HEIDELBERG UNIVERSITY

SEMINAR: THEORETICAL STATISTICAL PHYSICS

Langevin and Fokker-Planck equation

Gerrit Gerhartz

supervised by
Prof. Dr. Georg WOLSCHIN

June 9, 2023

Abstract

This summary is an outline of the talk given in the summer semester 2023 in the context of the statistical physics seminar by Prof. Dr. Wolschin in Heidelberg.

We give an introduction to the Fokker-Planck equation by an approximation to the Master equation with the Kramers-Moyal expansion. We give an argument why this approximation should be done by introducing Pawulas theorem and proofing it.

The Langevin equation is presented in a non-rigerous form, before formalizing it with the Wiener process and the notion of a stochastic differential equation. With Ito's lemma a mapping from a stochastic differential equation to the Fokker-Planck equation and vice versa is presented, where on sees that the two formalism basically describe the same systems.

In the end the Fokker-Planck formalism is used to analyze the Ornstein-Uhlenbeck process. It is shown that the process is stationary with a gaussian stationary function. Lastly it is presented, where the Ornstein-Uhlenbeck process is used in a modern physics application. The dynamics of the velocity distribution of heavy particles colliding with each other is a non equilibrium process and therefore a Fokker-Planck equation is used to describe it.

Contents

1 History **2**

2 Kramers-Moyal expansion **3**

2.1 Pawulas Theorem 4

3 Introduction to the Langevin Equation **6**

4 Stochastic processes **9**

4.1 Wiener Process 9

4.2 Stochastic differential equation 10

4.3 Ito's lemma and the Fokker-Planck equation 10

4.3.1 Formal proof: Ito's Lemma 12

5 Example processes **13**

5.1 Brownian motion 13

5.2 Ornstein-Uhlenbeck process 14

5.3 Application: Heavy Ion collisions 16

6 Conclusion **18**

References **19**

1 History

The theory of stochastic processes is a mathematical framework that allows for the modeling and analysis of random phenomena. One of the earliest examples of such a phenomenon is Brownian motion, which refers to the random movement of particles suspended in a fluid. The temperature of the fluid influences the motion of these particles, and the movement appears random because of the many interactions between the particles and the fluid molecules.

In 1827, Robert Brown first observed this phenomenon while examining pollen particles in water [1]. He noted that the particles did not follow predetermined trajectories but instead moved randomly. However, it wasn't until 1905 that Albert Einstein mathematically investigated the motion of these particles [2]. He modeled the water as consisting of individual particles and calculated the motion of the pollen particles. From his calculations, he could derive the random nature of the motion. His finding was experimentally confirmed by Jean Perrin in 1908, who was later awarded the Nobel Prize in 1926 "for his work on the discontinuous structure of matter" [3]. Einstein's work was also significant because it provided evidence for the theory of molecules and atoms, which was not widely accepted at that time.

In 1908, Paul Langevin learned of Einstein's work on Brownian motion and formalized it in the form of the Langevin equation [4] [5]. This equation is used to describe the motion of particles in a viscous fluid and has made it possible to formalize a whole category of physical and stochastic processes that utilize some form of random motion.

Adriaan Fokker developed the Fokker-Planck equation in his thesis in 1913 parallel to Max Planck as another formulation for solving stochastic problems [6]. This equation provides a broader context for solving Brownian motion problems and can also resolve mathematical ambiguities. Instead of solving for the trajectory for every molecule, the Fokker-Planck equation allows for the calculation of the probability distribution of particle positions over time. Therefore, a higher number of particles can be solved directly. Furthermore, we will show, that every Langevin equation can be equivalently formulated in terms of a Fokker-Planck equation, demonstrating the close relationship between the two equations.

In modern science the use of Langevin and Fokker-Planck equations is used a lot and has applications in statistical physics, financial sciences, machine learning, biology and more.

2 Kramers-Moyal expansion

In this section we derive the Fokker-Planck equation as an approximation to a master equation. To do this, the master equation is transformed into a partial differential equation, which can be approximated using the usual arguments of the order of the derivatives. We will start by reintroducing the master equation.

The master equation for a given Markov process (for given transition amplitudes $W_t(x|x')$) is:

$$\frac{\partial p(x,t)}{\partial t} = \int_{\mathbb{R}} \underbrace{W_t(x|x')p(x',t)}_{\text{gain}} - \underbrace{W_t(x'|x)p(x,t)}_{\text{loss}} \,\mathrm{d}x' \tag{1}$$

This equation describes the dynamics of the probability distribution $p(x,t)$ of a stochastic parameter x governing a Markov process. The time derivative of the probability distribution is given by a gain/loss structure like the Boltzmann equation. The first term in the integral can be interpreted as a probability flow from a point x' to x, and the second term is a probability flow away from x to x'. In the end the integral over these terms must be taken, since we have a continuous state space and every position needs to be hit.

This master equation is an integro-differential equation and is hard to solve analytically, as a consequence we want to transform it into a partial differential equation, to use the mathematical machinery, which is developed in PDEs. To do this the gain term in the master equation is Taylor expanded.

To make the Taylor expansion easier the transition amplitudes are redefined with $r := x - x'$ and $W'_t(r,x) = W_t(r+x|x)$, the master equation changes accordingly:

$$\frac{\partial p(x,t)}{\partial t} = \int_{\mathbb{R}} W'_t(r, x-r)p(x-r,t) - W'_t(r,x)p(x,t)\mathrm{d}r \tag{2}$$

The new transition rates W'_t can be understood to be the transition rates, by which a state x will deviate by r. The gain term is now Taylor expanded around x:

$$W'_t(r,x-r)p(x-r,t) = \sum_{n=0}^{\infty} \frac{(-r)^n}{n!}\frac{\partial^n}{\partial x^n} W'_t(r,x)p(x,t) = W'_t(r,x)p(x,t) + \sum_{n=1}^{\infty} \frac{(-r)^n}{n!}\frac{\partial^n}{\partial x^n} W'_t(r,x)p(x,t) \tag{3}$$

If this term is reinserted into the master equation, the constant order terms will drop out inside the integral, and we are left with:

$$\frac{\partial p(x,t)}{\partial t} = \int_{\mathbb{R}} \sum_{n=1}^{\infty} \frac{(-r)^n}{n!}\frac{\partial^n}{\partial x^n} W'_t(r,x)p(x,t)\mathrm{d}r \tag{4}$$

To get the probability distribution outside the integral, the integral can be exchanged with the sum, giving us now the PDE formulation of the master equation:

$$\frac{\partial p(x,t)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!}\frac{\partial^n}{\partial x^n}\left[p(x,t)\int_{\mathbb{R}} r^n W'_t(r,x)\mathrm{d}r\right] \tag{5}$$

Defining now the coefficients $\alpha_n(x)$ and reinserting the definition of the auxiliary variable r gives rise to

the so called Kramers-Moyal expansion [7] of the master equation:

$$\frac{\partial p(x,t)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} \left[p(x,t) \alpha_n(x) \right]; \quad \alpha_n(x) = \int_{\mathbb{R}} (x - x')^n W_t(x'|x) \mathrm{d}x' \tag{6}$$

In this form of a series expansion of the master equation, we can canonically define an approximation of the master equation, by cutting of after specific orders of derivatives. If we forget about every derivative order higher than two, we are left with the Fokker-Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x} \left[\alpha_1(x) p(x,t) \right] + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[\alpha_2(x) p(x,t) \right] \tag{7}$$

This equation now describes the approximate dynamics induced by the master equation. And indeed this approximation is really a good one. Because from mathematics we know the so called Pawula theorem [8]:

If a coefficient $\alpha_n = 0$ for one $n \geq 3$, then every coefficient $\alpha_i = 0, \forall i \geq 3$. And if a coefficient $\alpha_n \neq 0$ for one $n \geq 3$, then there are an infinity number of contributions.

A proof of this theorem can be found in the next subsection. The theorem basically means, that if we approximate the master equation by the Fokker-Planck equation we have two situations:

1. The coefficients α_n are zero for $n > 2$, then the Fokker-Planck equation is an exact solution to the problem.

2. There is an infinite amount of non-vanishing coefficients, here the Fokker-Planck equation is a good approximation.

In the end we have found that the dynamics of a master equation can be approximated by the Fokker-Planck equation, and in a lot of cases the Fokker-Planck equation is even an exact description of the system.

2.1 Pawulas Theorem

The proof of Pawulas theorem is a direct consequence of the Cauchy-Schwarz inequality in the case of probability theory:

$$\left[\int g(x) f(x) P(x) \mathrm{d}x \right]^2 \leq \int g^2(x) P(x) \mathrm{d}x \int f^2(x) P(x) \mathrm{d}x \tag{8}$$

If one now sets $f(x) = (x - x')^n$, $g(x) = (x - x')^{n+m}$ and $P(x) = W_t(x|x')$, then the inequality becomes an inequality in the Kramers coefficients given in the previous section:

$$\alpha_{2n+m}^2 \leq \alpha_{2n} \alpha_{2n+2m} \tag{9}$$

Now we insert the $n!$ from the Kramers-Moyal expansion into the definition of the coefficients and go from $\beta_n = \frac{\alpha_n}{n!}$. The inequality changes to be the following:

$$[(2n+m)! \beta_{2n+m}]^2 \leq (2n)!(2n+2m)! \beta_{2n} \beta_{2n+2m} \tag{10}$$

Now one argues two folds:

1. If $\beta_{2n} = 0$ then $\beta_{2n+m} = 0, \forall m > 1$
2. If $\beta_{2n+2m} = 0$ then $\beta_{2n+m} = 0, \forall m > 1$

The first point says, that every coefficient with index bigger than $2n$ is zero if one index $\beta_{2n} = 0$. Therefore the one direction of the argument is done.

Now we need to argue for indices smaller than $2n$. For this we use the second point. If one defines $r = n + m \geq 2$, with the ranges $r \geq 2$ and $r > m \geq 1$. The proposition in the second point changes to:

$$\alpha_{2r} = 0 \Rightarrow \alpha_{2r-m} = 0, \quad r > m \geq 1. \tag{11}$$

Recursively we can lower the index till $\alpha_n = 0$ with $n \geq 3$, for one $r \geq 2$ $\alpha_{2r} = 0$. Now we have lowered the index till α_2 and the statement is proven.

3 Introduction to the Langevin Equation

In this section we want to discuss an introductory model to describe the theory of Brownian motion. The situation is now, that we have bigger particles in a fluid consistent of molecules, which push the bigger particles around. The model, which we are introducing here, is not mathematically rigorous, and we make it more strict in the next section.

The idea behind the Langevin equation as a model to describe Brownian motion is the following: We want to generalize Newtons equation of motion in a way, to describe stochastic systems. To do this we will add a stochastic term into Newtons equation of motion as follows:

$$\underbrace{m\frac{\mathrm{d}v}{\mathrm{d}t} = -\xi v}_{\text{classic}} + \underbrace{\sigma\eta(t)}_{\text{stochastic}} \tag{12}$$

This is called the Langevin equation. Its classical part describes the motion of a particle in a fluid. The linear term in the differential equation describes the viscous force from the fluid with a drag coefficient ξ. (Stokes Law)

The added stochastic part of the equation describes a random force $\eta(t)$, which should describe the random pushes from the fluid molecules onto the particles. These random pushes of the fluid particles onto the bigger particles, gives arise to the Brownian motion of the bigger particles. The σ is now a measure of how large this random force is. The random force is defined implicitly by stating its statistical properties:

1. $\langle\eta(t)\rangle = 0$ the mean value of the random force is zero. This is a symmetry property, since we assume the pushes to come from each side equivalently

2. $\langle\eta(t)\eta(t')\rangle = 2\delta(t-t')$ the correlation function between the random force should be a delta distribution. This means that two pushes should be stochastically independent if they are happening at different times. Therefore, we assume the process to be a markovian process.

In the next step we want to analyze this model and gather statistical information about the Brownian motion.

In the case of the Langevin equation, one can give a formal solution

$$v(t) = e^{-\frac{t}{t_0}}\left[v_0 + \int_0^t \mathrm{d}s\, e^{\frac{s}{t_0}}\frac{\sigma}{m}\eta(s)\right], \tag{13}$$

where we defined a timescale $t_0 = \frac{m}{\xi}$. One should take this result with a grain of salt, because to show, that the formal solution is really a solution to the Langevin equation one needs to differentiate the integral. But because we don't know what mathematical object the random force is, it is not clear that the fundamental theorem of calculus holds. This and more issues are addressed in the end of this section.

Because we have only defined the random force implicitly by its statistical properties, the formal solution has not a direct mathematical meaning. To gather information about the model, we can only calculate statistical properties of the formal solutions. The first property is the mean velocity:

$$\langle v(t)\rangle = e^{-\frac{t}{t_0}}\left[v_0 + \int_0^t \mathrm{d}s\, e^{\frac{s}{t_0}}\frac{\sigma}{m}\langle\eta(s)\rangle\right] = v_0 e^{-\frac{t}{t_0}} \tag{14}$$

Here the expectation value is put inside the integral since only the random force has statistical properties. By definition the statistical part vanishes, and we are left with the classical part of the solution to the Langevin equation.

The correlation function between the velocities at two different times t and t' can also be calculated, by using the defining properties of the random force:

$$\langle v(t)v(t')\rangle = v_0^2 e^{-\frac{t+t'}{t_0}} + \left(\frac{\sigma}{m}\right)^2 e^{-\frac{t+t'}{t_0}} \int_0^t \mathrm{d}s \int_0^{t'} \mathrm{d}s'\, e^{\frac{s+s'}{t_0}} 2\delta(s-s') \tag{15}$$

$$= e^{-\frac{t+t'}{t_0}} \left(v_0^2 - \frac{\sigma^2}{m\xi}\right) + \frac{\sigma^2}{m\xi} e^{\frac{t'-t}{t_0}} \tag{16}$$

To understand the correlation function in more detail, we make some approximation arguments. First we assume the times $t + t' \gg t_0$ to be a lot larger than the timescale t_0. In this approximation the classical part of the Langevin equation can be neglected, and the correlation function is only the consequence of the statistical properties. In this case the first term in the correlation function vanishes, because of the negative exponent. Furthermore, we calculate the same time $t = t'$ correlation function (mean of the velocity squared):

$$\langle v(t)^2\rangle = \frac{\sigma^2}{m\xi} \tag{17}$$

In the next step we use this property of $\langle v(t)^2\rangle$ to make a connection between the random force and a statistical physics property namely the temperature. To do this, we use the equipartition theorem of statistical physics in one dimension and insert $\langle v^2\rangle = \frac{\sigma^2}{m\xi}$:

$$\frac{1}{2}m\langle v^2\rangle = \frac{1}{2}k_B T \quad \Rightarrow \quad \sigma^2 = \xi k_B T \tag{18}$$

Here T is the temperature and k_B the Boltzmann constant. Here we found a connection between a statistical quantity, the temperature T, with the random force strength σ. This is what one would expect, since we think of the Brownian motion as an emerging phenomenon caused by the kinetic energy of the fluid particles, namely the temperature.

In the next step the trajectories of the particles are investigated. Here the same approximation is made as in the last calculation $t \gg t_0$. As a consequence the Langevin equation is now

$$\xi v = \sigma\eta(t), \tag{19}$$

which can be solved by a formal solution:

$$x(t) = x_0 + \frac{1}{\xi}\int_0^t \mathrm{d}s\, \sigma\eta(s) \tag{20}$$

Similar to the velocity case, we are calculating the mean position

$$\langle x(t)\rangle = x_0 \tag{21}$$

and not the correlation function but the so-called mean squared difference:

$$\left\langle (x(t) - x_0)^2 \right\rangle = \frac{1}{\xi^2} \int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \, 2\sigma^2 \delta(t' - t'') = \frac{1}{\xi^2} 2\sigma^2 t \tag{22}$$

This quantity is a measure of how far away from the mean value the particles can be found. Here we see that this quantity rises linearly by the time, therefore the particles can be found further and further away from the mean position. Such a process is called a diffusion process. And if we identify the mean squared difference with the diffusion coefficient D

$$\left\langle (x(t) - x_0)^2 \right\rangle = 2Dt \Rightarrow D = \frac{\sigma^2}{\xi^2} = \frac{k_B T}{\xi}. \tag{23}$$

We get the Einstein relation for the diffusion coefficient, which he already calculated in [2].

Now that we have seen, how the random forces arises from the temperature and that the Brownian motion is a diffusion process, we state the mathematical problems, which arise in this formal calculation. Basically we don't know what kind of mathematical object the random force η is, but we use a lot of mathematical results from calculus, which we don't know hold in this case. To make this model more rigorous we want to switch out the notion of a random force by a mathematical well-defined object. In the next section this problem is answered.

4 Stochastic processes

In the last section we have seen, how one can use the Langevin equation to model stochastic systems like the system of Brownian motion. It was only a formal calculation since we have never defined the random force, we used. In this section the Langevin equation is made mathematical rigorous, by exchanging the notion of a random force by a stochastic process. The stochastic used here is the so-called Wiener process, which is defined in the next subsection.

4.1 Wiener Process

The Wiener Process $W_t \in \mathbb{R}$, $t \in \mathbb{R}_0^+$ is a real valued stochastic process with positive time parameter. Analogous to the random force it is defined implicitly by its mathematical properties:

1. $W_0 = 0$ The startpoint of the process is the origin
2. for $0 \leq t_0 < t_1 < t_2 < t_3$ the increments $W_{t_0} - W_{t_1}, W_{t_2} - W_{t_3}$ are stochastically independent
3. For $0 \leq s < t$ the increments $W_t - W_s \propto \mathcal{N}(0, t-s)$, this means the increments are normal distributed, with variance $t - s$
4. the paths are continuous in the time parameter

Point two in the definition is the generalization of the correlation function of the random force. And in point three we can see the mean value set to be zero, like in the random force case. Also, the variance should rise with longer time differences, which mimics the diffusion property we calculated with the Langevin equation.

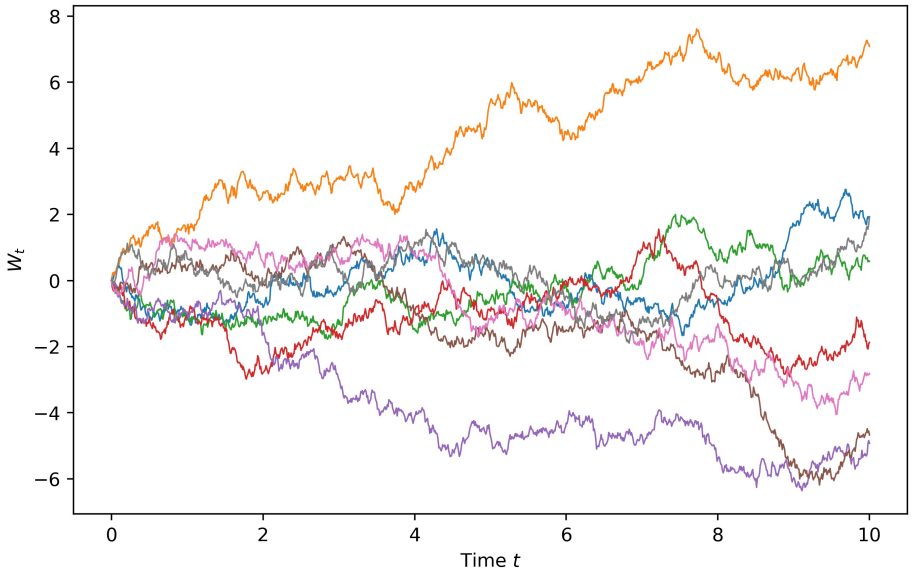

Figure 1: Simulation of the Wiener process

In figure 4.1 you can see how such a process might look like. There the trajectories were simulated by a python program.

4.2 Stochastic differential equation

With the definition of the Wiener process, we can generalize the notion of ordinary differential equations to stochastic differential equations, like we did in the Langevin case. A general stochastic differential equation of a stochastic parameter x with respect to the Wiener process W_t is written in the following way:

$$\mathrm{d}x = A(x,t)\mathrm{d}t + B(x,t)\mathrm{d}W_t \tag{24}$$

For a physicist it is probably more intuitive if one writes it in the form of a differential equation by dividing the $\mathrm{d}t$ out:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = A(x,t) + B(x,t)\frac{\mathrm{d}W_t}{\mathrm{d}t} \tag{25}$$

One should keep in mind, that this formulation as a differential equation should just be intuition, since the Wiener process almost nowhere differentiable. As an example: the Langevin equation can now be written in the form of a stochastic differential equation:

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = -\xi v + \sigma\sqrt{2}\frac{\mathrm{d}W_t}{\mathrm{d}t} \tag{26}$$

Here one can see that the random force is switched with the Wiener process.

As another side note: one should notice that these two equations are also mathematically not really rigorous but a shorthand notation for writing out the following integral equation:

$$x(t) - x(s) = \int_s^t A(x,\tau)\mathrm{d}\tau + \int_s^t B(x,t)\mathrm{d}W_t \tag{27}$$

Here one sees that the difference in the stochastic parameter x is given by a part, which is the Lebesgue integral over the function A and another part, which is an integral with respect to the Wiener process W_t of B. The second integral is a so-called stochastic integral (e.g. Ito-integral), and one needs to define it, to make finish the definition of a stochastic differential equation. But in favor of the scope of this summary this field is not further explained.

With the definition of the stochastic differential equation one could in principle go back to the calculation of the Langevin equation and see that the results are now mathematical rigorous.

4.3 Ito's lemma and the Fokker-Planck equation

In the next section we derive a connection between a general stochastic differential equations and the Fokker-Planck equation. The stochastic differential equation describes the particles trajectories, but if one would like to model the system as a whole one needs to find a differential equation, which describes the dynamics of the probability distribution of the system. In this part the dynamics of the probability distribution is derived from a general stochastic differential equation.

To make the connection between the probability distribution and the stochastic differential equation, we need a mathematical tool called Ito's Lemma [9]:

Now f is a function of the process x and the time t. The process x is given by:

$$\mathrm{d}x = A(x,t)\mathrm{d}t + B(x,t)\mathrm{d}W_t \tag{28}$$

Then the differential $\mathrm{d}f$ *is given as:*

$$\mathrm{d}f = \left[\frac{\partial f}{\partial t} + A(x,t)\frac{\partial f}{\partial x} + \frac{B^2(x,t)}{2}\frac{\partial^2 f}{\partial x^2}\right]\mathrm{d}t + B(x,t)\frac{\partial f}{\partial x}\mathrm{d}W_t \tag{29}$$

This is basically a generalization of the chain from calculus, now reformulated in the sense of probability theory. This can be seen more clearly if one writes it as a differential equation:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \left[\frac{\partial f}{\partial t} + A(x,t)\frac{\partial f}{\partial x} + \frac{B^2(x,t)}{2}\frac{\partial^2 f}{\partial x^2}\right] + B(x,t)\frac{\partial f}{\partial x}\frac{\mathrm{d}W_t}{\mathrm{d}t} \tag{30}$$

To introduce the probability distribution in the picture, first the Ito Lemma is used on a general time-independent function $f(x)$. Then the expectation of the differential equation is taken:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle f\rangle = \left\langle\frac{\mathrm{d}f}{\mathrm{d}t}\right\rangle = \underbrace{\left\langle\frac{\partial f}{\partial t}\right\rangle}_{=0} + \left\langle A(x,t)\frac{\partial f}{\partial x}\right\rangle + \left\langle\frac{B^2(x,t)}{2}\frac{\partial^2 f}{\partial x^2}\right\rangle \tag{31}$$

Here we used The Expectation value of $\langle dW_t\rangle = 0$ is zero per definition and the stochastic part therefore vanishes.

If one writes down the expectation value explicitly the probability distribution $p(x,t)$ can be identified:

$$\int_{\mathbb{R}}\mathrm{d}x\, f(x)\partial_t p(x,t) = \int_{\mathbb{R}}\mathrm{d}x\, A(x,t)\frac{\partial f}{\partial x}p(x,t) + \int_{\mathbb{R}}\mathrm{d}x\,\frac{B^2(x,t)}{2}\frac{\partial^2 f}{\partial x^2}p(x,t) \tag{32}$$

Here we used the fact, that the total derivative can be put inside the integral. Since we want to know about the dynamics of the probability distribution, we need an equation with the derivatives of the probability distribution. Consequently, we want to shift the derivatives of the function f onto the probability distribution p. This can be done with partial integration and the fact that $p(x,t), \partial_x p(x,t), \partial_x^2 p(x,t) \to 0$ with $x \to \pm\infty$:

$$\int_{\mathbb{R}}\mathrm{d}x\, f(x)\partial_t p(x,t) = \int_{\mathbb{R}}\mathrm{d}x\, f(x)\left[-\partial_x(A(x,t)p(x,t)) + \frac{1}{2}\partial_x^2(B^2(x,t)p(x,t))\right] \tag{33}$$

Since the function f(x) was arbitrary, we can infer with the fundamental lemma of calculus of variations, that the integrands of the equation above must be equal:

$$\partial_t p(x,t) = -\partial_x(A(x,t)p(x,t)) + \frac{1}{2}\partial_x^2(B^2(x,t)p(x,t)) \tag{34}$$

This is now the equation, which describes the dynamics of the probability distribution under a general stochastic differential equation. If we compare this equation with the Fokker-Planck equation (7), we can see a connection between the coefficients α_1, α_2 and the functions $A(x,t)$ and $B(x,t)$:

$$\alpha_1 = A(x,t), \quad \alpha_2 = B^2(x,t) \tag{35}$$

With this identification we have made a one to one connection between the systems, a stochastic differential equation can describe, and the systems a Fokker-Planck equation can describe. This means these two approaches to describe stochastic systems are equivalent. The difference is that a stochastic

differential equation describes trajectories in state space and the Fokker-Planck equation describes the dynamics of the system as a whole.

4.3.1 Formal proof: Ito's Lemma

In this subsection we give a formal non-rigorous proof to the Ito's lemma. Basically the statement is a direct consequence of a Taylor series and the definition of a Wiener process.

The proof starts with the function $f(x,t)$ of which we want to calculate the differential. We write $\mathrm{d}f$ as a Taylor series:

$$\mathrm{d}f = \frac{\partial f}{\partial t}\mathrm{d}t + \frac{1}{2}\frac{\partial f^2}{\partial t^2}\mathrm{d}t^2 + \ldots \frac{\partial f}{\partial x}\mathrm{d}x + \frac{1}{2}\frac{\partial f^2}{\partial x^2}\mathrm{d}x^2 + \ldots \tag{36}$$

Now we can insert the stochastic process for $\mathrm{d}x = A(x,t)\mathrm{d}t + B(x,t)\mathrm{d}W_t$. And we get:

$$\mathrm{d}f = \frac{\partial f}{\partial t}\mathrm{d}t + \frac{1}{2}\frac{\partial f^2}{\partial t^2}\mathrm{d}t^2 + \cdots + \frac{\partial f}{\partial x}\left(A(x,t)\mathrm{d}t + B(x,t)\mathrm{d}W_t\right) \tag{37}$$

$$+\frac{1}{2}\frac{\partial f^2}{\partial x^2}\left(A(x,t)^2\mathrm{d}t^2 + 2A(x,t)B(x,t)\mathrm{d}t\mathrm{d}W_t + B(x,t)^2\mathrm{d}W_t^2\right) + \ldots \tag{38}$$

Now in the Limit $\mathrm{d}t \to 0$ the terms $\mathrm{d}t^2$ and $\mathrm{d}t\mathrm{d}W_t$ go to zero faster than $\mathrm{d}W_t^2$, which is of Order $\mathrm{d}t$. The higher order terms are set to zero. And if we use that $\mathrm{d}W_t^2 = \mathrm{d}t$, by the definition of the Wiener process, we get the result:

$$\mathrm{d}f = \left[\frac{\partial f}{\partial t} + A(x,t)\frac{\partial f}{\partial x} + \frac{B^2(x,t)}{2}\frac{\partial^2 f}{\partial x^2}\right]\mathrm{d}t + B(x,t)\frac{\partial f}{\partial x}\mathrm{d}W_t \tag{39}$$

5 Example processes

Now that we have introduced the mathematical tools, which describes stochastic differential equations and build the connection between stochastic differential equations and the Fokker-Planck equation, we calculate the Langevin equation from the point of view of the Fokker-Planck equation. In the SDE formulation the Langevin equation was given as follows:

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = -\xi v + \sigma\sqrt{2}\frac{\mathrm{d}W_t}{\mathrm{d}t} \tag{40}$$

Before we investigate the full equation in the Ornstein-Uhlenbeck process we discuss the limit where the inertial term in the SDE can be neglected. This Situation is called Brownian motion.

5.1 Brownian motion

As already stated, the case of Brownian motion arises from the Langevin equation if the inertial term of the equation can be neglected and only the stochastic part matters. This is analogous to the non-rigerous case, where we demanded that the time should be a lot bigger than the timescale t_0. If we neglect the inertial term the differential equation becomes:

$$\xi\frac{\mathrm{d}x}{\mathrm{d}t} = \sqrt{2}\sigma\frac{\mathrm{d}W_t}{\mathrm{d}t} \Rightarrow \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\sqrt{2}\sigma}{\xi}\frac{\mathrm{d}W_t}{\mathrm{d}t} \tag{41}$$

In the next step we want to reformulate the equation into Fokker-Planck form. If we compare (41) with the general stochastic process we get the functions:

$$A(x,t) = 0, \quad B(x,t) = \frac{\sqrt{2}\sigma}{\xi} \tag{42}$$

With this identification we can calculate the Fokker-Planck coefficients for this specific situation:

$$\alpha_1 = 0, \quad \alpha_2 = \frac{2\sigma^2}{\xi^2} \tag{43}$$

With these coefficients, the corresponding Fokker-Planck equation is given as:

$$\frac{\partial p(x,t)}{\partial t} = \frac{\sigma^2}{\xi^2}\frac{\partial^2 p(x,t)}{\partial x^2} \tag{44}$$

This is now the equation, which defines the dynamics of the probability distribution of the system. And indeed we got back the diffusion properties, we calculated in the non-rigerous case. If we compare this, to the general diffusion equation:

$$\frac{\partial p(x,t)}{\partial t} = D\frac{\partial^2 p(x,t)}{\partial x^2}, \tag{45}$$

we can identify the diffusion coefficient directly as $D = \frac{\sigma^2}{\xi^2}$, which coincides with the result we have gotten in the first section.

The diffusion property of the specific Fokker-Planck equation, can be understood by investigating

the solutions to this system. The model can be solved by a fundamental solution and is given as:

$$p(x,t) = \frac{1}{\sqrt{4\pi D t}} \int_{\mathbb{R}} \mathrm{d}x \, \exp\left[-\frac{(x-y)^2}{4Dt}\right] f(x) \tag{46}$$

Where $f(x) = p(x,0)$ is the initial condition of the probability distribution. Here one can see the gaussian property of the solution and one sees that the variance rises with further time, which directly leads to the diffusion property.

Now that the simpler Brownian motion case has been calculated by the Fokker-Planck equation, the whole Langevin problem is discussed in the next subsection.

5.2 Ornstein-Uhlenbeck process

In this subsection the whole Langevin problem is calculated with the Fokker-Planck formalism. This setting is also called the Ornstein-Uhlenbeck process, where the SDE is given as:

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = -\frac{\xi v}{m} + \frac{\sqrt{2}\sigma}{m}\frac{\mathrm{d}W_t}{\mathrm{d}t} \tag{47}$$

Analogous to the Brownian motion case, we can identify the functions from the general stochastic differential equation:

$$A(v,t) = -\frac{\xi v}{m}, \quad B(v,t) = \frac{\sqrt{2}\sigma}{m} \tag{48}$$

Such that the Fokker Planck formulation of the problem is the following:

$$\frac{\partial p(v,t)}{\partial t} = \frac{\xi}{m}\frac{\partial}{\partial v}\left[v p(v,t)\right] + \frac{\sigma^2}{m^2}\frac{\partial^2 p(v,t)}{\partial v^2} \tag{49}$$

In the next step we want to calculate, first if the process is stationary and second how the stationary solution looks like. To calculate the stationary distribution, one can set the partial derivative of the probability distribution with respect to time, equal to zero and is left with the following differential equation:

$$0 = \frac{\partial}{\partial v}\left[v p(v)\right] + \frac{\sigma^2}{m\xi}\frac{\partial^2 p(v)}{\partial v^2} \tag{50}$$

If this equation has a solution we know that the process is stationary. This equation can be integrated to get the following equation:

$$C = v p(v) + \frac{\sigma^2}{m\xi}\frac{\partial p(v)}{\partial v} \tag{51}$$

Here we needed to introduce an integration constant C, which can be set zero by the arguments in the following. Now we use that the solution to this equation is a probability distribution, therefore we know:

$$p(v) \geq 0, \forall v \in \mathbb{R}; \quad \int_{\mathbb{R}} \mathrm{d}v \, p(v) = 1 \tag{52}$$

We also know that the function $p(v)$ is differentiable at every point in $\mathbb{R}$, therefore we know the function is bounded. With the condition, that the function is bounded, and the probability distribution is symmetric around v, we know that the integration constant must be $C = 0$.

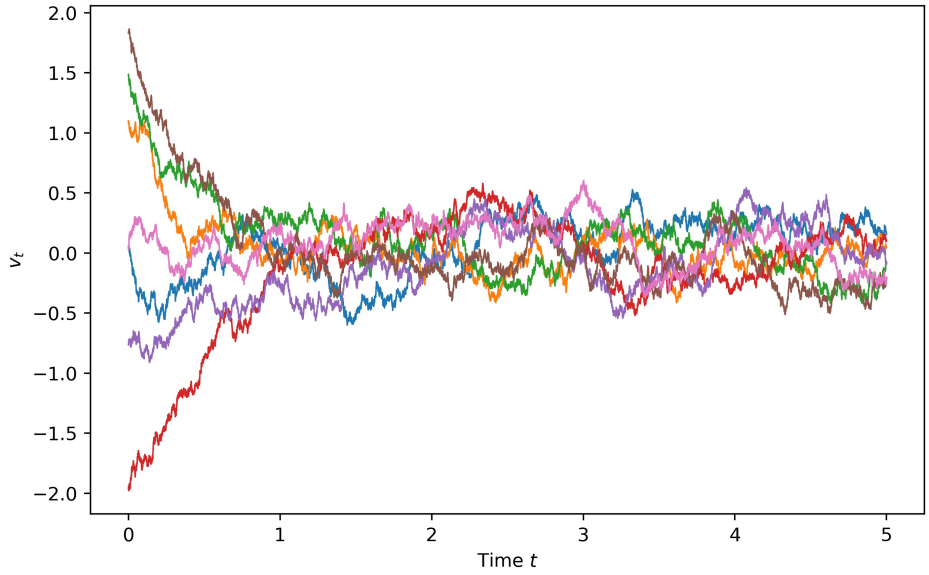

Figure 2: Simulation of the Ornstein-Uhlenbeck process

Therefore, our differential equation simplifies to:

$$\frac{\partial p(v)}{\partial v} + \frac{m\xi}{\sigma^2} v p(v) = 0 \tag{53}$$

The solution to this differential equation is a Gaussian:

$$p(v) = \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left(-\frac{v^2}{2\gamma^2}\right), \quad \gamma = \frac{\sigma}{\sqrt{m\xi}} \tag{54}$$

therefore we found, that the Ornstein-Uhlenbeck process is a stationary process with a gaussian stationary function. And essentially this is what one would expect, because we know from statistical physics is distributed via the Maxwell-Boltzmann distribution, and if we investigate the stationary solution we see, that they coincide. The stationary solution to the Ornstein-Uhlenbeck process is the Maxwell-Boltzmann distribution.

The stationarity can be seen in figure 5.2 where this process is simulated with various initial conditions. For completeness the general solution for the Ornstein-Uhlenbeck process is written here in the form of a Greens function.

$$P(v,t|v',t') = \sqrt{\frac{\theta}{2\pi D(1-e^{-2\theta(t-t')})}} \exp\left[-\frac{\theta}{2D}\frac{(v-v'e^{-\theta(t-t')})^2}{1-e^{-2\theta(t-t')}}\right] \tag{55}$$

Here $\theta = \frac{\xi}{m}$ and $D = \frac{\sigma^2}{\xi^2}$. This result is used in the next section, where the Ornstein-Uhlenbeck process is used to model velocity distributions of colliding heavy particles.

5.3 Application: Heavy Ion collisions

Now, we want to introduce an application to stochastic processes and specifically the Ornstein-Uhlenbeck-process. In the case of relativistic heavy ion collisions, the velocity distribution of the colliding particles after the collision can be modeled by a distribution following an Ornstein-Uhlenbeck-process. For this we follow the work done by Hoelck and Wolschin in [10].

The Situation is the following: We have heavy ion particles which a colliding into one another. Because they are colliding relativistically, a lot of new particles are generated. The kinetic energy of the colliding particles decreases as side effect therefore we can expect the distribution of the velocities parallel to the beam axis to be in a non-equilibrium state. Since we expect the velocities to have a small component in the transverse direction, we assume the velocities in this direction to be in equilibrium.

The idea is now to describe the dynamics of the longitudinal velocities by a Fokker-Planck equation which follows an Ornstein-Uhlenbeck process. Since the velocity of a particle in special relativity is a bounded parameter, we can not use the standard notion of stochastic processes to describe the velocity directly. The Wiener process is an unbounded process and can therefore only describe non-bounded quantities. We need to find a quantity describing the velocity of a relativistic particle, which is not bounded and additive.

The rapidity

$$y = \frac{1}{2}\ln\left(\frac{E + p_{\|}}{E - p_{\|}}\right) \tag{56}$$

is a parameter with the demanded properties and can therefore be used to be described by a stochastic differential equation. (Here $p_{\|}$ is the momentum parallel to the beam axis and E the energy of the particle) The Fokker-Planck equation of the Ornstein-Uhlenbeck process describing the rapidity distribution $R(y,t)$ is now given by

$$\frac{\partial R(y,t)}{\partial t} = \frac{1}{\tau_y}\frac{\partial}{\partial y}\left[yR(y,t)\right] + \frac{\partial^2}{\partial y^2}\left[D_y R(y,t)\right] \tag{57}$$

Here one can see the linear drift term with parameter $\frac{1}{\tau_y}$ and the Diffusion coefficient D_y. If one solves the equation with the initial condition $R(y, t=0) = \delta(y \pm y_b)$, which describes two particle bunches with the same velocity magnitude and opposite direction, one gets

$$R(y,t) = \left[\sqrt{2\pi}2\sigma_y(t)\right]^{-1}\left(\exp\left[-\frac{\left(y + y_b\exp\left(-\frac{t}{\tau_y}\right)\right)^2}{2\sigma_y^2(t)}\right] + \exp\left[-\frac{\left(y - y_b\exp\left(-\frac{t}{\tau_y}\right)\right)^2}{2\sigma_y^2(t)}\right]\right), \tag{58}$$

with the variance:

$$\sigma_y^2(t) = D_y\tau_y\left[1 - \exp\left(-\frac{2t}{\tau_y}\right)\right] \tag{59}$$

This solution describes two gaussian functions, which drift towards zero and diffuse. The Drifting behavior can be seen in the exponential function in the exponent of the Gaussian function. If the time gets larger and larger $y_b\exp\left(-\frac{t}{\tau_y}\right)$ gets smaller and smaller. This is the behavior one would expect. Since energy is converted into particles the kinetic energy of the particles gets lower and therefore the velocities get smaller.

In the end, this is a model, which needs to be tested with experimental data. If one fits the model to

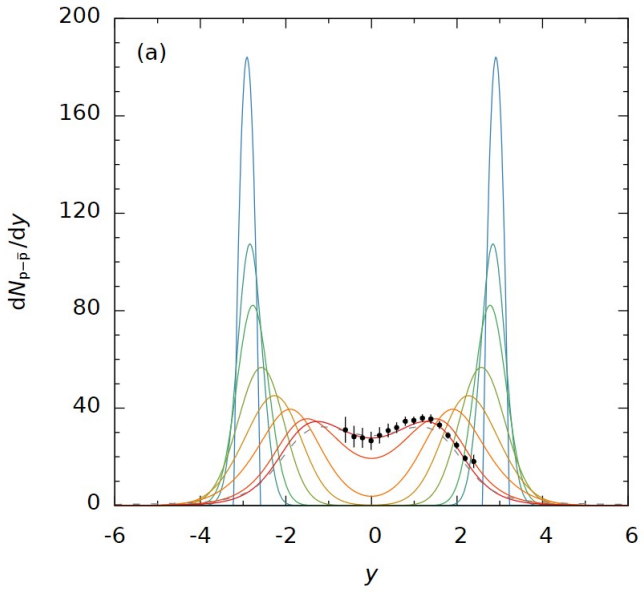

Figure 3: Time evolution of net-proton rapidity distribution function in $^{208}\mathrm{Pb}$ collisions [10]. Here the numerical solution to the real stochastic differential equation is plotted and not the approximation by an Ornstein-Uhlenbeck process

the data provided by a particle accelerator one can see that the data can be described well with this kind of stochastic process. This can be seen in figure 5.3. In the color gradient one sees the dynamical evolution of the initial rapidity distributions in bright blue. The initial conditions are more sophisticated than a simple delta function, and for the scope of this summary this is not described further. For more detail you can read up in [10]. Another point one should note is, that in the paper [10] the stochastic process is not exactly an Ornstein-Uhlenbeck process but a more sophisticated equation, which can be approximated by an Ornstein-Uhlenbeck process.

At larger times the data points line up with the model and can conclude that a diffusion coefficient with linear drift can accurately describe this kind of data.

6 Conclusion

We have started with a short historical introduction into the field of Brownian motion and stochastic differential equations. Before we derived the Fokker-Planck equation as an approximation to the Master equation. The Master equation was transformed from the integro-differential form into the form of a partial differential equation, the Kramers-Moyal expansion. Here one could canonically define an approximation, where the Fokker-Planck equation arrived. With the Pawulas theorem this approximation was justified, and we saw that the Fokker-Planck approximation was in a lot of cases even an exact description of the situation.

The Langevin equation was introduced in a non-rigerous form with the notion of a random force. We used this model to calculate the Einstein relation to the diffusion coefficient and saw that the Langevin equation gives rise to a diffusion process. Because this approach was mathematical not well-defined, we used the notion of stochastic processes and especially the Wiener process to exchange the random force with a mathematical well-defined object. This also gave arise to the formulation of stochastic problems as stochastic differential equations, which are generalizations of ordinary differential equations, and can be used to model a whole range of interesting systems. Thereafter, Ito's lemma was used to establish a connection between stochastic differential equation and the dynamics of the probability distribution by the Fokker-Planck equation. We saw that the models can be translated into each other, therefore they can be used to model the same systems. The differences are that the SDE models trajectories of stochastic parameters and the Fokker-Planck formulation directly models the dynamics of the whole system by describing its probability density.

In the last step the Langevin problem was analyzed with the Fokker-Planck formulation. We saw that the Brownian motion case gave arise to a diffusion PDE. We showed that the Ornstein-Uhlenbeck process was stationary before describing the general solution to the problem in the Form of a Greens function. And at the end the Ornstein-Uhlenbeck process was used to model the non-equilibrium dynamics of the velocity distribution in heavy ion collisions.

To conclude one can say that the theory of Brownian motion gave arise to a rich and powerful theory, using stochastic differential equations and the Fokker-Planck equation. And today it is indispensable in modern research in physics, biophysics, finance and even machine learning.

References

[1] B. Robert. Xxvii. a brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine*, 4(21):161–173, 1828.

[2] A. Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[3] J. Perrin. Le mou-vement brownien et la réalité moleculaire. *Ann. Chimi.Phys.*, 18:5–114, 1910.

[4] P. Langevin. Sur la théorie du mouvement brownien. *C. R. Acad. Sci.* (*Paris*), (146):530–533, 1908.

[5] Don Lemons and Anthony Gythiel. Paul langevin's 1908 paper "on the theory of brownian motion". *American Journal of Physics - AMER J PHYS*, 65:1079–1081, 01 1997.

[6] A. D. Fokker. Die mittlere energie rotierender elektrischer dipole im strahlungsfeld. *Annalen der Physik*, 348(5):810–820, 1914.

[7] H.A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[8] R. Pawula. Generalizations and extensions of the fokker- planck-kolmogorov equations. *IEEE Transactions on Information Theory*, 13(1):33–41, 1967.

[9] Kiyosi Itô. Stochastic integral. *Proceedings of the Imperial Academy*, 20(8):519 – 524, 1944.

[10] Johannes Hoelck and Georg Wolschin. Baryon stopping as a relativistic markov process in phase space. *Physical Review Research*, 2(3), sep 2020.