Chapter 2

Random Walks

A problem, which is closely related to Brownian motion and which we will examine in this chapter, is that of a *random walker*. This concept was introduced into science by Karl pearson in a letter to Nature in 1905:

> *A man starts from a point 0 and walks ℓ yards in a straight line; he then turns through any angle whatever and walks another ℓ yards in a straight line. He repeats this process n times. I require the probability that after these n stretches he is at a distance between r and $r + \delta r$ from his starting point 0.*

The solution to this problem was provided in the same volume of Nature by Lord Rayleigh (1842-1919), who told him that he had solved this problem 25 years earlier when studying the superposition of sound waves of equal frequency and amplitude but with random phases.

The random walker, however, is still with us today.

2.1 The Random Walk on a Line

Let us assume that a walker can sit at regularly spaced positions along a line that are a distance Δx apart (see fig. 2.1) so we can label the positions by the set of whole numbers m. Furthermore we require the walker to be at position 0 at time 0. After fixed time intervals Δt the walker either jumps to the right with probability p or to the left with probability $q = 1 - p$, so we work with discrete time points $N \geq 0$.

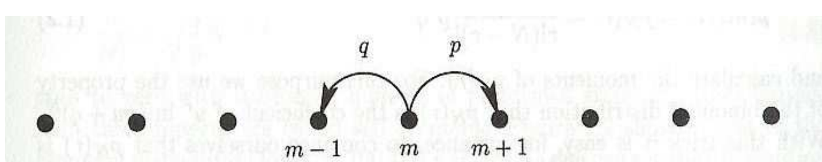

Figure 2.1: *A random walker on a 1-dimensional lattice of sites that are a fixed distance Δx apart. The walker jumps to the right with probability p and to the left with probability $q = 1 - p$.*

Our aim is to answer the following question: what is the probability $p(m, N)$ that the walker will be at position m after N steps?

For $m < N$ there are many ways to start at 0 and go through N jumps to nearest-neighbor sites and end up at m. But since all these possibilities are independent we have to add up their probabilities. For all these ways we know that the walker must have made $n_1 = m + n_2$ jumps to the right and n_2 jumps to the left, and since $n_1 + n_2 = N$, the walker must have made

$$\begin{aligned} n_1 &= \frac{1}{2}(N+m) \quad \text{jumps to the right and} \\ n_2 &= \frac{1}{2}(N-m) \quad \text{jumps to the left} \end{aligned}$$

The probability for a sequence of left and right jumps is the product of the probabilities of the individual jumps. Since the individual jumps are independent, all paths starting at 0 and ending at m have the same overall probability. This follows since the probability for making n_1 jumps to the right and n_2 jumps to the left must contain n_1 factors p and n_2 factors q and becomes

$$p^{n_1} q^{n_2} = p^{\frac{1}{2}(N+m)} q^{\frac{1}{2}(N-m)}$$

This probability must be multiplied by the total number of paths with n_1 steps to the right and n_2 steps to the left. This is given by the number of ways to put n_1 objects out of N in one box and $n_2 = N - n_1$ objects in another box. The first object can be choosen in N ways, the second in $N-1$ ways and so on. The total number of choosing n_1 object is therefore $N(N-1)(N-2)\cdots(N-n_1-1)$. But the n_1 objects are identical and can be arranged in any order. The total number of distinguishable ways to have n_1 steps to the right and n_2 to the left therefore becomes

$$\frac{N!}{n_1! n_2!} = \frac{N!}{n_1!(N-n_1)!}$$

The probability of being at position m after N jumps is therefore given as

$$p(m, N) = \frac{N!}{\left(\frac{N+m}{2}\right)! \left(\frac{N-m}{2}\right)!} p^{\frac{1}{2}(N+m)} q^{\frac{1}{2}(N-m)} \tag{2.1}$$

which is the so called binomial distribution.

If we know the probability distribution $p(m, N)$ we can calculate all the moments of m at any fixed time N. From the number of steps to the right $n = (N+m)/2$ we can write

$$p(m, N) = p_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n} = \begin{pmatrix} N \\ n \end{pmatrix} p^n q^{N-n} \tag{2.2}$$

and calculate the various moments of $p_N(n)$. Noting that

$$(pu+q)^N = \sum_{n=0}^{N} \begin{pmatrix} N \\ n \end{pmatrix} u^n p^n q^{N-n}$$

we see that $p_N(n)$ is the coefficient of u^n in this binomial expansion. Then

$$\sum_{n=0}^{N} p_N(n) = \left[(pu+q)^N\right]_{u=1} = 1 \tag{2.3}$$

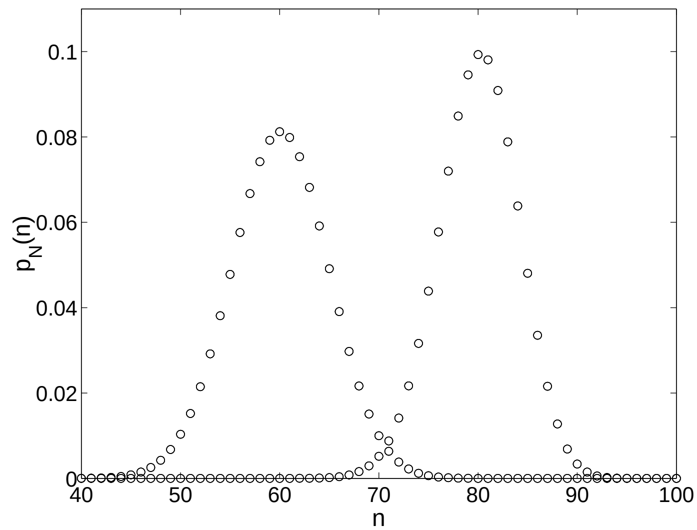

Figure 2.2: *Plot of the binomial distribution for a number of steps $N = 100$ and the probability of a jump to the right $p = 0.6$ and $p = 0.8$. We see that the maximum is located close to Np.*

which shows that $p_N(n)$ is properly normalized to one.

The first moment or *expectation value* of n is:

$$
\begin{aligned}
\mathrm{E}\left[n\right] = \langle n \rangle = \sum_{n=0}^{N} n p_N(n) &= \sum_{n=0}^{N} n \left[\binom{N}{n} u^n p^n q^{N-n} \right]_{u=1} \\
= \sum_{n=0}^{N} \left[\binom{N}{n} u \frac{\mathrm{d}}{\mathrm{d}u} \left(u^n p^n q^{N-n} \right) \right]_{u=1} &= \left[u \frac{\mathrm{d}}{\mathrm{d}u} \sum_{n=0}^{N} \binom{N}{n} u^n p^n q^{N-n} \right]_{u=1} \\
= \left[u \frac{\mathrm{d}}{\mathrm{d}u} (pu + q)^N \right]_{u=1}
\end{aligned}
$$

This leads to

$$\mathrm{E}\left[n\right] = \langle n \rangle = Np \tag{2.4}$$

In the same manner we can derive the second moment:

$$\mathrm{E}\left[n^2\right] = \langle n^2 \rangle = \left[\left(u \frac{\mathrm{d}}{\mathrm{d}u} \right)^2 (pu + q)^N \right]_{u=1} = Np + N(N-1)p^2 \tag{2.5}$$

The *variance* of the variable n is defined by

$$\mathrm{Var}\left[n\right] = \sigma^2 = \langle (n - \langle n \rangle)^2 \rangle = \langle n^2 \rangle - \langle n \rangle^2 \tag{2.6}$$

σ^2 is a measure of the width of the distribution and from (2.4) and (2.5) we find

$$\sigma^2 = Npq \tag{2.7}$$

The relative width of the distribution

$$\frac{\sigma}{\langle n \rangle} = \sqrt{\frac{q}{p}} \frac{1}{\sqrt{N}} \tag{2.8}$$

goes to zero with incrreasing number of steps N. Distributions with this property are called *self-averaging*. Figure 2.2 shows a plot of the binomial distribution for $N = 100$ and $p = 0.6$ and 0.8. As one can see the distribution has a bell-shaped form with a maximum occuring around the average value $\langle n \rangle = Np$.

The position of the walker after N steps is now given by

$$\langle m \rangle = 2\langle n \rangle - N = N(p - q) \tag{2.9}$$

and for the second moment

$$\langle m^2 \rangle = 4\langle n^2 \rangle - 4\langle n \rangle N + N^2 = 4\sigma^2 + \langle m \rangle^2 \tag{2.10}$$

The variance of the displacement m is therefore

$$\sigma_m^2 = 4\sigma^2 = 4Npq \tag{2.11}$$

For the special case of a symmetric walker with $p = q = 1/2$ this reduces to

$$\langle m \rangle = 0, \quad \text{and} \quad \langle m^2 \rangle = N$$

This behaviour, in which the square of the distance traveled is proportional to time is called *free diffusion*.

Introducing the real displacement $x = m\Delta x$ and the time $t = N\Delta t$, we see that (2.11) implies

$$\langle (x - \langle x \rangle)^2 \rangle = 4pq(\Delta x)^2 N = 4pq\frac{(\Delta x)^2}{\Delta t} N\Delta t = 2Dt \tag{2.12}$$

where the diffusion coefficient is defined as $D = 2pq(\Delta x)^2/\Delta t$. The averaged distance traveled by the walker is $\langle x \rangle = \langle m \rangle \Delta x = (p - q)\Delta x N = vt$, which defines the drift velocity $v = (p - q)\Delta x/\Delta t$. For a symmetric walk the average position is zero and this is also the case for v.

2.2 Gaussian distribution

We are interested in the behaviour of the binomial distribution for $Np \to \infty$ i.e. for long times. Assuming $N \gg 1$ we can use Stirling's formula to approximate the factorials in the binomial distribution

$$\ln N! = \left(N + \frac{1}{2}\right) \ln N - N + \frac{1}{2} \ln 2\pi + O\left(\frac{1}{N}\right) \tag{2.13}$$

From this formula we get

$$\begin{aligned} \ln p(m, N) &= \left(N + \frac{1}{2}\right) \ln N - \left(\frac{N+m}{2} + \frac{1}{2}\right) \ln \frac{N+m}{2} - \left(\frac{N-m}{2} + \frac{1}{2}\right) \ln \frac{N-m}{2} \\ &+ \frac{N+m}{2} \ln p + \frac{N-m}{2} \ln q - \frac{1}{2} \ln 2\pi \end{aligned}$$

For large values of N we expect $p(m,N)$ to be sharply peaked around the maximun value, which is close to the average value $\langle m\rangle = N(p-q)$ as seen in fig 2.2. We write

$$m = \langle m\rangle + \delta m = N(p-q) + \delta m$$

which leads to

$$\frac{N+m}{2} = Np + \frac{\delta m}{2}, \quad \frac{N-m}{2} = Nq - \frac{\delta m}{2}$$

With these relations we find

$$\begin{aligned}
\ln p(m,N) &= \left(N+\frac{1}{2}\right)\ln N - \frac{1}{2}\ln 2\pi - \left(Np + \frac{1+\delta m}{2}\right)\ln\left[Np\left(1+\frac{\delta m}{2Np}\right)\right] \\
&- \left(Nq + \frac{1-\delta m}{2}\right)\ln\left[Np\left(1-\frac{\delta m}{2Np}\right)\right] + \left(Np+\frac{\delta m}{2}\right)\ln p + \left(Nq - \frac{\delta m}{2}\right)\ln q \\
&= -\frac{1}{2}\ln 2\pi Npq - \left(Np + \frac{1+\delta m}{2}\right)\ln\left(1+\frac{\delta m}{2Np}\right) - \left(Nq + \frac{1-\delta m}{2}\right)\ln\left(1-\frac{\delta m}{2Np}\right)
\end{aligned}$$

We can expand the logarithm

$$\ln(1\pm x) = \pm x - \frac{1}{2}x^2 + O(x^3)$$

to find

$$\begin{aligned}
\ln p(m,N) &= -\frac{1}{2}\ln 2\pi Npq + \frac{(p-q)}{4Npq}\delta m - \frac{1}{2}\frac{1}{4Npq}(\delta m)^2 \\
&= -\frac{1}{2}\ln 2\pi Npq + \frac{(p-q)}{\sigma_m^2}\delta m - \frac{1}{2\sigma_m^2}(\delta m)^2 \qquad (2.14)
\end{aligned}$$

where we used that $\sigma_m^2 = 4Npq$ is the variance of m. We notice that for $(\delta m)^2 \propto \sigma_m^2 \propto N$ the second term in (2.14) is of order $O((Np)^{-1/2})$ which therefore can be neglected for $Np \to \infty$. We therefore find

$$p(m,N) \to \frac{2}{\sqrt{2\pi\sigma_m^2}}\exp\left(-\frac{1}{2}\frac{(\delta m)^2}{\sigma_m^2}\right) \qquad (2.15)$$

which is a Gaussian distribution.

We can perform a continous limit and introduce

$$x = m\Delta x, \quad t = N\Delta t, \quad D = 2pq\frac{(\Delta x)^2}{\Delta t}, \quad \langle x\rangle = (p-q)N = vt$$

We then find

$$p(m\Delta x, N\Delta t) = \frac{2}{\sqrt{2\pi\sigma_m^2}}\exp\left(-\frac{1}{2}\frac{(x-\langle x\rangle)^2}{2Dt}\right)$$

for the probability to find the random walk in an interval of width $2\Delta x$ around the position x at time t. In the limit

$$\Delta x \to 0, \quad \Delta t \to 0, \quad D = 2pq\frac{(\Delta x)^2}{\Delta t} = \text{const.}, \quad v = (p-q)\frac{\Delta x}{\Delta t} = \text{const.} \qquad (2.16)$$

we find the probability density to find the walker in an interval $(x, x + \mathrm{d}x)$ as

$$p(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x - vt)^2}{4Dt}\right) \tag{2.17}$$

with the starting condition

$$p(x,t) = \delta(x)$$

Any asymmetry in the transition rates ($p \neq q$) produces a net drift velocity of the walker.

2.3 Rate Equation

The combinatorial method used above to derive the expression for the probability distribution $p(m, N)$ or the corresponding probability density $p(x, t)$ is not easily generelized to higher dimensions or to more complex walks. Therefore we now want to derive an evolution equation for the probability density starting from the discrete walker. This *rate equation* can straightforwardly be generalized to other situations.

Let's assume that the walker arrives at position m at time $N + 1$. The walker has to jump to position m either from the position to the left or to the right of m with respective probabilities q and p. This gives

$$p(m, N+1) = pp(m-1, N) + qp(m+1, N) \tag{2.18}$$

This is an example of a *master equation* fo a stochastic process. We can subtract $p(m, N)$ from both sides and divide by Δt to obtain

$$\frac{p(m,N+1) - p(m,N)}{\Delta t} = -p\frac{\Delta x}{\Delta t}\frac{p(m,N) - p(m-1,N)}{\Delta x} + q\frac{\Delta x}{\Delta t}\frac{p(m+1,N) - p(m,N)}{\Delta x}$$

From (2.16) we can solve for $\Delta x/\Delta t$ as

$$\frac{\Delta x}{\Delta t} = v + 2q\frac{\Delta x}{\Delta t} = 2p\frac{\Delta x}{\Delta t} - v$$

Inserting these in the first two terms on the right hand side respectively gives

$$\begin{aligned}\frac{p(m,N+1) - p(m,N)}{\Delta t} &= -vp\frac{p(m,N) - p(m-1,N)}{\Delta x} - vq\frac{p(m+1,N) - p(m,N)}{\Delta x} \\ &+ 2pq\frac{(\Delta x)^2}{\Delta t}\frac{p(m+1,N) - 2p(m,N) + p(m-1,N)}{(\Delta x)^2}\end{aligned}$$

We now perform the continuum limit of this equation keeping v and $D = 2pq(\Delta x)^2/\Delta t$ constant and arrive at

$$\frac{\partial}{\partial t}p(x,t) = -v\frac{\partial}{\partial x}p(x,t) + D\frac{\partial^2}{\partial x^2}p(x,t) \tag{2.19}$$

This is the diffusion equation with a drift term. It is straightforward to show that the expression in (2.17) satisfies this equation.

The probability $p(x,t)\mathrm{d}x$ gives the probability to find the walker or particle in the interval $\mathrm{d}x$ around the possition x at time t, given that it starts at $x = 0$ at $t = 0$.

Assuming now that we have a collection of particles like macromolecules swimming in a solution of small molecules. It is clear that the density of molecules at x is given by $n(x,t) = N_0 p(x,t)$, with N_0 the total number of particles. The density fluctuations therefore also satisfies the equation (2.19)

$$\frac{\partial}{\partial t} n(x,t) = -v \frac{\partial}{\partial x} n(x,t) + D \frac{\partial^2}{\partial x^2} n(x,t)$$

This equation is easily generalized to higher dimensions and is just the continuity equation

$$\frac{\partial}{\partial t} n(\boldsymbol{r},t) + \nabla \cdot \boldsymbol{j}(\boldsymbol{r},t) = 0 \tag{2.20}$$

This equation holds if the number of particles in the system is constant in time. This is true if no chemical reactions can take place. Here $\boldsymbol{j}(\boldsymbol{r},t)$ is the flux or current of particles, which is the number of particles passing through a unit area perpendicular to $\boldsymbol{j}$ per unit time. Particles will tend to move from regions of high concentration to regions of low concentration, and Fick's law gives the flux

$$\boldsymbol{j}(\boldsymbol{r},t) = -D\nabla n(\boldsymbol{r},t)$$

with D the diffusion coefficient. If the particles also have a constant streaming velocity $\boldsymbol{v}$ there will be an additional flux of particles $\boldsymbol{v}n(\boldsymbol{r},t)$, which would exist even in the absence of particle diffusion, so that

$$\boldsymbol{j}(\boldsymbol{r},t) = \boldsymbol{v}n(\boldsymbol{r},t) - D\nabla n(\boldsymbol{r},t) \tag{2.21}$$

Substitutiing the flux $\boldsymbol{j}$ from (2.21) into (2.20) yields the equation

$$\frac{\partial}{\partial t} n(\boldsymbol{r},t) = -\boldsymbol{v} \cdot \nabla n(\boldsymbol{r},t) + D\nabla^2 n(\boldsymbol{r},t) \tag{2.22}$$

which generalizes ((2.19) to higher dimensions.

If we subtract the uniform density n from $n(\boldsymbol{r},t)$ to get the fluctuations $\delta n(\boldsymbol{r},t)$, and multiply with $\delta n(\mathbf{0}, \mathbf{0}$ we find after averaging the equation for the density-density correlation function $C(\boldsymbol{r},t) = \langle \delta n(\boldsymbol{r},t)\delta n(\mathbf{0},\mathbf{0})\rangle$

$$\frac{\partial}{\partial t} C(\boldsymbol{r},t) + \boldsymbol{v} \cdot \nabla C(\boldsymbol{r},t) - D\nabla^2 C(\boldsymbol{r},t) = 0 \tag{2.23}$$

with the initial condition $C(\boldsymbol{r},0) = \delta(\boldsymbol{r})$.

2.4 Poisson distribution

The Gaussian distribution is not the only limiting distribution which can be derived from the binomial distribution. The Gaussian distribution follow in the limit $N \to \infty, Np \to \infty$. There is another possibility:

$$N \to \infty, \quad p \to 0, \quad Np = \lambda = \text{const.}$$

The probability p for a step is therefore assumed to be very small. Now

$$\begin{aligned} p_N(n) &= \frac{N!}{n!(N-n)!}p^n q^{N-n} = \frac{N(N-1)\dots(N-n+1)}{n!}p^n(1-p)^{(N-n)} \\ &= \left(1-\frac{1}{N}\right)\dots\left(1-\frac{n-1}{N}\right)\frac{(Np)^n}{n!}\left(1-\frac{Np}{N}\right)^{N-n} \end{aligned}$$

Now $p_N(n)$ is sharply peaked around the mean value i.e $n \approx Np$ and so $n \ll N$. Also $\lim_{N\to\infty}(1-\lambda/N)^N = \exp(-\lambda)$. Therefore the limit gives the Poisson distribution in the limit $N \to \infty, Np = \lambda$

$$p(n) = \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda} \tag{2.24}$$

This distribution is completely characterized by the parameter λ. For the first moment we get

$$\langle n\rangle = \sum_{n=0}^{\infty} np(n) = \mathrm{e}^{-\lambda}\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!} = \lambda$$

In general one can derive the recursion relation

$$\langle n^k\rangle = \sum_{n=0}^{\infty} n^k\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda} = \sum_{n=0}^{\infty} n^{k-1}\left[\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\lambda^n}{n!}\right]\mathrm{e}^{-\lambda} = \lambda\left(\frac{\mathrm{d}}{\mathrm{d}\lambda}+1\right)\langle n^{k-1}\rangle$$

The second moment is then

$$\langle n^2\rangle = \lambda + \lambda^2, \quad \to \quad \sigma^2 = \lambda$$

In fig 2.3a and 2.3b we compare the binomial distribution with the Gaussian and Poisson distributions for $N = 1000$. In fig. 2.3a we use $p = 0.8$ giving $Np = 800$. The binomial and the Gaussian distributions are practically indentical, while the Poisson distribution is much broader and cannot describe the binomial distribution. In fig 2.3b $p = 0.01$ and $Np = \lambda = 10$. The situation is reversed, the Poisson distribution is now the better approximation to the binomial distribution.

2.5 Examples

In Light or neutron scattering experiments one measure $C(\boldsymbol{r}, t)$ or rather the spatial and temporal Fourier transform

$$F(\boldsymbol{q}, t) = \int \mathrm{d}\boldsymbol{r}\,\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}C(\boldsymbol{r}, t) \tag{2.25}$$

and

$$S(\boldsymbol{q}, \omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}t\,\mathrm{e}^{-\mathrm{i}\omega t}F(\boldsymbol{q}, t) = \frac{1}{\pi}\mathrm{Re}\int_0^{\infty}\mathrm{d}t\,\mathrm{e}^{-\mathrm{i}\omega t}F(\boldsymbol{q}, t) \tag{2.26}$$

The spatial Fourier transform of (2.23) is

$$\frac{\partial}{\partial t}F(\boldsymbol{q}, t) = \mathrm{i}\boldsymbol{q}\cdot\boldsymbol{v}F(\boldsymbol{q}, t) - q^2DF(\boldsymbol{q}, t) \tag{2.27}$$

which with the initial condition $F(\boldsymbol{q}, t = 0) = 1$ have the solution

$$F(\boldsymbol{q}, t) = \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{v}t - q^2Dt} \tag{2.28}$$

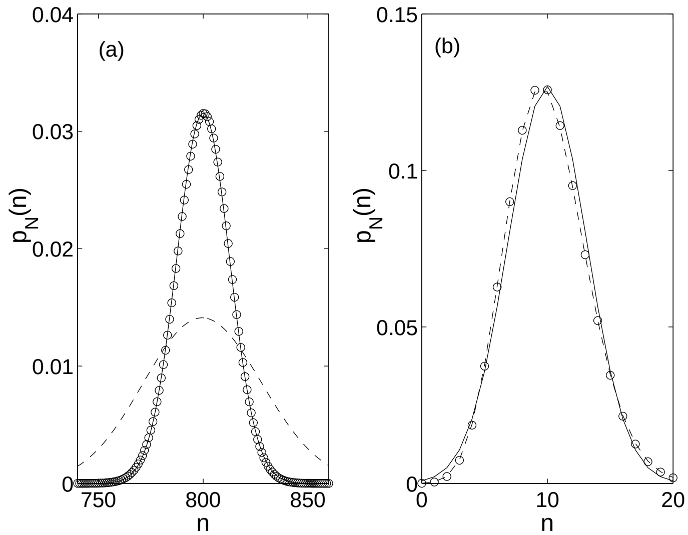

Figure 2.3: *(a) Plot of the binomial distribution for $N = 1000$ and $p = 0.8$ (*open circles*). This is compared with the Gaussian approximation with the same mean and width (*solid curve*) and the Poisson distribution with the same mean (*dashed curve*). (b) Same as in (a) but with $p = 0.01$.*

Diffusion

For the case that there is no drift velocity, $\boldsymbol{v} = 0$, then (2.28) reduces to the diffusion result

$$F(\boldsymbol{q}, t) = \mathrm{e}^{-q^2 D t} \tag{2.29}$$

The corresponding spectral density becomes

$$S(\boldsymbol{q}, \omega) = \frac{1}{\pi}\mathrm{Re}\int_0^\infty \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t - q^2 D t} = \frac{1}{\pi}\mathrm{Re}\frac{1}{\mathrm{i}\omega + q^2 D} = \frac{1}{\pi}\frac{q^2 D}{\omega^2 + (q^2 D)^2} \tag{2.30}$$

which is a so called Lorentzian lineshape with a maximun at $\omega = 0$ and width $q^2 D$.

From these formulas and with light scattering experiments there have been many measurements of the diffusion constant of macromolecules. From this one can also obtain information about the radius of the diffusing particle. The Stokes-Einstein relation gives

$$D = \frac{k_{\mathrm{B}} T}{6\pi\eta R} \tag{2.31}$$

where η is the viscosity of the solvent and R is the radius of the macromolecules.

An example of such measurements are shown in fig. 2.4, which show $F(\boldsymbol{q}, t)$ for highly monodisperse sample of single-stranded circular DNA from the *fd bacteriophage*. These data can be fitted to an exponential function.

Alternatively one can measure the spectrum $S(\boldsymbol{q}, \omega)$ and determine the diffusion constant from the half-width of the Lorentzian in (2.30).

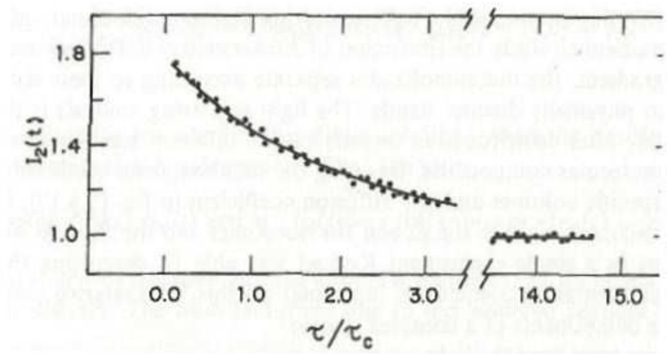

Figure 2.4: *Correlation function obtained with light scattering at a scattering angle of 60° for a solution containing 0.17 mg/cm³ of fd DNA in SCC (0.15 m NaCl, 0.015 m Nacitrate, pH=8) as a function of τ/τ_c where $\tau_c = (q^2D)^{-1}$ is the correlation time.* (From Newman, J., Swinney, H. L., Berkowitz, S. A. and Day, L. A. *Biochem.* **13**, 4832 (1974).)

Motile microorganisms

In the last decades there have been a lot of efforts to try to understand the motion of motile microorganisms. Several studies have been made of the motion of *E. coli* which swim by beating flagella.

Motile microorganisms are far more complex than molecules. They move in a very complicated way. It seems to be true that once a microorganism starts moving in a given direction it persistes with constant velocity in that direction for a distance long compared with typical values of q^{-1}. That is to say the “meah free path” of a microorganism is in general long compared with q^{-1}. In a dilute solution one can therefore ignore “collisions”, i.e changes in swimming direction.

Neglecting the collision term proportional to D in (2.28) we find

$$F(\boldsymbol{q},t) = \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{v}t} \tag{2.32}$$

We expect the swimming velocity $\boldsymbol{v}$ to have some distribution $P(\boldsymbol{v})$ and we should therefore average (2.32) over all velocites. In general $P(\boldsymbol{v})$ is not the equilibrium Maxwell-Boltzmann distribution.

The measured correlation function is then

$$F(\boldsymbol{q},t) = \int \mathrm{d}\boldsymbol{v} P(\boldsymbol{v})\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{v}t} \tag{2.33}$$

We expect that it is equally likely for a bacterium to move off in one direction as another. Therefore $P(\boldsymbol{v})$ should not depend on the direction of $\boldsymbol{v}$ but only on its magnitude $|\boldsymbol{v}|$. The angular integrations in (2.33) can thus be carried through yielding

$$\begin{aligned} F(\boldsymbol{q},t) &= \int_0^\infty \int_0^\pi \int_0^{2\pi} P(v)\mathrm{e}^{\mathrm{i}qv\cos\theta t} v^2 \sin\theta \mathrm{d}v\mathrm{d}\theta\mathrm{d}\phi \\ &= 2\pi \int_0^\infty \int_{-1}^1 P(v)\mathrm{e}^{\mathrm{i}qvxt} v^2 \mathrm{d}v\mathrm{d}x = 4\pi \int_0^\infty v^2 P(v)\frac{\sin qvt}{qvt}\mathrm{d}v \end{aligned} \tag{2.34}$$

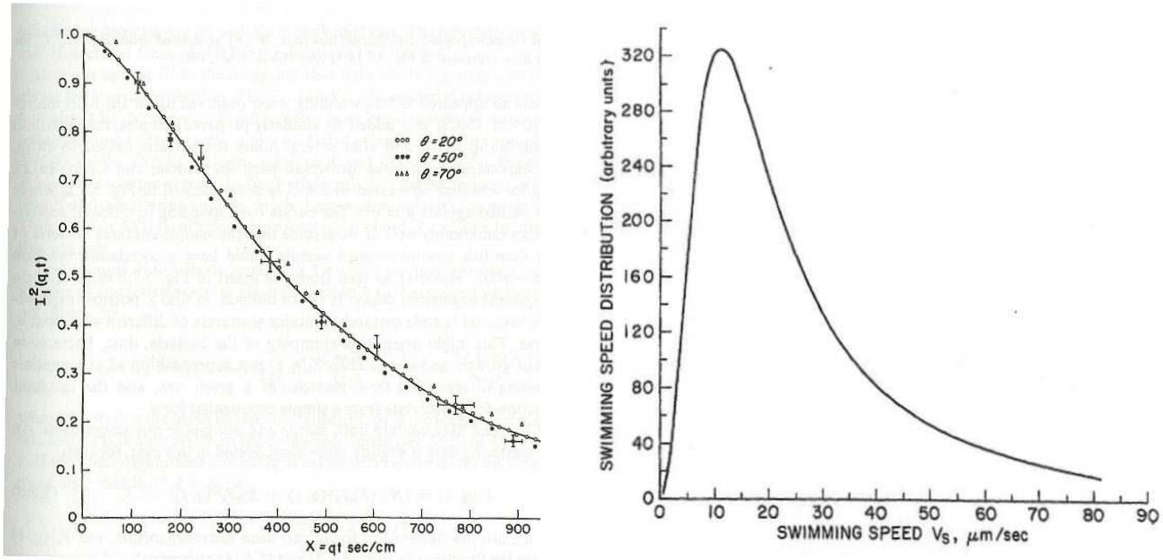

Figure 2.5: *(a) The correlation function obtained from light scattering for motile E. coli at $T = 25$ C taken at different scattering angles, and plotted versus $x = qt$. Concentration of bacteria $\approx 10/cm^3$. Number of bacteria in the scattering volume ≈ 10* (From Nossal, R., Chen, S. H., and Lai, C. C. *Opt Comm.* **4**, 35 (1971).)

The quantity

$$W(v)\mathrm{d}v = 4\pi v^2 P(v)\mathrm{d}v$$

is simply the probability that an organism will be found with a speed between v and $v+\mathrm{d}v$. It should be noted that $F(\boldsymbol{q}, t)$ depends on q and t only through the combination $x = qt$, i.e.

$$F(\boldsymbol{q}, t) = F(x) = \int_0^\infty W(x) \frac{\sin xv}{xv} \mathrm{d}v \tag{2.35}$$

By taking a Fourier sine transform of this function we get the swimming speed distribution

$$W(v) = \frac{2v}{\pi} \int_0^\infty xF(x)\mathrm{d}x \tag{2.36}$$

Thus measuring $F(\boldsymbol{q}, t)$ we can determine the speed distribution function of motile microorganisms.

An application is illustrated in fig. 2.5a and 2.5b, which shows $F(x)$ for three different wavevectors q, and the extracted speed distribution respectively.

Molecules in uniform motion

For particles which both flow and spread out via diffusion we have the full equation in (2.28) with the solution

$$F(\boldsymbol{q}, t) = \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{v}t - q^2 Dt}$$

The corresponding spectrum is

$$S(\boldsymbol{q}, \omega) = \frac{1}{2\pi} \left[\frac{q^2 D}{(\omega - \boldsymbol{q}\cdot\boldsymbol{v})^2 + (q^2 D)^2} + \frac{q^2 D}{(\omega + \boldsymbol{q}\cdot\boldsymbol{v})^2 + (q^2 D)^2} \right] \tag{2.37}$$

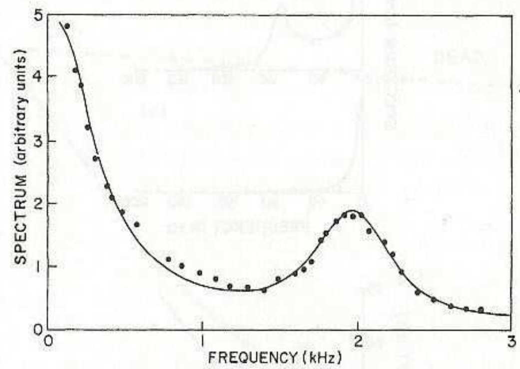

Figure 2.6: *Power spectrum of laser light scattered from silver chloride colloids in water. The circles are for the actual experimental data. The solid line is the sum of two Lorentzians, one centered at zero frequency, the second one at 1.97 kHz. The half-width of the first Lorentzian is 0.35 kHz and that of the second, 0.29 kHz.* (From Ben-Yosef, N., Zeigenbaum, S., and Weitz, A., *Appl. Phys. Lett.* **21**, 436 (1972))

There is now a shifted peak at

$$\omega(\boldsymbol{q}) = \pm \boldsymbol{q} \cdot \boldsymbol{v} \tag{2.38}$$

Measurements of particles in a flow have been performed for various systems. In fig. 2.6 we show experimental data for colloidal AgCl solutions. We see a shift at a frequency expected from (2.38).

The streaming motion of protoplasm in the alga Nitellar flexes are shown in fig. 2.7a. The Doppler shift in (2.38) could be studied as a function of wavevector q, and in fig. 2.7b is demonstrated that the peak frequency follow (2.38). Other examples are studies of blood flow. Here it is important to recognize that the blood-flow velocity is not uniform across a vein but varies from zero near the vein wall to a maximum in the center. Thus the spectrum observed should be an average over the distribution of velocities:

$$S(\boldsymbol{q}, \omega) = \frac{1}{2\pi} \int \mathrm{d}\boldsymbol{v} P(\boldsymbol{v}) \left[\frac{q^2 D}{(\omega - \boldsymbol{q} \cdot \boldsymbol{v})^2 + (q^2 D)^2} + \frac{q^2 D}{(\omega + \boldsymbol{q} \cdot \boldsymbol{v})^2 + (q^2 D)^2} \right] \tag{2.39}$$

where $\Gamma = q^2 D$. This average can give a spectrum which does not exhibit shifted peaks.

2.6 General random walks

Step-length distribution

Let's denote the probability to find a random walker at position m at step j with $p_j(m)$. The rate equation in (2.18) can then be written

$$p_{j+1}(m) = \sum_k f(m-k) p_j(k) \tag{2.40}$$

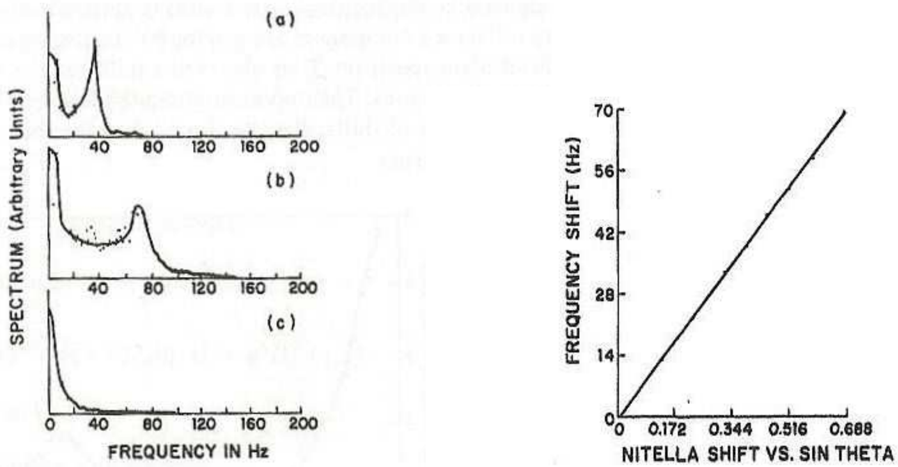

Figure 2.7: *Spectra of light scattered from the protoplasm of Nitella. The horizontal axis is frequency in Hz and the vertical is relative intensity. Spectra (a) and (b) was taken at scattering angles of* 19.5 *deg and* 36.1 *deg respectively. Spoectra (c) was taken immediately after addition of parachloromercuribenzoate, a streaming inhibitor. The right figure is a plot of the magnitude of the Doppler shift as a function of the sine of the scattering angle.* (From Mustachich, R. V. and Ware, B. R., *Phys. Rev. Lett.* **33**, 617 (1974))

where $f(k)$ is the probability density for the displacement of the walker at each step and for the simple walker in (2.18) is given by

$$f(k) = p\delta_{k,1} + q\delta_{k,-1}$$

i.e. the walker can only go one step to the right or left. Clearly $f(k)$ is normalized $\sum_k f(k) = 1$.

We can now generalize this and let $f(k)$ be an arbitrary probability distribution. Moreover in the continous limit $x = m\Delta x$ but still discrete time steps, (2.40) becomes

$$p_{j+1}(x) = \int \mathrm{d}x' f(x - x') p_j(x')$$

or in arbitrary dimension

$$p_{j+1}(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}' f(\boldsymbol{r} - \boldsymbol{r}') p_j(\boldsymbol{r}') \tag{2.41}$$

For a walker initially at the origin $p_0(\boldsymbol{r}) = \delta(\boldsymbol{r})$. We can solve (2.41) by taking the Fourier transform

$$p_j(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} p_j(\boldsymbol{r}), \quad f(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} f(\boldsymbol{r}) \tag{2.42}$$

Then

$$p_j(\boldsymbol{q}) = f(\boldsymbol{q}) p_{j-1}(\boldsymbol{q}) = [f(\boldsymbol{q})]^j \tag{2.43}$$

By Fourier inversion of (2.43) we find

$$p_j(\boldsymbol{r}) = \int \frac{\mathrm{d}\boldsymbol{q}}{(2\pi)^d} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} \, [f(\boldsymbol{q})]^j \tag{2.44}$$

Many important properties of random variables are derived from the generating function

$$G(\boldsymbol{r}, z) = \sum_{j=0}^{\infty} p_j(\boldsymbol{r}) z^n = \int \frac{\mathrm{d}\boldsymbol{q}}{(2\pi)^d} \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} \frac{1}{1 - z f(\boldsymbol{q})} \tag{2.45}$$

Then $p_j(\boldsymbol{r})$ is the coefficient of z^n in (2.45). From (2.42) we have for small $\boldsymbol{q}$-values $\boldsymbol{q} \to 0$

$$f(\boldsymbol{q}) = \int \mathrm{d}\boldsymbol{q} \left[1 + \mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r} - \frac{1}{2}(\boldsymbol{q}\cdot\boldsymbol{r})^2 + \dots\right] f(\boldsymbol{r}) = 1 + \mathrm{i}\boldsymbol{q}\cdot\langle\boldsymbol{r}\rangle - \frac{1}{2}\sum_{\alpha}\sum_{\beta} q^{\alpha} q^{\beta} \langle r^{\alpha} r^{\beta}\rangle + \dots \tag{2.46}$$

For an isotropic walker $\langle r^{\alpha} r^{\beta}\rangle = \delta_{\alpha\beta}\langle r^2\rangle/2d$ we find

$$f(\boldsymbol{q}) = 1 + \mathrm{i}\boldsymbol{q}\cdot\langle\boldsymbol{r}\rangle - \frac{1}{2d} q^2 \langle r^2\rangle \approx \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\langle\boldsymbol{r}\rangle - q^2\sigma^2/2d}$$

This gives for the probability density as a function of time $t = j\Delta t$

$$p_j(\boldsymbol{q}) = \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{v}t - q^2 D t} \tag{2.47}$$

where $\boldsymbol{v} = \langle\boldsymbol{r}\rangle/\Delta t$ is the mean drift velocity and $D = \sigma^2/(2d\Delta t)$ the diffusion coefficient.

Waiting time distribution

The discussion above was concerned with random walks generated by steps taken at regular intervals $\Delta t, 2\Delta t, 3\Delta t, \dots$. We can generalize the random walks further, by introducing a probability density function $\psi(t)$ for pausing times between successive steps in the walk. The Laplace transform of $\psi(t)$ is

$$\hat{\psi}(z) = \int_0^{\infty} \mathrm{d}t \, \mathrm{e}^{-zt} \psi(t)$$

Let $\psi_j(t)$ be the probability density function for the time execution of the j th step, i.e. $\psi_j(t)\mathrm{d}t$ is the probability that step number j takes place in the interval $(t, t + \mathrm{d}t)$. The probability density for the first step is $\psi_1(t) = \psi(t)$ and $\psi_j(t)$ satisfies the recurrence relation

$$\psi_j(t) = \int_0^t \mathrm{d}\tau \psi(t - \tau)\psi_{j-1}(\tau)$$

This follows since if step $j - 1$ occurs at time τ the probability that the next step occurs at time $(t - \tau)$ is $\psi(t - \tau)$. Integrating over all intervening times τ gives the result. By the convolution theorem of Laplace transforms we find

$$\hat{\psi}_j(z) = \hat{\psi}(z)\hat{\psi}_{j-1}(z) = \left[\hat{\psi}(z)\right]^j$$

We want to calculate $p(\boldsymbol{r}, t)$ which is the probability density that a walker originally at the origin is at $\boldsymbol{r}$ at time t. To calculate $p(\boldsymbol{r}, t)$ we first consider the function $Q(\boldsymbol{r}, t)$ that is defined as the probability that a walker is at point $\boldsymbol{r}$ at time τ immediately after a step has been taken. Having arrived at $\boldsymbol{r}$ at time τ the walker can stay at $\boldsymbol{r}$ for a random interval $t-\tau$, and the probability density that the walker remains fixed during an interval $(0,t)$ is

$$\phi(t) = 1 - \int_0^t \mathrm{d}t'\psi(t') = \int_t^\infty \mathrm{d}t'\psi(t')$$

The relation between $p(\boldsymbol{r},t)$ and $Q(\boldsymbol{r},t)$ is

$$p(\boldsymbol{r},t) = \int_0^t \mathrm{d}\tau\phi(t-\tau)Q(\boldsymbol{r},\tau) \tag{2.48}$$

i.e. if a walker arrives at $\boldsymbol{r}$ at time τ and remains there for a time $(t-\tau)$, $p(\boldsymbol{r},t)$ is the average over all arrival times with $0<\tau<t$.

The quantity $Q(\boldsymbol{r},t)\mathrm{d}t$ is the probability that a walker may arrive at $\boldsymbol{r}$ during the time interval $(t,t+\mathrm{d}t)$. This may happen after the first step, after the second step etc. The probability to arrive at $\boldsymbol{r}$ at step j is $p_j(\boldsymbol{r})$ in (2.44) and the probability that step j occurs at time t is $\psi_j(t)$. Therefore provided that displacements and timesteps are independent variables, we find the probability that after step j the walker is at $\boldsymbol{r}$ at time t is $\psi_j(t)p_j(\boldsymbol{r})$. Summing over all steps j gives

$$Q(\boldsymbol{r},t) = \sum_{j=0}^\infty \psi_j(t)p_j(\boldsymbol{r})$$

A Fourier and Laplace transform then gives

$$Q(\boldsymbol{q},z) = \sum_{j=0}^\infty \hat{\psi}_j(z)p_j(\boldsymbol{q}) = \sum_{j=0}^\infty \left[\hat{\psi}(z)f(\boldsymbol{q})\right]^j = \frac{1}{1-\hat{\psi}(z)f(\boldsymbol{q})}$$

and then from (2.48)

$$p(\boldsymbol{q},z) = \hat{\phi}(z)Q(\boldsymbol{q},z) = \frac{1-\hat{\psi}(z)}{z}\frac{1}{1-\hat{\psi}(z)f(\boldsymbol{q})} = \frac{1}{z+(1-f(\boldsymbol{q}))m(z)} \tag{2.49}$$

with $m(z) = z\hat{\psi}(z)/(1-\hat{\psi}(z))$.

For the waiting time distribution we have

$$\hat{\psi}(z) = \int_0^\infty \mathrm{d}t\mathrm{e}^{-zt}\psi(t) = 1 - z\langle t\rangle + \ldots$$

provided $\langle t\rangle < \infty$. In the limit $z\to 0$ and $\boldsymbol{q}\to 0$ we then find

$$p(\boldsymbol{q},z) = \frac{\langle t\rangle}{1-(1-z\langle t\rangle)(1+\mathrm{i}\boldsymbol{q}\cdot\langle\boldsymbol{r}\rangle - q^2\langle r^2\rangle/2d)} = \frac{1}{z-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{v}+q^2D} \tag{2.50}$$

which is again the result (2.28), i.e. a simple diffusion with a drift term.

By taking more general distributions for $\psi(t)$ and $f(\boldsymbol{r})$ we can model more complex situations.