X(t)) = \left[ \frac{\partial u(t, X(t))}{\partial t} + U(t, X(t)) \cdot Du(t, X(t)) \right] \Delta t + \sigma(t, X(t)) \cdot Du(t, X(t)) \circ dW(t). \quad (27)$$

To see this, we can redo the discrete version, and to simplify, use a function of  $X$  only,

$$u(X^{k+1}) = u(X^k) + Du(X^k) \left[ \Delta t U(t^k, X^k) + \sqrt{\Delta t} \sigma \left( \frac{X^k + X^{k+1}}{2} \right) Y^k \right] + \frac{\Delta t}{2} D^2 u(X^k) \sigma \left( \frac{X^k + X^{k+1}}{2} \right)^2 (Y^k)^2 + O(\Delta t^{3/2}),$$

which we may ‘center’ to follow the semi-implicit rule

$$\begin{aligned} u(X^{k+1}) = u(X^k) + \Delta t Du(X^k) U(t^k, X^k) + \sqrt{\Delta t} Du \left( \frac{X^k + X^{k+1}}{2} \right) \sqrt{\Delta t} \sigma \left( \frac{X^k + X^{k+1}}{2} \right) Y^k \\ - \frac{\sqrt{\Delta t}}{2} D^2 u(X^k) (X^{k+1} - X^k) \sigma \left( \frac{X^k + X^{k+1}}{2} \right)^2 Y^k \\ + \frac{\Delta t}{2} D^2 u(X^k) \sigma \left( \frac{X^k + X^{k+1}}{2} \right)^2 (Y^k)^2 + O(\Delta t^{3/2}) \end{aligned}$$

and because the dominant term in  $X^{k+1} - X^k$  is  $\sqrt{\Delta t} \sigma \left( \frac{X^k + X^{k+1}}{2} \right) Y^k$ , we finally end up with

$$u(X^{k+1}) = u(X^k) + \Delta t Du(X^k) U(t^k, X^k) + \sqrt{\Delta t} Du \left( \frac{X^k + X^{k+1}}{2} \right) \sigma \left( \frac{X^k + X^{k+1}}{2} \right) Y^k + O(\Delta t^{3/2}),$$

which is the discrete version of (27).

## 6 The Kolmogorov Equation

To go further and make the link with the Fokker-Planck equation, we recall the

**Definition 6.1.** The probability density  $n(x)$  of a random variable  $X$  is defined by the identity

$$\int u(x) n(x) dx = \mathbb{E} [u(X(\omega))],$$

for all smooth test functions  $u$  with compact support.

We give the probability density  $n^0(x)$  of  $X^0$  and we denote by  $n(t, x)$  the probability density of the process  $X(t, \omega)$ . Then, we can deduce from the Itô formula that

**Theorem 6.1** (Fokker-Planck-Kolmogorov Equation). *For the process (13), the limiting probability density  $n(t, x)$  satisfies the Fokker-Planck equation (1).*

## 7 Numerical Implementations

## References

- [1] Benoît Perthame. “Parabolic equations in biology”. In: *Parabolic Equations in Biology*. Springer, 2015, pp. 1–21.

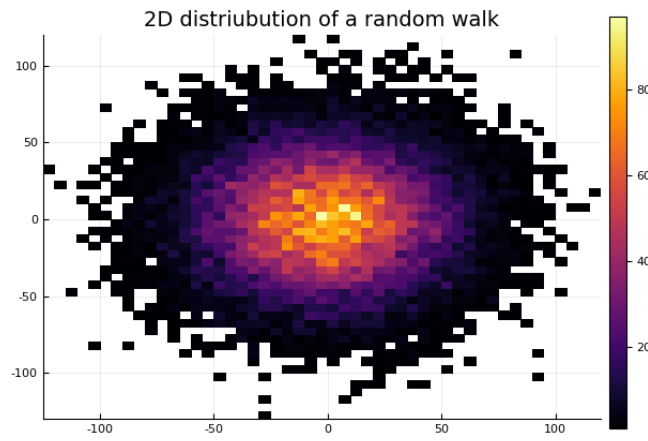


Figure 1: Numerical simulation of a 2D random walk

rw.jl

- [2] Sandro Salsa. *Equazioni a derivate parziali: Metodi, modelli e applicazioni*. Vol. 98. Springer, 2016.
- [3] Mark Tuckerman. *Statistical mechanics: theory and molecular simulation*. Oxford University Press, 2010.