A sum rule for unbiased Brownian motion, and generalisations of Fick's law for space-dependent diffusivity

A sum rule for unbiased Brownian motion, and generalisations of Fick's law for space-dependent diffusivity

Robert Collins† and Tadashi Takemori‡

† Physics Department, University of York, Heslington, York YO1 5DD, UK
‡ Institute of Applied Physics, University of Tsukuba, Tsukuba City, Ibaraki 305, Japan

Abstract. For a Brownian particle moving in a succession of independent random hops with no local directional bias, its expectation position is shown to be constant in time, even when the hopping statistics vary from point to point. This gives rise to a simple sum rule on the probability density $w(r, t)$ of the particle. For a position-dependent diffusivity D derived from such a locally unbiased random process, this sum rule is shown to be consistent with the relation (i) $j = -\mathrm{grad}(Dw)$ as the appropriate generalisation of Fick's Law, but inconsistent with the alternative form (ii) $j = -D\,\mathrm{grad}\,w$ often used. Properties of w derived from (i) are developed, and its applicability to diffusion processes in real solids is discussed. It is shown that the sum rule also holds in the presence of absorbers, and in some cases can give eventual absorption probabilities without the need to solve the full time-dependent problem.

1. Introduction

For some time there have been two alternative interpretations of the term 'diffusivity', denoted here by D following majority usage. In cases where D depends on position, these give rise to two alternative forms for the corresponding diffusion equation, which can give qualitatively different results. Accordingly we need criteria to decide which equation (if either) should be used in any given case. In the present paper we discuss a simple extension of the Chandrasekhar 'random flights' model of Brownian motion under no external forces. This gives rise to a simple sum rule on the particle probability density, based on the fact that the expectation particle position remains constant in time. Although the sum rule is rather intuitive, it does not seem to have been formally stated, nor its implications noted. For diffusion processes to which the model does apply, the sum rule gives a clear indication that one of the diffusion equations considered is definitely incorrect, and hence for these processes at least it helps to resolve the problem of deciding which equation to apply.

The sum rule is also of interest in its own right, since it still applies when absorbers are present. It could be used to monitor the accuracy of numerical calculations, and in some cases it also gives direct answers to questions about absorption probabilities (a simple example is given in § 7).

Although the formalism used is general, only applications to solids will be considered here. In many diffusion processes, the migrating particle motion consists of an irregular statistical part superimposed on a systematic motion through the solid induced by local vector fields. For solids these fields are usually gravitational or electric. Usual practice is to identify the irregular component as 'diffusion' and the systematic component as 'drift'. In cases where the diffusion statistics are independent of position there is no difficulty in identifying the separate contributions of these two components. If the diffusion statistics *do* depend on the position of the migrating particle, however, there can be a corresponding drift contribution additional to those of external fields. For the first treatment given in the present paper, we consider only simple systems in which there are no external vector fields present. However, external *scalar* fields (in particular temperature gradients) *are* considered.

2. Sum rule for locally unbiased Brownian motion

The 'random flights' model of Brownian motion (Chandrasekhar 1943) has formed the basis for many studies of diffusion. The trajectory γ of a migrating Brownian particle M is taken to consist of successive displacements (hops) $h_1, h_2, h_3, \ldots$. The starting point of M is r_0, so that its position vector r_i after i hops is given by

$$r_i = r_0 + h_1 + h_2 + \ldots + h_i. \tag{2.1}$$

Suppose that M has reached r on its last hop. The probability density $\psi(h|r)$ for the next hop h is taken to depend only on r, and to be otherwise independent of the previous motion of M. It is also taken to be 'unbiased' in the sense that the expectation value $\langle h \rangle$ is the null vector. Hence we assume the relations (with $h = |h|$)

$$\int d^3h \, \psi(h|r) = 1 \qquad (\text{all } r) \tag{2.2}$$

$$\int d^3h \, h\psi(h|r) = 0 \qquad (\text{all } r). \tag{2.3}$$

Let $w_i(r)$ denote the probability density for the position vector of M after i hops. Then we have

$$w_{i+1}(r) = \int d^3r' \, \psi(r - r'|r')w_i(r'). \tag{2.4}$$

Let $\langle \ \rangle$ denote expectation values taken over the whole ensemble E of possible independent single-particle migration paths γ. Then we have

$$\langle r_{i+1} \rangle = \int d^3r \, r w_{i+1}(r)$$

$$= \int d^3r \, r \int d^3r' \, \psi(r - r'|r')w_i(r')$$

$$= \int d^3r' \, w_i(r') \int d^3r \, r\psi(r - r'|r')$$

$$= \int d^3r' \, w_i(r') \int d^3h \, (r' + h)\psi(h|r').$$ (2.5)

Using (2.2) and (2.3) this becomes

$$\langle r_{i+1} \rangle = \int d^3r' \, (r' + 0)w_i(r') = \langle r_i \rangle.$$ (2.6)

By repeated recurrence this gives

$$\langle r_i \rangle = r_0 \qquad \text{(all } i\text{).}$$ (2.7)

So far we have followed the Chandrasekhar model in which the hop index i was taken proportional to time t. We now introduce into the Brownian motion an additional time structure in the form of a position-dependent *mean hop rate* $\nu(r)$ defined as follows. Suppose M arrives at r at time t. We now take the *dwell time* τ before its next hop to be a random variable, which is independent of t and depends only on r, being otherwise independent of the previous hops of M. Let $f(\tau|r) \, d\tau$ be the probability that the dwell time at r lies between τ and $\tau + d\tau$, where $f(\tau|r)$ is supposed differentiable in τ. Then $\nu(r)$ is defined to be the reciprocal of the expectation dwell time $\langle \tau(r) \rangle$ and is given by

$$1/\nu(r) = \langle \tau(r) \rangle = \int_0^\infty d\tau \, \tau f(\tau|r).$$ (2.8)

Because of the variation of dwell time with position, there is no longer a one-to-one correspondence between time t and number i of hops, so that the number of hops in a given time depends on the particular path γ taken by M. Averaging over all possible γ in E, let $W_i(t)$ be the probability that, in time t, M makes exactly i hops. Then the positional probability density $w(r, t)$ of M at t is given by

$$w(r, t) = \sum_i W_i(t)w_i(r)$$ (2.9)

so that at t the expectation particle position $\langle r(t) \rangle$ is given by

$$\langle r(t) \rangle = \int d^3r \, rw(r, t) = \sum_i W_i(t)\langle r_i \rangle.$$ (2.10)

Using (2.7) this reduces to

$$\langle r(t) \rangle = \int d^3r \, rw(r, t) = r_0 \qquad \text{(all } t\text{).}$$ (2.11)

Relation (2.11) is the required sum rule on $w(r, t)$. Because of the assumed differentiability of $f(\tau|r)$ with respect to τ, it can be written in the alternative form

$$\frac{d}{dt} \langle r(t) \rangle = 0$$ (2.12)

which is more appropriate when we are only given an initial probability distribution for M and not a known starting point. In the particular case of spatially *uniform* diffusivity the sum rule has been noted before (p. 168 of Reichl 1980) but not apparently as the general result given here.

To summarise, the assumptions made in deriving this result are:

The individual hops are unbiased in
the sense specified by (2.3).   (2.13a)

The hopping process is Markovian in that
both the space and time statistics of the
next hop depend only on the present position   (2.13b)
of the particle and not on its previous
history nor explicitly on the time.

Both the above conditions are necessary for the foregoing argument to be valid. In particular (2.13a) by itself is not enough. (The authors are indebted to one of the referees for pointing this out, and thus enabling them to replace an earlier incorrect derivation of (2.12.).)

What we have shown is that (subject to assumptions (2.13)), irrespective of other details of the positional dependence of both the spatial distribution (as given by $\psi(h|r)$) and the time structure (as given by $f(\tau|r)$) of the local hopping process, the expectation position of the migrating particle remains constant in time.

It remains to consider to what extent this model is applicable to real diffusion processes in solids. It seems reasonably well established (at least for low densities of the migrating species) that most of the modes of diffusion through solids can be represented as independent hopping processes satisfying (2.13b)—for a general discussion see for example Manning (1968). Correlation effects can be taken into account by reducing the effective local hop rate $\nu(r)$ by the usual correlation factor f (cf. p 75 of Manning 1968) without affecting the basic argument.

Assumption (2.13a) is more difficult, and in applications must be justified on a case-by-case basis. This point is taken up again in § 6.

3. Diffusion equations for locally unbiased Brownian motion

3.1. Diffusion in an unbounded region

Consider a particle migrating through a solid by Brownian motion in the absence of external vector fields. The motion is not, however, supposed homogeneous, so that the diffusion rate may vary from point to point.

At any point r at time t, let $j(r, t)$ denote the net probability transfer current density of the migrating particle—cf. for example, p 62 of Stratonovich (1963) or p 72 of Risken (1984). Conservation of probability then gives the relation

$$\partial w/\partial t = -\operatorname{div} j. \tag{3.1}$$

To get a diffusion equation equation for w in terms of t, we need a second relation between w and j. For Brownian diffusion, which is locally unbiased in the sense already described, there are two alternative forms, both extensively used. The first (Seitz 1951, Crank 1975, de Groot and Mazur 1983, Manning 1968) is

$$j = -D_F \operatorname{grad} w \tag{3.2}$$

giving the diffusion equation

$$\partial w/\partial t = \operatorname{div}(D_F \operatorname{grad} w). \tag{3.3}$$

Here $D_F(r)$ is called the *diffusivity* or *diffusion coefficient* at r. It is here given the

suffix F because (3.2) has the same form as the Fick equation for diffusion through a homogeneous medium (diffusivity independent of position). The alternative relation to (3.2) is given by equating to zero the drift term in the Fokker–Planck equation (Risken 1984, Feller 1971, Montroll and West 1987) to give

$$j = -\mathrm{grad}(D_2 w) \tag{3.4}$$

resulting in a diffusion equation of the form

$$\partial w / \partial t = \nabla^2 (D_2 w). \tag{3.5}$$

The quantity $D_2(r)$ is defined as the limit of the second moment of small displacements, and is *also* called the diffusivity. For h small but finite, a first approximation for $D_2(r)$ is given by

$$D_2(r) \simeq \tfrac{1}{6} \nu(r) \int \mathrm{d}^3 h \, h^2 \, \psi(h|r) \tag{3.6}$$

where $\nu(r)$ is defined by (2.8).

For homogeneous media where D_2 is independent of r, then $D_F = D_2$ and both (3.2) and (3.4) reduce to the usual Fick equation, so that there is no inconsistency. For an inhomogeneous medium with a position-dependent $D_2(r)$, however, equations (3.2) and (3.4) are inconsistent and give qualitatively different results. Moreover, for a given $D_2(r)$ there is in general no way of choosing $D_F(r)$ to remove this discrepancy. To see this, consider a case in which $w(r, t)$ ultimately tends to a limiting steady-state form $w^*(r)$, as $t \to \infty$. The necessary and sufficient *steady-state condition* is

$$j = 0. \tag{3.7}$$

Relation (3.4) then gives immediately

$$w^*(r) = \frac{\mathrm{const}}{D_2(r)} \neq \mathrm{const} \tag{3.8}$$

whereas the only possibility from (3.2) is

$$w^* = \mathrm{const} \tag{3.9}$$

which is inconsistent with (3.8) whatever the functional form of $D_F(r)$. It follows that for a diffusive motion satisfying the condition (2.13a) of no local bias, at least one of (3.2) and (3.4) is definitely wrong.

The dilemma of choosing between (3.2) and (3.4) has been noted by van Kampen (1981, p 291) and discussed in detail by Landauer (1978, 1983) and Landsberg (1984). In these treatments the discussion is complicated by the presence of drift terms produced by applied electric fields. These make it difficult to distinguish the underlying physics from those aspects which are purely statistical. For the particular case of locally unbiased diffusion, it turns out that the sum rule already derived can be used to show that it is relation (3.2) which is (except fortuitously) incorrect.

We should here note the argument of Manning (1968) that a position-dependent diffusivity *always* induces a local bias. If true, this would make inconsistent our simultaneous assumptions of no bias and position-dependent D. However, Manning only considers spatial variations of D which arise from a corresponding variation of the *heights* of the energy barriers over which the diffusing particle must hop (cf. figure 1–9(c) of Manning 1968). Clearly if a particle is to hop out of a potential well with one side higher

than the other, the hop probabilities will be biased in favour of the lower side. For physical processes of this type, therefore, Manning's argument is clearly correct. However, as is shown in § 6, real physical systems *do* exist in which a particle executes Brownian motion which simultaneously satisfies the conditions of position-dependent D and also no local bias in the strict sense of equation (2.3), It is to such systems that the results of the present discussion apply.

3.2. Region restricted by a reflecting boundary

Since no diffusing particle can cross a reflecting surface S_{ref}, the corresponding reflecting surface *boundary condition* there is

$$\boldsymbol{\Omega} \cdot \boldsymbol{j} = 0 \tag{3.10}$$

where $\boldsymbol{\Omega}$ is the unit vector normal to S_{ref}. Even in cases where (3.4) *does* hold exactly away from S_{ref}, the sum rule (2.7) is in general no longer true. This is most easily seen by considering the steady-state function w^* in the one-dimensional case where there are reflecting surfaces at $x = 0$ and $x = L$. Then (3.8) becomes

$$w^*(x) = \frac{1}{D_2(x)} \left(\int_0^L \frac{dx}{D_2(x)} \right)^{-1} \tag{3.11}$$

and hence

$$\lim_{t \to \infty} \langle x \rangle = \int_0^L \frac{x \, dx}{D_2(x)} \left(\int_0^L \frac{dx}{D_2(x)} \right)^{-1} \tag{3.12}$$

irrespective of x_0. In (3.11) and (3.12) it is assumed that the integrals converge so that a limiting w^* actually exists.

The outer surface of a solid usually acts as a reflecting surface unless there is appreciable evaporation and/or sputtering of the diffusing atomic species, or the formation of a segregated surface phase. In the latter case the surface becomes (at least in part) an *absorbing* boundary. The general case of an absorbing boundary is discussed in § 7.

4. Use of the sum rule to discriminate between diffusion equations

Equations (3.2) and (3.4) are not the only possibilities for diffusion in the absence of external vector fields. However, they are the ones most commonly assumed, particularly when not enough is known about the detailed physics of the migration process of the Brownian particle. In the present section we show that for processes that do satisfy conditions (2.13) (so that the sum rule (2.12) holds) and also have position-dependent D, then (3.2) is definitely wrong. Equation (3.4) is not necessarily correct in every such case, but it is at least consistent with the sum rule, as shown by the following argument.

For the moment we assume that the particle's motion is unrestricted by confining surfaces, so that it is free to move over all space (this assumption will be relaxed in § 6).

Also we assume that both D_F and D_2 are bounded everywhere by some maximum value D_M, so that we can write

$$D_F < D_M \qquad D_2 < D_M. \tag{4.1}$$

To find bounds on w and j at large distances, we consider the case where the particle starts within a fixed distance a of the origin, i.e.

$$\begin{aligned} w(r, 0) &> 0 \qquad (r < a) \\ w(r, 0) &= 0 \qquad (a < r). \end{aligned} \tag{4.2}$$

Then, for given t, we have

$$w, j \sim \exp(-r^2/2D_M t) \qquad (r \to \infty) \tag{4.3}$$

which gives the relations

$$\lim_{r \to \infty} r^\lambda w(r, t) = 0 \qquad (0 < \lambda) \tag{4.4}$$

and

$$\lim_{r \to \infty} r^\lambda j(r, t) = \mathbf{0} \qquad (0 < \lambda). \tag{4.5}$$

Apart from the condition (4.2), $w(r, 0)$ is arbitrary. Let I denote the region interior to a large spherical surface S of radius R, centred at the origin. Then we have

$$\langle r \rangle = \int_{(\infty)} d^3r \, rw = \lim_{R \to \infty} \langle r \rangle_R \tag{4.6}$$

where

$$\langle r \rangle_R = \int_I d^3r \, rw. \tag{4.7}$$

Using (3.1) we have

$$\frac{d}{dt} \langle r \rangle_R = \int_I d^3r \, r \frac{\partial w}{\partial t} = -\int_I d^3r \, r \operatorname{div} j = \int_I d^3r \, j - \int_S dS \, r(\mathbf{\Omega} \cdot j) \tag{4.8}$$

where $\mathbf{\Omega}$ is the unit outward vector $\mathbf{R}/R$ normal to dS. Using (4.3) we have

$$\int_S dS \, r(\mathbf{\Omega} \cdot j) = O[4\pi R^3 \exp(-R/2D_M t)^2] \to 0 \qquad (R \to \infty) \tag{4.9}$$

so that in the limit we need only consider the first term on the RHS of (4.8). If we assume relation (3.4) then we have

$$\int_I d^3r \, j = -\int_I d^3r \, \operatorname{grad}(D_2 w) = \int_S dS \, \mathbf{\Omega} D_2 w \to 0 \qquad (R \to \infty) \tag{4.10}$$

so that we have

$$\frac{d}{dt} \langle r \rangle = \lim_{R \to \infty} \frac{d}{dt} \langle r \rangle_R = 0 \tag{4.11}$$

in agreement with (2.12). Hence the form (3.4) for the diffusion equation automatically ensures that the sum rule is satisfied. In contrast, if we assume relation (3.2) we get

$$\frac{d}{dt} \langle r \rangle_R = - \int_I d^3r \, D_F \, \text{grad } w$$

$$= - \int_I d^3r \, [\text{grad} \, (D_F w) - w \, \text{grad } D_F]. \tag{4.12}$$

The integral of the first term vanishes in the limit $R \to \infty$ from (4.3) as before, giving

$$\frac{d}{dt} \langle r \rangle = \int_{(\infty)} d^3r \, w \, \text{grad } D_F. \tag{4.13}$$

Hence for a $w(r, t)$ that is arbitrary (subject only to (4.2)) it follows that $d\langle r \rangle/dt$ can only be zero when

$$\text{grad } D_F = 0$$

i.e.

$$D_F = \text{const.} \tag{4.14}$$

Hence, the sum rule (2.11) cannot be satisfied by a relation of the form (3.2) when D_F is position-dependent. It follows that for unbiased diffusion of the type considered here, the sum rule shows that (3.2) must be wrong. Hence, in cases where (3.2) and (3.4) are the *only* possibilities, the foregoing argument points to (3.4) as the generalisation of Fick's equation when the diffusivity is position-dependent.

In one dimension, equation (3.6) takes the form

$$D = \nu \langle h^2 \rangle / 2. \tag{4.15}$$

In the case of one-dimensional symmetric hops between lattice planes with constant spacing h and position-dependent ν we can show directly (without invoking the sum rule) that the continuum limit leads unequivocally to (3.4) as the correct diffusion equation. The argument is given in Appendix 1. (The result has been obtained earlier (Collins 1981) but this derivation is much shorter.) Accordingly the remainder of this discussion will be devoted to the properties of solutions of (3.4). For this reason we now drop the suffix 2 on D.

5. Some properties of the generalised diffusion equation

For cases where (3.4) (and hence (3.5)) is correct, we may note some properties of the resulting w which for position-dependent D are qualitatively different from those with constant D. From (3.4) we have

$$j = -D \, \text{grad } w - w \, \text{grad } D. \tag{5.1}$$

It follows that even when w is spatially uniform, a non-uniform D induces a progressive transfer of sampling probability w from regions of high D to regions of low D. It is
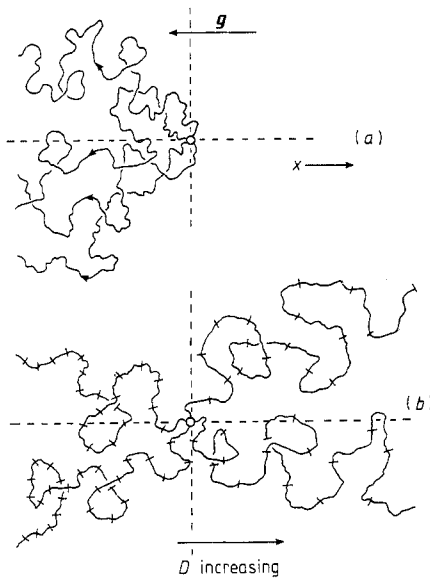
Figure 1. Typical Brownian sample paths. (*a*) Spatially uniform diffusivity *D*; constant gravitational field in the negative *x* direction. (*b*) No gravitational field; *D* increasing monotonically with *x*.

important to note that this is purely a *time-sharing* effect—it is *not* the result of a non-zero contribution v_E to the ensemble average drift velocity $d\langle r\rangle/dt$ of the particle, of the type induced by an external gravitational field or (for a charged particle) an electric field. ($d\langle r\rangle/dt$ does exist in general because of the assumed differentiability of $f(\tau|r)$ with respect to τ, even in a mathematical model where individual paths γ are discontinuous sequences of discrete hops.)

The difference is shown schematically in figures 1(*a*) and (*b*), where a diffusing particle starts from rest at 0. In figure 1(*a*) a constant gravitational field **g** in the negative *x* direction induces a corresponding negative value of $(v_E)_x$, so that the particle sample paths are biased in the direction of **g** as shown. As a result, the region of high probability density *w* moves steadily to the left and we have

$$\frac{d}{dt}\langle x\rangle < 0. \tag{5.2}$$

In contrast, figure 1(*b*) illustrates a case of the type under discussion, where there is no external field acting but *D* increases steadily with *x*. The individual sample paths are now unbiased in direction and their *spatial* structure is symmetrically distributed about 0 (the four shown would have equal probability). The increasing asymmetry of *w* about 0 arises from the asymmetry in the *times* taken to traverse the different paths in figure 1(*b*). The marks on each path correspond to equal time increments. In regions of lower *D* the particle is (on average) moving more slowly, so that the marks are closer together. Consequently the sampling probability density *w* for the particle along these paths is higher for $x < 0$ (low *D*) than for $0 < x$ (high *D*), simply because the particle spends more time in the low-*D* paths. The effect increases with *t* as more paths find their way into the low-*D* $(x < 0)$ region, resulting in a steady transfer of sampling probability to the left. It is only in this sense that the increasing asymmetry of *w* can be described as a

'drift' since, unlike the case of figure 1(a), $\langle x \rangle$ remains zero. This is true since relation (2.11) still holds, which here becomes

$$\langle x \rangle = \int_0^\infty dx \, xw(x) = 0 \qquad g = 0 \qquad \text{(all } t\text{)}. \tag{5.3}$$

The reason that $\langle x \rangle$ does not move to the left into the high-w region is that, although the particle is more likely to be found to the left of 0 than to the right, if it *is* found to the right of 0 then it will on average be found *further* away from 0 because it has been moving faster. Hence the contributions to $\langle x \rangle$ of the regions $x < 0$ (large w but small $|x|$) and $0 < x$ (small w but large $|x|$) exactly cancel, giving relation (5.3). The limiting case of this occurs when absorbers ($D = 0$) are present, illustrated by the second example discussed in § 7.

The second feature of systems where D is spatially non-uniform is that if D decreases smoothly to zero in any region, there may not exist any limiting steady-state probability density $w^*(r)$. For example, consider the one-dimensional case of a particle diffusing along the x axis between reflecting boundaries at $x = 0$, $x = c$ under no external forces and a non-uniform diffusivity given by

$$D = D_0(x/c)^\mu \qquad (0 < x < c). \tag{5.4}$$

If it exists, $w^*(x)$ must have the form given by (3.7), i.e.

$$A/x^\mu \qquad (A \text{ const}). \tag{5.5}$$

However, since the integral

$$\int_0^c dx/x^\mu \tag{5.6}$$

converges for $\mu < 1$ and diverges for $1 \leq \mu$, it follows that a $w^*(x)$ correctly normalised to unity over $0 < x < c$ exists only for $\mu < 1$. For the divergent case $1 \leq \mu$ the interpretation is that as t increases, the sampling probability accumulates nearer to $x = 0$ as more and more sample paths get trapped for long periods in the very slow region where D is nearly zero. However, for the same reason the rate at which this happens itself becomes progressively slower, and for $1 \leq \mu$ this latter effect is dominant and prevents a limiting $w^*(x)$.

6. Applicability of the model to diffusion in real solids

So far the analysis has related to the migration of a single isolated particle. For most practical applications we need to consider the simultaneous migration of N such particles, where N is large enough for fractional fluctuations (of order $1/\sqrt{N}$) to be negligible, so that w and j may be scaled to give respectively the macroscopic number density n and current density J according to the relations

$$n = Nw \tag{6.1}$$

$$J = Nj. \tag{6.2}$$

However, n must be small enough to justify the assumption that each migrating particle
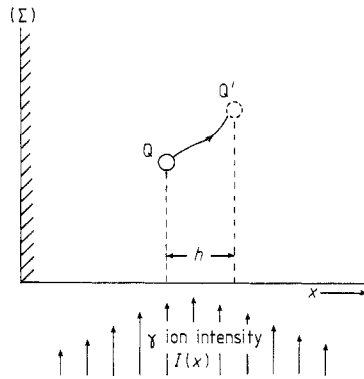
Figure 2. Random depth relocation of atoms by inhomogeneous ion beam parallel to solid surface Σ.

is unaffected by the $N-1$ others. For diffusion through a host species of atomic density ρ, this implies the relation

$$n \ll \rho. \tag{6.3}$$

Since the total number of host atoms in a typical solid specimen is of order 10^{20}, we can take $N \sim 10^{14}$–10^{16} to satisfy (6.3) and still keep $1/\sqrt{N}$ negligibly small. Assuming that (6.3) is in fact satisfied, then for systems satisfying conditions (2.13) the sum rule takes the form

$$\langle r(t) \rangle \equiv (1/N) \int d^3 r\, rn(r, t) = c \qquad \text{(const)} \tag{6.4}$$

and equations (3.4) and (3.5) become

$$J = -\text{grad}(Dn) \tag{6.5}$$

$$\partial n/\partial t = \nabla^2(Dn). \tag{6.6}$$

The first application we consider is rather special but of some technological importance. It occurs as a limiting case in the analysis of implant depth profiles in solids by secondary-ion molecular spectroscopy (SIMS). The system considered is shown in figure 2. A small number of implant (α) atoms is distributed inside a solid matrix of host (β) atoms with a plane surface Σ at $x = 0$. A beam of high-speed ions (γ) with depth-dependent current density $I(x)$ is incident on the solid. Although the magnitude of I varies with depth x, its direction is constant and parallel to Σ. An α atom at Q which is struck by a γ ion eventually comes to rest at some point Q'. Relative to a given Q, in general we expect the distribution of possible points Q' to have a bias in the I direction. However, its projection on the x axis must be symmetric about Q, and hence its probability density function $\psi(h|x)$ is symmetric in h. (The $\psi(h|x)$ used here is analogous to the *ballistic relocation function* $F(x, h)$ introduced by Sigmund and Gras-Marti to describe ballistic relocation *parallel* to I (Sigmund and Gras-Marti 1980, 1981, Gras-Marti and Sigmund 1981).) For a given α atom, successive such displacements h (from different γ ions) are clearly independent, and so here we have a one-dimensional case where both the conditions (2.13) are satisfied.

Consequently, provided that the α atoms are far enough from Σ to have a negligible chance of reaching it, the sum rule (6.4) is satisfied in the form

$$\langle x(t) \rangle \equiv (1/N) \int_0^\infty \mathrm{d}x \, xn(x, t) = \text{const} \tag{6.7}$$

where $n(x, t) \, \mathrm{d}x$ now denotes the total number of α atoms lying in the depth layer $(x, x + \mathrm{d}x)$ at t. Hence, provided $\langle h^2 \rangle$ is small enough to justify a diffusion approximation at all (cf. Collins *et al* 1988) we expect n to satisfy the equation

$$\partial n/\partial t = \partial^2 (Dn)/\partial t^2 \tag{6.8}$$

with $D(x) \propto |I(x)|$. In fact it has been shown (Collins and Jimenez-Rodriguez 1982; $V = 0$ in the notation used there) that in this particular case equation (6.8) can be derived directly from the integral equation for ballistic relocation. We describe the system shown in figure 2 as a 'limiting case' because, in current experimental rigs, I cannot be *exactly* parallel to Σ. Also, in practice, implant location is affected by other factors. However, I directions approaching grazing incidence on Σ are used in analysing the profiles of shallow implant layers, and calculating the effects of ballistic relocation alone helps to assess its importance relative to the other factors involved (Collins 1986).

More generally we seek to identify migration processes in solids which satisfy conditions (2.13) and also have a position-dependent diffusivity. These must involve a spatially inhomogeneous distributed energy source for the hopping particle M which produces a correspondingly inhomogeneous hop frequency $\nu(r)$, without introducing appreciable asymmetry into the individual potential wells between which M hops. The most obvious candidates of this type are thermal and radiation-induced diffusion in which there are respectively spatial gradients of temperature and radiation density.

For an atom M migrating through a solid by thermal diffusion, the hops consist of relatively sharp transitions between sites of local equilibrium. Between hops, M oscillates about a local equilibrium centre. The oscillations are not usually simple harmonic but can be characterised by a dominant frequency $\nu_{\mathrm{vib}} \sim 10^{13} \, \mathrm{Hz}$. Typical values of ν_{vib}/ν lie in the range $10 \rightarrow 10^{13} \, \mathrm{Hz}$. Typical values of ν_{vib}/ν lie in the range $10 \rightarrow 10^4$ where ν is deduced from measured values of D. It follows that successive hops of M are separated by an interval of between 10 and 10^4 local oscillations, which will effectively remove any 'memory' of the previous hop. (The 'immediate reversal' possibility has already been dealt with within this model by scaling ν by the correlation factor f.) Consequently we may assume relation (2.13b). The case of (2.13a) is less clear-cut. For present purposes, a temperature gradient may be regarded as having two opposite effects:

  (i) a spatial variation of ν leading to a j directed from high to low T;
  (ii) an induced asymmetry in the hops, giving a j contribution from low to high T.

Effects (i) and (ii) are shown schematically in figure 3, where we consider interstitial diffusion of particles of type M parallel to the x axis when T increases with increasing x. Consider the oscillations of particles M_1 and M_2. M_2 is hotter than M_1, so oscillates faster and will (on average) jump sooner, producing a net contribution to j in the *negative* x direction (effect (i)) by the mechanism already considered. The second effect arises because atoms B are hotter than atoms A, and are oscillating faster in the y and z directions. Hence a temporary gap between B_1 and B_2 (allowing M_1 to move to the right) is likely to occur earlier than a similar gap between A_1 and A_2 (allowing M_1 to move to the left). The result is an induced asymmetry in the hopping
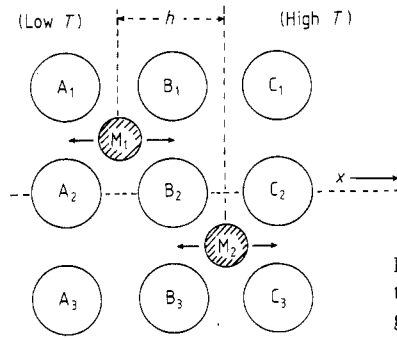
Figure 3. Interstitial diffusion of particles M through a crystalline solid with a temperature gradient but no applied vector field.

probability of M_1 favouring hops to the right, and hence a contribution to j in the *positive x* direction (effect (ii)). In cases where (i) is dominant then both the sum rule (in the absence of reflecting surfaces) and relation (3.4) for D, would be expected to be good approximations. On the other hand, if (ii) is dominant (the case considered by Manning (1968)) then we would expect (3.2) to hold (equation (5–11) of Manning (1968) with $F = 0$). In any given case a full discussion is needed of the kinetics of the energy transfer from the host atoms to the diffusing atoms, and we do not attempt it here. However, it is of some interest to investigate some of the features of systems where (i) *is* the dominant effect, and D is given by (3.4).

In particular we consider the Soret effect of a diffusive mass transfer by a temperature gradient—cf. p 273 of de Groot and Mazur (1983). If D only varies with r because of its variation with T (via ν) then (6.5) becomes

$$J = -D \operatorname{grad} n - n \frac{\mathrm{d}D}{\mathrm{d}T} \operatorname{grad} T. \tag{6.9}$$

The usual assumption made in practice is that the variation of D with T does in fact arise from the temperature variation of ν, i.e. through effect (i), justifying the assumption of form (3.4) for D in this case. For many solids, the variation of D with T is typically (Seitz 1951, Manning 1968) of the form

$$D = D_0 \exp(-T_c/T) \tag{6.10}$$

where $T_c = Q/k_B$ is a constant threshold temperature for thermal diffusion: Q is an activation energy and k_B is the Boltzmann constant. For ranges of T of most experimental interest, the variation of D_0 is small and it can be taken as a constant. Using (6.10), equation (6.9) now becomes

$$J = -D[\operatorname{grad} n + (nT_c/T^2) \operatorname{grad} T]. \tag{6.11}$$

For a finite solid without absorbers we may expect a limiting steady state with $J = 0$ everywhere, when we have

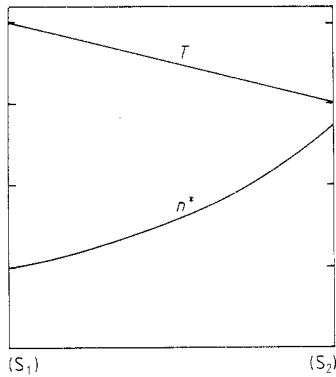$$\operatorname{grad} n^* = -(n^* T_c/T^2) \operatorname{grad} T \tag{6.12}$$

which integrates immediately to give

$$n^* \propto \exp(T_c/T) \tag{6.13}$$

independently of D_0. For a solid with parallel flat faces S_1 and S_2 maintained at constant temperatures T_1 and T_2 we have

$$n_1^*/n_2^* = \exp[T_c(1/T_1 - 1/T_2)] \tag{6.14}$$

irrespective of the distance between S_1 and S_2. The time to achieve the steady state

Figure 4. Steady-state distribution for thermal diffusion in a linear temperature gradient across a solid with opposite faces at temperatures $T_c/3$ and $T_c/4$ (from equation (6.13)).

increases exponentially as T falls below T_c, so in practice T_1 and T_2 cannot be chosen very small. The largest practical temperature range is typically $(T_c/4 < T < T_c/3)$. Putting

$$T_1 = T_c/3 \qquad T_2 = T_c/4 \tag{6.15}$$

we get $n_2^*/n_1^* = \exp(4 - 3) \simeq 2.7$, which suggests that a 3:1 density ratio of the diffusant between the faces should be achievable. The spatial variation of n^* in one dimension for case (6.15) and a linear temperature gradient is shown in figure 4.

For the case of diffusion induced by spatially inhomogeneous radiation (particularly from a laser) much depends on the relative magnitudes of the excitation of the migrating particles and that of the host atoms. When the host excitations are small compared with those of the diffusing particles then effect (ii) will be correspondingly small compared with (i) and conditions (2.13) will effectively be satisfied.

7. Extension of the sum rule to systems containing absorbers

The sum rule also applies to systems containing surface or volume distributions of absorbers (but not, as noted in § 3.2, when *reflecting* surfaces are present). This follows trivially from the fact that (provided the assumptions (2.13) are satisfied) the sum rule holds irrespective of the magnitude of local mean hop rate $\nu(r)$. In particular it holds for arbitrarily small $\nu(r)$. Since an absorbing site (trap) r_a is one for which $\nu(r_a) \to 0$ (i.e. the dwell time $\tau(r_a) \to \infty$) this is the required result.

In a region containing absorbing material, at any time t the ensemble E of possible sample paths consists of an absorbed set $E_a[t]$, which have already terminated on an absorber somewhere, together with an unabsorbed set $E_u[t]$, corresponding to motions in which the particle is still migrating. We can now write relation (2.11) in the form

$$\int_R d^3r\, r w(r, t) = r_0 \tag{7.1}$$

where R denotes the total region of space (E_u and E_a) accessible to the migrating particle. (For example, if the only absorber present is a simple closed surface S and r_0 is inside S, then R is the region I interior to S, together with S itself.)

Although relation (7.1) gives the most concise statement of the sum rule with absorbers present, it is not necessarily the best form to use in calculations. For most

purposes it is more convenient to rewrite (7.1) by separating the E_a and E_u contributions to $\langle r \rangle$, as follows (for simplicity we treat only the case of absorbing surfaces; the extension to systems containing three-dimensional absorbing distributions is straightforward).

Let Σ_a denote the set of all absorbing surfaces present. At time t, let $P(t)$ denote the probability that the particle has already been absorbed somewhere on Σ_a, and let $p(r_a, t) \, dS_a$ denote the probability that this absorption has occurred on the small element dS_a of surface surrounding the point r_a on one of the absorbers. Then we have

$$P(t) = \int_{\Sigma_a} dS_a \, p(r_a, t). \tag{7.2}$$

Let $w_u(r, t) \, d^3r$ denote the probability that, at time t, the particle has not yet been absorbed and is located in the volume element d^3r surrounding point r in R. Conservation of probability then gives

$$\int_R d^3r \, w_u(r, t) = 1 - P(t). \tag{7.3}$$

The form (7.1) for the sum rule can now be replaced by the relation

$$\int_{\Sigma_a} dS_a \, r_a p(r_a, t) + \int_R d^3r \, r w_u(r, t) = r_0. \tag{7.4}$$

If the particle is certain to be eventually absorbed ($P(t) \rightarrow 1$ as $t \rightarrow \infty$) and *also* (it does not follow automatically) if

$$\lim_{t \to \infty} \int_R d^3r \, r w_u(r, t) = 0 \tag{7.5}$$

then relation (7.4) gives

$$\int_{\Sigma_a} dS_a \, r_a p^*(r_a) = r_0 \tag{7.6}$$

where $p^*(r_a) \, dS_a$ is the probability that the particle will eventually be absorbed on dS_a.

In some simple cases, relation (7.6) by itself can give precise answers for ultimate absorption probabilities. This is shown by the following simple example.

Consider the case (illustrated in figure 5) of a particle diffusing in the region R between infinite parallel absorbing planes S_0 (at $x = 0$) and S_L (at $x = L$). At $t = 0$ the particle starts from rest at a point B at $x = b$. Then γ_u is a typical path from the subset E_u still unabsorbed at a later time t. The problem is to determine the probabilities P_0^* and P_L^* that the particle will eventually be absorbed on S_0 or S_L respectively. For simplicity we consider the usual case where the probability P^* of eventual absorption of the particle is unity, so that

$$P^* = P_0^* + P_L^* = 1. \tag{7.7}$$

The diffusion process is assumed to be (as before) locally unbiased, but the diffusivity D is not necessarily spatially uniform nor even a function of x only. It is assumed to
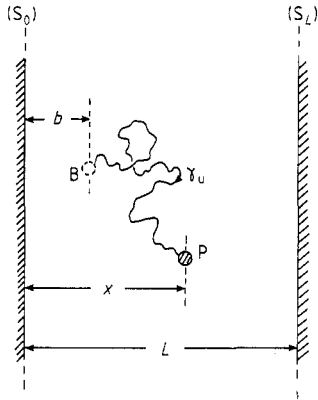
Figure 5. Typical Brownian path for unbiased diffusion between two infinite parallel absorbing planes (prior to particle absorption).

be non-zero throughout R (so that the particle cannot get trapped anywhere except S_0 or S_L). For definiteness we assume some minimum value $D_{\min}$ such that

$$D_{\min} < D(r) \qquad (r \text{ in } R). \tag{7.8}$$

Since the probability of eventual absorption is known to be unity in the uniform case $D = D_{\min}$ everywhere in R, it follows that the same is true if (7.8) is satisfied. For finite L, (7.5) is satisfied and (7.6) here takes the simple form (integrated over y and z)

$$0(1 - P_L^*) + LP_L^* = b$$

giving

$$P_L^* = b/L \tag{7.9}$$

irrespective of the detailed behaviour of $D(r)$ between the plates. Even in the case of constant D, which is relatively easy to calculate, this derivation of (7.9) is far quicker than the treatment usually given (for comparison this is summarised in Appendix 2).

Even in cases where there is probability $P^* = 1$ of eventual absorption, relation (7.5) is not trivial. For example, consider the limiting case (illustrated in figure 6) of the previous example in which $L \to \infty$, i.e. the second absorbing surface S_L disappears from the system. For simplicity we again consider the case of constant D. The region
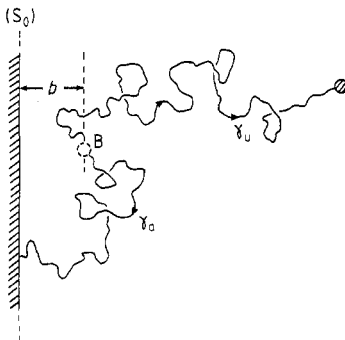
Figure 6. Sample Brownian paths for unbiased diffusion near a single infinite absorbing plane.

R available to the diffusing particle is now the semi-infinite plane $0 < x$. The Kelvin image method (see for example p 341 of Feller (1971)) gives the solution

$$w_u = \frac{1}{2\sqrt{(\pi Dt)}} \{\exp[-(x - b)^2/4Dt]$$

$$- \exp[-(x + b)^2/4Dt]\} \qquad (0 < x, 0 < t) \qquad (7.10)$$

so that the probability $P(t)$ of absorption on S_0 *before* time t is given by

$$P = 1 - \int_0^\infty dx \, w_u = \operatorname{erfc}\left(\frac{b}{2\sqrt{(Dt)}}\right) = 1 - \frac{b}{\sqrt{(\pi Dt)}} + O(t^{-3/2}) \qquad (7.11)$$

i.e.

$$P^* = \lim_{t \to \infty} P(t) = 1 \qquad (7.12)$$

so that the particle is eventually absorbed on S_0 with unit probability. At first sight this appears to imply that $\langle x \rangle \to 0$ as $t \to \infty$. However, from (7.10) we have

$$\langle x \rangle = \int_0^\infty dx \, x w_u = \frac{b}{2}\left[\operatorname{erfc}\left(\frac{b}{2\sqrt{(Dt)}}\right) + \operatorname{erfc}\left(\frac{-b}{2\sqrt{(Dt)}}\right)\right]$$

$$= b \qquad (\text{all } t) \qquad (7.13)$$

in agreement with (7.1). From (7.12), with probability unity any given sample path eventually ends at $x = 0$ like the γ_a shown in figure 6. However, the decreasing *numbers* of unabsorbed paths γ_u are exactly compensated by their steadily increasing *distance* from S_0 measured by the increasing mean value $\langle x \rangle_u$ $(= b/(1 - P))$ taken over $E_u[t]$. The resulting distribution (an increasing component $P(t)\delta(x)$, where $\delta(x)$ is the Dirac delta functions, together with a long 'tail' of decreasing total probability $1 - P(t)$ but extending to larger and larger values of x) is a limiting case of the $w(x, t)$ asymmetry discussed in § 5.

8. Summary and conclusions

In this paper we have considered a class of Brownian motions modelled as Markov hopping processes with local symmetry but with position-dependent hop rate $\nu(r)$. We have shown that all such motions involve a simple sum rule on the probability density w of the migrating particle, which remains valid in the presence of absorbers. Of the two forms of diffusion equation in current use for diffusion with no external vector fields but position-dependent diffusivity D, one (3.4) satisfies this sum rule and the other (3.2) does not. It follows that the two equations are inconsistent and lead to qualitatively different results. The axioms underlying the model have been shown to apply to some diffusion processes, and to be implicit in many diffusion calculations even when not expressly assumed. The form (3.2) due to Fick (and much more commonly used for diffusion in solids) has often been used in cases even when $D(r)$ is calculated using essentially the same model as the one considered here, with which (3.2) has now been shown to be inconsistent. The conclusion of the present paper is that, in these cases at least, (3.4) should be used instead. More generally, in any particular case where D is a function of r the choice of diffusion equation must depend

quite sensitively on the mathematical model assumed. It seems quite likely that in cases of force-free diffusion in solids where the sum rule does fail, then neither (3.2) nor (3.4) will be adequate, and some more complicated equation must be used.

In addition to its value in discriminating between the two diffusion equations, the sum rule has been shown to be a useful theoretical tool, particularly in assigning final particle distributions over systems of absorbers.

Acknowledgments

Appendix 1. Symmetric hopping between equispaced lattice planes

Consider an infinite set of equispaced parallel lattice planes $\{S_k\}$ $\{k = 0, \pm1, \pm2, \ldots\}$ normal to the x axis, with separation h as shown in figure A1. A particle M migrates parallel to the x axis by hopping randomly between the planes. From S_k, M can hop directly only to S_{k-1} or S_{k+1}, each with probability $1/2$. The mean hop rate from S_k is ν_k, so that if M is on S_k at time t, the probability that it has *not* hopped away at $t + \Delta t$ is $\exp(-\nu_k \Delta t)$. Let $W_k(t)$ be the probability that M is on S_k at time t. Then the probability currents going right and left from S_k are both $\nu_k W_k/2$. Similarly the current going left from S_{k+1} is $\nu_{k+1} W_{k+1/2}$. Hence the net current $j_{k+1/2}$ going right through the intermediate (virtual) plane $S_{k+1/2}$ is given by

$$j_{k+1/2} = \tfrac{1}{2}(\nu_k W_k - \nu_{k+1} W_{k+1}). \tag{A1.1}$$

In the continuum approximation we replace $j_k(t)$, ν_k and $W_k(t)$ by smoothly varying $j(x, t)$, $\nu(x)$ and $w(x, t)$ given by

$$j_k(t) = j(x_k, t) \tag{A1.2}$$

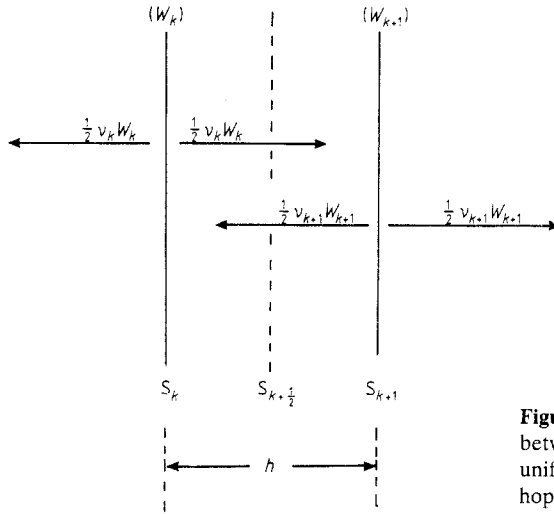$$\nu_k = \nu(x_k) \tag{A1.3}$$

$$W_k(t) = hw(x_k, t). \tag{A1.4}$$

Relation (A.1) then becomes

$$(2/h)j(x_k + h/2, t) = \nu(x_k)w(x_k, t) - \nu(x_k + h)w(x_k + h, t). \tag{A1.5}$$

We now put $x_k = x + h/2$ and expand νw in powers of h to give

Figure A1. Unbiased hopping between parallel lattice planes with uniform spacing h but non-uniform hop rates ν_k.

$$j(x, t) = \frac{\partial}{\partial x}\left[\tfrac{1}{2}h^2 \nu(x) w(x, t)\right] + \mathrm{O}(h). \tag{A1.6}$$

We now make the standard substitution

$$\nu(x) = 2D(x)/h^2 \qquad (D(x) \text{ finite}) \tag{A1.7}$$

and let $h \to 0$, when (A1.6) becomes

$$j = -\partial(Dw)/\partial x \tag{A1.8}$$

which is of the form (3.4).

Appendix 2. Diffusion with uniform D between two parallel absorbing planes

We have to solve the equation

$$\partial w_{\mathrm{u}}/\partial t = D\partial^2 w_{\mathrm{u}}/\partial x^2 \qquad (0 < x < L, 0 < t; D \text{ const}) \tag{A2.1}$$

subject to the boundary conditions

$$w_{\mathrm{u}} = 0 \qquad (x = 0, L; 0 < t) \tag{A2.2}$$

and the initial condition

$$w_{\mathrm{u}}(x, 0) = \delta(x - b). \tag{A2.3}$$

We expand w in terms of the functions $\sin(q\pi x/L)$ $(q = 1, 2, \ldots)$ which satisfy (A2.2). Relations (A2.1) and (A2.3) then reduce to

$$w_{\mathrm{u}} = (2/L) \sum_{q=1}^{\infty} \sin(q\pi b/L) \sin(q\pi x/L) \exp(-q^2\pi^2 Dt/L^2). \tag{A2.4}$$

The probability current j_L into the plane S_L is then given by

$$j_L = -D(\partial w_{\mathrm{u}}/\partial x)_L$$

$$= (2\pi D/L^2) \sum_{q=1}^{\infty} (-1)^{q+1} q \sin(q\pi b/L) \exp(-q^2\pi^2 Dt/L^2) \qquad \text{(A2.5)}$$

and the probability P_L^* of eventual absorption on S_L is given by

$$P_L^* = \int_0^{\infty} \mathrm{d}t\, j_L = (2/\pi) \sum_{q=1}^{\infty} [(-1)^{q+1}/q] \sin(q\pi b/L) = b/L \qquad \text{(A2.6)}$$

as required. (For the last step see for example p 38 of Gradshteyn and Ryzhik (1965).)

References

Chandrasekhar S 1943 *Rev. Mod. Phys.* **15** 1–89 (reissued in Wax N (ed.) 1954 *Selected Papers on Noise and Stochastic Processes* (New York: Dover))
Collins R 1981 *Radiat. Eff. Lett.* **58** 133–7
—— 1986 *Radiat. Eff.* **98** 1–20
Collins R and Jimenez-Rodriguez J J 1982 *Radiat. Eff. Lett.* **68** 19
Collins R, Jimenez-Rodriguez J J, Wadsworth M and Badheka R 1988 *J. Appl. Phys.* **64** 1120–4
Crank J 1975 *Mathematics of Diffusion* 2nd edn (Oxford: Clarendon)
de Groot S R and Mazur P 1983 *Non-Equilibrium Thermodynamics* (New York: Dover)
Feller W 1971 *An Introduction to Probability Theory and its Applications* vol 2 (New York: Wiley)
Gradshteyn I S and Ryzhik I W 1965 *Tables of Integrals, Series and Products* (New York: Academic)
Gras-Marti A and Sigmund P 1981 *Nucl. Instrum. Methods* **180** 211
Landauer R 1978 *Phys. Lett.* **68A** 15–16
—— 1983 *Helv. Phys. Acta* **56** 847–61
Landsberg P T 1984 *J. Appl. Phys.* **56** 1119–22
Manning J R 1968 *Diffusion Kinetics for Atoms in Crystals* (Princeton, NJ: Van Nostrand)
Montroll E W and West B J 1987 *Fluctuation Phenomena* ed. E W Montroll and J L Lebowitz (Amsterdam: North-Holland)
Reichl L E 1980 *A Modern Course in Statistical Physics* (Texas: Texas University Press)
Risken H 1984 *The Fokker–Planck Equation* (Berlin: Springer)
Seitz F 1951 *Phase Transformations in Solids* ed. R Smoluchowski, J E Mayer and W A Weyl (New York: Wiley)
Sigmund P and Gras-Marti A 1980 *Nucl. Instrum. Methods* **168** 38
—— 1981 *Nucl. Instrum. Methods* **182/183** 25
Stratonovich R L 1963 *Topics in the Theory of Random Noise* vol 12 (New York: Gordon and Breach)
Van Kampen N G 1981 *Stochastic Processes in Physics and Chemistry* (Amsterdam: North-Holland)