1. Brownian motion

Brownian motion is a useful first example of nonequilibrium statistical mechanics because it shows, in a particularly transparent setting, how the description of a physical system changes when different time scales and different sets of variables are resolved.

Consider a Brownian particle immersed in a fluid. At the microscopic level the complete state contains the position and momentum of the Brownian particle together with the coordinates and momenta of the surrounding molecules. Schematically,

$$\Gamma_{\mathrm{tot}} = (x, p; \mathbf{r}_1, \mathbf{p}_1, \dots, \mathbf{r}_N, \mathbf{p}_N).$$

The variables x and p are therefore genuine mechanical variables of the complete system. If the microscopic state of the fluid is included, their evolution forms part of the ordinary microscopic dynamics. For an isolated particle-fluid system governed by time-reversal-invariant mechanics, reversing all momenta reverses the complete trajectory.

Knowing x and p alone, however, does not specify the complete microscopic state. At the same values of x and p, many different microscopic configurations of the fluid are possible, and these generally exert different instantaneous forces on the Brownian particle. Thus x and p may be measured with arbitrary precision and still fail to constitute a closed mechanical state by themselves. This distinction will become important when we discuss relaxation.

The Brownian particle evolves on time scales much longer than the microscopic fluctuations of the force exerted by the fluid. The separation of these time scales makes it possible to pass successively from a detailed mechanical description to reduced statistical descriptions.

Throughout this section we keep a strict distinction between **elapsed time** and **temporal resolution**. The symbol t denotes the physical time elapsed after preparation of the system. The quantity Δt_{res} denotes the width of the temporal window with which the motion is resolved. Conditions on t describe how the system evolves. Conditions on Δt_{res} describe which temporal structures are visible in the description. These two notions are not interchangeable.

1.1 Langevin equation

For simplicity, consider one-dimensional motion of a spherical Brownian particle of mass m. Its momentum is

$$p = mv.$$

For slow motion through a viscous fluid, the systematic force is the Stokes drag

$$F_{\mathrm{drag}} = -\gamma v,$$

where $\gamma = 6\pi\eta R$ for a sphere of radius R in a fluid of viscosity η. The remaining rapidly varying part of the force exerted by the fluid will be denoted by $F(t)$. The equation of motion is written

$$\dot{p} + \Gamma p = F(t), \qquad \Gamma \equiv \frac{\gamma}{m}.$$

The corresponding momentum-relaxation time is

$$\tau_p \equiv \Gamma^{-1} = \frac{m}{\gamma}.$$

The fluctuating force changes on a much shorter microscopic correlation time τ,

$$\tau \ll \tau_p.$$

This separation of time scales will organise the whole discussion.

For the initial condition $p(0) = p_0$, the formal solution is

$$p(t) = p_0 e^{-\Gamma t} + \int_0^t e^{-\Gamma(t-t')}\, F(t')\, dt'.$$

The first term carries the effect of the prepared initial momentum. The second contains the accumulated effect of the force exerted by the surrounding fluid.

Characteristic microscopic and Brownian time scales

Before choosing a temporal resolution, it is useful to estimate the physical time scales that enter the Brownian-particle problem. Consider, for orientation, a particle of radius $R \approx 10^{-4}\, cm = 1\, \mu m$. The force $F(t)$ is produced by a very large number of molecular interactions with the surrounding medium. Two different microscopic times must first be distinguished.

Let τ denote the duration of one molecule–particle interaction. If d is the characteristic range of the intermolecular interaction and v_{th} is a typical molecular thermal speed, then the duration is of order

$$\tau \approx d/v_{th}\,.$$

Using $d \approx 3 \times 10^{-8}\, cm$ and $v_{th} \approx 3 \times 10^4 - 5 \times 10^4\, cm\, s^{-1}$ gives

$$\tau \approx 10^{-12}\, s\,.$$

The same order of magnitude characterises the width of the force-correlation function: force values separated by times much longer than one interaction duration are no longer appreciably correlated through that particular molecular encounter.

A second microscopic time, denoted by τ', is the mean interval between the beginnings of successive molecular interactions with the Brownian particle. It is much shorter than τ because a micrometre-sized particle presents a large surface to the surrounding molecules. For a dilute gas of number density n, the molecular flux incident on a unit area is of order $j \approx n\, v_{th}/4$. The total rate at which molecules strike a sphere of radius R is therefore

$$\nu_{coll} \approx (n\, v_{th}/4)(4\pi R^2) = \pi R^2\, n\, v_{th}\,.$$

The corresponding average interval between successive impacts is

$$\tau' \approx 1/\nu_{coll} = 1/(\pi R^2\, n\, v_{th})\,.$$

For air under normal conditions, a mean intermolecular spacing $a \approx 3.3 \times 10^{-7}\, cm$ corresponds to $n \approx a^{-3} \approx 3 \times 10^{19}\, cm^{-3}$. With $R \approx 10^{-4}\, cm$ and $v_{th} \approx 5 \times 10^4\, cm\, s^{-1}$,

$$\nu_{coll} \approx 5 \times 10^{16}\, s^{-1}, \quad \tau' \approx 2 \times 10^{-17}\, s\,.$$

Since the estimate is only at the level of orders of magnitude, it is convenient to write

$$\tau' \approx 10^{-17} - 10^{-16}\, s\,.$$

Hence $\tau' \ll \tau$. During one interaction time τ, roughly $\tau/\tau' \approx 10^4 - 10^5$ other molecular interactions begin at different points on the surface of the Brownian particle. The total force $F(t)$ is therefore not a sequence of isolated kicks. At every instant it is the superposition of many overlapping microscopic force contributions.

The third characteristic time is associated with the response of the Brownian particle itself. From the Langevin equation, the momentum-relaxation time is

$$\tau_p = \Gamma^{-1} = m/\gamma\,.$$

For a spherical particle, $\gamma = 6\pi\eta R$ and $m = \left(\frac{4}{3}\right)\pi R^3 \rho_p$, where η is the dynamic viscosity and ρ_p is the particle density. Therefore

$$\tau_p = 2\rho_p R^2/(9\eta)\,.$$

For $R \approx 10^{-4}\,cm, \rho_p$ of order $1\ g\ cm^{-3}$, and a sufficiently viscous surrounding medium with η of order $1-2\ Pa\cdot s$, this gives τ_p of order $10^{-10}\,s$. The precise value of τ_p is strongly medium dependent; the estimate is used only to illustrate the separation between the microscopic force-correlation time and the slower momentum relaxation.

The time τ_p marks the relaxation of the momentum and therefore the onset of the position-only diffusive description. It should not be confused with the generally much longer time required for the particle position to spread over a macroscopic distance. That second time scale will be introduced below after the spatial diffusion coefficient has been obtained.

$$\tau' \approx 10^{-17} - 10^{-16}\,s \ll \tau \approx 10^{-12}\,s \ll \tau\_p \approx 10^{-10}\,s\,.$$

These are intrinsic physical time scales of the system. The temporal resolution introduced below has a different status: it is a chosen resolution of the description. For the first coarse description we shall use the resolution $\Delta t_{res}^{(1)}$ defined below, chosen so that $\tau \ll \Delta t_{res}^{(1)} \ll \tau_p$. This choice suppresses the detailed structure of the microscopic force while retaining the slower evolution of the Brownian momentum.

Microscopic temporal resolution

If the motion is resolved on a scale

$$\Delta t_{\mathrm{res}} \lesssim \tau,$$

the detailed variation of the microscopic force is visible. On this scale $F(t)$ is an ordinary, rapidly varying mechanical force generated by the instantaneous microscopic state of the fluid. The complete particle-fluid state evolves according to the underlying mechanical equations.

We shall not use this microscopic description to derive the statistical laws below. It serves as the reference level from which the first coarse description is constructed.

First coarse temporal resolution

We now choose a fixed temporal resolution $\Delta t_{\mathrm{res}}^{(1)}$ such that

$$\boxed{\tau \ll \Delta t_{\mathrm{res}}^{(1)} \ll \tau_p.}$$

For a rapidly varying quantity $A(t)$, its value on this coarse time scale is represented by the average

$$\overline{A}(t) = \frac{1}{\Delta t_{\mathrm{res}}^{(1)}} \int_{t-\Delta t_{\mathrm{res}}^{(1)}/2}^{t+\Delta t_{\mathrm{res}}^{(1)}/2} A\,(s)\,ds.$$

The overbar will denote this coarse-time average in the Brownian-motion discussion below. The averaging window is long compared with the force-correlation time, so many microscopic force fluctuations occur inside one resolution interval, but it is short compared with τ_p, so the slower evolution of the Brownian momentum remains resolved.

For a fluid in equilibrium there is no preferred sign of the rapidly varying force, and on this scale

$$\overline{F}(\mathrm{t}) = 0.$$

Averaging the formal solution gives

$$\boxed{\overline{p}(t) = p_0 e^{-\Gamma t}.}$$

This equation describes the systematic part of the Brownian momentum on the chosen coarse scale. The actual momentum $p(t)$ continues to fluctuate around it.

Define

$$\Delta p(t) = p(t) - \overline{p}(t).$$

Then

$$\Delta p(t) = \int_0^t e^{-\Gamma(t-t')}\, F(t')\, dt'.$$

For a stationary fluid the correlation between force values at two times depends only on their separation,

$$\overline{F(t_1)F(t_2)} = \varphi(t_1 - t_2).$$

The correlation function $\varphi(s)$ is appreciable only for $|s| \lesssim \tau$ and becomes negligible when $|s| \gg \tau$. Thus, the microscopic force has a short memory.

The momentum variance is therefore

$$\overline{(\Delta p)^2} = \int_0^t d\, t_1 \int_0^t d\, t_2\, e^{-\Gamma(2t-t_1-t_2)} \varphi(t_1 - t_2).$$

The integrand differs appreciably from zero only in a narrow strip of width of order τ around $t_1 = t_2$. Across this narrow strip the exponential varies only weakly because its characteristic time is $\tau_p \gg \tau$. To leading order, the detailed shape of φ therefore enters only through its integrated strength. Denoting the characteristic amplitude of the short correlation by φ as well, this integrated strength is of order $\varphi\tau$, and one obtains

$$\overline{(\Delta p)^2} \simeq \frac{\varphi\tau}{2\Gamma}(1 - e^{-2\Gamma t}).$$

At sufficiently long elapsed time the Brownian momentum must approach thermal equilibrium. Writing

$$\theta \equiv k_B T,$$

one-dimensional equipartition gives

$$\overline{p^2}_{\mathrm{eq}} = m\theta.$$

This fixes the strength of the force correlations,

$$\varphi\tau = 2\Gamma m\theta = 2\gamma\theta,$$

and hence

$$\boxed{\overline{(\Delta p)^2} = m\theta(1 - e^{-2\Gamma t}).}$$

Evolution at the fixed first coarse resolution

From this point until explicitly stated otherwise, the temporal resolution remains fixed at

$$\tau \ll \Delta t_{\mathrm{res}}^{(1)} \ll \tau_p.$$

We now vary the **elapsed time** t, not the resolution.

For resolved elapsed times satisfying

$$\tau \ll t \ll \tau_p,$$

we have $\Gamma t \ll 1$ and therefore

$$1 - e^{-2\Gamma t} \simeq 2\Gamma t.$$

The momentum variance becomes

$$\boxed{\overline{(\Delta p)^2} \simeq 2\Gamma m\theta\, t = 2\gamma\theta\, t.}$$

Thus, at the already chosen coarse resolution, the momentum fluctuations grow linearly with the elapsed time. This is diffusion in momentum space. Nothing has been changed in the temporal

resolution in obtaining this result: the linear law is the **time evolution of the momentum variance at fixed coarse resolution**.

At the same time,

$$\overline{p}(t) = p_0 e^{-\Gamma t} \simeq p_0$$

for $t \ll \tau_p$. The mean momentum has not yet relaxed appreciably even though its fluctuations are already spreading diffusively.

As the elapsed time approaches τ_p, friction becomes important. For

$$t \gg \tau_p,$$

we obtain

$$\overline{p}(t) \to 0,$$

while

$$\overline{(\Delta p)^2} \to m\theta.$$

The Brownian momentum has then lost the systematic memory of the prepared value p_0 and has reached its equilibrium variance.

Coordinate dynamics at the same coarse resolution

The coordinate satisfies

$$\dot{x} = \frac{p}{m}.$$

Integrating the formal solution for the momentum gives

$$x(t) = x_0 + \frac{v_0}{\Gamma}(1 - e^{-\Gamma t}) + \frac{1}{m\Gamma}\int_0^t \left[1 - e^{-\Gamma(t-s)}\right] F(s)\, ds,$$

where $v_0 = p_0/m$. Its coarse-time average is

$$\boxed{\overline{x}(t) = x_0 + \frac{v_0}{\Gamma}(1 - e^{-\Gamma t}).}$$

The fluctuations around this mean satisfy

$$\overline{(x - \overline{x})^2} = \frac{2\theta}{m\Gamma}\left[t - \frac{2}{\Gamma}(1 - e^{-\Gamma t}) + \frac{1}{2\Gamma}(1 - e^{-2\Gamma t})\right].$$

For

$$\tau \ll t \ll \tau_p,$$

the expansion of the exponential gives

$$\overline{x}(t) - x_0 \simeq v_0 t,$$

and

$$\overline{(x - \overline{x})^2} \simeq \frac{2\theta\Gamma}{3m}\, t^3.$$

The total mean-square displacement from the initial position is therefore

$$\overline{(x - x_0)^2} = (\overline{x} - x_0)^2 + \overline{(x - \overline{x})^2} \simeq v_0^2 t^2 + \frac{2\theta\Gamma}{3m}\, t^3.$$

For a particle prepared with non-zero v_0, the leading contribution is the familiar ballistic term,

$$\boxed{\overline{(x - x_0)^2} \simeq v_0^2 t^2.}$$

Thus, there is an intermediate range of elapsed times in which the momentum fluctuations already have a diffusive variance proportional to t, while the coordinate still displays essentially mechanical, ballistic motion. This difference is a consequence of the fact that position contains an additional time integration of the momentum.

For

$$t \gg \tau_p,$$

the initial velocity contributes only a finite displacement v_0/Γ, while the fluctuating part continues to spread. The leading long-time result is

$$\overline{(x - x_0)^2} \simeq \frac{2\theta}{m\Gamma} t.$$

Since $m\Gamma = \gamma$, define the spatial diffusion coefficient

$$\boxed{D \equiv \frac{\theta}{\gamma} = \frac{k_B T}{\gamma}.}$$

Then

$$\boxed{\overline{(x - x_0)^2} \simeq 2Dt, \qquad t \gg \tau_p.}$$

The long-time coordinate dynamics is therefore diffusive.

Macroscopic diffusion time

Once the momentum has relaxed, the Brownian coordinate continues to evolve diffusively. If L is the characteristic spatial scale over which we want to follow this redistribution, the corresponding diffusion time is of order

$$\boxed{\tau_{diff} \sim \frac{L^2}{D}.}$$

The meaning of τ_{diff} is therefore different from that of the momentum-relaxation time τ_p. The time $\tau_p = m/\gamma$ is an intrinsic property of the Brownian particle and its coupling to the surrounding fluid. By contrast, τ_{diff} also depends on the spatial scale L being considered.

For a macroscopic length L,

$$\tau_{diff} \gg \tau_p.$$

Thus, momentum can already be equilibrated while the position distribution is still evolving appreciably over the spatial scale L.

Comparison with the microscopic-resolution behaviour

Only now do we compare the coarse result with what would be seen if the force were resolved on the microscopic scale. This is a **change of temporal resolution**, not another elapsed-time regime within the coarse description above.

If

$$\Delta t_{\mathrm{res}} \lesssim \tau,$$

and one examines a sufficiently short microscopic interval $t \ll \tau$, the force is smooth over that interval,

$$F(t) = F(0) + \dot{F}(0)t + \cdots.$$

Since also $t \ll \tau_p$, the viscous correction is subleading over such a microscopic interval, and

$$p(t) - p_0 \simeq F(0)t.$$

Consequently,

$$\overline{[p(t) - p_0]^2} \propto t^2.$$

At microscopic resolution the local momentum change therefore has the ordinary mechanical t^2 law. At the first coarse resolution, by contrast, many microscopic force fluctuations are contained in one resolved interval and the accumulated momentum variance grows linearly with elapsed time. These statements refer to **different resolutions** and must not be read as two consecutive time regimes of one and the same coarse description.

What does the relaxation mean physically?

The relaxation of $\overline{p}(t)$ does not mean that the actual momentum of an individual Brownian particle literally decays smoothly to zero. The exact $p(t)$ continues to fluctuate because the particle continuously exchanges momentum with the surrounding fluid. The statistical character of the reduced Brownian description is not a consequence of limited measurement precision. Even if $x(t)$ and $p(t)$ were measured with arbitrary precision, they would still constitute only part of the complete state

$$\Gamma_{\mathrm{tot}} = (x, p; \Gamma_{\mathrm{fluid}}).$$

At a given pair (x, p), different microscopic states Γ_{fluid} produce different instantaneous forces on the Brownian particle, and therefore different subsequent trajectories. This means that (x, p) alone is not a closed mechanical state: its future evolution cannot be determined from (x, p) without additional information about the surrounding fluid. The reduced description becomes statistical because the fluid degrees of freedom are not retained explicitly. Their influence is represented instead through the fluctuating force and its statistical properties.

If the complete particle-fluid system is isolated and its microscopic dynamics is time-reversal invariant, the full trajectory remains reversible. To retrace it one must reverse the momenta of the Brownian particle **and** those of the surrounding fluid molecules. Reversing the Brownian momentum alone does not reconstruct the previous fluid microstate and therefore does not make the Brownian trajectory retrace itself.

The relaxation seen in the Brownian variables has a simple mechanical meaning: momentum initially associated with the distinguished particle is redistributed among the many degrees of freedom of the surrounding fluid. The complete microscopic dynamics continues, while the retained Brownian variables approach their stationary statistics.

Temporal coarse-graining does not create this transfer of momentum. Its role is to determine when the microscopic force history need not be resolved explicitly and when a simpler statistical description becomes appropriate.

Second coarse temporal resolution

At the first coarse resolution, momentum remains an explicitly resolved variable even after its statistical distribution has approached equilibrium. We can now make a further reduction by choosing a second, coarser temporal resolution satisfying

$$\boxed{\tau_p \ll \Delta t_{res}^{(2)} \ll \tau_{diff}.}$$

The lower inequality means that momentum relaxes many times within one resolution interval; its instantaneous value is therefore a fast variable and need not be retained explicitly. The upper inequality means that the Brownian position changes only gradually over the spatial scale L during one resolution interval, so its spatial evolution remains resolved. Position therefore becomes the natural state variable of the next description.

This is a genuine second coarse-graining step:

$$(x, p) \quad \rightarrow \quad x.$$

It leads naturally to a description in terms of transition probabilities for the position alone.

1.2 Chapman-Kolmogorov-Smoluchowski equation

At the second coarse resolution, let

$$P(x, t \mid x_0, t_0)$$

denote the conditional probability density for finding the Brownian particle at position x at time t, given that it was at x_0 at time t_0. It is normalized according to

$$\int_{-\infty}^{\infty} P\,(x, t \mid x_0, t_0)\, dx = 1,$$

and at the initial time

$$P(x, t_0 \mid x_0, t_0) = \delta(x - x_0).$$

Choose an intermediate time t_1 with

$$t_0 < t_1 < t.$$

If the present position is sufficient to specify the probability of future positions at this coarse resolution, then the probability of reaching x from x_0 is obtained by summing over all possible intermediate positions x_1,

$$\boxed{P(x, t \mid x_0, t_0) = \int_{-\infty}^{\infty} P\,(x, t \mid x_1, t_1) P(x_1, t_1 \mid x_0, t_0)\, dx_1.}$$

This is the Chapman-Kolmogorov composition law. In the historical Brownian-motion literature it is also associated with the Smoluchowski description.

Its physical content is the Markov property. Once the present coarse state x_1 is specified, the earlier trajectory is not needed to construct the probability distribution of future positions.

The qualification “at this coarse resolution” is essential. On time scales that resolve the momentum, position alone is not a sufficient state variable. Two realizations with the same x but opposite momenta have different short-time futures. On the first coarse scale the appropriate state is therefore (x, p). At the second coarse resolution, $\Delta\, t_{res}^{(2)} \gg \tau_p$, the momentum relaxes on a time scale shorter than the one being resolved. After this fast momentum dynamics has been eliminated, the position x can be used as the state variable of an approximately Markovian reduced description.

Short-time transition moments of the reduced process

The Chapman-Kolmogorov equation by itself does not determine the short-time form of the transition probability. It is satisfied by a broad class of Markov processes. To obtain the Fokker-Planck equation for Brownian motion, we now use the additional physical information already obtained from the Langevin description developed in Section 1.1.

Consider the already constructed position-only process and a short increment Δt. This is not the temporal resolution window $\Delta t_{\mathrm{res}}^{(2)}$ used to define the second coarse description. The increment Δt is the mathematical step used to describe evolution within that reduced process. Physically, it is short compared with the resolved positional evolution, while the momentum-relaxation dynamics on the scale τ_p has already been eliminated.

Let

$$\Delta x = x' - x$$

be the displacement during this increment. For a particle known to be at x at time t, define the conditional transition moments

$$M_n(x,\Delta t) = \int_{-\infty}^{\infty} (\Delta x)^n\, P(x+\Delta x, t+\Delta t \mid x, t)\, d\Delta x = \langle (\Delta x)^n \rangle_{x,t}.$$

The average in this equation is a conditional ensemble average over all realizations that start from the same coarse state x at time t. In Section 1.1, however, the corresponding displacement moments were obtained from the Langevin dynamics after applying the coarse-time averaging to the rapidly fluctuating microscopic motion. The overbar used there and the ensemble average defined here are therefore different averaging operations.

To make the probabilistic description represent the same reduced Brownian dynamics, we adopt the ergodic identification used in the Langevin-to-Smoluchowski construction: at the same coarse level of description, and for the same preparation, the corresponding moments are assumed to coincide,

$$M_n(x,\Delta t) = \langle (\Delta x)^n \rangle_{x,t} \overset{\text{erg}}{=} \overline{(\Delta x)^n}.$$

This identification is an additional statistical hypothesis. It does not follow from the Chapman-Kolmogorov equation. It is the step that matches the time coarse-grained Langevin description to the ensemble description in terms of the transition probability P.

The Langevin calculation now supplies the physical short-time constraints on P. At the second coarse resolution the momentum has already relaxed, and the remaining coordinate dynamics is diffusive. Consequently, for sufficiently small Δt within this reduced description,

$$M_1(x,\Delta t) = A(x)\,\Delta t + O((\Delta t)^2),$$

$$M_2(x,\Delta t) = 2B(x)\,\Delta t + O((\Delta t)^2),$$

whereas the higher moments have no contribution proportional to Δt,

$$M_n(x,\Delta t) = O((\Delta t)^2), \qquad n \geq 3.$$

Thus the systematic displacement and the mean-square displacement both survive to first order in Δt, while every higher conditional moment vanishes after division by Δt in the differential limit. In the equations above, $A(x) = -\frac{1}{\gamma}\frac{dU}{dx}$ (with U describing an external potential) and $B(x) = D$ are the drift and diffusion coefficients already determined from the coarse-grained Langevin dynamics in Section 1.1 (there we assumed $U = 0$). Here, we require the transition probability P to reproduce those same coefficients through its first two conditional moments:

$$\lim_{\Delta t\to 0} \frac{M_1(x,\Delta t)}{\Delta t} = A(x), \qquad \lim_{\Delta t\to 0} \frac{M_2(x,\Delta t)}{2\Delta t} = B(x).$$

The notation $\Delta t \to 0$ is the differential limit of the reduced position-only theory. It must not be interpreted as returning to microscopic times shorter than τ_p, where momentum is again a resolved variable and the coordinate motion is ballistic. The limit means that Δt is made small relative to the slow positional evolution while the separation from the already eliminated momentum-relaxation scale is maintained.

The Chapman-Kolmogorov equation therefore supplies the composition law for P, while the moment conditions inherited from the Langevin dynamics, through the ergodic identification above, select the Brownian diffusive process that leads to the Fokker-Planck equation. A general Markov transition probability need not satisfy these conditions; for a jump process, for example, higher moments may also contain terms proportional to Δt.

1.3 Fokker-Planck equation

The Chapman-Kolmogorov-Smoluchowski equation supplies the composition law for the transition probability

$$P(x, t \mid x_0, t_0).$$

The Brownian short-time moment conditions derived above provide the additional physical input needed to obtain the local differential evolution equation satisfied by this same transition probability. No additional probability distribution is introduced.

Let $F(x)$ be an arbitrary sufficiently smooth function of position and define

$$\mathcal{F}(t) = \int_{-\infty}^{\infty} F(x)\, P(x, t \mid x_0, t_0)\, dx.$$

Its time derivative may be written as

$$\frac{d\mathcal{F}}{dt} = \lim_{\Delta t \to 0} \int_{-\infty}^{\infty} d\,x'\, F(x') \frac{P(x', t+\Delta t \mid x_0, t_0) - P(x', t \mid x_0, t_0)}{\Delta t}.$$

Here Δt is the small increment of the **already constructed reduced time variable**. It is not the temporal resolution window $\Delta t_{\mathrm{res}}^{(2)}$. The latter has already been chosen so that the momentum relaxation is unresolved and the position alone forms the Markov state.

For the first term in the numerator, the Chapman-Kolmogorov-Smoluchowski equation gives

$$P(x', t+\Delta t \mid x_0, t_0) = \int_{-\infty}^{\infty} d\,x\, P(x', t+\Delta t \mid x, t) P(x, t \mid x_0, t_0).$$

For the second term we may use the zero-time transition probability,

$$P(x', t \mid x, t) = \delta(x' - x),$$

and therefore

$$P(x', t \mid x_0, t_0) = \int_{-\infty}^{\infty} d\,x\, P(x', t \mid x, t) P(x, t \mid x_0, t_0).$$

Substitution gives

$$\frac{d\mathcal{F}}{dt} = \int_{-\infty}^{\infty} d\,x\, P(x, t \mid x_0, t_0) \lim_{\Delta t \to 0} \int_{-\infty}^{\infty} d\,x'\, F(x') \frac{P(x', t+\Delta t \mid x, t) - P(x', t \mid x, t)}{\Delta t}.$$

This is the point at which the short-time transition moments introduced at the end of Section 1.2 enter the derivation. The Brownian short-time moment conditions established above imply that the displacement becomes small as Δt decreases within the reduced position-only description. The transition probability $P(x', t+\Delta t \mid x, t)$ is therefore concentrated near $x' = x$. We may therefore expand

$$F(x') = F(x) + (x' - x)F'(x) + \frac{1}{2}(x' - x)^2 F''(x) + \cdots.$$

The zeroth-order term vanishes because both transition probabilities are normalized,

$$\int_{-\infty}^{\infty} P(x', t+\Delta t \mid x, t)\, dx' = \int_{-\infty}^{\infty} P(x', t \mid x, t)\, dx' = 1.$$

For the zero-time transition, all non-zero displacement moments vanish because

$$P(x', t \mid x, t) = \delta(x' - x).$$

The remaining terms are determined by the short-time moments of the same transition probability. In the notation introduced in Section 1.2,

$$A(x) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (x' - x)\, P(x', t+\Delta t \mid x, t)\, dx',$$

and

$$2B(x) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (x' - x)^2\, P(x', t+\Delta t \mid x, t)\, dx'.$$

The Brownian moment conditions inherited from the Langevin description give, for the higher moments,

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{-\infty}^{\infty}(x'-x)^n\,P(x',t+\Delta t\mid x,t)\,dx'=0,\qquad n\geq 3.$$

Consequently,

$$\frac{d\mathcal{F}}{dt}=\int_{-\infty}^{\infty}d\,x\,P(x,t\mid x_0,t_0)[A(x)F'(x)+B(x)F''(x)].$$

On the other hand, from the definition of $\mathcal{F}(t)$,

$$\frac{d\mathcal{F}}{dt}=\int_{-\infty}^{\infty}d\,x\,F(x)\frac{\partial P(x,t\mid x_0,t_0)}{\partial t}.$$

Integrating the term containing $F'(x)$ once by parts and the term containing $F''(x)$ twice gives (recall that $P\to 0$ as $x\to\infty$)

$$\int_{-\infty}^{\infty}d\,x\,F(x)\left\{\frac{\partial P}{\partial t}+\frac{\partial}{\partial x}[A(x)P]-\frac{\partial^2}{\partial x^2}[B(x)P]\right\}=0,$$

where, in this equation only, P is written as a shorthand for $P(x,t\mid x_0,t_0)$. Since $F(x)$ is arbitrary, the expression in braces must vanish. Hence

$$\boxed{\frac{\partial P(x,t\mid x_0,t_0)}{\partial t}=-\frac{\partial}{\partial x}[A(x)P(x,t\mid x_0,t_0)]+\frac{\partial^2}{\partial x^2}[B(x)P(x,t\mid x_0,t_0)].}$$

This is the one-dimensional Fokker-Planck equation. It is not an independent postulate added to the Chapman-Kolmogorov-Smoluchowski description. It is the local differential form obtained from the same transition probability when the first two short-time displacement moments survive to order Δt and the higher moments do not.

It may also be written as a continuity equation for probability,

$$\frac{\partial P}{\partial t}+\frac{\partial J}{\partial x}=0,$$

with probability current

$$\boxed{J=A(x)P-\frac{\partial}{\partial x}[B(x)P].}$$

From the transition probability to a one-time density

Up to this point, the Fokker–Planck equation has been derived for the transition probability density $P(x,t\mid x_0,t_0)$, with the initial position x_0 and initial time t_0 fixed. This quantity gives the probability density of finding the Brownian particle at x at time t, provided that it was at x_0 at t_0.

A more general preparation need not place every realisation at the same initial position. Let the initial positions be distributed according to

$$\rho_0(x_0)=\rho(x_0,t_0).$$

The probability density at the later time is then obtained by summing over all possible initial positions, weighted by their initial probability:

$$\rho(x,t)=\int dx_0\,P(x,t\mid x_0,t_0)\,\rho_0(x_0).$$

The initial coordinate x_0 has therefore not been discarded. For the transition probability $P(x,t\mid x_0,t_0)$, x_0 is fixed. For a general initial preparation, all possible values of x_0 are included through the integral above, with weight $\rho_0(x_0)$.

The Fokker–Planck differential operator acts on the final variables x and t, while x_0 appears only as a parameter of the transition probability. Since the Fokker–Planck equation is linear, we may multiply the equation for P by $\rho_0(x_0)$, integrate over x_0, and move the x-derivatives through the x_0-integral. The resulting equation is therefore the same Fokker–Planck equation for the one-time density $\rho(x,t)$:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x}[A(x)\rho] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[B(x)\rho].$$

The initial preparation now enters through the initial condition

$$\rho(x,t_0) = \rho_0(x).$$

A definite initial position is recovered as a special case. If every realisation starts at x_0,

$$\rho_0(x) = \delta(x - x_0),$$

and the integral above reduces to

$$\rho(x,t) = P(x,t \mid x_0,t_0).$$

Thus P and ρ obey the same Fokker–Planck equation but play different roles. $P(x,t \mid x_0,t_0)$ is the transition probability for a specified initial point, whereas $\rho(x,t)$ is the one-time probability density generated by an arbitrary initial preparation.

For N statistically identical, noninteracting Brownian particles prepared with the same initial distribution, the mean particle-number density is

$$n(x,t) = N\rho(x,t),$$

and, because the equation is linear, $n(x,t)$ obeys the same Fokker–Planck equation.

Brownian particle in an external potential

The coefficients $A(x)$ and $B(x)$ are not introduced independently of the Brownian dynamics discussed in Section 1.1. They are fixed by the systematic drift and by the long-time spreading of the Brownian coordinate.

Let the particle move in a slowly varying potential $U(x)$. On the second coarse scale the momentum has already relaxed, so the systematic motion is overdamped. The deterministic force balance is

$$-\gamma v - U'(x) = 0,$$

and therefore

$$\boxed{A(x) = v_{\text{drift}}(x) = -\frac{1}{\gamma}U'(x).}$$

For a homogeneous fluid the spreading coefficient is independent of position. It is the same spatial diffusion coefficient obtained from the long-time Brownian mean-square displacement,

$$\boxed{B(x) = D.}$$

The Fokker-Planck equation becomes

$$\boxed{\frac{\partial P}{\partial t} = \frac{1}{\gamma}\frac{\partial}{\partial x}[U'(x)P] + D\frac{\partial^2 P}{\partial x^2}.}$$

The corresponding current is

$$J = -\frac{U'(x)}{\gamma}P - D\frac{\partial P}{\partial x}.$$

The first term is the systematic drift produced by the external force; the second is the diffusive flux produced by the thermal fluctuations of the surrounding fluid.

Einstein relation from equilibrium

The value of D found from the Langevin dynamics can also be recovered directly from the equilibrium limit of the same Fokker-Planck equation. At long times, for an ergodic Brownian process in a confining potential, the transition probability loses memory of the initial position and approaches a stationary distribution. In equilibrium the probability current vanishes,

$$J_{\mathrm{eq}} = 0.$$

Hence

$$-\frac{U'}{\gamma} P_{\mathrm{eq}} - D\frac{dP_{\mathrm{eq}}}{dx} = 0,$$

or

$$\frac{d\ln P_{\mathrm{eq}}}{dx} = -\frac{U'(x)}{\gamma D}.$$

Integration gives

$$P_{\mathrm{eq}}(x) \propto \exp\left[-\frac{U(x)}{\gamma D}\right].$$

Thermal equilibrium requires the Boltzmann form

$$P_{\mathrm{eq}}(x) \propto \exp\left[-\frac{U(x)}{k_B T}\right].$$

Comparison yields

$$\boxed{D = \frac{k_B T}{\gamma}.}$$

This is precisely the coefficient obtained earlier from the long-time coordinate dynamics. The agreement is physically important: the diffusive spreading and the viscous relaxation are two consequences of the coupling of the Brownian particle to the same thermal environment.

Free diffusion

For a free Brownian particle, $U(x) = 0$, and the Fokker-Planck equation reduces to

$$\boxed{\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial x^2}.}$$

With the initial condition already built into the transition probability,

$$P(x, t_0 \mid x_0, t_0) = \delta(x - x_0),$$

the solution is

$$\boxed{P(x, t \mid x_0, t_0) = \frac{1}{\sqrt{4\pi D(t - t_0)}} \exp\left[-\frac{(x - x_0)^2}{4D(t - t_0)}\right].}$$

The distribution remains normalized and centred at x_0. Its second moment is

$$\int_{-\infty}^{\infty} (x - x_0)^2 \, P(x, t \mid x_0, t_0)\, dx = 2D(t - t_0).$$

Thus, the same transition probability that satisfies the Chapman-Kolmogorov-Smoluchowski composition law also obeys the diffusion equation and reproduces Einstein's linear growth of the mean-square displacement.

In three dimensions,

$$\int r^2 P(\mathbf{r}, t \mid \mathbf{r}_0, t_0)\, d^3 r = 6D(t - t_0),$$

where $r = |\mathbf{r} - \mathbf{r}_0|$.

The free diffusion result also recovers the macroscopic diffusion time introduced in Section 1.1. A displacement of order L requires

$$\tau_{diff} \sim \frac{L^2}{D},$$

confirming explicitly that the position can continue to redistribute over the scale L on times much longer than τ_p.

The Brownian problem therefore contains the hierarchy

$$\boxed{\tau' \ll \tau \ll \tau_p \ll \tau_{diff}.}$$

Here τ', τ, and τ_p are characteristic microscopic or Brownian time scales of the particle-fluid system, whereas τ_{diff} is scale-dependent because it also contains the chosen length L. The hierarchy therefore compares intrinsic relaxation scales with the slower positional evolution on the spatial scale of interest.

1.4 Random walks

A random walk is a simple stochastic model for motion produced by a sequence of elementary displacements. It provides a direct route from microscopic random steps to macroscopic diffusion. We first consider a walk in which a jump occurs after every fixed time interval. We then consider a continuous-time walk in which the jump times themselves are random.

Throughout this subsection, angle brackets denote ensemble averages over many independent realizations of the same random walk. This is different from the coarse-time average denoted by an overbar in Section 1.1.

1.4.1 Discrete-time symmetric random walk

Consider a particle on a one-dimensional lattice with spacing a. The particle starts at the origin, $x_0 = 0$. At each step it moves either one lattice spacing to the right or one lattice spacing to the left, with equal probability. If ξ_i is the displacement during the ith step, then

$$\xi_i = \begin{cases} +a, & \text{with probability } 1/2, \\ -a, & \text{with probability } 1/2. \end{cases}$$

Let the time between two successive steps be the fixed interval τ_s. After N steps,

$$t = N\tau_s,$$

and the total displacement is

$$x_N = \sum_{i=1}^{N} \xi_i.$$

The subscript N reminds us that, in this discrete-time model, exactly N jumps have occurred.

Probability distribution after a fixed number of steps

Let N_R be the number of jumps to the right and N_L the number of jumps to the left. They satisfy

$$N_R + N_L = N.$$

The displacement after these N jumps is

$$x_N = a(N_R - N_L).$$

Introduce the integer

$$m = N_R - N_L,$$

so that $x_N = ma$. Solving the two equations

$$N_R + N_L = N, \qquad N_R - N_L = m$$

for N_R and N_L gives

$$N_R = \frac{N+m}{2}, \qquad N_L = \frac{N-m}{2}.$$

Therefore N and m must have the same parity. For example, after an even number of jumps only even values of m are possible.

A particular sequence containing N_R right jumps and N_L left jumps has probability $(1/2)^N$. The number of distinct sequences containing the same numbers N_R and N_L is

$$\frac{N!}{N_R!\,N_L!}.$$

Hence the probability of finding the walker at $x_N = ma$ after exactly N steps is

$$P_N(m) = \frac{N!}{[(N+m)/2]!\,[(N-m)/2]!}\left(\frac{1}{2}\right)^N.$$

This is the binomial distribution written in terms of the displacement.

Mean and mean-square displacement

For one step,

$$\langle \xi_i \rangle = \frac{1}{2}a + \frac{1}{2}(-a) = 0,$$

and

$$\langle \xi_i^2 \rangle = \frac{1}{2}a^2 + \frac{1}{2}a^2 = a^2.$$

Therefore

$$\langle x_N \rangle = \left\langle \sum_{i=1}^{N} \xi_i \right\rangle = \sum_{i=1}^{N} \langle\, \xi_i \rangle = 0.$$

To calculate the mean-square displacement, first expand the square:

$$x_N^2 = \left(\sum_{i=1}^{N} \xi_i \right)^2 = \sum_{i=1}^{N} \xi_i^2 + 2 \sum_{i<j} \xi_i\, \xi_j.$$

Taking the ensemble average gives

$$\langle x_N^2 \rangle = \sum_{i=1}^{N} \langle\, \xi_i^2 \rangle + 2 \sum_{i<j} \langle\, \xi_i \xi_j \rangle.$$

Different steps are statistically independent. Therefore, for $i \neq j$,

$$\langle \xi_i \xi_j \rangle = \langle \xi_i \rangle \langle \xi_j \rangle = 0.$$

Only the first sum remains:

$$\langle x_N^2 \rangle = \sum_{i=1}^{N} a^2 = Na^2.$$

Thus

$$\boxed{\langle x_N^2 \rangle = N a^2.}$$

Using $N = t/\tau_s$,

$$\langle x_N^2 \rangle = \frac{a^2}{\tau_s} t.$$

The one-dimensional diffusion law is conventionally written as

$$\langle x^2(t) \rangle = 2Dt.$$

Comparison with the random-walk result identifies

$$\boxed{D = \frac{a^2}{2\tau_s}.}$$

The linear growth of the mean-square displacement therefore follows directly from independent symmetric steps.

Large-N form of the spatial distribution

For large N, the binomial distribution can be approximated by a Gaussian. Starting from

$$P_N(m) = \frac{N!}{[(N+m)/2]!\,[(N-m)/2]!} 2^{-N},$$

use Stirling's approximation

$$\ln n! \simeq n \ln n - n + \frac{1}{2}\ln(2\pi n).$$

Set

$$\varepsilon = \frac{m}{N}.$$

Then

$$\frac{N \pm m}{2} = \frac{N}{2}(1 \pm \varepsilon).$$

Substituting the Stirling form into $\ln P_N(m)$ and collecting the terms that depend on m gives

$$\ln P_N(m) \simeq -\frac{1}{2}\ln\left(\frac{\pi N}{2}\right) - \frac{N}{2}[(1+\varepsilon)\ln(1+\varepsilon) + (1-\varepsilon)\ln(1-\varepsilon)].$$

For displacements much smaller than the maximum possible displacement, $|m| \ll N$, we have $|\varepsilon| \ll 1$. Using

$$\ln(1+\varepsilon) = \varepsilon - \frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3} - \frac{\varepsilon^4}{4} + O(\varepsilon^5),$$

$$\ln(1-\varepsilon) = -\varepsilon - \frac{\varepsilon^2}{2} - \frac{\varepsilon^3}{3} - \frac{\varepsilon^4}{4} + O(\varepsilon^5),$$

one obtains

$$(1+\varepsilon)\ln(1+\varepsilon) + (1-\varepsilon)\ln(1-\varepsilon) = \varepsilon^2 + \frac{\varepsilon^4}{6} + O(\varepsilon^6).$$

Keeping the leading term in the central region $|m| \ll N$ gives

$$P_N(m) \simeq \sqrt{\frac{2}{\pi N}} \exp\left(-\frac{m^2}{2N}\right).$$

For fixed N, allowed values of m differ by 2, because m must have the same parity as N. The corresponding allowed positions therefore differ by

$$\Delta x = 2a.$$

Let $\rho(x,t)$ be the continuous probability density on length scales large compared with a. The probability associated with one allowed lattice position is then approximately

$$P_N(m) = \rho(x,t)\,\Delta x \simeq \rho(x,t)(2a).$$

Hence

$$\rho(x,t) \simeq \frac{P_N(m)}{2a}.$$

Using $x = ma$ and $Na^2 = 2Dt$ gives

$$\boxed{\rho(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right).}$$

This is the same Gaussian diffusion kernel obtained in Section 1.3.

1.4.2 Continuous-time symmetric random walk

In the discrete-time model, jumps occur at the predetermined times $\tau_s, 2\tau_s, 3\tau_s, \ldots$. We now remove this restriction. Jumps occur at random times, while the spatial jump remains $+a$ or $-a$ with equal probability.

We consider the simplest continuous-time case: jumps occur independently with a constant total rate ν. The meaning of the rate ν is fixed by the short-time probabilities. During a sufficiently small time interval Δt,

$$P(\text{one jump during } \Delta t) = \nu\Delta t + O((\Delta t)^2),$$

$$P(\text{no jump during } \Delta t) = 1 - \nu\Delta t + O((\Delta t)^2),$$

and

$$P(\text{two or more jumps during } \Delta t) = O((\Delta t)^2).$$

These relations mean that, to first order in Δt, at most one jump must be considered.

Number of jumps occurring by time t

Let $P_n(t)$ denote the probability that exactly n jumps have occurred between time 0 and time t. At $t = 0$ no jump has yet occurred, so

$$P_0(0) = 1, \qquad P_n(0) = 0 \quad (n \geq 1).$$

To have exactly $n \geq 1$ jumps at time $t + \Delta t$, there are two possibilities to first order in Δt:

1. there were already n jumps at time t and no jump occurred during the next interval Δt;
2. there were $n - 1$ jumps at time t and one jump occurred during the next interval Δt.

Therefore

$$P_n(t + \Delta t) = P_n(t)(1 - \nu\Delta t) + P_{n-1}(t)\nu\Delta t + O((\Delta t)^2).$$

Subtracting $P_n(t)$ from both sides gives

$$P_n(t + \Delta t) - P_n(t) = \nu\Delta t\,[P_{n-1}(t) - P_n(t)] + O((\Delta t)^2).$$

Dividing by Δt and taking the limit $\Delta t \to 0$ gives

$$\boxed{\frac{dP_n}{dt} = \nu P_{n-1} - \nu P_n, \qquad n \geq 1.}$$

For $n = 0$, the only way to remain in the state with zero jumps is for no jump to occur. Hence

$$P_0(t + \Delta t) = P_0(t)(1 - \nu\Delta t) + O((\Delta t)^2),$$

which gives

$$\frac{dP_0}{dt} = -\nu P_0.$$

With $P_0(0) = 1$,

$$\boxed{P_0(t) = e^{-\nu t}.}$$

For $n \geq 1$, rewrite the differential equation as

$$\frac{dP_n}{dt} + \nu P_n = \nu P_{n-1}.$$

Multiplication by the integrating factor $e^{\nu t}$ gives

$$\frac{d}{dt}(e^{\nu t} P_n) = \nu e^{\nu t} P_{n-1}.$$

For $n = 1$, using $P_0(t) = e^{-\nu t}$,

$$\frac{d}{dt}(e^{\nu t} P_1) = \nu.$$

Since $P_1(0) = 0$,

$$P_1(t) = e^{-\nu t} \nu t.$$

Now suppose that the solution for $n - 1$ has the form

$$P_{n-1}(t) = e^{-\nu t} \frac{(\nu t)^{n-1}}{(n-1)!}.$$

Then

$$\frac{d}{dt}(e^{\nu t} P_n) = \nu \frac{(\nu t)^{n-1}}{(n-1)!} = \frac{\nu^n t^{n-1}}{(n-1)!}.$$

Integrating from 0 to t, and using $P_n(0) = 0$, gives

$$e^{\nu t} P_n(t) = \frac{\nu^n}{(n-1)!} \int_0^t s^{n-1}\, ds = \frac{(\nu t)^n}{n!}.$$

Therefore, for every $n = 0,1,2, \ldots$,

$$\boxed{P_n(t) = e^{-\nu t} \frac{(\nu t)^n}{n!}.}$$

This is the Poisson distribution for the number of jumps occurring by time t.

Its normalization follows directly:

$$\sum_{n=0}^{\infty} P_n(t) = e^{-\nu t} \sum_{n=0}^{\infty} \frac{(\nu t)^n}{n!} = e^{-\nu t} e^{\nu t} = 1.$$

The mean number of jumps must also be calculated from this distribution:

$$\langle n(t) \rangle = \sum_{n=0}^{\infty} n\, P_n(t).$$

Substituting the Poisson probability,

$$\langle n(t) \rangle = e^{-\nu t} \sum_{n=1}^{\infty} n \frac{(\nu t)^n}{n!}.$$

Since $n/n! = 1/(n-1)!$,

$$\langle n(t)\rangle = e^{-\nu t}(\nu t)\sum_{n=1}^{\infty}\frac{(\nu t)^{n-1}}{(n-1)!}.$$

With $m = n - 1$,

$$\langle n(t)\rangle = e^{-\nu t}(\nu t)\sum_{m=0}^{\infty}\frac{(\nu t)^{m}}{m!}.$$

The series is $e^{\nu t}$, so

$$\boxed{\langle n(t)\rangle = \nu t.}$$

Mean-square displacement in continuous time

Each jump is still either $+a$ or $-a$ with equal probability. Consider first the subset of realizations in which exactly n jumps have occurred. For these realizations the spatial calculation is identical to the discrete-time calculation above, with N replaced by n. The mean-square displacement within this subset is therefore

$$na^2.$$

At time t, the probability of belonging to this subset is $P_n(t)$. The mean-square displacement over all realizations is obtained by multiplying the mean-square displacement for each value of n by its probability and then summing over all possible n:

$$\langle x^2(t)\rangle = \sum_{n=0}^{\infty} P_n\,(t)\,na^2.$$

Factor out a^2:

$$\langle x^2(t)\rangle = a^2\sum_{n=0}^{\infty} n\,P_n(t).$$

The sum is exactly the mean number of jumps calculated above. Therefore

$$\boxed{\langle x^2(t)\rangle = \nu a^2 t.}$$

Comparison with $\langle x^2(t)\rangle = 2Dt$ gives

$$\boxed{D = \frac{\nu a^2}{2}.}$$

Exact lattice evolution equation

Let $W_j(t)$ denote the probability that the walker is at the lattice site

$$x_j = ja$$

at time t. During a short interval Δt, the walker can reach site j in three ways to first order in Δt:

1. it was already at j and no jump occurred;
2. it was at $j - 1$ and made one jump to the right;
3. it was at $j + 1$ and made one jump to the left.

Because the total jump rate is ν and the two directions are equally likely, the rate for a right jump is $\nu/2$ and the rate for a left jump is $\nu/2$. Hence

$$W_j(t+\Delta t) = (1-\nu\Delta t)W_j(t) + \frac{\nu\Delta t}{2}W_{j-1}(t) + \frac{\nu\Delta t}{2}W_{j+1}(t) + O((\Delta t)^2).$$

Subtract $W_j(t)$, divide by Δt, and take $\Delta t \to 0$:

$$\boxed{\frac{dW_j}{dt} = \frac{\nu}{2}\left[W_{j-1} - 2W_j + W_{j+1}\right].}$$

This equation is exact for the continuous-time lattice walk defined above.

Continuum space limit

The lattice equation describes probabilities at discrete positions separated by a. To obtain a continuous spatial description, consider variations on length scales much larger than a. Introduce a probability density $\rho(x,t)$ by

$$W_j(t) \simeq a\rho(x,t), \qquad x = ja.$$

This relation is required because W_j is a probability, whereas ρ is a probability per unit length: the probability in one lattice interval of width a is approximately $a\rho$.

At the neighbouring sites,

$$W_{j\pm1}(t) \simeq a\rho(x \pm a, t).$$

For a density that varies slowly over one lattice spacing,

$$\rho(x+a,t) = \rho(x,t) + a\frac{\partial\rho}{\partial x} + \frac{a^2}{2}\frac{\partial^2\rho}{\partial x^2} + O(a^3),$$

$$\rho(x-a,t) = \rho(x,t) - a\frac{\partial\rho}{\partial x} + \frac{a^2}{2}\frac{\partial^2\rho}{\partial x^2} + O(a^3).$$

Adding these two expansions and subtracting $2\rho(x,t)$ gives

$$\rho(x-a,t) - 2\rho(x,t) + \rho(x+a,t) = a^2\frac{\partial^2\rho}{\partial x^2} + O(a^4).$$

Substitution into the lattice equation yields

$$\frac{\partial\rho}{\partial t} = \frac{\nu a^2}{2}\frac{\partial^2\rho}{\partial x^2} + O(a^4\nu).$$

The diffusive continuum limit is taken as

$$a \to 0, \qquad \nu \to \infty, \qquad \frac{\nu a^2}{2} = D \quad \text{fixed.}$$

In this limit the correction term vanishes and the continuum equation becomes

$$\boxed{\frac{\partial\rho}{\partial t} = D\frac{\partial^2\rho}{\partial x^2}.}$$

The continuous-time random walk therefore gives the same diffusion equation as the Brownian Langevin description.

Connection with the short-time transition moments of Section 1.2

The continuous-time lattice walk also shows explicitly why a Markov process need not already have the Fokker-Planck form.

Start the walker at a lattice site x and consider a short interval Δt. To first order in Δt,

$$\Delta x = 0 \quad \text{with probability} \quad 1 - \nu\Delta t,$$

$$\Delta x = +a \quad \text{with probability} \quad \frac{\nu\Delta t}{2},$$

$$\Delta x = -a \quad \text{with probability} \quad \frac{\nu\Delta t}{2}.$$

The probability of two or more jumps contributes only terms of order $(\Delta t)^2$.

The first transition moment is therefore

$$M_1 = 0(1 - \nu\Delta t) + a\frac{\nu\Delta t}{2} - a\frac{\nu\Delta t}{2} + O((\Delta t)^2) = O((\Delta t)^2).$$

Hence

$$A = \lim_{\Delta t \to 0} \frac{M_1}{\Delta t} = 0.$$

The second transition moment is

$$M_2 = 0^2(1 - \nu\Delta t) + a^2\frac{\nu\Delta t}{2} + a^2\frac{\nu\Delta t}{2} + O((\Delta t)^2),$$

so

$$M_2 = \nu a^2 \Delta t + O((\Delta t)^2).$$

Therefore, in the notation of Section 1.2,

$$\boxed{2B = \lim_{\Delta t \to 0} \frac{M_2}{\Delta t} = \nu a^2 = 2D.}$$

At finite lattice spacing, higher even moments also contain terms proportional to Δt. For example,

$$M_4 = \nu a^4 \Delta t + O((\Delta t)^2),$$

and therefore

$$\lim_{\Delta t \to 0} \frac{M_4}{\Delta t} = \nu a^4.$$

This quantity is not zero for a jump process with finite a. In the diffusive continuum limit, however,

$$\nu a^2 = 2D$$

is held fixed while $a \to 0$. Consequently,

$$\nu a^4 = (\nu a^2)a^2 = 2D\,a^2 \to 0.$$

For any even order $2m$ with $m \geq 2$, the two one-jump contributions add:

$$M_{2m} = \nu a^{2m} \Delta t + O((\Delta t)^2).$$

Therefore

$$\lim_{\Delta t \to 0} \frac{M_{2m}}{\Delta t} = \nu a^{2m} = (\nu a^2)a^{2m-2} = 2D\,a^{2m-2} \to 0$$

in the diffusive continuum limit.

For any odd order $2m + 1$, the contributions from the right and left jumps cancel:

$$M_{2m+1} = \frac{\nu\Delta t}{2} a^{2m+1} + \frac{\nu\Delta t}{2}(-a)^{2m+1} + O((\Delta t)^2) = O((\Delta t)^2).$$

Hence

$$\lim_{\Delta t \to 0} \frac{M_{2m+1}}{\Delta t} = 0.$$

In the diffusive continuum limit, only the first and second transition moments can therefore remain proportional to Δt. This is the short-time condition used in Section 1.2 to pass from the Chapman-Kolmogorov-Smoluchowski description to the Fokker-Planck equation. It is worth keeping in mind this important distinction: the lattice master equation already describes a Markov process, but a Markov process does not automatically obey a Fokker-Planck equation. The Fokker-Planck form emerges here only after the additional long-wavelength, diffusive scaling has been imposed.

For the symmetric walk considered here,

$$A = 0, \qquad B = D.$$

For the thermal Brownian particle studied in Section 1.1,

$$D = \frac{k_B T}{\gamma},$$

so the same coefficient is

$$\boxed{B = \frac{k_B T}{\gamma}.}$$

Connection with Section 3: random processes

The random-walk examples introduced above will reappear in Section 3 in a more general setting. There the object of study will no longer be Brownian motion itself, but random processes and the probability laws that describe their evolution in time.

The discrete time walk already contains the basic structure of a Markov chain: the probability of the next position is determined by the present position and by a transition rule. The continuous-time walk adds a random event time. If $N(t)$ denotes the random number of jumps accumulated by time t, then $N(t)$ is a Poisson process with rate ν,

$$\langle N(t) \rangle = \nu t, \qquad \mathrm{Var}[N(t)] = \nu t.$$

For $\nu t \gg 1$, its probability distribution approaches the Gaussian form derived above. The spatial random walk gives the corresponding result for the sum of many independent displacements: its binomial distribution approaches the Gaussian diffusion kernel.

Section 3 will formulate these ideas without reference to a particular Brownian model. Transition probabilities, the Markov property, master equations, moments, correlation functions, counting processes and continuous limits will be treated as general properties of stochastic processes. The random walks of this section will then serve as concrete examples for that broader formalism.

Transition to Section 2: from Brownian motion to kinetic equations

Brownian motion has introduced several ideas that will now be used in a more systematic many-particle setting.

The complete Brownian system has a microscopic state

$$\Gamma_{\mathrm{tot}} = (x, p; \Gamma_{\mathrm{fluid}}).$$

The Brownian variables x and p are genuine components of this state, but they are not sufficient to determine their own future evolution: the force on the Brownian particle also depends on the microscopic state of the fluid. At the first coarse resolution the detailed force history is represented statistically while (x, p) remain resolved. At the second coarse resolution the momentum is itself fast relative to spatial evolution, and the state is reduced further to x alone.

This illustrates a general problem of statistical mechanics: we are often interested in only a small set of variables, while their exact dynamics is coupled to many other degrees of freedom. A useful reduced theory must identify which variables are retained, which are not followed explicitly, and under what scale separation the retained variables admit a closed statistical evolution.

Section 2 begins with the corresponding problem for an interacting system of N classical particles. The complete microscopic state is

$$\Gamma = (\mathbf{r}_1, \mathbf{p}_1, \dots, \mathbf{r}_N, \mathbf{p}_N),$$

a point in the $6N$-dimensional phase space. Unlike the Brownian discussion, we shall first write the **exact** statistical evolution in this full phase space. The N-particle probability density

$$\rho_N(\Gamma, t)$$

obeys the Liouville equation. At this level there is no kinetic approximation and no irreversible collision term: the probability distribution is transported by the same reversible Hamiltonian flow that generates the microscopic trajectories.

The next step is to ask for a reduced description of one particle. We define the one-particle distribution by integrating over the remaining $N-1$ particles,

$$f_1(1,t) = N\int d2\cdots dN\, \rho_N(1,2,\dots,N,t),$$

where $1 \equiv (\mathbf{r}_1, \mathbf{p}_1)$, $2 \equiv (\mathbf{r}_2, \mathbf{p}_2)$, and so on.

This operation is closely analogous in spirit to retaining the Brownian variables while not following every fluid molecule. But an important new feature appears. For interacting particles, the force acting on particle 1 depends on the state of other particles. Consequently, the exact evolution equation for f_1 cannot in general be written in terms of f_1 alone: it contains the two-particle distribution f_2. The equation for f_2 contains f_3, and the resulting sequence forms the BBGKY hierarchy.

Thus, the central question of kinetic theory is not simply how to write an equation for a one-particle distribution. The exact reduced equation is readily obtained. The real question is **how and under what physical conditions that equation can be closed**.

Different kinetic equations correspond to different physical mechanisms and different controlled reductions. In a long-range system, collective interactions may be represented by a self-consistent mean field, leading to the Vlasov equation. In a dilute gas with short-range interactions, isolated binary collisions and the statistical structure of incoming pairs lead to the Boltzmann equation. The Boltzmann H theorem will then show explicitly how irreversible relaxation can appear in a closed kinetic description even though the underlying N-particle Hamiltonian dynamics remains reversible.

Relaxation hierarchy and the kinetic time scale

The Brownian example also suggests a more specific organisation of the many-particle problem. Bogoliubov formulated the evolution of a many-particle system in terms of successive relaxation stages: a microscopic stage, a kinetic stage and, at still longer times, a hydrodynamic stage. The physical variables required at each stage become progressively fewer as the faster processes have already relaxed.

For Brownian motion the corresponding hierarchy is

$$\boxed{\tau' \ll \tau \ll \tau_p \ll \tau_{diff}.}$$

Here τ', τ, and τ_p are characteristic microscopic or Brownian time scales of the particle-fluid system, whereas $\tau_{diff} \sim L^2/D$ depends also on the spatial scale L over which the position is followed. The two reduced descriptions therefore use temporal resolution windows that lie between the faster processes being eliminated and the slower evolution that remains resolved,

$$\tau \ll \Delta t_{res}^{(1)} \ll \tau_p,$$

and

$$\tau_p \ll \Delta t_{res}^{(2)} \ll \tau_{diff}.$$

For a dilute gas the analogous physical hierarchy can be written as

$$\tau_{\mathrm{int}} \ll \tau_{\mathrm{loc}} \ll \tau_H, \qquad \tau_{\mathrm{loc}} \sim \tau_{\mathrm{mfp}}.$$

Here τ_{int} is the duration of an individual binary molecular interaction. The mean-free-flight time τ_{mfp} is the microscopic time between successive collisions of one molecule. The local-equilibration

time τ_{loc} is the time over which collisions drive the local momentum distribution toward a local Maxwell form. For a dilute gas τ_{loc} is of the order of τ_{mfp}, although the two quantities need not be numerically identical because several collisions may be required for local equilibration.

For representative dilute-gas conditions,

$$\tau_{\mathrm{int}} \sim 10^{-12}\ \mathrm{s}, \qquad \tau_{\mathrm{loc}} \sim \tau_{\mathrm{mfp}} \sim 10^{-10}\ \mathrm{s}.$$

Once $t \gtrsim \tau_{\mathrm{loc}}$, the local momentum distribution is approximately Maxwellian and a hydrodynamic description becomes possible. This onset of the hydrodynamic regime must be distinguished from the much longer macroscopic evolution time τ_H, over which the fields $n(\mathbf{r},t)$, $\mathbf{u}(\mathbf{r},t)$ and $T(\mathbf{r},t)$ vary appreciably. The value of τ_H is system- and process-dependent; it may be set by an acoustic propagation time, a viscous relaxation time, a thermal-diffusion time or another macroscopic transport time.

As in the Brownian problem, the chosen temporal resolutions are separate from these intrinsic physical times. A kinetic description may use

$$\tau_{\mathrm{int}} \ll \Delta t_{\mathrm{kin}} \ll \tau_{\mathrm{loc}}.$$

At this resolution the detailed continuous force history during one binary interaction is not resolved. A collision can be represented by its net scattering effect on the incoming momenta, while the slower kinetic evolution of the one-particle distribution remains resolved.

After local Maxwellisation, a hydrodynamic description may use a still coarser temporal resolution,

$$\boxed{\tau_{loc} \ll \Delta t_{hyd} \ll \tau_H,}$$

and a corresponding spatial resolution

$$\boxed{\lambda_{mfp} \ll \Delta x_{hyd} \ll L.}$$

At this resolution the local-Maxwellisation process itself is no longer resolved. The detailed momentum dependence of $f_1(\mathbf{r},\mathbf{p},t)$ is fast relative to the hydrodynamic evolution, while its low-order moments—the number density $n(\mathbf{r},t)$, the flow velocity $\mathbf{u}(\mathbf{r},t)$ and the temperature $T(\mathbf{r},t)$—remain slow.

The resulting sequence can be written schematically as

$$\boxed{\rho_N(\Gamma,t) \ \rightarrow\ \{f_1, f_2, \dots\} \ \rightarrow\ f_1 \ \rightarrow\ \{n, \mathbf{u}, T\}.}$$

The first arrow should not be confused with a kinetic approximation. Integrating the full N-particle distribution over unobserved particle variables produces the exact reduced distributions $f_1, f_2, \dots$ and leads to the exact BBGKY hierarchy. The kinetic approximation enters only when the separation of relaxation times is used to express the faster correlations in terms of the slower one-particle distribution and thereby close that hierarchy. The subsequent passage from f_1 to hydrodynamic fields becomes possible after local equilibration on the time τ_{loc} and uses the separation between that local-relaxation scale and the much slower hydrodynamic evolution time τ_H.

This is the many-particle counterpart of the two coarse descriptions encountered for the Brownian particle. The correspondence is not literal: the physical variables and the closure assumptions are different. The common structure is the identification of intrinsic relaxation times, followed by the choice of temporal and spatial resolutions that retain the slower dynamics while faster processes are represented through their accumulated effect.

The Brownian example therefore prepares the logic of Section 2 without replacing it. It has shown, in the simplest possible setting, that

complete microscopic state → retained variables → statistical evolution
→ closed reduced equation

requires a careful separation of dynamical scales and a precise statement of what has been retained and what has been omitted. Section 2 will develop the same logic for the full hierarchy of many-particle distribution functions, starting from exact Liouville dynamics before introducing any kinetic closure.