Chapter 5

Diffusion and Hopping

In this chapter, we describe the motion of a particle hopping at random on a lattice. We shall use this very simple situation to introduce the description of dynamical motions in mechanics. Our analysis will show that the typical motion of the particle is one in which it only slowly moves outward from its starting point. It goes forward and back and retraces its path many times so that, in the end, its net motion is quite slow indeed. This slow type of motion is called *diffusion*, and this chapter is all about diffusive motion. Thus random hopping and diffusion provide alternative ways of viewing the selfsame situation.

One way we shall analyze this motion is to look at the probability that the diffusing particle might, at a given time, be found at a particular point in space. In doing this analysis we shall follow a general prescription which has proven to be a very powerful method in the analysis of dynamical situations: First we define the conservation laws obeyed by the system, then identify a likely behavior for the current of the conserved quantity, and finally put that current into the conservation law to find a hydrodynamic equation. In the case considered here, the conserved quantity is the total probability of finding the particle someplace, and the resulting hydrodynamic equation is the partial differential equation called the *diffusion equation.*

5.1 Random Walk on a Lattice

Newton emphasized the importance of momentum conservation in the description of isolated systems. The conservation of momentum is, as we shall see, a very important ingredient in determining the motion of a Brownian particle. However, in this chapter, we shall not include the effects of momentum conservation. (Those effects will come in the next chapter.) Instead, we consider the somewhat more elementary problem posed by the motion of impurities on a lattice. Imagine a regular hypercubic lattice (as in Fig. 5.1) and a particle which is allowed to sit on any one of the lattice sites

$$\mathbf{r} = a(n_1, n_2, \ldots, n_\mu, \ldots, n_d)\,, \tag{5.1}$$

where the n's are integers, d is the dimensionality of the system, and a is the lattice constant, i.e. the distance between nearest neighbor lattice sites. An impurity particle in a real solid may get trapped for a while at some lattice site and then come into equilibrium with its environment. Then, after it has been sitting on the site for a time, it may hop to a neighboring site, where it sits for a while and equilibrates. As it equilibrates, the momentum which may have been gained to help it hop is lost so that it has no memory of its past. When it hops again, it jumps to a neighboring site picked at random from the nearest neighbors. This process is repeated again and again and the particle moves through the lattice.

Notice that we have not defined the mechanism which causes its hopping. The most likely process is one in which, while the particle is sitting on the site, a bunch of thermally excited sound waves arrive together at the site and thus accidentally bring the particle enough energy to hop to its neighbor. Alternatively the process may be quantum in nature and involve, in part, quantum tunneling. What causes the hopping will be not too relevant to our analysis. What we need instead is a good phenomenological description of the hopping process.

To analyze this situation in its simplest form, we introduce a simplified model in which the particle is assumed to hop to the next site after the time interval, τ. On each hop, the particle moves from a site to a randomly chosen nearest neighbor site. The change in position on the jth hop is written ξ_j, where ξ_j is a vector of magnitude a directed parallel to the possible vectors connecting nearest neighbors. (See Fig. 5.1.) There are $2d$ possible values of ξ_j, which are $\pm a\mathbf{e}_\mu$ with $\mathbf{e}_\mu$ being a unit vectors drawn parallel to the μth axis. These random variables, ξ_j, are picked independently for each different value of j. At each j, the different possible value of ξ_j are chosen with equal probability. Thus they obey the condition that ξ_j and ξ_k are uncorrelated for different hops, that is if j is different from k.

$$\langle \xi_j \cdot \xi_k \rangle = \langle \xi_j \rangle \cdot \langle \xi_k \rangle \,. \tag{5.2a}$$

Further, because we choose ξ_j in such a way that a vector and its negative are equally likely, it follows at once that

$$\langle \xi_j \rangle = 0 \,. \tag{5.2b}$$

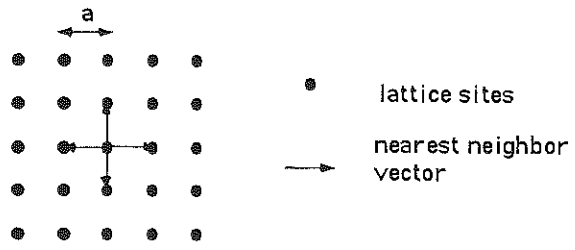

Fig. 5.1. The simple square lattice in two dimensions.

Thus, if j is different from k, the two vectors, ξ_j and ξ_k are uncorrelated. Since the length of all the ξ_j is a, we can find the average of $(\xi_j)^2$ and thus obtain

$$\langle \xi_j \cdot \xi_k \rangle = a^2 \delta_{j,k} \,. \tag{5.2c}$$

In the time interval $M\tau$, the particle hops M times and its displacement is

$$\mathbf{R} = \sum_{j=1}^{M} \xi_j \,. \tag{5.3}$$

Of course, the average displacement vector is zero

$$\langle \mathbf{R} \rangle = \sum_{j=1}^{M} \langle \xi_j \rangle = 0 \,. \tag{5.4a}$$

However, the mean squared displacement is certainly nonzero.

$$\langle \mathbf{R} \cdot \mathbf{R} \rangle = \sum_{j,k=1}^{M} \langle \xi_j \cdot \xi_k \rangle = a^2 \sum_{j,k=1}^{N} \delta_{j,k} \,.$$

Thus the answer is that

$$\langle \mathbf{R} \cdot \mathbf{R} \rangle = a^2 M \,. \tag{5.4b}$$

Equations (5.4a) and (5.4b) together imply that for large M the particle really makes very little net progress. In M steps, it only progresses a distance which is of the order of $aM^{1/2}$. This slow progress should be compared with the behavior which arises when the different directions are weighted unequally so that

$$\langle \xi_j \rangle = \mathbf{u}\tau$$

with $\mathbf{u}$ being the average velocity of the particle. In that case, $\langle \mathbf{R} \rangle = \mathbf{u}\tau N$, so that the displacement varies linearly with N. If there were a nonzero average of ξ_j this 'velocity term' would dominate the long term motion. But, in the situation we have picked, this regular motion is absent and we only have the much slower motion cased by the random walk. Figure 5.2 shows some random walks in two dimensions and lets us compare the resulting random motion with a regular 'up and back' walk which covers the same distance. Notice how much more compact are the random walks. They really are compressed into a small region of space.

5.2 Formulating This Problem

Let $\rho_{r/a,t/\tau}$ be the probability of finding a walker at the position r after a step which has been completed at time t. One can formally define this probability as

$$\rho^{L}_{\mathbf{r}/a,t/\tau} = \langle \delta_{\mathbf{R}(t),\mathbf{r}} \rangle \,. \tag{5.5}$$

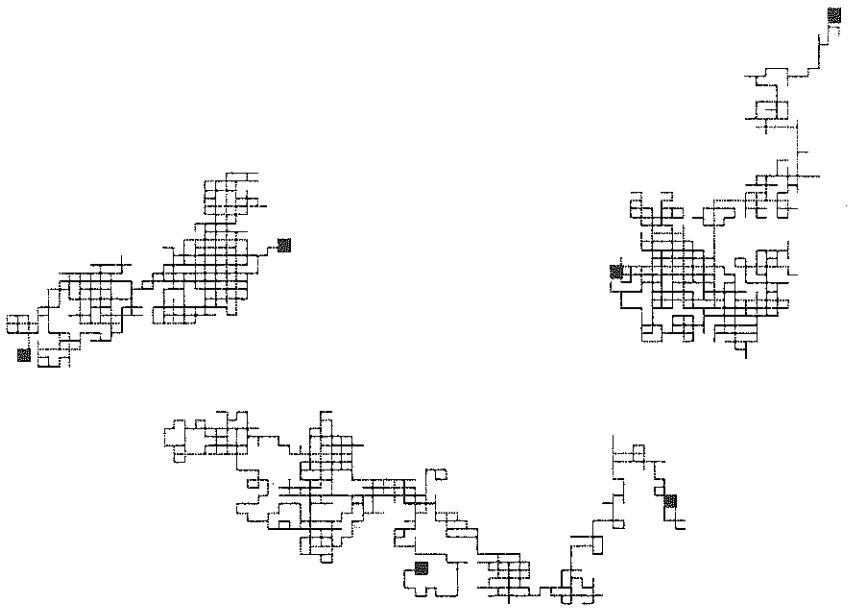

Fig. 5.2. Two-dimensional Random Walks. We see three different random walks. The dark squares represent the beginning and the end of the walk. Each walk contains one thousand steps, with the same step-length. The walks each extend over a relative small distance, start to finish. This small distance should be contrasted with the distance which would be covered if the particles would move in a straight line. The straight line path it would cover sixteen page widths! Thus, random walks are so convoluted that they tend to fall into a quite small region of space.

Here $\mathbf{R}(t)$ is a vector which takes on values depending upon the values of the ξ_j. It is what is called a random variable. On the other hand $\mathbf{r}$ is a prescribed vector which defines the position of the particle. It takes on values which are the sites on our lattice. The hypercubic lattice is defined by taking the lattice sites $\mathbf{r}$ to be $a\mathbf{n}$, where a is the lattice constant and $\mathbf{n}$ is a vector composed of d integers. Similarly the time t has values which are τ times an integer. Thus, we shall usually write this probability as $\rho^L_{\mathbf{n},M}$. The delta symbol on the right hand side of Eq. (5.5) gives a value one whenever the random vector falls on the lattice site, $\mathbf{r}$, and zero otherwise. The average of a quantity which gives a one, when a specific thing happens, and zero otherwise is just the probability of the thing happening. Thus, Eq. (5.5) defines the probability $\rho^L_{\mathbf{n},M}$ that our hopping particle will appear at the lattice site with a position vector $\mathbf{r} = a\mathbf{n}$ at the time $t = M\tau$.

The superscript 'L' in Eq. (5.5) reminds us that we are working with a lattice and that ρ^L is a probability defined on the lattice points. We shall be most interested in the behavior of this probability in the situation in which it is slowly varying in the spatial coordinate, $\mathbf{r}$. To handle that case, we introduce a continuum *probability density*, $\rho(\mathbf{r}, t)$. Now let us assume that these densities are slowly varying, and then calculate the total probability in a region Ω from both probabilities. The region Ω is large enough to include many lattice sites, but small enough so that the probabilities vary very little over the region.

We have:

$$\int_{\mathbf{r}\,\text{in}\,\Omega} d\mathbf{r}\,\rho(r,t) = \sum_{\mathbf{r}\,\text{in}\,\Omega} \rho^L_{\mathbf{r}/a,t/\tau}\,.$$

We take the two ρ's out from under the summations and integrations that they are connected by a factor of the number of sites per unit volume. Since one site has a volume a^d

$$\rho(\mathbf{r},t) = a^{-d}\rho^L_{\mathbf{r}/a,t/\tau}\,. \tag{5.6}$$

We can solve Eq. (5.5) quite explicitly to obtain the time dependence of the probability. To get started we need to know the value of ρ at some beginning time, say $t = 0$. Let us take as our initial data the statement that at $t = 0$ the particle is certainly at a specified point. For simplicity let us take this point to be $\mathbf{r} = 0$. In symbols

$$\rho^L_{\mathbf{n},0} = \delta_{\mathbf{n},0}\,.$$

At the time $t = M\tau$, with M being a positive integer, the particle has gone through M random steps and has a new position vector which is given by Eq. (5.3). Given this form of $\mathbf{R}(t)$ it is relatively easy to compute the average (5.5). To simplify the notation, we do this calculation in one dimension. It is easily generalized to any dimensionality. In one dimension each ξ takes on the two values $\pm a$. To take advantage of statistical mechanics tools, we would like to recast the right hand side of Eq. (5.5) as a sum of exponentials. This is easy since there is a fourier representation of the one-dimensional delta symbol as:

$$\delta_{n,m} = \int_{-\pi}^{\pi} \frac{dq}{2\pi} e^{iq(n-m)}\,, \tag{5.7}$$

whenever n and m are integers. Thus Eq. (5.5) may be written as

$$\rho^L_{r/a,t/\tau} = \int_{-\pi}^{\pi} \frac{dq}{2\pi} \langle e^{iq(R(t)-r)-a}\rangle = \int_{-\pi}^{\pi} \frac{dq}{2\pi} e^{-iqr/a} \left\langle \prod_{j=1}^{M} e^{iq\xi_j/a} \right\rangle .$$

Since each ξ_j is independent of the others, the average of the product is just the product of the averages and that can computed at once giving

$$\rho^L_{n,M} = \int_{-\pi}^{\pi} \frac{dq}{2\pi} e^{-iqn}[\langle e^{iq\xi_j/a}\rangle]^M = \int_{-\pi}^{\pi} \frac{dq}{2\pi} e^{-iqn}[\cos q]^M\,. \tag{5.8}$$

You are asked to evaluate this integral in homework Problem 5.7.

Equation (5.8) provides a nice but not fully explicit result. To get a little further notice that we can see one very sharp result from doing this integral. In this one-dimensional example, each of the M steps takes us either to the left or to the right. If M is even, we will have in total traversed an even number of steps; If M is odd we will have traversed an odd number. Then if $r = ma$ and $t = M\tau$, we get a nonzero result only for $\rho(r,t)$ only if n and M have the same evenness or oddness. Thus, quite explicitly:

$$\rho^L_{n,M} = 0 \quad \text{if} \quad n \neq M \mod 2\,. \tag{5.9a}$$

If the result is nonzero, the integral in (5.8) covers the same ground twice since the regions from $-\pi$ to $-\pi/2$ and the one from $\pi/2$ to π duplicate the integrand from $-\pi/2$ to $\pi/2$. When the integral is nonzero, it is

$$\rho^L_{n,M} = 2\int_{-\pi/2}^{\pi/2} \frac{dq}{2\pi} e^{-iqn}[\cos q]^M \,. \tag{5.9b}$$

The result can be worked out further in the limit as the number of steps, M, gets to be very large. We do the calculation as a steepest descent integral, with the exponent being $iqn + M \ln\cos q$. This exponent has maxima at $q = 0$ and $q = \pi$. The steepest descent calculation applied to Eq. (5.9b) gives:

$$\rho^L_{n,M} = 2\sqrt{\frac{1}{2\pi M}} \exp\left(-\frac{n^2}{2M}\right) \qquad \text{for } n = M \mod 2\,. \tag{5.10a}$$

The equivalent expression in terms of dimensional variables is:

$$\rho^L_{r/a,t/\tau} = 2\sqrt{\frac{\tau}{2\pi t}} \exp\left[-\frac{r^2\tau}{2ta^2}\right] \qquad \text{for } n = M \mod 2\,. \tag{5.10b}$$

The result (5.10) indicates that after a time t the probability is spread out over a distance, r, of the order of

$$r = a\sqrt{\frac{t}{\tau}}\,. \tag{5.11}$$

Notice how this distance grows as the square root of the time. This growth rate, which we first saw in Eq. (5.4b), is characteristic of random walk behavior.

5.3 The Diffusion of Probability and Particles

We now look at precisely this same process from a different point of view. Recall that each site, $\mathbf{r}$, contains walkers which have come from the neighboring sites $\mathbf{r}+\xi$, where ξ is one of the $2D$ nearest neighbor vectors. Each hopping vector, ξ, has a likelihood $1/(2d)$ of being picked. This hopping took place during the proceeding time interval τ. In symbols, what we have just said is the equation,

$$\rho^L_{\mathbf{r}/a,1+t/\tau} = \frac{1}{2d}\sum_{\xi} \rho^L_{(\mathbf{r}-\xi)/a,t/\tau}\,. \tag{5.12}$$

We can use this expression to compute the probabilities in a very explicit fashion. For example, let us focus on a two-dimensional example in which, as before, at time 0 the site $\mathbf{r} = 0$ is the only one occupied. (See Fig. 5.3) Then at this time $\rho^L_{\mathbf{n},0} = \delta_{\mathbf{n},0}$. At the next time, all four of the sites neighboring to this one will have occupation probability 1/4, with the result shown in Fig. 5.3 in which

$$\rho^L_{\mathbf{n},1} = \begin{cases} 1/4 & \text{for } \mathbf{n} = \pm(1,0) \text{ or } \pm(0,1)\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

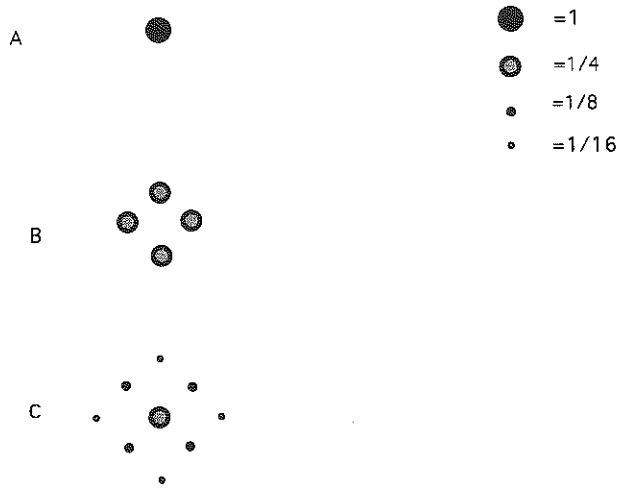

Fig. 5.3. Probabilities for occupation of a given site at three successive times. At the earliest time, A, the particle starts out at the center. After one step B is has a probability 1/4 of being at the neighboring sites. The next step, C, brings further outward motion.

The next few times are also shown on this figure. Notice that, as time goes on, ρ^L begins to spread out over the lattice. Once again all sites are divided into two groups: those with an even sum of n_μ and those with an odd sum. On even time steps ρ^L is nonzero at the first kind of sites; on odd steps it is nonzero at odd sum sites. Aside from this odd-even vanishing, after a few time steps ρ^L will become a slowly varying function of its spatial argument. It also changes very little with each successive time-step. (See Eq. (5.10)).

One can make use of this slow variation to convert Eq. (5.12) into a partial differential equation. We shall do it in a moment. It is easy. But before we do that, it is important to say why we are going to this partial differential equation. Imagine an experimentalist observing the diffusion of something, say Copper through some lattice of another material, say Gold. Our experimentalist is interested in the total amount of copper which gets from one region of the material to another. So this person wishes to take a macroscopic view in which instead of observing a particular lattice site, the observation is of a whole region containing many lattice sites. Similarly instead of looking for one hop, our imagined experimentalist is interested in the net result of many hops. So the person might well measure an average over many of our time intervals, τ. Because the diffusion process eventually leads to a flow over large regions of space, covering many lattice constants, it is important to have a good description of what is going on. But because the observer may be ignorant of or uninterested in the details of lattice structure, this information should not be $\rho^L_{\mathbf{n},M}$. Instead the right information is contained in a density $\rho(\mathbf{r},t)$ defined so that $\rho(\mathbf{r},t)d\mathbf{r}$ is the probability of finding a particular hopping atom in the element $d\mathbf{r}$ about the spatial point $\mathbf{r}$. A sharp connection between these quantities can be made when both vary slowly in space and time. In Eq. (5.6), we write

$$\rho(\mathbf{r},t)a^d = \rho^L_{\mathbf{n},M} \qquad \text{if } \mathbf{r} = a\mathbf{n},\ t = M\tau\,. \tag{5.13}$$

In order to have this work with every cell in the lattice we must have $\rho^L_{\mathbf{n},M}$ be a slowly varying in its arguments.[1] We shall henceforth assume that ρ^L is indeed slowly varying.

To see the consequences of ρ's smoothness multiply Eq. (5.12) by a^d and get the consequence

$$\rho(\mathbf{r}, t+\tau) - \rho(\mathbf{r}, t) = \sum_{\xi} \frac{\rho(\mathbf{r}-\xi, t) - \rho(\mathbf{r}, t)}{2d} \,. \tag{5.14}$$

No approximations have been made so far. To get the diffusion approximation, expand both side of the equation in the 'small' numbers τ and ξ and hold on to the first non-vanishing term on each side of the equation. On the left hand side, the lowest order non-vanishing term is first order. On the right one expands in a power series in ξ. The expansion gives

$$\tau \frac{\partial}{\partial t}\rho(\mathbf{r}, t) = \sum_{\xi} \frac{-\xi\cdot\nabla + (\xi\cdot\nabla)^2/2}{2d}\rho(\mathbf{r}, t) \,.$$

The sum on the right is easily calculated. The first term just cancels out since $\langle\xi\rangle = 0$; the second order term is nonzero since the components of a given step obey Eq. (5.2c), which has the alternative expression

$$2d\langle\xi_\mu\xi_\upsilon\rangle = \sum_{\xi}\xi_\mu\xi_\upsilon = 2\delta_{\mu,\upsilon} \,.$$

Thus the continuum version of our equation of motion for the probability distribution reads

$$\tau\frac{\partial}{\partial t}\rho(\mathbf{r}, t) = \frac{a^2}{2d}\nabla^2\rho(\mathbf{r}, t) \,. \tag{5.15}$$

This conclusion is a consequence of the specific assumptions we have made, but there is a much more general conclusion which can be stated:

> In a random walk, in the long run, the probability of finding the walker at the position $\mathbf{r}$ at the time t obeys
>
> $$\frac{\partial}{\partial t}\rho(\mathbf{r}, t) = \lambda\nabla^2\rho(\mathbf{r}, t) \,. \tag{5.16}$$

An equation of the form (5.16) is called a *diffusion equation.* The coefficient of the Laplacian term, λ, is called a *diffusivity.* Under the particular assumption we have made the value of the diffusivity is

$$\lambda = \frac{a^2}{2d\tau} \,. \tag{5.17}$$

[1] Thus the configuration shown in Fig. 5.3 is unacceptable since at each step half the lattice sites will remain unoccupied. However, this is a trivial difficulty, easily fixed by averaging the initial occupation over some region of the lattice containing more than a few lattice sites, or by averaging the observation of ρ over some interval of space or some interval of time.

The diffusion equation, (5.16) arises quite often in physical systems. By definition, conserved quantities, like the probability discussed here, have the property that the total amount of the quantity remains independent of time. If the conserved quantity moves from place to nearby place, in response to gradients of that quantity, then it will almost always obey a diffusion equation.

5.4 From Conservation to Hydrodynamic Equations

A 'hydrodynamic equation' is a continuum partial differential equation which describes the limit of slowly varying flows in a nonequilibrium system. One example is the diffusion equation, discussed in the previous section. In this section, we shall discuss how the hydrodynamic equations follow from a conservation law plus an additional piece of information, called a *constitutive equation*, used to prescribe the current in the conservation law. We get two different kinds of hydrodynamics, one for charged particles, and the other for uncharged ones. Then we describe once more, but this time in d dimensions, the form of solutions to the diffusion equation.

Electrodynamics contains a conservation law for charge which can be expressed in the form

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) + \nabla \cdot \mathbf{j}(\mathbf{r},t) = 0, \qquad \text{conservation law}\,. \tag{5.18}$$

Here, ρ is the charge density and j is the current density. Exactly the same form of conservation law applies in our hopping case, except that ρ now has the meaning of the probability per unit volume for observing a hopping particle and j represents the probability current for such hoppers. Of course, the integrated form of the conservation law

$$\frac{d}{dt}\int_\Omega d\mathbf{r}\ \rho(\mathbf{r},t) = -\int_{\partial\Omega} d\mathbf{S}\cdot\mathbf{j}(\mathbf{r},t) \tag{5.19}$$

follows directly from integrating Eq. (5.18) over a volume, Ω which is bounded by a surface $\partial\Omega$. Equation (5.19) says that the time derivative of the total amount of the conserved stuff in Ω is equal to the negative of the flux of stuff flowing out.

Notice that neither Eq. (5.18) nor Eq. (5.19) provide a theory which will enable us to find ρ. To get that one needs an additional relation called a constitutive equation which defines $\mathbf{j}$ in terms of ρ. Thus we are required to understand the process which is driving a current through the system. In the electrodynamics of materials, there are several different constitutive equations, describing qualitatively different situations. Perhaps the most famous is Ohm's law $\mathbf{j} = \sigma\mathbf{E}$, where σ is the electrical resistivity. This law effective describes the low frequency response of a conductor. To describe the high frequency response, one needs to include the inertia of the charged particles. A charged particle gas is called a *plasma*. Consider a plasma with number density N/Ω and a charge per particle e and mass per particle m which is being accelerated by a electric field in an almost lossless