RANDOM WALKS, DIFFUSION AND OTHER STATISTICAL MODELS
The familiar continuum equations of motions are actually statistical models and deserve examination from this point of view. They are tremendously successful in reducing the information required to describe fluid dynamics. There is also one interesting fundamental aspect of these equations, time irreversibility, which can be understood only within the statistical context.

1. Irreversibility

Consider the foundations of the familiar Navier-Stokes equations used to describe viscous fluid motion. Imagine a fluid composed of N molecules, each identified by the subscript n, with position $\mathbf{r}$, velocity $\mathbf{v}$ and identical masses m. These molecules obey an equation of motion of the form

$$m\frac{d^2}{dt^2}\mathbf{r}_n \;=\; m\frac{d}{dt}\mathbf{v}_n \;=\; \mathbf{F}(\mathbf{r}_n|\mathbf{r}_1,\mathbf{r}_2,.....,\mathbf{r}_N) \tag{1}$$

where $\mathbf{F}$ describes the intermolecular force field. There are something like $N = 10^{48}$ molecules in the world ocean so clearly some reduction in complexity is required. To do this the fundamentally statistical concept of a fluid continuum is introduced. It is based on the hypothesis that all the particles within a macroscopic volume are statistically equivalent; the volume then defines an ensemble and continuum properties are the statistics of the molecules in the volume. Thus the continuum velocity is defined as

$$\mathbf{u}(\mathbf{x}, t) \;=\; \sum_{vol(\mathbf{x})} m\mathbf{v}_n \;/\; \sum_{vol(\mathbf{x})} m \tag{2}$$

where $vol(\mathbf{x})$ is a small volume surrounding $\mathbf{x}$. Thus $\mathbf{u}$ is the average particle velocity near $\mathbf{x}$. In the following section we will examine the statistical mechanics approach to deriving evolution equations for statistical variables like $\mathbf{u}$. Whether this is done with statistics or heuristic arguments about scale separation, the result (*cf.* Batchelor, An Introduction to Fluid Dynamics, 3.3) are the Navier-Stokes and advection-diffusion equations

$$\frac{d}{dt}\mathbf{u} \;=\; -\frac{1}{\rho}\nabla p \;+\; \nabla\cdot\nu\nabla\mathbf{u} \qquad\qquad \frac{d}{dt}\theta \;=\; \nabla\cdot\alpha\nabla\theta \tag{3}$$

where ρ is the density (from the average number of particles in $vol(\mathbf{x})$) while ν and α are the kinematic viscosity and diffusivity.

Comparison of (1) and (3) shows immediately that something fundamental has been changed in the continuum description. The original equation is strictly reversible with respect to changing the sense of time. If the velocities of all particles were reversed at the end of an experiment he system would return to its initial state. The continuum equations, in contrast, have a definite sense of time since kinetic energy and θ^2 are lost as time increases. The origin of this paradox is that the continuum equations describe the spatial average $\mathbf{u}$. While the mean square velocity of all particles (molecular kinetic energy) may be conserved the mean square of the average velocity is not conserved. Dissipation of continuum energy is associated with generation of heat which increases the molecular velocity fluctuations around the mean $\mathbf{u}$, thus allowing $|\mathbf{u}|^2$ to decrease while conserving total molecular kinetic energy. Time irreversibility in continuum equations results from the fact that these are equations for averages; it is closely related to the sense of time flow provided by the Second Law of Thermodynamics (the Thermodynamic laws are also statistical models).

The smallest scales of motion in the ocean are of the order one millimeter. If this is used to define the continuum averaging volume, the complexity of the ocean is reduced to something like 10^{28} degrees of freedom. Unfortunately, extending the kind of averaging used to develop the

continuum equation to further reduce complexity does not work well. The conventional approach is to imagine that an experiment can be repeated many times and to define the ensemble average velocity as

$$\langle \mathbf{u}(\mathbf{x}, t) \rangle = \frac{1}{N} \sum_n^N \mathbf{u}_n(\mathbf{x}, t) \,, \qquad \mathbf{u}' = \mathbf{u} - \langle \mathbf{u} \rangle \tag{4}$$

where each subscript n refers to one repeated realization of the flow in which time t is reinitialized at the experiment start. Substitution of (4) into (3) followed by accurate use of the linear operator defining the average $\langle \cdot \rangle$ and noting $\nabla \cdot \mathbf{u} = 0$ yields

$$\partial_t \langle \mathbf{u} \rangle + \langle \mathbf{u} \rangle \cdot \nabla \langle \mathbf{u} \rangle = -\frac{1}{\rho} \nabla \langle p \rangle - \nabla \cdot \nu \cdot \nabla \langle \mathbf{u} \rangle - \nabla \cdot \langle \mathbf{u}' \mathbf{u}' \rangle \tag{5}$$

The Reynolds stress appearing on the right of (5) is something like the viscous stress in the continuum equation in that it is an average stress produced by variations of velocity about the average. Unfortunately, it has not been possible to "close" this equation with an appropriate statistical relation between this stress and the mean velocity.

In fluid dynamics courses the derivations of the molecular flux and stress laws are usually based on an experimentally supported hypothesis that these fluxes are linearly related to first derivatives of the associated potential (*e.g.* temperature for heat flux). Actually, we already have the tools to develop a more dynamically satisfying model based on molecular random walks. The term "random walk" refers to a class of processes which is important in physics and which is central to understanding how sample averages converge as the number of samples increases. It is a model of the molecular dynamics which lead to diffusive transport.

2. Discrete random walks

Consider the one-dimensional motion of "molecules" which are bumped every time interval Δ by intermolecular collisions and take on new velocities which are maintained over the next time interval. Define times $t_n = n\Delta$ to be the times when collisions occur, $X(t_n)$ be the particle position at time t_n and $V(t_n)$ be the velocity over the interval $t_{n-1} < t < t_n$. The position of a particle is then

$$X(t_n) = X(t_{n-1}) + \Delta V(t_n) = X(0) + \Delta \sum_{m=1}^n V(t_m) \tag{6}$$

This completely deterministic process must be treated as a random one if each $V(t_m)$ is not known. It is still possible, however, to say a great deal about the behavior of "molecules" if the statistics of V are known. The statistics of X are defined by imagining that the process (6) is carried out many times starting with the same point, $X(0) = 0$ say, but with a new sequence of V's in each realization. The collection of such realizations is the ensemble over which probability and averages are defined. A random walk is the special case of this process when $V(t_m)$ has stationary (independent of the absolute time t_m statistics. A strict random walk also requires $V(t_n)$ to be statistically independent of $V(t_m)$ for all $n \neq m$. Because statistics are stationary the mean velocity

$$\langle V(t_m) \rangle = \frac{1}{N} \sum_n^N V_n(t_m) = \langle V \rangle \tag{7}$$

is a constant which here will be taken to vanish.

It is easy to examine the typical behavior of X because the averaging operator is linear and passes through other linear operators. For example, the mean position and variance are

$$\langle X(t_n) \rangle = X(0) + \Delta \sum_{m=1}^n \langle V(t_m) \rangle = n\Delta V = 0 \tag{8}$$

$$\langle X^2(t_n)\rangle = \Delta^2 \sum_{m=1}^{n} \sum_{k=1}^{n} \langle V(t_m)\, V(t_k)\rangle. \tag{9}$$

Because V has stationary statistics, its covariance depends only on the difference of the two associated times:

$$\langle V(t_m)\, V(t_k)\rangle = C_{VV}(t_k - t_m) = \langle V(t_{k-m})\, V(0)\rangle \tag{10}$$

It follows from the requirement that successive values of V are independent that $C_{VV}(t_{n-m}) = \langle V^2\rangle\delta(n,m)$ so that

$$\langle X^2(t_n)\rangle = \langle V^2\rangle\Delta^2 \sum_{m=1}^{n} \sum_{k=1}^{n} \delta(k,m) = \langle V^2\rangle\Delta^2 \cdot n = \langle V^2\rangle\Delta \cdot t_n \tag{11}$$

and the variance of displacement grows linearly with t. The analog of the molecular diffusivity is (see below)

$$\kappa = \frac{1}{2}\frac{d\langle X^2\rangle}{dt} = \frac{1}{2}\Delta\langle V^2\rangle \ \text{ so } \ \langle X^2\rangle = 2\kappa t. \tag{12}$$

The statistics of $X(t)$ are clearly not stationary so that the probability density of position, $F_{X(t)}(r)$, depends on t. Because X is a sum, the Central Limit Theorem shows the probability density to be Gaussian:

$$F_{X(t)}(r) = [2\pi\langle X^2\rangle]^{-1/2} \ \exp[-r^2/2\langle X^2\rangle] \tag{13}$$

From this pdf it is straightforward to find the mean concentration of some tracer released at $r = 0$. According to the definition of a pdf, the probability of $r < X(t) < r + dr$ is $F_{X(t)}(r)\, dr$. Thus if M is the mass of tracer released in each realization and the variance X^2 is $2\kappa t$, then the *average* concentration of tracer is

$$\Gamma(x,t) = M\, [4\pi\kappa t]^{-1/2} \ \exp[-r^2/4\kappa t] \tag{14}$$

where $M = \int dx\, \Gamma$. Two- and three-dimensional analogs are simple extensions of (13–14).

Direct calculation shows that Γ evolves according to the diffusion equation. The motivation for extending this to the advection-diffusion equation involves one main hypothesis: the molecular processes occur over such small scales that the diffusive fluxes are not changed by advective shear. It is the analog of this hypothesis that causes trouble when one applies the same arguments to turbulence.

A general result that follows from the example above is that the variance of a sum of independent, equally distributed numbers grows as the number of elements in the sum and as that number becomes large the sum has a normal distribution. Since averages, which are defined in terms of infinite sums, are usually approximated by finite sums, this result is central to estimating the accuracy of sample (realizable) statistics.

3. Continuous random walks

To see how molecular diffusion differs from turbulent transport, consider the differential limit to (6), $\frac{dX}{dt} = V(t)$, still with $X(0) = 0$ and the restriction of stationarity on V. In this case (11) becomes

$$\frac{1}{2}\frac{d}{dt}\langle X^2\rangle = \langle V(t)X(t)\rangle = \int_0^t \langle V(t)V(\tau)\rangle d\tau = \int_0^t C_{VV}(\tau - t)d\tau \tag{15}$$

$$\langle X^2(t)\rangle = 2\int_0^t d\tau \int_0^\tau d\tau'\, C_{VV}(\tau') = 2\int_0^t d\tau(t-\tau)C_{VV}(\tau) \tag{16}$$

Note that the growth of $\langle X^2 \rangle$ comes from a correlation of V and X. These are correlated because X includes a contribution from recent velocities which are correlated with present velocity.

It is easy to see how $\langle X^2 \rangle$ varies in the limits of short time and long time. From (16) it is clear that for times much less than the scale on which C_{VV} varies, *i.e.* for which C_{VV} is essentially constant,

$$\langle X^2(t) \rangle \approx V^2 \, t^2 \tag{17}$$

On the other extreme the velocity covariance is expected to vanish at time lags long compared with the **decorrelation time**. From (15) it is clear that for times large compared with the decorrelation time

$$\langle X^2(t) \rangle \approx 2t \int_0^\infty d\tau \, C_{VV}(\tau) \tag{18}$$

In accord with (12), a diffusivity can be defined as

$$\kappa(t) = \frac{1}{2} \frac{d\langle X^2(t) \rangle}{dt} = \int_0^t d\tau \, C_{VV}(\tau), \quad \kappa^\infty = \int_0^\infty d\tau \, C_{VV}(\tau). \tag{19}$$

According to (19) dispersion results from a random walk, possibly one with serially uncorrelated velocity increments. Even if the time over which velocity is correlated approaches zero there is dispersion with the asymptotic diffusivity given by κ^∞ (19). Does this mean that diffusion can be caused by arbitrarily short time scales? The covariance is related to the spectrum:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^\infty dt \, C(t) \, \exp(-i\omega t) \tag{20}$$

from which we see that the diffusivity is

$$\kappa^\infty = \frac{1}{2} \int_{-\infty}^\infty dt \, C_{VV}(t) = \pi \, S_{VV}(\omega = 0). \tag{21}$$

This shows that diffusion is caused by the lowest frequency of variability.

The classical random walk velocity covariance has an infinitesimal correlation scale and, consequently, corresponds to a white spectrum. Perhaps for this reason, the random walk is sometimes said to be associated with velocities of arbitrarily short time scale. This is clearly not quite the whole story since it is the low (zero) frequency component of particle velocity which leads to diffusion. In the same way it will later be shown that that the error in estimating the mean of $Y(t)$ as $\frac{1}{T} \int_0^T dt \, Y(t)$ comes from variability of Y with frequencies of $O(1/T)$.

The relations (17–19) play an important role in our understanding of dispersion of material by eddy-like motions in the ocean. The original discussion of these equations was by G.I. Taylor in the 1921 Proc. London Math. Soc. **20**. Freeland, Rhines and Rossby discussed their application to the ocean in J. Marine Res. **33** in 1975. For times long compared with the decorrelation time of V the variance $\langle X^2 \rangle$ grows approximately as $2\kappa^\infty t$ and, since X is the integral of V over many decorrelation times, the pdf should still be Gaussian. In the limit of large times compared with the decorrelation time we still expect the mean concentration to evolve according to a diffusion equation if there is no advection.

4. Problems
(1) When is a particle's position X normally distributed? If $\langle X \rangle = Ut$ and $\langle X'^2 \rangle = 2\kappa t$ then when will the mean concentration of particles whose position is X obey the advection-diffusion equation?

(2) Suppose particle motion evolves according to

$$X(n) \; = \; X(n-1) \; + \; \Delta V(n) \,, \qquad V(n) \; = \; (1-\alpha) \, V(n-1) \; + \; R(n) \qquad (22)$$

where R is independent of $V(n)$, R has stationary statistics and is serially uncorrelated. Find the general solution for $V(n)$. Use it to find the diffusivity $\kappa^{\infty} \; = \; \lim_{t\to\infty} \frac{1}{2} \frac{d}{dt} \langle X^2(t) \rangle$ in terms of α and $\langle R^2 \rangle$. Will the concentration of $X-$particles obey a diffusion equation? Why?

Answers to Random Walk problems

(1) When the position is the sum of many independent steps (discrete or continuous) the pdf of X will be normal. In this case the mean tracer concentration, which differs from the pdf only by a normalization, will be of the form

$$\Gamma(x,t) = \frac{Q}{[2\pi 2\kappa t]^{1/2}} \exp\left[-\frac{(x-Ut)^2}{4\kappa t}\right] \tag{23}$$

where Q is the total amount of tracer. Direct calculation shows this to evolve according to $\partial_t \Gamma + U \cdot \partial_x \Gamma = \kappa \partial_{xx}^2 \Gamma$.

(2) The stated equation of motion and its solution are

$$V(n) = (1-\alpha) \cdot V(n-1) + R(n) = \sum_{p=0}^{n} R(n-p) \cdot (1-\alpha)^p \tag{24}$$

The covariance of the nth and m^{th} velocity is

$$C_{VV}(n,m) = \langle V(n)V(m)\rangle = \sum_{p=0}^{n}\sum_{p=0}^{m} \langle R(n-p)R(m-q)\rangle \tag{25}$$

and since the different $R(n)$ are uncorrelated (taking $n \geq m$)

$$C_{VV}(n,m) = \sum_{q=0}^{m}(1-\alpha)^{|n-m|+2q}\langle R^2\rangle = (1-\alpha)^{|n-m|}\langle R^2\rangle \sum_{q=0}^{m}(1-\alpha)^{2q} \tag{26}$$

and that as n and m approach infinity

$$C_{VV}(n,m) = (1-\alpha)^{|n-m|}\langle R^2\rangle \sum_{q=0}^{\infty}(1-\alpha)^{2q} = (1-\alpha)^{|n-m|}\langle R^2\rangle \frac{1}{1-(1-\alpha)^2} \tag{27}$$

The particle position is the sum $X(n) = \sum_{p=0}^{n} V(p)$ so that the position variance is

$$X(n)^2 = \sum_{p=0}^{n}\sum_{q=0}^{n}[\langle V(p)V(q)\rangle = C_{VV}(p,q) = C_{VV}(p-q)] = \sum_{p=-n}^{n}(n-p)\cdot C_{VV}(p) \tag{28}$$

As shown above, as $p \to \infty$ the covariance $C_{VV}(p) \to 0$ so that as $n \to \infty$

$$\langle X(n)^2\rangle \to n\sum_{p=-\infty}^{\infty} C_{VV}(p) = n\frac{\langle R^2\rangle}{\alpha(2-\alpha)}\left[\sum_{-\infty}^{\infty}(1-\alpha)^p = 2\sum_{p=0}^{\infty}(1-\alpha)^p - 1 = \frac{2}{\alpha} - 1\right] = n\frac{\langle R^2\rangle}{\alpha^2} \tag{29}$$

Thus even though the successive steps are serially correlated, the large time behavior of $\langle X(n)^2\rangle$ is linear in time.