MATH0488 – Elements of stochastic processes

Molecular dynamics: Langevin equation

Maarten Arnst
Marco Lucio Cerquaglia, Adrien Crovato, Joffrey Coheur, Thomas Lambert, and Cédric Laruelle

March 13, 2018

Outline

- Einstein's findings.
- Langevin equation.
- Brief review from ordinary differential equations.
- Stochastic calculus.
- Solution to Langevin equation.
- Numerical approximation.
- Integrating the displacement.
- Assignment.

Einstein's findings

1. The **evolution of the probability density function** of any component of the position of a particle relative to its initial position is governed by a **diffusion equation**:

$$\frac{\partial \rho}{\partial t}(x,t) - D\frac{\partial^2 \rho}{\partial x^2}(x,t) = 0.$$

The **second moment** of any component of the position of a particle relative to its initial position increases **proportionally to the elapsed time**:

$$\overline{x^2} = 2Dt.$$

2. In the case of spherical particles, the diffusion coefficient can be expressed as follows as a function of **experimentally observable quantities**:

$$D = \frac{RT}{N}\,\frac{1}{6\pi k a}.$$

Langevin equation

- Under normal conditions, in a liquid, a particle will suffer about 10^{21} collisions per second, Thus, there is a **slow time scale** associated with the effects of inertia of the particle and friction between the particle and the fluid and a **very fast time scale** associated with the molecular collisions.
- Because the **forces exerted by molecular collisions** on a particle happen on a much faster time scale than the effects of inertia of the particle and friction between the particle and the fluid, one can argue that these forces exerted by molecular collisions on a particle can be appropriately modeled as a **white noise**, and therefore as the **time derivative of a Wiener process**.
- This leads to the following stochastic differential equation (SDE) for (one component of the) motion of a Brownian particle, called the **Langevin equation**,

$$\underbrace{m\frac{dV}{dt}(t)}_{\text{inertia}} = \underbrace{-\mu V(t)}_{\text{friction}} + \underbrace{\sigma\frac{dW}{dt}(t)}_{\substack{\text{white noise}\\ \text{molecular collisions}}},$$

  where V is (the considered component of) the velocity of the particle, m is the mass of the particle, μ is the friction coefficient ($\mu = 6\pi ka$ in the case of a spherical particle and when Stokes's law is applicable), σ represents the magnitude of the forces exerted by the molecular collisions forces, and $\{W(t),\ t \in \mathbb{R}^+\}$ is the (scalar-valued) Wiener process.
- This is still a rather **phenomenological model**. The intermolecular interactions between the particle and the fluid molecules are not modeled explicitly; instead, their entire complexity is lumped into the friction coefficient μ, the magnitude σ, and the white-noise assumption.

Brief review from ordinary differential equations

- A first-order linear ordinary differential equation (ODE) with constant coefficients reads as

$$\frac{du}{dt}(t) = au(t) + b(t).$$

The theory of ODEs requires the right-hand side to be sufficiently regular, for example, the theorem of Cauchy-Lipschitz requires b to be a continuous function of time, among other conditions.

- The initial-value problem $\frac{du}{dt}(t) = au(t) + b(t)$ with $u(0) = u_0$ can be written in integral form as

$$u(t) = u_0 + \int_0^t au(s)ds + \int_0^t b(s)ds.$$

Both the Riemann and Lebesque theories of integration require integrands to be sufficiently regular, for example, continuity as a sufficient condition for Riemann integrability and measurability as a necessary condition for Lebesque integrability.

- The solution to $\frac{du}{dt}(t) = au(t) + b(t)$ with initial condition $u(0) = u_0$ has the analytical expression

$$u(t) = \exp(at)u_0 + \int_0^t \exp\big(a(t-s)\big)b(s)ds,$$

a result known in the literature as Duhamel's formula.

Stochastic calculus

- Unfortunately, **sample paths of white noise processes are not continuous or measurable functions**. Hence, SDEs involving white noise, such as the Langevin equation, cannot be studied within the theory of ODEs or by means of Riemann or Lebesque integrals.

- Still, SDEs involving white noise have found widespread use in many applications, for example, signal processing, biology, and finance, most often for modeling phenomena that happen on much faster time scales than other phenomena in the problem.

- The need for a **rigorous framework for studying SDEs** has led to the development of the theory of **stochastic calculus**, with important contributions from Kiyoshi Itô and many others.

Stochastic calculus

- Stochastic calculus defines a notion of stochastic integral that allows a rigorous meaning to be given to integrals involving white noise. Here, we will not look at this notion in its full generality, but we will limit ourselves to the definition of the stochastic integral of a continuous function of time, say, f, with respect to a white noise, that is, we will limit ourselves to a stochastic integral of the form

$$\int_0^t f(s)\frac{dW}{ds}(s)ds,$$

  which stochastic calculus also denotes alternatively (and more concisely) as follows:

$$\int_0^t f(s)dW(s).$$

- Not unlike how the theory of Riemann integration defines the integral $\int_0^t f(s)ds$ as $\lim_{n\to+\infty} \sum_{j=0}^{n-1} f(t_j)(t_{j+1}-t_j)$, stochastic calculus defines the **stochastic integral** $\int_0^t f(s)dW(s)$ as

$$\int_0^t f(s)dW(s) = \lim_{n\to+\infty} \sum_{j=0}^{n-1} f(t_j)\big(W(t_{j+1})-W(t_j)\big),$$

  where the limit has to be understood in the sense of the mean-square convergence of second-order random variables, and $0 = t_0 < t_1 < t_2 < \ldots < t_n = t$ is any partition of the interval $[0, t]$ such that $\max_{0\leq j\leq n-1}(t_{j+1}-t_j) \to 0$ as $n \to +\infty$. Whereas taking the Riemann integral of a continuous function returns a deterministic value, taking the stochastic integral of a continuous function with respect to a white noise returns a second-order random variable.

- The motivation underlying the choice of the mean-square convergence and the way in which the sum is formed, for example, the evaluation of f at t_j rather than at the mid-point $(t_{j+1} + t_j)/2$, is beyond the scope of this course. It is a methodological choice that can be justified by the level of generality of requisite regularity conditions, convenience for establishing theorems, adequacy to applications in science and engineering, convenience in numerical calculations, etc.

- The property that the Wiener process has bounded quadratic variation (Assignment 1) is key to defining the limit in the sense of the mean-square convergence of second-order random variables.

- The stochastic integral $\int_0^t f(s)dW(s)$ has the following **useful properties**:

$$E\left\{\int_0^t f(s)dW(s)\right\} = 0 \quad \text{and} \quad E\left\{\left(\int_0^t f(s)dW(s)\right)^2\right\} = \int_0^t \big(f(s)\big)^2 ds.$$

Proof of the aforementioned properties:

$$\begin{aligned} E\left\{\int_0^t f(s)dW(s)\right\} &= \lim_{n\to+\infty}\sum_{j=0}^{n-1} E\left\{f(t_j)\big(W(t_{j+1})-W(t_j)\big)\right\} \\ &= 0, \end{aligned}$$

$$\begin{aligned} E\left\{\left(\int_0^t f(s)dW(s)\right)^2\right\} &= \lim_{n\to+\infty}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} E\left\{f(t_i)\big(W(t_{i+1})-W(t_i)\big)f(t_j)\big(W(t_{j+1})-W(t_j)\big)\right\} \\ &= \lim_{n\to+\infty}\sum_{i=0}^{n-1}\big(f(t_j)\big)^2(t_{j+1}-t_j) \\ &= \int_0^t \big(f(s)\big)^2 ds, \end{aligned}$$

where the expectation operators and the limits may be interchanged by the continuity of the expectation operator on the space of second-order random variables, and $E\{W(t_{j+1})-W(t_j)\}=0$ and $E\{(W(t_{i+1})-W(t_i))(W(t_{j+1})-W(t_j))\}$ is equal to $(t_{j+1}-t_i)$ if i is equal to j and 0 otherwise by the definition of the Wiener process.

Solution to Langevin equation

- In order to give a rigorous meaning to the Langevin equation, stochastic calculus rewrites the stochastic initial-value problem

$$m\frac{dV}{dt}(t) = -\mu V(t) + \sigma\frac{dW}{dt}(t), \quad \text{with } V(0) = V_0,$$

where V_0 is a second-order random variable, in the integral form

$$V(t) = V_0 + \int_0^t -\frac{\mu}{m}V(s)ds + \int_0^t \frac{\sigma}{m}dW(s),$$

where the last integral is defined in the sense of the stochastic integral. Stochastic calculus then considers a stochastic process $\{V(t),\ t \in \mathbb{R}^+\}$ a solution if it solves this integral form.

- It can be shown that this stochastic initial-value problem has a solution $\{V(t),\ t \in \mathbb{R}^+\}$ that is unique up to probabilistic equivalence. This **solution has the analytical expression**

$$V(t) = \exp\left(-\frac{\mu}{m}t\right)V_0 + \int_0^t \exp\left(-\frac{\mu}{m}(t-s)\right)\frac{\sigma}{m}dW(s),$$

where the integral is again defined in the sense of the stochastic integral (please note the resemblance with Duhamel's formula obtained in the deterministic case).

Solution to Langevin equation

Proof of analytical expression of solution to Langevin equation:

- Let the stochastic process $\{V(t),\ t \in \mathbb{R}^+\}$ be the unique solution to

$$V(t) = V_0 + \int_0^t -\frac{\mu}{m} V(s)ds + \int_0^t \frac{\sigma}{m} dW(s).$$

- Consider then the stochastic process $\{Y(t),\ t \in \mathbb{R}^+\}$ such that

$$Y(t) = V(t) \exp\left(\frac{\mu}{m} t\right).$$

It can be shown that the sample paths of $\{V(t),\ t \in \mathbb{R}^+\}$ and therefore also those of $\{Y(t),\ t \in \mathbb{R}^+\}$ are continuous functions of time; hence, we have

$$Y(t) = V(0) \underbrace{\exp\left(\frac{\mu}{m} 0\right)}_{=1} + \sum_{j=0}^{n-1} \left(V(t_{j+1}) \exp\left(\frac{\mu}{m} t_{j+1}\right) - V(t_j) \exp\left(\frac{\mu}{m} t_j\right) \right),$$

where $0 = t_0 < t_1 < t_2 < \ldots < t_n = t$ is any partition of the interval $[0, t]$ such that $\max_{0 \leq j \leq n-1}(t_{j+1} - t_j) \to 0$ as $n \to +\infty$.

Solution to Langevin equation

Proof of analytical expression of solution to Langevin equation (continued):

- Recall the bivariate Taylor series of a function $(v,t) \mapsto g(v,t)$:

$$g(v+\Delta v, t+\Delta t) = g(v,t) + \frac{\partial g}{\partial v}(v,t)\Delta v + \frac{\partial g}{\partial t}(v,t)\Delta t + \frac{1}{2}\frac{\partial^2 g}{\partial v^2}(v,t)\Delta v^2 + \frac{\partial^2 g}{\partial v \partial t}(v,t)\Delta v \Delta t + \frac{1}{2}\frac{\partial^2 g}{\partial t^2}(v,t)\Delta t^2 + \dots$$

- With the help of a bivariate Taylor series of the function $(v,t) \mapsto v \exp\left(\frac{\mu}{m}t\right)$ up to the second order about $(V(t_j), t_j)$, we obtain

$$\begin{aligned} V(t_{j+1}) \exp\left(\frac{\mu}{m}t_{j+1}\right) = {} & V(t_j) \exp\left(\frac{\mu}{m}t_j\right) \\ & + \exp\left(\frac{\mu}{m}t_j\right)\left(V(t_{j+1}) - V(t_j)\right) \\ & + V(t_j)\frac{\mu}{m}\exp\left(\frac{\mu}{m}t_j\right)(t_{j+1} - t_j) \\ & + \frac{\mu}{m}\exp\left(\frac{\mu}{m}t_j\right)\left(V(t_{j+1}) - V(t_j)\right)(t_{j+1} - t_j) \\ & + \frac{1}{2}V(t_j)\left(\frac{\mu}{m}\right)^2 \exp\left(\frac{\mu}{m}t_j\right)(t_{j+1} - t_j)^2 \\ & + \text{remainder}. \end{aligned}$$

Solution to Langevin equation

Proof of analytical expression of solution to Langevin equation (continued):

- Because $V(t_{j+1}) - V(t_j) = \int_{t_j}^{t_{j+1}} -\frac{\mu}{m} V(s)ds + \frac{\sigma}{m}\big(W(t_{j+1}) - W(t_j)\big)$, it follows from the definitions of the Riemann integral and the stochastic integral that

$$\lim_{n\to+\infty} \sum_{j=0}^{n-1} \exp\left(\frac{\mu}{m} t_j\right) \left(V(t_{j+1}) - V(t_j)\right) = \int_0^t -\frac{\mu}{m} V(s) \exp\left(\frac{\mu}{m} s\right) ds + \int_0^t \frac{\sigma}{m} \exp\left(\frac{\mu}{m} s\right) dW(s),$$

$$\lim_{n\to+\infty} \sum_{j=0}^{n-1} V(t_j) \frac{\mu}{m} \exp\left(\frac{\mu}{m} t_j\right) (t_{j+1} - t_j) = \int_0^t \frac{\mu}{m} V(s) \exp\left(\frac{\mu}{m} s\right) ds$$

  and that the sums of the second-order terms and the remainder tend to 0 as $n \to +\infty$.

- Putting everything together, we get

$$Y(t) = V(0) + \int_0^t \frac{\sigma}{m} \exp\left(\frac{\mu}{m} s\right) dW(s),$$

  and therefore, since $Y(t) = V(t) \exp\left(\frac{\mu}{m} t\right)$, and upon dividing by $\exp\left(\frac{\mu}{m} t\right)$,

$$V(t) = \exp\left(-\frac{\mu}{m} t\right) V(0) + \int_0^t \exp\left(-\frac{\mu}{m}(t-s)\right) \frac{\sigma}{m} dW(s).$$

Solution to Langevin equation

- Let the random variable V_0 representing the initial velocity be statistically independent of the Wiener process whose derivative represents the molecular collisions. Then, by the properties of the stochastic integral stated previously, we obtain for the **mean and variance of the velocity**:

$$E\{V(t)\} = \exp\left(-\frac{\mu}{m}t\right) E\{V_0\},$$

$$E\left\{\left(V(t) - E\{V(t)\}\right)^2\right\} = \exp\left(-2\frac{\mu}{m}t\right) E\left\{\left(V_0 - E\{V_0\}\right)^2\right\} + \int_0^t \exp\left(-2\frac{\mu}{m}(t-s)\right)\left(\frac{\sigma}{m}\right)^2 ds$$

$$= \exp\left(-2\frac{\mu}{m}t\right) E\left\{\left(V_0 - E\{V_0\}\right)^2\right\} + \frac{\sigma^2}{2m\mu}\left(1 - \exp\left(-2\frac{\mu}{m}t\right)\right).$$

- Thus, we obtain $\lim_{t\to+\infty} E\{V(t)\} = 0$ and $\lim_{t\to+\infty} E\left\{\left(V(t)\right)^2\right\} = \frac{\sigma^2}{2m\mu}$; in other words, the **mean kinetic energy of the particle tends to become constant as** $t \to +\infty$:

$$\lim_{t\to+\infty} E\left\{\frac{m\left(V(t)\right)^2}{2}\right\} = \frac{\sigma^2}{4\mu}.$$

- Note that because of the definition of the stochastic integral in the analytical expression of the solution, because the increments of the Wiener process are Gaussian, because the sum of Gaussian random variables is Gaussian, and because convergence in mean square implies convergence in distribution, we have that if V_0 is either deterministic or Gaussian, then $V(t)$ will be Gaussian at any t, and if V_0 is random but not Gaussian, then $V(t)$ **will tend to become Gaussian as** $t \to +\infty$.

Solution to Langevin equation

- So far we have not dealt with **temperature**. However, we can intuitively expect that the kinetic energies of the particles and the fluid molecules will increase with temperature.
- In fact, in solutions of particles in fluids, the **collisions** will tend to lead to **kinetic energy transfer** from particles and fluid molecules with higher kinetic energy to those with lower kinetic energy. As a result, over time, a **thermal equilibrium will tend to set in**, wherein all the particles and fluid molecules will tend to have **on average the same kinetic energy, dependent on temperature**.
- In ideal solutions of particles in fluids, the thermal equilibrium is characterized by a linear relation between the mean kinetic energy and the (absolute) temperature (in Kelvin). Specifically, the mean kinetic energy associated with any component of motion is $\frac{1}{2}k_{\mathsf{B}}T$, so that the mean kinetic energy associated with all components involved is $\frac{3}{2}k_{\mathsf{B}}T$; here, k_{B} is the Boltzmann constant, which is related to the ideal gas constant R and Avogadro's number N as $k_{\mathsf{B}} = \frac{R}{N}$.
- This "law of equipartition of kinetic energy" is **consistent with the long-time behavior of Langevin's equation**, which provides us in turn a way of **calibrating the coefficient** σ, namely,

$$\frac{\sigma^2}{4\mu} = \frac{1}{2}k_{\mathsf{B}}T \implies \sigma = \sqrt{2k_{\mathsf{B}}T\mu}.$$

- We note that this "law of equipartition of kinetic energy" is intimately related to Van 't Hoff's law for the osmotic pressure that suspended particles develop in fluids, which Einstein used in his paper. Both laws hold under the same idealized conditions.

Numerical approximation

- Numerically approximating sample paths of the solution $\{V(t),\ t \in \mathbb{R}^+\}$ to the Langevin equation typically entails seeking to approximate its values at discrete time instants. We thus set $t_n = n\,\Delta t$, with n the index of the time instant and Δt the time step.
- We let V_n denote the approximation to $V(t_n)$ at t_n. The simplest time integration method, namely, the **explicit Euler method**, involves marching forward in time as $V_{j+1} = V_j + \left(-\frac{\mu}{m} V_j\right) \Delta t + \frac{\sigma}{m}\big(W(t_{j+1}) - W(t_j)\big)$, with $j = 0, 1, 2, \ldots$, where each term on the right-hand side approximates the corresponding term on the right-hand side of the integral form on Slide 10/20.
- Thus, given a sample path $\{w_0, w_1, w_2, \ldots\}$ of the values taken by the Wiener process at the discrete time instants $t_0, t_1, t_2, \ldots$, one begins with setting v_0 and then proceeds with computing the corresponding sample path $\{v_0, v_1, v_2, \ldots\}$ of approximations to the values taken by the solution to the Langevin equation at the discrete time instants $t_0, t_1, t_2, \ldots$ as follows:

$$v_{j+1} = v_j + \left(-\frac{\mu}{m} v_j\right) \Delta t + \frac{\sigma}{m}(w_{j+1} - w_j), \quad j = 0, 1, 2, \ldots,$$

- The explicit Euler method is **convergent** in that accuracy increases with a decrease in the time step Δt when integrating over finite intervals of time.
- As in the case of ODEs, the use of the explicit Euler method for the time-integration of SDEs is only **conditionally absolutely stable**. To have absolute stability, the time step must satisfy $\Delta t < \frac{2m}{\mu}$.

Integrating the displacement

- So far we have dealt with the velocity but not yet with the displacement.

- The **displacement can be integrated by considering the system of SDEs**

$$\begin{cases} \dfrac{dU}{dt}(t) = V(t), \\ m\dfrac{dV}{dt}(t) = -\mu V(t) + \sigma \dfrac{dW}{dt}(t), \end{cases} \quad \text{with} \quad \begin{cases} U(0) = U_0, \\ V(0) = V_0. \end{cases}$$

- It can be shown that this stochastic initial-value problem has a solution $\{(U(t), V(t)),\ t \in \mathbb{R}^+\}$ that is unique up to probabilistic equivalence and that this **solution has the analytical expression**

$$\begin{bmatrix} U(t) \\ V(t) \end{bmatrix} = \exp\left(\begin{bmatrix} 0 & 1 \\ 0 & -\frac{\mu}{m} \end{bmatrix} t \right) \begin{bmatrix} U_0 \\ V_0 \end{bmatrix} + \int_0^t \exp\left(\begin{bmatrix} 0 & 1 \\ 0 & -\frac{\mu}{m} \end{bmatrix} (t-s) \right) \begin{bmatrix} 0 \\ \frac{\sigma}{m} \end{bmatrix} dW(s).$$

  Please note that $\exp$ denotes here the matrix exponential (not the same as simply taking the exponential of the entries), and please note again the resemblance with Duhamel's formula.

- It is interesting to compare the behavior of this system of SDEs with Einstein's findings and Perrin's experiment, which pertained to the displacement. . . but let's leave that to you for the assignment :)

Assignment

1. Determine the expression of the mean vector and the covariance matrix of the solution $(U(t), V(t))$ at time t to the stochastic initial-value problem stated on Slide 17/20. Assume that the "law of equipartition of kinetic energy" holds and interpret the expression you find for the variance of $U(t)$ in the light of Einstein having found that the second moment of any component of the position of a particle relative to its initial position increases proportionally to the elapsed time.

2. It can be shown that the stochastic process $\{(U(t), V(t)),\ t \in \mathbb{R}^+\}$ that solves the stochastic initial-value problem stated on Slide 17/20 is a Markov process. Explain in a few words what is meant with the Markov property and why it is indeed to be expected that $\{(U(t), V(t)),\ t \in \mathbb{R}^+\}$ possesses the Markov property. Use your results for Question 1 to determine the expression of the transition probability density function, that is, the conditional probability density function of the solution $(U(t), V(t))$ at time t to the system of stochastic differential equations stated on Slide 17/20 given that the solution takes the deterministic value $(u(s), v(s))$ at a preceding time s.

3. Implement the explicit Euler method stated on Slide 16/20. Set $m = \mu = \sigma = 1$ and generate a few sample paths of the solution with $V(0) = 0$. Carry out a few tests to verify that your results make sense (suitability of the time step, correspondance with theoretical results about the mean and the variance,...). Change the values of m, μ, and σ. Interpret your results (role of m, μ, and σ).

Assignment

4. Extend the explicit Euler method on Slide 16/20 to the case of the system of SDEs on Slide 17/20. For a few of the cases that you considered under Question 2, please repeat your simulation, this time with $(U(0), V(0)) = (0, 0)$, and generate a few sample paths of the displacement. Interpret.
5. Extract from Perrin's paper values of m, μ, and σ that correspond to Perrin's experiment. Use the explicit Euler method implemented under Question 4 to generate for these values of m, μ, and σ a few sample paths of the solution to the stochastic initial-value problem stated on Slide 17/20 with $(U(0), V(0)) = (0, 0)$ for a sufficiently short time interval. Explain how the stability restriction on the time step prevents the explicit Euler method from being used with reasonable computational cost for the simulation of sample paths up to 120 s as in Perrin's experiment.
6. The availability of the analytical expression of the transition probability density function, which you determined under Question 2, enables an alternative and less computationally costly simulation of Perrin's experiment. Read on https://en.wikipedia.org/wiki/Multivariate_normal_distribution the section about drawing samples from a multivariate Gaussian probability density function. Imagine, explain, and implement a method that uses draws of samples from the transition probability density function to generate sample paths of the values taken at the time instants 0, 30 s, 60 s, 90 s, 120 s by the solution to the stochastic initial-value problem stated on Slide 17/20 with $(U(0), V(0)) = (0, 0)$ for the values of m, μ, and σ relevant to Perrin's experiment.
7. Imagine, explain, and implement how the simulation of Perrin's experiment of Question 6 can be used to characterize the accuracy with which Perrin was able to determine Avogadro's number. Were his estimates of Avogadro's number correct up to the accuracy limitations that you obtain?

References

Suggested reading material

- V. Denoel. Eléments de processus stochastiques. ULg, 2014–2015.
- A. Einstein. Investigations on the theory of the Brownian movement. www.bnpublishing.com, 2011.
- J. B. Perrin. Mouvement Brownien et réalité moléculaire. Annales de Chimie et de Physique, 19:5–104, 1909.

Additional references also consulted to prepare this lecture

- F. Comets and T. Meyre. Calcul stochastique et modèles de diffusions. Dunod, 2006.
- R. Feynman. The Feynman lectures on physics. Volume 1: Mainly mechanics, radiation, and heat. Basic Books, 2010.
- D. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. SIAM Review, 43:525–546, 2001.
- W. Liu, E. Karpov, and H. Park. Nano mechanics and materials. Wiley, 2006.
- B. Oksendal. Stochastic differential equations: an introduction with applications. Springer, 2010.
- C. Soize. The Fokker–Planck equation for stochastic dynamical systems and its explicit steady state solutions. World Scientific Publishing, 1994.
- G. Uhlenbeck and L. Ornstein. On the theory of the Brownian motion. Physical Review, 36:823–841, 1930.