Langevin Methods

B. Dünweg
Max Planck Institute for Polymer Research
Ackermannweg 10
D–55128 Mainz
Germany

Abstract. The lecture outlines the most important mathematical facts about stochastic processes which are described by a Langevin equation (stochastic differential equation, SDE) or (equivalently) a Fokker–Planck equation (FPE) comprising both drift and diffusion terms. The importance of the short–time behavior of the moments (mean displacement, mean square displacement) is stressed, and the problem of interpretation of SDEs (Ito vs. Stratonovich) is explained. The simplest integration scheme (Euler) is a straightforward consequence of this theory. For the simulation of thermal systems, drift and diffusion must balance each other in a well–defined way which fixes the temperature (fluctuation–dissipation theorem). The application of the general framework is then discussed for various methods commonly used in classical statistical physics (Brownian dynamics, stochastic dynamics, dissipative particle dynamics, force-biased Monte Carlo).

1. Introduction

A Langevin equation is typically written down when one wants to describe the dynamics of a system that (more or less naturally) can be decomposed into *fast* and *slow* degrees of freedom. The archetype of such a system (but by far not the only example) is slow Brownian particles immersed into fast solvent particles. The idea is that on sufficiently long time scales, the motion of the Brownian particles can be described just as random hops — the actual and complicated dynamical processes which bring these hops about are deliberately discarded from our attention, and the positions of the particles are only noticed at the flashes of some imaginary "stroboscope". On a somewhat shorter time scale, the influence of the solvent particles is replaced by *friction* and *noise* which "kicks" the Brownian particles randomly. The first picture gives rise to the method of Brownian Dynamics (BD), where one simulates a stochastic process just in terms of the particle coordinates, while the latter picture implies the method of stochastic dynamics (SD) , where one keeps both the positions and the momenta of the Brownian particles, and introduces friction and noise as additional forcing terms.

Such descriptions, which are based on the reduction of the number of degrees of freedom, and which therefore are both conceptually and technically much simpler than the original system, certainly involve

approximations. The purpose of this lecture is *not* to discuss if and under what circumstances this is a valid procedure, and how accurate such an approximation is. This is the topic of transport theory [8, 17] and would be well beyond the scope of the present contribution, which is rather intended as an elementary introduction into the mathematical background of stochastic processes involving both a continuous state space and continuous time [1, 11, 15, 9]. This is needed in order to understand what is actually meant when a Langevin equation is written down, and to understand how this is implemented on the computer with a finite discretization time step. Langevin simulations thus can be seen as somehow in between Monte Carlo (MC) and Molecular Dynamics (MD), sharing the element of randomness with the further, and of continuous trajectories in phase space with the latter. It turns out that the mathematical equivalence of the Langevin equation with the Fokker–Planck equation (FPE), which describes the evolution of the probability distribution in phase space, is particularly fruitful: This allows us to describe the stochastic process in terms of well–known concepts of probability theory, and to actually *define* what is meant by a Langevin equation. Furthermore, the concept of detailed balance, which plays such a central role in equilibrium MC (see contribution by A. Milchev), is here replaced by the rather analogous concept of the fluctuation–dissipation theorem (FDT): A simulation which runs in thermal equilibrium should have the Boltzmann distribution as (only) stationary distribution. This is also very easily checked in the Fokker–Planck picture, and we will see several examples below.

At this point, it should be emphasized that Langevin methods are very useful even when the underlying picture of fast vs. slow degrees of freedom does *not* apply, i. e. even when it is *impossible* to identify certain "hidden" degrees of freedom, which are supposed to be modeled by the random noise. Such a simulation should then not necessarily been viewed as a realistic description of the dynamics of the system, but it *will* produce the correct statics of the canonical ensemble if the FDT holds. It is thus just another MC procedure to generate the desired distribution. This is extremely convenient if one has a running MD code of the system available; a straightforward addition of just a few (typically of order ten, plus maybe a few hundred for a sophisticated random number generator) lines of code turns this into an SD simulation. This results in a change of ensemble (from constant energy to constant temperature), plus in a nice feature of numerical stabilization: As we will see below, the temperature is the ratio between noise strength and friction, and this may be viewed as a feedback control procedure for every single degree of freedom: A particle that is "too hot" will be cooled down because the friction term dominates, while

a particle that is too cold will be heated up by the noise. Such events easily occur as a result of the discretization errors of MD, but are "corrected" by the thermostat, which prevents such local catastrophes from building up further, and, in particular, from spreading throughout the system. Typically, the simulation of a dense Lennard–Jones fluid can thus be run with a time step which is two or even three times larger than what is appropriate for the constant–energy ensemble. Conversely, deterministic thermostats like the Nose–Hoover procedure rely on a *global* feedback for the whole system (see contribution by D. Rapaport), and hence do not have that good stabilization properties.

If one applies the Langevin equation with such a motivation, one should however be *very* careful when interpreting the results in terms of the dynamics of the system. Not much can be said beyond this rather general warning, since this depends very much on the system under consideration. One particular case shall however be discussed in some detail: The SD algorithm is absolutely useless for the simulation of hydrodynamic phenomena, even in the limit of rather weak noise. The reason is that it breaks Galilean invariance, as we will see below, and that it prevents the build–up of hydrodynamic correlations, and of (*physical*!) hydrodynamic instabilities, beyond a certain length scale ξ, the hydrodynamic screening length, which we shall discuss below. The same property which is a blessing for numerical stabilization thus turns out to be a curse for hydrodynamics. This problem has led to the modification of the SD thermostat to the so–called "dissipative particle dynamics" (DPD) algorithm, which cures it by restoring Galilean invariance, while still having quite nice stabilization properties.

The outline of this contribution is as follows: The first part is devoted to the discussion of the mathematical theory of Fokker–Planck processes and its lowest–order implementation, the Euler algorithm, at the end of which we will be able to formulate the FDT. We will then discuss various standard Langevin methods (BD, Force Biased MC, SD, and DPD) and, in particular, demonstrate the validity of the FDT for each of them. We will conclude with some remarks on higher–order algorithms.

2. Theory of Fokker–Planck Processes

We consider a Markov process with continuous state space, denoting the state space variable with x. Usually, the state space is multi–dimensional; however, for the ease of notation and discussion we will focus on the one–dimensional case — the generalization to the multi–dimensional case is straightforward. The time variable will be denoted

with t. In contrast to the standard MC method (see contribution A. Milchev), where one considers a Markov *chain* with discrete time, we are here concerned with the case of continuous time. This means that the Markov property, which states that the process has no memory (i. e. the future behavior depends only on the present state, but not on the previous history), must hold for *arbitrarily small* time intervals. This is a very strong property, which, together with the continuousness of x, allows us to derive an equation of motion for the so–called "propagator" $P(x,t|x_0,t_0)$, i. e. the conditional probability density for the event that the process is in state x at time t if it was at x_0 at some earlier time $t_0 < t$. This equation of motion, the generalized FPE, is a partial differential equation involving derivatives with respect to both x and t; the aim of what follows (the so–called Kramers–Moyal expansion [15]) is to establish how its coefficients are related to the short–time behavior of the process (or of $P(x,t|x_0,t_0)$).

We begin by noting a few elementary properties of $P(x,t|x_0,t_0)$, normalization,

$$\int dx P(x,t|x_0,t_0) = 1, \tag{1}$$

the initial condition,

$$P(x,t_0|x_0,t_0) = \delta(x-x_0), \tag{2}$$

and the Chapman–Kolmogorov equation

$$P(x,t|x_0,t_0) = \int dx_1 P(x,t|x_1,t_1) P(x_1,t_1|x_0,t_0) \tag{3}$$

for times $t_0 < t_1 < t$. This latter equation simply states that by summing over all intermediate states x_1, one gets the full probability to go from x_0 to x. The factorization within the integral expresses the Markov property, i. e. the statistical independence of the future from the past.

It is obvious that $P(x,t|x_0,t_0)$ at short times ($\tau := t - t_0 > 0$ small) is very sharply peaked (see also Eq. 2). This means in turn that the moments

$$\begin{aligned} \mu_n(t;x_0,t_0) &:= \left\langle (x-x_0)^n \right\rangle (t,t_0) \\ &= \int dx\ (x-x_0)^n\, P(x,t|x_0,t_0) \end{aligned} \tag{4}$$

(mean displacement, mean square displacement, etc.) will also be very small for small τ (except, of course, for μ_0, which is trivially identical to unity for all times). We therefore write for $n \geq 1$

$$\left\langle (x-x_0)^n \right\rangle (t_0+\tau,t_0) = n! D^{(n)}(x_0,t_0)\tau + o(\tau), \tag{5}$$

where $o(\tau)$ denotes terms of order *higher* than linear, while the $D^{(n)}$ are the so–called Kramers–Moyal coefficients. Here we have assumed that all moments exist (the theory is only valid for such processes). However, it is well–known that in this case the probability density P is uniquely determined by its moments. The standard proof of this fact relies on the Fourier transform of P, whose Taylor coefficients are just the moments. Formally, this one–to–one correspondence is directly expressed via

$$P(x,t|x_0,t_0) = \sum_{n=0}^{\infty} \left(-\frac{\partial}{\partial x}\right)^n \delta(x-x_0)\frac{1}{n!}\mu_n(t;x_0,t_0), \tag{6}$$

as is easily verified by taking the nth moment of both left–hand and right–hand side, using partial integration.

We now consider the Chapman–Kolmogorov equation, Eq. 3, for the special case that the intermediate time t_1 is very close to the final time t, $t_1 = t - \tau$. We then insert the expansion Eq. 6 for the propagator corresponding to the short time interval τ:

$$\begin{aligned} P(x,t|x_0,t_0) &= \int dx_1 \sum_{n=0}^{\infty} \left(-\frac{\partial}{\partial x}\right)^n \delta(x-x_1)\frac{1}{n!}\mu_n(t;x_1,t-\tau) \\ &\quad P(x_1,t-\tau|x_0,t_0) \\ &= \sum_{n=0}^{\infty} \left(-\frac{\partial}{\partial x}\right)^n \frac{1}{n!}\mu_n(t;x,t-\tau)P(x,t-\tau|x_0,t_0) \end{aligned} \tag{7}$$

or (after subtraction of the $n=0$ term)

$$\begin{aligned} &\frac{1}{\tau}\left[P(x,t|x_0,t_0) - P(x,t-\tau|x_0,t_0)\right] \\ &= \frac{1}{\tau}\sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x}\right)^n \frac{1}{n!}\mu_n(t;x,t-\tau)P(x,t-\tau|x_0,t_0). \end{aligned} \tag{8}$$

Within linear order in τ we can write

$$\begin{aligned} \mu_n(t;x,t-\tau) &= n!D^{(n)}(x,t-\tau)\tau + o(\tau) \\ &= n!D^{(n)}(x,t)\tau + o(\tau) \end{aligned} \tag{9}$$

and

$$P(x,t-\tau|x_0,t_0) \approx P(x,t|x_0,t_0). \tag{10}$$

Inserting these expressions into Eq. 8, we arrive in the limit $\tau \to 0$ at the *generalized FPE*

$$\frac{\partial}{\partial t}P(x,t|x_0,t_0) = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x}\right)^n D^{(n)}(x,t)P(x,t|x_0,t_0), \tag{11}$$

which is often written in the shorthand notation

$$\frac{\partial}{\partial t} P(x,t|x_0,t_0) = \mathcal{L} P(x,t|x_0,t_0), \tag{12}$$

where $\mathcal{L}$ is the Fokker–Planck operator.

The processes can further be classified by the order at which the expansion stops. The simplest case is where all $D^{(n)}$ vanish. In this case there is obviously no dynamics at all; the process just stays at $x = x_0$. The next case is where only $D^{(1)}$ is nonzero, while all other $D^{(n)}$ vanish. This is the case of *deterministic* dynamics. Indeed, for deterministic dynamics we have a well–defined trajectory $x(t)$ starting at $x = x_0$ at time $t = t_0$, i. e. $P = \delta(x - x(t))$, and

$$\frac{\partial}{\partial t} P = -\frac{\partial}{\partial x} \dot{x}(t) \delta(x - x(t)). \tag{13}$$

In the case of Hamiltonian deterministic dynamics, the corresponding FPE is just the Liouville equation. Usual Fokker–Planck (or diffusion) processes correspond to the case where all $D^{(n)}$ are zero from $n = 3$ on, while $D^{(2)}$ is nonzero. In this case $D^{(2)}$ is called diffusion coefficient, while $D^{(1)}$ is called the drift coefficient which describes the deterministic part of the dynamics. Finally, there is the case that there are even nonzero coefficients for $n \geq 3$. In this case, the expansion actually does not stop at *any* finite order. This so–called Pawula theorem is an interesting mathematical result which can be proven rather straightforwardly using the positivity of P (for details, see Ref. [15]). This means that a truncation of the expansion after, say, the $n = 4$ term would result in solutions which are negative for some times and some regions of state space, which is of course completely unphysical. From now on, we will only consider usual Fokker–Planck processes.

We can thus define a Langevin simulation as a procedure which generates stochastic trajectories for such a process with a discretization time step τ. The physics is then specified by the details of $D^{(1)}$ and $D^{(2)}$. In the multi–variable case, $D^{(1)}$ is a vector, while $D^{(2)}$ is a second–rank tensor; they are related to the short–time behavior of the (vectorial) mean displacement, and the displacement covariance matrix, respectively.

The *Euler algorithm* is the simplest procedure, which is derived directly from what we have already learned:

$$\begin{aligned} \langle \Delta x_i \rangle &= D_i^{(1)}(x,t)\tau + o(\tau) \\ \langle \Delta x_i \Delta x_j \rangle &= 2D_{ij}^{(2)}(x,t)\tau + o(\tau) \\ \langle (\Delta x)^n \rangle &= o(\tau) \quad n \geq 3; \end{aligned} \tag{14}$$

this is satisfied by the updating rule

$$x_i(t+\tau) = x_i(t) + D_i^{(1)}\,\tau + \sqrt{2\tau}\,r_i, \tag{15}$$

where the r_i are random variables with $\langle r_i \rangle = 0$ and $\langle r_i r_j \rangle = D_{ij}^{(2)}$ (all higher moments existing).

This latter requirement is usually rather easy to satisfy, since in many cases the diffusion tensor is either constant or diagonal, or both. The most demanding case is where $D_{ij}^{(2)}$ has nonzero off–diagonal elements, which moreover depend on the stochastic variables. This occurs for systems of Brownian particles with hydrodynamic interactions, where the stochastic displacements are highly correlated (see contribution by A. Ladd). In that case, one has to calculate the "square root" of a large matrix every single time step. This is conveniently done by the algorithm first developed by Ermak and McCammon [4]. For a recent large–scale application to the dynamics of a single polymer chain in solution, see Ref. [12]. In this lecture, we will not further discuss this case, and refer the interested reader to the original papers.

It should be emphasized that on the level of the Euler algorithm it is *not* necessary to use random numbers with a Gaussian distribution, although this is sometimes stated in the literature. As a matter of fact, the derivation of the theory has so far not used Gaussian distributions at all, but rather the properties written down in Eq. 14. From this, one directly sees that the requirement is to just satisfy these, and a uniform distribution, which is computationally more efficient, will be just as good [3].

Nevertheless, the Gaussian distribution does have a prominent role. The reason is that in the limit $\tau \to 0$ the stochastic term strongly dominates. Therefore, very many stochastic displacements have already taken place before the deterministic drift is felt. A sum of very many independent random variables with identical distribution is however Gaussian (central limit theorem). For this reason, the stochastic term is often called "Gaussian white noise" (the word "white" hints, roughly spoken, to the fact that the process is supposed to be Markovian on arbitrarily short time scales, or, in other words, that it does not have an intrinsic time or frequency scale). However, from the standpoint of implementation this means that "Gaussianity" is nothing the programmer should worry about — the procedure will rather produce it automatically.

The very special role of the Gaussian distribution is also seen from the fact that it actually is the exact solution of the FPE for the case that both $D^{(1)}$ and $D^{(2)}$ are constants, as is easily verified. In this case

the updating rule Eq. 15 is exact for arbitrarily large time steps (and, of course, there is no point in running such a simulation).

We now proceed by rewriting Eq. 15 as

$$\frac{1}{\tau}\left(x_i(t+\tau) - x_i(t)\right) = D_i^{(1)} + \sqrt{\frac{2}{\tau}} r_i. \tag{16}$$

Attempting to take the limit $\tau \to 0$ would generate an object which does not exist in the sense of conventional calculus. Indeed, the stochastic trajectories are only continuous, but not differentiable; this implies that the left hand side diverges (as is also seen from the $\tau^{-1/2}$ term on the right hand side). Therefore, one writes down the so–called *Langevin equation*

$$\frac{d}{dt} x_i = D_i^{(1)} + f_i(t), \tag{17}$$

with deterministic drift $D_i^{(1)}$ and Gaussian white noise f_i, based on the understanding that this means nothing but a formal way of writing down the Euler updating rule, Eq. 15. The requirements on the random numbers r_i then translate into

$$\langle f_i \rangle = 0 \tag{18}$$

and

$$\langle f_i(t) f_j(t') \rangle = 2 D_{ij} \delta(t - t'). \tag{19}$$

This specification is needed for consistency; it makes sure that integrating the Langevin equation over a small time interval τ just recovers the Euler rule (up to irrelevant terms of order $o(\tau)$), such that the fundamental properties Eq. 14 are satisfied. Similarly, one must specify the higher–order moments of f_i such that $\int_0^\tau dt\, f_i(t)$ is a Gaussian random variable (for details, see Ref. [15]).

One subtlety remains. As the Langevin equation is not an object of ordinary calculus, one obviously has to define what is meant by it. Unfortunately, the definition given above (the so–called *Ito interpretation*) is not the only definition used in the literature. Another common convention is the so–called *Stratonovich interpretation*. While Eqs. 17–19 are left unchanged, the Stratonovich interpretation uses a different prescription how to proceed from the Langevin equation to the updating rule at finite time step. This prescription is inspired by the idea that one should first interpret the Langevin equation in the sense of ordinary calculus, and take the limit of vanishing correlation time of the noise at the very end. One thus arrives at a different updating rule, which actually means a different stochastic process and a different FPE. Fortunately, the difference only occurs for the case of so–called *multiplicative noise*, where the noise strength (or the diffusion tensor)

depends on the stochastic variable (the "usual" case where the noise strength is constant is termed *additive noise*).

To understand the difference, let us thus consider the one–dimensional Langevin equation with multiplicative noise

$$\frac{d}{dt}x = F(x) + \sigma(x)f(t), \tag{20}$$

where F denotes the deterministic part, while $\langle f \rangle = 0$ and $\langle f(t)f(t') \rangle = 2\delta(t - t')$. Now, Ito and Stratonovich give different answers to the question how one should evaluate the stochastic term

$$(\Delta x)_{st} = \int_0^\tau dt\, \sigma(x(t))f(t). \tag{21}$$

While the Ito interpretation simply prescribes

$$(\Delta x)_{st} \to \sigma(x(0)) \int_0^\tau dt\, f(t), \tag{22}$$

resulting in

$$\left\langle \int_0^\tau dt\, \sigma(x(t))f(t) \right\rangle = 0, \tag{23}$$

the Stratonovich interpretation rather Taylor expands $\sigma(x)$ within the integral,

$$\begin{aligned} & \int_0^\tau dt\, \sigma(x(t))f(t) \qquad (24) \\ \to\ & \sigma(x(0)) \int_0^\tau dt\, f(t) + \frac{d\sigma}{dx} \int_0^\tau dt\, \Delta x(t) f(t) + \dots \\ =\ & \sigma(x(0)) \int_0^\tau dt\, f(t) + \sigma \frac{d\sigma}{dx} \int_0^\tau dt \int_0^t dt'\, f(t')f(t) + \dots, \end{aligned}$$

such that we now obtain an effective drift term (often called "spurious drift")

$$\left\langle \int_0^\tau dt\, \sigma(x(t))f(t) \right\rangle = \sigma \frac{d\sigma}{dx}\tau + o(\tau). \tag{25}$$

This means that the Ito interpretation leads to a FPE where only F occurs as drift term, while in the Stratonovich interpretation the drift term is $F + \sigma(d\sigma/dx)$.

3. The Fluctuation–Dissipation Theorem

So far, we have only studied the mathematics of Fokker–Planck processes. In statistical physics, these processes are most commonly used

to describe the fluctuations of a system in thermal equilibrium, or the relaxation from a non–equilibrium state into equilibrium. This has two important consequences: (i) Drift and diffusion coefficients cannot depend explicitly on time, since otherwise the requirement of time translational invariance were violated. (ii) The Boltzmann distribution

$$\rho(x) = Z^{-1} \exp\left(-\beta \mathcal{H}(x)\right), \tag{26}$$

where $\mathcal{H}(x)$ is the Hamiltonian of the system, $\beta = 1/(k_B T)$, T the absolute temperature, k_B Boltzmann's constant, and $Z = \int dx \, \exp\left(-\beta\mathcal{H}\right)$ the partition function, must be a stationary solution of the FPE,

$$\mathcal{L} \exp\left(-\beta\mathcal{H}\right) = 0. \tag{27}$$

This results in a relation between drift and diffusion coefficient; in what follows we will derive these relations for BD, SD, and DPD.

4. Common Simulation Methods

In BD, we consider a system of particles with coordinates $\vec{r}_i$, friction coefficients ζ_i, and diffusion coefficients D_i. We denote the (effective) potential which describes the interaction between the Brownian particles with U. This is the system Hamiltonian which governs the Boltzmann distribution. The forces are then given by $\vec{F}_i = -\partial U / \partial \vec{r}_i$. The BD algorithm is to simulate the Langevin equations

$$\frac{d}{dt}\vec{r}_i = \frac{1}{\zeta_i}\vec{F}_i + \vec{\delta}_i \tag{28}$$

$$\left\langle \vec{\delta}_i \right\rangle = 0 \tag{29}$$

$$\left\langle \vec{\delta}_i(t) \otimes \vec{\delta}_j(t') \right\rangle = 2D_i \overset{\leftrightarrow}{1} \delta_{ij}\delta(t-t'); \tag{30}$$

we here have focused on the case without hydrodynamic interaction, where the stochastic displacements are uncorrelated. One thus can read off

$$\mathcal{L} = -\sum_i \frac{\partial}{\partial \vec{r}_i} \frac{1}{\zeta_i} \vec{F}_i + \sum_i D_i \left(\frac{\partial}{\partial \vec{r}_i}\right)^2, \tag{31}$$

and Eq. 27 results in

$$\sum_i \frac{\partial}{\partial \vec{r}_i} \left[\frac{1}{\zeta_i} \frac{\partial \mathcal{H}}{\partial \vec{r}_i} - \beta D_i \frac{\partial \mathcal{H}}{\partial \vec{r}_i} \right] \exp(-\beta\mathcal{H}) = 0. \tag{32}$$

This is satisfied if the *Einstein relation*

$$D_i = \frac{k_B T}{\zeta_i} \tag{33}$$

holds.

This method can be easily combined with the standard MC method. The idea is simply to use the BD step as a MC *trial* move, and to accept or reject it by the standard Metropolis criterion. In order to do this correctly, on must of course satisfy the condition of detailed balance (see contribution by A. Milchev). This, in turn, requires to take into account that the trial move is *biased* ("force biased MC"), i. e. that the *a priori* probability for the reaction $\vec{r} \to \vec{r}'$ *differs* from that for $\vec{r}' \to \vec{r}$. These *a priori* probabilities are of course nothing but the probabilities to generate the appropriate random numbers. For these, one should take a Gaussian distribution in order to avoid zero values. One thus arrives at a modified Metropolis criterion where the Boltzmann factor is multiplied by the ratio of the Gaussian functions for the two reactions. In the limit of vanishing time step, the acceptance rate of this procedure tends to unity.

Stochastic dynamics is a simulation method where one starts from a Hamiltonian dynamical system, and augments the momentum equation (Newton's equation of motion) by a friction and a noise term. The development is most transparent, and most general, if we start from Hamilton's equations of motion

$$\frac{d}{dt}q_i = \frac{\partial \mathcal{H}}{\partial p_i} \tag{34}$$

$$\frac{d}{dt}p_i = -\frac{\partial \mathcal{H}}{\partial q_i} \tag{35}$$

where the q_i denote the generalized coordinates, and the p_i the generalized canonically conjugate momenta. We then add friction and noise,

$$\frac{d}{dt}q_i = \frac{\partial \mathcal{H}}{\partial p_i} \tag{36}$$

$$\frac{d}{dt}p_i = -\frac{\partial \mathcal{H}}{\partial q_i} - \zeta_i \frac{\partial \mathcal{H}}{\partial p_i} + \sigma_i f_i; \tag{37}$$

here ζ_i is again a friction coefficient (note that $\partial \mathcal{H}/\partial p_i$, for usual Cartesian coordinates, is nothing but the velocity), σ_i denotes the noise strength, while $\langle f_i \rangle = 0$ and $\langle f_i(t) f_j(t') \rangle = 2\delta_{ij}\delta(t - t')$. We can even allow that the friction constants ζ_i and the noise strengths σ_i depend on the coordinates q_i (but not on the momenta p_i!). For this system of Langevin equations we can again read off the Fokker–Planck operator,

$$\mathcal{L} = \mathcal{L}_H + \mathcal{L}_{SD}, \tag{38}$$

where the first part refers to the Hamiltonian part of the dynamics,

$$\begin{aligned} \mathcal{L}_H &= -\sum_i \frac{\partial}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i} + \sum_i \frac{\partial}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i} \\ &= -\sum_i \frac{\partial \mathcal{H}}{\partial p_i}\frac{\partial}{\partial q_i} + \sum_i \frac{\partial \mathcal{H}}{\partial q_i}\frac{\partial}{\partial p_i}, \end{aligned} \tag{39}$$

with

$$\mathcal{L}_H \exp\left(-\beta\mathcal{H}\right) = 0, \tag{40}$$

while the second part is due to friction and noise,

$$\mathcal{L}_{SD} = \sum_i \frac{\partial}{\partial p_i}\left[\zeta_i \frac{\partial \mathcal{H}}{\partial p_i} + \sigma_i^2 \frac{\partial}{\partial p_i}\right], \tag{41}$$

such that Eq. 27 results in

$$\sum_i \frac{\partial}{\partial p_i}\left[\zeta_i \frac{\partial \mathcal{H}}{\partial p_i} - \beta\sigma_i^2 \frac{\partial \mathcal{H}}{\partial p_i}\right] \exp\left(-\beta\mathcal{H}\right) = 0. \tag{42}$$

Hence the relation

$$\sigma_i^2 = k_B T \zeta_i \tag{43}$$

must hold. The temperature is thus controlled as the ratio between noise strength and friction.

As already discussed in the Introduction, this is a useful and convenient way to stabilize a standard MD simulation. In the limit of weak friction, $\zeta_i \to 0$, the dynamics does not differ very much from the original Hamiltonian dynamics. In that case, the system can be simulated by just taking a good integrator for Hamiltonian dynamics (usually the Verlet or leapfrog algorithm, see contribution by D. Rapaport), and adding friction and random force just to the deterministic force whenever the latter occurs. It should however be noted that, strictly spoken, in this case the use of uniform random numbers reduces the order of the algorithm from second (Verlet) to first (Euler). The reason is that an overall accuracy up to second order would require to accurately sample the moments of the noise up to fourth order, while uniform random numbers only sample the first and second moment correctly. On the other hand, this loss of accuracy is only minor for weak friction, since in this case the behavior is dominated by the deterministic part. The issue of higher–order integrators will be discussed below.

As also mentioned in the Introduction, the SD algorithm is useless for studying hydrodynamic phenomena. The reason is that Galilei invariance and momentum conservation are among the most important properties of hydrodynamics, and both are violated in SD. The overall

momentum is not conserved, and the algorithm dampens the *absolute* velocities, thus labeling the "laboratory frame" as special, which is of course unphysical. In a real Galilei invariant fluid, the internal friction (viscosity) rather dampens velocity gradients, i. e. *relative* velocities. More quantitatively, it is easy to see that the change from MD to SD corresponds, on the hydrodynamic scale, to a change from the usual incompressible Navier–Stokes equation to a modified incompressible Navier–Stokes equation: The usual term $\eta\nabla^2\vec{u}$ (η viscosity, $\vec{u}$ velocity flow field), which describes the amount of internal friction per unit volume, is changed to $\eta\nabla^2\vec{u} - n\zeta\vec{u}$, where n is the particle density. From this, one directly reads off a typical length scale $l = [\eta/(n\zeta)]^{1/2}$, which is the screening length beyond which hydrodynamic correlations are broken up. For more details, and a more formal derivation, see Ref. [2].

Dissipative particle dynamics (DPD) has been developed to cure this problem, and to simulate hydrodynamic phenomena in fluids on a mesoscopic scale. DPD, as it is usually described in the literature, consists of two parts: (i) Introduction of very soft interparticle potentials in order to facilitate a large time step, and (ii) introduction of a Galilei invariant thermostat, which is similar to SD, but dampens *relative* velocities, and applies the stochastic kicks to *pairs* of particles such that Newton's third law (i. e. momentum conservation) is satisfied. As the procedure is also completely local, it is therefore suitable for the description of (isothermal) hydrodynamics. Unfortunately, it is often not made sufficiently clear that these two parts are *completely unrelated*, i. e. that one can use the DPD thermostat with "conventional" hard potentials, and that one can go from a working MD code to DPD, just as one would go to SD. We will from now on exclusively focus on the thermostat aspect of DPD. As Espanol and Warren [5] have shown, the structure of the FDT for DPD is very similar to the SD case. A particularly useful application of the DPD thermostat, which is just presently being appreciated, is its use in *nonequilibrium* studies like the simulation of steady–state Couette flow. Nonequilibrium steady states are characterized by a constant nonzero rate of entropy production, usually showing up as viscous heat. This produced entropy must be removed from the system, and therefore such simulations are usually coupled to a thermostat (an alternative approach, which rather removes the entropy by a Maxwell demon, has recently been developed by Müller–Plathe [13]). Before the advent of DPD, it was a non–trivial problem to introduce the thermostat in such a way that it would not prefer a certain profile (so–called "profile–unbiased thermostats", see Ref. [6]). The DPD thermostat solves this problem in a very natural and straightforward way [16].

In practice, DPD simulations are done as follows: We first define two functions, $\zeta(r)$, the relative friction coefficient for particle pairs with interparticle distance r, and $\sigma(r)$, the noise strength for a stochastic kick applied to the same particle pair. We will show below that the FDT implies the relation

$$\sigma^2(r) = k_B T \zeta(r), \tag{44}$$

in close analogy to SD. The function has a finite range, such that only near neighbors are taken into account.

Defining $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j = r_{ij}\hat{r}_{ij}$, we then obtain the friction force on particle i by projecting the relative velocities on the interparticle axes:

$$\vec{F}_i^{(fr)} = -\sum_j \zeta(r_{ij}) \left[(\vec{v}_i - \vec{v}_j)\cdot\hat{r}_{ij}\right]\hat{r}_{ij}; \tag{45}$$

it is easy to see that the relation $\sum_i \vec{F}_i^{(fr)} = 0$ holds. Similarly, we get the stochastic forces along the interparticle axes:

$$\vec{F}_i^{(st)} = \sum_j \sigma(r_{ij})\,\eta_{ij}(t)\,\hat{r}_{ij}, \tag{46}$$

where the noise η_{ij} satisfies the relations $\eta_{ij} = \eta_{ji}$, $\langle\eta_{ij}\rangle = 0$, and $\langle\eta_{ij}(t)\eta_{kl}(t')\rangle = 2(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\delta(t-t')$, such that different pairs are statistically independent. As before, one easily shows $\sum_i \vec{F}_i^{(st)} = 0$. The equations of motion,

$$\frac{d}{dt}\vec{r}_i = \frac{1}{m_i}\vec{p}_i \tag{47}$$

$$\frac{d}{dt}\vec{p}_i = \vec{F}_i + \vec{F}_i^{(fr)} + \vec{F}_i^{(st)} \tag{48}$$

therefore indeed conserve the total momentum, as the conservative forces $\vec{F}_i$ satisfy Newton's third law. The Fokker–Planck operator can then be written as

$$\mathcal{L} = \mathcal{L}_H + \mathcal{L}_{DPD}, \tag{49}$$

where $\mathcal{L}_H$ again describes the Hamiltonian part with $\mathcal{L}_H \exp(-\beta\mathcal{H}) = 0$ (cf. Eq. 39), and $\mathcal{L}_{DPD}$ is given by

$$\begin{aligned}\mathcal{L}_{DPD} &= \sum_{ij}\zeta(r_{ij})\hat{r}_{ij}\cdot\frac{\partial}{\partial\vec{p}_i}\left[\hat{r}_{ij}\cdot\left(\frac{\partial\mathcal{H}}{\partial\vec{p}_i} - \frac{\partial\mathcal{H}}{\partial\vec{p}_j}\right)\right] \\ &- \sum_{i\neq j}\sigma^2(r_{ij})\left(\hat{r}_{ij}\cdot\frac{\partial}{\partial\vec{p}_i}\right)\left(\hat{r}_{ij}\cdot\frac{\partial}{\partial\vec{p}_j}\right)\end{aligned}$$

$$+ \sum_i \sum_{j(\neq i)} \sigma^2(r_{ij}) \left(\hat{r}_{ij} \cdot \frac{\partial}{\partial \vec{p}_i} \right)^2$$

$$= \sum_i \sum_{j(\neq i)} \hat{r}_{ij} \cdot \frac{\partial}{\partial \vec{p}_i} \left[\zeta(r_{ij}) \hat{r}_{ij} \cdot \left(\frac{\partial \mathcal{H}}{\partial \vec{p}_i} - \frac{\partial \mathcal{H}}{\partial \vec{p}_j} \right) \right.$$

$$\left. + \sigma^2(r_{ij}) \hat{r}_{ij} \cdot \left(\frac{\partial}{\partial \vec{p}_i} - \frac{\partial}{\partial \vec{p}_j} \right) \right]. \qquad (50)$$

In the stochastic term, we have first taken into account the off–diagonal terms (cross–correlations, which are actually anti–correlations between the neighbors). The prefactors for the diagonal terms are given by the sum of all the mean square noise strengths from all the neighbors. Applying this operator to $\exp(-\beta\mathcal{H})$, we find that the FDT is satisfied if $\sigma^2(r) = k_B T \zeta(r)$.

5. Higher–Order Algorithms

Beyond the simple Euler method, one can try to develop algorithms which are of higher order. For the case of *additive* noise, this can be done in a rather systematic fashion via *operator factorization*. Assuming that the Fokker–Planck operator does not explicitly depend on time, the formal solution of the FPE is $P = \exp(\mathcal{L}t)\,\delta(x - x_0)$, where $x(t = 0) = x_0$. Calculating the exponential operator is nothing but actually solving the FPE. In the interesting cases where simulations are required, this is of course impossible. However, if we can decompose $\mathcal{L}$ as $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$ in such a way that both $\exp(\mathcal{L}_1 t)$ and $\exp(\mathcal{L}_1 t)$ are known, then we may use the relation

$$\exp(\mathcal{L}t) = \exp(\mathcal{L}_1 t/2) \exp(\mathcal{L}_1 t) \exp(\mathcal{L}_1 t/2) + O(t^3). \qquad (51)$$

Each of the three propagations corresponds to an exact solution and, as such, can be cast into an exact updating procedure. Such a method is then accurate up to second order. As a matter of fact, each of the three updates must only be accurate up to second order, too. For example, we can use the decomposition

$$\mathcal{L} = \mathcal{L}_{det} + \mathcal{L}_{stoch} \qquad (52)$$

into deterministic and stochastic updates. Here $\exp(\mathcal{L}_{stoch} t)$ corresponds to the standard Gaussian propagator, while $\exp(\mathcal{L}_{det} t)$ is just a deterministic update, which can be handled by a conventional method for ordinary differential equations up to any desired order. It is even

possible to go up to fourth order, which is however somewhat cumbersome, since this involves the evaluation of higher-order derivatives of the interaction potential. For more details, see Ref. [7].

In the case of *multiplicative* noise, things become much more involved, because even for the pure stochastic update $\exp(\mathcal{L}_{stoch}t)$ there is no general closed solution of the FPE. A second–order algorithm has been developed (see, e. g., Ref. [14]); however, for the interesting case of hydrodynamic interactions the method becomes so complicated that it is practically not useful. The present author therefore agrees with Ref. [10], which recommends to just use the Euler scheme, perhaps combined with an extrapolation to zero time step.

References

1. Chandrasekhar, S.: 1943. *Rev. Mod. Phys.* **15**, 1.
2. Dünweg, B.: 1993. *J. Chem. Phys.* **99**, 6977.
3. Dünweg, B. and W. Paul: 1991. *Int. Journ. Mod. Phys. C* **2**, 817.
4. Ermak, D. L. and J. A. McCammon: 1978. *J. Chem. Phys.* **69**, 1352.
5. Espanol, P. and P. Warren: 1995. *Europhys. Lett.* **30**, 191.
6. Evans, D. J. and G. P. Morriss: 1990, *Statistical Mechanics of Non–Equilibrium Liquids.* London: Academic Press.
7. Forbert, H. A. and S. A. Chin: 2000. *Phys. Rev. E* **63**, 016703.
8. Forster, D.: 1975, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions.* Reading, Mass.: Benjamin–Cummings.
9. Gardiner, C. W.: 1985, *Handbook of Stochastic Methods for Physics, Chemistry and Natural Science.* Berlin: Springer–Verlag.
10. Greiner, A., W. Strittmatter, and J. Honerkamp: 1988. *J. Stat. Phys.* **51**, 95.
11. Lax, M.: 1966. *Rev. Mod. Phys.* **38**, 541.
12. Liu, B. and B. Dünweg: 2003. To appear in J. Chem. Phys.
13. Müller-Plathe, F.: 1999. *Phys. Rev. E* **59**, 4894.
14. Öttinger, H. C.: 1996, *Stochastic Processes in Polymeric Fluids: Tools and Examples for Developing Simulation Algorithms.* Berlin: Springer–Verlag.
15. Risken, H.: 1984, *The Fokker–Planck Equation.* Berlin: Springer–Verlag.
16. Soddemann, T., B. Dünweg, and K. Kremer. In preparation.
17. van Kampen, N.: 2001, *Stochastic Processes in Physics and Chemistry (revised edition).* Amsterdam: North–Holland.