Physics 127b: Statistical Mechanics

Brownian Motion

Brownian motion is the motion of a particle due to the buffeting by the molecules in a gas or liquid. The particle must be small enough that the effects of the discrete nature of matter are apparent, but large compared to the molecular scale (pollen in the early experiments, various plastic beads these days). It is a convenient example to display the residual effects of molecular noise on macroscopic degrees of freedom. I will use this example to investigate the type of physics encountered, and the tools used to treat the fluctuations.

Random Walk

The first observation of Brownian motion is that the particle under the microscope appears to perform a "random walk", and it is first useful to study this aspect in its simplest form.

Lets consider first a one dimensional random walk, consisting of n jumps of $\pm l$ along the x axis. We take n to be even (the odd case is essentially the same, but differs in minor details). For the particle after n jumps to be at $x = ml$ there must have been $\frac{1}{2}(n+m)$ forward jumps, and $\frac{1}{2}(n-m)$ backwards jumps (in any order), and m must be even. The probability of arriving at $x = ml$ is therefore

$$p_n(m) = \frac{n!}{\left[\frac{1}{2}(n-m)\right]!\left[\frac{1}{2}(n+m)\right]!}. \tag{1}$$

For large m, n Stirling's approximation $n! \simeq (2\pi n)^{1/2}(n/e)^n$ gives

$$p_n(m) = \frac{2}{\sqrt{2\pi n}} e^{-m^2/2n}. \tag{2}$$

This is a Gaussian probability centered around $m = 0$ (the most probable and mean position is the origin) and the mean square displacement $\langle m^2 \rangle = n$, or

$$\langle x^2 \rangle = nl^2. \tag{3}$$

For large n the discreteness of the displacements is unimportant compared to the root mean square distance of the walk. Transforming to a continuous variable x and a probability density $p(x,t)$ using $p_n(m) = p(x) \times 2l$ (since the interval between the discrete results is $dx = 2l$) and introducing time supposing there are n jumps in time t

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

where we have written

$$nl^2/2t = D. \tag{5}$$

We recognize that this is the expression for diffusion, with p satisfying

$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x^2}, \qquad p(x, t=0) = \delta(x) \tag{6}$$

with the diffusion constant D. In terms of D

$$\langle x^2 \rangle = 2Dt. \tag{7}$$

These results are readily extended to 3 dimensions, since we can consider a walk with steps $(\pm l, \pm l, \pm l)$ for example, so that the walk is the product of walks in each dimension. The mean square distance gone after n walks is again $\langle r^2 \rangle = nL^2$ with $L = \sqrt{3}l$ the length of each step. The probability distribution $p(\vec{x}, t)$ satisfies the $3d$ diffusion equation

$$\frac{\partial p}{\partial t} = \frac{1}{(4\pi Dt)^{3/2}} \exp\left(-\frac{r^2}{4Dt}\right) \tag{8}$$

with $r^2 = x^2 + y^2 + z^2$. This equation is simply the product of three $1d$ diffusion equations with $D = nl^2/2t$ as before. The means square distance is

$$\langle r^2 \rangle = 6Dt. \tag{9}$$

(The results in $2d$ can similarly be constructed.)

The fact that the mean displacement is zero, and the mean square displacement grows linearly in time can be derived by very simple arguments. Lets consider the two dimensional case of a random walk consisting of n vectors of length s but with arbitrary angles θ_i taken from a uniform probability distribution. The total displacement in the x direction is

$$X = \sum_i s \cos\theta_i. \tag{10}$$

Clearly $\langle X \rangle = 0$ since $\cos\theta_i$ is equally likely to be positive or negative. On the other hand

$$\langle X^2 \rangle = \left\langle \left(\sum_i s \cos\theta_i \right)^2 \right\rangle = s^2 \left\langle \sum_i \cos\theta_i \sum_j \cos\theta_j \right\rangle \tag{11}$$

$$= s^2 \left\langle \sum_i (\cos\theta_i)^2 \right\rangle = ns^2/2 \tag{12}$$

where we have used the fact that $\left\langle \sum_{i,j\neq i} \cos\theta_i \cos\theta_j \right\rangle = 0$ since again each $\cos\theta_i$ is equally likely to be positive or negative. Thus the mean square distance is

$$\langle R^2 \rangle = \langle X^2 + Y^2 \rangle = ns^2. \tag{13}$$

This specific result is useful in adding complex numbers with random phases: the average amplitude is zero, and the mean square magnitude (the "intensity") scales linearly with the number of vectors.

Some general nomenclature

The position $x(t)$ in a one dimensional random walk forms a one dimensional *random process*—in general a scalar function $y(t)$ for which the future data is not determined uniquely by the known initial data.

The random process is in general characterized by probability distributions p_1, $p_2 \ldots$ such that

$$p_n(y_1, t_1; y_2, t_2 \ldots; y_n, t_n) dy_1 dy_2 \ldots dy_n \tag{14}$$

is the probability that a single process drawn from the ensemble of processes will take on a value between y_1 and $y_1 + dy_1$ at t_1 etc. The different p_n are related by

$$\int_{-\infty}^{\infty} p_n dy_j \to p_{n-1}. \tag{15}$$

Ensemble averages are determined by integrating over the appropriate distribution, e.g. for the *mean* and the *two point correlation function*

$$\langle y(t_1) \rangle = \int_{-\infty}^{\infty} y_1 p_1(y_1, t_1) dy_1, \tag{16}$$

$$\langle y(t_1) y(t_2) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 y_2 \, p_2(y_1, t_1; y_2, t_2) dy_1 dy_2. \tag{17}$$

Higher order correlation functions require the knowledge of higher order distribution functions. In the random walk we have just looked at p_1.

A *stationary* random process is one for which the p_n depend only on time differences, or

$$p_n(y_1, t_1 + \tau; y_2, t_2 + \tau; \ldots; y_n, t_n + \tau) = p_n(y_1, t_1; y_2, t_2 \ldots; y_n, t_n). \tag{18}$$

I have chosen to formulate the random walk as starting a particle from a particular position at time $t = 0$, so that $x(t)$ is not stationary. Alternatively we could have considered a stationary process (e.g. the field of vision of a microscope with many Brownian particles) and then calculated the *conditional probability* $P_2(x_1, t_1 | x_2, t_2)$ which is the probability of the particle being at x_2 at time t_2 given that it was at x_1 at time t_1. Then $P_2(0, 0 | x, t)$ takes the diffusive form that we have calculated and the p_n all just depend on the time difference ($p_1(x, t)$ is just constant, for example).

Means and correlation functions are defined with respect to the ensemble average. For a stationary random process we usually assume *ergodicity*, and replace the ensemble average by a time average, e.g.

$$\langle y \rangle = \overline{y(t)} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} y(t) dt. \tag{19}$$

The probability distribution for the random walk is a Gaussian function. A *Gaussian process* in general is one in which *all* the probability distributions are Gaussian

$$p_n(y_1, t_1; y_2, t_2 \ldots; y_n, t_n) = A \exp \left[- \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk} (y_j - \langle y \rangle)(y_k - \langle y \rangle) \right] \tag{20}$$

where $\langle y \rangle$ is the mean of y, α_{jk} is a positive definite matrix and A is a normalization constant. For a stationary process $\langle y \rangle$ is time independent and α and A depend only on time differences.

Gaussian processes are important in physics because of the *central limit theorem*: if

$$Y = \frac{1}{N} \sum y_i \tag{21}$$

with y_i a random process or variable with arbitrary distribution but with finite mean $\langle y \rangle$ and variance σ_y^2 then for N large Y is a Gaussian process or variable

$$p(Y) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} \exp\left[-\frac{(Y - \langle Y \rangle)^2}{2\sigma_Y^2} \right] \tag{22}$$

with $\langle Y \rangle = \langle y \rangle$ and $\sigma_Y = \sigma_y/\sqrt{N}$. The central limit theorem is why the Gaussian distribution of the random walk is independent of the details of the step (e.g. fixed length, or varying length) providing the mean is zero and the variance is finite.

Equation (6) for the time evolution of the probability distribution is actually the *Fokker-Planck* equation for this random process. We will return to this topic in more detail later.

Spectral description of a random process

For a conventional function $y(t)$ a convenient definition of the Fourier transform is

$$\tilde{y}(f) = \int_{-\infty}^{\infty} y(t) e^{i2\pi f t} dt, \tag{23a}$$

$$y(t) = \int_{-\infty}^{\infty} \tilde{y}(f) e^{-i2\pi f t} df. \tag{23b}$$

The correctness of the inverse is shown from the result

$$\int_{-\infty}^{\infty} e^{i2\pi x y} dx = \lim_{x\to\infty} \frac{\sin 2\pi x y}{\pi y} = \delta(y). \tag{24}$$

For a real function $y(t)$ we have $\tilde{y}^*(f) = \tilde{y}(-f)$.

For a stationary random process the integral defining $\tilde{y}(f)$ diverges, so we instead define the auxiliary process

$$y_T(t) = \begin{cases} y(t) & -T/2 < t < T/2 \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

and then use the finite $\tilde{y}_T(f)$.

Parseval's theorem tells us

$$\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} [y(t)]^2 = \lim_{T\to\infty} \frac{2}{T} \int_0^{\infty} |\tilde{y}_T(f)|^2 \, df. \tag{26}$$

Here and elsewhere we use $\tilde{y}^*(f) = \tilde{y}(-f)$ to restrict the frequency domain to positive values.

With these preliminaries in mind, we define the spectral density of the random process $y(t)$ as

$$G_y(f) = \lim_{T\to\infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} [y(t) - \bar{y}] e^{i2\pi f t} dt \right|^2. \tag{27}$$

where $\bar{y}$ is the time average over T. Why do we use this expression? Lets suppose that the mean has been subtracted off of y, so $\bar{y} = 0$. The quantity inside the $||$ is the Fourier transform of the

process $y_T(t)$. How does this grow with T? We can estimate this by supposing the interval T to be formed of N subintervals of length τ. The Fourier transform $\tilde{y}_T$ is then the sum of N transforms $\tilde{y}_\tau$ of processes $y_\tau(t)$defined over the interval τ. For a random process we would expect each of the $\tilde{y}_\tau$ to be of similar magnitude, but with arbitrary phase, since the latter depends sensitively on the phasing of the $e^{i2\pi ft}$ with respect to the time start of the signal. Adding N complex numbers with random phase gives a number of magnitude $\propto \sqrt{N}$ and random phase. Thus the transform of $y(t) - \bar{y}$ grows as $\sqrt{T}$ and the phase varies over all values as T changes. The spectral density $G_y(f)$ is constructed to be *independent of* T and to contain all the useful information. Parseval's theorem now gives us

$$\int_0^\infty G_y(f)df = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} [y(t) - \bar{y}]^2 dt = \sigma_y^2, \tag{28}$$

so that the frequency integral of the spectral density is the variance of the signal.

The spectral density is directly related to the Fourier transform of the correlation function $C_y(\tau)$. Let's set the mean $\bar{y}$ to zero for simplicity. Then, using assumption of ergodicity to replace the ensemble average by a time average, the correlation function is

$$C_y(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} dt\; y(t)y(t+\tau) \tag{29}$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_{-\infty}^{\infty} dt\; y_T(t)y_T(t+\tau) \tag{30}$$

where the small error in replacing $y(t+\tau)$ by $y_T(t+\tau)$ is unimportant in the limit. Now inserting the Fourier transforms and using $\tilde{y}^*(f) = \tilde{y}(-f)$

$$C_y(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \tilde{y}_T(f)\tilde{y}_T(f') e^{-i2\pi f'\tau} e^{i2\pi(f+f')t}. \tag{31}$$

The t integrations is $\delta(f+f')$, and using $\tilde{y}^*(f) = \tilde{y}(-f)$ gives

$$C_y(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-\infty}^{\infty} df\; |\tilde{y}_T(f)|^2\, e^{-i2\pi f\tau} \tag{32}$$

$$= \lim_{T\to\infty} \frac{2}{T} \int_{0}^{\infty} df\; |\tilde{y}_T(f)|^2 \cos 2\pi f\tau \tag{33}$$

$$= \int_0^\infty G_y(f)\cos(2\pi f\tau)df. \tag{34}$$

Thus we have the inverse pair

$$C_y(\tau) = \int_0^\infty G_y(f)\cos(2\pi f\tau)df \tag{35a}$$

$$G_y(f) = 4\int_0^\infty C_y(\tau)\cos(2\pi f\tau)d\tau \tag{35b}$$

(since C_y and G_f are both even functions, we have written the results as cosine transforms only involving the positive domain). These equations are known as the *Wiener-Khintchine* theorem.

A particularly simple spectral density is a flat one, independent of frequency. We describe such a random process as being *white*. The corresponding correlation function is a delta function, i.e. no correlations except for time differences tending to zerp. One strength parameter g is needed to specify the force

$$G_F(f) = g, \tag{36a}$$

$$C_F(\tau) = \frac{g}{2}\delta(\tau). \tag{36b}$$

The Einstein Relation

Einstein showed how to relate the diffusion constant, describing the random fluctuations of the Brownian particle, to its mobility μ, the systematic response to an externally applied force.

Under an applied force $-dV/dx$ the drift velocity of the particle is (the definition of the mobility)

$$u_d = -\mu \frac{dV}{dx}. \tag{37}$$

For a sphere of radius a in a liquid the viscosity η the mobility is given by the Stokes expression $\mu = (6\pi\eta a u_d)^{-1}$, and so μ is related to the *dissipation* in the fluid.

Consider now the thermodynamic equilibrium of a density $n(x)$ of independent Brownian particles in the potential $V(x)$. We can dynamically understand the equilibrium in terms of the cancelling of the particle currents due to diffusion and mobility

$$-D\frac{dn}{dx} + n\left(-\mu\frac{dV}{dx}\right) = 0. \tag{38}$$

Equilibrium thermodynamics on the other hand tells us $n(x) \propto \exp[-V(x)/kT]$. Substituting into Eq. (38) gives the Einstein identity

$$D = kT\mu. \tag{39}$$

Note the use of *equilibrium* constraints to relate *fluctuation* quantities (the diffusion constant which gives us $\langle x^2(t)\rangle$) and *dissipation* coefficients (μ or η). This is an example of a general approach known as *fluctuation dissipation* theory, that we will take up again later. The fact that the fluctuations and dissipation of a Brownian particle are related should not be unexpected: both are a reflection of the molecular buffeting, the dissipation given by the net force due to the systematic component of the collisions coming from the drift of the particle relative to the equilibrium molecular velocity distribution, and the fluctuations coming from the random component.

Fluctuation-Dissipation Theory

The relationship between the dissipation coefficient and the fluctuations is made more explicit by directly evaluating D in terms of the fluctuations producing the random walk

$$D = \lim_{t\to\infty} \frac{1}{2t}\left\langle [x(t) - x(0)]^2 \right\rangle. \tag{40}$$

Expressing the displacement as the integral of the stochastic velocity u

$$x(t) - x(0) = \int_0^t u(t_1) dt_1 \tag{41}$$

leads to

$$D = \lim_{t\to\infty} \frac{1}{2t} \int_0^t dt_1 \int_0^t dt_2 \, \langle u(t_1)u(t_2)\rangle \,, \tag{42}$$

which depends on the *velocity correlation function*. The integrand is symmetric in t_1, t_2 and we can replace the integral over the square by twice the integral over the triangle $0 < t_1 < t, t_1 < t_2 < t$, and then introducing the time difference $\tau = t_2 - t_1$

$$D = \lim_{t\to\infty} \frac{1}{t} \int_0^t dt_1 \int_0^{t-t_1} d\tau \, \langle u(t_1)u(t_1+\tau)\rangle \,. \tag{43}$$

Since the correlation function $\langle u(t_1)u(t_1+\tau)\rangle$ decays to zero in some finite relaxation time τ_r, as $t \to \infty$ the limit of the second integral can be replaced by infinity for almost all values of t_1 in the first integration. Further, $\langle u(t_1)u(t_1+\tau)\rangle = C_u(\tau)$ is independent of t_1 ($u(t)$ is a stationary random process if external conditions are fixed). Hence

$$D = \int_0^\infty d\tau \, \langle u(0)u(\tau)\rangle \tag{44}$$

and

$$\mu = \frac{1}{kT} \int_0^\infty d\tau \, \langle u(0)u(\tau)\rangle \tag{45}$$

directly relating a *dissipation coefficient* to a *correlation function*.