

Charles R. Doering\* and Camilla Nobili  
**Lectures on Stirring, Mixing and Transport**

**Abstract:** These lectures introduce some basic concepts and mathematical approaches to the quantitative study of passive tracer mixing, dispersion and transport resulting from stirring by incompressible flows. Qualitative and quantitative similarities and differences between mixing and transport are explored. We focus on notions of *effective diffusion* and its dependence on the strength of stirring relative to molecular diffusion.

**Keywords:** Mixing, transport, dispersion, advection-diffusion equation, diffusion processes, stochastic differential equations, Itô calculus

## 1.1 Stirring, mixing and transport

The words stirring, mixing and transport are often interchanged but they describe very different processes. Let us distinguish the notions:

*Stirring:* the flow of a fluid.

*Mixing:* the intermingling of distinct materials or fluid properties that were originally separated in space.

*Transport:* the displacement of material or fluid properties from one place to another.

There are various levels at which mixing can be observed. At the smallest scales mixing happens via molecular diffusion, and this may be a very slow process. Molecular diffusion is described by the heat equation

$$\dot{T} = \kappa \Delta T, \quad (1.1)$$

where  $T(x, t)$  is the concentration of a material or density or intensity of a property (such as temperature) in the fluid, and  $\kappa$  is the *molecular diffusion coefficient* which is a property of the solvent and the material or property that is dissolved in it. The heat equation (1.1) must also be augmented with suitable initial and boundary conditions which depend on the specific application. In these lectures we will consider  $T$  to be the density of a *passive tracer*, a concentration of particles (or markers) that do not feed back and affect the motion of the fluid.

For simplicity let  $x \in \mathbb{R}$  (what follows can be easily generalized to  $\mathbf{x} \in \mathbb{R}^n$ ) and suppose that at the initial time  $t = 0$  the concentration  $T$  satisfies

$$T(x, 0) = \delta(x),$$

---

\***Corresponding Author: Charles R. Doering:** Center for the Study of Complex Systems, University of Michigan, U.S.A., E-mail: doering@umich.edu

**Camilla Nobili:** Departement Mathematik und Informatik, Universität Basel, Switzerland, E-mail: camilla.nobili@unibas.ch

where  $\delta(x)$  is the delta distribution centered at  $x = 0$ . Then the solution is

$$T(x, t) = \frac{1}{\sqrt{4\pi\kappa t}} \exp\left(-\frac{x^2}{4\kappa t}\right).$$

We first observe that the integral over the whole space is preserved at each time

$$\int T(x, t) dx = 1, \quad (1.2)$$

and in particular  $\frac{d}{dt} \int T dx = 0$ . Moreover, since there is no drift, the center of mass location does not move with time,

$$\int xT(x, t) dx = 0. \quad (1.3)$$

To see this, multiply (1.1) by  $x$  and integrate by parts:

$$\frac{d}{dt} \int xT dx = \int x\partial_t T dx = \kappa \int x\partial_x^2 T dx = -\kappa \int \partial_x T dx = 0.$$

On the other hand the width of the distribution, i.e., the variance, grows linearly in time:

$$\int x^2 T(x, t) dx = 2\kappa t. \quad (1.4)$$

To see this, multiply (1.1) by  $x^2$  and integrate by parts twice:

$$\frac{d}{dt} \int x^2 T dx = \int x^2 \partial_t T dx = \kappa \int x^2 \partial_x^2 T dx = 2\kappa \int T dx = 2\kappa.$$

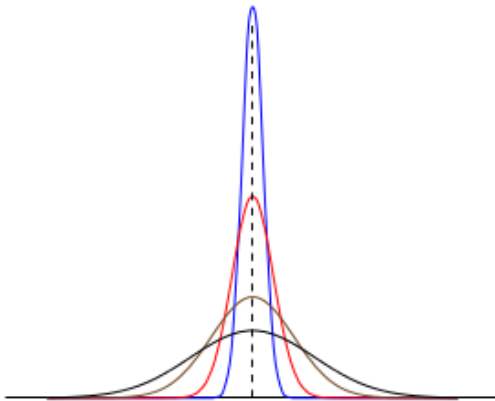


Fig. 1.1. Gaussian distributions.

Molecular diffusion coefficients are typically small numbers in familiar physical units. For thermal diffusion in water, for example,  $\kappa \approx 10^{-5} \frac{m^2}{s}$  while for, say, ink in water  $\kappa \approx 10^{-7} \frac{m^2}{s}$ .

Now consider a stochastic process  $X_t$  describing the position of the particle at time  $t$ . Suppose that the process starts at 0 at  $t = 0$ , i.e.,  $X_0 = 0$ , and subsequently evolves according to

$$dX_t = \sqrt{2\kappa} dW_t, \quad (1.5)$$

where  $W_t$  is the Wiener process (see Appendix). This stochastic dynamics corresponds to the heat equation in the sense that the time-dependent probability density for  $X_t$  satisfies (1.1). Using the stochastic calculus we can directly compute the expectations  $\mathbb{E}(X_t)$  and  $\mathbb{E}(X_t^2)$  and see that they coincide with (1.3) and (1.4).

First, since the Wiener process satisfies

$$\mathbb{E}(dW_t) = 0, \quad (1.6)$$

by (1.5) we have

$$\mathbb{E}(dX_t) = d\mathbb{E}(X_t) = \sqrt{2\kappa} \mathbb{E}(dW_t) = 0,$$

which implies  $\mathbb{E}(X_t) = \mathbb{E}(X_0) = 0$ .

Then using the Itô formula (see Appendix) to write the stochastic increment of  $X_t^2$ ,

$$d(X_t^2) = 2X_t dX_t + \frac{1}{2} 2 (dX_t)^2, \quad (1.7)$$

we have

$$\mathbb{E}(d(X_t^2)) = 2 \mathbb{E}(X_t dX_t) + \frac{1}{2} 2 \mathbb{E}((dX_t)^2). \quad (1.8)$$

The first term of the right hand side vanishes thanks to (1.5) and the property of Itô equations that, for any function  $\phi(\cdot)$ ,

$$\mathbb{E}(\phi(X_t) dW_t) = \mathbb{E}(\phi(X_t)) \mathbb{E}(dW_t) = 0. \quad (1.9)$$

The second term is  $\mathbb{E}((dX_t)^2) = 2\kappa \mathbb{E}(dW_t^2) = 2\kappa dt$  where we call on the property of the Wiener process that

$$\mathbb{E}(dW_t^2) = dt. \quad (1.10)$$

Therefore  $\mathbb{E}(d(X_t^2)) = d\mathbb{E}(X_t^2) = 2\kappa dt$  so that  $\mathbb{E}(X_t^2) = 2\kappa t$ .

This simple example illustrates how the stochastic differential equation formalism can, at least in some cases, produce quick and clean computations of moments.

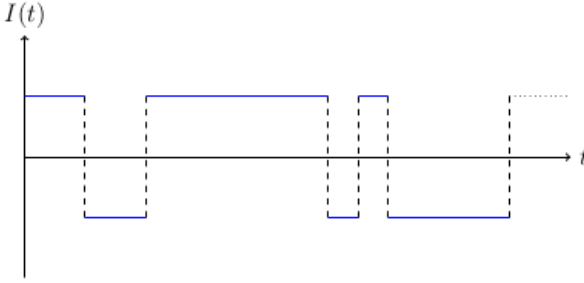
## 1.2 Advection as diffusion

Consider the two-level Markov process  $I_t = \pm 1$ , which consists of randomly switching between the (+1) and (-1) states. The waiting times between jumps are independent

and exponentially distributed. If  $\mathbb{E}(I_0) = 0$  then

$$\mathbb{E}(I_t I_s) = \exp(-2\gamma|t-s|) = \exp\left(-\frac{|t-s|}{\tau}\right), \quad (1.11)$$

where  $\tau = \frac{1}{2\gamma}$  is the process' correlation time.



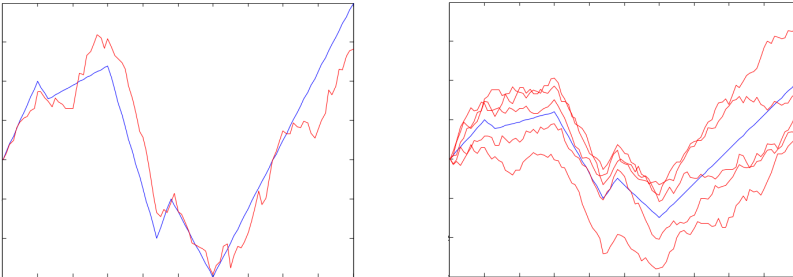
**Fig. 1.2.** The two-level (dichotomous) Markov process that jumps between the  $(\pm 1)$  states.

We denote the probability of being in the  $(\pm 1)$  state at time  $t$  by  $p^\pm(t) \in [0, 1]$ . The rate of change of the probabilities is summarized by the ODEs

$$\frac{d}{dt} \begin{bmatrix} p^+(t) \\ p^-(t) \end{bmatrix} = \begin{bmatrix} -\gamma & \gamma \\ \gamma & -\gamma \end{bmatrix} \begin{bmatrix} p^+ \\ p^- \end{bmatrix}.$$

Now consider a particle whose velocity switches randomly between  $\pm U$  (where  $U$  is a real – without loss of generality positive – number) in the two-level Markov process so that its position as a function of time  $X_t$  satisfies the differential equation

$$dX_t = U I_t dt. \quad (1.12)$$



**Fig. 1.3.** Particle trajectories: in blue the piecewise constant velocity path and in red an illustrative Brownian motion (left) and many Brownian motions (right).

The probability densities  $\rho^\pm(x, t)$  — the densities for the particle's position  $X_t$  conditioned on its velocity being  $\pm U$  at time  $t$  — evolve according to

$$\frac{\partial}{\partial t} \begin{bmatrix} \rho^+(x, t) \\ \rho^-(x, t) \end{bmatrix} = \begin{bmatrix} -U\partial_x - \gamma & \gamma \\ \gamma & U\partial_x - \gamma \end{bmatrix} \begin{bmatrix} \rho^+(x, t) \\ \rho^-(x, t) \end{bmatrix}. \quad (1.13)$$

If the particle is in the + state it is advected to the right (the  $-U\partial_x$  term) or it jumps down to the – state at rate  $\gamma$ . In the – state the particle advects to the left or switches to the + state at rate  $\gamma$ .

The *marginal distribution* of the particle position is  $\rho(x, t) = \rho^+(x, t) + \rho^-(x, t)$ . Consider as well the auxiliary quantity  $q(x, t) = \rho^+(x, t) - \rho^-(x, t)$ . Adding and subtracting the differential equations in (1.13) we obtain evolution equations for  $\rho$  and for  $q$ :

$$(A) \quad \partial_t \rho = -U\partial_x q \quad \text{and,}$$

$$(B) \quad \partial_t q = -U\partial_x \rho - 2\gamma q.$$

If  $q(x, 0) = 0$  (meaning that  $\rho^+ = \rho^-$  initially) the solution of (B) is

$$q(x, t) = -U \int_0^t e^{-2\gamma(t-s)} \partial_x \rho(x, s) ds = -U\partial_x \int_0^t e^{-2\gamma(t-s)} \rho(x, s) ds$$

and inserting this into (A) we find

$$\frac{\partial \rho}{\partial t} = U^2 \partial_x^2 \int_0^t e^{-2\gamma(t-s)} \rho(x, s) ds,$$

which is, in a certain sense, “close” to a diffusion equation.

Indeed, multiplying and dividing the expression above by  $2\gamma$  we obtain

$$\frac{\partial \rho}{\partial t} = \frac{U^2}{2\gamma} \partial_x^2 \int_0^t 2\gamma e^{-2\gamma(t-s)} \rho(x, s) ds,$$

where, on time scales larger than the velocity process' correlation time  $\tau$ , the kernel mimics a delta distribution in  $s$  concentrated around  $t$ . More to the point, if we send  $\gamma \rightarrow \infty$  and set  $\frac{U^2}{2\gamma} = \kappa$ , we are left with the diffusion equation

$$\frac{\partial \rho}{\partial t} = \kappa \frac{\partial^2 \rho}{\partial x^2}.$$

Alternatively, we can directly compute the variance of the position process  $X_t = \int_0^t U I_s ds$  finding

$$\mathbb{E}(X_t^2) = \frac{U^2}{2\gamma} \left( 2t - \frac{1}{\gamma} (1 - \exp(-2\gamma t)) \right). \quad (1.14)$$

This calculation proceeds as follows. First note that

$$X_t^2 = \left( U \int_0^t I_s ds \right) \left( U \int_0^t I_r dr \right) = U^2 \int_0^t \int_0^t I_s I_r ds dr.$$

Then take the expectation to obtain

$$\begin{aligned} \mathbb{E}(X(t)^2) &= U^2 \mathbb{E} \int_0^t \int_0^t I_s I_r ds dr \\ &= U^2 \int_0^t \int_0^t \mathbb{E}(I_s I_r) ds dr \\ &= U^2 \int_0^t \int_0^t \exp(-2\gamma|s-z|) ds dr. \end{aligned}$$

Decompose the domain of integration as  $\{0 \leq s \leq t, 0 \leq z \leq s\} \cup \{0 \leq z \leq t, 0 \leq s \leq z\}$  and observe that by the symmetry of the integrand

$$\begin{aligned} \int_0^t \int_0^t \exp(-2\gamma|s-r|) ds dr &= \int_0^t \int_0^r \exp(-2\gamma(-s+r)) ds dr \\ &\quad + \int_0^t \int_0^s \exp(-2\gamma(s-r)) dr ds \\ &= 2 \int_0^t \int_0^r \exp(-2\gamma(-s+r)) ds dr \\ &= 2 \frac{1}{2\gamma} \int_0^t (1 - \exp(-2\gamma r)) dr \\ &= 2 \frac{1}{2\gamma} \left( t + \frac{1}{2\gamma} (\exp(-2\gamma t) - 1) \right) \\ &= \frac{1}{2\gamma} \left( 2t + \frac{1}{\gamma} (\exp(-2\gamma t) - 1) \right), \end{aligned}$$

which coincides with (1.14).

Then it is clear that  $E(X(t)^2) \sim \frac{U^2}{\gamma}(t - \tau)$  as  $t \rightarrow \infty$  where  $\tau = \frac{1}{2\gamma}$ . Thus, for  $t \gg \tau$  this random process disperses particles like a diffusion process (albeit with a short delay) with *effective diffusivity*

$$\kappa_{\text{eff}} := U^2 \frac{1}{2\gamma} = U^2 \tau.$$

The correlation time  $\tau$  is the typical time it takes before changing velocity, a.k.a. the mean free time. It is also natural to define the mean free length  $\ell = U\tau$  — the typical distance traveled before changing velocity — and write

$$\kappa_{\text{eff}} = U\ell.$$

This example shows how advection — in this case advection by a random-in-time velocity field — can disperse tracer particles diffusively, at least on sufficiently long time scales. In such systems the theoretical challenge is often to deduce an effective diffusivity in terms of details of the advecting flow field.

### 1.3 Can transport enhance diffusion?

Here we illustrate how a combination of advection and diffusion can enhance particle dispersion beyond the effect of either alone.

#### 1.3.1 Shear flow

Consider a uniform shear flow  $\mathbf{u} = \mathbf{i}Sy$  where  $\mathbf{i}$  is the horizontal unit vector,  $S$  is the shear and  $y$  is the vertical component of the position vector  $\mathbf{x} = \mathbf{i}x + \mathbf{j}y$ .

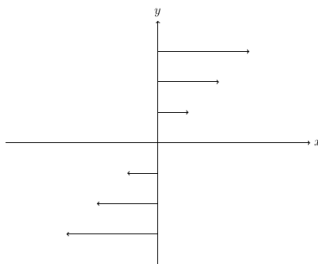


Fig. 1.4. Shear flow.

First consider motion in this flow field without diffusion. The horizontal and vertical positions of a passive tracer particle satisfy

$$\begin{aligned} dX_t &= S Y_t dt, \\ dY_t &= 0, \end{aligned}$$

with initial data which we will take to be  $X_0 = 0$  and  $Y_0 = y_0$ . Integrating the differential equations we find  $X_t = S y_0 t$  and  $Y_t = y_0$  from which we easily see the variance

of the position to be  $\mathbb{E}(X_t^2) = S^2 Y_0^2 t^2$ . This sort of dispersion,  $\mathbb{E}(X_t^2) \sim t^2$ , is what we expect from ballistic motion.

Now include diffusion in the differential equations — making them truly stochastic differential equations — and consider

$$\begin{aligned} dX_t &= SY_t dt + \sqrt{2\kappa} dW_t^{(1)}, \\ dY_t &= \sqrt{2\kappa} dW_t^{(2)}, \end{aligned}$$

where  $W_t^{(1)}$  and  $W_t^{(2)}$  are independent Wiener processes. Then the time-dependent joint probability density of particle's position  $\rho(x, y, t)$  evolves according to the advection-diffusion partial differential (also known in various communities as a Fokker-Plank or Forward Kolmogorov) equation

$$\frac{d\rho}{dt} = -\frac{\partial}{\partial x}(SY\rho) + \kappa \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \rho,$$

a linear but non-constant coefficient partial differential equation.

We can easily compute the expectation values of  $X_t$  and  $Y_t$ . For simplicity suppose that  $X_0 = 0 = Y_0$ . Using the initial condition we have  $Y_t = \sqrt{2\kappa} W_t$  and, since the expectation of the Wiener process is zero,  $\mathbb{E}(Y_t) = 0$ . From this it is easy to deduce that  $\mathbb{E}(X_t) = 0$  as well.

The variance of  $X_t$  and  $Y_t$  follow from the Itô formula (1.7). Applied to  $Y_t^2$  we see, as before, that

$$\mathbb{E}(Y_t^2) = 2\kappa t. \quad (1.15)$$

On the other hand applying the Itô formula to  $X_t^2$  yields

$$\begin{aligned} \mathbb{E}(dX_t^2) &= 2\mathbb{E}(X_t(SY_t dt + \sqrt{2\kappa}dW_t)) + 2\kappa dt \\ &= 2S\mathbb{E}(X_t Y_t dt) + 2\sqrt{2\kappa}\mathbb{E}(X_t dW_t) + 2\kappa dt \\ &= 2S\mathbb{E}(X_t Y_t)dt + 2\kappa dt. \end{aligned} \quad (1.16)$$

It is left to compute  $\mathbb{E}(X_t Y_t)$ . Using the Leibniz rule and (1.9) we find

$$\begin{aligned} \mathbb{E}(d(X_t Y_t)) &= \mathbb{E}(Y_t dX_t + X_t dY_t + dX_t dY_t) \\ &= S\mathbb{E}(Y_t^2)dt + \sqrt{2\kappa}\mathbb{E}(Y_t dW_t) + \sqrt{2\kappa}\mathbb{E}(X_t dW_t) \\ &= 2S\kappa t dt. \end{aligned}$$

This implies that

$$\mathbb{E}(X_t Y_t) = S\kappa t^2.$$

Inserting this into (1.16) we find

$$\mathbb{E}(dX_t^2) = d\mathbb{E}(X_t^2) = 2S^2\kappa t^2 dt + 2\kappa dt,$$

which upon integration yields

$$\mathbb{E}(X_t^2) = \frac{2}{3}S^2\kappa t^3 + 2\kappa t.$$



At early times molecular diffusion dominates particle dispersion, i.e.,  $\mathbb{E}(X_t^2) \sim t$ , but later, in the presence of shear, the variance of the (horizontal) position grows  $\sim t^3$ . This super-diffusive motion is faster even than ballistic dispersion  $\sim t^2$ . Advection and diffusion together can disperse the particles faster than either alone!

### 1.3.2 Advection-diffusion equation

We have seen that mixing — thought of as passive tracer particle dispersion measured by position variance — can be modeled in two equivalent ways. For one, given an initial position, the stochastic trajectory of a passive tracer in  $\mathbb{R}^d$  advected by a given incompressible flow field  $\mathbf{u}$  (i.e., with  $\nabla \cdot \mathbf{u} = 0$ ) may be described by the stochastic differential equation

$$d\mathbf{X}_t = \mathbf{u}(X_t, t) dt + \sqrt{2\kappa} d\mathbf{W}_t.$$

Equivalently a diffusing concentration field (density) field  $T(\mathbf{x}, t)$  stirred by the incompressible fluid evolves according to the advection-diffusion equation

$$\dot{T} + \mathbf{u} \cdot \nabla T = \kappa \Delta T. \quad (1.17)$$

Previously we saw that if  $\mathbf{u} = 0$  then the variance of the particle position is  $\mathbb{E}(X_t^2) \sim 2\kappa t$ . The density distribution of particles gets wider and wider. In this section we want to consider an alternative measure of mixing, namely the *suppression of the variance of the density*.

Consider the problem when the scalar field  $T$  satisfies the equation (1.17) in a bounded (and regular) domain  $\Omega \subset \mathbb{R}^3$  with no-flux boundary conditions,

$$\mathbf{n} \cdot \mathbf{u} = 0 \quad \text{and} \quad \mathbf{n} \cdot \nabla T = 0 \quad \text{on } \partial\Omega, \quad (1.18)$$

where  $\mathbf{n}$  is the outward normal to  $\partial\Omega$ .

Thanks to the incompressibility condition, we can rewrite (1.17) as

$$\dot{T} + \nabla \cdot \mathbf{J} = 0, \quad (1.19)$$

where  $\mathbf{J} = \mathbf{u}T - \kappa \nabla T$ . The field vector  $\mathbf{J}$  has the units of a current,

$$[\mathbf{J}] \approx \frac{\text{stuff}}{\text{area} \times \text{time}},$$

when the density  $T$  has units

$$[T] \approx \frac{\text{stuff}}{\text{volume}}$$

and the molecular diffusivity has units

$$[\kappa] = \frac{\text{length}^2}{\text{time}}.$$

If we isolate a small surface area in the domain and identify its normal unit vector  $\mathbf{n}$ , then  $\mathbf{J} \cdot \mathbf{n}$  is the flux, the amount of stuff per unit time and per unit area, passing through that surface. (Similar considerations apply in lower dimensions: in  $\mathbb{R}^2$  volumes are replaced by areas and area elements by line elements while in  $\mathbb{R}^1$  “volumes” are line segments and “area elements” are their end points.)

Define the spatial average

$$\langle \cdot \rangle_{\mathbf{x}} = \frac{1}{|\Omega|} \int_{\Omega} (\cdot) d\mathbf{x}$$

and, for the purposes of this discussion, consider  $\Omega \subset \mathbb{R}^n$  to be a spatially periodic domain  $[0, 2\pi)^n$ . Since there is no flux at the boundary — there are no boundaries — (1.17) implies

$$\frac{d}{dt} \langle T(\cdot, t) \rangle_{\mathbf{x}} = 0,$$

so that  $\langle T(\cdot, t) \rangle_{\mathbf{x}}$  is constant in time. Then denote  $\langle T \rangle_{\mathbf{x}} := \langle T(\cdot, t) \rangle_{\mathbf{x}}$  and decompose  $T$  into the stationary spatial mean  $\langle T \rangle_{\mathbf{x}}$  and the deviation  $\theta$  such that  $T(\mathbf{x}, t) = \langle T \rangle_{\mathbf{x}} + \theta(\mathbf{x}, t)$ . The deviation satisfies the same advection-diffusion equation

$$\dot{\theta} + \mathbf{u} \cdot \nabla \theta = \kappa \Delta \theta, \quad (1.20)$$

albeit with spatial mean zero at all times,  $\langle \theta(\cdot, t) \rangle_{\mathbf{x}} = 0$ .

When the density  $T$  is constant in space, i.e., when tracer particles are uniformly distributed and hence perfectly well mixed, the deviation  $\theta = 0$ . On the other hand when the solution is *not* well mixed then  $T$  is *not* constant and  $\theta \neq 0$ .

Therefore the spatial variance of  $T$ , i.e.,  $\langle (T(\cdot, t) - \langle T \rangle_{\mathbf{x}})^2 \rangle_{\mathbf{x}} = \langle \theta(\cdot, t)^2 \rangle_{\mathbf{x}}$  is a meaningful measure of the degree of “mixedness” of the tracers throughout  $\Omega$ . The smaller the value of  $\langle \theta^2 \rangle_{\mathbf{x}}$  is, the better mixed the tracers are.

Testing the advection-diffusion equation (1.20) with  $\theta$  and integrating by parts using the no-flux condition at the boundary, we obtain an expression for the evolution of the concentration variance

$$\frac{d}{dt} \langle \theta(\cdot, t)^2 \rangle_{\mathbf{x}} = -2\kappa \langle |\nabla \theta|^2 \rangle_{\mathbf{x}}. \quad (1.21)$$

Note specifically two things:

- molecular diffusion is necessary to make the concentration variance change (at least for sufficiently smooth densities such that  $\nabla T = \nabla \theta$  is spatially square integrable), and
- the stirring flow field  $\mathbf{u}$  does not explicitly appear in the variance evolution equation (1.21).

Suppose that  $\mathbf{u} = 0$  and choose the initial data  $\theta_0(\mathbf{x}) = A \sin(\mathbf{k} \cdot \mathbf{x})$  where  $\mathbf{k}$  is a vector with integer components. Then the solution of (1.20) is  $\theta(\mathbf{x}, t) =$

$A \exp(-\kappa k^2 t) \sin(\mathbf{k} \cdot \mathbf{x})$  where  $k = |\mathbf{k}|$ . Using the fact that  $\langle \sin(\mathbf{k} \cdot \mathbf{x})^2 \rangle_{\mathbf{x}} = \frac{1}{2}$ , we see that the variance decays exponentially,

$$\frac{\langle \theta(\cdot, t)^2 \rangle_{\mathbf{x}}}{\langle \theta_0^2 \rangle_{\mathbf{x}}} = \exp(-2\kappa k^2 t), \quad (1.22)$$

with decay rate proportional to the diffusivity and inversely proportional to (the square of) the length scale in the density deviation. The question is, what role might stirring, i.e., advection with  $\mathbf{u} \neq 0$ , play in the variance suppression game? In fact stirring may greatly enhance the variance dissipating power of molecular diffusion to amplify mixing.

We illustrate this with a simple example. Focus for the moment on  $\Omega = \mathbb{R}^2$  and consider a uniform shear flow

$$\mathbf{u} = \mathbf{i}S\mathbf{y}.$$

Then the advection-diffusion equation (1.20) is

$$\dot{\theta} + S\mathbf{y}\partial_x\theta = \kappa(\partial_x^2 + \partial_y^2)\theta, \quad (1.23)$$

and, when  $\theta(x, y, 0) = A \sin kx$ , the solution is (exercise: check this)

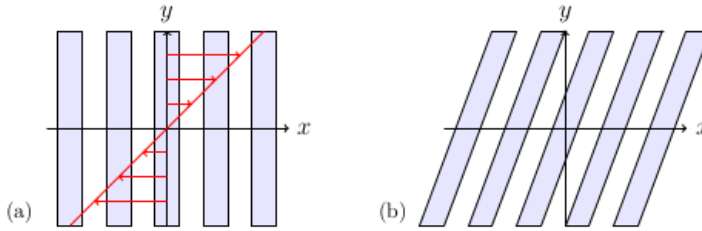
$$\theta(x, y, t) = A \exp\left(-\kappa k^2\left(t + \frac{1}{3}S^2 t^3\right)\right) \sin(k(x - Sty)).$$

The solution shows that the spatial structure of the density deviation remains sinusoidal but with a time-dependent wave-vector,  $(\mathbf{i}k) \rightarrow (\mathbf{i}k - \mathbf{j}kSt)$ . The shear flow transfers density fluctuations from the initial length scale  $\sim k^{-1}$  to increasingly smaller scales  $\sim \left(k \times \sqrt{1 + S^2 t^2}\right)^{-1}$ . The concentration variance is

$$\langle \theta(\cdot, t)^2 \rangle_{\mathbf{x}} \approx \frac{A^2}{2} \exp(-2\kappa k^2(t + \frac{1}{3}S^2 t^3)), \quad (1.24)$$

where we have used  $\langle \sin^2(k(x - Sty)) \rangle_{\mathbf{x}} \approx \frac{1}{2}$  in sufficiently large and suitably shaped volumes  $\Omega \subset \mathbb{R}^2$ .

Shear stirring greatly enhances the exponential decay of the variance compared to the effect of diffusion alone. The flow field accomplishes this by tilting level sets of the initial condition which were aligned with the  $y$ -axis, forcing the wavelength to become smaller and smaller which increases concentration gradients to amplify the effect of molecular diffusion and increase the rate of variance dissipation.



**Fig. 1.5.** The level set of the initial condition  $\theta_0(x, y)$  are aligned with the  $y$ -axis and subsequently tilted by the shear flow.

Observe as well that if  $\theta_0(x, y) = \sin ky$ , i.e., if level sets of the concentration are initially aligned along the  $x$  axis, then the solution would be the same as for the diffusion-sans-advection equation. In that case this particular flow field plays *no* role in the concentration evolution. This example illustrates the importance of details of the alignment of density gradients and the flow — more precisely directions associated with rates of strain in the flow — in mixing processes as characterized by variance dissipation.

More generally, consider the evolution of the *concentration gradient*. The gradient of the advection-diffusion equation (1.20) is

$$\frac{d}{dt} \nabla \theta + \mathbf{u} \cdot \nabla \nabla \theta + \nabla \mathbf{u} \cdot \nabla \theta = \kappa \nabla \Delta \theta,$$

and, testing this with  $\nabla \theta$  followed with integrations by parts, implies (when boundary terms vanish) that

$$\frac{d}{dt} \langle |\nabla \theta|^2 \rangle_x = -\kappa \langle |\Delta \theta|^2 \rangle_x - \langle \nabla \theta \cdot \nabla \mathbf{u} \cdot \nabla \theta \rangle_x. \quad (1.25)$$

The first term on the right hand side above, the negative (semi)definite term, results from the effect of diffusion which strongly dissipates density fluctuations at small scales. The second term is the effect of stirring, showing how the symmetric part of  $(\nabla \mathbf{u})$  — the *rate of strain* tensor  $(\nabla \mathbf{u})_{\text{sym}}$  — affects the gradient variance.

For a two-dimensional velocity field  $\mathbf{u} = i\mathbf{u} + j\mathbf{v}$  the rate of strain tensor is

$$(\nabla \mathbf{u})_{\text{sym}} = \begin{bmatrix} u_x & \frac{v_x + u_y}{2} \\ \frac{v_x + u_y}{2} & v_y \end{bmatrix}.$$

At each point in space this matrix is symmetric and, thanks to the incompressibility condition  $u_x + v_y = 0$ , trace-free. Therefore whenever (generically) it is not identically zero, the rate of strain matrix has real and distinct eigenvalues  $\lambda_{\pm}$  with orthogonal eigenvectors. The eigenvalues sum to zero so at generic non-degenerate points let  $\lambda_+ >$

0 and  $\lambda_- < 0$ . The eigenvector associated with  $\lambda_+$  is the *expansive* direction while the eigenvector associated with  $\lambda_-$  is the *contractive* direction.

In order for  $(\nabla \mathbf{u})_{\text{sym}}$  to generate gradient variance (and thus contribute to concentration variance dissipation and increased mixing)  $\nabla \theta$  should be predominantly aligned along the *contractive* directions. That is, level sets of  $\theta$  should tend to be locally aligned with the *contractive* directions. Then  $-\langle \nabla \theta \cdot \nabla \mathbf{u} \cdot \nabla \theta \rangle$  on the right hand side of (1.25) is positive which adds positively to  $\frac{d}{dt} \langle |\nabla \theta|^2 \rangle$ .

For the shear flow  $\mathbf{u} = iSy$

$$\nabla \mathbf{u} = \begin{bmatrix} 0 & 0 \\ S & 0 \end{bmatrix},$$

and the rate of strain tensor is

$$(\nabla \mathbf{u})_{\text{sym}} = \frac{S}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

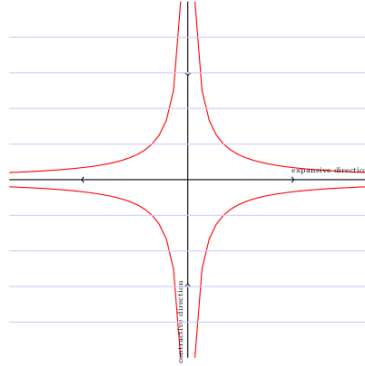
The eigenvectors are  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  and  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$  corresponding, respectively, to the maximally expanding and contractive directions. This steady shear flow is not an ideal gradient amplifier, however, because it changes the direction of the concentration gradients: sometimes it aligns concentration gradients along the contractive direction, but often the flow reorients them away as well.

### 1.3.3 Uniform strain flow

The *pure strain* flow, a divergence-free vector field of the form  $\mathbf{u} = i\Gamma x - j\Gamma y$  with  $\Gamma > 0$ , has uniform gradient amplification properties. The trajectories of tracer particles are hyperbole and the rate of strain tensor is

$$(\nabla \mathbf{u})_{\text{sym}} = \Gamma \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The origin is a saddle point of the flow, the contractive direction is the  $y$ -direction and the expansive direction coincides with the  $x$ -direction.



**Fig. 1.6.** The trajectories of the particles in the pure strain flow are hyperbole.

Maximal gradient amplification – and consequently maximal variance dissipation – is realized by initial data of the form  $\theta_0(x, y) = A \sin k_0 y$  with level sets of the concentration fluctuations aligned with the expansive directions. With the pure strain flow in (1.20) the advection-diffusion equation becomes

$$\dot{\theta} + \Gamma x \frac{\partial}{\partial x} \theta - \Gamma y \frac{\partial}{\partial y} \theta = \kappa \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \theta. \quad (1.26)$$

To solve it we make the ansatz  $\theta(x, y, t) = \Theta(t) \sin(k(t)y)$  where we highlight the dependence of the wavenumber on time. Inserting the ansatz into the equation we find the exact solution (exercise: check this)

$$\theta(x, y, t) = A \exp \left( -\frac{\kappa k_0^2}{2\Gamma} (\exp(2\Gamma t) - 1) \right) \sin(k_0 \exp(\Gamma t) y).$$

Tracer trajectories along the  $y$ -axis are squeezed exponentially into the origin and the wavelength  $k(t)$  increases exponentially in time. This exponential amplification results in the double exponential decay of the amplitude  $\Theta(t)$ .

An alternative analysis proceeds by considering the system of two linear stochastic differential equations

$$dX_t = +\Gamma X_t dt + \sqrt{2\kappa} dW_t^{(1)}, \quad (1.27)$$

$$dY_t = -\Gamma Y_t dt + \sqrt{2\kappa} dW_t^{(2)}, \quad (1.28)$$

describing a particle moving and diffusing in the strain flow, where  $W_t^{(1)}$  and  $W_t^{(2)}$  are independent Wiener processes. The solutions  $X_t$  and  $Y_t$  are *Ornstein-Uhlenbeck processes* and the moments are easy to compute. Suppose initial conditions  $X_0$  and  $Y_0$ . Then  $\mathbb{E}(dX_t) = d\mathbb{E}(X_t) = \Gamma \mathbb{E}(X_t) dt$  and  $\mathbb{E}(dY_t) = d\mathbb{E}(Y_t) = -\Gamma \mathbb{E}(Y_t) dt$  so

$$\mathbb{E}(X_t) = \mathbb{E}(X_0) \exp(\Gamma t) \quad \text{and} \quad \mathbb{E}(Y_t) = \mathbb{E}(Y_0) \exp(-\Gamma t).$$

The Itô analysis (1.7-1.10) then implies

$$\mathbb{E}(dX_t^2) = d\mathbb{E}(X_t^2) = +2\Gamma\mathbb{E}(X_t^2)dt + 2\kappa dt,$$

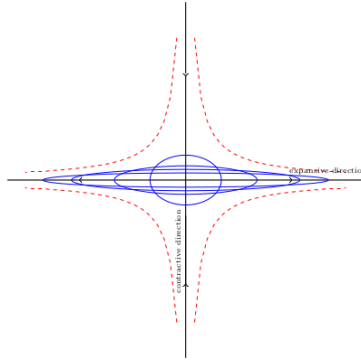
$$\mathbb{E}(dY_t^2) = d\mathbb{E}(Y_t^2) = -2\Gamma\mathbb{E}(Y_t^2)dt + 2\kappa dt.$$

Thus if we specialize to  $X_0 = 0 = Y_0$ , then

$$\mathbb{E}(X_t^2) = \frac{\kappa}{\Gamma}(\exp(+2\Gamma t) - 1),$$

$$\mathbb{E}(Y_t^2) = \frac{\kappa}{\Gamma}(1 - \exp(-2\Gamma t)).$$

The variance of  $X_t$  grows exponentially while the variance of  $Y_t$  converges not to zero, but to  $\frac{\kappa}{\Gamma}$  at  $t \rightarrow \infty$ . This means that a source point at the origin will be squeezed and stretched and eventually distributed over a thin layer of width  $(\frac{\kappa}{\Gamma})^{\frac{1}{2}}$  along the  $x$ -axis.



**Fig. 1.7.** A distribution of tracers is squeezed and stretched by the flow to be eventually distributed over a strip of vertical width  $(\frac{\kappa}{\Gamma})^{\frac{1}{2}}$ .

### 1.3.4 The smallest scale of mixing

It is of great interest to understand the smallest dynamical scales in turbulent fluid flows. This remains a great challenge for turbulent solutions of the three dimensional Navier-Stokes equations but for other (simpler) problems the smallest scales can be identified. We ask, *what is the smallest length-scale we expect to see in passive tracer mixing?* In absence of diffusion “good” stirring creates very fine filaments which, due to incompressibility, become thinner and thinner as they are stretched longer and longer. In presence of diffusion, however, these filaments reach a minimal width; the

filaments cannot get too thin because molecular diffusion broadens them. In the case of a pure strain flow the minimal width of the stretched blob is

$$\lambda_B = \left( \frac{\kappa}{\Gamma} \right)^{\frac{1}{2}}, \quad (1.29)$$

known as the *Batchelor scale*. The Batchelor scale quantifies the balance of strain (which stretches filaments and tends to make them thinner) with diffusion (which tends to broaden them).

For (homogeneous isotropic) turbulence we can relate the typical *strain rate*  $\Gamma$  with a quantity that characterizes the intensity of turbulence. Turbulence does not maintain itself; a source of energy must be supplied to sustain the motion and the relevant magnitude is the *energy dissipation rate per unit mass*

$$\epsilon := \nu \langle |\nabla u|^2 \rangle_{\mathbf{x}}. \quad (1.30)$$

(Note: denoting the turbulent energy dissipation rate per unit mass by  $\epsilon$  does *not* imply smallness in any sense; this is simply the traditional symbol for it in turbulence theory.) In turbulent flows at high Reynolds numbers — “small” viscosity — the flow stretches itself in order to increase its own velocity gradients which also enhances the effect of dissipation of energy at small scales due to viscosity. The energy dissipation rate  $\epsilon$  is the power (per unit mass) required to maintain (statistical) steady state turbulence.

Identifying the typical rate of strain with the root mean square rate of strain, i.e., writing  $\Gamma = \langle |\nabla u|^2 \rangle_{\mathbf{x}}^{1/2}$ , we have

$$\Gamma = (\epsilon \nu)^{\frac{1}{2}}.$$

Thus we can rewrite the Batchelor scale in terms of the energy dissipation rate

$$\lambda_B = \frac{\kappa^{\frac{1}{2}}}{\Gamma^{\frac{1}{2}}} = \frac{\kappa^{\frac{1}{2}} \nu^{\frac{1}{4}}}{\epsilon^{\frac{1}{4}}} = \left( \frac{\kappa}{\nu} \right)^{\frac{1}{2}} \left( \frac{\nu^3}{\epsilon} \right)^{\frac{1}{4}}.$$

Both  $\nu$  and  $\kappa$  have the units of length<sup>2</sup>/time, so  $\frac{\kappa}{\nu}$  is nondimensional. The energy dissipation rate per unit mass  $\epsilon$  has units length<sup>2</sup>/time<sup>3</sup> and  $(\nu^3/\epsilon)^{\frac{1}{4}}$  has units of a length. This quantity,

$$\lambda_k = \left( \frac{\nu^3}{\epsilon} \right)^{\frac{1}{4}},$$

is the *Kolmogorov scale*. It is supposedly the smallest length scale for the velocity field in turbulent flows.

The ratio  $\frac{\nu}{\kappa}$  is called the *Prandtl number* (when  $\kappa$  is a thermal diffusion coefficient) or the *Schmidt number* (when  $\kappa$  is a mass diffusion rate). Prandtl numbers vary greatly according to the fluid under consideration. Liquid metals have Prandtl numbers  $\ll 1$  while thick oils have Prandtl numbers  $\gg 1$ . Schmidt numbers are often high, though, and certainly so when the dissolved material has heavy molecular weight relative to



the solvent. Thus in many applications the Batchelor scale  $\lambda_B = \left(\frac{\kappa}{\nu}\right)^{\frac{1}{2}} \lambda_k$  is much smaller than the Kolmogorov scale. This means that one can have efficient and effective stirring even in mildly turbulent flows—and even at very small Reynolds numbers.

### 1.3.5 Quantitative indicators of mixing efficacy

We now consider quantification of the mixing enhancement of stirring and examine how such measures might depend on the strength of the stirring.

Suppose the goal is to reduce the scalar variance from its initial value as much as possible in a certain time interval and reconsider the initial data  $\theta_0(\mathbf{x}) = A \sin(\mathbf{k}_0 \cdot \mathbf{x})$ . In absence of flow, i.e.,  $\mathbf{u} = 0$ , the solution of (1.20) is  $\theta(\mathbf{x}, t) = A \exp(-\kappa k_0^2 t) \sin(\mathbf{k}_0 \cdot \mathbf{x})$  where  $k_0 = |\mathbf{k}_0|$  and the variance decays exponentially,

$$\frac{\langle \theta(\cdot, t)^2 \rangle_{\mathbf{x}}}{\langle \theta(\cdot, 0)^2 \rangle_{\mathbf{x}}} = \exp(-2\kappa k_0^2 t).$$

We can turn this formula inside out and note that molecular diffusivity may be expressed in terms of the time  $t_\varepsilon$  that it takes to reduce the variance to  $\varepsilon$  times its initial value:

$$\kappa = \frac{\ln \frac{1}{\varepsilon}}{2t_\varepsilon k_0^2}. \quad (1.31)$$

In the presence of stirring we may still use (1.31) to define an *effective diffusion coefficient* in terms of the time  $t_\varepsilon$  that it takes to reduce the variance to  $\varepsilon$  times its initial value. That is, we gauge the effect of stirring by identifying the value of molecular diffusion that would be required to achieve the same level of variance suppression in time  $t_\varepsilon$ :

$$\kappa_{\text{eff}} = \frac{\ln \frac{1}{\varepsilon}}{2t_\varepsilon k_0^2}. \quad (1.32)$$

In dimensionless terms we may consider *enhancement factor*

$$E := \frac{\kappa_{\text{eff}}}{\kappa}. \quad (1.33)$$

The task is to understand the dependence (i.e., scaling) of the dimensionless enhancement factor  $E$  on a dimensionless measure of the stirring strength commonly chosen to be the *Péclet number*

$$\text{Pe} = \frac{U\ell}{\kappa}, \quad (1.34)$$

where  $U$  is a relevant stirring speed, e.g, the root mean square stirring speed  $U = \langle \langle |\mathbf{u}|^2 \rangle_{\mathbf{x}} \rangle_t^{\frac{1}{2}}$ , and  $\ell$  is a suitable length scale.

We have seen that the variance of the concentration in the case of the shear flow decays super-exponentially, i.e., when  $\mathbf{k}_0 = i\mathbf{k}_0$

$$\frac{\langle \theta(\cdot, t)^2 \rangle_{\mathbf{x}}}{\langle \theta(\cdot, 0)^2 \rangle_{\mathbf{x}}} = \exp(-2\kappa k_0^2 t(1 + S^2 t^2)).$$

Focusing on relatively long times ( $St \gg 1$ ) we set  $\varepsilon = \exp(-2\kappa k_0^2 S^2 t_\varepsilon^3) \ll 1$  so that

$$t_\varepsilon = \frac{(\ln \varepsilon)^{\frac{1}{3}}}{(2\kappa k_0^2 S^2)^{\frac{1}{3}}}.$$

Then by (1.33)

$$\kappa_{\text{eff}} = \frac{(2S \ln \frac{1}{\varepsilon})^{\frac{2}{3}} \kappa^{\frac{1}{3}}}{k_0^{\frac{4}{3}}} \quad \text{and} \quad E = \frac{\kappa_{\text{eff}}}{\kappa} = \left( \frac{2S \ln \frac{1}{\varepsilon}}{k_0^2 \kappa} \right)^{\frac{2}{3}}.$$

In the case of pure shear flow there is no natural length scale in the velocity field – it is entirely characterized by the shear strength  $S$  – so we must utilize the length scale in the initial concentration perturbation ( $\ell \sim k_0^{-1}$ ) and use  $U = \frac{S}{k_0}$  to identify the Péclet number  $\text{Pe} = \frac{U\ell}{\kappa} = \frac{S}{k_0^2 \kappa}$ . Then we observe

$$E = \frac{\kappa_{\text{eff}}}{\kappa} \sim \text{Pe}^{\frac{2}{3}} \quad \text{as} \quad \text{Pe} \rightarrow \infty.$$

In this case, since  $2/3 < 1$ ,  $\kappa_{\text{eff}} = \kappa E \rightarrow 0$  as  $\kappa \rightarrow 0$ . That is, there is no residual mixing (by this measure) in the limit of zero molecular diffusivity.

We remark that if there is a definition of effective diffusivity  $\kappa_{\text{eff}}$  and a stirring such that the enhancement factor  $E \sim \text{Pe}$ , then  $\kappa_{\text{eff}} \sim U\ell$  remains non-zero even in the singular limit of zero molecular diffusivity (all other parameters held fixed). In such situations  $\kappa_{\text{eff}}$  would be independent of the molecular diffusivity in the  $\kappa \rightarrow 0$  limit. Then stirring would produce diffusive effects even if molecular diffusion is ostensibly negligible.

The pure strain flow, characterized by a constant rate of strain  $\Gamma$ , dissipates scalar variance faster than shear flow. Indeed, for pure strain flow and  $\mathbf{k}_0 = \mathbf{j}k_0$  the effective diffusion definition in (1.33) yields  $E = \frac{\kappa_{\text{eff}}}{\kappa} \sim \frac{\text{Pe}}{\ln \text{Pe}}$ . To see this recall that the variance for the concentration associated to the pure strain flow is

$$\frac{\langle \theta(\cdot, t)^2 \rangle_{\mathbf{x}}}{\langle \theta(\cdot, 0)^2 \rangle_{\mathbf{x}}} = \exp\left(-\frac{\kappa k_0^2}{\Gamma} (\exp(2\Gamma t) - 1)\right).$$

Again focusing on relatively long times ( $\Gamma t \gg 1$ ) we set  $\varepsilon = \exp(-\kappa k_0^2 \exp(2\Gamma t_\varepsilon)/\Gamma)$  so that  $t_\varepsilon = \frac{1}{2\Gamma} \ln\left(\frac{\ln(\frac{1}{\varepsilon})2\Gamma}{\kappa k_0^2}\right)$  and the effective diffusion definition (1.33) yields

$$\kappa_{\text{eff}} = \frac{\ln(\frac{1}{\varepsilon})\Gamma}{k_0^2 \ln\left(\frac{\ln(\frac{1}{\varepsilon})2\Gamma}{\kappa k_0^2}\right)}.$$

Then with  $\ell = \frac{1}{k_0}$  and  $U = \frac{\Gamma}{k_0}$  so that  $\text{Pe} = \frac{\Gamma}{k_0^2 \kappa}$  we obtain

$$E = \frac{\kappa_{\text{eff}}}{\kappa} \sim \frac{\text{Pe}}{\ln \text{Pe}}.$$

This is only logarithmically less than the  $E \sim \text{Pe}$  scaling necessary to produce finite non-zero residual effective diffusion in the limit of vanishing molecular diffusion.

Later in these lectures we will see that there are definitions of effective diffusivity and stirring flow fields for which  $E \sim \text{Pe}^2$ , meaning that effective diffusion can actually *increase* as molecular diffusivity decreases!

## 1.4 Sources and sinks

### 1.4.1 Internal source-sink distributions

Suppose now that  $\Omega$  is a regular bounded domain of  $\mathbb{R}^n$  and the velocity field  $\mathbf{u}$  satisfies

$$\nabla \cdot \mathbf{u} = 0 \quad \text{and} \quad \mathbf{n} \cdot \mathbf{u} = 0 \quad \text{on } \partial\Omega,$$

where  $\mathbf{n}$  is the outward normal to the domain. Let the scalar field  $T(\mathbf{x}, t)$  satisfy

$$\begin{cases} \dot{T} + \mathbf{u} \cdot \nabla T = \kappa \Delta T + S & \text{in } \Omega \times (0, \infty), \\ \mathbf{n} \cdot \nabla T = 0 & \text{in } \partial\Omega \times (0, \infty), \\ T = 0 & \text{in } \Omega \times \{t = 0\}, \end{cases} \quad (1.35)$$

where  $S = S(\mathbf{x})$  is a steady scalar source distribution.

Averaging in space using the no-flux boundary conditions we have

$$\frac{d}{dt} \langle T(\cdot, t) \rangle_{\mathbf{x}} = \langle S \rangle_{\mathbf{x}},$$

so that  $\langle T(\cdot, t) \rangle_{\mathbf{x}} = \langle S \rangle_{\mathbf{x}} t$ . Therefore it is natural to consider the decomposition

$$T(\mathbf{x}, t) = \langle S \rangle_{\mathbf{x}} t + \theta(\mathbf{x}, t),$$

where the deviation from the spatial average distribution satisfies

$$\dot{\theta} + \mathbf{u} \cdot \nabla \theta = \kappa \Delta \theta + s(\mathbf{x}), \quad (1.36)$$

with  $s(\mathbf{x}) = S(\mathbf{x}) - \langle S \rangle_{\mathbf{x}}$ . Therefore  $\langle s \rangle_{\mathbf{x}} = 0$ ,  $\theta(\mathbf{x}, 0) = 0$  and, subsequently,  $\langle \theta(\cdot, t) \rangle_{\mathbf{x}} = 0$ .

The efficiency of stirring to suppress the variance of the scalar concentration in this case is related to its ability to minimize  $\theta$  (in the mean square sense). The inhomogeneous term  $s(\mathbf{x})$  in (1.36) is signed — it is source in some places and a sink in others — and effective flows will transport tracers from source regions where  $s(\mathbf{x}) > 0$  to sink regions where  $s(\mathbf{x}) < 0$ .

When  $\mathbf{u} = 0$ , (1.36) reduces to the inhomogeneous heat equation

$$\dot{\theta} = \kappa \Delta \theta + s(\mathbf{x}),$$

with steady state  $\bar{\theta}(\mathbf{x}) = -\frac{1}{\kappa} (\Delta^{-1} s)(\mathbf{x})$  where  $\Delta^{-1} = \nabla^{-2}$  denotes the appropriate inverse Laplacian operator. The variance of the scalar concentration in the “unstirred” steady state is

$$\langle \bar{\theta}^2 \rangle_{\mathbf{x}} = \frac{\langle (\Delta^{-1} s)^2 \rangle_{\mathbf{x}}}{\kappa^2}.$$

When the flow is non zero, i.e.,  $\mathbf{u} \neq 0$ , the variance of the passive tracer concentration is

$$\langle\langle \theta^2 \rangle_x \rangle_t = \frac{\langle\langle \theta^2 \rangle_x \rangle_t}{\langle\langle \theta^2 \rangle_x \rangle_t} \langle\langle \bar{\theta}^2 \rangle_x \rangle_t = \frac{\langle\langle (\Delta^{-1}s)^2 \rangle_x \rangle_t}{\kappa_{\text{eff}}^2}, \quad (1.37)$$

where  $\langle \cdot \rangle_t$  is the time average and we have defined an effective diffusion coefficient

$$\kappa_{\text{eff}} := \sqrt{\frac{\langle\langle (\Delta^{-1}s)^2 \rangle_x \rangle_t}{\langle\langle \theta^2 \rangle_x \rangle_t}} = \kappa_{\text{mol}} E,$$

and the *enhancement factor*

$$E := \sqrt{\frac{\langle\langle \bar{\theta} \rangle_x \rangle_t}{\langle\langle \theta^2 \rangle_x \rangle_t}}.$$

The effective diffusion coefficient—and its dimensionless version in terms of the enhancement factor—is a quantitative gauge of how much the stirring contributes to mixing as measured by concentration variance suppression. A basic challenge in the mixing business is to understand, predict and/or control this enhancement in terms of properties of the stirring velocity vector field. Another basic question is to determine the extent to which such effective diffusions or enhancement factors are meaningful as properties of a given flow field. That is, the degree to which  $\kappa_{\text{eff}}$  and  $E$  are properties of  $\mathbf{u}$  alone—or how much they depend on the details of the task that the field is asked to perform e.g. the source-sink distribution  $s(\mathbf{x})$ , is not *a priori* evident.

#### 1.4.2 Source-sink distribution at the boundary

Consider a rectangular box  $\Omega = [0, L] \times [0, h]$  and, as before, suppose that  $\mathbf{u}$  is given and satisfies

$$\nabla \cdot \mathbf{u} = 0 \quad \text{and} \quad \mathbf{n} \cdot \mathbf{u} = 0 \text{ in } \partial\Omega$$

and that the scalar field  $T(\mathbf{x}, t)$  satisfies the homogeneous advection-diffusion equation with inhomogeneous boundary conditions:

$$\left\{ \begin{array}{ll} \dot{T} + \mathbf{u} \cdot \nabla T = \kappa \Delta T & \text{in } [0, L] \times [0, h] \times (0, \infty), \\ \nabla T \cdot \mathbf{n} = 0 & \text{at } y \in \{0, h\}, \\ T = T_0 + \delta T & \text{at } x = 0, \\ T = T_0 & \text{at } x = L. \end{array} \right. \quad (1.38)$$

In the absence of fluid flow, i.e., when  $\mathbf{u} = \mathbf{i}u + \mathbf{j}v = 0$ , then the steady solution is the linear profile  $T_0 + \frac{\delta T}{L}(L - x)$  and the flux of the scalar, a.k.a. the current  $\mathbf{J} \equiv \mathbf{i}J_x + \mathbf{j}J_y$ , is  $\mathbf{J} = \mathbf{i}J_x = \mathbf{i}\kappa G$  where the imposed gradient is  $G = \frac{\delta T}{L}$ . The linear relation between the current and the imposed gradient is known as Fick's law, and the coefficient of proportionality between the flux and the gradient is precisely the diffusion coefficient.

When the flow is non-zero then the local instantaneous current is

$$\mathbf{J}(\mathbf{x}, t) = \mathbf{i}J_x + \mathbf{j}J_y = \mathbf{u}(\mathbf{x}, t)T(\mathbf{x}, t) - \kappa \nabla T$$

and the space-time average of  $J_x$  is

$$\langle\langle J_x \rangle_x \rangle_t = \langle\langle uT - \kappa \partial_x T \rangle_x \rangle_t = \langle\langle uT \rangle_x \rangle_t + \kappa G, \quad (1.39)$$

where in the last step we used the boundary conditions for  $T$ . The term  $\langle\langle uT \rangle_x \rangle_t$  is the correlation between the horizontal velocity and  $T$ .

Again we decompose the function  $T(\mathbf{x}, t)$  into the steady no-flow state and deviation function  $\theta$ :

$$T(\mathbf{x}, t) = T_0 + \frac{\delta T}{L}(L - x) + \theta(\mathbf{x}, t).$$

The deviation satisfies homogeneous boundary conditions,  $\theta = 0$  at  $x = 0$  and  $x = L$ , with a space and time dependent source-sink distribution proportional to the horizontal velocity  $u$ :

$$\begin{cases} \dot{\theta} + \mathbf{u} \cdot \nabla \theta &= \kappa \Delta \theta + Gu & \text{in } [0, L] \times [0, h] \times (0, \infty), \\ \theta &= 0 & \text{at } x \in \{0, L\}. \end{cases} \quad (1.40)$$

Note that for divergence-free  $\mathbf{u}$  and the boundary conditions,  $\langle u\phi \rangle_x = 0$  for all the functions  $\phi$  that depend only on  $x$ . Using this fact we find that the space-time average of  $J_x$  is

$$\langle\langle J_x \rangle_x \rangle_t = \langle\langle u\theta \rangle_x \rangle_t + \kappa G$$

and we can define the *effective diffusivity* as the ratio of the flux and the gradient:

$$\kappa_{\text{eff}} := \frac{\langle\langle J_x \rangle_x \rangle_t}{G} = \frac{\langle\langle u\theta \rangle_x \rangle_t + \kappa G}{G} = \kappa + \frac{\langle\langle u\theta \rangle_x \rangle_t}{G}. \quad (1.41)$$

The *enhancement factor* in this context is

$$E := \frac{\kappa_{\text{eff}}}{\kappa} = 1 + \frac{\langle\langle u\theta \rangle_x \rangle_t}{\kappa G}. \quad (1.42)$$

Testing the equation (1.40) with  $\theta$  and using the incompressibility condition and the boundary conditions, we see that the evolution of the variance of  $\theta$  is

$$\frac{d}{dt} \langle \theta(\cdot, t)^2 \rangle_x = -\kappa \langle |\nabla \theta|^2 \rangle_x + G \langle u\theta \rangle_x.$$

By the maximum principle for the temperature  $T$  (and therefore also  $\theta$ ) is suitably bounded so that

$$\limsup_{t_0 \rightarrow \infty} \frac{1}{t_0} \int_0^{t_0} \frac{d}{dt} \langle \theta(\cdot, t)^2 \rangle_x dt = \limsup_{t_0 \rightarrow \infty} \frac{\langle \theta(\cdot, t_0)^2 \rangle_x - \langle \theta(\cdot, 0)^2 \rangle_x}{t_0} = 0.$$

Therefore  $\kappa \langle\langle |\nabla \theta|^2 \rangle_x \rangle_t = G \langle\langle u\theta \rangle_x \rangle_t$  and we can re-express the enhancement factor

$$E = \frac{\kappa_{\text{eff}}}{\kappa} = 1 + \frac{\langle\langle u\theta \rangle_x \rangle_t}{\kappa G} = 1 + \frac{\langle\langle |\nabla \theta|^2 \rangle_x \rangle_t}{G^2}.$$

This tells us that the rate of transport of the scalar by the flow is proportional to  $\langle\langle |\nabla \theta|^2 \rangle_x \rangle_t$ , and in particular it ensures that  $E \geq 1$ . That is, flow of any sort can only *increase* the boundary-to-boundary scalar flux (on average).

## 1.5 Model stirring as diffusion

In this final section we develop a case study to address the following questions:

- How should we gauge the effectiveness of a flow as a mixer?
- How might we model advection and stirring as diffusion?
- What properties of a flow field make it a good mixer?

We saw that in order to investigate the mixing properties of a flow field we can consider the dispersion of tracer particles moving according to

$$d\mathbf{X}_t = \mathbf{u}(\mathbf{X}_t, t) dt + \sqrt{2\kappa} d\mathbf{W}_t,$$

or we can study the concentration of the tracer particles and study solutions of

$$\partial_t T + \mathbf{u} \cdot \nabla T - \kappa \Delta T = S(\mathbf{x}),$$

where  $S(\mathbf{x})$  is a scalar source distribution.

We might sensibly seek to quantify the mixing efficiency of stirring by replacing the flow field and diffusion with an effective diffusion that achieves the same goal. Symbolically, we seek to make sense of the substitution

$$\mathbf{u} \cdot \nabla - \kappa \Delta \rightarrow -\partial_i \kappa_{ij}^{\text{eff}} \partial_j,$$

and include the effect of stirring as an effective—usually *enhanced*—diffusion. (Not unexpectedly, a diffusion tensor would generally be necessary to model anisotropic flow features.) This approach raises several other fundamental questions:

- Which particular properties of “mixing” do we want to capture in the notion of an effective diffusion?
- Do different criteria for measuring mixing produce qualitatively different effective diffusivities?

Three reasonable ways to parametrize stirring as diffusion are (1) in terms of particle dispersion yielding an effective diffusion coefficient  $\kappa^{\text{PD}}$ , (2) as the ratio of a flux given a gradient in which case  $\kappa^{\text{FG}}$ , or (3) via concentration variance reduction with  $\kappa^{\text{VR}}$ . The second question above may be rephrased, for a given flow field are  $\kappa^{\text{PD}}$ ,  $\kappa^{\text{FG}}$  and  $\kappa^{\text{VR}}$  the same? Or are they different?

In the first case the effective diffusivity is the time prefactor of the square-mean-displacement — at least if the time dependence is eventually linear:

$$\mathbb{E}\{(X_i(t) - X_i(0))(X_j(t) - X_j(0))\} \sim 2\kappa_{ij}^{\text{eff}} t.$$

In the second case we decompose the function  $T$  into a steady profile  $T_0 - Gx$  and a deviation function  $\theta = T - T_0 + Gx$  that satisfies

$$\dot{\theta} + \mathbf{u} \cdot \nabla \theta = \kappa \Delta \theta + G \mathbf{u} \cdot \mathbf{i}, \quad (1.43)$$

where  $\mathbf{u} = u \cdot \mathbf{i}$  and  $\mathbf{i}$  is the horizontal unit vector; recall (1.40). The source-sink here is proportional to the velocity in the direction of the gradient flux and the effective diffusivity is defined

$$\kappa_{11}^{\text{eff}} := \kappa + \frac{\langle \langle u \theta \rangle_x \rangle_t}{G}. \quad (1.44)$$

As shown before, testing (1.43) with  $\theta$  and integrating by parts implies this is

$$\kappa_{11}^{\text{eff}} = \kappa \left( 1 + \frac{\langle \langle |\nabla \theta|^2 \rangle_x \rangle_t}{G^2} \right). \quad (1.45)$$

In the third case we look at the deviation of the concentration variance in presence of a steady source-sink distribution,  $\partial_t \theta + \mathbf{u} \cdot \nabla \theta = \kappa \Delta \theta + s(\mathbf{x})$ , and define

$$\kappa^{\text{eff}} := \sqrt{\frac{\langle \langle (\Delta^{-1} s)^2 \rangle_x \rangle_t}{\langle \langle \theta^2 \rangle_x \rangle_t}}. \quad (1.46)$$

This measure of effective diffusion expresses how much a flow reduces the scalar variance that is sustained by the inhomogeneous sources and sinks.

The Péclet number  $Pe$  is the non-dimensional parameter that measures the strength of the fluid flow relative to the molecular diffusion (see (1.34)). Large values of the Péclet number, i.e.,  $Pe \gg 1$ , means that the stirring is strong while low values of Péclet,  $Pe \ll 1$ , indicate strong molecular diffusion and/or weak stirring. The goal of theory and analysis — and also often computation, simulation and/or experimentation — is to determine the enhancement factor as a function of the Péclet number. We are particularly interested in the limit  $\kappa \rightarrow 0$  to see in which cases the effective diffusivity might become independent of the molecular diffusivity. But here a dilemma occurs. There are flows for which

$$E(Pe) = \frac{\kappa^{\text{PD}}}{\kappa} = \frac{\kappa^{\text{FG}}}{\kappa} \sim Pe^2 \quad \text{as } Pe \rightarrow \infty,$$

while on the other hand [1] in terms of variance reduction

$$E(Pe) = \frac{\kappa^{\text{VR}}}{\kappa} \leq Pe^1 \quad \text{as } Pe \rightarrow \infty.$$

The remainder of these lectures will demonstrate these facts and discuss resolution of the quandary.

To see the latter, consider a periodic box of size  $L$  with  $\theta$  satisfying

$$\dot{\theta} + \mathbf{u} \cdot \nabla \theta = \kappa \Delta \theta + s(\mathbf{x}).$$

Multiply by a smooth test function  $\phi(\mathbf{x})$  and average in  $\mathbf{x}$  and  $t$ . Take the long time average in  $t$  (exercise: the first term vanishes by the suitable boundedness of  $\theta$ ) to obtain

$$\begin{aligned}\langle s\phi \rangle_{\mathbf{x}} &= -\langle \langle \theta(\mathbf{u} \cdot \nabla \phi + \kappa \Delta \phi) \rangle_{\mathbf{x}} \rangle_t \\ &\leq \langle \langle \theta^2 \rangle_{\mathbf{x}} \rangle_t^{\frac{1}{2}} \langle \langle (\mathbf{u} \cdot \nabla \phi + \kappa \Delta \phi)^2 \rangle_{\mathbf{x}} \rangle_t^{\frac{1}{2}}\end{aligned}$$

using Cauchy-Schwarz in the last step. Then according to definition (1.46), for any test function  $\phi$

$$E = \frac{\kappa^{\text{eff}}}{\kappa} \leq \frac{1}{\kappa} \frac{\langle \langle (\Delta^{-1}s)^2 \rangle_{\mathbf{x}} \rangle_t^{\frac{1}{2}} \langle \langle (\mathbf{u} \cdot \nabla \phi + \kappa \Delta \phi)^2 \rangle_{\mathbf{x}} \rangle_t^{\frac{1}{2}}}{\langle s\phi \rangle_{\mathbf{x}}}. \quad (1.47)$$

To get the best upper bound, we would consider the infimum over all test functions

$$E \leq \inf_{\phi} \left\{ \frac{1}{\kappa} \frac{\langle \langle (\Delta^{-1}s)^2 \rangle_{\mathbf{x}} \rangle_t^{\frac{1}{2}} \langle \langle (\mathbf{u} \cdot \nabla \phi + \kappa \Delta^{-1}\phi)^2 \rangle_{\mathbf{x}} \rangle_t^{\frac{1}{2}}}{\langle s\phi \rangle_{\mathbf{x}}} \right\}, \quad (1.48)$$

but for our purposes it suffices to choose  $\phi = \Delta^{-2}s$ . (For some alternative analyses see [2].) When  $\kappa$  is small the dominant term on the right hand side is  $\langle \langle (\mathbf{u} \cdot \nabla \phi) \rangle_{\mathbf{x}} \rangle_t$  and, since  $\frac{\nabla^{-3}s}{\Delta^{-1}s}$  has the unit of a length, we can employ it as  $\ell$  to conclude that  $E \lesssim \frac{U\ell}{\kappa} = \text{Pe}^1$ , where  $U$  is the root mean square stirring speed.

We reiterate that in this case the “suitable” length-scale  $\ell$  employed in the definition of the Péclet number comes from the source-sink distribution function.

### 1.5.1 Single scale source-sink distribution

We now focus on the most fundamental example of a single scale source-sink distribution being stirred by a single-scale flow and explicitly compute some effective diffusions. Consider tracer particles diffusing in a steady sinusoidal shear flow satisfying the coupled stochastic (Itô) equations

$$\begin{aligned}dX_t &= \sqrt{2}U \sin(k_u Y_t) dt + \sqrt{2\kappa} dW_t^{(1)}, \\ dY_t &= \sqrt{2\kappa} dW_t^{(2)},\end{aligned}$$

where  $W_t^{(1)}$  and  $W_t^{(2)}$  are independent Wiener processes.

One the one hand, for a source-sink distribution  $s(\mathbf{x}) \sim \sin k_s x$ , it was shown in [3] that the effective diffusion defined in terms of variance reduction according to (1.46) is actually  $\kappa^{\text{VR}} \sim \kappa \left(\frac{k_u}{k_s}\right)^{1/3} \text{Pe}^{5/6}$ . (Here  $\text{Pe} = U/\kappa k_s$  using the length scale in the source-sink distribution.) This corresponds to an enhancement factor  $E \lesssim \text{Pe}^1$  in accord with the theorem.

On the other hand for the flux-gradient relation we need to compute the steady state solution of (1.43) with velocity field of the form  $\mathbf{u} = \sqrt{2}U \sin(k_u y)\mathbf{i}$ :

$$\partial_t \theta + \sqrt{2}U \sin(k_u y) \partial_x \theta = \kappa \Delta \theta + G \sqrt{2}U \sin(k_u y).$$



The steady-state solution is

$$\theta_s(y) = \frac{\sqrt{2}Gu}{\kappa k_u^2} \sin(k_u y),$$

and the enhancement factor can be easily computed by the definition (see (1.45)):

$$E = \frac{\kappa^{FG}}{\kappa} = 1 + \frac{\langle \langle |\nabla \theta_s|^2 \rangle_x \rangle_t}{G^2} = 1 + \text{Pe}^2.$$

Moreover, we can compute the (enhanced) particle dispersion in the  $x$  direction exactly. Noting that

$$dX_t^2 = 2X_t dX_t + \frac{1}{2} 2 (dX_t)^2,$$

and  $\mathbb{E}((dX_t)^2) = 2\kappa \mathbb{E}((dW_t^{(1)})^2) + o(dt) = 2\kappa dt$ , taking the expectation value of the expression above, we have

$$\mathbb{E}(dX_t^2) = 2\mathbb{E}(X_t dX_t) + 2\kappa dt = 2\sqrt{2} U \mathbb{E}(X_t \sin(k_u Y_t)) dt + 2\kappa dt.$$

In order to compute the term  $\mathbb{E}(X_t \sin(k_u Y_t))$  we use again the Itô formula to compute

$$\begin{aligned} d(X_t \sin(k_u Y_t)) &= \sin(k_u Y_t) dX_t + X_t (k_u \cos(k_u Y_t)) dY_t - \frac{1}{2} X_t k_u^2 \sin(k_u Y_t) 2\kappa dt \\ &= \sin(k_u Y_t) (\sqrt{2} U \sin(k_u Y_t) dt + \sqrt{2\kappa} dW_t^{(1)}) \\ &\quad + X_t (k_u \cos(k_u Y_t)) \sqrt{2\kappa} dW_t^{(2)} - \frac{1}{2} X_t k_u^2 \sin(k_u Y_t) 2\kappa dt. \end{aligned}$$

Taking the expectation value of the last expression we have

$$\mathbb{E}(d(X_t \sin(k_u Y_t))) = \sqrt{2} U \mathbb{E}(\sin^2 k_u Y_t) dt - \kappa k_u^2 \mathbb{E}(X_t \sin(k_u Y_t)) dt.$$

Since  $\mathbb{E}(\sin^2(k_u Y_t)) = \frac{1}{2}$ , we are left to solve

$$\mathbb{E}(d(X_t \sin(k_u Y_t))) = d\mathbb{E}(X_t \sin(k_u Y_t)) = (U/\sqrt{2}) dt - \kappa k_u^2 \mathbb{E}(X_t \sin(k_u Y_t)) dt.$$

The solution is

$$\mathbb{E}(X_t \sin(k_u Y_t)) = \exp(-\kappa k_u^2 t) + U/(\sqrt{2} \kappa k_u^2).$$

In the long-time limit  $\mathbb{E}(X_t \sin(k_u Y_t)) \rightarrow \frac{U}{\sqrt{2\kappa k_u^2}}$ , so the variance of  $X_t$  satisfies

$$\mathbb{E}(dX_t^2) = d\mathbb{E}(X_t^2) \rightarrow \frac{2U^2}{\kappa k_u^2} dt + \kappa dt$$

and

$$\mathbb{E}(X_t^2) \rightarrow 2 \left( \frac{U^2}{\kappa k_u^2} + 2\kappa \right) t.$$

Thus  $\kappa_{11}^{\text{eff}} = \kappa + \frac{U^2}{\kappa k_u^2}$ , and using  $\ell \sim k_u^{-1}$  as the reference length-scale so that  $\text{Pe} = \frac{U^2}{\kappa k_u^2}$ , we have  $E = \frac{\kappa_{11}^{\text{eff}}}{\kappa} = \frac{\kappa^{PD}}{\kappa} = 1 + \text{Pe}^2$ .

The origin of the different behavior of the different definitions of effective diffusion — i.e.,  $E \sim \text{Pe}^1$  for variance reduction versus  $E \sim \text{Pe}^2$  for the flux-gradient and particle dispersion measures — resides in a combination of the different choices for the length scale  $\ell$  and consideration of the different times scales over which the mixing processes take place.

Indeed, it takes time  $\sim (\kappa k_u^2)^{-1}$  for tracers to develop an effective diffusivity  $\sim \text{Pe}^2$  since trajectories must diffuse across various shear “streams” before the flow-induced dispersion enhancement to emerge. If this time is as long or longer than the time it takes for the flow to transport tracer from sources to sinks, i.e.,  $(Uk_s)^{-1}$ , then then  $\text{Pe}^2$  enhancement is irrelevant for variance reduction. That is,  $E = \mathcal{O}(\text{Pe}^2)$  enhancement is only possible when  $\text{Pe} < \frac{k_u}{k_s}$  (where the Péclet number is defined with  $\ell \sim k_u^{-1}$ ).

Ultimately the difference arises due to fundamental physical features of *transient* mixing, such as might be measured by tracer particle dispersion, and *transport* properties of a flow that are relevant to steady state variance suppression in the presence of sources and sinks. For further discussion see [4].

**Acknowledgment:** The authors would like to thank Anna Mazzucato and Gianluca Crippa for organizing the very interesting Summer School in Levico Terme, and the Fondazione Bruno Kessler und the CIRM for their administrative support and hospitality. Much of the research reported here was supported in part by US National Science Foundation Division grants including, most recently, PHY-1205219 and DMS-1515161. The first author (CRD) gratefully acknowledges support as a Simons Fellow in Theoretical Physics and as a Fellow of the John Simon Guggenheim Foundation during, respectively, the presentation of these lectures and preparation of these notes.

## Appendix: Itô Formula

The Wiener process is a Gaussian distribution with  $\mathbb{E}(W_t) = 0$ ,  $\mathbb{E}(W_t^2) = t$  and  $\mathbb{E}(W_t, W_s) = \min(t, s)$ . Consider the independent increments

$$\Delta W_t = W_{t+\Delta t} - W_t,$$

and compute their expectation values:

$$\begin{aligned} \mathbb{E}(\Delta W_t^2) &= \mathbb{E}(W_{t+\Delta t} - W_t)^2 \\ &= \mathbb{E}(W_{t+\Delta t}^2 - 2W_{t+\Delta t}W_t + W_t^2) \\ &= t + \Delta t - 2t + t = \Delta t. \end{aligned} \tag{1.49}$$

Therefore,

$$\mathbb{E}(dW_t^2) = dt.$$

Now consider a stochastic process defined by

$$dX_t = f(X_t)dt + g(X_t)dW_t,$$

where  $W_t$  is a Wiener process, interpreted via the forward Euler scheme

$$\Delta X_t = X_{t+\Delta t} - X_t = f(X_t)\Delta t + g(X_t)\Delta W_t.$$

Note that  $\Delta W_t$  is independent of  $X_t$ .

Let us compute the variance of the increment  $\Delta X$ . We have

$$\begin{aligned}\mathbb{E}(\Delta X_t^2) &= \mathbb{E}(f(X_t)^2(\Delta t)^2) + 2\mathbb{E}(f(X_t)g(X_t)\Delta W_t\Delta t) + \mathbb{E}(g(X_t)^2\Delta W_t^2) \\ &= \mathbb{E}(f(X_t)^2)(\Delta t)^2 + \mathbb{E}(g(X_t)^2)\Delta t \\ &\approx \mathbb{E}(g(X_t)^2)\Delta t.\end{aligned}\tag{1.50}$$

The Itô formula states that the differential of any function  $F$  of the process  $X_t$  is

$$dF(X_t) = F'(X_t)dX_t + \frac{1}{2}F''(X_t)(dX_t)^2,\tag{1.51}$$

keeping only the  $\mathcal{O}(dt)$  terms, i.e., those in  $dX_t$  and using  $(dX_t)^2 \approx g(X_t)^2 dt$ .

For example we use the Itô formula (1.51) applied to the function  $F : X_t \rightarrow X_t^2$  in order to compute the variance of stochastic processes. Namely,

$$d(X_t^2) = 2X_t dX_t + \frac{1}{2}2(dX_t)^2.$$

Taking the expectation value of (1.7) we have

$$\mathbb{E}(dX_t^2) = 2\mathbb{E}(X_t dX_t) + \mathbb{E}(dX_t)^2.$$

## Bibliography

- [1] Jean-Luc Thiffeault, Charles R. Doering, and John D. Gibbon. A bound on mixing efficiency for the advection–diffusion equation. *Journal of Fluid Mechanics*, 521:105–114, 2004.
- [2] Charles R. Doering and Jean-Luc Thiffeault. Multiscale mixing efficiencies for steady sources. *Physical Review E*, 74:025301R, 2006.
- [3] Tiffany A. Shaw, Jean-Luc Thiffeault, and Charles R. Doering. Stirring up trouble: Multi-scale mixing measures for steady scalar sources. *Physica D*, 231:143–164, 2007.
- [4] Zhi Lin, Katarína Bodová, and Charles R. Doering. Models and measures of mixing and effective diffusion. *Discrete and Continuous Dynamical Systems*, 28:259–274, 2010.