M.Phys. C6 Theory Option
Lecture Notes for 2020-2021

Lectures given by Sid Parameswaran in Michaelmas 2020
Lecture Notes written by F.H.L. Essler with contributions by Sid Parameswaran
The Rudolf Peierls Centre for Theoretical Physics
Oxford University, Oxford OX1 3PU, UK

September 23, 2020

These notes cover Part V of the lectures (Random Systems and Stochastic Processes), which is only part of M.Phys. C6 Theory Option.
Please report errors and typos to sid.parameswaran@physics.ox.ac.uk.
©2015 F.H.L. Essler.

Contents

V Random Systems and Stochastic Processes **3**

1 Random Variables **3**
1.1 Some Definitions . . . 3
1.2 Discrete-time random walk . . . 4
1.3 The central limit theorem . . . 5

2 Stochastic Processes: Definitions **6**

3 Markov Processes **6**
3.1 Examples of Markov Processes . . . 7
3.2 Markov Chains . . . 7

4 Brownian Motion **9**
4.1 Langevin Equation . . . 9
4.2 Fokker-Planck Equation . . . 10
4.3 Diffusion Equation . . . 13

5 Brownian motion via the Path Integral **14**
5.1 Homework Questions 22-26 . . . 15

Some general remarks:

These notes aim to be self-contained. Homework questions are marked in red, and are placed at appropriate positions in the text, i.e. to work them out you will require only the preceeding material. Passages marked in blue give details on derivations we don't have time to go through in the lectures, or present material that goes beyond the core of the course. In some cases this material will be very useful for particular homework problems. All of the material covered in the course can be found in some form or other in a variety of books. These is no book that covers everything. Some useful references for this section are

- Basic ideas of probability, kinetic theory, etc.

  M. Kardar, *Statistical Physics of Particles*, Cambridge.

- Stochastic Processes

  N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland/Elsevier.

Part V
Random Systems and Stochastic Processes

In nature there are many phenomena, in which some quantity varies in a random way. An example is *Browninan motion*, which refers to the motion of a small particle suspended in a fluid. The motion, observed under a microscope, looks *random*. It is hopeless to try to compute the position in detail, but certain *average*

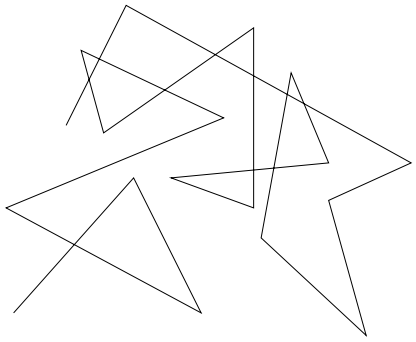


Figure 1: Cartoon of a path of a particle undergoing Brownian motion.

features obey simple laws. Averaging over a suitable time interval is difficult and one therefore replaces time averaging of a single irregularly varying function of time by averaging over an *ensemble of functions*. The latter must of course be chosen in such a way that the two results agree

$$\text{time average} \longrightarrow \text{ensemble average}. \tag{1}$$

1 Random Variables

A random variable is an object X defined by

1. a set $\{x_j\}$ of possible values (either discrete or continuous);
2. a probability distribution $P_X(x_j)$ over this set

$$P_X(x_j) \geq 0 \ , \qquad \sum_j P_X(x_j) = 1 \text{ (this becomes an integral in the continuous case.)}. \tag{2}$$

Example: Let X be the number of points obtained by casting a die

$$\{x_j\} = \{1, 2, 3, 4, 5, 6\} \ , \qquad P_X(x_j) = \frac{1}{6}. \tag{3}$$

1.1 Some Definitions

The probability distribution $P_X(x)$ can be characterized by the moments of X

$$\boxed{\langle X^n \rangle \equiv \int dx \ x^n \ P_X(x).} \tag{4}$$

The average value of X is the first moment $\langle X \rangle$, while the variance is $\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$.

The Fourier transform of $P_X(x)$ is called the characteristic function

$$\boxed{\phi_X(k) = \int dx\ P_X(x) e^{ikx} = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle \equiv \langle e^{ikX} \rangle.} \tag{5}$$

The last equality shows that the characteristic function is the generating function for the moments

$$\langle X^n \rangle = (-i)^n \frac{d^n}{dk^n}\Big|_{k=0} \phi_X(k). \tag{6}$$

The cumulants of $P_X(x)$ are defined as

$$\boxed{C_n(X) = (-i)^n \frac{d^n}{dk^n}\Big|_{k=0} \ln\big(\phi_X(k)\big).} \tag{7}$$

The first few cumulants are

$$C_1(X) = \langle X \rangle\ , \quad C_2(X) = \langle X^2 \rangle - \langle X \rangle^2\ , \quad C_3(X) = \langle X^3 \rangle - 3\langle X^2 \rangle \langle X \rangle + 2\langle X \rangle^3. \tag{8}$$

The Gaussian distribution is defined by

$$P_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_0)^2}{2\sigma^2}}. \tag{9}$$

Its characteristic function is

$$\phi_X(k) = e^{ikx_0 - \frac{1}{2}k^2\sigma^2}. \tag{10}$$

The cumulants are $C_1(X) = x_0$, $C_2(X) = \sigma^2$, $C_{n>2}(X) = 0$. Hence the Gaussian distribution is completely determined by its first two cumulants.

1.2 Discrete-time random walk

To see the above definitions in action we consider a tutor on (without loss of generality) his way back to Summertown after a long night out (needless to say it must be a humanities tutor). He moves along Banbury road, by making each second a step forward/backward with equal probability. Modelling Banbury road by a line, his possible positions are all integers $-\infty < n < \infty$ (assuming for simplicity that Banbury road is infinitely long, which is probably how it feels to our tutor), and we want to know the probability $P_N(n)$ for her to be at position n after N steps, starting from $n = 0$.

- Each step is a random variable X_j $(j = 1, \ldots, N)$ taking the values ± 1 with probabilities $1/2$.
- The position after r steps is $Y = \sum_{j=1}^{N} X_j$.

Clearly we have

$$\langle Y \rangle = 0\ , \qquad \langle Y^2 \rangle = N \langle X_j^2 \rangle = N\ , \tag{11}$$

where we have used that the steps are mutually independent. To obtain $p_N(n)$ we employ the characteristic function

$$\phi_{X_j}(k) = P_{X_j}(1) e^{ik1} + P_{X_j}(-1) e^{ik(-1)} = \cos(k). \tag{12}$$

The characteristic function of the random variable Y is

$$\begin{aligned} \phi_Y(k) &= \langle e^{ikY} \rangle = \langle e^{ik\sum_{j=1}^{N} X_j} \rangle = \prod_{j=1}^{N} \langle e^{ikX_j} \rangle = \big(\phi_{X_j}(k)\big)^N \\ &= \frac{1}{2^N} \left(e^{ik} + e^{-ik}\right)^N = \frac{1}{2^N} \sum_{r=0}^{N} \binom{N}{r} e^{ik(N-2r)}. \end{aligned} \tag{13}$$

On the other hand we have by definition of the characteristic function

$$\phi_Y(k) = \sum_n P_Y(n) e^{ikn} = \sum_{n=-N}^{N} p_N(n) e^{ikn}. \tag{14}$$

Equating (13) with (14) gives

$$p_N(n) = \frac{1}{2^N} \binom{N}{\frac{N-n}{2}}, \tag{15}$$

where the binomial coefficient is taken to be zero if $(N-n)/2$ is not an integer between 0 and N.

1.3 The central limit theorem

Let X_j $j = 1, \ldots, N$ be independent random variables with identical distributions $P_X(x)$ such that $\langle X \rangle$ and $\langle X^2 \rangle$ are finite. and consider their sum

$$S_N = \sum_{j=1}^{N} X_j. \tag{16}$$

The central limit theorem states that, for large N, the random variable

$$Y \equiv \frac{S_N - N\langle X \rangle}{\sqrt{N}} \tag{17}$$

is distributed as a Gaussian with zero mean and variance $\langle X^2 \rangle - \langle X \rangle^2$, irrespective of the form of $P_X(x)$, provided that $\langle X \rangle$ and $\langle X^2 \rangle$ are finite.

To see that this is the case, let us consider the characteristic function of Y:

$$\phi_Y(k) = \langle e^{i\frac{k}{\sqrt{N}} S_N} \rangle e^{-ik\sqrt{N}\langle X \rangle} \tag{18}$$

Now,

$$\langle e^{i\frac{k}{\sqrt{N}} S_N} = \langle e^{i\frac{k}{\sqrt{N}} \sum_{j=1}^{N} X_j} \rangle = \langle e^{ik\frac{X}{\sqrt{N}}} \rangle^N = \left(\phi_X\big(\frac{k}{\sqrt{N}}\big) \right)^N, \tag{19}$$

where we have used that the random variables are independent and have the same distribution. Thus, we have

$$\phi_Y(k) = \left(\phi_X\big(\frac{k}{\sqrt{N}}\big) \right)^N e^{-ik\sqrt{N}\langle X \rangle} \tag{20}$$

The cumulants of Y are

$$C_n(Y) = (-i)^n \frac{d^n}{dk^n}\Big|_{k=0} \left[N \ln \left(\phi_X\big(\frac{k}{\sqrt{N}}\big) - ik\sqrt{N}\langle X \rangle \right] \right). \tag{21}$$

We see that these cumulants are related to the cumulants $C_n(X)$ of $P_X(x)$ by

$$\boxed{C_n(Y) = N^{1-n/2} C_n(X) - \delta_{n,1}\sqrt{N}\langle X \rangle.} \tag{22}$$

In particular, we see that $C_1(Y) = 0$, $C_2(Y) = C_2(X)$, and all higher moments $C_{n>2}(Y) = N^{1-n/2} C_n(X)$ vanish for $N \gg 1$. Hence, for very large N, $P_Y(y)$ tends to a Gaussian distribution with average 0 and variance $C_2(X) = \sigma^2$, i.e.

$$\boxed{P_Y(y) \longrightarrow \left(\frac{1}{2\pi\sigma^2} \right)^{\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2} y^2 \right].} \tag{23}$$

Note that by rescaling, this can also be phrased in terms of the distribution of the arithmetic mean of the X_js (which is a Gaussian with $O(1)$ mean and $O(1/N)$ variance, rather than the $O(N)$ mean and a $O(1)$ variance of the sum), as in previous editions of these notes. However, in my opinion the proof for the sum — particularly the vanishing of higher cumulants — is cleaner.

2 Stochastic Processes: Definitions

A function $Y_X(t)$ of time t and a random variable X is called a *stochastic process* (SP). Examples are the position $x(t)$ or the velocity $v(t)$ of a particle undergoing Brownian motion. A SP is characterized by probability densities

$$\begin{aligned} P_1(y_1,t_1) &= \text{probability that } Y(t_1)=y_1; \\ &\vdots \\ P_n(y_1,t_1;\ldots;y_n,t_n) &= \text{joint probability that } Y(t_1)=y_1,\ldots,Y(t_n)=y_n. \end{aligned} \tag{24}$$

These are

- normalized
$$\int dy_1\ldots dy_n\ P_n(y_1,t_1;\ldots;y_n,t_n)=1\ ; \tag{25}$$
- reducible
$$\int dy_n\ P_n(y_1,t_1;\ldots;y_n,t_n)=P_{n-1}(y_1,t_1;\ldots;y_{n-1},t_{n-1}). \tag{26}$$

Another way of characterizing a SP is through *correlation functions*

$$\begin{aligned} \langle Y(t_1)\rangle &= \int dy_1\ y_1\ P(y_1,t_1)\ , \\ \langle Y(t_1)Y(t_2)\rangle &= \int dy_1 dy_2\ y_1y_2\ P_2(y_1,t_1;y_2,t_2)\ , \\ &\ldots \end{aligned} \tag{27}$$

A stochastic process is called *stationary* if for all n and any t

$$P_n(y_1,t_1;\ldots;y_n,t_n)=P_n(y_1,t_1+\tau;\ldots;y_n,t_n+\tau). \tag{28}$$

The *conditional probability* that $Y=y_2$ at time t_2, given that $Y=y_1$ at time t_1 is denoted by

$$P_{1|1}(y_2,t_2|y_1,t_1). \tag{29}$$

This is normalized

$$\int dy_2\ P_{1|1}(y_2,t_2|y_1,t_1)=1\ , \tag{30}$$

and related to unconditional probabilities by

$$P_2(y_1,t_1;y_2,t_2)=P_{1|1}(y_2,t_2|y_1,t_1)\ P_1(y_1,t_1). \tag{31}$$

The *conditional probability* that $Y=y_n$ at time t_n, given that $Y=y_j$ at time t_j for $j=1,\ldots,n-1$ is denoted by

$$P_{1|n-1}(y_n,t_n|y_1,t_1;y_2,t_2;\ldots;y_{n-1},t_{n-1}). \tag{32}$$

3 Markov Processes

Perhaps the most important stochastic processes are so-called *Markov processes* (MP). Their defining property is that $t_1<t_2<\ldots<t_n$

$$\boxed{P_{1|n-1}(y_n,t_n|y_1,t_1;y_2,t_2;\ldots;y_{n-1},t_{n-1})=P_{1|1}(y_n,t_n|y_{n-1},t_{n-1}).} \tag{33}$$

This means that at time t_{n-1} one can predict the state of the system at time t_n on the basis of *present* information, i.e. y_{n-1}, only! The history of how the system arrived at y_{n-1} at time t_{n-1} is irrelevant.

A Markov process is completely determined by the two functions $P_1(y_1,t_1)$ and $P_{1|1}(y_2,t_2|y_1,t_1)$, and this makes Markov processes tractable. Any $P_1(y_1,t_1)$ and $P_{1|1}(y_2,t_2|y_1,t_1)$ define a MP, provided that they fulfil the following two consistency conditions

- Chapman-Kolmogorov equation

$$\boxed{P_{1|1}(y_3,t_3|y_1,t_1) = \int dy_2 \; P_{1|1}(y_3,t_3|y_2,t_2)P_{1|1}(y_2,t_2|y_1,t_1) \; , \quad t_3 > t_2 > t_1.} \tag{34}$$

- Evolution equation

$$\boxed{P_1(y_2,t_2) = \int dy_1 \; P_{1|1}(y_2,t_2|y_1,t_1)P_1(y_1,t_1) \; , \quad t_2 > t_1.} \tag{35}$$

A MP is called *stationary* if $P_1(y,t)$ is time-independent and $P_{1|1}(y_2,t_2|y_1,t_1)$ depends only on the time difference $t_2 - t_1$ (and $y_{1,2}$).

3.1 Examples of Markov Processes

- The discrete-time random walk is a MP.
- The *Wiener process*, defined by

$$\begin{aligned} P_{1|1}(y_2,t_2|y_1,t_1) &= \frac{1}{\sqrt{2\pi(t_2-t_1)}} e^{-\frac{(y_2-y_1)^2}{2(t_2-t_1)}} \; , \\ P_1(y,t) &= \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} \; , \end{aligned} \tag{36}$$

  is a non-stationary Markov process. It was originally invented for describing the stochastic behaviour of the position of a Brownian particle, and is also called a continuous-time random walk.

3.2 Markov Chains

A Markov chain is a MP, in which the random variable only takes a finite number of values and involves discrete time steps. The probability $P_1(y,t)$ can then be represented as a N-component vector, and $P_{1|1}(y_2,t_2|y_1,t_1) \equiv T$ as an $N \times N$ matrix. The normalization condition (30) becomes

$$\boxed{\sum_{j=1}^{N} T_{jk} = 1,} \tag{37}$$

which is often referred to as probability conservation.

Example: Two state process with $Y = 1$ or $Y = 2$, with conditional probabilities

$$\begin{aligned} P_{1|1}(1,t+1;1,t) &= 1-q \; , \\ P_{1|1}(2,t+1;1,t) &= q \; , \\ P_{1|1}(1,t+1;2,t) &= r \; , \\ P_{1|1}(2,t+1;2,t) &= 1-r \; . \end{aligned} \tag{38}$$

Introducing a two-component vector

$$\vec{p}(t) = \begin{pmatrix} P_1(1,t) \\ P_1(2,t) \end{pmatrix} \ , \tag{39}$$

the evolution equation for the process can be expressed as a vector equation

$$\boxed{\vec{p}(t+1) = \underbrace{\begin{pmatrix} 1-q & r \\ q & 1-r \end{pmatrix}}_{T} \vec{p}(t).} \tag{40}$$

T is a square matrix with non-negative entries, that is in general not symmetric. The matrix elements T_{ij} are the rates for transitions from state j to state i. It is useful to rewrite the equation in components

$$p_n(t+1) = \sum_j T_{nj} p_j(t). \tag{41}$$

Then

$$p_n(t+1) - p_n(t) = \sum_j T_{nj} p_j(t) - \underbrace{\sum_j T_{jn}}_{1} \ p_n(t) = \sum_j T_{nj} p_j(t) - T_{jn} p_n(t). \tag{42}$$

Now consider our time interval to be small instead of 1. Then (42) turns into a differential equation, called a *Master equation*

$$\boxed{\frac{dp_n(t)}{dt} = \sum_j T_{nj} p_j(t) - T_{jn} p_n(t).} \tag{43}$$

This has a nice physical interpretation as a "loss/gain" equation for probabilities: the first term on the right-hand side is the rate of transitions from state j to state n times the probability of j being realized, i.e. the total gain of probability for state n. The second term on the right-hand side is the rate of transitions from state n to state j times the probability of n being realized, i.e. the total loss of probability for state n. Let us now return to the discrete form (40), which can be iterated to give

$$\vec{p}(t+1) = T^{t+1} \vec{p}(0). \tag{44}$$

While T is generally not symmetric, it is nevertheless often diagonalizable. Then there exist left and right eigenvectors such that

$$T|R_\alpha\rangle = \lambda_\alpha |R_\alpha\rangle \ , \quad \langle L_\alpha | T = \lambda_\alpha \langle L_\alpha | \ , \quad \langle L_\alpha | R_\beta \rangle = \delta_{\alpha,\beta} \ , \tag{45}$$

and T can be represented in the form

$$T = \sum_\alpha \lambda_\alpha |R_\alpha\rangle\langle L_\alpha|. \tag{46}$$

Using the orthonormality of left and right eigenvectors we have

$$\boxed{T^{t+1} = \sum_\alpha \lambda_\alpha^{t+1} |R_\alpha\rangle\langle L_\alpha|.} \tag{47}$$

In our example

$$\lambda_1 = 1 \ , \quad \langle L_1| = (1,1) \ , \quad |R_1\rangle = \frac{1}{1+r/q} \begin{pmatrix} r/q \\ 1 \end{pmatrix}. \tag{48}$$

$$\lambda_2 = 1-q-r \ , \quad \langle L_2| = (-\frac{q}{r}, 1) \ , \quad |R_2\rangle = \frac{1}{1+r/q} \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{49}$$

So for large t we have

$$T^{t+1} \approx |R_1\rangle\langle L_1| = \frac{1}{r+q}\begin{pmatrix} r & r \\ q & q \end{pmatrix}, \tag{50}$$

and hence

$$\vec{p}(\infty) = \frac{1}{r+q}\begin{pmatrix} r & r \\ q & q \end{pmatrix}\ \vec{p}(0). \tag{51}$$

4 Brownian Motion

We now want to think of Brownian motion as a Markov process. Let $v_1, v_2, \ldots$ be the velocities of the particle at different time steps. Then v_{k+1} depends only on v_k, but not on $v_1, \ldots, v_{k-1}$.

4.1 Langevin Equation

One approach to Brownian motion is via a *stochastic differential equation*, the *Langevin equation* for the velocity $v(t)$ (more precisely the velocity in $D = 1$ or a component of the velocity in $D > 1$)

$$\boxed{\frac{dv(t)}{dt} = -\gamma v(t) + \eta(t).} \tag{52}$$

Here the first term on the right hand side is damping term linear in v, while the second term represents the remaining *random force* with zero average $\langle \eta(t) \rangle = 0$. This is often referred to as "noise". For simplicity we will assume collisions to be instantaneous, so that forces at different times are uncorrelated

$$\langle \eta(t)\eta(t') \rangle = \Gamma \delta(t - t'). \tag{53}$$

Given our assumptions about the noise, we can calculate noise-averaged quantities quite easily. We have

$$\frac{d}{dt'}\left(v(t')e^{\gamma t'}\right) = \left(\frac{dv}{dt'} + \gamma v\right)e^{\gamma t'} = \eta(t')e^{\gamma t'}, \tag{54}$$

where in the last step we used the Langevin equation (52). Integrating both sides of this equation between $t = 0$ and $t' = t$, we obtain

$$\boxed{v(t) = v(0)e^{-\gamma t} + \int_0^t dt'\ \eta(t')e^{-\gamma(t-t')}.} \tag{55}$$

Averaging this over the noise, we find the average velocity

$$\boxed{\langle v(t) \rangle = v(0)e^{-\gamma t} + \int_0^t dt'\ \underbrace{\langle \eta(t') \rangle}_{=0} e^{-\gamma(t-t')} = v(0)e^{-\gamma t}.} \tag{56}$$

Similarly we have

$$\begin{aligned} \langle v^2(t) \rangle &= \left\langle \left[v(0)e^{-\gamma t} + \int_0^t dt'\ \eta(t')e^{-\gamma(t-t')} \right] \left[v(0)e^{-\gamma t} + \int_0^t dt''\ \eta(t'')e^{-\gamma(t-t'')} \right] \right\rangle \\ &= v^2(0)e^{-2\gamma t} + e^{-2\gamma t}\int_0^t dt'dt''\ \underbrace{\langle \eta(t')\eta(t'') \rangle}_{\Gamma\delta(t'-t'')} e^{\gamma(t'+t'')} \end{aligned} \tag{57}$$

Carrying out the time integrals this becomes

$$\boxed{\langle v^2(t) \rangle = v^2(0)e^{-2\gamma t} + \frac{\Gamma}{2\gamma}(1 - e^{-2\gamma t}).} \tag{58}$$

The displacement of the particle is

$$x(t) - x(0) = \int_0^t dt'\ v(t') = \frac{v(0)}{\gamma}(1 - e^{-\gamma t}) + \int_0^t dt' \int_0^{t'} dt''\ \eta(t'')e^{-\gamma(t'-t'')}. \tag{59}$$

Assuming $x(0) = 0$, the average position of the particle is

$$\boxed{\langle x(t)\rangle = \frac{v(0)}{\gamma}(1 - e^{-\gamma t}).} \tag{60}$$

Finally, we want to determine the particle's mean square deviation

$$\langle \big[x(t) - \langle x(t)\rangle\big]^2\rangle = \langle x^2(t)\rangle - \langle x(t)\rangle^2. \tag{61}$$

Substituting (59) and using again that $x(0) = 0$ this becomes

$$\int_0^t dt_1 \int_0^{t_1} dt_2 e^{-\gamma(t_1-t_2)} \int_0^t dt_1' \underbrace{\int_0^{t_1'} dt_2' e^{-\gamma(t_1'-t_2')}\langle \eta(t_2)\eta(t_2')\rangle}_{\Gamma\Theta(t_1'-t_2)e^{-\gamma(t_1'-t_2)}} = \int_0^t dt_1 \int_0^{t_1} dt_2 e^{-\gamma(t_1-2t_2)} \underbrace{\int_{t_2}^t dt_1' e^{-\gamma t_1'}}_{\gamma^{-1}(e^{-\gamma t_2}-e^{-\gamma t})}. \tag{62}$$

Carrying out the remaining two integrals we find

$$\boxed{\langle \big[x(t) - \langle x(t)\rangle\big]^2\rangle = \frac{\Gamma}{\gamma^2}t - \frac{\Gamma}{\gamma^3}(1 - e^{-\gamma t}) - \frac{\Gamma}{2\gamma^3}(1 - e^{-\gamma t})^2.} \tag{63}$$

So at very late times we have

$$\boxed{\langle x^2(t)\rangle = \frac{\Gamma}{\gamma^2}t + \ldots} \tag{64}$$

The displacement grows like $\sqrt{t}$, which is characteristic of *diffusion*. Finally, we may relate Γ/γ to the temperature of the fluid by noting that

$$\langle v^2(t \to \infty)\rangle = \frac{\Gamma}{2\gamma}. \tag{65}$$

On the other hand, by equipartition we have

$$\frac{m}{2}\langle v^2\rangle \sim \frac{k_B T}{2} \tag{66}$$

Combining these two equations, we arrive at

$$\boxed{\frac{\Gamma}{2\gamma} = \frac{k_B T}{m}.} \tag{67}$$

4.2 Fokker-Planck Equation

We now want to derive a differential equation for the probability $P_1(v,t)$ of our particle having velocity v at time t from the Langevin equation

$$v(t) = v(0)e^{-\gamma t} + \int_0^t dt'\ e^{-\gamma(t-t')}\eta(t'). \tag{68}$$

Our starting point is the general evolution equation

$$P_1(v, t + \Delta t) = \int du P_{1|1}(v, t + \Delta t|u, t)\ P_1(u, t). \tag{69}$$

It is convenient to consider the integral

$$\Omega = \int dv \, [P_1(v, t + \Delta t) - P_1(v, t)] \, h(v), \tag{70}$$

where $h(v)$ is is test function (infinitely many time differentiable, $h(v)$ and all of its derivatives going to zero at infinity etc). On the one hand, we have to linear order in Δt

$$\int dv \, [P_1(v, t + \Delta t) - P_1(v, t)] \, h(v) = \int dv \; \frac{\partial P_1(v, t)}{\partial t} \Delta t \; h(v). \tag{71}$$

On the other hand, using (69) we have

$$\Omega = \int du \; dv \; h(v) \; P_{1|1}(v, t + \Delta t|u, t) \; P_1(u, t) - \int dv \; h(v) \; P_1(v, t). \tag{72}$$

Relabelling the integration variable from v to u in the second term, and using that normalization condition $\int dv \; P_{1|1}(v, t + \Delta t|u, t) = 1$, we obtain

$$\Omega = \int du \; P_1(u, t) \int dv \; P_{1|1}(v, t + \Delta t|u, t) \, [h(v) - h(u)] \, . \tag{73}$$

Expanding $h(v)$ around u in a Taylor series gives

$$\Omega = \int du \; P_1(u, t) \sum_{n=1}^{\infty} h^{(n)}(u) \underbrace{\int dv \; P_{1|1}(v, t + \Delta t|u, t) \frac{(v - u)^n}{n!}}_{D^{(n)}(u)} \, . \tag{74}$$

Integrating the n'th term in the sum n times by parts, and using the nice properties of the function $h(u)$, then leads to the following expression

$$\Omega = \int du \; h(u) \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial u} \right)^n P_1(u, t) D^{(n)}(u). \tag{75}$$

Using that (71) and (75) have to be equal for *any* test function $h(u)$, we conclude that

$$\boxed{\frac{\partial P_1(v, t)}{\partial t} \Delta t = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial v} \right)^n P_1(v, t) D^{(n)}(v).} \tag{76}$$

This starts looking like our desired differential equation. What remains is to determine the quantities $D^{(n)}(v)$

$$\begin{aligned} D^{(n)}(v) &= \int dw \; P_{1|1}(w, t + \Delta t|v, t) \frac{(w - v)^n}{n!} = \int dz \; P_{1|1}(v + z, t + \Delta t|v, t) \frac{z^n}{n!} \\ &= \frac{1}{n!} \langle [v(t + \Delta t) - v(t)]^n \rangle. \end{aligned} \tag{77}$$

We see that $D^{(n)}(v)$ are related to the moments of the velocity difference distribution! We can use the Langevin equation to determine them, and the result is

$$\begin{aligned} \langle v(t + \Delta t) - v(t) \rangle &= -\gamma v(0) e^{-\gamma t} \Delta t \; , \\ \langle \big(v(t + \Delta t) - v(t) \big)^2 \rangle &= \Gamma \Delta t + \mathcal{O}\big((\Delta t)^2 \big) \; , \\ \langle \big(v(t + \Delta t) - v(t) \big)^n \rangle &= \mathcal{O}\big((\Delta t)^2 \big) \; , \quad n \geq 3. \end{aligned} \tag{78}$$

Substituting these into (76) and then taking the limit $\Delta t \to 0$, we arrive at the *Fokker-Planck equation*

$$\boxed{\frac{\partial}{\partial t}P_1(v,t) = \gamma\frac{\partial}{\partial v}vP_1(v,t) + \frac{\Gamma}{2}\frac{\partial^2}{\partial v^2}P_1(v,t).} \tag{79}$$

This is a second order linear PDE for $P_1(v,t)$ and can be solved by standard methods. For initial conditions $P_1(v,0) = \delta(v - v_0)$ we find

$$\boxed{P_1(v,t) = \frac{1}{\sqrt{2\pi\sigma^2(t)}}\exp\left(-\frac{\big(v-\bar{v}(t)\big)^2}{2\sigma^2(t)}\right),} \tag{80}$$

where

$$\sigma^2(t) = \frac{\Gamma}{2\gamma}(1 - e^{-2\gamma t}) \ , \quad \bar{v}(t) = v_0 e^{-\gamma t}. \tag{81}$$

In the limit $t \to \infty$ this turns into the *Maxwell distribution*

$$P_1(v,t) = \sqrt{\frac{\gamma}{\pi\Gamma}}e^{-\frac{\gamma v^2}{\Gamma}} = \sqrt{\frac{m}{2\pi k_B T}}e^{-\frac{mv^2}{2k_B T}}. \tag{82}$$

1: Moments of the Velocity Difference Probability Distribution

Let us see how to derive (78), starting from the Langevin equation

$$v(t) = v(0)e^{-\gamma t} + \int_0^t dt' \ e^{-\gamma(t-t')}\eta(t'). \tag{83}$$

For a very small time interval Δt we have

$$\begin{aligned} v(t+\Delta t) &= v(0)e^{-\gamma(t+\Delta t)} + \int_0^{t+\Delta t} dt' \ e^{-\gamma(t+\Delta t-t')}\eta(t') \\ &= v(0)e^{-\gamma t}(1-\gamma\Delta t) + \int_0^t dt' \ e^{-\gamma(t-t')}\eta(t')(1-\gamma\Delta t) + \int_t^{t+\Delta t} dt' \ e^{-\gamma(t-t')}\eta(t') + \mathcal{O}\big((\Delta t^2)\big) \\ &= (1-\gamma\Delta t)\ v(t) + \int_t^{t+\Delta t} dt' \ e^{-\gamma(t-t')}\eta(t') + \mathcal{O}\big((\Delta t^2)\big). \end{aligned} \tag{84}$$

Hence $\Delta v(t) = v(t+\Delta t) - v(t)$ is given by

$$\Delta v = -\gamma v\Delta t + \int_t^{t+\Delta t} dt' \ e^{-\gamma(t-t')}\eta(t') + \mathcal{O}\big((\Delta t^2)\big). \tag{85}$$

To derive (78) we rewrite (85) in the form

$$\Delta v(t) = v_0(t)\Delta t + \int dt' K(t,t') \ \eta(t'), \tag{86}$$

where

$$\begin{aligned} v_0(t) &= -\gamma v(0)e^{-\gamma t} \ , \\ K(t,t') &= -\gamma\Theta(t-t')\Theta(t')e^{-\gamma(t-t')}\Delta t + \Theta(t'-t)\Theta(t+\Delta t-t')e^{-\gamma(t-t')}. \end{aligned} \tag{87}$$

Let us now assume that $\eta(t)$ is Gaussian distributed. Then the probability distribution for the noise is the functional

$$P[\eta(t)] = e^{-\frac{1}{2\Gamma}\int dt \ \eta^2(t)}. \tag{88}$$

Averages are then given by the path integral

$$\langle \eta(t_1)\dots\eta(t_n)\rangle = \int \mathcal{D}\eta(t)\ \eta(t_1)\dots\eta(t_n)\ e^{-\frac{1}{2\Gamma}\int dt'\ \eta^2(t')}. \tag{89}$$

As this is Gaussian, we may use Wick's theorem to calculate averages

$$\begin{aligned}
\langle \eta(t)\rangle &= 0\ , \\
\langle \eta(t)\eta(t')\rangle &= \Gamma\delta(t-t')\ , \\
\langle \eta(t_1)\eta(t_2)\eta(t_3)\rangle &= \langle \eta(t_1)\eta(t_2)\rangle\langle \eta(t_3)\rangle + \langle \eta(t_3)\eta(t_1)\rangle\langle \eta(t_2)\rangle + \langle \eta(t_2)\eta(t_3)\rangle\langle \eta(t_1)\rangle = 0\ , \\
\langle \eta(t_1)\eta(t_2)\eta(t_3)\eta(t_4)\rangle &= \langle \eta(t_1)\eta(t_2)\rangle\langle \eta(t_3)\eta(t_4)\rangle + \dots
\end{aligned} \tag{90}$$

The probability distribution for $\Delta v(t)$ can be obtained from the generating function

$$Z(\lambda) = \langle e^{\lambda\Delta v(t)}\rangle = e^{\lambda v_0(t)\Delta t}\int \mathcal{D}\eta(t)\ e^{-\frac{1}{2\Gamma}\int dt'\left[\eta^2(t')-2\lambda\Gamma K(t,t')\eta(t')\right]}. \tag{91}$$

Changing variables to

$$\tilde\eta(t') = \eta(t') - \lambda\Gamma K(t,t'), \tag{92}$$

the generating function becomes

$$Z(\lambda) = e^{\lambda v_0(t)\Delta t + \frac{\lambda^2\Gamma}{2}\int dt'\ K^2(t,t')}. \tag{93}$$

It is then a straightforward matter to calculate the moments

$$\begin{aligned}
\langle \Delta v\rangle &= \frac{dZ}{d\lambda}\Big|_{\lambda=0} = v_0(t)\Delta t\ , \\
\langle (\Delta v)^2\rangle &= \frac{d^2Z}{d\lambda^2}\Big|_{\lambda=0} = \Gamma\int dt'\ K^2(t,t') = \Gamma\int_t^{t+\Delta t} dt'\ e^{-\gamma(t-t')} + \mathcal{O}\big((\Delta t)^2\big) \\
&= \Gamma\Delta t + \mathcal{O}\big((\Delta t)^2\big)\ , \\
\langle (\Delta v)^n\rangle &= \frac{d^nZ}{d\lambda^n}\Big|_{\lambda=0} = \mathcal{O}\big((\Delta t)^2\big)\ , \quad n\geq 3.
\end{aligned} \tag{94}$$

4.3 Diffusion Equation

Finally, we would like to obtain a differential equation for $P_1(x,t)$. The difficulty is that $x(t)$ is *not* a Markov process, because $x(t+\Delta t)$ depends on both $x(t)$ and $v(t)$. On can treat the combined evolution of $x(t)$ and $v(t)$, but we will follow a simpler route. Recalling that

$$v(t) = v(0)e^{-\gamma t} + \int_0^t dt'\ e^{-\gamma(t-t')}\eta(t'), \tag{95}$$

we see that $v(t)$ becomes a random variable for $t \gg \gamma^{-1}$ with

$$\langle v(t)\rangle = 0\ , \qquad \langle v(t)v(t')\rangle = \frac{\Gamma}{2\gamma}e^{-\gamma|t-t'|}. \tag{96}$$

Now let us imagine that we observe the Brownian particle only at sufficiently long time intervals $t,t' \gg \gamma^{-1}$, and describe only these coarse grained positions. Then we may replace

$$\langle v(t)v(t')\rangle = \frac{\Gamma}{2\gamma}e^{-\gamma|t-t'|} \to \frac{\Gamma}{\gamma^2}\ \delta(t-t'). \tag{97}$$

This is because $\frac{\Gamma}{2\gamma}e^{-\gamma|t-t'|}$ is substantially different from zero only if $|t-t'| < \gamma^{-1}$, and

$$\int_{-\infty}^{\infty} dt \frac{\Gamma}{2\gamma} e^{-\gamma|t-t'|} = \frac{\Gamma}{\gamma^2}. \tag{98}$$

The differential equation for the position

$$\frac{dx(t)}{dt} = v(t), \tag{99}$$

then turns into a special case (of no damping) of the Langevin equation we solved for the velocity. We therefore can use our previous results to conclude that

$$\boxed{\frac{\partial}{\partial t}P_1(x,t) = \frac{\Gamma}{2\gamma^2}\frac{\partial^2}{\partial x^2}P_1(x,t).} \tag{100}$$

This is the *diffusion equation* with diffusion coefficient

$$D = \frac{\Gamma}{2\gamma^2}. \tag{101}$$

Its solution for initial conditions $P_1(x,0) = \delta(x-x_0)$ is

$$\boxed{P_1(x,t) = \frac{1}{\sqrt{2\pi D|t-t_0|}} \exp\left(-\frac{(x-x_0)^2}{4D|t-t_0|}\right).} \tag{102}$$

Observe that the diffusion equation is a special case of the Fokker-Planck equation in the absence of a viscous damping term (the fact that there is a γ in the expression should not worry you – the equation for x has no damping term, which is the important thing.)

5 Brownian motion via the Path Integral

We now give a brief description of the link between stochastic processes and path integrals, focusing on Markov processes, and specifically working with the example of Brownian motion studied above.

Consider a Markov process; it is characterized by the two functions $P_1(y_1,t_1)$ and $P_{1|1}(y_2,t_2|y_1,t_1)$, satisfying the Chapman-Kolmogorov (34) and evolution (103) equations.

For a stochastic process the 'state' of a system at time t_0 is captured by the probability distribution function $P_1(y_0,t_0)$ of the stochastic variable at time t_0 and the evolution equation describes how this evolves to the state at time t:

$$P_1(y,t) = \int dy_0 \; P_{1|1}(y,t|y_0,t_0)P_1(y_1,t_1) \;, \quad t_2 > t_1. \tag{103}$$

In quantum mechanics, we have an analogous equation :

$$|\Psi(t)\rangle = U(t;t_0)|\Psi(t_0)\rangle \tag{104}$$

$$\langle x|\Psi(t)\rangle = \int dx_0 \langle x|U(t;t_0)|x_0\rangle\langle x_0\Psi(t_0)\rangle \tag{105}$$

where in the second step, we have chosen to work in the position basis. Evidently, there is a formal resemblance between equations (103) and (105): namely, $P_{1|1}(y,t|y_0,t_0)$ plays a role analogous to the propagator $U(t;t_0)$ in quantum mechanics. Given the state of the system at a given time t_0, it 'propagates' this forward in time to time $t > t_0$. Of course, (105) describes probability *amplitudes* rather than probabilities, and

therein lies a key difference between quantum and classical physics. However, we can exploit the formal resemblance between the classical evolution of probabilities in a Markov process and the quantum evolution of amplitudes under unitary time evolution to recast the former in path integral form. We now carry out this procedure, specializing to the case of Brownian motion.

Recall that the fundamental idea in the path integral approach to QM was to write the time evolution operator as a path integral by inserting a complete set of states after each infinitesimal time step. The analogous idea for the Markov process is to repeatedly use the Chapman-Kolmogorov equation to split up the time evolution into infinitesimal chunks. Let us split the time interval from t_0 to t into N equal time slices, defining $\epsilon = \frac{t-t_0}{N}$, $t_n = t_0 + n\epsilon$, so that $t_N = t$. Repeatedly applying (34), we have

$$P_{1|1}(y,t|y_0,t_0) = \int dy_1 dy_2 \dots dy_{N-1}\, P_{1|1}(y,t|y_{N-1},t_{N-1}) \dots P_{1|1}(y_2,t_2|y_1,t_1) P_{1|1}(y_1,t_1|y_0,t_0) \quad (106)$$

$$= \int dy_1 dy_2 \dots dy_{N-1} \prod_{n=0}^{N-1} P_{1|1}(y_{n+1},t_{n+1}|y_n,t_n). \quad (107)$$

Just as for the QM case, we have broken up the evolution into infinitesimal evolutions. The reason is that it's frequently easier to compute the *infinitesimal* time evolution. Let's consider the case of Brownian motion, where we have from (36) that

$$P_{1|1}(y_{n+1},t_{n+1}|y_n,t_n) = \frac{1}{\sqrt{2\pi(t_{n+1}-t_n)}} e^{-\frac{(y_{n+1}-y_n)^2}{2(t_{n+1}-t_n)}} = \frac{1}{\sqrt{2\pi\epsilon}} e^{-\frac{(y(t_{n+1})-y(t_n))^2}{2\epsilon}}, \quad (108)$$

where in the second step we have used the fact that in our case $t_n - t_{n-1} = \epsilon$ by construction, and defined $y_n = y(t_n)$. Substituting this into (107) and rewriting, we have

$$P_{1|1}(y,t|y_0,t_0) = \int dy_1 dy_2 \dots dy_{N-1} \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\epsilon}} e^{-\frac{(y(t_{n+1})-y(t_n))^2}{2\epsilon}}$$

$$= \left(\frac{1}{\sqrt{2\pi\epsilon}}\right)^N \int dy_1 dy_2 \dots dy_{N-1}\, e^{-\sum_{n=0}^{N-1} \frac{(y(t_{n+1})-y(t_n))^2}{2\epsilon}}. \quad (109)$$

We now let $N \to \infty$, which defines a trajectory or path $y(t)$. We recognise once again that $\frac{(y(t_{n+1})-y(t_n))}{\epsilon} \simeq \dot{y}(t_n)$, and hence that

$$\sum_{n=0}^{N-1} \frac{(y(t_{n+1})-y(t_n))^2}{2\epsilon} \simeq \int_0^t dt' \frac{\dot{y}(t')^2}{2}. \quad (110)$$

Thus we find that we have expressed the propagator for Brownian motion as a path integral,

$$P_{1|1}(y,t|y_0,t_0) = \mathcal{N} \int \mathcal{D}y(t') e^{-\int_0^t dt' \frac{\dot{y}(t')^2}{2}}. \quad (111)$$

which is the analogue of Eq.(53) of the first part of the notes.

Observe now that if we define an 'imaginary time' $\tilde{t} = -it$, we find that this is identical to the free-particle path integral (Eq. (64) from the first part) for $\hbar = m = 1$. Note that this transformation also maps between the transition probability for the Wiener process (36) and the free-particle propagator for $t_0 = 0$ (Eq. (70) of the first part). Amusingly, we see that this same transformation changes the diffusion equation (100) to the Schrödinger equation, which, more generally, links the Fokker-Planck and path integral descriptions. The course has indeed come full circle (pun intended).

5.1 Homework Questions 22-26

Question 22. This question is concerned with the central limit theorem.
(i) Show explicitly that for $N \gg 1, pN \gg 1$ the binomial distribution

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}, \qquad p+q=1$$

becomes

$$P_N(n) = \frac{1}{\sqrt{2\pi\sigma^2}} \quad \exp\left(\frac{-(n- < n >)^2}{2\sigma^2}\right)$$

where $\sigma^2 = Npq$. Check that the same result follows from the central limit theorem.
(ii) Consider a random walk in one dimension, for which the probability of taking a step of length $x \to x + dx$ is

$$f(x)dx = \frac{1}{\pi} \frac{\gamma}{x^2 + \gamma^2} \, dx.$$

Find the probability distribution for the total displacement after N steps. Does it satisfy the central limit theorem? Should it? What are the cumulants of this distribution?

Question 23. Let $y = \pm 1$. Show that

$$P_{1|1}(y, t \mid y', t') = \frac{1}{2}\left\{1 + e^{-2\gamma(t-t')}\right\}\delta_{y,y'} + \frac{1}{2}\left\{1 - e^{-2\gamma(t-t')}\right\}\delta_{y,-y'} \tag{112}$$

obeys the Chapman-Kolmogorov equation.

Show that

$$P_1(y, t) = \frac{1}{2}(\delta_{y,1} + \delta_{y,-1}) \tag{113}$$

is a stationary solution. Write $P_{1|1}$ as a 2×2 matrix and formulate the Chapman-Kolmogarov equation as a property of that matrix.

Question 24. This question is about a continuous random walk, also known as a Wiener process. Show that for $-\infty < y < \infty$ and $t_2 > t_1$ the Chapman-Kolmogarov equation is satisfied for

$$P_{1|1}(y_2, t_2 \mid y_1, t_1) = \frac{1}{\sqrt{2\pi(t_2 - t_1)}} \exp - \left\{\frac{(y_2 - y_1)^2}{2(t_2 - t_1)}\right\}. \tag{114}$$

Choose $P_1(y_1, 0) = \delta(y_1)$. Show that for $t > 0$

$$P_1(y, t) = \frac{1}{\sqrt{2\pi t}} \exp\left\{\frac{-y^2}{2t}\right\}. \tag{115}$$

Show that $P_1(y, t)$ obeys the diffusion equation

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial y^2} \tag{116}$$

for $D = \frac{1}{2}$. What is the solution for arbitrary $D > 0$?

Question 25. A particle suspended in a fluid undergoes Brownian motion in one dimension with position $x(t)$ and velocity $v(t)$. This motion is modelled by the Langevin equation

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\gamma v + \eta(t),$$

where $\eta(t)$ is a Gaussian random variable characterised completely by the averages $\langle \eta(t) \rangle = 0$ and $\langle \eta(t_1)\eta(t_2) \rangle = \Gamma\delta(t_1 - t_2)$. Discuss the physical origin of each of the terms in the Langevin equation.

What is meant by the term *Markov process*? Illustrate your answer by discussing which of the following are Markov processes: (a) $v(t)$ alone; (b) $x(t)$ alone; (c) $v(t)$ and $x(t)$ together.

Show that for $t > 0$

$$x(t) = \frac{v(0)}{\gamma}\left(1 - \mathrm{e}^{-\gamma t}\right) + \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \; \mathrm{e}^{-\gamma(t_1 - t_2)} \, \eta(t_2)$$

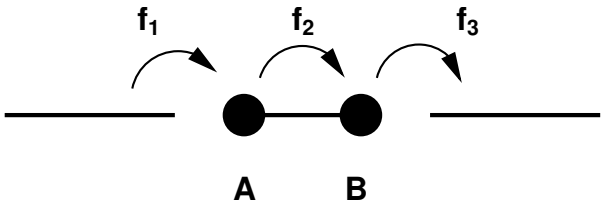
is a solution of the Langevin equation with initial condition $x(0) = 0$. Calculate the average $\langle x(t)\, v(t)\rangle$ and discuss its limiting behaviour at short and long times.

Question 26. The time evolution of a stochastic system is represented by a master equation of the form

$$\frac{\mathrm{d}p_n(t)}{\mathrm{d}t} = \sum_m W_{nm} p_m(t)\,.$$

Explain briefly the meaning of this equation and discuss the assumptions on which it is based. What general conditions should the matrix elements W_{nm} satisfy?

A molecule lies between two atomic-scale contacts and conducts charge between them. A simple model of this situation is illustrated below. The model has three states: the molecule may be uncharged, or may carry a single charge at either site A or site B but not both. Charges hop between these sites, and between the sites and the contacts, at the rates indicated in the figure. (For example, a charge at site A has probability f_2 per unit time of hopping to site B.)



Write down a master equation for this model. For the system in equilibrium, calculate the occupation probabilities of the three states, and show that the average number of charges flowing through the molecule per unit time is

$$\frac{f_1 f_2 f_3}{f_1 f_2 + f_1 f_3 + f_2 f_3}\,.$$

Consider the case $f_1 = f_2 = f_3 \equiv f$. The molecule is uncharged at time $t = 0$. Show that the probability $p(t)$ for it to be uncharged at a later time t is

$$p(t) = \frac{1}{3} + \frac{2}{3}\exp\left(-\frac{3}{2}ft\right)\cos\left(\frac{\sqrt{3}}{2}ft\right)\,.$$