Chapter 9

Theory of diffusion and viscosity

Diffusion processes occur in a fluid or gas whenever a property is transported in a manner resembling a random walk. For instance, consider the spreading of a contaminant in a fluid. The molecules of the contaminating substance get random kicks in collisions with the individual molecules of the fluid. In fact, this is how the random walk was discovered in nature: the historical example of *Brownian motion*. This example looks at the effects of the kicks on the position of the molecule.

Alternatively, one can think of particles subject to a random force with zero mean: in case the 'kick' changes the momentum of a particle. As a result, particles will diffuse in momentum in the sense that a group of identical particles, all starting out with the same initial momentum, will acquire a spread in momentum values after some time. This happens for instance when a beam of charged particles propagates through a plasma: the electrostatic interactions with the plasma particles (called *Coulomb collisions* as they result from the Coulomb interaction between free charges) will cause a spread in momentum of the beam particles.

The properties of a diffusion process can be studied using a simple statistical model. More complicated discussions can be found in the book by Zeldovich *et al.*[1] and in the book by Van Kampen[2]

9.1   The one-dimensional random walk

I will start the discussion with a simple example: a one-dimensional random walk in space. Consider a random walk in position x, which starts at a time $t = 0$. For simplicity I put the starting position at $x = 0$. The random walk consists of discrete steps forward and backward, with an equal probability, $P_+ = P_- = \frac{1}{2}$, and a constant step size λ.

---

[1]Ya.B. Zeldovich, A.A. Ruzmaikin & D.D. Sokoloff, 1990: *The Almighty Chance*, World Scientific Lecture Notes in Physics, Vol. 20, World Scientific, Singapore

[2]N.G. van Kampen, 1981: *Stochastic Processes in Physics and Chemistry*, North-Holland Publ. Company, Amsterdam

I will use a roman index (e.g. i, j etc.) to denote the step number in a sequence of steps.

In this simple random walk, the position shift Δx_i at step number i can be written in the form

$$\Delta x_i = \tilde{\sigma}_i \lambda \, . \tag{9.1.1}$$

A forward step has $\tilde{\sigma}_i = +1$, and a backward step $\tilde{\sigma}_i = -1$, so $\tilde{\sigma}$ is a direction parameter. After N such steps, the position of a participant in the random walk equals:

$$
\begin{aligned}
x(N) &= \Delta x_1 + \Delta x_2 + \cdots \cdot + \Delta x_N \\[2mm]
&= \lambda \times (\tilde{\sigma}_1 + \tilde{\sigma}_2 + \cdots \cdot + \tilde{\sigma}_N) \; .
\end{aligned}
\tag{9.1.2}
$$

An example of two statistically independent realizations of such a simple prescription are shown in the figure below. In order to be able to investigate the statistical properties of this random walk, I will assume that there is large number of participants, the random walkers, which all follow recipe (9.1.1).

Now consider a the whole *group* of random walkers. Each member of the group performs his random walk *independently*, without being influenced by the others. All start at the same position $x = 0$, and at the same time, $t = 0$. Let us first calculate the *average* position $\overline{x(N)}$ of that group, the 'center of mass'. In calculating that average position one has to realize that in the sequence of steps that leads to $x(N)$ the values $\tilde{\sigma}_i = +1$ and $\tilde{\sigma}_i = -1$ occur with equal probability, $P_+ = P_- = \frac{1}{2}$. Therefore, the average value of the direction parameter $\tilde{\sigma}_i$, when averaged over a large number of steps, or when averaged over a large group[3], is

$$\overline{\tilde{\sigma}_i} = P_+ \times (+1) + P_- \times (-1) = 0 \, . \tag{9.1.3}$$

As a result, the net value of the sum (9.1.2) vanishes:

$$\boxed{\overline{x(N)} = N \, \overline{\tilde{\sigma}} \, \lambda = 0 \, .}$$

$$\tag{9.1.4}$$

---

[3]This interchangeability of a group average and a time average frequently occurs in statistical physics
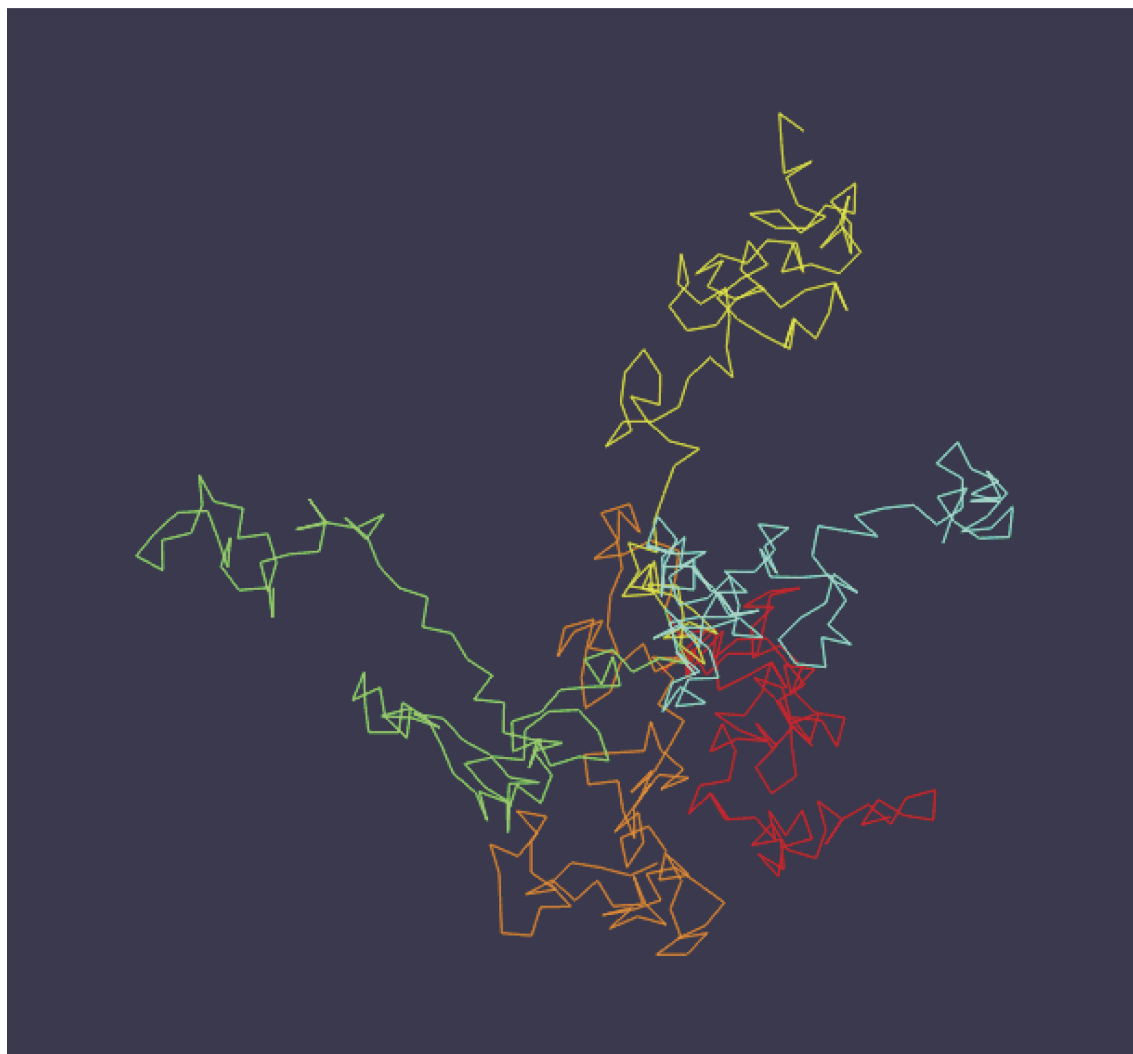
Figure 9.1: *An illustration of a two-dimensional random walk. Five statistically independent realizations (color-coded) of a random walk with the same mean free path are shown.*

After N steps, the average position $\overline{x(N)}$ is still located at the origin, the point where everybody started. This result does **not** imply that all participants are still at $x = 0$, as can be seen as follows. In general, each *individual* participant will not perform an equal number of forward and backward steps. As each participant is statistically independent from the others, the difference between the number of forward steps taken and the number of backward steps taken varies from participant to participant. As a result, the group spreads out along x with -on average- as many random walkers with a positive value x as with a negative value of x. This *dispersion* of the group reflects the difference between the path of *individual* group members and the *mean* path.

The amount of spreading of the random walkers increases with the number of steps. It can be quantified by the so-called *dispersion*, a quantity that follows directly from the average of x^2. From Eqn. (9.1.2) one has

$$
\begin{aligned}
x^2(N) \;=\; & \Delta x_1^2 + \Delta x_2^2 + \cdots \Delta x_N^2 + \\[2mm]
& + 2\Delta x_1 \Delta x_2 + 2\Delta x_1 \Delta x_3 + 2\Delta x_2 \Delta x_3 + \cdots .
\end{aligned}
\tag{9.1.5}
$$

This sum contains two contributions.

1. The first contribution to $\overline{x^2}$ consists of the squares of the *same* N steps, like Δx_1^2, Δx_2^2 etc. Using Eqn. (9.1.1) it follows for each step i that these terms in the expression for x^2 satisfy $\Delta x_i^2 = \tilde{\sigma}_i^2 \lambda^2 = \lambda^2$. After N steps, there are a total of N such terms, each of which contributes equally to the sum for x^2

2. The second contribution to x^2 comes from the terms that are the product of *different* steps: $\Delta x_i \Delta x_j = \tilde{\sigma}_i \tilde{\sigma}_j \lambda^2$ with $i \neq j$. There are a total of $N^2 - N = N(N-1)$ such terms in the sum for x^2. When $x^2(N)$ is averaged over the whole group, the value $\tilde{\sigma}_i \tilde{\sigma}_j = +1$ occurs as often for as does the value $\tilde{\sigma}_i \tilde{\sigma}_j = -1$. The probability $\mathcal{P}_+$ that $\tilde{\sigma}_i \tilde{\sigma}_j$ takes the value $+1$, and the probability $\mathcal{P}_-$ that it takes the value -1, are equal to:

$$
\begin{aligned}
\mathcal{P}_+ \;=\; & P_+ P_+ + P_- P_- , \\[2mm]
\mathcal{P}_- \;=\; & 2 P_- P_+ .
\end{aligned}
\tag{9.1.6}
$$

When $P_+ = P_- = \frac{1}{2}$ one automatically has $\mathcal{P}_+ = \mathcal{P}_- = \frac{1}{2}$. Therefore, this second contribution to x^2 vanishes upon averaging as $\overline{\tilde{\sigma}_i \tilde{\sigma}_j} = \mathcal{P}_+ - \mathcal{P}_- = 0$.

These considerations yield a very simple answer: after each of the random walkers has taken N steps, the dispersion of the whole group of random walkers is determined by:

$$\overline{x^2(N)} = N \times \lambda^2 \ .$$

$$(9.1.7)$$

Therefore, the typical distance $d(N)$ between an individual random walker and his point of origin equals after N steps:

$$d(N) \equiv \sqrt{\overline{x^2(N)}} = \sqrt{N} \times \lambda \ . \qquad (9.1.8)$$

This distance $d(N)$ is a measure of the dispersion. Note that it increases as $d(N) \propto N^{1/2}$. This indicates that a random walk is a very inefficient method of getting anywhere! A straight (directed) walk will take the walker a distance $d \propto N$ after N steps.

So far, I have described the random walk as a sequence of discrete steps of equal size. As a result, the number of steps N figures prominently in all equations. Usually however, one is not interested in the average position $\overline{x}$ and dispersion $\sqrt{\overline{x^2}}$ as a function of N, but as a function of time t. The simplest way of achieving this is by assuming that each participant in the random walk takes ν steps every second, with ν the 'step-frequency', so that the time-interval between two successive steps, the collision time, equals $\Delta t_{\rm c} = 1/\nu$. The typical number of steps taken in a time t is

$$N \approx \nu t \simeq t/\Delta t_{\rm c} \ , \qquad (9.1.9)$$

where the approximate sign is needed since the number of steps N can only take integer values. This relation allows us to write the dispersion as a function of time:

$$\overline{x^2}(t) = \nu \lambda^2 t \ . \qquad (9.1.10)$$

For a one-dimensional random walk, one defines the *diffusion coefficient* $\mathcal{D}$ by the relation:

$$\mathcal{D} \equiv \frac{\overline{x^2}(t)}{2t} = \frac{1}{2} \lambda^2 \nu \qquad (9.1.11)$$

In terms of this diffusion coefficient one can write

$$\overline{x^2}(t) = 2\mathcal{D}t .$$

(9.1.12)

Note that the diffusion coefficient is completely determined by the *intrinsic* properties of the random walk process: the step size λ and the step frequency ν. These definitions, and the derivation presented above, assume implicitly that the diffusion coefficient $\mathcal{D}$, and also λ and/or ν, do not depend on time or on the position x.

9.2   The isotropic three-dimensional random walk

The principles of the one-dimensional random walk can be generalized easily to the case of an *isotropic* three-dimensional random walk. In this context, 'isotropic' means that the random walk has no preferred direction: the diffusion proceeds equally rapid along the x, y and z axes. In addition, I will use a more realistic model for the individual steps in the random walk. Rather than assuming that each step in the random walk has the same length, $|\Delta \boldsymbol{x}| = \lambda$, I will allow different step sizes that are drawn from a probability distribution for the step size. This distribution has a well-defined average step size $\lambda = \langle |\Delta \boldsymbol{x}| \rangle$.

   Let us assume that the random walkers take a step in an arbitrary direction, with a step size equal to Δx_i at the i-th step (see figure below). This implies that the i-th step can be written in vector notation as

$$\Delta \boldsymbol{x}_i = \Delta x_i \, \hat{\boldsymbol{e}}_i .$$

(9.2.1)

Here $\hat{\boldsymbol{e}}_i$ is a vector of unit length, so that $|\hat{\boldsymbol{e}}_i| = 1$, with a direction that is chosen randomly at each step in the random walk. This unit vector plays the same role as the direction parameter $\tilde{\sigma}_i$ in the one-dimensional random walk of the previous Section. The assumption of isotropy implies that the *group average* of the random vector $\hat{\boldsymbol{e}}$ satisfies:

$$\overline{\hat{e}_x} = \overline{\hat{e}_y} = \overline{\hat{e}_z} = 0 \ , \ \ \overline{\hat{e}_x^2} = \overline{\hat{e}_y^2} = \overline{\hat{e}_z^2} = \frac{1}{3} \ , \ \ \overline{\hat{e}_x \hat{e}_y} = \overline{\hat{e}_x \hat{e}_z} = \overline{\hat{e}_y \hat{e}_z} = 0 \ .$$

(9.2.2)

As we compare this with the one-dimensional random walk of the previous section, it is easy to see that the first relation is the equivalent of $\overline{\tilde{\sigma}} = 0$, while the second set of
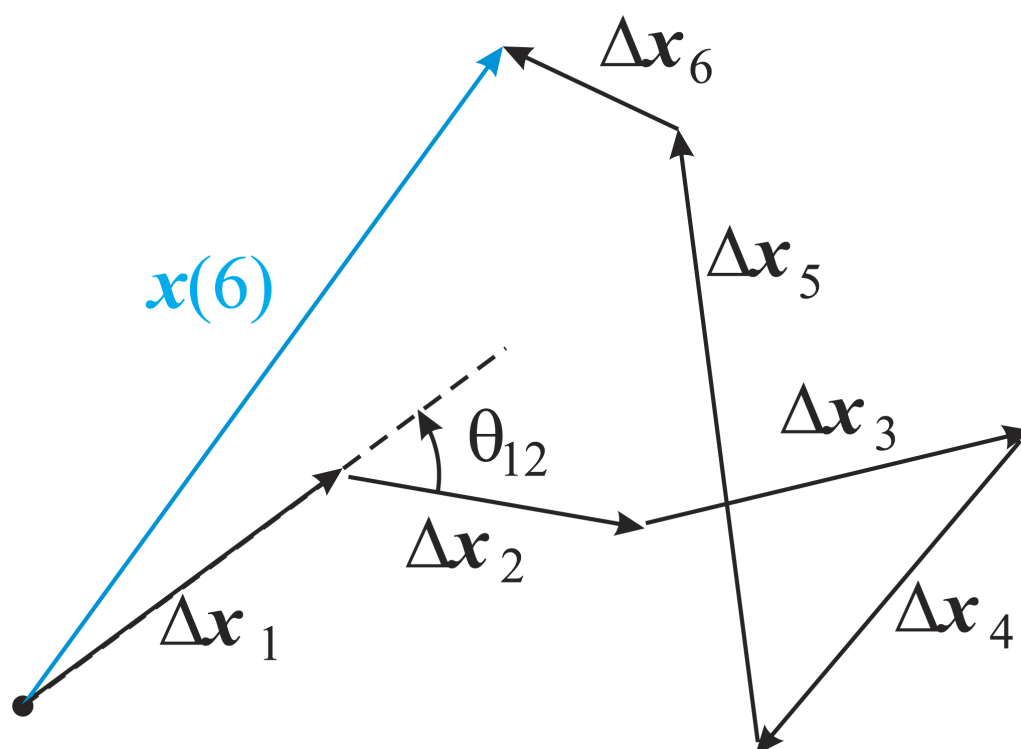
Figure 9.2: *A representation of a three-dimensional random walk with steps of different size. Shown are the first six steps Δx_1-Δx_6, and (in blue) the position vector $x(6)$ from the origin to the position of the random walker after these six steps. Also shown is the angle θ_{12} between the direction of the first and the second step.*

relations replaces $\overline{\tilde{\sigma}^2} = 1$. The second set of relations follows from the fact that $\hat{e}$ is a vector of unit length, and the assumption of isotropy.

The choice of direction of two different steps is arbitrary and, more importantly, uncorrelated. This implies that the *group average* of the step direction satisfies

$$\overline{\hat{\boldsymbol{e}}_i \cdot \hat{\boldsymbol{e}}_j} = \overline{\cos\theta_{ij}} = 0 \ \text{ for } i \neq j. \tag{9.2.3}$$

Here θ_{ij} is the angle between the unit vector of step i and the unit vector of step j. This replaces the relation $\overline{\tilde{\sigma}_i \tilde{\sigma}_j} = 0$ for $i \neq j$ in the one-dimensional random walk.

I will assume that the size of each individual step, $|\Delta\boldsymbol{x}| \equiv \Delta x$, and its direction is determined by some scattering process. The scattering limits the distance that a random walker can travel between two subsequent scatterings. To describe this scattering process, I introduce a *probability distribution* for the step size.

Let the probability $\mathrm{d}P$ that a given step has a length in the range Δx, $\Delta x + \mathrm{d}(\Delta x)$ be given by

$$\mathrm{d}P \equiv \mathcal{P}(\Delta x)\,\mathrm{d}(\Delta x) = \frac{\exp(-\Delta x/\lambda)}{\lambda}\,\mathrm{d}(\Delta x)\,. \tag{9.2.4}$$

The probability distribution $\mathcal{P}(\Delta x) = \mathrm{d}P/\mathrm{d}(\Delta x)$ that is defined in this manner has the following properties:

$$\int_0^\infty \mathrm{d}(\Delta x)\,\mathcal{P}(\Delta x) \ = \ 1\,;$$

$$\int_0^\infty \mathrm{d}(\Delta x)\,\Delta x\,\mathcal{P}(\Delta x) \ \equiv \ \langle\Delta x\rangle = \lambda\,; \tag{9.2.5}$$

$$\int_0^\infty \mathrm{d}(\Delta x)\,\Delta x^2\,\mathcal{P}(\Delta x) \ = \ \left\langle\Delta x^2\right\rangle = 2\lambda^2\,.$$

Here the average $\langle\cdots\rangle$ is an average over the step size distribution (9.2.4). The first relation ensures that the probability distribution is properly normalized. The second relation shows that λ is the *mean* distance traveled by a particle between two subsequent collisions. The third relation defines the mean square distance.

One can now repeat the analysis of the previous Section. We will consider the position of a random walker after N steps, and the average distance of the whole group of random walkers to the starting point of the random walk, which I take to be the origin.

The position of a random walker after N steps is

$$
\begin{aligned}
\boldsymbol{x}(N) &= \Delta\boldsymbol{x}_1 + \Delta\boldsymbol{x}_2 + \cdots + \Delta\boldsymbol{x}_N \\
&= \Delta x_1 \hat{\boldsymbol{e}}_1 + \Delta x_2 \hat{\boldsymbol{e}}_2 + \cdots + \Delta x_N \hat{\boldsymbol{e}}_N .
\end{aligned}
\tag{9.2.6}
$$

We now again take a group average, meaning that we average over all possible directions that the vector $\hat{\boldsymbol{e}}_i$ can take, **and** over all possible step sizes using the distribution (9.2.4). Just as happens in the one-dimensional case, the center-of-mass of the group of random walkers does not change. If all random walkers start at the origin $(x = y = z = 0)$, one has after N steps:

$$
\overline{\boldsymbol{x}(N)} = \langle \Delta x \rangle \sum_{i=1}^{N} \overline{\hat{\boldsymbol{e}}_i} = \lambda \sum_{i=1}^{N} \overline{\hat{\boldsymbol{e}}_i} = \boldsymbol{0} .
\tag{9.2.7}
$$

Here I have used that the group average of the step size at each step is the same, $\langle \Delta x_i \rangle = \lambda$, and the fact that $\overline{\hat{\boldsymbol{e}}_i} = \boldsymbol{0}$. In component form this means:

$$
\overline{x(N)} = \overline{y(N)} = \overline{z(N)} = 0 .
\tag{9.2.8}
$$

The group spreads out in three dimensions. As a result, a calculation similar to the one that leads to Eqn. (9.1.7) in the one-dimensional case can now be performed. For a given random walker one has after N steps, using $\hat{\boldsymbol{e}}_i \cdot \hat{\boldsymbol{e}}_i = 1$:

$$
\begin{aligned}
|\boldsymbol{x}(N)|^2 &= \Delta x_1^2 + \Delta x_2^2 + \cdots + \Delta x_N^2 \\
&\quad + \sum_{i \neq j} \Delta x_i \Delta x_j \, (\hat{\boldsymbol{e}}_i \cdot \hat{\boldsymbol{e}}_j) .
\end{aligned}
\tag{9.2.9}
$$

When we take the group average, the first N terms in this sum contribute equally since the mean value of the square of the step sizes is $\langle \Delta x_i^2 \rangle = 2\lambda^2$ for all $1 \leq i \leq N$. All $N(N-1)$ terms in the second group, involving different steps, vanish according to the averaging rule (9.2.3) for the unit vectors. One therefore finds:

$$
\overline{|\boldsymbol{x}(N)|^2} = N \left\langle \Delta x^2 \right\rangle = 2N\lambda^2 .
\tag{9.2.10}
$$

The dispersion of the group in a given direction, say the x-direction, follows in an analogous fashion. From

$$
\begin{aligned}
x^2(N) \;=\;& \Delta x_1^2\,\hat{e}_{1x}^2 + \Delta x_2^2\,\hat{e}_{2x}^2 + \cdots + \Delta x_N^2\,\hat{e}_{Nx}^2 \\[2mm]
& + \sum_{i\neq j} \Delta x_i \Delta x_j\,(\hat{e}_{ix}\,\hat{e}_{jx}) \;,
\end{aligned}
\tag{9.2.11}
$$

one finds from (9.2.2) that the group average of $x^2(N)$ equals:

$$
\overline{x^2(N)} = \frac{N}{3}\left\langle \Delta x^2 \right\rangle = \frac{2N\lambda^2}{3}\;.
\tag{9.2.12}
$$

The dispersion in the y and z direction can be calculated in a similar manner, and one has

$$
\overline{x^2(N)} = \overline{y^2(N)} = \overline{z^2(N)} = \frac{2N\lambda^2}{3}\;.
\tag{9.2.13}
$$

It is also easily checked using (9.2.3) that cross terms like $\overline{xy}$ or $\overline{xz}$ vanish. For instance:

$$
\overline{x(N)\,y(N)} = 2\lambda^2 \sum_{i=1}^{N} \overline{\hat{e}_{xi}\,\hat{e}_{yi}} = 0\;.
\tag{9.2.14}
$$

The typical distance traveled still satisfies a relation like Eqn. (9.1.8):

$$
d(N) \equiv \sqrt{\overline{x^2} + \overline{y^2} + \overline{z^2}} = \sqrt{N\left\langle \Delta x^2 \right\rangle} = \sqrt{2N}\,\lambda\;.
\tag{9.2.15}
$$

The only difference is that $\left\langle \Delta x^2 \right\rangle = 2\lambda^2$ in this case, whereas in the one-dimensional random walk with a fixed step size of the previous Section one has $\Delta x^2 = \lambda^2$ at each step.

9.2.1 The diffusion tensor

In three dimensions one can define a *diffusion tensor*. Formally it is a tensor with components[4]

$$D_{ij} \equiv \frac{\overline{x_i x_j}}{2t} \ .$$
(9.2.16)

Here (for Cartesian coordinates) one has $x_1 = x$, $x_2 = y$ and $x_3 = z$. Note that this tensor is symmetric: $D_{ij} = D_{ji}$. Now assume that the random walk has a step frequency $\nu = 1/\Delta t_{\rm c}$. In that case one has $N \sim t/\Delta t_{\rm c}$, and the results of the previous Section (Eqn. 9.2.13) show that for isotropic diffusion one must have:

$$D_{xx} = D_{yy} = D_{zz} = \frac{\nu \lambda^2}{3} \ , \quad D_{xy} = D_{xz} = D_{yz} = 0 \ .$$
(9.2.17)

One concludes that for isotropic diffusion the diffusion tensor takes a simple (diagonal) form:

$$\boldsymbol{D} = \frac{\nu \lambda^2}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ .$$
(9.2.18)

This means that isotropic diffusion can be characterized with a *single* (scalar) diffusion coefficient,

$$\mathcal{D} \equiv \frac{\nu \lambda^2}{3} \ ,$$
(9.2.19)

with

$$\overline{x^2}(t) = \overline{y^2}(t) = \overline{z^2}(t) = 2\mathcal{D}t \ ,$$
(9.2.20)

and by a diffusion tensor that is simply $\boldsymbol{D} = \mathcal{D} \, \boldsymbol{I}$, with $\boldsymbol{I} = \mathrm{diag}(1 \, , \, 1 \, , \, 1)$ the unit tensor in three dimensions.

---

[4]From this point onwards, roman indices $(i, \, j, \, \cdots)$ are used to denote vector components, **not** step numbers.

The typical distance from the origin of a random walker after a time t equals

$$\overline{|\boldsymbol{x}^2|}(t) = 6\mathcal{D}t \tag{9.2.21}$$

for an isotropic random walk in three dimensions.

Not all diffusion processes are isotropic. Sometimes there are preferred directions. For instance, when charged particles diffuse in a magnetized medium, the magnetic field influences particle motion through the Lorentz force. This force (which is proportional to $\boldsymbol{v} \times \boldsymbol{B}$, with $\boldsymbol{v}$ the particle velocity and $\boldsymbol{B}$ the magnetic field) restricts particle motion in the plane perpendicular to the magnetic field, while leaving the motion along the field unaffected. Without collisions, the particles would 'slide' along the field with velocity $v_\parallel = \boldsymbol{v} \cdot \boldsymbol{B}/|\boldsymbol{B}|$, while gyrating in a circular orbit in the plane perpendicular to the magnetic field. In this case, the diffusion is very anisotropic. As a result, the diffusion tensor has a complicated form. The diffusion along the magnetic field proceeds much faster than diffusion in the plane perpendicular to the field, at least as long as the collision frequency ν is so small that particles make at least one full gyration around the magnetic field between two collisions.

9.3 The diffusion equation

Whenever the carrier particles of some quantity perform a random walk, that quantity obeys a diffusion equation. As a specific example, I will consider diffusion in space, making use of the picture sketched above. However, the type of equation that will be derived for spatial diffusion has a characteristic form that applies to *any* diffusion process.

Let us assume that some quantity $\mathcal{C}$ is carried by particles (molecules or atoms) with a number density $n_c(\boldsymbol{x}, t)$. Due to collisions between molecules with other molecules or atoms in the gas it is diffusively transported in a fluid. For simplicity I will assume that the fluid itself is at rest, so that the mean flow velocity vanishes: $\boldsymbol{V} = 0$. The density n_c can represent anything: the concentration of some pollutant in a pool of water, or the density of photons that are propagating diffusively through a scattering medium like the interior of the Sun.

Let us assume that the diffusive propagation is characterized by a mean-free-path (average step size) λ, and that the 'random walkers' that carry (transport) this quantity have a *randomly* directed velocity vector $\boldsymbol{\sigma}$ of magnitude $|\boldsymbol{\sigma}| = \sigma$. If the collision frequency is $\nu = 1/\Delta t_c$, with Δt_c the mean time between collisions, the mean free path λ and the magnitude of the random velocity σ are related by

$$\lambda = \sigma \, \Delta t_c = \frac{\sigma}{\nu} \, . \tag{9.3.1}$$

This random velocity could be the thermal velocity of the atoms in a fluid, and equals the velocity of light in the case of diffusing photons.

Let us assume that the quantity $n_c(\boldsymbol{x}, t)$ varies only in the $x-$direction, but not in the $y-$ and $z-$ directions:

$$n_c(\boldsymbol{x}, t) = n_c(x, t) \tag{9.3.2}$$

Consider a flat surface in the $y - z$ plane, which is located at the position $x = x_0$. If no carrier particles are destroyed, the density n_c must satisfy a continuity equation of the form

$$\frac{\partial n_c}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{F} = 0 \tag{9.3.3}$$

Here $\boldsymbol{F}$ is the flux of the carrier particles, a quantity that is to be calculated shortly.

Let us consider the flux component F_x, the component of the flux that is directed perpendicular to our surface (see the figure below). The number of particles ΔN_c that passes the surface *per unit area* in a time interval Δt is (by definition) related to the flux F_x by

$$\Delta N_c = F_x \, \Delta t \ . \tag{9.3.4}$$

On average, a particles arriving at the surface (i.e. at $x = x_0$) within one collision time, with the velocity at an inclination angle $0 \le i \le 360^0$ with respect to the normal to that surface, must have originated at a position $x \le x_m$, with

$$x_m = x_0 - \Delta x \, \cos i \ . \tag{9.3.5}$$

Here Δx is the size of the last step in the random walk. The quantity $|x_m|$ is the maximum distance from the site of the last collision to the surface, measured along the normal. Of course not all particles that cross the surface in a collision time have originated from this maximum distance: *on average*, the normal distance from the site of the last collision and the surface equals for an inclination angle i:

$$|x - x_0| = \tfrac{1}{2}\Delta x \, \cos i \ . \tag{9.3.6}$$

The (infinitesimal) number of molecules that reaches the surface per unit area with a crossing angle in the interval i, $i + \mathrm{d}i$, and assuming a step size Δx, is equal to the column density of all particles with $x - x_0 \le \Delta x \, \cos i$ times the fraction of particles $\mathrm{d}n_c(i)/n_c$ that have the inclination angle in the interval $\mathrm{d}i$:

$$
\begin{aligned}
\mathrm{d}N_c &= \Delta x \cos i \times n_c(x_0 - \tfrac{1}{2}\Delta x \, \cos i) \times \frac{\mathrm{d}n_c(i)}{n_c} \\[2mm]
&\approx \left[\Delta x \cos i \, n_c(x_0) - \tfrac{1}{2}\Delta x^2 \cos^2 i \left(\frac{\partial n_c}{\partial x} \right)_{x=x_0} \right] \frac{\mathrm{d}n_c(i)}{n_c} \ .
\end{aligned}
\tag{9.3.7}
$$

Note that I have assumed that the average density n_c in the column corresponds to the value at the *mean* starting point of the last step in the random walk. Particles reaching the surface from the left side (see the figure below) have $0 \le i < 90^o$, while particles reaching the surface from the other side have $90^o \le i \le 180^o$

The *net* number of particles crossing a unit area on the surface at $x = x_0$ follows from two averages: one average over all possible inclination angles $0 \leq i \leq 2\pi$, *and* an average over all possible step sizes Δx, with the step size probability distribution given by Eqn. (9.2.4). We first perform the average over all possible arrival directions. Assuming that the particles have an isotropic velocity distribution where all values of i are equally probable, the fraction of particles residing in the interval $(i, i + \mathrm{d}i)$ is simply the fraction of the solid angle 4π that corresponds to this interval:

$$\frac{\mathrm{d}n_{\mathrm{c}}(i)}{n_{\mathrm{c}}} = \frac{2\pi \sin i \, \mathrm{d}i}{4\pi} = -\tfrac{1}{2} \, \mathrm{d}\cos i \ . \tag{9.3.8}$$

The average over all possible arrival directions (which is represented by the notation $\langle\langle \cdots \rangle\rangle$) then reduces to a simple integration over $\cos i$ between $\cos i = -1$ and $\cos i = +1$:

$$\langle\langle \cdots \rangle\rangle = \tfrac{1}{2} \int_{-1}^{+1} (\cdots) \, \mathrm{d}\cos i \ . \tag{9.3.9}$$

It is easy to show[5] that $\langle\langle \cos i \rangle\rangle = 0$ and $\langle\langle \cos^2 i \rangle\rangle = 1/3$. One sees that the first term in Eqn. (9.3.7) gives no net contribution, while the second term averages to:

$$\mathrm{d}N_{\mathrm{c}} = -\frac{\Delta x^2}{6} \left(\frac{\partial n_{\mathrm{c}}}{\partial x} \right) \ . \tag{9.3.10}$$

If we now average over all possible step sizes, using from (9.2.5) the relation $\langle \Delta x^2 \rangle = 2\lambda^2$, one finds:

$$\langle \mathrm{d}N_{\mathrm{c}} \rangle = -\frac{\langle \Delta x^2 \rangle}{6} \left(\frac{\partial n_{\mathrm{c}}}{\partial x} \right) = -\frac{\lambda^2}{3} \left(\frac{\partial n_{\mathrm{c}}}{\partial x} \right) \ . \tag{9.3.11}$$

If the collision time is $\Delta t_{\mathrm{c}} = 1/\nu$, one must have the following relation between the diffusive flux and the number of particles that cross in a collision time:

$$\langle \mathrm{d}N_{\mathrm{c}} \rangle = F_x \, \Delta t_{\mathrm{c}} = -\frac{\lambda^2}{3} \left(\frac{\partial n_{\mathrm{c}}}{\partial x} \right) \ . \tag{9.3.12}$$

---

[5]One can derive this result in another way. In this example we have chosen our axes so that $\cos i = \hat{e}_x$, the x-component of the unit vector $\hat{e}$ along a given step in the random walk. For an isotropic random walk we have $\overline{\hat{e}_x} = 0$ and $\overline{\hat{e}_x^2} = 1/3$, see Eqn. (9.2.2) of the previous section.

This relation immediately gives the diffusive flux:

$$F_x = -\frac{\nu \lambda^2}{3}\left(\frac{\partial n_{\mathrm{c}}}{\partial x}\right) = -\mathcal{D}\left(\frac{\partial n_{\mathrm{c}}}{\partial x}\right)$$

(9.3.13)

We conclude that in the diffusion approximation the net flux is proportional to the first derivative of the concentration of the diffusing quantity, with the diffusion coefficient $\mathcal{D}$ as the constant of proportionality.

In the more general three-dimensional case, where n_{c} also varies in the $y-$ and in the z-direction, the diffusive particle flux in these directions can be calculated in exactly the same way, provided one assumes isotropic diffusion.

The conclusion therefore is as follows: the diffusive flux of some quantity is proportional to the 'concentration gradient'. For isotropic diffusion in three dimensions it is given by

$$\boldsymbol{F} = -\mathcal{D}\,\boldsymbol{\nabla} n_{\mathrm{c}}\ .$$

(9.3.14)

The scalar diffusion coefficient is defined as (see Eqn 9.3.1)

$$\mathcal{D} = \frac{\nu \lambda^2}{3} = \frac{\sigma \lambda}{3}\ ,$$

(9.3.15)

and contains the information about the properties of the random walk that the particles perform. Substituting definition (9.3.14) into the continuity equation, one finds the diffusion equation in three dimensions:

$$\frac{\partial n_{\mathrm{c}}}{\partial t} = \boldsymbol{\nabla}\cdot\left(\mathcal{D}\,\boldsymbol{\nabla} n_{\mathrm{c}}\right)\ .$$

(9.3.16)

This equation is also known as the *heat conduction equation*: heat in a conducting medium[6] transports diffusively, and the temperature T satisfies an equation of the form (9.3.16). If the diffusion coefficient is uniform ($\boldsymbol{\nabla}\mathcal{D} = 0$), this equation reduces to the simpler form:

$$\frac{\partial n_{\mathrm{c}}}{\partial t} = \mathcal{D}\,\nabla^2 n_{\mathrm{c}}\ .$$

(9.3.17)

---

[6]For instance: in a metal rod the heat is carried by conduction electrons that collide with the atoms.
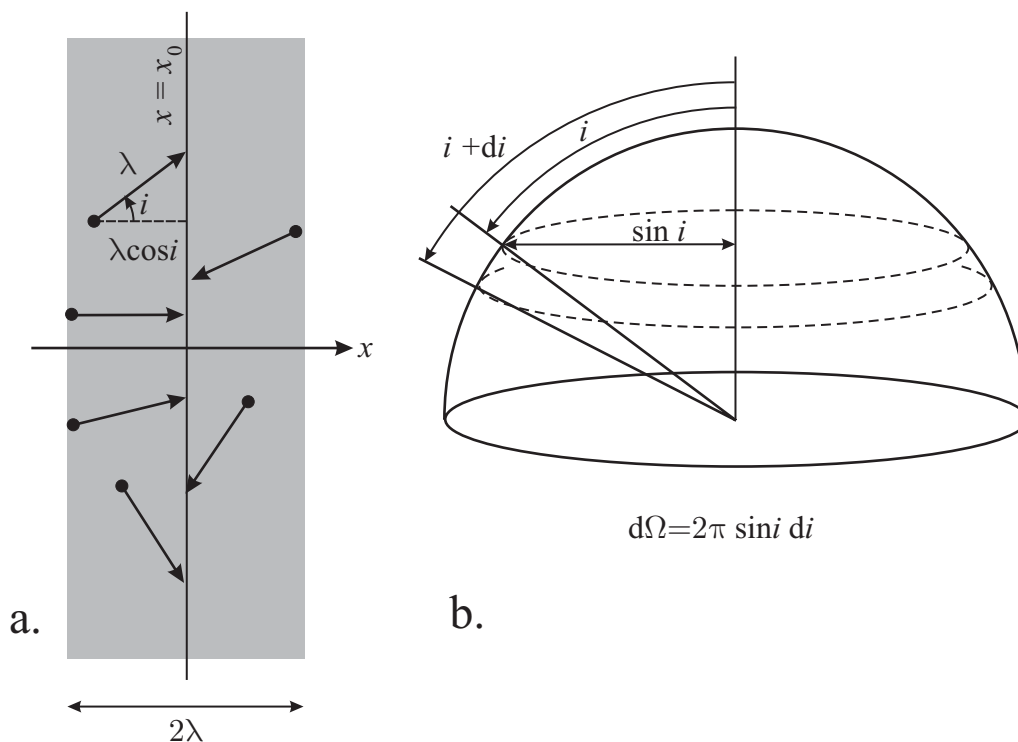
Figure 9.3: *The diffusive transport of some quantity (figure a). Only particles originating from the grey region of size 2λ around the surface $x = x_0$, with λ the mean free path, can carry fluid properties to the surface. If the inclination angle between the velocity of the particle and the normal of the surface equals i, the particle carries information about the properties of the fluid at a distance $\Delta x = \lambda \cos i$ from the surface.*

If distribution of the random ('thermal') velocities of the particles is isotropic, the number of particles incident on the surface at in some inclination angle interval i, $i + \mathrm{d}i$ is proportional to the solid angle corresponding to that interval (figure b). This solid angle is equal to the area on a unit sphere that lies between a cone with opening angle i and a cone with opening angle $i + \mathrm{d}i$. This area equals $\mathrm{d}\Omega = 2\pi \, \sin i \, \mathrm{d}i$.

9.3.1   Fundamental solutions of the diffusion equation

Consider the diffusion equation in one dimension,

$$\frac{\partial n_{\rm c}(x\,,\,t)}{\partial t} = \mathcal{D}\,\frac{\partial^2 n_{\rm c}(x\,,\,t)}{\partial x^2}\;. \tag{9.3.18}$$

Let us assume that at $t = 0$ the quantity is concentrated in a sharp spike at $x = 0$. We represent this spike by a *Dirac δ-function*[7] for the concentration $n_{\rm c}$:

$$n_{\rm c}(x\,,\,0) = \mathcal{C}_0\,\delta(x)\;. \tag{9.3.19}$$

With this initial condition the solution of the diffusion equation takes the following form for $t > 0$:

$$n_{\rm c}(x\,,\,t) = \frac{\mathcal{C}_0}{\sqrt{4\pi\mathcal{D}t}}\,\exp\left(-\frac{x^2}{4\mathcal{D}t}\right)\;. \tag{9.3.20}$$

The solution has a Gaussian shape, with a width $|\Delta x|$ that increases in time as

$$|\Delta x| \sim \sqrt{2\mathcal{D}t}\;. \tag{9.3.21}$$

The maximum of the density $n_{\rm c}(x\,,\,t)$ is always at $x = 0$. This clearly shows how this fundamental solution reflects the basic properties of the underlying random walk process.

The figure below shows a numerical simulation of diffusion in one dimension, together with the analytical solution (9.3.20). One can show that this solution satisfies the condition

$$\int_{-\infty}^{+\infty} {\rm d}x\, n_{\rm c}(x\,,\,t) = \mathcal{C}_0 \tag{9.3.22}$$

for all t. This reflects the fact the none of the diffusing quantity is lost: diffusion merely spreads it out over a larger and larger volume.

---

[7]e.g. Arfken & Weber, *Mathematical Methods for Physicists*, Ch. Sixth Ed., Ch. 1.15
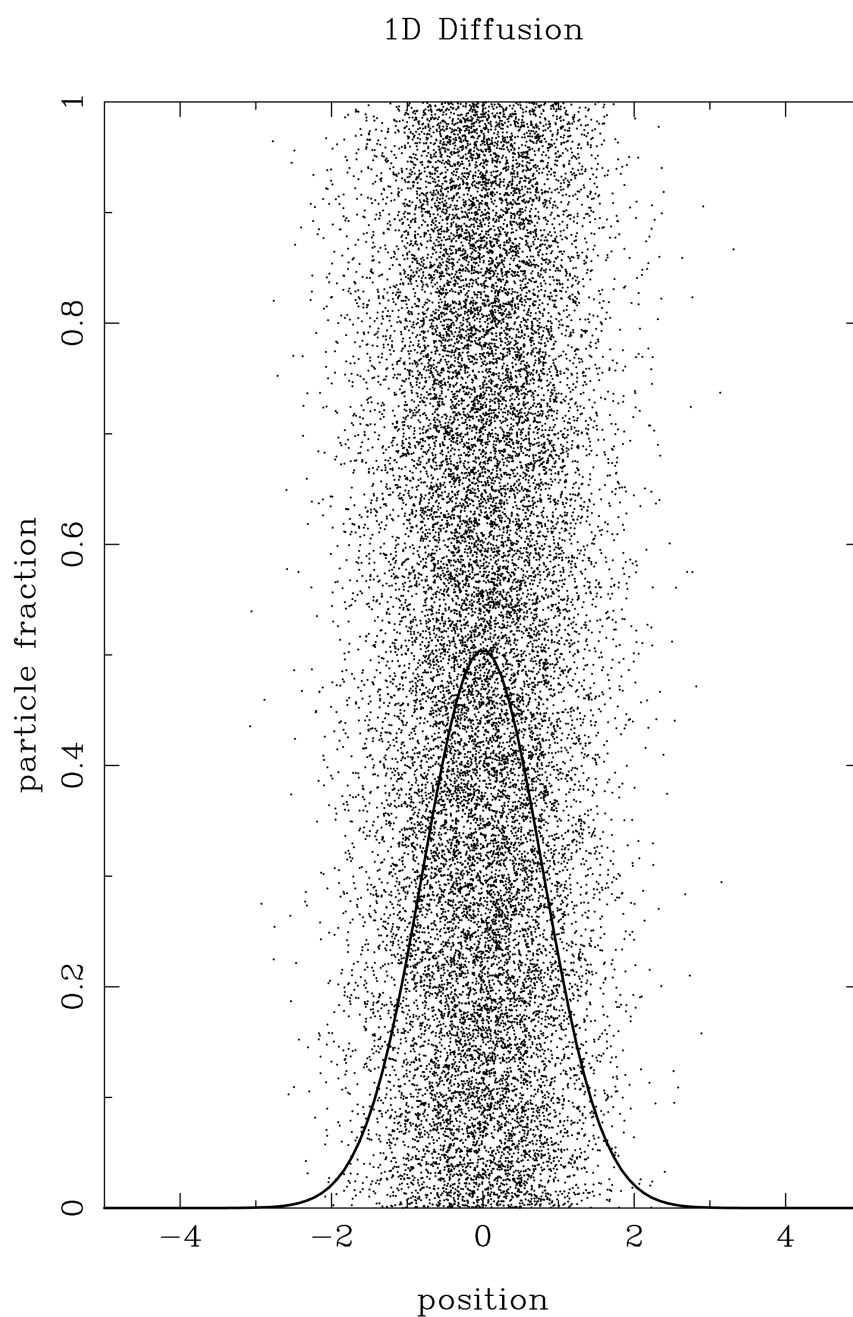
Figure 9.4: *A numerical simulation of a one-dimensional random walk. All particles started at $x = 0$, and have a random (but fixed) position y. The figure shows their position in the $x - y$ plane, together with the analytical solution for their density distribution, expressed in terms of the particle fraction: the number of particles at x divided by the total number of particles.*

For isotropic diffusion in three dimensions one can immediately guess the fundamental solution if the concentration at $t = 0$ satisfies

$$n_{\mathrm{c}}(\boldsymbol{x} , 0) = \mathcal{C}_0 \, \delta(x) \, \delta(y) \, \delta(z) . \qquad (9.3.23)$$

In that case one has:

$$\begin{aligned} n_{\mathrm{c}}(\boldsymbol{x} , t) &= \frac{\mathcal{C}_0}{(4\pi\mathcal{D}t)^{3/2}} \exp\left(-\frac{x^2 + y^2 + z^2}{4\mathcal{D}t}\right) \\[2mm] &= \frac{\mathcal{C}_0}{(4\pi\mathcal{D}t)^{3/2}} \exp\left(-\frac{r^2}{4\mathcal{D}t}\right) , \end{aligned} \qquad (9.3.24)$$

with $r^2 = x^2 + y^2 + z^2$. Apart from the normalization, this three-dimensional solution is simply the product of the three fundamental solutions for diffusion in one dimension: one Gaussian for the x-direction, one for the y-direction and one for the z-direction. This solution reflects the fact that diffusion proceeds independently in the three directions, but with the same diffusion coefficient, c.f. Eqn. (9.2.20). The three-dimensional analogue of (9.3.22) reads

$$\int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y \int_{-\infty}^{+\infty} \mathrm{d}z \, n_{\mathrm{c}}(\boldsymbol{x} , t) = \mathcal{C}_0 . \qquad (9.3.25)$$

Nothing of the diffusing quantity is lost, it is only redistributed in space!

9.3.2 Effects of a mean flow: the advection-diffusion equation

If one adds the effects of advection, where the particles are dragged along passively by a fluid that moves with a with velocity $\boldsymbol{V}(\boldsymbol{x}\,,\,t)$ in the laboratory frame, things become more complicated. One then has to add to the particle flux an extra term equal to

$$\boldsymbol{F}_{\mathrm{adv}}(\boldsymbol{x}\,,\,t) = \boldsymbol{V}(\boldsymbol{x}\,,\,t)\,n_{\mathrm{c}}(\boldsymbol{x}\,,\,t)\,. \tag{9.3.26}$$

This term described the advection of particles by the bulk flow. If the diffusion with respect to the fluid is isotropic, the total particle flux in the laboratory frame equals

$$\boldsymbol{F}(\boldsymbol{x}\,,\,t) = \underbrace{\boldsymbol{V}\,n_{\mathrm{c}}}_{\text{advection}} - \underbrace{\mathcal{D}\,\boldsymbol{\nabla}n_{\mathrm{c}}}_{\text{diffusion}}\,. \tag{9.3.27}$$

The equation resulting from Eqn. (9.3.3) which describes the changes in the concentration $C(\boldsymbol{x}\,,\,t)$ now becomes of the *advection-diffusion* type:

$$\frac{\partial n_{\mathrm{c}}}{\partial t} + \boldsymbol{\nabla}\cdot(\boldsymbol{V}\,n_{\mathrm{c}}) = \boldsymbol{\nabla}\cdot(\mathcal{D}\boldsymbol{\nabla}n_{\mathrm{c}})\,. \tag{9.3.28}$$

More general (and complicated) relations apply when diffusion proceeds at a different rate in different directions. In that case the diffusive contribution to the flux $\boldsymbol{F}$ must be defined in terms of a diffusion *tensor* $\mathbf{D}$ that is no longer diagonal (that is: of the form $\mathbf{D} = \mathcal{D}\,\mathbf{I}$):

$$\boldsymbol{F}_{\mathrm{diff}} = -\mathbf{D}\cdot\boldsymbol{\nabla}n_{\mathrm{c}}\,. \tag{9.3.29}$$

In component form, with the tensor components of $\mathbf{D}$ denoted by D_{ij}:

$$(F_i)_{\mathrm{diff}} = -D_{ij}\left(\frac{\partial n_{\mathrm{c}}}{\partial x_j}\right)\,. \tag{9.3.30}$$

The total flux describes the combined effect of advection and diffusion and equals:

$$\boldsymbol{F} = \boldsymbol{F}_{\mathrm{adv}} + \boldsymbol{F}_{\mathrm{diff}} = \boldsymbol{V}\,n_{\mathrm{c}} - (\mathbf{D}\cdot\boldsymbol{\nabla})\,n_{\mathrm{c}}\,. \tag{9.3.31}$$

Particle conservation reads

$$\frac{\partial n_{\mathrm{c}}}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{F} = 0 \ .$$
(9.3.32)

Substituting for the flux from (9.3.32), and re-arranging terms, one finds the general form of the *advection-diffusion equation:*

$$\frac{\partial n_{\mathrm{c}}}{\partial t} + \boldsymbol{\nabla} \cdot (\boldsymbol{V} \, n_{\mathrm{c}}) = \boldsymbol{\nabla} \cdot (\mathbf{D} \cdot \boldsymbol{\nabla} n_{\mathrm{c}}) \ .$$
(9.3.33)

In component notation, using Cartesian coordinates $x_1 = x$, $x_2 = y$ and $x_3 = z$, this is

$$\frac{\partial n_{\mathrm{c}}}{\partial t} + \frac{\partial}{\partial x_i} (V_i \, n_{\mathrm{c}}) = \frac{\partial}{\partial x_i} \left(D_{ij} \, \frac{\partial n_{\mathrm{c}}}{\partial x_j} \right) \ .$$
(9.3.34)

9.4 Application I: diffusive radiation transport

When a medium is optically thick (opaque), photons will propagate diffusively. The mean free path of the photons is determined by the *opacity* of the medium. If the photons are scattered (or absorbed) by particles with density n, mass m and a cross-section σ, the photon mean-free-path is equal to

$$\ell_{\mathrm{ph}} = \frac{1}{n\sigma} = \frac{1}{\kappa\rho} \ . \tag{9.4.1}$$

This quantity plays the same role as λ in the previous Sections. One defines the opacity κ (the cross-section per unit mass) as

$$\kappa = \sigma/m \ . \tag{9.4.2}$$

The probability that a photon travels a path length $s = ct$ without being scattered or absorbed, the 'survival probability', decays with increasing s as

$$\mathcal{P}_{\mathrm{surv}}(s) = 1 - \underbrace{\int_0^s \mathrm{d}(\Delta x) \, \frac{\exp(-\Delta x/\ell_{\mathrm{ph}})}{\ell_{\mathrm{ph}}}}_{\text{probability for scattering}} = e^{-s/\ell_{\mathrm{ph}}} \ . \tag{9.4.3}$$

We will first concentrate on the case where photons are simply scattered, excluding absorption processes where photons are actually destroyed.

Since photons move with the speed of light, the photon diffusion coefficient, given a photon mean free path $\lambda = \ell_{\mathrm{ph}}$ for scattering, equals

$$\mathcal{D}_{\mathrm{ph}} = \frac{c\ell_{\mathrm{ph}}}{3} = \frac{c}{3\kappa\rho} \ . \tag{9.4.4}$$

In principle, these quantities depend on the frequency of the photons involved, and the properties (e.g. density and temperature) of the medium. If the photons are in *thermal equilibrium* with the medium, with both the medium and the photons in a thermal distribution[8] with temperature T, the energy density of the radiation equals

$$U_{\mathrm{rad}} = a_{\mathrm{r}} \, T^4 \ . \tag{9.4.5}$$

Here $a_{\mathrm{r}} = 4\sigma_{\mathrm{sb}}/c$ is the radiation constant.

---

[8]For photons the *Planck Distribution*, and for the gas the *Maxwell-Boltzmann distribution*.

If there are temperature gradients, the diffusion of radiation will try to erase these gradients. Photons from hotter regions deposit extra energy, while photons arriving from a colder region have an energy deficit compared with photons originating locally. The resulting net energy flux carried by the radiation is approximated by the diffusion form derived in the previous Section:

$$\boldsymbol{F}_{\rm rad} = -\mathcal{D}_{\rm ph}\,\boldsymbol{\nabla} U_{\rm rad} = -\frac{4ca_{\rm r}T^3}{3\kappa\rho}\,\boldsymbol{\nabla} T \; . \tag{9.4.6}$$

This energy flux will lead to a net loss of energy from hotter regions to cooler regions: a form of *thermal conduction*. The radiation energy density then satisfies the transport equation

$$\frac{\partial U_{\rm rad}}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{F}_{\rm rad} = 0 \; . \tag{9.4.7}$$

Using the above definitions for $U_{\rm rad}$ and $\boldsymbol{F}_{\rm rad}$, this equation for the photon energy density can be written as a diffusion equation for the *temperature*:

$$\frac{\partial T}{\partial t} = \boldsymbol{\nabla} \cdot \left(\mathcal{K}_{\rm rad}\,\boldsymbol{\nabla} T \right) \; . $$

$$\tag{9.4.8}$$

The *radiative thermal conduction coefficient* defined in this equation equals the photon diffusion coefficient:

$$\mathcal{K}_{\rm rad} = \mathcal{D}_{\rm ph}(T) = \frac{c}{3\kappa(\rho, T)\,\rho} \; . \tag{9.4.9}$$

This description simplifies the treatment of radiation transport and the interaction between matter and radiation considerably. For slow speeds ($|\boldsymbol{V}| \ll c$) one can include the advection of photons by a flow approximately by adding an advection term to the photon transport equation as sketched in the preceding section. However, the theory of radiation transfer in moving media is very complicated, and such simple recipes for radiative transfer fail surprisingly fast.

The diffusion approximation for radiative transport has a number of limitations which I will list briefly.

1. As already stated, this diffusion approximation is only valid if the medium is optically thick. Roughly speaking, this means that the size L of the scattering medium (a cloud, a star or an accretion disk) must satisfy the requirement

$$\tau \sim \frac{L}{\ell_{\mathrm{ph}}} \gg 1 \; . \qquad (9.4.10)$$

The quantity τ is called the *optical depth* in radiation transfer theory. This condition ensures that the photon is scattered many times before reaching the edge of the medium where it can escape freely.

2. This simple description glosses over the differences between *scattering* and *true absorption*. In scattering a photon changes some of its properties (propagation direction and/or frequency), but remains 'intact'. In absorption the photon is destroyed. In principle, the opacity κ contains a contribution of both these processes:

$$\kappa(\rho, T) = \kappa_{\mathrm{s}} + \kappa_{\mathrm{a}} \; . \qquad (9.4.11)$$

As long as one assumes thermal equilibrium, this distinction between scattering and absorption is not important. The reason is that in this case a photon with similar properties is emitted for each photon that is absorbed. However, this is only true in an *average* sense. If one wants to calculate the detailed spectrum of the radiation, one has to do a fully-fledged radiation transport calculation in which this distinction *is* important, and where the diffusion approximation need not be true at all wavelengths.

The details of more complicated descriptions of radiation transport can be found in the book by Rybicki & Lightman or the first book in the two-volume set of Shu[9].

In the case of an optically thin medium ($\tau \ll 1$), where the coupling between the photons and the medium in which they propagate is relatively weak, the emerging photons are not necessarily 'thermalized' into a black-body spectrum with a temperature equal to the temperature of the medium. In that case, all kinds of *non-thermal* radiation processes, such as synchrotron radiation or Comptonization, get a chance to put their imprint on the spectrum of the emerging radiation: the spectrum is no longer a Planck spectrum.

---

[9]G.B. Rybicki & A.P. Lightman, 1979: *Radiative Processes in Astrophysics*, Chapter 1.7 and 1.8, John Wiley & Sons;
F.H. Shu, 1991, *The Physics of Astrophysics*, Volume 1: Radiation, Ch. 2, University Science Books

9.5   Diffusion and viscosity

Diffusive transport is not limited to scalar quantities like the concentration of some contaminant, or the energy density (temperature) of a gas. Vector quantities, such as the momentum density $\boldsymbol{M} = \rho\,\boldsymbol{V}$ of a fluid, can be transported diffusively. Such transport must lead to momentum changes, and therefore to a force. In this case one speaks of *viscous forces*.

The (diffusive) transport of a vector quantity, like the momentum density of a fluid or gas, must be described in terms of a **tensor**. This requirement is already apparent from our discussion of the conservative version of the momentum equation in Chapter 3. The components of such a tensor $\mathbf{T}$ are defined as

$$T_{ij} \equiv \text{flux in the } i\text{-direction of the } j\text{-component of momentum density} \qquad (9.5.1)$$

The amount of momentum density (a vector!) crossing some oriented surface $\mathrm{d}\boldsymbol{O}$ in a time-interval Δt equals

$$\Delta \boldsymbol{M} = \Delta t \,(\mathrm{d}\boldsymbol{O} \cdot \mathbf{T})\,, \qquad (9.5.2)$$

or in component form:

$$\Delta M_i = \Delta t \,(\mathrm{d}O_j \, T_{ji})\,. \qquad (9.5.3)$$

Now consider an infinitesimally small (closed) volume $\mathcal{V}$. The net amount of momentum $\Delta \boldsymbol{p}$, entering or leaving the volume across it's exterior surface in a time span Δt, equals

$$\Delta \boldsymbol{p} = \Delta \left(\int \mathrm{d}\mathcal{V}\, \boldsymbol{M} \right) = -\Delta t \left(\oint_{\partial \mathcal{V}} \mathrm{d}\boldsymbol{O} \cdot \mathbf{T} \right)\,. \qquad (9.5.4)$$

The extra minus sign introduced in this expression reflects the fact that (by convention) the normal of the oriented surface surrounding the volume is always pointed *outwards*: when $\mathrm{d}\boldsymbol{O} \cdot \mathbf{T} > 0$ there is net momentum *leaving* the volume! Using Stokes' theorem (Eqn. 3.1.9) one can rewrite the surface integral into a volume integral,

$$\oint_{\partial \mathcal{V}} \mathrm{d}\boldsymbol{O} \cdot \mathbf{T} = \int \mathrm{d}\mathcal{V}\,(\boldsymbol{\nabla} \cdot \mathbf{T})\,. \qquad (9.5.5)$$

The 'force' on the volume element, which is simply the change in total momentum per unit time, equals:

$$\boldsymbol{F} = \frac{\Delta \boldsymbol{p}}{\Delta t} = -\int \mathrm{d}\mathcal{V} \left(\boldsymbol{\nabla} \cdot \mathbf{T} \right) . \tag{9.5.6}$$

This expression gives the force on the whole (infinitesimal) volume as a volume integral. That implies that the force density, i.e. the force per unit volume that is associated with the tensor $\mathbf{T}$, must equal:

$$\boldsymbol{f} = -\boldsymbol{\nabla} \cdot \mathbf{T} . \tag{9.5.7}$$

This relation tells us how to calculate the force from the (tensor-)divergence of the momentum flux tensor. It can be directly applied to diffusive momentum transport.

9.5.1 A simple example: shear flow

As a simple example, consider a pure *shear flow* with straight stream lines. In a shear flow the fluid moves in a fixed direction along straight lines, but the fluid velocity changes in the direction perpendicular to the flow lines, see the figure below. We choose our coordinate system in such a way that the fluid moves along the y-axis, with a velocity that varies in the $x-$direction:

$$\boldsymbol{V} = V(x) \, \hat{\boldsymbol{y}} . \tag{9.5.8}$$

Such a shear flow may serve as a rough (local) approximation for the flow in a differentially rotating accretion disk. The momentum density of this fluid only has a y-component:

$$\boldsymbol{M} = \rho \boldsymbol{V} = (0 \, , \, \rho V(x) \, , \, 0) . \tag{9.5.9}$$

To keep things simple, I will assume that the mass density of the fluid is uniform.

Due to diffusion of the atoms or molecules in the fluid or gas, momentum is transported from one position to another. In this specific case, diffusion leads to a flux of the y-component of momentum $\left(M_y(x) = \rho V(x) \right.$ in the x-direction. This momentum flux is given by the standard expression for a diffusive flux:

$$T_{xy}^{\mathrm{visc}} = -\mathcal{D} \left(\frac{\partial M_y}{\partial x} \right) \sim -\rho \mathcal{D} \left(\frac{\partial V}{\partial x} \right) . \tag{9.5.10}$$
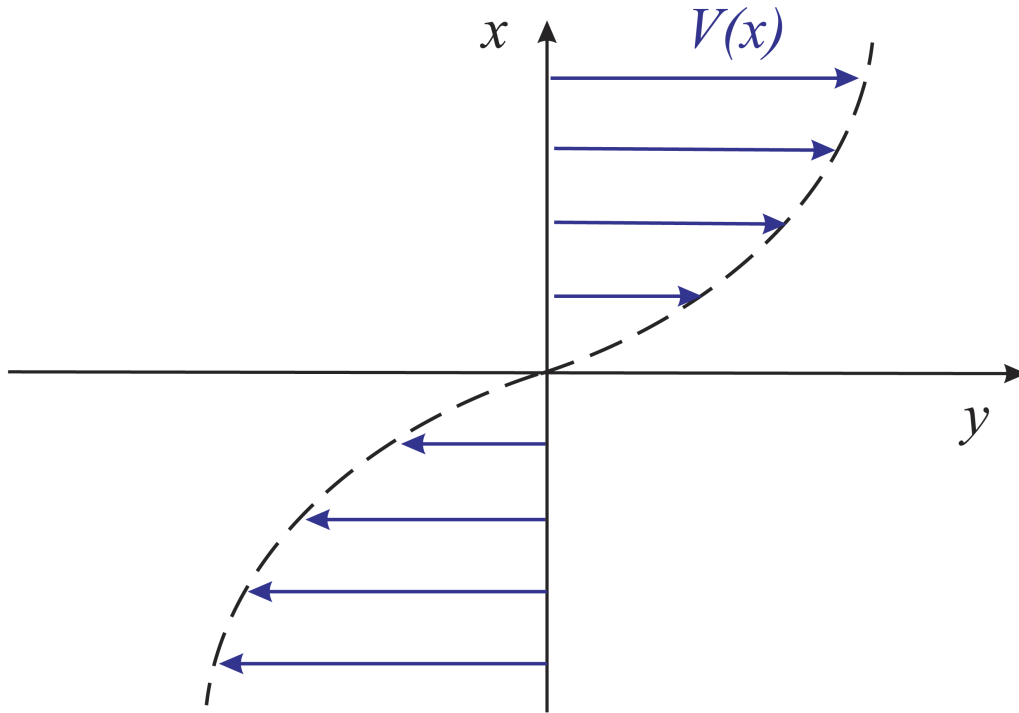
Figure 9.5: *A one-dimensional (linear) shear flow, where the flow velocity is along the y-axis, and the flow speed $V(x)$ varies in the direction perpendicular to the flow lines. In this case we have $V = 0$ at $x = 0$. The thermal motion of the atoms or molecules in the fluid will transport y-momentum in the x-direction, leading to a viscous flux $T_{xy}^{\mathrm{visc}} = -\eta \, (\partial V/\partial x)$, and a force density $\boldsymbol{f}^{\mathrm{visc}} = -\boldsymbol{\nabla} \cdot \mathbf{T}^{\mathrm{visc}} = \eta \, (\partial^2 V/\partial x^2)\hat{\boldsymbol{y}}$, with $\eta = \rho \mathcal{D}$ the coefficient of shear viscosity.*

This is the only non-vanishing of the diffusive momentum flux tensor for this flow. The negative divergence of T_{xy} gives the amount of $y-$momentum, deposited by the diffusing particles per unit time and unit volume, and therefore corresponds to a force density: the *viscous force*, which only has an y-component since there is only momentum in the y-direction:

$$f_y^{\mathrm{visc}} = -\frac{\partial T_{xy}^{\mathrm{visc}}}{\partial x} \sim \eta \left(\frac{\partial^2 V}{\partial x^2}\right) \qquad (9.5.11)$$

Here I have defined the *dynamic viscosity coefficient* η as

$$\eta \sim \rho \mathcal{D} \; . \qquad (9.5.12)$$

The viscous force is proportional to the second derivative of the velocity. It essentially gives the effect of friction between adjacent fluid layers.

One usually defines the *kinematic viscosity* ν of a fluid by:

$$\nu \equiv \frac{\eta}{\rho} = \tfrac{1}{3}\sigma\ell = \mathcal{D} \ . \tag{9.5.13}$$

This calculation shows, at least in this one-dimensional example, that the kinematic viscosity and the diffusion coefficient of the atoms or molecules involved are the same quantity.

9.5.2 The viscous stress tensor

The full expression for the viscosity tensor in a fully three-dimensional flow is somewhat more complicated. Just like in our simple example, the tensor $\mathbf{T}^{\text{visc}}$, the *viscous shear stress tensor*, must depend linearly on the velocity gradients in the fluid as it describes a diffusive flux of momentum. The viscous stress tensor must therefore be a linear function of the components of the *velocity gradient tensor*, which looks like

$$\boldsymbol{\nabla V} = \begin{pmatrix} \partial V_x/\partial x & \partial V_y/\partial x & \partial V_z/\partial x \\ \partial V_x/\partial y & \partial V_y/\partial y & \partial V_z/\partial y \\ \partial V_x/\partial z & \partial V_y/\partial z & \partial V_z/\partial z \end{pmatrix} \tag{9.5.14}$$

in cartesian coordinates. However, it turns out the simplest possible recipe, $\mathbf{T}^{\text{visc}} = \eta\,\boldsymbol{\nabla V}$, does not work in general, even though it gives the correct answer for the shear flow treated in the previous section.

A simple example shows why the simple recipe fails. Viscous stresses must vanish if a fluid that rotates uniformly, so that

$$\boldsymbol{V} = \boldsymbol{\Omega} \times \boldsymbol{x} \ , \tag{9.5.15}$$

with $\boldsymbol{\Omega}$ some constant angular velocity vector. In that case there is no friction: the rotational velocity in this case corresponds to *solid body rotation*, where the fluid elements at a different distance from the axis of rotation do do not move with respect to each other. This is most easily seen in the frame that co-rotates with $\boldsymbol{\Omega}$: in that *corotating frame* all fluid elements appear at rest! Therefore, there can be **no** viscous dissipation of momentum for solid body rotation .

If the rotation axis coincides with the $z-$ axis, the velocity for solid-body rotation is in the $x - y$-plane: it equals

$$V_x = -\Omega y \ , \ \ V_y = \Omega x \ , \ \ V_z = 0 \ . \tag{9.5.16}$$

The velocity gradient tensor for this case is

$$\boldsymbol{\nabla V} = \begin{pmatrix} 0 & \Omega & 0 \\ -\Omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{9.5.17}$$

The velocity gradient tensor does not vanish identically for solid-body rotation! There is no viscous force in this case, thus the viscous stress tensor can **not** be directly proportional to the velocity gradient tensor $\boldsymbol{\nabla V}$. However, the components of the *symmetrized* tensor,

$$\boldsymbol{\nabla V} + \boldsymbol{\nabla V}^\dagger \ , \tag{9.5.18}$$

do vanish identically for solid-body rotation. Here $\boldsymbol{\nabla V}^\dagger$ is the tensor one gets by interchanging the columns and rows of $\boldsymbol{\nabla V}$, i.e. by mirroring the velocity gradient tensor with respect to the diagonal. The components of this symmetrized tensor can be written in index notation as

$$\left(\boldsymbol{\nabla V} + \boldsymbol{\nabla V}^\dagger\right)_{ij} \equiv \frac{\partial V_j}{\partial x_i} + \frac{\partial V_i}{\partial x_j} \ . \tag{9.5.19}$$

Therefore, the components of this symmetrized tensor must appear in the general expression for the diffusive momentum flux and in the associated viscous force density.

There is, however, another possible contribution. If the fluid is isotropic so that it's properties are independent of direction, the only other tensor-quantity that can appear in the viscous stress tensor is the unit tensor $\boldsymbol{I}$, which has the components δ_{ij}. Since the viscous stress tensor is linear in the velocity gradients, the unit tensor has to appear in combination with the only scalar quantity that one can form which is linear in the velocity derivatives: the divergence $\boldsymbol{\nabla} \cdot \boldsymbol{V}$. Therefore, the viscous stress tensor in an isotropic fluid must take the general form

$$T_{ij}^{\text{visc}} = a \left(\frac{\partial V_j}{\partial x_i} + \frac{\partial V_i}{\partial x_j}\right) + b \left(\boldsymbol{\nabla} \cdot \boldsymbol{V}\right) \delta_{ij} \ . \tag{9.5.20}$$

Note that this expression involves **two** proportionality constants, a and b, yet to be determined! It is always possible to write (9.5.20) in a somewhat different form:

$$T_{ij}^{\text{visc}} = -\eta \left(\frac{\partial V_j}{\partial x_i} + \frac{\partial V_i}{\partial x_j} - \tfrac{2}{3} \left(\boldsymbol{\nabla} \cdot \boldsymbol{V} \right) \delta_{ij} \right) - \zeta \left(\boldsymbol{\nabla} \cdot \boldsymbol{V} \right) \delta_{ij} \ . \tag{9.5.21}$$

This expression involves two viscosity coefficients, the *shear viscosity* η and the *bulk viscosity* ζ. It is easily checked that the shear viscosity η is the same quantity as the viscosity coefficient calculated in the one-dimensional example of a shear flow in the previous Section. That pure shear flow is incompressible, with $\boldsymbol{\nabla} \cdot \boldsymbol{V} = 0$, and if one calculates T_{xy}^{visc} from recipe (9.5.21) for the shear flow (D.5.1), it is easily checked that one obtains the correct answer: $T_{xy}^{\text{visc}} = -\eta \left(\partial V_y / \partial x \right)$. The coefficient of bulk viscosity ζ is not so straightforwardly calculated in terms of the diffusivity of the flow. In most astrophysical applications one simply puts $\zeta = 0$.

According to the discussion at the beginning of this Chapter the viscous force density equals

$$\boldsymbol{f}^{\text{visc}} = -\boldsymbol{\nabla} \cdot \mathbf{T}^{\text{visc}} \ . \tag{9.5.22}$$

Still using cartesian coordinates, and substituting expression (9.5.21), one finds:

$$f_j^{\text{visc}} = -\frac{\partial T_{ij}^{\text{visc}}}{\partial x_i} = \eta \left(\frac{\partial^2 V_j}{\partial x_i \partial x_i} \right) + \left(\zeta + \tfrac{1}{3}\eta \right) \frac{\partial (\boldsymbol{\nabla} \cdot \boldsymbol{V})}{\partial x_j} \ . \tag{9.5.23}$$

This implies that the viscous force density can be written in vector form as

$$\boldsymbol{f}^{\text{visc}} = \eta \nabla^2 \boldsymbol{V} + \left(\zeta + \tfrac{1}{3}\eta \right) \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{V}) \ . \tag{9.5.24}$$

This vector expression is valid in *any* coordinate system by the principle of covariance.

If the flow is incompressible, so that $\boldsymbol{\nabla} \cdot \boldsymbol{V} = 0$, the viscous force density reduces to a more simple form:

$$\boldsymbol{f}^{\text{visc}} = \eta \nabla^2 \boldsymbol{V} \ . \tag{9.5.25}$$

The equation of motion for a fluid including viscous forces (but without gravity) becomes

$$\rho \left[\frac{\partial \boldsymbol{V}}{\partial t} + (\boldsymbol{V} \cdot \boldsymbol{\nabla}) \, \boldsymbol{V} \right] = -\boldsymbol{\nabla} P + \eta \nabla^2 \boldsymbol{V} + \left(\zeta + \tfrac{1}{3} \eta \right) \, \boldsymbol{\nabla} (\boldsymbol{\nabla} \cdot \boldsymbol{V}) \, . \qquad (9.5.26)$$

9.5.3   The Reynolds Number

For order-of-magnitude estimates it is useful to have a measure of the importance of viscous effects on a flow. I will assume that $\zeta = 0$. This measure is provided by a dimensionless number, the *Reynolds number*, which can be defined as:

$$\mathrm{Re} = \frac{\text{magnitude inertial force}}{\text{magnitude viscous force}} \sim \frac{|\rho \, (\boldsymbol{V} \cdot \boldsymbol{\nabla}) \boldsymbol{V}|}{|\eta \, \nabla^2 \boldsymbol{V}|} \, . \qquad (9.5.27)$$

Let U and L be the typical value of the velocity, and the gradient scale of the velocity field. In that case one may use the estimates

$$|(\boldsymbol{V} \cdot \boldsymbol{\nabla}) \boldsymbol{V}| \sim U^2/L \quad , \quad |\eta \nabla^2 \boldsymbol{V}| \sim \rho \nu U/L^2 \, . \qquad (9.5.28)$$

This implies that the Reynolds number has a typical magnitude

$$\boxed{\mathrm{Re} \sim \frac{UL}{\nu} \, .}$$

$$(9.5.29)$$

Here I have used that the shear viscosity satisfies $\eta = \rho \, \nu$. If the Reynolds number is large so that $\mathrm{Re} \gg 1$ one can neglect the viscous effects to lowest order. They become important when $\mathrm{Re} \leq 1$. The viscous forces will dominate the force balance in the flow if $\mathrm{Re} \ll 1$.

9.6   Viscous dissipation

Viscous forces do work on a flow. This work converts the kinetic energy of the flow, the energy of the *bulk motion*, into thermal energy: the kinetic energy of the thermal motion. If there are no losses, such as radiation losses, the internal friction provided by viscosity will heat the fluid.

The amount of viscous dissipation of the kinetic energy can be derived directly from the equation of motion (9.5.26). If one writes this equation as

$$\rho \left[\frac{\partial \boldsymbol{V}}{\partial t} + (\boldsymbol{V} \cdot \boldsymbol{\nabla}) \, \boldsymbol{V} \right] = -\boldsymbol{\nabla} P - \boldsymbol{\nabla} \cdot \mathbf{T}^{\mathrm{visc}} \, , \tag{9.6.1}$$

one immediately finds from the scalar product of this equation with $\boldsymbol{V}$, c.f. Eqn. (3.2.18):

$$\frac{\partial}{\partial t} \left(\frac{\rho V^2}{2} \right) + \boldsymbol{\nabla} \cdot \left(\rho \boldsymbol{V} \, \frac{V^2}{2} \right) = -(\boldsymbol{V} \cdot \boldsymbol{\nabla}) P - (\boldsymbol{\nabla} \cdot \mathbf{T}^{\mathrm{visc}}) \cdot \boldsymbol{V} \, . \tag{9.6.2}$$

The last term in this equation is the amount of work per unit volume done by viscous forces:

$$W^{\mathrm{visc}} \equiv \boldsymbol{f}^{\mathrm{visc}} \cdot \boldsymbol{V} = -(\boldsymbol{\nabla} \cdot \mathbf{T}^{\mathrm{visc}}) \cdot \boldsymbol{V} \tag{9.6.3}$$

We now use an identity valid for an arbitrary tensor $\mathbf{T}$ and vector $\boldsymbol{V}$:

$$\boldsymbol{\nabla} \cdot (\mathbf{T} \cdot \boldsymbol{V}) = (\boldsymbol{\nabla} \cdot \mathbf{T}) \cdot \boldsymbol{V} + \mathbf{T}^{\dagger} : \boldsymbol{\nabla} \boldsymbol{V} \, . \tag{9.6.4}$$

Here $\mathbf{T}^{\dagger}$ is the *transpose tensor* of $\mathbf{T}$ obtained from interchanging rows and columns so that

$$T_{ij}^{\dagger} = T_{ji} \, , \tag{9.6.5}$$

and the symbol ' **:** ' is a 'double contraction'[10], defined for two rank-2 tensors $\mathbf{T}$ and U as an operation which yields a scalar quantity defined as (remember the summation convention for double indices!)

---

[10]See the Mathematical Appendix on Internet

$$\mathbf{T} : \boldsymbol{U} = T_{ik}U_{ki} \ . \tag{9.6.6}$$

The above identity can be most simply proven in Cartesian coordinates where x_i denotes x for $i = 1$, y for $i = 2$ and z for $i = 3$. It is a simple consequence of the chain-rule for differentiation. In component form, using the summation convention, one has:

$$\begin{aligned}
\boldsymbol{\nabla} \cdot (\mathbf{T} \cdot \boldsymbol{V}) \equiv \frac{\partial}{\partial x_i} \left(T_{ik}V_k \right) &= \frac{\partial T_{ik}}{\partial x_i} \, V_k + T_{ik} \, \frac{\partial V_k}{\partial x_i} \\[2mm]
&= (\boldsymbol{\nabla} \cdot \mathbf{T})_k V_k + T_{ki}^{\dagger} \left(\boldsymbol{\nabla} \boldsymbol{V} \right)_{ik} \tag{9.6.7} \\[2mm]
&= (\boldsymbol{\nabla} \cdot \mathbf{T}) \cdot \boldsymbol{V} + \mathbf{T}^{\dagger} : \boldsymbol{\nabla} \boldsymbol{V} \ .
\end{aligned}$$

Since the viscous stress tensor is symmetric,

$$(\mathbf{T}^{\mathrm{visc}})^{\dagger} = \mathbf{T}^{\mathrm{visc}} \ , \tag{9.6.8}$$

this identity allows one to write the amount of work done by viscous forces as

$$W^{\mathrm{visc}} \;=\; - \underbrace{\boldsymbol{\nabla} \cdot (\mathbf{T}^{\mathrm{visc}} \cdot \boldsymbol{V})}_{\text{divergence of viscous energy flux}} \;+\; \underbrace{\mathbf{T}^{\mathrm{visc}} : \boldsymbol{\nabla} \boldsymbol{V}}_{\text{kinetic energy loss}}$$

$$\equiv \; -\boldsymbol{\nabla} \cdot \boldsymbol{F}^{\mathrm{visc}} - \mathcal{H}^{\mathrm{visc}} \ . \tag{9.6.9}$$

The first term on the left-hand side is the divergence of an energy flux that is associated with the diffusive transport of momentum:

$$\boldsymbol{F}^{\mathrm{visc}} \equiv \mathbf{T}^{\mathrm{visc}} \cdot \boldsymbol{V} \ . \tag{9.6.10}$$

This energy flux can **not** correspond to true dissipation. Like any flux, it simply redistributes energy! The second term, $\mathcal{H}^{\mathrm{visc}}$, corresponds to the net loss of kinetic energy per unit volume due to viscous dissipation. This term describes *true* dissipation. It is easily shown that $\mathcal{H}^{\mathrm{visc}}$ is a positive definite quantity, as it should be: dissipation is irreversible, and friction always leads to a loss of bulk kinetic energy.

Using (9.5.21) one finds that the amount of kinetic energy **lost** per unit time is

$$
\begin{aligned}
\mathcal{H}^{\text{visc}} &\equiv -\mathbf{T}^{\text{visc}} : \boldsymbol{\nabla} \boldsymbol{V} \\
\\
&= \eta \left(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i} \right) \frac{\partial V_i}{\partial x_j} + \left(\zeta - \tfrac{2}{3}\eta \right) (\boldsymbol{\nabla} \cdot \boldsymbol{V})^2 \quad\quad\quad (9.6.11) \\
\\
&= \tfrac{1}{2}\eta \left(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i} - \tfrac{2}{3} (\boldsymbol{\nabla} \cdot \boldsymbol{V}) \delta_{ij} \right)^2 + \zeta (\boldsymbol{\nabla} \cdot \boldsymbol{V})^2 \; .
\end{aligned}
$$

The last equality follows after some tensor algebra.

The viscous heating $\mathcal{H}^{\text{visc}}$ represents the heat generated by internal friction in the fluid or gas. This heating is an irreversible process, which leads to an increase of the entropy density s of a gas or fluid. The entropy increase in an ideal gas with entropy density

$$
s = c_{\text{v}} \ln \left(\frac{P}{\rho^\gamma} \right) \quad\quad\quad (9.6.12)
$$

is described by (see Eqn. 3.2.25):

$$
\rho c_{\text{v}} T \left(\frac{\partial}{\partial t} + \boldsymbol{V} \cdot \boldsymbol{\nabla} \right) \ln \left(\frac{P}{\rho^\gamma} \right) = \mathcal{H} \; . \quad\quad\quad (9.6.13)
$$

Here I have introduced the total heating rate:

$$
\mathcal{H} = \mathcal{H}^{\text{visc}} + \mathcal{H}^{\text{ext}} \; . \quad\quad\quad (9.6.14)
$$

Here $\mathcal{H}^{\text{visc}}$ gives the contribution of viscous dissipation as given by Eqn. (9.6.11), and $\mathcal{H}^{\text{ext}}$ includes all *external* sources (or sinks) of heat, such as radiation losses.

9.7   Conservative equations for a viscous fluid

The equations for a viscous fluid can be cast in a conservative form. The derivation is quite analogous to the case without viscosity as treated in Chapter 3, and I will not go through it in great detail. Rather, I will stress the influence of the viscous terms.

The equation of mass conservation is not affected by the presence of viscosity:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{V}) = 0 \ .$$

(9.7.1)

The conservation law for the fluid momentum density, $\boldsymbol{M} = \rho \boldsymbol{V}$, is also formally the same as for an ideal fluid, simply because the viscous force can be written as the (negative) divergence of the viscous shear tensor $\mathbf{T}^{\mathrm{visc}}$. One simply defines the fluid stress tensor in such a way that it includes the contribution of viscosity:

$$\mathbf{T} \equiv \rho \boldsymbol{V} \otimes \boldsymbol{V} + P \, \boldsymbol{I} + \mathbf{T}^{\mathrm{visc}} \ . \tag{9.7.2}$$

The conservative equation of motion therefore takes the same form as for an ideal (non-viscous) fluid:

$$\frac{\partial \boldsymbol{M}}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{T} = -\rho \, \boldsymbol{\nabla} \Phi \ .$$

(9.7.3)

Here I have re-introduced the effects of gravity.

The conservative form of the equation for the fluid energy density takes some thought. The equation for the kinetic energy (Eqn. 9.6.2) can be written as

$$\frac{\partial}{\partial t} \left(\frac{\rho V^2}{2} \right) + \boldsymbol{\nabla} \cdot \left(\rho \boldsymbol{V} \, \frac{V^2}{2} + \mathbf{T}^{\mathrm{visc}} \cdot \boldsymbol{V} \right) = -(\boldsymbol{V} \cdot \boldsymbol{\nabla}) P - \rho (\boldsymbol{V} \cdot \boldsymbol{\nabla}) \Phi - \mathcal{H}^{\mathrm{visc}} \ . \tag{9.7.4}$$

Here I have used tensor identity (9.6.4) once again, together with the definition of the viscous heating rate.

If one uses the thermodynamic relations (3.2.22) and (3.2.23),

$$\boldsymbol{\nabla} P = \rho \boldsymbol{\nabla} h - \rho T \boldsymbol{\nabla} s \ , \ \ \rho \frac{\partial e}{\partial t} = \rho T \frac{\partial s}{\partial t} - \frac{P}{\rho} \boldsymbol{\nabla} \cdot (\rho \boldsymbol{V}) \tag{9.7.5}$$

together with

$$\rho(\boldsymbol{V} \cdot \boldsymbol{\nabla})\Phi = \frac{\partial(\rho\Phi)}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{V} \ \Phi) - \frac{\partial \Phi}{\partial t} \ , \tag{9.7.6}$$

one finds that one can write this relationship as an equation for the total (thermal plus kinetic) energy density of the fluid:

$$\frac{\partial}{\partial t} \left(\tfrac{1}{2}\rho V^2 + \rho\, e + \rho\, \Phi \right) \ + \ \boldsymbol{\nabla} \cdot \left[\rho \boldsymbol{V} \left(\tfrac{1}{2}V^2 + h + \Phi \right) + \mathbf{T}^{\mathrm{visc}} \cdot \boldsymbol{V} \right] =$$

$$= \ \rho T \underbrace{\left(\frac{\partial s}{\partial t} + (\boldsymbol{V} \cdot \boldsymbol{\nabla})s \right)}_{\text{thermal energy gained}} - \underbrace{\left(\mathcal{H}^{\mathrm{visc}} + \rho\, (\mathbf{V} \cdot \boldsymbol{\nabla})\Phi \right)}_{\text{kinetic energy lost}} \ . \tag{9.7.7}$$

If we now substitute relation (9.6.13) for the entropy density, one finds that the viscous heating term involving $\mathcal{H}^{\mathrm{visc}}$ drops out of the equation. This is obvious from a physical point-of-view: the kinetic energy lost due to internal friction, as given in Eqn. (9.7.4), is added to the thermal energy of the gas. Since viscosity is an *internal* process, it can not change the total energy of the gas or fluid! The only viscous effect that remains in the conservative form of the energy equation is the viscous contribution to the energy flux, equal to $\boldsymbol{F}^{\mathrm{visc}} = \mathbf{T}^{\mathrm{visc}} \cdot \boldsymbol{V}$. The final form of the conservative energy equation for a viscous medium reads:

$$\frac{\partial}{\partial t} \left(\tfrac{1}{2}\rho\, V^2 + \rho e + \rho\Phi \right) + \boldsymbol{\nabla} \cdot \left[\rho \boldsymbol{V} \left(\tfrac{1}{2}V^2 + h + \Phi \right) + \mathbf{T}^{\mathrm{visc}} \cdot \boldsymbol{V} \right] = \mathcal{H}_{\mathrm{eff}} \ . \tag{9.7.8}$$

As before, the net heating term $\mathcal{H}_{\mathrm{eff}}$ contains only the heat added or removed irreversibly for the system by external processes such as radiation losses, and the violent relaxation term that occurs in a time-dependent gravitational field:

$$\mathcal{H}_{\mathrm{eff}} \equiv \mathcal{H}^{\mathrm{ext}} + \rho \frac{\partial \Phi}{\partial t} \ . \tag{9.7.9}$$