J. Fluid Mech. (2000), *vol.* 411, *pp.* 165–185. Printed in the United Kingdom

Intermittent mixing in strongly stratified fluids as a random walk

By J. VANNESTE† AND P. H. HAYNES

Centre for Atmospheric Science, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 9EW, UK

(Received 23 November 1998 and in revised form 28 September 1999)

In strongly stratified geophysical fluids such as the stratosphere and the ocean, the vertical mixing of tracers is largely due to patches of turbulence that are intermittent in time and space. Heuristic models for this type of mixing are studied which extend that of Dewan (1981*a*). The recognition that, in these models, fluid particles undergo continuous-time random walks allows the derivation of closed-form results for the particle-position statistics. The particle dispersion is shown generally to be diffusive in the long-time limit. However, the early-time, non-diffusive regime is also analysed, since a time-scale estimate indicates its practical importance, in particular for stratospheric mixing.

Because the restratification of fluid patches previously homogenized by turbulence takes a finite time, the probability for a fluid region to become turbulent may depend on the time elapsed since it has last been turbulent. This introduces a 'memory effect' whose consequences for the tracer mixing are analysed in detail using a simple non-Markovian model.

The heuristic models studied allow the large-scale dispersive effects of the turbulent patches to be inferred from the properties of individual patches. This highlights those properties that might most usefully be determined from investigations of the dynamics of the turbulent patches themselves.

1. Introduction

The fundamental problem of tracer stirring and mixing has a particular importance in geophysical flows, since it is central to such phenomena as atmospheric photochemistry and oceanic water-mass modification. Although stirring, i.e. differential advection by large-scale flows, can efficiently reduce the typical tracer scales, small-scale mixing is essential since it is responsible for tracer homogenization at a molecular level. This mixing is ultimately accomplished by molecular diffusion, but in geophysical flows the effect of molecular diffusion is considerably enhanced by the presence of three-dimensional turbulence. The density stratification that characterizes geophysical fluids strongly constrains vertical motions and hence deeply influences the nature of turbulence and mixing (e.g. Csanady 1973). This influence is particularly important for the two strongly stratified fluids that motivate the present study – the stratosphere and the ocean below the mixed layer.

Turbulence in strongly stratified fluids is intrinsically heterogeneous in the vertical: it is concentrated in thin, pancake-like layers separated by regions of strong density

† Present address: Department of Mathematics and Statistics, University of Edinburgh, James Clerk Maxwell Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, UK.

gradients (Fernando 1991; Barenblatt 1996; Kimura & Herring 1996, and references therein). This heterogeneity exists in simple settings in which energy is supplied more or less uniformly while a large-scale density gradient is maintained; it is much more pronounced, however, when the energy supply is intermittent. The latter situation is the relevant one for the stratosphere and the ocean where turbulence is believed to be generated by localized breaking of internal gravity waves and localized shear instability. Because of the intermittence of the generation mechanisms, stratospheric and oceanic turbulence can be regarded as a discontinuous process in which isolated patches of turbulence appear, and subsequently decay, at various instants and locations depending on the details of the internal-wave field and large-scale flow (Barenblatt 1996, and references therein).

Describing the mixing that results from such turbulence is particularly challenging. Turbulent patches have a complex life cycle, and their mixing properties are not well understood (Fernando 1991). The turbulence generation mechanisms are also complex and particularly difficult to model. A description of the mixing due to internal-gravity-wave breaking, for instance, requires a satisfactory description not only of the breaking itself, but also of the evolution of the wave field in the environment of the mixing region. In view of this complexity, it is clear that heuristic approaches based on phenomenological representations of the turbulence may be best suited to modelling the mixing in the stratosphere and in the ocean. Such an approach has been employed by Dewan (1981*a*): concentrating on the vertical transport, he proposed a simple one-dimensional model of mixing in strongly stratified fluids and studied its applicability to the stratosphere (see also Dewan 1980, 1981*b*). Dewan's model assumes that turbulent patches of random thicknesses form at random altitudes; only the net effect of each mixing event (i.e. the complete life cycle of a turbulent patch) is considered and is assumed to result in the perfect mixing of tracer within the patch.† Dewan (1980, 1981*a*) performed numerical simulations which showed that his model of turbulence leads to the vertical diffusion of tracer with a well-defined diffusivity.

It is intuitively clear that in Dewan's model particles undergo some sort of random walk in the vertical. One of the purpose of the present paper is to make this point precise. Specifically, we show that the evolution of the particle position can be represented by a random process known as a continuous-time random walk (Montroll & Weiss 1965). In § 2, we describe the general framework provided by the theory of continuous-time random walks and show how Dewan's model fits in that framework (§ 2.1). In fact, various extensions of Dewan's model may be devised which remain in the same framework. This is highly advantageous, since the theory of continuous-time random walks provides closed-form results from which the statistics characterizing tracer dispersion are derived (§ 2.2). From these results, it is easy to verify that the probability density function of the particle position evolves according to the diffusion equation (as is observed in Dewan's simulations) only for times much longer than the average time interval between two mixing events experienced by a given particle.

† Dewan's model might seem very crude when compared to others that attempt to reproduce vertical dispersion in stratified fluids by solving simplified equations for the motion of buoyant particles submitted to random pressure fluctuations (Csanady 1964; Pearson, Puttock & Hunt 1983; Kimura & Herring 1996). However, models of the latter type are not adapted for the highly intermittent and localized turbulence we are concerned with; they require a certain homogeneity of the turbulence and, in the geophysical context, probably apply better to mixing in the troposphere or in the ocean mixed layer.

The early-time evolution of tracer distributions depends on the details of the statistics of the mixing events. In § 3 we investigate this evolution using two simple models for such statistics. In the first model (§ 3.1), which is Markovian, successive mixing events experienced by a fluid particle are regarded as independent. In the second (§ 3.2), which is non-Markovian, memory is introduced which makes short time intervals between mixing events more likely than long ones. The dependence between mixing events is motivated by the idea that fluid particles which have experienced mixing in the recent past are in an environment that is only weakly stratified and is thus more susceptible to becoming turbulent. Using a simple model with three parameters, we study in detail the effect of the memory on particle dispersion.

In § 4 we estimate the average time interval between successive mixing events experienced by a particle in the deep ocean and in the stratosphere. In both the cases, we find it to be of the order of one day or so. As a result we argue that the non-diffusive behaviour that characterizes the early-time tracer evolution is important, particularly for stratospheric mixing. This highlights the practical importance of a better understanding of the mixing properties of turbulent patches, since some of these properties are crucial to the early-time behaviour. The properties that are most relevant may be clearly identified from the continuous-time random walk model which provides a rigorous framework relating the small-scale turbulent-patch dynamics to their impact on large-scale tracer distributions. The results described therefore have important implications for research on stratified fluids, much of which is motivated by potential applications to atmospheric and oceanic tracer evolution. Such implications are discussed in § 4.

In this paper, the focus is on one-dimensional modelling and on vertical mixing. More precisely, we consider vertical dispersion of fluid particles and exclusively calculate single-particle distributions. This is by itself of real interest in many physical contexts, but more generally, to consider vertical mixing, e.g. for chemically reacting flows, it will be necessary to calculate multi-particle statistics. We discuss this in the concluding section (§ 5).

2. Continuous-time random walks

In Dewan's model, fluid particles experience a random vertical displacement every time a mixing event takes place. The vertical position of fluid particles can therefore be described by a random walk with variable step size. In Dewan's simulations, the mixing events occur at regular time intervals, but it is more realistic to consider time intervals that are also random. In this case, the stochastic process describing the fluid-particle vertical positions falls into the class of continuous-time random walks introduced by Montroll & Weiss (1965) (see also Weiss & Rubin 1983; Hughes & Prager 1983, and references therein).

Continuous-time random walks are entirely defined by an associated probability density function $\Phi(l,\tau)$ which governs the probability that a (vertical) displacement of size l occurs τ time units after the previous displacement – in what follows we refer to l as the step and to τ as the waiting time. In the context of interest, $\Phi(l,\tau)$ completely describes the statistics of the vertical displacement experienced by a given fluid parcel as a result of intermittent turbulence; it is essentially a Lagrangian quantity which may not be related in an obvious way to the Eulerian characteristics of the turbulence. Deriving $\Phi(l,\tau)$ from Eulerian information is nonetheless possible in principle; for Dewan's model, it can be done explicitly as we now show.

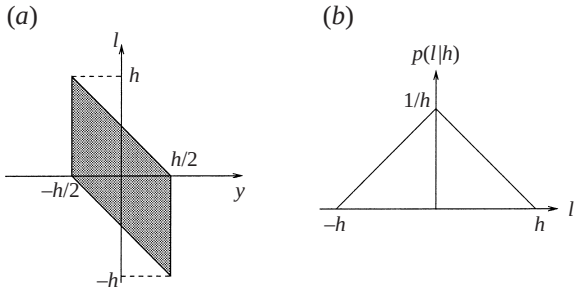

FIGURE 1. Calculation of $p(l|h)$: (a) the probability for the step to be l when the particle is at a distance y from the centre of the turbulent slab is distributed uniformly in the shaded area; (b) by integration of this probability over y, the form of $p(l|h)$ is derived.

2.1. *Dewan's model as a random walk*

A slightly generalized version of Dewan's model is defined as follows: mixing events, during which a slab of fluid of height h centred at altitude z is entirely homogenized by a turbulent patch, occur separated by random waiting times τ. The probability density functions of τ, h and z need to be specified: let these be $\psi(\tau)$, $\phi(h)$ and $q(z)$. A uniform distribution is assumed for z in $[0, H]$, where H is the total fluid height, so that $q(z) = 1/H$. For this model, it is clear that the vertical displacements are independent of the waiting times and therefore that $\Phi(l, \tau)$ factorizes according to

$$\Phi(l, \tau) = \psi(\tau) p(l). \tag{2.1}$$

We compute the step probability density $p(l)$ associated with a given $\phi(h)$ and denote it by $p_\phi(l)$. Neglecting boundary effects, we first note that any given particle has a probability h/H of being inside the turbulent slab. This implies that

$$p_\phi(l) = \int_0^\infty \left[\left(1 - \frac{h}{H}\right) \delta(l) + \frac{h}{H} p(l|h) \right] \phi(h)\, \mathrm{d}h, \tag{2.2}$$

where $p(l|h)$ is the step-size probability density given that the particle is inside a turbulent slab of height h. The form of $p(l|h)$ depends on the nature of the turbulent mixing. Following Dewan (1981a), we assume that the mixing is perfect, that is, we assume that the particles are redistributed uniformly inside the turbulent slab. If the distance between a particle and the centre of the slab is y (with $|y| < h/2$), the step size l for a given y is uniformly distributed in $[-h/2 - y, h/2 - y]$ (see figure 1). Since y itself is uniformly distributed (as is z), $p(l|h)$ takes the form displayed in figure 1(b), namely

$$p(l|h) = \frac{1}{h^2} \times \begin{cases} h + l, & -h \leqslant l \leqslant 0, \\ h - l, & 0 \leqslant l \leqslant h, \\ 0, & l \leqslant -h \ \text{ or } l \geqslant h. \end{cases}$$

Introducing this result into (2.1)–(2.2) provides the function $\Phi(l, \tau)$ that defines the continuous-time random walk completely.

Note that (2.2) implies a non-zero probability of steps with zero displacement, which we might call 'null steps'. This arises because we consider mixing events somewhere in the fluid column, which may not affect a given particle. To avoid this, we can instead consider only mixing events experienced by a given particle, thereby excluding null steps. This leads to new expressions for $\psi(\tau)$ and $p(l)$ which are derived explicitly in Appendix A.

We emphasize that continuous-time random walks provide a general framework which may be used to analyse a wide class of models of mixing by highly intermittent turbulence. Dewan's model, in particular, may easily be extended while remaining in this framework, for instance by modifying the form of $p(l|h)$ to relax the hypothesis of perfect mixing. But such models need not be derived from an Eulerian picture, such as that described in this section. Indeed, we believe that there are advantages to taking a Lagrangian perspective from the start and defining models directly from $\Phi(l,\tau)$. Such a definition might follow from *in situ* observations or laboratory experiments.

2.2. *General results*

Continuous-time random walks, although non-Markovian in general, are remarkably simple processes: using Laplace and Fourier transforms, $P(z,t)$, the probability for a particle starting initially at $z = 0$ to be at altitude z at time t, can be derived in closed form given any $\Phi(l,\tau)$ (e.g. Hughes & Prager 1983; Shlesinger, Klafter & Wong 1982). $P(z,t)$ completely describes the one-particle statistics associated with the mixing; it can be interpreted as the Green's function describing the vertical spreading of a passive tracer.

We now review briefly the properties of $P(z,t)$ for arbitrary $\Phi(l,\tau)$. For simplicity, we restrict our attention to functions $\Phi(l,t)$ that decouple according to (2.1). Defining the Laplace and Fourier transforms of, respectively, $\psi(\tau)$ and $p(l)$ as

$$\hat{\psi}(u) = \int_0^\infty \mathrm{e}^{-u\tau}\psi(\tau)\,\mathrm{d}\tau \quad \text{and} \quad \hat{p}(m) = \int_{-\infty}^\infty \mathrm{e}^{\mathrm{i}ml}p(l)\,\mathrm{d}l,$$

the following expression can be derived for the Fourier–Laplace transform of $P(z,t)$:

$$\hat{P}(m,u) = \hat{\Psi}_1(u) + \frac{\hat{\Psi}(u)\hat{\psi}_1(u)\hat{p}(m)}{1-\hat{\psi}(u)\hat{p}(m)}, \tag{2.3}$$

where $\hat{\Psi}(u)$ is the Fourier transform of the waiting-time distribution function $\Psi(\tau)$ defined by

$$\Psi(\tau) := \int_\tau^\infty \psi(\tau')\,\mathrm{d}\tau'$$

(e.g. Weiss & Rubin 1983). The functions $\psi_1(\tau)$ and $\Psi_1(\tau)$ whose Fourier transforms appear in (2.3) describe the statistics of the time elapsed between the origin of time, associated with the first observation, and the first mixing event. These statistics are different from those of the waiting times between successive mixing events (described by $\psi(\tau)$ and $\Psi(\tau)$) because an event does not necessarily occur at $t = 0$. Regarding the origin of time as a random time with no correlation with the times at which turbulent events occur, it may be shown that

$$\psi_1(\tau) = \frac{1}{\tau_\mathrm{m}}\Psi(\tau) \quad \text{and} \quad \Psi_1(\tau) = \frac{1}{\tau_\mathrm{m}}\int_\tau^\infty \Psi(\tau')\,\mathrm{d}\tau', \tag{2.4}$$

where

$$\tau_\mathrm{m} := \int_0^\infty \tau\psi(\tau)\,\mathrm{d}\tau = \int_0^\infty \Psi(\tau)\,\mathrm{d}\tau \tag{2.5}$$

is the average waiting time.

Equation (2.3) gives the closed form of $P(z,t)$; it can also be employed to obtain the evolution of the moments $\langle z^{2n}\rangle$ of the vertical position (n is an integer and $\langle\cdot\rangle$

denotes ensemble average). The Laplace transform of these moments is given by

$$\mathscr{L}\langle z^{2n}\rangle(u) = (-1)^n \left.\frac{\partial^{2n}}{\partial m^{2n}}\right|_{m=0} \hat{P}(m,u). \tag{2.6}$$

A remarkable conclusion is readily derived from this expression. Suppose that all the moments of $\psi(\tau)$ and $p(l)$ are finite, then $\hat{\psi}(u)$ and $\hat{p}(m)$ can be expanded in Taylor series according to

$$\hat{\psi}(u) = 1 - \tau_{\mathrm{m}} u + O(u^2), \quad \hat{p}(m) = 1 - \frac{\sigma^2}{2} m^2 + O(m^4), \tag{2.7}$$

where

$$\sigma^2 := \int_{-\infty}^{\infty} l^2 p(l)\,\mathrm{d}l$$

is the step variance and we assume that the step distribution is symmetric. The evolution of the moments $\langle z^{2n}\rangle$ for large time is governed by the limit of (2.6) as $u \to 0$. It can be verified that the terms neglected in (2.7) do not contribute to $\mathscr{L}\langle z^{2n}\rangle(u)$ in this limit, i.e.

$$\lim_{u\to 0} \mathscr{L}\langle z^{2n}\rangle(u) = \lim_{u\to 0} (-1)^n \left.\frac{\partial^{2n}}{\partial m^{2n}}\right|_{m=0} \hat{P}_{\mathrm{d}}(m,u)$$

where

$$\hat{P}_{\mathrm{d}}(m,u) := \frac{1}{u+\kappa m^2}, \quad \text{with} \quad \kappa := \frac{\sigma^2}{2\tau_m}. \tag{2.8}$$

Now, $\hat{P}_{\mathrm{d}}(m,u)$ can be recognised as the Fourier–Laplace transform of the Gaussian

$$P_{\mathrm{d}}(z,t) = \frac{1}{\sqrt{4\pi\kappa t}} \exp\left(-\frac{z^2}{4\kappa t}\right), \tag{2.9}$$

the Green's function associated with the diffusion equation with diffusivity κ. We therefore conclude that the vertical mixing processes modelled by continuous-time random walks are diffusive in the long-time limit, unless $\Phi(l,\tau)$ possesses infinite moments (cf. Hughes & Prager 1983, and references therein).

This conclusion – essentially a central-limit theorem – provides rigorous support for the numerical simulations of Dewan (1981*a*) which revealed a diffusive behaviour for the mixing. (In fact, we can give an exact expression for the diffusion coefficient in Dewan's 1981*a* model, correcting his Eq. (2); this is done in Appendix B.) The conclusion also gives precise conditions for the diffusive behaviour: it is only when $\psi(\tau)$ or $p(l)$ have infinite moments (or if τ_{m} or σ^2 vanish) that the form of the distribution $P(z,t)$ for long time can be significantly different from the Gaussian (2.9).

Infinite moments, which lead to Levy distributions and anomalous diffusion (e.g. Shlesinger *et al.* 1982), seem unlikely in the case we consider of vertical displacements induced by turbulent patches. This is clear for the displacement variance which is naturally bounded because of the stratification and of the finiteness of the energy supplied. It is less clear for the average waiting time, but it will be admitted as an hypothesis that may need re-examination. However, even with finite moments, it is important to recognise that the diffusive behaviour is obtained in the long-time limit only. This limit is valid for times long compared to a suitable time scale, which is provided by the average waiting time τ_{m} (although higher-order moments of $\psi(\tau)$

are also relevant if they differ significantly from τ_m). For shorter time scales, the moments $\langle z^{2n}\rangle$ differ from those obtained for a diffusive process, and the mixing is not represented accurately by the diffusion equation. Interestingly the second-order moment evolves as predicted by the diffusion equation for all time. Indeed, (2.3)–(2.6) give

$$\left.\frac{\partial^2}{\partial m^2}\right|_{m=0} \hat{P}(m,u) = \frac{\sigma^2}{\tau_m u^2}$$

and, after Laplace inversion,

$$\langle z^2\rangle = \frac{\sigma^2}{\tau_m} t \tag{2.10}$$

exactly.† It is easy to confirm that this is exceptional, and that higher-order moments behave diffusively only asymptotically for $t \gg \tau_m$. In particular, the fourth-order moment is found to satisfy

$$\mathscr{L}\langle z^4\rangle = \left.\frac{\partial^4}{\partial m^4}\right|_{m=0} \hat{P}(m,u) = \frac{1}{\tau_m}\left(\frac{6\sigma^4\hat{\psi}(u)}{u^3\hat{\Psi}(u)} + \frac{\mu^4}{u^2}\right), \tag{2.11}$$

where μ^4 is the fourth-order moment of the step distribution $p(l)$. This clearly differs from the expression $6\sigma^4/(\tau_m^2 u^3)$ found for diffusion (an expression which is of course recovered in the limit of small u corresponding to large t).

A striking difference between continuous-time random walk and diffusion on short time scales is the fact that, unlike $P_d(z,t)$, $P(z,t)$ is not a smooth function of z for $t > 0$. Indeed, it follows from the Fourier inversion formula applied to $\hat{P}(m,u)$, that $P(z,t)$ has a contribution proportional to $\delta(z)$ with amplitude $T(t)$ given by

$$T(t) = \mathscr{L}^{-1}[\lim_{m\to\infty} \hat{P}(m,u)],$$

where $\mathscr{L}^{-1}$ denotes the inverse Laplace transform. When $p(l)$ is smooth, $\lim_{m\to\infty} \hat{p}(m) = 0$, and (2.3) indicates that $\lim_{m\to\infty} \hat{P}(m,u) = \hat{\Psi}_1(u)$, yielding

$$T(t) = \Psi_1(t). \tag{2.12}$$

The contribution to $P(z,t)$ proportional to $\delta(z)$ is associated with particles that are deterministically trapped at the initial position. Physically, the trapping results from the finiteness of the waiting time between mixing events which guarantees that, at any time, a (decreasing) number of particles have not been affected by the mixing events. This implies, for example, that tracer concentrations in certain locations will remain at their initial value for finite times.

The average waiting time τ_m provides an estimate of the time scale required for diffusion to be a valid approximation to the more complex process of mixing by turbulent patches. In geophysical fluids, this time should be compared with other dynamically relevant time scales in order to assess the importance of the short-time, non-diffusive behaviour. Order-of-magnitude estimates detailed in § 4.1 suggest that in the stratosphere, and to a lesser extent in the deep ocean, the relevant time scales are not much larger than τ_m. This makes the short-time, non-diffusive regime of the vertical dispersion by turbulent patches highly relevant to geophysical mixing. This regime is unfortunately non-universal, with the behaviour of $P(z,t)$ and $\langle z^{2n}\rangle$, $n \neq 1$,

† This does not hold when no difference is made between the statistics of the first step and of the subsequent ones, i.e. when the origin of time is assumed to coincide with a turbulent event, except in the Markovian case.

depending on the details of $\psi(\tau)$ and $p(l)$. In the next section we employ very simple models for these probability density functions to gain insight into the short-time regime, comparing the form of $P(z,t)$ in particular cases with $P_{\mathrm{d}}(z,t)$ given in (2.9).

3. Short-time behaviour

A natural choice for the waiting-time distribution is that of a Poisson distribution

$$\psi(\tau) = \alpha \mathrm{e}^{-\alpha\tau}, \tag{3.1}$$

where α is a fixed parameter. This distribution appears provided that the successive turbulent events experienced by a given particle are independent. We argue below that a certain dependence between turbulent events is likely to exist because of the physical mechanisms leading to the formation of turbulent patches and we propose a simple parameterization of this effect. However we start by considering the Poisson distribution (3.1) as a reference case.

3.1. *Randomized random walks*

With a Poisson distribution for the waiting times, the continuous-time random walk becomes a Markovian process known as randomized random walk (e.g. Hughes & Prager 1983). The parameter α in (3.1) determines the time scale of the process and is the inverse of the average waiting time, $\tau_{\mathrm{m}} = \alpha^{-1}$. An important aspect of randomized random walks concerns the statistics of the first turbulent event which are the same as the statistics of subsequent events. This is easily established by introducing the Laplace transform $\hat{\psi}(u) = \alpha/(u+\alpha)$ of (3.1) into the transform of (2.4), yielding $\hat{\psi}_1(u) = \hat{\psi}(u)$.

The probability density function $P(z,t)$ is easily calculated from (3.1) and (2.3); it takes the form

$$P(z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left\{-\mathrm{i}mz - \alpha t[1-\hat{p}(m)]\right\} \mathrm{d}m, \tag{3.2}$$

once the Laplace inversion is performed. Clearly, this expression is different from $P_{\mathrm{d}}(z,t)$, unless $\hat{p}(m) = 1 - \sigma^2 m^2/2$ (which corresponds to a non-smooth $p(l)$).

As an illustration, let us consider a Gaussian distribution for the steps, namely

$$p(l) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{l^2}{2\sigma^2}\right) \quad \rightarrow \quad \hat{p}(m) = \exp\left(-\frac{\sigma^2 m^2}{2}\right). \tag{3.3}$$

Equation (3.2) then becomes

$$P(z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left[-\mathrm{i}mz - \alpha t(1-\mathrm{e}^{-\sigma^2 m^2/2})\right] \mathrm{d}m.$$

Isolating the contribution proportional to $\delta(z)$, we find

$$P(z,t) = \delta(z)\mathrm{e}^{-\alpha t} + \frac{\mathrm{e}^{-\alpha t}}{\pi}\int_0^{\infty} \cos(zm)[\exp(\mathrm{e}^{-\sigma^2 m^2/2}\alpha t) - 1] \mathrm{d}m. \tag{3.4}$$

In particular, the fraction of particles trapped at the origin is given by $T(t) = \exp(-\alpha t)$, in agreement with (2.12), (3.1) and (2.4).

To confirm the formula (3.4), we have carried out numerical simulations of the randomized random walk for 60 000 realisations and for t up to $20\alpha^{-1}$. We have computed the evolution of the variance $\langle z^2 \rangle$ as well as approximations to $P(z,t)$ at regular time intervals. To estimate $T(t)$, we have simply counted the number of

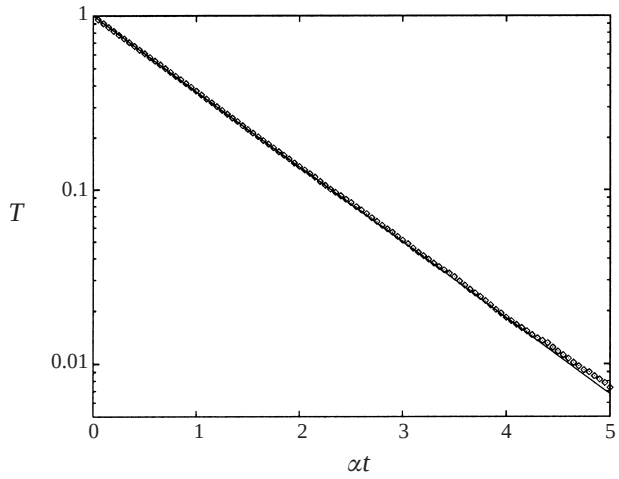

FIGURE 2. Evolution of T, the fraction of particles trapped at the origin. The result of a numerical simulation (symbols) is compared with the theoretical prediction $\exp(-\alpha t)$ (solid line).

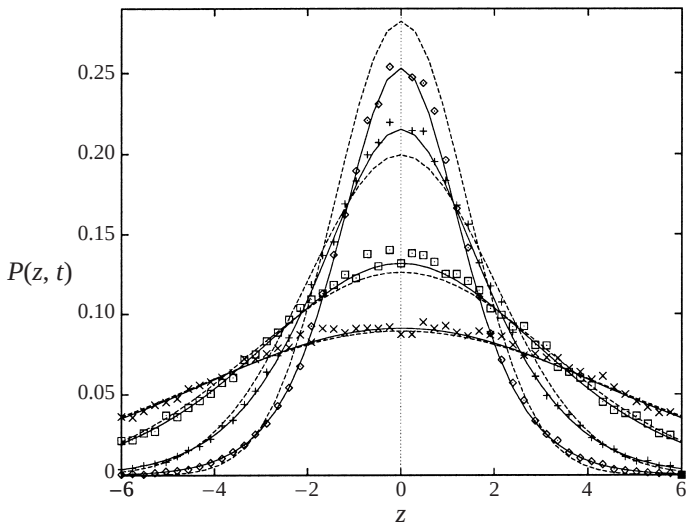

FIGURE 3. Probability density function $P(z,t)$ for the randomized random walk with Gaussian step distribution: the theoretical formula (3.4) (solid lines) is compared with the results of a numerical simulation (symbols) and with the Gaussian (2.9) (dashed lines) for $\alpha t = 2, 4, 10$ and 20. z is normalized by the standard deviation σ.

particles that are in a narrow interval around $z = 0$ and normalized this number by the number of realisations. The corresponding result is displayed in figure 2 which shows that the expected exponential behaviour is recovered accurately. Figure 3 compares the smooth part of $P(z,t)$ given by (3.4) with the results of the numerical simulation and with the density $P_{\mathrm{d}}(z,t)$ corresponding to a diffusive process with the same variance and average waiting time. As expected, $P(z,t)$ and $P_{\mathrm{d}}(z,t)$ become identical for large time, but they are significantly different – in particular near the origin – at earlier times. Note that we have verified that the variance derived from the simulation (not shown) increases linearly with time for all time as predicted by

(2.10). We have also confirmed the analytic result for $\langle z^4 \rangle$ derived from (2.11), (3.1) and (3.3); this result is conveniently written in terms of the flatness (or excess), which vanishes for a diffusion process but is here found to be

$$\frac{\langle z^4 \rangle}{\langle z^2 \rangle^2} - 3 = \frac{3\tau_{\mathrm{m}}}{t} > 0. \tag{3.5}$$

The extension of the Dewan model discussed in § 2.1 can be considered using the Poisson distribution (3.1) for the waiting times. The corresponding probability density function of z is given by (3.2), with $\hat{p}(m)$ replaced by the Fourier transform of (2.2), explicitly given by

$$\hat{p}_\phi(m) = 1 - \frac{\overline{h}}{H} + \frac{2}{H} \int_0^\infty \frac{1 - \cos(mh)}{hm^2} \phi(h)\, \mathrm{d}h, \tag{3.6}$$

where $\overline{h}$ is the average height of the turbulent layers. The resulting expression, as expected, also implies a contribution proportional to $\delta(z)$ in $P(z,t)$ with amplitude $T(t) = \exp(-\alpha \overline{h} t/H)$.

3.2. *Memory effect*

As already mentioned, a Poisson distribution for $\psi(\tau)$ is obtained under the assumption that the mixing events are independent. However, it is likely that geophysical mixing events are not independent for two reasons: (i) previous mixing reduces stratification, hence the Richardson number for a given shear; and (ii) the sharp density gradients which are formed at the boundary of mixed regions lead to strong shear and low Richardson numbers when tilted by the large-scale flow (Thorpe 1973; Bühler, McIntyre & Scinocca 1999). Since low Richardson numbers are associated with flow destabilization by shear instability, both processes lead to a certain memory effect: turbulent patches are more likely to form in regions which have been affected by turbulence in the recent past. In Lagrangian terms, this means that the probability for a particle to experience a new mixing event decreases with the time elapsed since the last event.

To investigate this effect, let us construct a simple model that introduces some memory in the probability of mixing events. Consider a random point process such that the probability of an event in a time interval $\mathrm{d}\tau$ is $[\alpha + \beta \exp(-\gamma\tau)]\mathrm{d}\tau$, where τ is the time since the previous event and α, β, γ are positive constants. It is easy to verify that the probability $\Psi(\tau)$ for the waiting time to be larger than τ is

$$\Psi(\tau) = \exp\left[-\alpha\tau - \frac{\beta}{\gamma}(1 - \mathrm{e}^{-\gamma\tau})\right], \tag{3.7}$$

from which the probability density function $\psi(\tau)$ is derived according to

$$\psi(\tau) = -\frac{\mathrm{d}\Psi}{\mathrm{d}\tau}. \tag{3.8}$$

For both short and long time, this density is well approximated by an exponential, but with a rate changing from $\alpha + \beta$ for $\gamma\tau \ll 1$ to α for $\gamma\tau \gg 1$, with γ controlling the time scale of the transition between these two rates. Physically, we can interpret α as the probability per unit time that a mixing event occurs when the fluid is stratified, $\alpha + \beta$ as the probability per unit time that a mixing event occurs when the fluid has been homogenized by prior turbulence, and γ as the inverse of the time required for the restratification of turbulent patches.

The long-time behaviour of the process is characterized by the average waiting

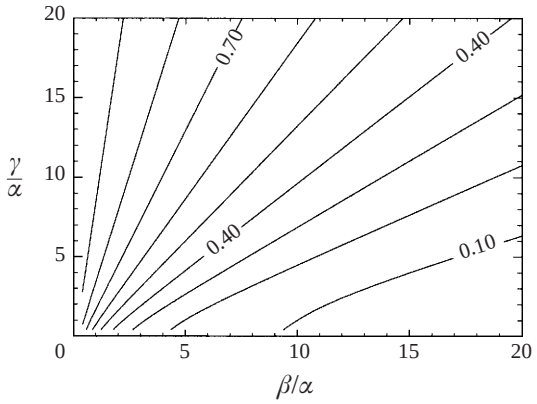

FIGURE 4. Non-dimensional average waiting time $\alpha\tau_m$ as a function of β/α and γ/α for the process with memory defined by (3.7)–(3.8).

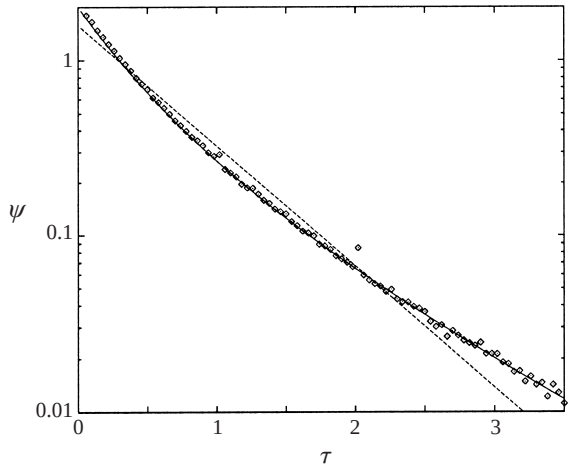

FIGURE 5. Probability density function $\psi(\tau)$ for the process with memory (3.7)–(3.8) with $\beta/\alpha = \gamma/\alpha = 1$ (solid line) compared with the probability density function for a Poisson process with the same average waiting time (dashed line). The symbols result from a direct numerical simulation of the process with memory.

time τ_m, which is conveniently derived from (3.7) by integration with respect to τ. Figure 4 displays $\alpha\tau_m$ as a function of γ/α and β/α (these non-dimensional variables appear when α^{-1} is chosen as reference time). It may be verified that τ_m is in the range $[(\alpha+\beta)^{-1}, \alpha^{-1}]$, tending to the lower (upper) bound as γ tends to 0 (∞). Once the value of τ_m corresponding to given α, β, γ is calculated, one can compare the probability density function (3.8) defining the process with memory with the exponential obtained for a (memoryless) Poisson process with the same τ_m. This is done in figure 5 in the case $\beta/\alpha = \gamma/\alpha = 1$ for which the average waiting time is given by $\alpha\tau_m = 0.633$. As anticipated, the effect of the memory is to introduce two different decay rates for short and long time (corresponding to two different slopes in the linear-log coordinates of figure 5); τ_m^{-1} can be regarded as the properly averaged decay rate.

The memory effect that exists when $\beta \neq 0$ and $\gamma < \infty$ makes our model non-Markovian. As a consequence, the statistics of the first waiting time differ from those of the subsequent waiting times: $\psi_1(\tau) \neq \psi(\tau)$. To characterize this difference, we can

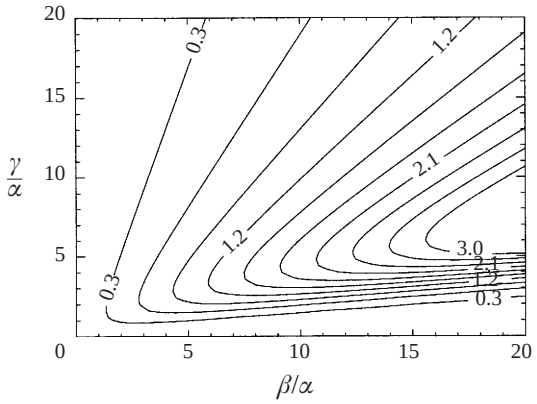

FIGURE 6. Non-dimensional parameter η defined by (3.9) as a function of β/α and γ/α.

compare the average of the first waiting time, $\tau_{1\mathrm{m}}$ say, with the average waiting time τ_{m} and define

$$\eta := \frac{\tau_{1\mathrm{m}} - \tau_{\mathrm{m}}}{\tau_{\mathrm{m}}} = \frac{\int_0^\infty \tau\psi_1(\tau)\,\mathrm{d}\tau - \tau_{\mathrm{m}}}{\tau_{\mathrm{m}}} = \frac{\int_0^\infty \tau\Psi(\tau)\,\mathrm{d}\tau}{\tau_{\mathrm{m}}^2} - 1, \tag{3.9}$$

where the third equality follows from (2.4).

Other consequences of the memory effect are also characterized by η. The memory effect modifies the fraction of particles trapped at the origin $T(t)$. This can be quantified by comparing the average time spent by particles at the origin for the non-Markovian model with that obtained for a memoryless process with the same average waiting time. From (2.12), the first average time is found to be $\int_0^\infty \Psi_1(\tau)\,\mathrm{d}\tau = \tau_{1\mathrm{m}}$, whereas the second is simply τ_{m}. It is thus seen from (3.9) that η precisely measures the relative difference between these two average times. This gives a second sense in which η characterizes the effect of the memory. A third sense appears when one considers the flatness. Substituting the expansion $\hat{\Psi}(u) = \tau_{\mathrm{m}}[1 + (\eta + 1)u] + O(u^2)$, which is derived from (2.5) and (3.9), into (2.11), and assuming a Gaussian distribution for the steps leads to the following asymptotic expression for the flatness:

$$\frac{\langle z^4\rangle}{\langle z^2\rangle^2} - 3 = \frac{3(1 + 2\eta)\tau_{\mathrm{m}}}{t} + O(t^{-2}). \tag{3.10}$$

Comparison with the expression (3.5) derived for the Poisson process indicates that η measures the leading-order change to the flatness that results from the memory effect.

In view of its multiple interpretations, it is worthwhile investigating the behaviour of η for the process defined by (3.7). It can be shown analytically that $\eta \geqslant 0$ in general, with $\eta = 0$ only when $\beta = 0$, i.e. when the process is Markovian. Therefore, $\tau_{1\mathrm{m}} \geqslant \tau_{\mathrm{m}}$; this implies that the memory effect increases (i) the first waiting time compared with the subsequent ones; (ii) the number of particles that are trapped at the origin; and (iii) the flatness of the particle-position probability density function $P(z, t)$. Figure 6, showing η as a function of β/α and γ/α, gives an idea of the region of parameter space where the memory effect plays a significant role. This region corresponds to large β/α and γ/α, and we argue in § 5 that it is likely to be the relevant one in the context of stratospheric mixing. Asymptotic calculations for large β and γ detailed in Appendix C show that the maximum value for η is attained when β and γ obey

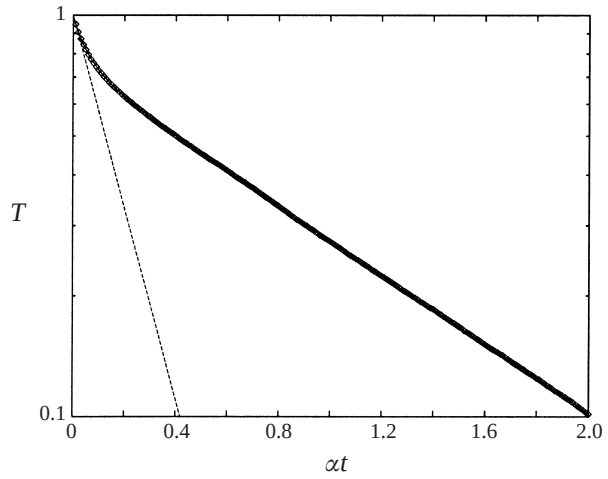

FIGURE 7. Evolution of T, the fraction of particles trapped at the origin. Results from a numerical simulation (symbols) are compared with the theoretical prediction $\Psi_1(t)$ (solid line) and with the corresponding density for a Poisson process (dotted line).

the relation $\gamma \ln(\beta/\alpha) \approx \beta$ and is given by $\eta \approx \beta/(4\alpha)$. Although these expressions are quite inaccurate for moderate values of β (the error scales like $1/\ln(\beta/\alpha)$ and thus decreases slowly as β increases), they may prove useful to assess roughly whether the memory effect is significant for given values of α, β and γ.

We have carried out a numerical simulation of the process with memory so as to confirm the theoretical results. A Gaussian distribution (3.3) has been taken for the step size, and the origin of time has been fixed randomly. The parameters $\beta/\alpha = 20$ and $\gamma/\alpha = 10$ have been chosen. With this choice, the average waiting time is given by $\alpha\tau_m = 0.182$, while $\eta = 3.19$; 60 000 realisations were employed to obtain a smooth approximation to the probability density function $P(z,t)$.

Figure 7 shows the non-smooth contribution to $P(z,t)$, i.e. the fraction of particles trapped at the origin (in practice in a small neighbourhood of the origin), as a function of time, compared with the theoretical prediction $\Psi_1(t)$. It also shows the function $\exp(-t/\tau_m)$ which gives the fraction of trapped particles for a Poisson process with average waiting time τ_m and thus it illustrates the increase in the number of trapped particles that results from the memory effect. In particular, it is seen from the definition of η (3.9) that the average time spent by particles at $z = 0$ is $\tau_{1m} = (1+\eta)\tau_m = 4.19\tau_m$.

The smooth part of $P(z,t)$ obtained from the numerical simulation is displayed in figure 8. It is compared with the probability density function (3.4) that we derived in § 2 taking a Poisson distribution for the waiting times. To make the comparison meaningful, the substitution $\alpha \to \alpha/0.182$ has been introduced in the probability density function (3.1) defining the Poisson process, so that the two processes under consideration have the same average waiting time. The figure clearly demonstrates the dramatic impact that the memory has on the spreading of particles: whilst the number of particles deterministically trapped at the origin increases because of the memory effect, the amplitude of the smooth part of $P(z,t)$ near the origin decreases significantly; at the same time the probability for long displacements ($z \gg \sigma$) is somewhat larger, consistently with the fact that the variance $\langle z^2 \rangle$ is unchanged by the memory effect. (It has been verified that $\langle z^2 \rangle$ evolves linearly according to (2.10) and that the flatness evolves according to (3.10).) The qualitative differences between the

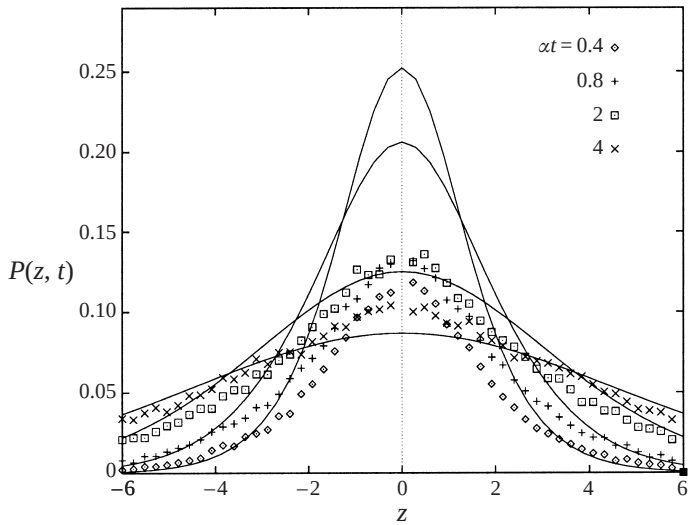

FIGURE 8. Probability density function $P(z,t)$ for the process with memory, assuming a Gaussian step distribution: the results of a numerical simulation (symbols) are compared with the theoretical result (3.4) obtained for a (memoryless) Poisson process with the same average waiting time $\tau_m = 0.182\alpha^{-1}$ (solid line) for $\alpha t = 0.4, 0.8, 2$ and 4. z is normalized by the standard deviation σ.

distributions obtained with and without memory can be understood by noting that the memory effect increases the probability of long displacements (due to series of steps separated by short waiting times) as well as the probability of zero displacement. Of course, in the long-time limit, the two distributions become identical as both tend to a Gaussian shape (cf. figure 3).

4. Implications

The continuous-time random-walk theory employed above is fruitful in relating the mixing at small scale, which is associated with single turbulent events, to its impact on the large-scale dispersion of tracers, which results from the superposition of a large number of such events. In particular, it suggests that the information essential to describe tracer dispersion is encapsulated in the two probability density functions $p(l)$ and $\psi(\tau)$. At present, little is known about these functions, either in the stratosphere or in the deep ocean, and one has to rely on heuristic models such as Dewan's perfect-mixing model. These are sufficient to illustrate some implications of the random-walk results and study the sensitivity of dispersion to such phenomena as the memory effect; but, clearly, accurate estimates for $p(l)$ and $\psi(\tau)$ are needed to model vertical mixing in the stratosphere or the ocean realistically. In this section, we suggest how information about $p(l)$ and $\psi(\tau)$ could be obtained either from geophysical observations or from laboratory or numerical experiments. We start by estimating the average waiting time τ_m using (2.8) and recent estimates for the oceanic and stratospheric diffusivity. As already mentioned, τ_m has a crucial importance in practice, since it allows one to assess the time scale after which the diffusive approximation applies.

4.1. *Stratospheric and oceanic* τ_m

For the oceans, the order of magnitude of the vertical diffusivity κ has been estimated using various methods which have provided values in the range 10^{-4}–$10^{-5}\,\mathrm{m^2\,s^{-1}}$

(Ledwell, Watson & Law 1993; Toole, Polzin & Schmitt 1994; Polzin *et al.* 1997, and references therein), while the height of typical turbulent layers suggest that the vertical displacements are of the order of a few metres (Gibson 1986; Polzin 1996). Taking $\sigma^2 = 10\,\mathrm{m}^2$ gives an average waiting time τ_m in the range 5×10^4–5×10^5 s, i.e. 0.6–6 days. The diffusive behaviour, which requires $t \gg \tau_\mathrm{m}$, therefore emerges for time scales longer than a week or so. This is quite short compared to the typical time scales relevant to large-scale ocean dynamics as given, e.g., by the inverse of a typical strain rate (which is estimated to be of the order of 40 days by Ledwell *et al.* 1993). So, we may expect diffusion to model oceanic mixing satisfactorily in many circumstances. However, important oceanic phenomena involve much shorter time scales (e.g. tides, biological activity); when these interact with small-scale mixing, the non-diffusive nature of the mixing can become critical.

For the stratosphere, the vertical diffusivity has recently been estimated as about $10^{-2}\,\mathrm{m}^2\,\mathrm{s}^{-1}$ (Balluch & Haynes 1997; Waugh *et al.* 1997). If, following Dewan (1980), we take typical vertical displacements of the order of 30 m, i.e. $\sigma^2 \approx 10^3\,\mathrm{m}^2$, we find that the average waiting time τ_m is about 1 day. Thus, in contrast to the oceanic situation, the time scale of a week or so that might be required for the diffusive behaviour to appear in the stratosphere is not short compared to other relevant time scales. For instance, dynamical time scales are provided by the inverse of the average strain rate, which has been estimated to be of the order of three days in the lower stratosphere (Haynes & Anglade 1997), and by the Lagrangian correlation time of the strain, estimated as one day or so (Ngan & Shepherd 1999). Similarly, photochemical time scales of stratospheric chemical species are frequently shorter than a week (Brasseur & Solomon 1986). This makes the short-time, non-diffusive regime of the vertical dispersion by turbulent patches potentially highly relevant to stratospheric mixing.

4.2. *Estimation of $\psi(\tau)$ and $p(l)$*

The waiting-time probability density function $\psi(\tau)$ which describes the (Lagrangian) frequency of turbulent events does not solely depend on the dynamics of turbulent patches. Indeed, it is also related to the spatio-temporal distribution of the forcing (gravity-wave activity in particular) which causes these turbulent events and to the manner in which this distribution is sampled by Lagrangian trajectories. Thus $\psi(\tau)$ integrates a variety of phenomena, many of which depend crucially on the specific geophysical situation. We can therefore argue that $\psi(\tau)$ would be best estimated by direct, preferably Lagrangian, geophysical observations. In the stratosphere, for instance, instruments detecting turbulence carried by long-duration balloons could provide data to approximate $\psi(\tau)$. Such an approach would allow direct estimation of τ_m and assessment of the importance of the memory effect. More sophisticated instruments could also provide information on the structure of turbulence within the patches.

In contrast to $\psi(\tau)$, the step probability density function $p(l)$ can be essentially regarded as a characteristic of turbulent patch dynamics. If it is the large-scale effects that are of interest, we suggest that extraction of information on $p(l)$ should be a priority for laboratory and numerical experiments on turbulent patches. It is clear that $p(l)$ describes the net irreversible effect of turbulent events on particle displacements. This suggests that the transient behaviour of the turbulent patches can be disregarded. In particular, direct knowledge of the patch lifetime is not necessary (assuming a certain separation between this time and τ_m), although it is an essential factor in the dynamical determination of both $p(l)$ and $\psi(\tau)$.

One may also hope that $p(l)$ depends on only a few parameters characterizing irreversible mixing – the difference between the potential energy of the stratification before and after the mixing, or perhaps more precisely the difference in the background potential energy (i.e. the potential energy of the stable density profile obtained by adiabatic rearrangement of fluid particles, Winters *et al.* 1995), is one such parameter. A challenge would then be to relate these parameters to quantities characterizing the large-scale flow such as the Richardson number or the gravity-wave activity. In a similar spirit, laboratory or numerical experiments could prove useful to establish relationships between $p(l)$ for a single patch on the one hand, and gross measures of the intensity of turbulence in this patch on the other hand. Information of this type would make it possible to estimate $p(l)$ using low-cost instruments carried on balloons or floats; this is important since a large number of instrumented balloons or floats would be necessary if reliable statistics for $p(l)$ were to be obtained.

When estimating the mixing properties of turbulent patches from observations, it is common to compute a diffusivity based on measured energy dissipation rates and large-scale stratification. This suggests a simple extension to Dewan's perfect mixing model in which $p(l)$ is derived by solving the diffusion equation inside the patch for a finite time, possibly with a time-dependent diffusivity. Although diffusion is at best a crude model for the instantaneous mixing taking place inside turbulent patches, it is plausible that it provides a satisfactory model when only $p(l)$ is required. Whether this is the case could be tested in direct numerical simulations.

There are a number of other important issues relevant to the models discussed earlier that could be examined using numerical or laboratory experiments, provided that these concerned not isolated patches but an ensemble of patches, thus allowing analysis of both the spatial and temporal aspects of the mixing in idealized situations. In particular, the importance of the memory effect studied in § 3.2 and the validity of the hypothesis of decoupled probability density function (2.1) could be assessed.

5. Discussion

In this paper, we have studied heuristic models of vertical mixing by intermittent turbulence using a framework provided by the theory of continuous-time random walks. As we have emphasized, a variety of models can be devised within this framework, and so a variety of specific physical effects may be accounted for – the memory effect is just one of these. The advantage of continuous-time random walks is their inherent simplicity, which leads to a closed-form expression for the probability density function of the particle position $P(z,t)$. However, if one wishes to regard $P(z,t)$ as the mean concentration of a tracer in a geophysical fluid, it can be argued that this closed-form expression is not particularly useful: since the tracer concentration is governed, not only by vertical mixing, but also by other phenomena (advection and chemical reactions in particular), what is needed is an evolution equation for $P(z,t)$ in which these phenomena can be included. As noted in § 1, $P(z,t)$ is generally not sufficient to describe chemical reactions, which depend on two-particle or more probability density functions, but in principle advection can be treated using $P(z,t)$ only. In general this is fairly delicate (see Compte 1997 for an *ad hoc* treatment); we therefore leave the study of the interaction between advection and mixing for future work.

Our approach in this paper is to use probabilistic models to relate the properties of single patches to their effect on large scales. Here, the precise effect studied is dispersion rather than genuine mixing. In the context of random walks, genuine

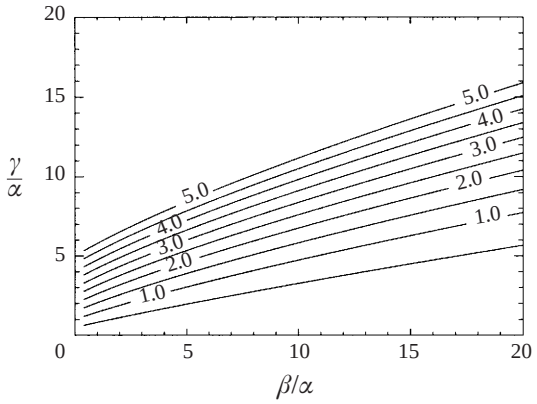

FIGURE 9. Non-dimensional parameter $\tau_m\gamma$ as a function of β/α and γ/α. A rough estimate suggests that $\tau_m\gamma$ for the stratosphere could be in the range [1, 5].

mixing would be implied by statistical independence of the walk undergone by each particle. Inside each patch this independence is likely to be a useful approximation, since the flow is three-dimensionally turbulent. However the independence will not hold when the effect of an ensemble of patches is considered. This is why our single-particle results can be taken, strictly speaking, only to describe dispersion, not genuine mixing. Nonetheless, it is clear that the overall approach used in this paper can be extended to quantify genuine mixing effects.

Returning to more specific issues raised by the models in this paper, the memory effect, discussed in § 3.2, associated with the finite time for complete restratification of turbulent patches can have an impact on tracer dispersion. Our simple three-parameter model allows us to assess the conditions under which this impact is most significant. In view of the importance of the transient, non-diffusive regime for mixing in the stratosphere, it would be useful to estimate 'stratospheric' values of the parameters α, β and γ. It is not an easy task, however. A first piece of information is provided by the estimate for τ_m given in § 4.1; a second may be derived by interpreting γ^{-1} as the typical time taken by turbulent patches to restratify. Arguments based on gravity-wave propagation indicate that, at mid-latitudes, this time might be of the order of several hours (O. Bühler, personal communication; see also Bühler *et al.* 1999). Taking τ_m of the order of a few days then suggests that realistic values of $\tau_m\gamma$ are in the range 1 to 5. A last piece of information is necessary to determine α, β and γ completely. In principle, this could be provided by the ratio β/α, which measures the relative increase in the mixing-event probability that appears when the fluid is homogenized by prior turbulence; unfortunately we do not have simple means of evaluating this ratio.

However, irrespective of the precise value of β/α, an interesting conclusion can be drawn from figure 9 which displays $\tau_m\gamma$ as a function of β/α and γ/α. When compared to figure 6, this figure indicates that η is greater than one, i.e. that the enhancement of dispersion due to memory is significant, when $\tau_m\gamma$ is in the realistic range, provided that β/α is not too small. The same conclusion may also be reached using the asymptotic expressions for τ_m and η developed in Appendix C.

We conclude this paper with two remarks that emphasize the importance of realistic representations of the small-scale mixing for the modelling of large-scale tracer distributions. First, when the tracer stirring is dominated by the large-scale

flow, as is thought to be the case in the stratosphere, the time necessary for tracer scales to cascade down to the mixing scale remains dependent on the strength of the mixing as this strength becomes vanishingly small (e.g. Pierrehumbert 1995). As a consequence, it is important to represent the small-scale mixing accurately even if it is weak. Second, recent work on stratospheric tracer spectra using diffusion to parameterize the mixing indicates that horizontal scales as large as few hundred kilometres are affected by diffusion when a realistic value is taken for the diffusivity (Vanneste & Haynes 2000). The same conclusion can be expected to hold when the small-scale mixing is represented by more sophisticated models such as that considered in this paper.

It is a pleasure to acknowledge helpful conversations with O. Bühler. We thank E. Hernández–Garcia who pointed out the need for particular first-step statistics and an anonymous referee for a number of constructive comments. The work of the Centre for Atmospheric Science is supported by the UK Natural Environmental Research Council through the UK Universities Global Atmospheric Modelling Programme and by the Isaac Newton Trust. J.V. was funded by the European Commission through grant FMBICT972004.

Appendix A. Two interpretations of Dewan's model

It is possible to express Dewan's model in terms of continuous-time random walks in two different ways. In § 2, we have taken the waiting-time probability density function, in essence a Lagrangian quantity, as equal to the probability density function $\psi(\tau)$ for the time between successive turbulent events in the fluid column, an Eulerian quantity. By doing so, we include in the description null events that do not affect a chosen particle (because it is outside the turbulent patch); this results in a term proportional to $\delta(l)$ in the displacement probability density function (2.2).

An alternative interpretation is obtained by considering only events that do affect a chosen particle. The corresponding waiting-time probability density function, $\psi^\star(\tau)$ say, is then related to $\psi(\tau)$. To derive this relationship, we use the associated probability distribution functions $\Psi^\star(\tau)$ and $\Psi(\tau)$. $\Psi^\star(\tau)$ is the probability that the chosen particle encounters no turbulent patch in the time interval $[0,\tau]$. If n turbulent events occur in $[0,\tau]$, this is achieved provided that the particle lies outside each of the n patches. (Correspondingly, $\Psi(\tau)$ is the probability that there are no turbulent events in the fluid column.) Since the average probability for the particle to be outside a patch is $(1-\overline{h}/H)$, we find by summing over the number n of turbulent events that

$$\Psi^\star(\tau) = \Psi(\tau) + \left(1-\frac{\overline{h}}{H}\right)\int_0^\tau \Psi(\tau-\tau')\psi(\tau')\,\mathrm{d}\tau' + \left(1-\frac{\overline{h}}{H}\right)^2 \int_0^\tau\int_0^{\tau'} \Psi(\tau-\tau')\psi(\tau'-\tau'')\psi(\tau'')\,\mathrm{d}\tau'\mathrm{d}\tau'' + \cdots.$$

(Note that the above expression assumes that a turbulent event occurs at $\tau = 0$ as is required to compute the waiting-time between steps other than the first one.) Using the Laplace transform, we obtain, after summing the resulting series, the compact expression

$$\hat{\Psi}^\star(u) = \frac{\hat{\Psi}(u)}{1-(1-\overline{h}/H)\hat{\psi}(u)}$$

from which the expression

$$\hat{\psi}^\star(u) = \frac{\overline{h}\hat{\psi}(u)}{H[1-(1-\overline{h}/H)\hat{\psi}(u)]}$$

for the Laplace-transformed probability density function follows easily. The forms of the first-step statistics $\psi_1^\star(\tau)$ and $\Psi_1^\star(\tau)$ are derived by a similar reasoning, or alternatively by using $\psi_1^\star(\tau) = \Psi^\star(\tau)/\tau_{\rm m}^\star$, where $\tau_{\rm m}^\star = H\tau_{\rm m}/\overline{h}$ is the average waiting time for a given particle (or, equivalently, at a given altitude).

The displacement probability density function associated with $\psi^\star(\tau)$ is then

$$p_\phi^\star(l) = \int_0^H \frac{h}{\overline{h}} p(l|h)\phi(h)\,{\rm d}h.$$

It has no contribution proportional to $\delta(l)$ since it assumes that the particle is inside a turbulent patch.

It is somewhat surprising that the two interpretations do not lead to the same joint distribution: $\psi(\tau)p_\phi(l) \neq \psi^\star(\tau)p_\phi^\star(l)$. However, it can be checked using formula (2.3) that both interpretations lead to the same particle-dispersion statistics $P(z,t)$. This demonstrates the consistency of the two interpretations. Note that the second interpretation avoids the use of the fluid-column height H.

Appendix B. Exact expression for the diffusivity in Dewan's model

Using (2.2), the step variance σ^2 is readily derived for Dewan's model. We find

$$\sigma^2 = \frac{1}{6H}\int_0^\infty h^3\phi(h)\,{\rm d}h =: \frac{\overline{h^3}}{6H}$$

which involves the third-order moment of the turbulent layer height. (This may also be derived by considering the $O(m^2)$ term in (3.6).) Introducing this result in (2.8) provides the exact expression for the diffusivity

$$\kappa = \frac{\overline{h^3}}{12H\tau_{\rm m}}. \tag{B 1}$$

Noting that Dewan (1981*a*) uses constant waiting times Δt leads to $\tau_{\rm m} = \Delta t$, so that using our notation, we can rewrite his expression (2) for the diffusivity as

$$\kappa_{\rm D} = \frac{\overline{h}\,\overline{h^2}}{8H\tau_{\rm m}}, \tag{B 2}$$

where $\overline{h}$ and $\overline{h^2}$ are the first- and second-order moments of turbulent layer height.† Dewan (1980) obtained an alternative expression for $\kappa_{\rm D}$ with a factor 12 replacing 8 in the denominator using a heuristic derivation based on the flux of a tracer; however Dewan (1981*a*) points out that this formula compares less well with the results of numerical simulations. It is now clear from (B 1) that neither expression for $\kappa_{\rm D}$ is correct for general height distributions. The fact that (B 2) compares well with simulations is coincidental and can be explained by Dewan's choice of $\phi(h)$. Indeed, using his distribution of height we find $\overline{h} = 4.84$, $\overline{h^2} = 29.64$ and $\overline{h^3} = 222.76$ and note that these values are such that $\overline{h^3}/12 = 18.56 \approx \overline{h}\,\overline{h^2}/8 = 17.93$.

† For practical use of (B 1), one can follow Dewan (1981*a*) and use the relationship $\tau_{\rm m}H = \tau_{\rm t}\overline{h}/F$, where $\tau_{\rm t}$ is the typical lifetime of turbulent patches and F is the fraction of the fluid column that is turbulent at a given time.

Appendix C. Asymptotic form of η for large β and γ

In this Appendix, we derive a simplified expression for η defined in (3.9) under the assumption that $\beta/\alpha \gg 1$ and $\gamma/\alpha \gg 1$. For simplicity of notation, we temporarily non-dimensionalize all variables using α^{-1} as a time scale – formally this is equivalent to letting $\alpha = 1$. The final results will however be stated using the dimensional variables.

The two parameters β and γ can tend to infinity in various ways, depending on the ratio β/γ. It turns out that the distinguished limit which captures situations in which $\beta \gg \gamma$ as well as those in which $\gamma \gg \beta$ and intermediate cases corresponds to $\beta \sim \gamma \ln \gamma$. We therefore let

$$\beta = s\gamma \ln \gamma,$$

where s is a parameter which is assumed to be of order unity. Substituting this expression for β and changing the integration variable from τ to $x := \exp(-\gamma\tau)$, we rewrite the two integrals involved in the definition of η as

$$\tau_{\mathrm{m}} = \int_0^\infty \Psi(\tau)\,\mathrm{d}\tau = \frac{1}{\gamma^{s+1}} \int_0^1 x^{1/\gamma - 1} \exp(s \ln \gamma\, x)\,\mathrm{d}x,$$

and

$$\chi := \int_0^\infty \tau\Psi(\tau)\,\mathrm{d}\tau = -\frac{1}{\gamma^{s+2}} \int_0^1 x^{1/\gamma - 1} \ln x \exp(s \ln \gamma\, x)\,\mathrm{d}x,$$

with $\eta = \chi/\tau_{\mathrm{m}}^2 - 1$. For large γ, the two integrals are dominated by the contribution of the neighbourhoods of the endpoints. Using Laplace's method we find at leading order

$$\tau_{\mathrm{m}} = \frac{1}{\gamma^s} + \frac{1}{s\gamma \ln \gamma} + O[(s^2\gamma \ln^2 \gamma)^{-1}],$$

and

$$\chi = \frac{1}{\gamma^s} + \frac{1}{s^2\gamma^2 \ln^2 \gamma} + O[(s^3\gamma^2 \ln^3 \gamma)^{-1}].$$

It is clear that either of the two terms in the expression for τ_{m} and in that of χ can be dominant, with the transition corresponding to $s = 1$ for τ_{m} and $s = 2$ for χ. In terms of the original dimensional variables, the average waiting time is rewritten as

$$\tau_{\mathrm{m}} \approx \alpha^{-1}[\exp(-\beta/\gamma) + \alpha\beta^{-1}],$$

while η has the form

$$\eta \approx \frac{\exp(\beta/\gamma)}{[1 + \alpha\beta^{-1} \exp(\beta/\gamma)]^2}.$$

For a given (large) β, the maximum value of η is attained when $\gamma \approx \beta/\ln(\beta/\alpha)$ and is given by $\beta/(4\alpha)$.

REFERENCES

Balluch, M. & Haynes, P. H. 1997 Quantification of lower stratospheric mixing processes using aircraft data. *J. Geophys. Res.* **102**, 23487–23504.

Barenblatt, G. I. 1996 *Scaling, Self-similarity, and Intermediate Asymptotics.* Cambridge University Press.

Brasseur, G. & Solomon, S. 1986 *Aeronomy of the Middle Atmosphere*, 2nd edn. Reidel.

Bühler, O., McIntyre, M. E. & Scinocca, J. F. 1999 On shear-generated gravity waves that reach the mesosphere. Part I: wave generation. *J. Atmos. Sci.* **56**, 3749–3763.

Compte, A. 1997 Continuous time random walks on moving fluids. *Phys. Rev.* E **55**, 6821–6831.

Csanady, G. T. 1964 Turbulent diffusion in stratified environment. *J. Atmos. Sci.* **21**, 439–447.

Csanady, G. T. 1973 *Turbulent Diffusion in the Environment*. D. Reidel.

Dewan, E. M. 1980 A one-dimensional vertical diffusion parameter for extremely inhomogeneous layered turbulence in stratified fluids. *Tech. Rep.* AFGL-TR-80-0186. US Air Force Geophysics Laboratory.

Dewan, E. M. 1981*a* Turbulent vertical transport due to thin intermittent mixing layers in the stratosphere and other stable fluids. *Science* **211**, 1041–1042.

Dewan, E. M. 1981*b* Vertical transport by small scale stratospheric turbulence: a critical review. *Tech. Rep.* AFGL-TR-81-0051. US Air Force Geophysics Laboratory.

Fernando, H. J. S. 1991 Turbulent mixing in stratified fluids. *Ann. Rev. Fluid Mech.* **23**, 455–493.

Gibson, C. H. 1986 Internal waves, fossil turbulence and composite ocean microstructure spectra. *J. Fluid Mech.* **168**, 89–117.

Haynes, P. H. & Anglade, J. 1997 The vertical-scale cascade in atmospheric tracers due to large-scale differential advection. *J. Atmos. Sci.* **54**, 1121–1136.

Hughes, B. D. & Prager, S. 1983 Random processes and random systems: an introduction. In *Mathematics and Physics of Disordered Media* (ed. B. D. Hughes & B. W. Ninham). Lecture Notes in Mathematics, vol. 1035, pp. 1–108. Springer.

Kimura, Y. & Herring, J. R. 1996 Diffusion in stably stratified turbulence. *J. Fluid Mech.* **328**, 253–269.

Ledwell, J. R., Watson, A. J. & Law, C. S. 1993 Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. *Nature* **364**, 701–703.

Montroll, E. W. & Weiss, G. H. 1965 Random walks on lattices. II. *J. Math. Phys.* **6**, 167–181.

Ngan, K. & Shepherd, T. G. 1999 A closer look at chaotic advection in the stratosphere. Part I. Geometric structure. *J. Atmos. Sci.* **56**, 4134–4152.

Pearson, H. J., Puttock, J. S. & Hunt, J. C. R. 1983 A statistical model of fluid-element motions and vertical diffusion in a homogeneous stratified turbulent flow. *J. Fluid Mech.* **129**, 219–250.

Pierrehumbert, R. T. 1995 Tracer microstructure in the large-eddy dominated regime. In *Chaos Applied to Fluid Mixing* (ed. H. Aref & M. S. El Naschie), pp. 347–366. Pergamon.

Polzin, K. 1996 Statistics of the Richardson number – mixing models and fine structure. *J. Phys. Oceanogr.* **26**, 1409–1425.

Polzin, K. L., Toole, J. M., Ledwell, J. R. & Schmitt, R. W. 1997 Spatial variability of turbulent mixing in the abyssal ocean. *Science* **276**, 93–96.

Shlesinger, M. F., Klafter, J. & Wong, Y. M. 1982 Random walks with infinite spatial and temporal moments. *J. Statist. Phys.* **27**, 499–512.

Thorpe, S. A. 1973 Experiments on instability and turbulence in a stratified shear flow. *J. Fluid Mech.* **61**, 731–751.

Toole, J. M., Polzin, K. L. & Schmitt, R. W. 1994 Estimates of diapycnal mixing in the abyssal ocean. *Science* **264**, 1120–1123.

Vanneste, J. & Haynes, P. H. 2000 The role of diffusion and vertical shear for stratospheric tracer spectra. *J. Atmos. Sci.* In preparation.

Waugh, D. W., Plumb, R. A., Elkins, J. W. *et al.* 1997 Mixing of polar air into middle lattitudes as revealed by tracer-tracer scatter plots. *J. Geophys. Res.* **102**, 13119–13134.

Weiss, G. H. & Rubin, R. J. 1983 Random walks: theory and selected applications. *Adv. Chem. Phys.* **52**, 363–505.

Winters, K. B., Lombard, P. N., Riley, J. J. & D'Asaro, E. A. 1995 Available potential energy and mixing in density-stratified fluids. *J. Fluid Mech.* **289**, 115–128.