Particle dispersion by nonlinearly damped random waves

Oliver Bühler[1],† **and Yuan Guo**[1]

[1]Center for Atmosphere Ocean Science at the Courant Institute of Mathematical Sciences,
New York University, New York, NY 10012, USA

We present a theoretical study of the dispersion of particles along quasi-horizontal stratification surfaces induced by small-amplitude internal gravity waves that are forced by white noise and dissipated by linear or nonlinear damping. This extends previous studies in which only linear damping was considered. The damping itself is a toy model for the nonlinear processes that would attenuate a wave mode in a broad spectrum of internal waves such as the Garrett–Munk spectrum for ocean internal waves. We compute the velocity covariance using an eigenfunction expansion of the Kolmogorov backward equation and investigate how the degree of nonlinearity affects the scaling of diffusivity with wave amplitude. We find a simple new expression for the diffusivity that is valid in both the linear and nonlinear cases, and we consider the likely quantitative importance of these process in the context of data from field experiments on small-scale ocean tracer dispersion.

---

1. Introduction

We are interested in the horizontal spreading of particles along stratification surfaces due to the advection by random small-amplitude gravity waves. On the one hand this process poses a fundamental problem in geophysical fluid dynamics in its own right (e.g. Herterich & Hasselmann 1982; Sanderson & Okubo 1988; Weichman & Glazman 2000; Balk, Falkovich & Stepanov 2004; Balk 2006), and on the other hand it might also be relevant to explain part of the actual tracer diffusion observed at very small horizontal scales in the ocean (e.g. Ledwell, Watson & Law 1998; Garrett 2006). The simplest quantity that can quantify this process is the one-particle diffusivity D of Taylor (1921), which in a horizontally isotropic situation can be computed based on a single Cartesian velocity component $U(t)$, say, which the Lagrangian velocity following a fluid particle. If $U(t)$ is a stationary zero-mean random function with covariance function

$$C(s) = C(-s) = \mathbb{E}[U(t)U(t+s)] \quad \text{then } D = \int_0^\infty C(s)\, \mathrm{d}s. \qquad (1.1)$$

† Email address for correspondence: obuhler@cims.nyu.edu

Here $\mathbb{E}$ denotes probabilistic expectation and (1.1) makes obvious how the problem hinges on the statistics of the Lagrangian velocity field $U(t)$.

In a recent paper, Bühler, Grisouard & Holmes-Cerfon (2013) (hereafter BGHC13), it was pointed out that the scaling of D with wave amplitude $a \ll 1$ depends crucially on the presence of explicit wave damping. Specifically, in the unforced and undamped case studied in rotating shallow water in Bühler & Holmes-Cerfon (2009) and in a three-dimensional rotating Boussinesq system in Holmes-Cerfon, Bühler & Ferrari (2011), the diffusivity scaling was $D = O(a^4)$, whilst in the forced–dissipative nonlinear numerical simulations presented in BGHC13 it was clearly found that $D = O(a^2)$. This significant difference could be satisfactorily explained by considering toy models for the forced–dissipative wave dynamics based on stochastic differential equations (SDEs). In these toy models a single wave mode with frequency ω was forced by white noise and dissipated by linear damping. Such linear models can be solved exactly and this allowed the computation of D directly from the theory, and good agreement with the numerical simulations was found.

However, as a model for the real ocean internal wave field, say, white-noise forcing and linear damping are of course poor models for the actual physical processes that underpin the spectral energy cascade through a broad spectrum of internal waves (e.g McComas & Bretherton 1977; Munk 1981). Hence the present paper investigates the robustness of the simple linear models by considering nonlinear variants of the damping function associated with a single wave mode. Crudely, such a damping function seeks to capture the nonlinear energy transfer from the given mode to all other modes of the wave spectrum, and there is no particularly good reason to assume that such a transfer should be linear in wave amplitude.

Our aim here is to investigate this issue by considering a family of nonlinear stochastic wave models (see (3.1) in § 3 below) that have an adjustable degree of nonlinearity indicated by a parameter $p \geqslant 0$. The technically difficult part is that nonlinear SDEs cannot be solved directly anymore, so the covariance function $C(t)$ must be found by other means in order to compute D. In this paper we investigate both an eigenfunction expansion based on the Kolmogorov backward equation (KBE) in § 2.1, which allows $C(t)$ to be computed exactly, and also a simple (but surprisingly accurate) approximation for $C(t)$ in § 2.2.

Overall, we find that the role of the nonlinearity strongly affects the scaling of D with wave amplitude $a \gg 1$, but that its explicit impact on D can be removed if one chooses the expected energy flux and the expected kinetic wave energy as variables to describe the wave mode statistics. In fact, rather unexpectedly, in this way the linear theory can be generalized to nonlinear damping functions. This leads to a simple new expression for D (see (3.7) and (3.10) in § 3 below) that generalizes the earlier results in BGHC13 and indicates their robustness. This simple new formula also allows the oceanic situation to be looked at in quantitative detail, as will be shown in § 4. Concluding remarks are offered in § 5.

2. Analysis of a nonlinear Ornstein–Uhlenbeck process

To introduce the analysis tools we first consider a nonlinear variant of the classical Ornstein–Uhlenbeck (OU) process in one and two dimensions and thereafter add the important physical feature of wave oscillations to the two-dimensional equations in § 3. The one-dimensional OU-process for a Lagrangian velocity component $U(t)$ is defined by the SDE

$$\mathrm{d}U = -\gamma U |U|^p \, \mathrm{d}t + \beta \, \mathrm{d}W. \tag{2.1}$$

Here $W(t)$ is a standard Wiener process (Brownian motion) and the parameters $\gamma > 0$ and $\beta \neq 0$ measure the strength of the damping and of the random forcing. The exponent $p \geqslant 0$ describes the form of the nonlinear damping term and for $p = 0$ the process (2.1) reduces to the linear OU-process studied in BGHC13. For $p > 0$ the instantaneous damping rate varies as $\gamma |U|^p$ and consequently the dimensions of γ depend on p (see (2.8a−d) below). We will be most interested in the cases $p = 1$ (quadratic nonlinearity) and $p = 2$ (quadratic instantaneous damping rate). The former is suggested by the elementary quadratic nonlinearity of the fluid equations, whilst the latter is often invoked in models for a spectrum of inertia–gravity waves in the ocean (e.g. McComas & Bretherton 1977; Munk 1981).

In the linear case $p = 0$ a closed-form solution for $U(t)$ is available, but this does not extend to the nonlinear case $p > 0$. Hence indirect methods are required in order to compute the covariance $C(t)$ and the diffusivity D. First, for all values of $p \geqslant 0$ a stationary state of (2.1) can be readily computed (see § A.1), which is described by the probability density function (p.d.f.)

$$\rho(u) = \frac{1}{Z} \exp\left(-\frac{2}{2+p}\frac{\gamma}{\beta^2}|u|^{2+p}\right) \quad \text{with } Z = \left(\frac{\beta^2}{\gamma}\right)^{1/(2+p)} \Gamma\left(\frac{1}{2+p}\right)\left(\frac{2}{2+p}\right)^{(1+p)/(2+p)}. \tag{2.2}$$

Here $\Gamma(\cdot)$ is the gamma function (see (A 3)). Compared to the linear case $p = 0$, the p.d.f. for $p > 0$ is flattened near the origin and then decreases more sharply for larger values of u, which indicates that the intermittency of U depends on p. The stationary moments of U are

$$\mathbb{E}[|U|^m] = \frac{\Gamma\left(\dfrac{1+m}{2+p}\right)}{\Gamma\left(\dfrac{1}{2+p}\right)} \left(\frac{2+p}{2}\frac{\beta^2}{\gamma}\right)^{m/(2+p)} \quad \text{for } m \geqslant 0. \tag{2.3}$$

In particular, the variance

$$C(0) = \mathbb{E}[U^2] = \frac{\Gamma\left(\dfrac{3}{2+p}\right)}{\Gamma\left(\dfrac{1}{2+p}\right)} \left(\frac{2+p}{2}\frac{\beta^2}{\gamma}\right)^{2/(2+p)}. \tag{2.4}$$

Notably, if $m = 2 + p$ then the identity $\Gamma(x+1) = x\Gamma(x)$ yields the simple result

$$\mathbb{E}[|U|^{2+p}] = \frac{\beta^2}{2\gamma}. \tag{2.5}$$

This is the fluctuation–dissipation relation belonging to (2.1) and it can be derived directly from the SDE by multiplying it by U to form an equation for the energy budget. It is peculiar to stochastic calculus that in this derivation one must retain quadratic terms such as $\mathrm{d}U\,\mathrm{d}U$ in order to get the correct answer (e.g. Gardiner 1997). This yields

$$U\,\mathrm{d}U = \mathrm{d}\left(\frac{U^2}{2}\right) - \frac{\mathrm{d}U\,\mathrm{d}U}{2} = \mathrm{d}\left(\frac{U^2}{2}\right) - \frac{\beta^2}{2}\,\mathrm{d}t = -\gamma|U|^{2+p}\,\mathrm{d}t + \beta U\,\mathrm{d}W \tag{2.6}$$

after using (2.1) and Itô's formula in the form $dW\,dW = dt$. The stationary expectation of (2.6) then yields (2.5) because $\mathbb{E}[U\,dW] = 0$. This makes it obvious that $\beta^2/2$ is the expected generation rate of energy per unit time, whilst $\gamma\mathbb{E}[|U|^{2+p}]$ is the expected energy dissipation rate per unit time, which balances the generation rate in a stationary state. The conventional notation for the energy dissipation rate is ϵ, so

$$\epsilon = \frac{\beta^2}{2} = \gamma\mathbb{E}[|U|^{2+p}] = \gamma O(a^{2+p}). \tag{2.7}$$

This provides a simple and accurate formula that relates the amplitude of U monotonically to the parameter β^2/γ.

Before moving on we note how dimensional analysis can be used to define length and time scales (L, T) from (β, γ). Because U is a velocity, this yields

$$L = \beta^{(2-2p)/(2+p)}\gamma^{-3/(2+p)}, \quad T = (\beta^p\gamma)^{-2/(2+p)} \quad \text{and} \quad \beta = \frac{L}{T^{3/2}}, \quad \gamma = \frac{T^{p-1}}{L^p}. \tag{2.8a--d}$$

The intrinsic velocity and diffusivity scales then follow as

$$\frac{L}{T} = \left(\frac{\beta^2}{\gamma}\right)^{1/(2+p)} \quad \text{and} \quad D \propto \frac{L^2}{T} = \left(\frac{\beta^{2-p}}{\gamma^2}\right)^{2/(2+p)} = \left(\frac{L}{T}\right)^{2-p}\frac{1}{\gamma} = O(a^{2-p})\frac{1}{\gamma}. \tag{2.9a,b}$$

For fixed variance and hence fixed $L/T = O(a)$ the diffusivity D is simply proportional to $1/\gamma$, which extends to all $p \geqslant 0$, a result quoted for the linear OU-process in BGHC13. Interestingly, in the case of quadratic damping rate $p = 2$, the diffusivity becomes independent of the velocity amplitude. However, the nonlinear OU-process of course lacks any wave-like oscillations and hence not much relevance to oceanic waves can be read into these scaling relations.

2.1. *Eigenfunction expansion for the covariance function*

The exact covariance function can be determined via an eigenfunction expansion of the generator of the SDE (2.1), which is the differential operator

$$L = -\gamma u|u|^p\frac{\partial}{\partial u} + \frac{\beta^2}{2}\frac{\partial^2}{\partial u^2}. \tag{2.10}$$

The generator L is less frequently used in physical applications than its adjoint $L^{\dagger}$, which appears in the Fokker–Planck equation (a.k.a. the Kolmogorov forward equation)

$$\frac{\partial\rho}{\partial t} = L^{\dagger}\rho = \frac{\partial}{\partial u}\left(\gamma u|u|^p\rho + \frac{\partial}{\partial u}\left(\frac{\beta^2}{2}\rho\right)\right) \tag{2.11}$$

for the p.d.f. $\rho(u, t)$ of the process. The invariant p.d.f. $\rho(u)$ in (2.2) is in the null-space of the adjoint: $L^{\dagger}\rho = 0$. From a mathematical perspective the generator L is fundamental and has a number of attractive properties (e.g. Oksendal 2013), the most important of which for the present purpose is its role in the KBE. For an autonomous SDE such as (2.1) the KBE takes the following form: consider a function $f(u, t)$ defined for $t \geqslant 0$ by the initial-value problem

$$\frac{\partial f}{\partial t} = Lf \quad \text{and} \quad f(u, 0) = g(u). \tag{2.12a,b}$$

It then follows from probability theory that

$$f(u, t) = \mathbb{E}[g(U(t)) \mid U(0) = u]. \tag{2.13}$$

In other words, if $f(u, t)$ solves the KBE problem in (2.12) then it equals the expected value of $g(U(t))$ at time $t \geqslant 0$ conditioned on $U(0) = u$ at the initial time $t = 0$. Here $g(\cdot)$ is an arbitrary function of u. Hence the KBE allows time-dependent expected values to be computed directly, i.e. without finding the time-dependent p.d.f. of the process first, which is a significant simplification.

This can be used to compute $C(t)$ as follows. We first use the KBE to compute the expected value of $U(t)$ conditioned on starting at a specific location $U(0) = u$. Thereafter we aggregate this result (times u) over all values of u according to the invariant measure. In equations, for $t \geqslant 0$ this means

$$C(t) = \mathbb{E}[U(0)U(t)] = \int_{-\infty}^{+\infty} u \mathbb{E}[U(t) \mid U(0) = u] \rho \, \mathrm{d}u = \int_{-\infty}^{+\infty} u f(u, t) \rho \, \mathrm{d}u \tag{2.14}$$

where $f(u, t)$ solves (2.12) with $g(u) = u$. The actual computation of f utilizes the eigenfunctions $y_k(u)$ and eigenvalues λ_k of L as defined by

$$L y_k = -\lambda_k y_k. \tag{2.15}$$

Multiplication by $\rho(u)$ and using $L^\dagger \rho = 0$ puts (2.15) into self-adjoint form:

$$\frac{\mathrm{d}}{\mathrm{d}u} \left(\frac{\beta^2}{2} \rho \frac{\mathrm{d}y_k}{\mathrm{d}u} \right) + \lambda_k \rho \, y_k = 0. \tag{2.16}$$

This is a standard Sturm–Liouville problem with weight $\rho(u)$, so the eigenfunctions are orthogonal under the inner product with that weight and can be normalized such that

$$\int_{-\infty}^{+\infty} y_k y_l \, \rho \, \mathrm{d}u = \delta_{kl}. \tag{2.17}$$

From here it is straightforward to find (y_k, λ_k) numerically by using a suitable software package such as MATSLISE for matlab (Ledoux, Daele & Berghe 2005). Because $\rho(u)$ decays rapidly at large $|u|$ this can be done using a finite computational domain.

Each eigenfunction generates a separable solution $y_k(u) \exp(-\lambda_k t)$ of the KBE and therefore $f(u, t)$ can be expanded as

$$f(u, t) = \sum_k a_k y_k(u) \mathrm{e}^{-\lambda_k t} \quad \text{where} \quad a_k = \int_{-\infty}^{+\infty} u y_k \rho \, \mathrm{d}u. \tag{2.18}$$

From the second part of (2.14) the covariance and the diffusion are then given by

$$t \geqslant 0: \quad C(t) = \sum_k a_k^2 \mathrm{e}^{-\lambda_k t} \quad \text{and} \quad D = \sum_k \frac{a_k^2}{\lambda_k}. \tag{2.19a,b}$$

The main result (2.19) shows that the covariance is always a sum of exponential functions, albeit with different decay rates given by the eigenvalues λ_k. Numerical results of this eigenfunction expansion for $p = 1$ and $p = 2$ are given in table 1 and illustrated in figure 1. The sum for $C(t)$ in (2.19) converges rapidly and this suggests a simple approximation based on a single exponential function $C_*(t)$, say. This is pursued in the next section.
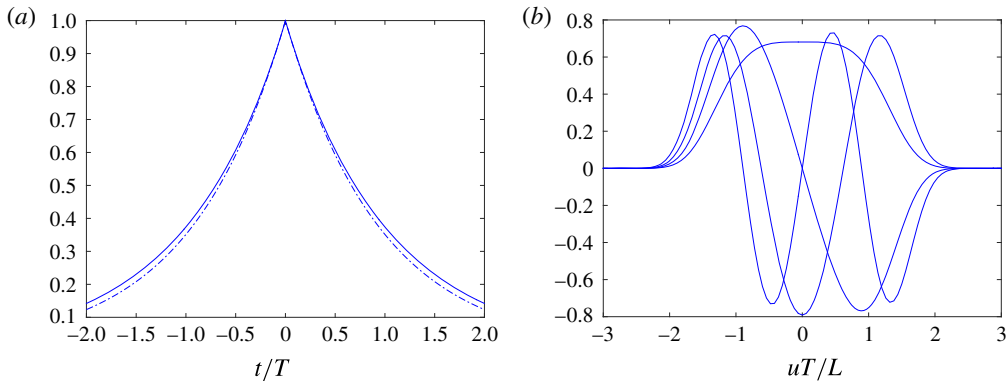
FIGURE 1. (Colour online) (*a*) Non-dimensional exact $C(t)$ computed in §2.1 (solid line) and approximate $C_*(t)$ computed in §2.2 (dot-dashed line) for the process (2.1) with $p = 2$. The approximation is very accurate except for a small underestimate at large lag times. (*b*) Illustration of the first four non-dimensional eigenfunctions y_k for $k \in \{0, 1, 2, 3\}$ for the same case. The weighted functions $y_k\sqrt{\rho}$ are plotted, which are orthogonal under the standard inner product.

	$p = 1$					$p = 1$				
k	λ_k	a_k^2	$\mathbb{E}[U^2]$	a_k^2/λ_k	D	λ_k	a_k^2	$\mathbb{E}[U^2]$	a_k^2/λ_k	D
1	0.98	0.48	0.48	0.49	0.492	0.97	0.47	0.47	0.49	0.486
3	4.70	0.005	0.49	0.001	0.493	5.84	0.007	0.48	0.001	0.488
5	9.32	10^{-6}	0.49	10^{-7}	0.493	12.6	0.0002	0.48	10^{-5}	0.488

TABLE 1. The eigenfunction expansion for $p = 1$ and $p = 2$, and $\beta = \gamma = 1$, which produces non-dimensional eigenvalues and other parameters. Dimensional values are obtained by multiplying the eigenvalues λ_k by $1/T$ and so on. Eigenvalues with even k are neglected because their corresponding eigenfunctions are orthogonal to the initial condition $f(u, 0) = u$ and hence make no contribution to $C(t)$. For $p = 1$ the approximate $D_* = 0.479$ and for $p = 2$ it is $D_* = 0.457$, indicating the remarkable accuracy of the exponential approximation (2.24*a,b*).

2.2. *Exponential approximation for the covariance function*

A natural idea for an approximate covariance function would be to truncate the sum in (2.19) after just one term. However, this would have little practical advantage because one would still need the eigenfunction expansion to compute the parameters a_1 and λ_1. Also, such an approximation would not capture the exact variance $C(0)$ of the process, which is the sum over all a_k^2. It is therefore a better idea to pose an exponential approximation for $C(t)$ at the outset and fit its parameters directly from the behaviour of the exact $C(t)$ near $t = 0$. This side-steps the need for an eigenfunction expansion altogether and turns out to be surprisingly accurate as well. We therefore consider

$$C_*(t) = Ae^{-B|t|} \quad \text{and} \quad D_* = \frac{A}{B} \qquad (2.20a,b)$$

with suitable constants (A, B) determined from the behaviour of $C(t)$ near $t = 0$. Specifically, we require

$$C_*(0) = C(0) \quad \text{and} \quad C'_*(0) = C'(0), \qquad (2.21a,b)$$

where the slope in the second equation is evaluated as $t \to 0+$, say. This is necessary because C' jumps at $t = 0$. The exact slope at the origin can be computed by combining a Taylor expansion for small $\mathrm{d}t > 0$ with the fluctuation–dissipation relation (2.7) as follows:

$$C(\mathrm{d}t) = C(0) + \mathrm{d}t C'(0) = \mathbb{E}[U(t)U(t + \mathrm{d}t)] = \mathbb{E}[U^2] + \mathbb{E}[U\,\mathrm{d}U] \qquad (2.22)$$

$$\Rightarrow \quad C'(0) = \frac{\mathbb{E}[U\,\mathrm{d}U]}{\mathrm{d}t} = -\gamma \mathbb{E}[|U|^{2+p}] = -\frac{\beta^2}{2} = -\epsilon. \qquad (2.23)$$

Hence $A = \mathbb{E}[U^2]$ and $B = \epsilon/\mathbb{E}[U^2]$, which finally yields

$$C_*(t) = \mathbb{E}[U^2] \exp\left(-\frac{\epsilon |t|}{\mathbb{E}[U^2]}\right) \quad \text{and} \quad D_* = \frac{(\mathbb{E}[U^2])^2}{\epsilon}. \qquad (2.24a,b)$$

The corresponding approximation for the auto-correlation time is

$$\tau_* = \frac{D_*}{\mathbb{E}[U^2]} = \frac{\mathbb{E}[U^2]}{\epsilon}. \qquad (2.25)$$

Of course, the functional expressions for D_* and τ_* are merely as one would expect from dimensional analysis based on the kinetic energy $\mathbb{E}[U^2]/2$ and on the energy dissipation ϵ, but the interesting point is that these formulae yield the correct diffusivity values to within a few percent, as was seen in § 2.1. Notably, there are no pre-factors involving the nonlinearity parameter $p \geqslant 0$, so in this approximation two processes with different values of p but equal values of $\mathbb{E}[U^2]$ and ϵ will yield the same diffusivity. This is in fact a robust finding even for wave-like processes, as we shall see.

The comparison of $C(t)$ and $C_*(t)$ in figure 1 shows that the latter matches the former very well, with a small underestimate for large values of t. Overall we find that the exponential approximation (2.24a,b) is surprisingly accurate, with errors less than 5 % or so. This unexpected accuracy of the exponential approximation $C_*(t)$ allows us to simplify a great number of relations below, but we did check that using the exact $C(t)$ would yield essentially the same results in all cases we considered.

2.3. *Two-dimensional nonlinear OU-process*

A natural two-dimensional isotropic generalization of the process (2.1) is

$$\mathrm{d}U = -\gamma U(U^2 + V^2)^{p/2}\,\mathrm{d}t + \beta\,\mathrm{d}W_1, \qquad (2.26a)$$

$$\mathrm{d}V = -\gamma V(U^2 + V^2)^{p/2}\,\mathrm{d}t + \beta\,\mathrm{d}W_2. \qquad (2.26b)$$

Here $\boldsymbol{U} = (U, V)$ is the two-dimensional horizontal velocity vector and (W_1, W_2) are two independent Wiener processes. No new dimensional parameters have been introduced so the dimensional analysis leading to (2.8a−d) and below remains unchanged. The nonlinear drift is a gradient flow with isotropic potential

$$\Phi(U, V) = -\frac{\gamma}{2+p}(U^2 + V^2)^{(2+p)/2}, \qquad (2.27)$$

and hence the two-dimensional invariant p.d.f. has the form (e.g Gardiner 1997)

$$\rho(u, v) = \frac{1}{Z} \exp\left(\frac{2}{\beta^2}\Phi(u, v)\right) \quad \text{with} \quad Z = \pi \left(\frac{\beta^2}{\gamma}\right)^{2/(2+p)} \Gamma\left(\frac{2}{2+p}\right)\left(\frac{2}{2+p}\right)^{p/(2+p)}. \qquad (2.28)$$

Clearly, $\rho(u, v)$ depends only on the velocity magnitude $r = \sqrt{u^2 + v^2}$, but it does not factorize and hence U and V are not independent, except in the linear case $p = 0$. For example, the conditional expectations

$$\mathbb{E}[V \mid U = u] = 0 \quad \text{but} \quad \mathbb{E}[V^2 \mid U = u] \approx \frac{\beta^2}{2\gamma |u|^p} \tag{2.29a,b}$$

for large $|u|$. The remaining analysis works much as it did for the one-dimensional model. First, the stationary moments of U are (see §A.1)

$$\mathbb{E}[|U|^m] = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\dfrac{1+m}{2}\right) \Gamma\left(\dfrac{2+m}{2+p}\right)}{\Gamma\left(\dfrac{2+m}{2}\right) \Gamma\left(\dfrac{2}{2+p}\right)} \left(\frac{2+p}{2} \frac{\beta^2}{\gamma}\right)^{m/(2+p)} \quad \text{for } m \geqslant 0. \tag{2.30}$$

The variance is

$$C(0) = \mathbb{E}[U^2] = \mathbb{E}[V^2] = \frac{1}{2} \frac{\Gamma\left(\dfrac{4}{2+p}\right)}{\Gamma\left(\dfrac{2}{2+p}\right)} \left(\frac{2+p}{2} \frac{\beta^2}{\gamma}\right)^{2/(2+p)}. \tag{2.31}$$

Second, the fluctuation–dissipation relation is

$$\mathbb{E}[\gamma U^2 (U^2 + V^2)^{p/2}] = \mathbb{E}[\gamma V^2 (U^2 + V^2)^{p/2}] = \frac{\beta^2}{2} = \epsilon = \gamma O(a^{2+p}). \tag{2.32}$$

Here ϵ is the energy dissipation per kinetic energy component $U^2/2$ or $V^2/2$, so the total energy dissipation associated with $(U^2 + V^2)/2$ is twice that, or $\beta^2 = 2\epsilon$. Third, the approximate covariance C_* and diffusion D_* take the same functional forms as in (2.24a,b). This is because (2.22) adjusts to

$$C'(0) = \frac{\mathbb{E}[U \, \mathrm{d}U]}{\mathrm{d}t} = -\gamma \mathbb{E}[U^2 (U^2 + V^2)^{p/2}] = -\frac{\beta^2}{2} = -\epsilon, \tag{2.33}$$

using the fluctuation–dissipation relation once more. Finally, the KBE (2.12) now applies to functions $f(\boldsymbol{u}, t)$ in an obvious fashion:

$$f(\boldsymbol{u}, 0) = g(\boldsymbol{u}) \quad \text{and} \quad \frac{\partial f}{\partial t} = Lf \quad \Rightarrow \quad f(\boldsymbol{u}, t) = \mathbb{E}[g(\boldsymbol{U}(t)) \mid \boldsymbol{U}(0) = \boldsymbol{u}]. \tag{2.34a,b}$$

The eigenfunction expansion is now based on the two-dimensional generator

$$L = -\gamma (u^2 + v^2)^{p/2} \left(u \frac{\partial}{\partial u} + v \frac{\partial}{\partial v}\right) + \frac{\beta^2}{2} \left(\frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2}\right), \tag{2.35}$$

which in polar coordinates takes the form

$$L = -\gamma r^{p+1} \frac{\partial}{\partial r} + \frac{\beta^2}{2} \left(\frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2}\right). \tag{2.36}$$

The eigenfunctions of L are separable in (r, θ) so we can consider functions of the form $y_k(r) \exp(\mathrm{i}n\theta)$ with integer n, which yields

$$-\gamma r^{p+1}\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\beta^2}{2}\left(\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}r\frac{\mathrm{d}}{\mathrm{d}r} - \frac{n^2}{r^2}\right)y_k = -\lambda_k y_k. \tag{2.37}$$

Again, multiplication with the weight $r\rho(r)$ puts this equation into self-adjoint form:

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\beta^2}{2}r\rho\frac{\mathrm{d}y_k}{\mathrm{d}r}\right) + \left(\lambda_k - \frac{n^2\beta^2}{2r^2}\right)r\rho y_k = 0. \tag{2.38}$$

The boundary condition at $r = 0$ is

$$\frac{\mathrm{d}y_k}{\mathrm{d}r} = 0 \quad \text{if } n = 0 \text{ and } y_k = 0 \text{ otherwise.} \tag{2.39}$$

For any given n the radial eigenfunctions are orthogonal and normalized such that

$$\int_0^\infty r\rho y_k y_l \, \mathrm{d}r = \delta_{kl}. \tag{2.40}$$

In order to compute the expected value of U we consider

$$f(r, \theta, 0) = u = r\cos\theta, \tag{2.41}$$

which is just the $n = 1$ mode. Hence

$$f(r, \theta, t) = \cos\theta \sum_k a_k y_k(r)\mathrm{e}^{-\lambda_k t} \quad \text{where } a_k = \int_0^\infty r^2\rho y_k \, \mathrm{d}r \tag{2.42}$$

and the y_k are the eigenfunctions of (2.38) with $n = 1$. These are again easily found numerically. The covariance function sought is

$$C(t) = \mathbb{E}[U(0)U(t)] = \int_0^\infty \int_0^{2\pi} r\cos\theta f \rho r \, \mathrm{d}r \, \mathrm{d}\theta \tag{2.43}$$

$$= \int_0^{2\pi} \cos^2\theta \, \mathrm{d}\theta \int_0^\infty \sum_k a_k y_k(r)\mathrm{e}^{-\lambda_k t}\rho r^2 \, \mathrm{d}r = \pi \sum_k a_k^2 \mathrm{e}^{-\lambda_k t} \tag{2.44}$$

and the two-dimensional diffusion coefficient is

$$D = \pi \sum_k \frac{a_k^2}{\lambda_k}. \tag{2.45}$$

We have computed the eigenfunction expansion for $p = 1$ and $p = 2$ but found the same result as in the one-dimensional case: the series for the diffusivity in (2.45) converges rapidly, the exponential approximation to it is accurate to within 5%, and the exponential covariance $C_*(t)$ slightly underestimates the true $C(t)$ for large values of t.

The summary of this investigation is that the exponential approximation gives an excellent prediction of the true covariance in both the one-dimensional and two-dimensional cases and that the value of the nonlinearity parameter $p \geqslant 0$ does not enter the expressions for $C_*(t)$ and D_* when these quantities are expressed in terms of $\mathbb{E}[U^2]$ and ϵ. Notably, ϵ is the energy dissipation rate per energy component $U^2/2$, so in the two-dimensional case the total dissipation range is 2ϵ.

3. Two-dimensional nonlinear process with wave oscillation

We now augment the two-dimensional nonlinear process (2.26*a*) with a simple model for a wave-like oscillation with frequency $\omega \geqslant 0$:

$$\left.\begin{array}{l} \mathrm{d}U = -\gamma U (U^2 + V^2)^{p/2}\,\mathrm{d}t + \omega V\,\mathrm{d}t + \beta\,\mathrm{d}W_1, \\ \mathrm{d}V = -\gamma V (U^2 + V^2)^{p/2}\,\mathrm{d}t - \omega U\,\mathrm{d}t + \beta\,\mathrm{d}W_2. \end{array}\right\} \tag{3.1}$$

In the context of inertia–gravity waves this model can be viewed either as a model for a clockwise inertial oscillation based on a Coriolis parameter $f = \omega$, or as a generic evolution equation for a wave mode in spectral space with frequency ω, say. The former is directly relevant to storm-forced oscillations in the upper ocean whilst the latter is relevant to nonlinear forced–dissipative inertia–gravity wave models as in § 4 of BGHC13. Either way, in terms of the intrinsic scales (L, T) based on (β, γ) in (2.8*a*−*d*), the new frequency ω introduces a non-dimensional parameter ωT, which significantly weakens the conclusions that can be drawn from dimensional analysis alone.

3.1. *Covariance reduction to the non-oscillatory case*

The linear case $(p = 0)$ was considered in BGHC13 and there the outcome was that the covariance of U for $\omega \neq 0$ was simply $\cos(\omega t)$ times the covariance of U for $\omega = 0$. However, this result was derived using the exact solution for the linear system, which does not extend to the nonlinear case. Hence we need to check carefully how the nonlinearity affects the situation. First of all, for all values of ω the system (3.1) has the same invariant measure (2.28) because the oscillatory drift $(\omega V, -\omega U)$ is along lines of constant potential Φ in (2.27). Hence although the Fokker–Planck equation for (3.1) contains new terms, these terms are zero if evaluated on the invariant measure (2.28). For further analysis it is convenient to replace U by the back-rotated complex variable

$$A = A_r + \mathrm{i}A_i = (U + \mathrm{i}V)\,\mathrm{e}^{+\mathrm{i}\omega t} \quad \text{such that} \quad (U, V) = (\mathrm{Re}\,, \mathrm{Im})A\,\mathrm{e}^{-\mathrm{i}\omega t}. \tag{3.2}$$

The system (3.1) then takes the form

$$\mathrm{d}A + \gamma A (A_r^2 + A_i^2)^{p/2}\,\mathrm{d}t = \mathrm{e}^{+\mathrm{i}\omega t}\beta(\mathrm{d}W_1 + \mathrm{i}\mathrm{d}W_2) = \beta(\mathrm{d}W_1 + \mathrm{i}\mathrm{d}W_2). \tag{3.3}$$

Here the second equality holds in the sense that the unitary factor $\exp(\mathrm{i}\omega t)$ can be absorbed without loss of generality in the definition of the complex-valued white noise. This is because the generator of the respective diffusions is identical (see § A.2). But this means that (A_r, A_i) satisfy exactly the $\omega = 0$ nonlinear equations (2.26*a*), for which we know how to solve for the covariance structure. Hence we can write

$$\begin{aligned} C_\omega(t) = \mathbb{E}_\omega[U(0)U(t)] &= \mathbb{E}_\omega[(\mathrm{Re}\,A(0))(\mathrm{Re}\,A(t)\,\mathrm{e}^{-\mathrm{i}\omega t})] \\ &= \cos\omega t\,\mathbb{E}_0[A_r(0)A_r(t)] + \sin\omega t\,\mathbb{E}_0[A_r(0)A_i(t)] \\ &= \cos\omega t\,\mathbb{E}_0[U(0)U(t)] + \sin\omega t\,\mathbb{E}_0[U(0)V(t)], \end{aligned} \tag{3.4}$$

where the subscripts indicate whether the expectation is taken with respect to the dynamics with $\omega \neq 0$ or with $\omega = 0$. The key question is whether $\mathbb{E}_0[U(0)V(t)] = 0$. In the linear case studied in BGHC13 this term was zero because if $\omega = 0$ then U and V were independent zero-mean functions. In the nonlinear case $p > 0$ they are

not independent, but it is still true that their cross-covariance is zero at all time lags. This can either be seen by inspection of the symmetries of the $\omega = 0$ equations, or it can be derived formally from the KBE (2.34). In the latter case we would integrate a solution starting from $f(r, \theta, 0) = v = r \sin \theta$ against a term $u = r \cos \theta$ in (2.43), which vanishes after integration over θ. The upshot is that

$$C_\omega(t) = \cos \omega t \, C_0(t) \tag{3.5}$$

holds exactly for all values of p. From (2.43) we find that

$$t \geqslant 0: \quad C_\omega(t) = \cos \omega t \sum_k \pi a_k^2 e^{-\lambda_k t} \quad \Rightarrow \quad D = \sum_k \pi a_k^2 \frac{\lambda_k}{\lambda_k^2 + \omega^2}. \tag{3.6}$$

This infinite sum for D generalizes the single-term expression (3.10) derived in BGHC13 for the linear case. The corresponding approximate versions are

$$C_*(t) = \mathbb{E}[U^2] \cos \omega t \, e^{-|t|/\tau^E} \quad \text{and} \quad D_* = \frac{\mathbb{E}[U^2]\tau^E}{1 + (\omega \tau^E)^2} \quad \text{where} \quad \tau^E = \frac{\mathbb{E}[U^2]}{\epsilon} \tag{3.7a,b}$$

is the energy damping time scale and $\epsilon = \beta^2/2$. As before, the dependence on the nonlinearity parameter p has disappeared when the model parameters (β, γ) are eliminated in favour of $\mathbb{E}[U^2]$ and ϵ. Notably, if $\omega \neq 0$ then the auto-correlation time

$$\tau_* = \frac{D_*}{\mathbb{E}[U^2]} = \frac{\tau^E}{1 + (\omega \tau^E)^2} \tag{3.8}$$

is always shorter than the energy damping time scale τ^E.

3.2. *High-frequency waves*

The practically relevant regime in the ocean is that of high-frequency waves in the sense that the wave period is short compared with the energy decay time scale of the wave field. This corresponds to

$$(\omega \tau^E)^2 \gg 1 \tag{3.9}$$

and then we obtain the limiting forms

$$\tau_* \approx \frac{1}{\tau^E} \frac{1}{\omega^2} = \frac{\epsilon}{\omega^2 \mathbb{E}[U^2]} \quad \text{and} \quad D_* \approx \frac{\epsilon}{\omega^2}. \tag{3.10a,b}$$

Finally, the fluctuation–dissipation relation (2.32) applied to (3.10) yields the simple scalings

$$\tau^E = O(a^{-p}), \quad \tau_* = O(a^p) \quad \text{and} \quad D_* = O(a^{2+p}), \tag{3.11a-c}$$

with respect to the $O(a)$ wave amplitude. This also recovers the linear results of BGHC13 if $p = 0$.

4. Ocean internal waves

We now investigate how these simple stochastic models for wave-induced diffusion might apply to the kind of small-scale horizontal diffusion in the ocean that was first observed in the NATRE experiment (Ledwell *et al.* 1998) and subsequently confirmed in a number of additional observational campaigns (e.g. Sundermeyer & Ledwell 2001; Shcherbina *et al.* 2014). A suitable reference point is the Garrett–Munk (GM) spectrum for the observed background ocean spectrum of internal waves (e.g. Munk 1981).

Three salient elements are important here: first, there is a systematic nonlinear energy flux ϵ_{GM}, say, that flows through the modal components of the GM-spectrum from small to large vertical wavenumbers. There are several different physical mechanisms for this flux, such as resonant triad interactions or the differential advection of small-scale waves by large-scale waves, but the upshot is that the energy flux ϵ_{GM} eventually leads to instability of the highest vertical wavenumber waves and therefore to wave breaking and wave dissipation. The small vertical wavenumber waves do not break, but their oscillatory dynamics is of course affected by the energy flux ϵ_{GM} flowing through them. This is the kind of nonlinear energy flux through the spectrum that is modelled by the forced–dissipative balance in our model.

Second, all evidence suggests that ϵ_{GM} is quadratic in the wave energy density E_{GM}, say, of the GM-spectrum. In other words, the instantaneous damping rate of the wave field is roughly proportional to its energy. In the context of our model this selects the value $p = 2$ for the nonlinearity parameter in (3.1) and consequently $D = O(a^4)$ according to (3.11c). This is the same scaling of D with wave amplitude as in the unforced inviscid weakly nonlinear study in Holmes-Cerfon *et al.* (2011), but the pre-factor is very different, as we shall see.

Third, typical values for E_{GM} and ϵ_{GM} at equilibrium in the main ocean thermocline are (e.g. Munk 1981; Watanabe & Hibiya 2005)

$$E_{GM} = 0.002 \text{ m}^2 \text{ s}^{-2} \quad \text{and} \quad \epsilon_{GM} = 5 \times 10^{-9} \text{ m}^2 \text{ s}^{-3} \quad \Rightarrow \quad \tau_{GM}^E = \frac{E_{GM}}{\epsilon_{GM}} \approx 5 \text{ days.}$$

$$(4.1a,b)$$

This, together with $\omega \geqslant f$, shows that the high-frequency regime (3.9) is clearly relevant here, because the energy decay time is of the order of a week, which is long compared to the wave periods.

Combing these three elements we assume that there is a common $\epsilon \propto \epsilon_{GM}$ for the modal components of the wave spectrum, that $p = 2$, and that (3.10) together with (4.1) suggests that the wave-induced particle diffusion is dominated by the near-inertial waves $\omega \approx f$. Notably, in the GM-spectrum the near-inertial waves are also the most energetic, but the present argument is based on the constant spectral energy flux ϵ rather than the variable spectral energy density. Let us consider only the diffusion induced by waves with $\omega \approx f$. To use (3.10) for the diffusion we need to relate ϵ to ϵ_{GM}. For near-inertial waves the wave energy is dominated by the horizontal kinetic energy and therefore ϵ_{GM} is the energy flux associated with $U^2/2 + V^2/2$. Our ϵ is the energy flux associated with $U^2/2$ alone, so by isotropy we can argue that

$$\epsilon = \frac{\epsilon_{GM}}{2} \quad \Rightarrow \quad D = \frac{\epsilon}{\omega^2} \approx \frac{\epsilon_{GM}}{2f^2} = 0.25 \text{ m}^2 \text{ s}^{-1} \qquad (4.2)$$

using $f = 0.0001 \text{ s}^{-1}$ and (4.1). On the one hand, this is larger by a factor of ten or so than the $O(a^4)$ inviscid estimate in Holmes-Cerfon *et al.* (2011) (or to the similarly

small estimate based on shear dispersion, see Young, Rhines & Garrett 1982). On the other, it is still somewhat modest when compared to the observed small-scale diffusivity value of 1–2 m^2 s^{-1} in the NATRE experiment (Ledwell *et al.* 1998).

An alternative formula for D that links it directly to the diapycnal diffusivity κ can also be derived, although this runs the risk of making quantities look related that physically have little to do with each other. This is based on the often-used empirical formula

$$\kappa = \Gamma_m \frac{\epsilon_{GM}}{N^2} \quad \Rightarrow \quad D = \frac{\kappa}{2\Gamma_m} \frac{N^2}{f^2} \tag{4.3}$$

where Γ_m is the mixing efficiency of stratified turbulence and N is the buoyancy frequency. Using the typical values $\kappa = 10^{-5}$ m^2 s^{-1}, $\Gamma_m = 0.2$, and $N = 0.01$ s^{-1} then recovers the numerical estimate for D in (4.2).

As noted above, this makes it look as if there was a close link between shear dispersion and the dispersion process considered here, because modulo the pre-factor (4.3) turns out to be identical to the deep-ocean version for shear dispersion quoted in §4d of Young *et al.* (1982). But we believe this is not the case. For example, one could imagine a situation where wave energy sloshes back and forth between different wave modes in a conservative fashion, say between the members of a resonant internal wave triad. In this situation there would be zero shear dispersion because there is no vertical diapycnal diffusion to begin with. However, the process described in this paper would lead to a non-zero horizontal dispersion because the wave–wave interactions upset the exquisite reversibility of the linear wave motion, with consequent particle dispersion in the horizontal.

Finally, it is tempting to speculate on what would happen to these estimates if one were to take the diffusivity due to waves with other frequencies than $\omega = f$ into account. For example, if this were a purely additive process then one could add up contributions of the type ϵ/ω^2 for each octave in frequency between $\omega = f$ and $\omega = N$. However, the rapid decay of $1/\omega^2$ with increasing ω suggests that this would not make a material difference to the estimate in (4.2); numerically we found that it might increase the estimate by some 30 % or so, which is not a big difference.

Overall it seems that internal-wave-induced dispersion can make a noticeable but modest contribution to the observed small-scale horizontal diffusivity in the ocean, leaving the bulk of the phenomenon to be attributed to other processes such as the small-scale tail of the potential-vorticity-controlled balanced flow (e.g. Polzin & Ferrari 2004). Interestingly, as suggested by a referee, the stochastic models studied here with frequency $\omega = 0$ might be a useful modelling tool for the diffusion achieved by such balanced flows.

5. Concluding remarks

In hindsight it is easy to see the connection between the $D = O(a^2)$ scaling of BGHC13 and the $D = O(a^4)$ scaling obtained here: in the high-frequency regime the former estimate for D was multiplied by the linear damping rate and divided by the wave frequency squared. The present estimate allows for a nonlinear amplitude-dependent damping rate of size $1/\tau^E = O(a^p)$ (cf. (3.11*a*)), and hence for $p = 2$ the scaling $D = O(a^4)$ is consistent with the BGHC13 result. As noted before, this is the same scaling as in the inviscid theory in Holmes-Cerfon *et al.* (2011), but with a much larger pre-factor.

The main restriction of the present toy model is of course that it considers only a single wave mode in isolation, with the presence of all other modes crudely

represented by white-noise forcing and nonlinear damping. It would be attractive to have a firm theoretical approach at hand to improve on this, but applications of many-mode theories such as weak wave turbulence theory have been difficult to carry out successfully for internal waves, especially in the presence of the strong Coriolis forces that are inevitable in oceanic applications. This remains an open theoretical question.

We have experimented with direct nonlinear numerical simulations of the rotating shallow-water equations under the constraint of uniform potential vorticity in order to discern the relevant statistics for a single wave mode in the spectrum, but this has not led to clear results so far.

An important theoretical point in this connection is that the wave–wave interactions that are crudely modelled by the nonlinear damping term in (3.1) do not impact on the potential vorticity of the fluid at all. This is in stark contrast to viscous wave dissipation or nonlinear wave breaking, both of which of course do change the potential vorticity locally (e.g. Bühler 2000). A consistent theoretical many-mode model that encompasses both wave and vortex modes would therefore only allow for actual wave dissipation on the smallest scales to impact on the potential-vorticity-controlled vortex flow. In other words, the bulk of the wave energy cascade through a spectrum of internal waves is invisible to the potential vorticity even though there is effective forcing and damping of each wave mode in the cascade. Clearly, it would be highly desirable to have a simple ocean spectral wave and vortex model that is functioning under these constraints. This would be useful not just for studies of diffusivity, but also for the exchange of energy between waves and vortices in general.

Acknowledgements

Appendix A

A.1. *Computations involving invariant measures*

The Fokker–Planck equation for the p.d.f. $\rho(u, t)$ belonging to (2.1) is (e.g. Gardiner 1997)

$$\frac{\partial \rho}{\partial t} = L^{\dagger} \rho = \frac{\partial}{\partial u} \left(\gamma u (u^2)^{p/2} \rho + \frac{\partial}{\partial u} \left(\frac{\beta^2}{2} \rho \right) \right) \tag{A 1}$$

and it follows that (2.2) is its time-independent equilibrium solution. The value of Z follows from the integral identity

$$\int_0^{\infty} x^c \, e^{-ax^b} \, dx = \frac{1}{b} \Gamma \left(\frac{1+c}{b} \right) \left(\frac{1}{a} \right)^{(1+c)/b}, \tag{A 2}$$

which is valid for $a, b > 0$ and $c > -1$. Here $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(1+t) = \int_0^{\infty} x^t \, e^{-x} \, dx. \tag{A 3}$$

To derive (2.30) for the two-dimensional system (2.26a) we start with polar coordinates such that $u = r \cos \theta$ and hence

$$\mathbb{E}[|U|^m] = \frac{1}{Z} \int_0^{2\pi} \int_0^\infty |r \cos \theta|^m e^{-(2\gamma/(2+p)\beta^2)r^{2+p}} r \, dr \, d\theta. \tag{A4}$$

The azimuthal integral is

$$\int_0^{2\pi} |\cos \theta|^m \, d\theta = 4 \int_0^{\pi/2} (\cos \theta)^m \, d\theta = 2\sqrt{\pi} \frac{\Gamma\left(\dfrac{1+m}{2}\right)}{\Gamma\left(\dfrac{2+m}{2}\right)} \tag{A5}$$

and the radial part is an instance of (A 2). Combining with the second equation in (2.28) then yields (2.30).

A.2. *Equivalent diffusions in (3.3)*

We ignore the identical drift terms. If n random variables are forced by m noise terms via

$$dX_i = \sum_{k=1}^m b_{ik} \, dW_k \quad \text{where } X \in R^n, b \in R^{n \times m}, W \in R^m \tag{A6}$$

then the generator is given by (e.g. Oksendal 2013)

$$L = \sum_{i=1}^n \sum_{j=1}^n \frac{c_{ij}}{2} \frac{\partial^2}{\partial x_i \partial x_j} \quad \text{where } c = bb^{\mathrm{T}} \in R^{n \times n}. \tag{A7}$$

In the present case $m = n = 2$, $(X_1, X_2) = (A_r, A_i)$, and hence (3.3) holds because

$$b = \beta \begin{pmatrix} \cos f t & -\sin f t \\ \sin f t & \cos f t \end{pmatrix} \quad \Rightarrow \quad c = bb^{\mathrm{T}} = \beta^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{A8}$$

REFERENCES

BALK, A. M. 2006 Wave turbulent diffusion due to the Doppler shift. *J. Stat. Mech.* P08018.

BALK, A. M., FALKOVICH, G. & STEPANOV, M. G. 2004 Growth of density inhomogeneities in a flow of wave turbulence. *Phys. Rev. Lett.* **92**, 244504.

BÜHLER, O. 2000 On the vorticity transport due to dissipating or breaking waves in shallow-water flow. *J. Fluid Mech.* **407**, 235–263.

BÜHLER, O., GRISOUARD, N. & HOLMES-CERFON, M. 2013 Strong particle dispersion by weakly dissipative random internal waves. *J. Fluid Mech.* **719**, R4.

BÜHLER, O. & HOLMES-CERFON, M. 2009 Particle dispersion by random waves in rotating shallow water. *J. Fluid Mech.* **638**, 5–26.

GARDINER, C. W. 1997 *Handbook of Stochastic Methods.* Springer.

GARRETT, C. 2006 Turbulent dispersion in the ocean. *Prog. Oceanogr.* **70** (2), 113–125.

HERTERICH, K. & HASSELMANN, K. 1982 The horizontal diffusion of tracers by surface waves. *J. Phys. Oceanogr.* **12**, 704–712.

HOLMES-CERFON, M., BÜHLER, O. & FERRARI, R. 2011 Particle dispersion by random waves in the rotating Boussinesq system. *J. Fluid Mech.* **670**, 150–175.

LEDOUX, V., DAELE, M. V. & BERGHE, G. V. 2005 Matslise: a matlab package for the numerical solution of Sturm–Liouville and Schrödinger equations. *ACM Trans. Math. Softw. (TOMS)* **31** (4), 532–554.

LEDWELL, J., WATSON, A. & LAW, C. 1998 Mixing of a tracer in the pycnocline. *J. Geophys. Res.* **103** (C10), 21499–21529.

MCCOMAS, C. H. & BRETHERTON, F. P. 1977 Resonant interaction of oceanic internal waves. *J. Geophys. Res.* **82** (9), 1397–1412.

MUNK, W. 1981 Internal waves and small-scale processes. In *Evolution of Physical Oceanography* (ed. B. A. Warren & C. Wunsch), pp. 264–291. MIT Press.

OKSENDAL, B. 2013 *Stochastic Differential Equations: An Introduction with Applications*. Springer Science & Business Media.

POLZIN, K. & FERRARI, R. 2004 Isopycnal dispersion in NATRE. *J. Phys. Oceanogr.* **34**, 247–257.

SANDERSON, B. G. & OKUBO, A. 1988 Diffusion by internal waves. *J. Geophys. Res.* **93**, 3570–3582.

SHCHERBINA, A. Y., SUNDERMEYER, M. A., KUNZE, E., D'ASARO, E., BADIN, G., BIRCH, D., BRUNNER-SUZUKI, ANNE-MARIE, E. G., CALLIES, J., KUEBEL, C., BRANDY, T., CLARET, M. *et al.* 2014 The latmix summer campaign: Submesoscale stirring in the upper ocean. A part of ams special collection latmix: studies of submesoscale stirring and mixing. *Bull. Am. Meteorol. Soc.* **96**, 1257–1279.

SUNDERMEYER, M. A. & LEDWELL, J. R. 2001 Lateral dispersion over the continental shelf: analysis of dye release experiments. *J. Geophys. Res.: Oceans* **106** (C5), 9603–9621.

TAYLOR, G. I. 1921 Diffusion by continuous movements. *Proc. Lond. Math. Soc.* **20**, 196–212.

WATANABE, M. & HIBIYA, T. 2005 Estimates of energy dissipation rates in the three-dimensional deep ocean internal wave field. *J. Oceanogr.* **61** (1), 123–127.

WEICHMAN, P. & GLAZMAN, R. 2000 Passive scalar transport by travelling wave fields. *J. Fluid Mech.* **420**, 147–200.

YOUNG, W. R., RHINES, P. B. & GARRETT, C. J. 1982 Shear-flow dispersion, internal waves and horizontal mixing in the ocean. *J. Phys. Oceanogr.* **12**, 515–527.