J. Fluid Mech. (2002), *vol.* 467, *pp.* 163–203.
DOI: 10.1017/S0022112002001337 Printed in the United Kingdom

Anomalous behaviour of a passive tracer in wave turbulence

By ALEXANDER M. BALK

Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, UT 84112, USA

(Received 27 March 2001 and in revised form 3 April 2002)

We consider the behaviour of a passive tracer in multiscale velocity field, when there is no separation of scales; the energy spectrum of the velocity field extends into the region of long waves and even can be singular there. We suppose that the velocity field is a superposition of random waves. The turbulence of various ocean or atmospheric waves provides examples. We find anomalous diffusion (sub- and super-diffusion), anomalous drift (super-drift), and anomalous spreading of a passive tracer cloud. For the latter we find the existence of two regimes: (i) 'close' passive tracer particles diverge sub- or supper-exponentially in time, and (ii) a 'large' passive tracer cloud spreads as a power-law in time. The exponents, as well as the corresponding pre-factors, are found. The theory is confirmed by numerical simulations.

1. Introduction

Many studies (see e.g. Bensoussan, Lions & Papanicolaou 1978; Sheng 1995; Majda & Kramer 1999; Cherkaev 2000; Milton 2000; Sobczyk & Kirkner 2001) consider media with microstructure, which has variations on some microscale l, whereas the macroproperties, on some macroscale L, are under investigation. It is assumed that $l \ll L$. To study such situations, the effective medium approximation is developed. Sometimes the medium has several microscales, all of them being much smaller than the macroscale L. Sometimes the variations on the macroscale are also included, which are taken into account by some procedures, such as WKB. What if the medium has variations on all scales from microscale l to macroscale L? This situation occurs in several practical problems (see also §8). This paper considers such a situation. We chose the passive tracer problem because of the simplicity of the underlying equation. When the molecular diffusion is negligible, the passive tracer equation can be reduced to an ordinary differential equation, and so, the theoretical predictions can be effectively tested by numerical simulations.

Consider the dynamics of a passive tracer in a random velocity field $\boldsymbol{v}(\boldsymbol{x}, t)$, so that the concentration $\varphi(\boldsymbol{x}, t)$ of the tracer obeys the equation

$$\frac{\partial \varphi}{\partial t} + \nabla[\boldsymbol{v}(\boldsymbol{r}, t)\varphi] = \kappa \nabla^2 \varphi, \tag{1.1}$$

where κ is the molecular diffusion coefficient. In this paper we assume that the molecular diffusion is negligible (infinite Péclet number) – as often happens in practical problems. Then the theoretical predictions can be tested by numerical simulations of the ordinary differential equation $\dot{\boldsymbol{r}} = \boldsymbol{v}(\boldsymbol{r}, t)$. The statistics of the velocity field $\boldsymbol{v}(\boldsymbol{x}, t)$ is given. The problem is to find the statistics of the passive tracer field $\varphi(\boldsymbol{x}, t)$, in particular, its first statistical moments.

This problem of passive tracer was intensively studied for many years (see e.g. reviews Batchelor & Townsend 1956; Majda & Kramer 1999; and references cited therein), often under the assumption of well-separated scales. The long-time large-scale behaviour of the passive tracer concentration is considered while the velocity field is short-range correlated. In particular, several authors studied the white noise model (Kraichnan 1968), when the velocity field is delta-correlated in time. Some studies also consider the opposite limit of well-separated scales. The passive tracer is observed during 'short' time, much shorter than the characteristic times of the velocity field (see e.g. Batchelor & Townsend 1956).

In this paper, we consider the intermediate situation. The observation time t is well inside the range of time scales of the velocity field. The energy spectrum of the velocity field has long inertial range, so that some time scales of the velocity field are shorter than t, and some are longer than t. As time t progresses, more and more time scales of the velocity field become shorter than t. We call this situation anomalous; in particular, it leads to the anomalous diffusion and anomalous drift.

Avellaneda & Majda (1990, 1992) have introduced a shear flow model in order to analyse the effect of large scales of the velocity field on the behaviour of the passive tracer. Unlike their study, we do not assume any specific geometry of the fluid flow. However, we assume that the velocity field is a superposition of random waves

$$\boldsymbol{v}(\boldsymbol{r},t) = \mathrm{Re}\int c_{\boldsymbol{k}}\boldsymbol{b}_{\boldsymbol{k}}\exp(\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\Omega_{\boldsymbol{k}}t))\,\mathrm{d}\boldsymbol{k}. \tag{1.2}$$

Here $\boldsymbol{k}$ is a wave vector, $\Omega_{\boldsymbol{k}}$ is the dispersion law; $c_{\boldsymbol{k}}$ are the wave amplitudes, which are time independent random variables; and $\boldsymbol{b}_{\boldsymbol{k}}$ is the polarization vector, which defines the multidimensional motion of the fluid particles. In particular, the flow can be compressible (if the flow is incompressible, then $\boldsymbol{k}\cdot\boldsymbol{b}_{\boldsymbol{k}} = 0$). The dimension d of the medium can be arbitrary; Re denotes the real part.

The behaviour of a passive tracer in a wave field was considered by Herterich & Hasselmann (1982), Weichman & Glazman (1999, 2000, 2002), Balk & McLaughlin (1999) under the assumption of well-separated scales, when the characteristic wavelength is much smaller than the transport distance of a tracer particle.

In this paper, we consider the situation when the energy spectrum of the velocity field extends into the region of long waves (and even can be singular there), as well as in the region of short waves; so, there is no separation of scales. We can say that our model attempts to resolve the delta function $\delta(t-t')$ of the velocity correlations in the white noise model (if the observation time t is much larger than the longest time-scale of the velocity field, then we can assume that the velocity field is delta-correlated in time; for shorter observation time t we must take into account the details of the time correlations).

The assumption that the velocity field is a superposition of waves (1.2) allows us to approach this problem with perturbation methods. We assume that the waves travel fast whereas the fluid particles move slowly. As usual in dynamical problems, the perturbational approach is not straightforward, since the small perturbations can lead to large effects over long times. In this paper, we develop the statistical near-identity transformation. It is motivated by the usual near-identity transformation (see e.g. Bogoliubov & Mitropolsky 1961; Sanders & Verhulst 1985) from the theory of dynamical systems and by the Wiener–Hermite expansion (see e.g. Wiener 1958; Eftimiu 1990).

This approach can be applied not only to the passive tracer problem (1.1), but to various wave turbulence problems. Moreover, the calculations in the general situation turn out to be simpler since we do not need to pay attention to the details of the specific situation of the passive tracer. We consider the passive tracer equation (1.1) as a representative of a class of linear dynamical equations with random coefficients (§ 2). We show that this class includes, for example, the Schrödinger equation with random potential.

After we describe the statistical near-identity transformation in § 2, we discuss the evolution of the average concentration in § 3 and the behaviour of the two point (same time) correlation function in §§ 4, 5, 6 and 7. Section 4 contains material applicable to the general situation (described in § 2.1.3), and in § 5 we specify the general results to the spreading of the passive tracer cloud. We find two regimes of anomalous spreading: (i) for small clouds (§ 6) and (ii) for large clouds (§ 7).

In the following subsections, we discuss physical questions, the basic scales characteristic to the passive tracer problem, and our main physical predictions.

1.1. *Physical questions*

Our investigation centres around the following two scenarios.

(i) Imagine that somewhere in the ocean we have put a small piece of wood. How will it be transported (on average) by the turbulence of the ocean waves (of various types)? Alternatively, we can think about an air balloon put somewhere in the atmosphere. How will it be transported by the turbulence of atmospheric waves?

This is the question about the evolution of the average concentration $\langle\phi(\boldsymbol{r},t)\rangle$, $\langle\ldots\rangle$ denotes the average over the ensemble of the velocity fields.

We normalize the function $\varphi(\boldsymbol{r},t)$ so that it is a probability of finding a particle at point $\boldsymbol{r}$ at instant t:

$$\int \varphi(\boldsymbol{r},t)\,\mathrm{d}\boldsymbol{r} = 1. \tag{1.3}$$

We consider the mean displacement of a tracer particle

$$\boldsymbol{R} = \langle \boldsymbol{r}\rangle = \int \boldsymbol{r}\langle\varphi(\boldsymbol{r},t)\rangle\,\mathrm{d}\boldsymbol{r} \tag{1.4}$$

and the variance tensor of the mean square displacement

$$\boldsymbol{D} = \langle(\boldsymbol{r}-\boldsymbol{R})(\boldsymbol{r}-\boldsymbol{R})'\rangle = \int (\boldsymbol{r}-\boldsymbol{R})(\boldsymbol{r}-\boldsymbol{R})'\langle\varphi(\boldsymbol{r},t)\rangle\,\mathrm{d}\boldsymbol{r} \tag{1.5}$$

(the prime denotes a transposed matrix; so that $\boldsymbol{r}'$ is a row, while $\boldsymbol{r}$ is a column). The mean square displacement is

$$\sigma^2 = \langle|\boldsymbol{r}-\boldsymbol{R}|^2\rangle = \mathrm{Trace}(\boldsymbol{D}). \tag{1.6}$$

We find the anomalous diffusion and the anomalous drift of the tracer.

(ii) Imagine that oil is spilled in the ocean. How will the oil spot spread? Alternatively, some pollution was released in the atmosphere. How will the polluted area spread owing to the various atmospheric waves? What is the characteristic size of the polluted area as a function of time?

These questions are related to the evolution of the two point (same time) correlation function $\langle\phi(\boldsymbol{r}_1,t)\phi(\boldsymbol{r}_2,t)\rangle$. In particular, we consider the variance tensor of relative displacement between two particles

$$\boldsymbol{S} = \langle(\boldsymbol{r}_2-\boldsymbol{r}_1)(\boldsymbol{r}_2-\boldsymbol{r}_1)'\rangle = \int (\boldsymbol{r}_2-\boldsymbol{r}_1)(\boldsymbol{r}_2-\boldsymbol{r}_1)'\langle\varphi(\boldsymbol{r}_1,t)\varphi^*(\boldsymbol{r}_2,t)\rangle\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2. \tag{1.7}$$

Although in the passive tracer case the function φ is real, we use complex conjugated φ^* in this equation, so that the matrix $\boldsymbol{S}$ is real in the general situation (see §§ 2.1.2 and 2.1.3). The diameter ρ of the passive tracer cloud is

$$\rho^2 = \langle |\boldsymbol{r}_2 - \boldsymbol{r}_1|^2 \rangle = \mathrm{Trace}(\boldsymbol{S}). \tag{1.8}$$

The main physical goal of this paper is to find the evolution of the quantities $\boldsymbol{R}(t)$, $\boldsymbol{D}(t)$, and $\boldsymbol{S}(t)$. We study the anomalous situation, when the observation time t is well inside the range of time scales of the velocity field.

1.2. *Basic scales*

In order to use perturbations, we assume that the velocity of the fluid particles is much smaller than the velocity of waves. If the velocity field were characterized by some length scale, then we would naturally define the small parameter as

$$\epsilon = \frac{\text{characteristic speed of the fluid particles}}{\text{characteristic group velocity of the waves}}. \tag{1.9}$$

Such a situation takes place when the energy spectrum $\varepsilon_{\boldsymbol{k}}$ of the velocity field has a clear pick near some wavenumber K_0 and vanishes away from K_0, e.g. $\varepsilon_{\boldsymbol{k}} = 0$ when $k < \frac{1}{2}K_0$ or $k > 2K_0$. Then as a characteristic wave speed, we can take the group velocity $\partial\Omega/\partial k$ evaluated at the point K_0, while as a characteristic fluid speed we can take the root mean square velocity u,

$$u^2 = \mathscr{E} = \int \varepsilon_{\boldsymbol{k}} \, \mathrm{d}\boldsymbol{k}.$$

However, in several practical situations the energy spectrum extends through a wide range of wavenumbers. What is the small parameter in such situations?

For example, the dispersion law $\Omega_{\boldsymbol{k}}$ and the energy spectrum $\varepsilon_{\boldsymbol{k}}$ can be power-law functions of the wavenumber $k = |\boldsymbol{k}|$ (with some possible angular dependence) in a large inertial range:

$$\Omega_{\boldsymbol{k}} = \mathscr{A} k^{\alpha} \Phi\left(\frac{\boldsymbol{k}}{k}\right) \quad (\alpha > 0), \tag{1.10}$$

$$\varepsilon_{\boldsymbol{k}} = \mathscr{C} k^{-\gamma} \Psi\left(\frac{\boldsymbol{k}}{k}\right) \text{ in some inertial interval } K_{min} \ll k \ll K_{max}. \tag{1.11}$$

We assume that the exponent α is positive, but the exponent γ can be arbitrary (positive, or negative, or zero). Since we have introduced the constants $\mathscr{A}$ and $\mathscr{C}$, we can assume that the functions Φ and Ψ (defining the angular dependence) are dimensionless, and their integrals over the angular variables are equal to one.

The formulae (1.10) and (1.11) introduce dimensional parameters $\mathscr{A}$ and $\mathscr{C}$ ($\mathscr{A}$ has dimension m^{α}/s, and $\mathscr{C}$ has dimension $m^{2+d-\gamma}/s^2$). By dimensional considerations these constants define the length scale ξ and the time scale τ

$$\xi = (\mathscr{C}/\mathscr{A}^2)^{1/\delta}, \quad \tau = (1/\mathscr{A})\xi^{\alpha} \quad \text{where} \quad \delta = 2 + d - \gamma - 2\alpha. \tag{1.12}$$

The applicability of the perturbational approach depends on the relation of the length scale ξ to the largest and smallest length scales of the velocity field, namely largest and smallest wavelengths:

$$\Lambda_{max} = \frac{1}{K_{min}}, \quad \Lambda_{min} = \frac{1}{K_{max}}. \tag{1.13}$$

In other words, the applicability of the perturbational approach depends on the relation of the time scale τ to the largest and smallest time scales of the velocity field, namely the largest and smallest wave periods:

$$T_{max} = \frac{1}{\Omega_{min}} = \frac{1}{\mathscr{A} K_{min}^{\alpha}}, \quad T_{min} = \frac{1}{\Omega_{max}} = \frac{1}{\mathscr{A} K_{max}^{\alpha}}. \tag{1.14}$$

In accordance with (1.9), we could require the root mean square velocity u to be much smaller than the slowest wave speed. However, this requirement is too strict. For instance, if the energy $\mathscr{E}$ diverges at small scales, the particle velocity u is infinite. However, small scales give only an insignificant contribution to the turbulent transport, and so, the perturbational approach can be applicable.

It turns out to be sufficient to compare the fluid velocity and the wave velocity only at the same scale. So we define the small parameter ϵ in the following way. Let $\mathscr{E}_K$ be the energy of scale K, i.e. the energy in the shell $\frac{1}{2}K < |\boldsymbol{k}| < 2K$:

$$\mathscr{E}_K = \int_{K/2<|\boldsymbol{k}|<2K} \varepsilon_{\boldsymbol{k}} \, \mathrm{d}\boldsymbol{k}.$$

Then ϵ^2 is the maximum of the ratio of this energy to the square group velocity of the waves with the same wave number:

$$\epsilon^2 = \max_{K_{min}<K<K_{max}} \frac{\mathscr{E}_K}{(\partial\Omega/\partial K)^2}.$$

In the power law situation this definition gives

$$\epsilon^2 = \frac{\mathscr{C}}{\mathscr{A}^2} K^{\delta} \frac{1}{\alpha^2(d-\gamma)} \left(2^{d-\gamma} - \frac{1}{2^{d-\gamma}}\right) \quad \text{where} \quad K = \begin{cases} K = K_{min} & \text{if } \delta < 0, \\ K = K_{max} & \text{if } \delta > 0. \end{cases}$$

The sign of the exponent δ determines which side of the inertial range (of the velocity field) the small parameter depends on. If $\delta < 0$, then ϵ depends on the large scale cutoff; if $\delta > 0$, then ϵ depends on the small scale cutoff. In this paper, we concentrate on the situation of negative δ (Balk (2000) contains an example of the opposite situation, when $\delta > 0$). So, without the constant factor, the small parameter is

$$\epsilon^2 = \frac{\mathscr{C}}{\mathscr{A}^2} K_{min}^{\delta} = (\xi K_{min})^{\delta}. \tag{1.15}$$

Even this definition of a small parameter can often be too restrictive. For instance, it may happen that the perturbational approach is applicable only during some time. During the time t a tracer particle 'feels' random waves with wavenumbers k greater than $K(t) = (1/\mathscr{A}t)^{1/\alpha}$. They are given by the condition $1/\Omega_{\boldsymbol{k}} < t$. So, instead of K_{min} in (1.15), we can put $K(t)$:

$$\epsilon^2 = [\xi K(t)]^{\delta} = \left(\frac{t}{\tau}\right)^{-\delta/\alpha}. \tag{1.16}$$

In § 5.1.1, we suggest another applicability condition.

The perturbational approach is applicable even if the total energy $\int \varepsilon_{\boldsymbol{k}} \, \mathrm{d}\boldsymbol{k}$ diverges at large scales, and the characteristic particle velocity u is infinite. This happens when $\gamma > d$. In this situation, the mean square displacement, (1.6), is infinite, but the mean square distance between two tracer particles can be finite (since large scales convect the two particles in a similar way). So the characteristic size of a tracer cloud can remain finite, and we find its evolution.

1.3. *Physical predictions*

The velocity field is characterized by a wide range of the space scales Λ and time scales T:

$$\Lambda_{min} \ll \Lambda \ll \Lambda_{max}, \quad T_{min} \ll T \ll T_{max} \tag{1.17}$$

(see (1.13) and (1.14)). The behaviour of the passive tracer crucially depends on the observation scales.

If we observe the passive tracer long enough, during the time $t \gg T_{max}$, then we have normal diffusion, similar to Brownian motion. This is the situation of well-separated scales; long-time behaviour of the passive tracer in a short-range correlated velocity field was considered by Herterich & Hasselmann (1982), Weichman & Glazman (1999, 2000, 2002) and Balk & McLaughlin (1999). In the first order ϵ, fluid particles in wave motion have closed orbits (this fact is well-known for sea waves); the particle displacement over the wave period is of order ϵ^2, and $\boldsymbol{D}$ should be of order ϵ^4. So, in the order ϵ^2 – which we consider in this paper – the mean square displacement stays constant: $\boldsymbol{D} = \text{const}$.

If we observe the passive tracer during the short time $t \ll T_{min}$, then the passive tracer 'feels' that it is transported by a constant (but random) velocity field, and so, $\boldsymbol{D}$ behaves like t^2, $\boldsymbol{D} \propto t^2$ (cf. Batchelor & Townsend 1956).

If we observe the passive tracer during the intermediate time t, $T_{min} \ll t \ll T_{max}$, then the tracer exhibits anomalous behaviour, which we study in this paper. Practical situations are provided by 'large-scale waves', such as Rossby waves. They have time scales from months to a few years (see e.g. Gill 1982). In this situation, the observation time t could be of the order of seasons.

During the intermediate time, short waves, i.e. waves with periods $T \ll t$, lead to normal diffusion. On the other hand, long waves, i.e. waves with periods $T \gg t$, convect passive tracer particles. So, the orbits of the fluid particles are not closed even at the lowest order ϵ. Therefore, the displacement of the passive tracer particles is of the order ϵ for each realization of the velocity field. However, when we average over the ensemble of velocity fields with random phases (of the complex wave amplitudes), the mean displacement vanishes at order ϵ and turns out to be of the order of ϵ^2.

1.3.1. *Anomalous diffusion*

While the normal diffusion vanishes at order ϵ^2 and appears only at order ϵ^4, the anomalous diffusion displays itself already at order ϵ^2.

Thus, at order ϵ^2, the mean square displacement, (1.6), behaves in the following way. If $t \ll T_{min}$, then $\sigma^2 \propto t^2$. If $t \gg T_{max}$, then $\sigma^2 \propto t^2$. In the intermediate time interval $T_{min} \ll t \ll T_{max}$ we can expect anomalous diffusion with $\sigma^2 \propto t^\lambda$, $0 < \lambda < 2$.

We find that the mean square displacement, (1.6), behaves like

$$\sigma^2 = \begin{cases} \sigma_1^2 & \text{if} \quad \gamma \leqslant d - 2\alpha, \\ \sigma_1^2 t^\lambda & \text{if} \quad d - 2\alpha \leqslant \gamma \leqslant d \\ \infty & \text{if} \quad \gamma \geqslant d. \end{cases} \qquad \text{where } \lambda = 2 + \frac{\gamma - d}{\alpha},$$

We have sub-diffusion when $0 < \lambda < 1$, and super-diffusion when $1 < \lambda < 2$.

Along with the exponent λ, we find the pre-factor σ_1^2. We also obtain a formula for the entire variance tensor (1.5). See formulae (3.8) for the one-dimensional case and (3.17) for the two-dimensional case. In a general situation (when the dispersion law and the energy spectrum are not necessarily power-law functions) the variance tensor is given by the formulae (3.4) or (3.2); the latter allows for the frequency spreading around $\omega = \Omega_{\boldsymbol{k}}$.

Balk (2001*a*, *b*) presented the numerical confirmation of this prediction. Figure 1 shows the numerical confirmation of our prediction over a five-decade time interval, during which σ increases 10^4 times. We have agreement not only in the exponent λ, but also in the pre-factor σ_1^2.

1.3.2. *Anomalous drift*

Besides the anomalous diffusion, we find anomalous drift as well.

Even if the average velocity is zero, the mean displacement (1.4) can be non-zero. For this it is necessary that the velocity field is anisotropic and compressible (see § 3.2.2).

If $t \gg T_{max}$, then we have the situation of well-separated scales and normal drift with $\dot{\boldsymbol{R}} \propto t$, like in anisotropic Brownian motion.

If $t \ll T_{min}$, then the passive tracer is transported by an 'almost' constant velocity field, and therefore, $\dot{\boldsymbol{R}} \propto t^2$. To see this, we solve the initial-value problem

$$\dot{\boldsymbol{r}} = \boldsymbol{v}(\boldsymbol{r}, t), \quad \boldsymbol{r}(0) = \boldsymbol{r}_0;$$

by Picard's iterations

$$\boldsymbol{r}(t) = \boldsymbol{r}_0 + \int_0^t \boldsymbol{v}(\boldsymbol{r}_0, \eta)\,\mathrm{d}\eta + O(t^2)\ (t \to 0).$$

Now we average this expression over the ensemble of the velocity fields. Assuming that the velocity field has zero mean ($\langle \boldsymbol{v}(\boldsymbol{r}_0, \eta)\rangle = 0$), we have $\langle \boldsymbol{r}(t)\rangle = \boldsymbol{r}_0 + O(t^2)$.

In the intermediate time interval $T_{min} \ll t \ll T_{max}$, we can expect anomalous drift with $\boldsymbol{R} \propto t^{\mu}$, $1 < \mu < 2$.

We find that the mean displacement, (1.4), grows like

$$\boldsymbol{R} = \begin{cases} \boldsymbol{R}_1 t & \text{if} \quad \gamma \leqslant d + 1 - \alpha \\ \boldsymbol{R}_1 t^{\mu} & \text{if} \quad d + 1 - \alpha \leqslant \gamma \leqslant d + 1 \end{cases} \qquad \text{where } \mu = 2 + \frac{\gamma - d - 1}{\alpha} = \lambda - \frac{1}{\alpha}.$$

Since $\mu \geqslant 1$, we can have only a super-drift. It occurs only when there is super-diffusion ($1 < \lambda < 2$), and $\alpha < 1$.

Along with the exponent μ, we also find the constant pre-factor $\boldsymbol{R}_1$, see formulae (3.10) for the one-dimensional case and (3.18) for the two-dimensional case. In a general situation (when the dispersion law and the energy spectrum are not necessarily power-law functions) the drift is given by the formulae (3.4) or (3.2); the latter allows for the frequency spreading around $\omega = \Omega_{\boldsymbol{k}}$.

Balk (2001) presented the numerical confirmation of this result.

The mean displacement, (1.4), and the mean square displacement, (1.5), are characteristics of the one-particle statistics. The two-particle statistics, and its primary characteristic – the relative mean square displacement (1.7) and (1.8) are also practically important.

The behaviour of two-particle statistical quantities crucially depends on the relation of the distance s between the particles and the observation time t to the scales (1.17) of the velocity field.

If $s \gg \Lambda_{max}$ then the two tracer particles are statistically independent, and therefore

$$\dot{\mathsf{S}} = 2\dot{\boldsymbol{D}} \Rightarrow \dot{\rho}^2 = 2\dot{\sigma}^2.$$

If $s \ll \Lambda_{min}$ then the two particles do not 'feel' any turbulence; the distance between them is just stretched by the flow. If we observe the particles during short time $t \ll T_{min}$, then the diameter of the passive tracer cloud grows exponentially with t^2:

$$\rho(t) = \rho_0 \exp\{\mathscr{B}t^2\},$$

where ρ_0 is the initial diameter of the passive tracer cloud (at $t = 0$), and $\mathscr{B}$ is a constant, see § 5.2.2.

For the intermediate range of the separations s, $\Lambda_{min} \ll s \ll \Lambda_{max}$, we can expect to see anomalous spreading of the passive tracer cloud. As the cloud spreads, more and more scales Λ of the velocity field become less than the diameter ρ of the cloud; the short scales $\Lambda \ll \rho$ contribute to the normal diffusion, while the large scales $\Lambda \gg \rho$ contribute to the convection of the cloud. We find the existence of two regimes of this anomalous behaviour; they are determined by the length scale ξ (defined by the constants $\mathscr{A}$ and $\mathscr{C}$ in (1.12)): (i) $\rho \ll \xi$ and (ii) $\rho \gg \xi$.

1.3.3. *Sub- or super-exponential divergence of 'close' tracer particles*

When the cloud is sufficiently small $\rho \ll \xi$, the diameter $\rho(t)$ of the cloud grows sub- or super-exponentially:

$$\rho(t) = \rho_0 \exp\left\{\mathscr{B}(t/\tau)^{\beta}\right\} \qquad \text{where} \qquad \beta = 2 + \frac{\gamma - d - 2}{\alpha} = \lambda - \frac{2}{\alpha},$$

where $\rho_0 = \rho(0)$ is the initial diameter (see § 6). Along with the exponent β, we also determine the constant pre-factor $\mathscr{B}$; see (6.5) for the two-dimensional case. Formula (6.3) describes a general situation, when the energy spectrum and/or the dispersion law are not necessarily power-law functions of the wavenumber k.

The sub- or super-exponential divergence of the tracer particles saturates at a certain instant. At this time, the diameter of the cloud reaches the value of order ξ and 'forgets' the initial value ρ_0.

1.3.4. *Power-law spreading of a 'large' passive tracer cloud*

If the initial diameter ρ_0 of the passive tracer cloud is large compared with the scale ξ, the diameter $\rho(t)$ grows according to the following ordinary differential equation

$$\frac{\tau}{\xi^2}\rho\dot{\rho} = a\left(\frac{t}{\tau}\right)\left(\frac{\rho}{\xi}\right)^{\gamma-d} + b\left(\frac{t}{\tau}\right)^{1+(\gamma-d)/\alpha}$$

see § 7. We find the value of the constant b, but the determination of the constant a would require the solution of an integral-differential equation, (5.2), which we derive in this paper.

If $\gamma < d$, then $a < 0$ and $b > 0$; if $\gamma > d$, then $a > 0$ and $b < 0$. For sufficiently long time t ($\tau \ll t \ll T_{max}$) the cloud diameter ρ behaves like a power of t:

$$\frac{\rho(t)}{\xi} \sim \begin{cases} \sqrt{\dfrac{2b}{\lambda}}\left(\dfrac{t}{\tau}\right)^{\lambda/2}, & \lambda = 2 + \dfrac{\gamma - d}{\alpha} \qquad \text{if} \quad \gamma < d, \\ \left(\dfrac{a}{\nu}\right)^{\nu/2}\left(\dfrac{t}{\tau}\right)^{\nu}, & \nu = \dfrac{2}{d + 2 - \gamma} \qquad \text{if} \quad \gamma > d. \end{cases}$$

2. Statistical near-identity transformation

The derivation of equations for the averaged behaviour of the passive tracer is simpler and more transparent if we generalize this case and ignore specific details of the passive tracer problem. In this paper we develop the statistical near-identity transformation for a general linear dynamical system with random coefficients. This system is a generalization, in particular, of the following two physical examples: the passive tracer problem (which we reformulate more precisely in § 2.1.1) and the Schrödinger equation with random potential (§ 2.1.2). The general situation is considered in § 2.1.3, and the statistical near-identity transformation is described in § 2.2. In § 2.3, we find the statistical Green's function $G(\boldsymbol{r}, t)$ that relates the average $\langle \varphi(\boldsymbol{r}, t) \rangle$ to the initial condition $\varphi_0(\boldsymbol{r}) = \varphi(\boldsymbol{r}, 0)$.

2.1. *The equation*

2.1.1. *Passive advection*

When the molecular diffusion is negligible ($\kappa = 0$), the evolution of the passive tracer concentration $\varphi(\boldsymbol{r}, t)$ is described by the equation

$$\frac{\partial \varphi}{\partial t} + \nabla[\boldsymbol{v}(\boldsymbol{r}, t)\varphi] = 0. \tag{2.1}$$

We assume that the velocity field $\boldsymbol{v}(\boldsymbol{r}, t)$ is a superposition of waves with some dispersion law $\Omega_{\boldsymbol{k}}$. If the dispersion law is odd, we can rewrite (1.2) in the form

$$\boldsymbol{v}(\boldsymbol{r}, t) = \int c_{\boldsymbol{k}} \boldsymbol{b}_{\boldsymbol{k}} \exp(\mathrm{i}(\boldsymbol{k} \cdot \boldsymbol{r} - \Omega_{\boldsymbol{k}} t))\, \mathrm{d}\boldsymbol{k}; \tag{2.2}$$

without the real part Re; the velocity field, (2.2), is real provided

$$c_{\boldsymbol{k}} = c^*_{-\boldsymbol{k}}, \quad \boldsymbol{b}_{\boldsymbol{k}} = \boldsymbol{b}^*_{-\boldsymbol{k}}, \quad \Omega_{\boldsymbol{k}} = -\Omega_{-\boldsymbol{k}}.$$

The representation (2.2) takes place, for example, for Rossby waves

$$\Omega_{\boldsymbol{k}} = \Omega(k_x, k_y) = k_x/(1 + k_x^2 + k_y^2).$$

If the dispersion law is even, $\Omega_{\boldsymbol{k}} = \Omega_{-\boldsymbol{k}}$, we have the following representation of the velocity field

$$\boldsymbol{v}(\boldsymbol{r}, t) = \frac{1}{\sqrt{2}} \int c_{\boldsymbol{k}} \boldsymbol{b}_{\boldsymbol{k}} \exp(\mathrm{i}(\boldsymbol{k} \cdot \boldsymbol{r} - \Omega_{\boldsymbol{k}} t))\, \mathrm{d}\boldsymbol{k} + \frac{1}{\sqrt{2}} \int c^*_{-\boldsymbol{k}} \boldsymbol{b}^*_{-\boldsymbol{k}} \exp(\mathrm{i}(\boldsymbol{k} \cdot \boldsymbol{r} + \Omega_{\boldsymbol{k}} t))\, \mathrm{d}\boldsymbol{k} \tag{2.3}$$

(this $1/\sqrt{2}$ normalization is chosen from the energy considerations, see below). The representation (2.3) takes place, for example, for gravity waves, $\Omega_{\boldsymbol{k}} = \sqrt{|\boldsymbol{k}|}$.

In the case of an arbitrary dispersion law, we represent the velocity field in the form

$$\boldsymbol{v}(\boldsymbol{x}, t) = \int A_q \boldsymbol{b}_{\boldsymbol{k}} \exp(\mathrm{i}(\boldsymbol{k} \cdot \boldsymbol{x} - \omega t))\, \mathrm{d}\boldsymbol{k}\, \mathrm{d}\omega \quad q = (\boldsymbol{k}, \omega) \tag{2.4}$$

where the function $A_q = A(\boldsymbol{k}, \omega)$ is 'concentrated' along some surface in the $(\boldsymbol{k}, \omega)$-space, defined by the dispersion relation of the medium. In particular,

$$\text{in the case (2.2)}, \; A_q = c_{\boldsymbol{k}} \delta(\omega - \Omega_{\boldsymbol{k}}),$$

$$\text{in the case (2.3)}, \; A_q = \frac{1}{\sqrt{2}} c_{\boldsymbol{k}} \delta(\omega - \Omega_{\boldsymbol{k}}) + \frac{1}{\sqrt{2}} c^*_{-\boldsymbol{k}} \delta(\omega + \Omega_{\boldsymbol{k}}).$$

The formula (2.4) works as well when the dispersion law is neither odd nor even. The representation (2.4) can also describe situations when there is some 'frequency spreading'. The function A_q is large in some 'vicinity' of the surface, defined by the dispersion relation (not just on the surface itself); it approaches zero when we move away from this surface.

For the velocity field to be real, we require that

$$A_q = A^*_{-q}, \quad \boldsymbol{b}_k = \boldsymbol{b}^*_{-k}.$$

We assume that the velocity $\boldsymbol{v}(\boldsymbol{r}, t)$ is a Gaussian random field with zero mean, statistically homogeneous in space and in time. It means that the wave amplitudes A_q are Gaussian random variables with zero mean $\langle A_q \rangle = 0$ and variance

$$\langle A_q A^*_{q_1} \rangle = E_q \delta(\boldsymbol{k} - \boldsymbol{k}_1)\delta(\omega - \omega_1) \quad (E_q = E_{-q}).$$

The spectrum $E_q = E(\boldsymbol{k}, \omega)$ (together with the polarization vector) completely defines the statistical ensemble of the velocity fields.

The problem is to find the statistical properties of the tracer field $\varphi(\boldsymbol{r}, t)$, in particular, the mean displacement (1.4), the variance tensor (1.5), and the variance tensor of relative displacement (1.7).

We normalize the polarization vector to a unit, $|\boldsymbol{b}_k| = 1$; then $E_q = E(\boldsymbol{k}, \omega)$ is the energy spectrum.

When the waves are linear, i.e. their frequency is defined by the linear dispersion relation (without frequency spreading), we consider the spectrum $\varepsilon_{\boldsymbol{k}}$

$$\langle c_1 c_2^* \rangle = \varepsilon_1 \delta(\boldsymbol{k}_1 - \boldsymbol{k}_2).$$

For odd and even dispersion law we have, respectively,

$$\text{in the case (2.2)},\ E_q = E(\boldsymbol{k}, \omega) = \varepsilon_{\boldsymbol{k}} \delta(\omega - \Omega_{\boldsymbol{k}}), \quad \varepsilon_{\boldsymbol{k}} = \varepsilon_{-\boldsymbol{k}}, \tag{2.5}$$

$$\text{in the case (2.3)},\ E_q = E(\boldsymbol{k}, \omega) = \tfrac{1}{2}\varepsilon_{\boldsymbol{k}} \delta(\omega - \Omega_{\boldsymbol{k}}) + \tfrac{1}{2}\varepsilon_{\boldsymbol{k}} \delta(\omega + \Omega_{\boldsymbol{k}}). \tag{2.6}$$

In both cases, the total energy is

$$E \equiv \int E(\boldsymbol{k}, \omega)\, \mathrm{d}\boldsymbol{k}\, \mathrm{d}\omega = \int \varepsilon_{\boldsymbol{k}}\, \mathrm{d}\boldsymbol{k}.$$

We write the passive tracer equation (2.1) in the Fourier representation

$$\dot{f}_1 + \int \mathrm{i}\boldsymbol{k}_1 \cdot \boldsymbol{b}_3 \exp(-\mathrm{i}\omega_3 t) f_2 A_3 \delta(-\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) d_{23} = 0, \tag{2.7}$$

where the function $f_{\boldsymbol{k}}(t)$ is the Fourier transform of $\varphi(\boldsymbol{r}, t)$ with respect to the spatial variables

$$\varphi(\boldsymbol{r}, t) = \int f_{\boldsymbol{k}}(t) \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r})\, \mathrm{d}\boldsymbol{k} \quad (f_{\boldsymbol{k}} = f^*_{-\boldsymbol{k}}).$$

Here and below we use notations of the following type: for any function h, depending on $(\boldsymbol{k}, \omega)$ or just on $\boldsymbol{k}$, we write h_j instead of $h(\boldsymbol{k}_j, \omega_j)$ $(j = 1, 2, \ldots)$, e.g. $f_1 = f_{\boldsymbol{k}_1}$, $f_2 = f_{\boldsymbol{k}_2}$, $\boldsymbol{b}_3 = \boldsymbol{b}_{\boldsymbol{k}_3}$ $A_3 = A(\boldsymbol{k}_3, \omega_3)$.

A similar equation in the Fourier representation also appears in the following problem.

2.1.2. *Schrödinger equation with random potential*

This problem is described by the equation

$$\mathrm{i}\frac{\partial \psi}{\partial t} + \Delta\psi - U(\boldsymbol{r}, t)\psi = 0, \tag{2.8}$$

where the potential $U(\boldsymbol{r}, t)$ is a random function.

We represent the potential in the form

$$U(\boldsymbol{r},t) = \int A_q \exp(\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t))\,\mathrm{d}\boldsymbol{k}\omega \quad q = (\boldsymbol{k},\omega). \tag{2.9}$$

In particular, the potential can be stationary (time-independent), which is the case in many studies (see e.g. Sheng 1995 and references cited therein). In that case, A_q does not depend on the frequency ω, and we can integrate over ω. However, we would like to keep consideration of a more general case, when the potential can depend on time. That is, we assume that $U(\boldsymbol{r},t)$ is a superposition of waves (2.9), which means that the function A_q is 'concentrated' along some surface (defined by the dispersion relation of the waves). As in the previous problem of passive advection, the representation (2.9) can describe situations when there is some 'frequency spreading' (the function A_q is large in some 'vicinity' of the surface, defined by the dispersion relation, not just on the surface itself). If the potential is stationary, the dispersion law of the waves just equals zero identically.

For the potential to be real, we require that $A_q = A^*_{-q}$.

We assume that the potential is a random Gaussian field with zero mean, statistically homogeneous in space and in time. It means that the wave amplitudes A_q are Gaussian random variables with zero mean $\langle A_q\rangle = 0$ and variance

$$\langle A_q A^*_{q_1}\rangle = E_q \delta(\boldsymbol{k}-\boldsymbol{k}_1)\delta(\omega-\omega_1) \quad (E_q = E_{-q}).$$

Thus, the spectrum $E_q = E(\boldsymbol{k},\omega)$ completely defines the statistical ensemble of the potentials.

If there were no potential ($U = 0$), the solution ψ would be merely a superposition of linear waves

$$\psi(\boldsymbol{x},t) = \int f_{\boldsymbol{k}} \exp(\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x} - k^2 t))\,\mathrm{d}\boldsymbol{k} \tag{2.10}$$

with wave amplitudes $f_{\boldsymbol{k}}$ being time independent. When there is a 'small' potential, the wave amplitudes become 'slow' functions of time, and we consider the formula (2.10) as the transformation to the new variable $f_{\boldsymbol{k}}(t)$. It satisfies the following equation

$$\dot{f}_1 + \int \mathrm{i}\exp(\mathrm{i}(k_1^2 - k_2^2 - \omega_3)t) f_2 A_3 \delta(-\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3)\,\mathrm{d}\boldsymbol{k}_2\,\mathrm{d}\boldsymbol{k}_3\,\mathrm{d}\omega_3 = 0. \tag{2.11}$$

2.1.3. *The general equation*

The two problems considered above (passive advection and the Schrödinger equation with random potential), as well as various other problems, can be formulated in terms of the following equation

$$\dot{f}_1 + \int \hat{W}_{-123} f_2 A_3\,\mathrm{d}_{23} = 0, \quad A_q = A^*_{-q}, \tag{2.12}$$

with the kernel

$$\begin{aligned}\hat{W}_{-123} &= W_{-123}\delta(-\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3)\\ &= U_{-123}\exp(-\mathrm{i}(-\sigma_1 + \sigma_2 + \omega_3)t)\,\delta(-\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3).\end{aligned} \tag{2.13}$$

Here '1' denotes $\boldsymbol{k}_1$, '2' denotes $\boldsymbol{k}_2$, and '3' denotes $q_3 = (\boldsymbol{k}_3,\omega_3)$; the hat symbol denotes the multiplication by the corresponding delta function.

In the case of the passive tracer

$$\sigma_{\boldsymbol{k}} \equiv 0, \quad U_{-123} = \mathrm{i}\boldsymbol{k}_1\cdot\boldsymbol{b}_3. \tag{2.14}$$

In the case of the Schrödinger equation

$$\sigma_{\boldsymbol{k}} = k^2, \quad U_{-123} = \mathrm{i}. \tag{2.15}$$

The function A_q is the Fourier representation of a random Gaussian field with zero mean $\langle A_q \rangle = 0$, statistically homogeneous in space and in time. Therefore, the spectrum $E_q = E(\boldsymbol{k}, \omega)$, defined by the formula

$$\langle A_q A^*_{q_1} \rangle = E_q \delta(\boldsymbol{k} - \boldsymbol{k}_1)\delta(\omega - \omega_1),$$

completely determines the statistical ensemble of the wave amplitudes A_q.

The problem is to find statistical properties of the field $f_{\boldsymbol{k}}(t)$.

We assume that the function A_q is 'concentrated' along some surface (defined by the dispersion relation) in the $(\boldsymbol{k}, \omega)$-space. The function A_q, as well as the spectrum E_q, is large in some 'vicinity' of this surface (they are delta-like functions with a ridge along this surface). For example, without the frequency spreading

$$A_q = \overline{A}_q \delta(L_q), \quad E_q = \overline{E}_q \delta(L_q),$$

where $L_q = L(\boldsymbol{k}, \omega)$ is the dispersion relation, e.g. $L_q = \omega^2 - \Omega^2_{\boldsymbol{k}}$.

2.2. *The transformation*

Our goal is to find statistical properties of the field $f_{\boldsymbol{k}}(t)$, in particular, the ensemble average $\langle f_{\boldsymbol{k}}(t) \rangle$. For this we introduce statistical near-identity transformation.

The form of this transformation is motivated by the usual near-identity transformation known in the theory of dynamical systems (e.g. Bogoliubov & Mitropolsky 1961; Sanders & Verhulst 1985) and by the Wiener–Hermite expansion (e.g. Wiener 1958; Eftimiu 1990).

2.2.1. *Recalling the usual near-identity transformation*

Let us first of all recall the usual near-identity transformation from the theory of dynamical systems. Suppose we want to solve the following ordinary differential equation

$$\ddot{u} + \omega^2 u = \epsilon h(u, \dot{u}, t)$$

with a small parameter ϵ ($\omega \neq 0$ is a given constant, h is given function, $u = u(t)$ is the unknown function). When $\epsilon = 0$, the solution is simple: $u = a\cos(\omega t + \phi)$, where a and ϕ are constants. When the parameter ϵ is positive, but 'small', the quantities a and ϕ become 'slow' functions of time: $a = a(t), \phi = \phi(t)$. We can change the variable $(u, \dot{u})$ to variables (a, ϕ); introducing the vector $\boldsymbol{x} = (a, \phi)$, we write the original equation in the form

$$\dot{\boldsymbol{x}} = \epsilon \boldsymbol{W}(\boldsymbol{x}, t). \tag{2.16}$$

This is the standard form, which is equivalent to the original form (no approximation has been made). The form of the function $\boldsymbol{W}$ is determined by the function h in the original equation. The standard form can arise from other equations, and the vector $\boldsymbol{x}$ can have any dimension. It is supposed that the function $\boldsymbol{W}$ is oscillatory in time, that is, it can be expanded in the form

$$\boldsymbol{W}(\boldsymbol{x}, t) = \sum_{v} \boldsymbol{W}_v(\boldsymbol{x})\,\mathrm{e}^{\mathrm{i}vt}.$$

To solve this equation by perturbation methods, we look for a near-identity transformation

$$\boldsymbol{x} = \boldsymbol{y} + \epsilon \boldsymbol{X}(\boldsymbol{y}, t),$$

where $\boldsymbol{X}$ is an undetermined function, bounded in $\boldsymbol{y}$ and t. (When $\epsilon = 0$, this transformation is an identity). We choose the function $\boldsymbol{X}$ so that the equation for the new variable $\boldsymbol{y}$ is the 'simplest' possible. We obtain the equation

$$\dot{\boldsymbol{y}} = \epsilon \boldsymbol{P}(\boldsymbol{y}) + \epsilon^2 \boldsymbol{Q}(\boldsymbol{y}, t).$$

So, we have 'pushed' the time dependence to the order ϵ^2. Neglecting the ϵ^2-term, we obtain the first approximation (The method of near-identity transformation is also called the method of averaging because the function $\boldsymbol{P}$ is the time-average of the function $\boldsymbol{W}$: $\boldsymbol{P}(\boldsymbol{y}) = \boldsymbol{W}_0(\boldsymbol{y})$.)

We can also obtain the second approximation: For this, we consider a near-identity transformation of the form

$$\boldsymbol{x} = \boldsymbol{y} + \epsilon \boldsymbol{X}(\boldsymbol{y}, t) + \epsilon^2 \boldsymbol{Y}(\boldsymbol{y}, t),$$

where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are undetermined functions. We choose these functions so that the equation for the new variable $\boldsymbol{y}$ is the 'simplest' possible. We obtain the equation

$$\dot{\boldsymbol{y}} = \epsilon \boldsymbol{P}(\boldsymbol{y}) + \epsilon^2 \boldsymbol{Q}(\boldsymbol{y}) + \epsilon^3 \boldsymbol{R}(\boldsymbol{y}, t). \tag{2.17}$$

So, we have 'pushed' the time dependence to the order ϵ^3. Neglecting the ϵ^3-term, we obtain the second approximation.

2.2.2. *Statistical transformation, motivated by the Wiener–Hermite expansion*

The method of near-identity transformation works well if there are no resonances (ideally when there is only one basic frequency ω; all other frequencies are integer multiples of ω). However, in the passive tracer situation we have a wide range of frequencies, and there are many resonances. We will use the randomness of the wave amplitudes in order to develop a near-identity transformation in this situation.

The Fourier transformation of the passive tracer equation (2.1) is similar to the reduction to the standard form; equation (2.12) is analogous to the ordinary differential equation (2.16) in the standard form. To approach equation (2.12) by perturbation methods, we construct the following near-identity transformation

$$f_1 = g_1 + \int \hat{X}_{-123} g_2 A_3 \mathrm{d}_{23} + \frac{1}{2} \int \hat{Y}_{-1234} g_2 [A_3 A_4 - E_3 \delta(q_3 + q_4)] \,\mathrm{d}_{234}. \tag{2.18}$$

Here, $\hat{X}$ and $\hat{Y}$ are some undetermined kernels; the hat symbol denotes the multiplication by the corresponding δ-function:

$$\hat{X}_{-123} = X(-\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3; t)\delta(-\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3),$$

$$\hat{Y}_{-1234} = Y(-\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4; t)\delta(-\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 + \boldsymbol{k}_4).$$

There is no explicit ϵ in the transformation (2.18), but we assume that the wave amplitudes are small, of the order ϵ, so that the first integral in (2.18) represents an ϵ-term, whereas the second integral represents an ϵ^2-term.

The transformation (2.18) contains the first two functions of the infinite sequence of the Wiener–Hermite polynomials H with respect to the Gaussian random field A_q:

$$H_1^{(1)} = A_1, \quad H_{12}^{(2)} = A_1 A_2 - E_1 \delta(q_1 + q_2),$$
$$H_{123}^{(3)} = A_1 A_2 A_3 - A_1 E_2 \delta(q_2 + q_3) - A_2 E_3 \delta(q_3 + q_1) - A_3 E_1 \delta(q_1 + q_2), \ldots .$$

The Wiener–Hermite polynomials are orthogonal with respect to the statistical averaging: $\langle H^{(i)} H^{(j)} \rangle = 0$ for two polynomials of different orders $i \neq j$. Besides, they form a complete set. A functional of the field A_q can be expanded in a series over the Wiener–Hermite polynomials. In particular, the solution of equation (2.12) can be represented by the Wiener–Hermite expansion

$$f_1 = M_1 f_1^0 + \int \hat{M}_{-123} f_2^0 H_3^{(1)} \, d_{23} + \tfrac{1}{2} \int \hat{M}_{-1234} f_2^0 H_{34}^{(2)} \, d_{234} + \\ + \frac{1}{3!} \int \hat{M}_{-12345} f_2^0 H_{345}^{(3)} \, d_{2345} + \cdots ,$$

with some kernels M_1, $\hat{M}_{-123}$, $\hat{M}_{-1234}$, $\hat{M}_{-12345}, \ldots$. The function $f_{\boldsymbol{k}}^0$ denotes the initial condition, $f_{\boldsymbol{k}}(0) = f_{\boldsymbol{k}}^0$. The transformation (2.18) could be considered as a truncation of the infinite Wiener–Hermite series. Herewith, we also change the term $M_{\boldsymbol{k}} f_{\boldsymbol{k}}^0$ to a new variable $g_{\boldsymbol{k}}(t)$.

The Wiener–Hermite polynomials motivated our form of the near-identity transformation (2.18). With the aid of this transformation we intend ‘to push’ the randomness to the ϵ^3-order, similar to ‘pushing’ time-dependence to the ϵ^3-order in the second approximation, (2.17), of the usual near-identity transformation method.

2.2.3. *Change of variables*

We consider equation (2.18) as a change of variables from the old variable $f_{\boldsymbol{k}}(t)$ to the new variable $g_{\boldsymbol{k}}(t)$.

Without loss of generality, we can assume that the kernel $\hat{Y}_{-1234}$ is symmetric with respect to the transposition of the last two indices: $\hat{Y}_{-1234} = \hat{Y}_{-1243}$ (this is the reason for introducing the normalization factor $\frac{1}{2}$ in front of the integral with $\hat{Y}_{-1234}$).

Substituting (2.18) into equation (2.12) for $f_{\boldsymbol{k}}(t)$, we arrive at the following equation for $g_{\boldsymbol{k}}(t)$

$$\dot{g}_1 + \int \dot{\hat{X}}_{-123} g_2 A_3 \, d_{23} + \tfrac{1}{2} \int \dot{\hat{Y}}_{-1234} g_2 [A_3 A_4 - E_3 \delta(q_3 + q_4)] \, d_{234} \\ + \int \hat{X}_{-123} \dot{g}_2 A_3 \, d_{23} + \tfrac{1}{2} \int \hat{Y}_{-1234} \dot{g}_2 [A_3 A_4 - E_3 \delta(q_3 + q_4)] \, d_{234} \\ + \int \hat{W}_{-123} \\ \times \left\{ g_2 + \int \hat{X}_{-256} g_5 A_6 \, d_{56} + \tfrac{1}{2} \int \hat{Y}_{-2567} g_5 [A_6 A_7 - E_6 \delta(q_6 + q_7)] \, d_{567} \right\} A_3 \, d_{23} = 0. \tag{2.19}$$

2.2.4. *The kernel choice*

In accordance with the general idea of the near-identity transformation, we choose the kernel X_{-123} so that the terms linear in the field A_q disappear from equation (2.19). This is possible if

$$\dot{X}_{-123} = -W_{-123} \quad \text{provided} \; -\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 = 0. \tag{2.20}$$

Then $\dot{g}_{\boldsymbol{k}} = O(\epsilon^2)$, and the integrals with time derivative $\dot{g}$ have the order ϵ^3. So, we obtain the following equation

$$\dot{g}_1 + \tfrac{1}{2} \int \dot{\hat{Y}}_{-1234} g_2 [A_3 A_4 - E_3 \delta(q_3 + q_4)] \, d_{234} \\ + \int \hat{W}_{-123} \hat{X}_{-256} g_5 A_6 A_3 \, d_{2356} + \epsilon^3 \{\ldots\} = 0. \tag{2.21}$$

Now we want to choose the Y-kernel to 'kill' the second-order Wiener–Hermite polynomial $H^{(2)}_{34} = A_3A_4 - E_3\delta(q_3 + q_4)$. For this, we make the first and the second integral in (2.21) look similar. First, we rename the integration variables in the second integral:

$$5 \to 2,\ 3 \to 3,\ 6 \to 4,\ 2 \to 5 \Rightarrow \int \hat{W}_{-153}\hat{X}_{-524}g_2A_4A_3\,\mathrm{d}_{2345};$$

then we symmetrize this integral with respect to transposition $3 \leftrightarrow 4$

$$\tfrac{1}{2}\int \left\{\hat{W}_{-153}\hat{X}_{-524} + \hat{W}_{-154}\hat{X}_{-523}\right\} g_2A_3A_4\,\mathrm{d}_{2345}$$

and write A_3A_4 in the latter integral as $A_3A_4 - E_3\delta(q_3+q_4) + E_3\delta(q_3+q_4)$. As a result, we rewrite equation (2.21) in the form

$$\begin{aligned} \dot{g}_1 &+ \tfrac{1}{2}\int \dot{\hat{Y}}_{-1234}g_2[A_3A_4 - E_3\delta(q_3 + q_4)]\,\mathrm{d}_{234} \\ &+ \tfrac{1}{2}\int \left\{\hat{W}_{-153}\hat{X}_{-524} + \hat{W}_{-154}\hat{X}_{-523}\right\} g_2[A_3A_4 - E_3\delta(q_3 + q_4)] \\ &+ \tfrac{1}{2}\int W_{-153}X_{-51-3}g_2E_3\delta(-\boldsymbol{k}_1 + \boldsymbol{k}_5 + \boldsymbol{k}_3)\,\mathrm{d}_{35} + \epsilon^3\{\ldots\} = 0. \end{aligned} \tag{2.22}$$

We choose the kernel Y_{-1234} so that

$$\dot{Y}_{-1234} = -\, W_{-153}X_{-524}|_{\boldsymbol{k}_5=\boldsymbol{k}_1-\boldsymbol{k}_3=\boldsymbol{k}_2+\boldsymbol{k}_4} - W_{-154}X_{-523}|_{\boldsymbol{k}_5=\boldsymbol{k}_1-\boldsymbol{k}_4=\boldsymbol{k}_2+\boldsymbol{k}_3}\,; \tag{2.23}$$

then the second-order Wiener–Hermite polynomial $H^{(2)}_{34} = A_3A_4 - E_3\delta(q_3 + q_4)$ disappears from equation (2.22). Thus, we find the following equation

$$\dot{g}_1 + g_1\int W_{-123}X_{-21-3}E_3\delta(-\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3)\,\mathrm{d}_{23} + \epsilon^3\{\ldots\} = 0. \tag{2.24}$$

So, we have pushed the randomness to the ϵ^3-term. Neglecting it, we obtain the approximation that we use to find the evolution of the mean $\langle f_{\boldsymbol{k}}(t)\rangle$.

We would like to emphasize that we do not use the series expansions. To derive equation (2.24), we do not work with series, and we ignore convergence (about the behaviour of nth term in the series as $n \to \infty$). Instead, we change the variables (2.18), which is a finite sum of three terms. We think that this approach via the statistical near-identity transformation would enable us to justify equation (2.24) rigorously.

2.2.5. *Simplified transformation*

The kernel Y_{-1234} does not affect equation (2.24). It is usual for a near-identity transformation that the highest-order term in such a transformation does not affect the corresponding approximation (see e.g. Bogoliubov & Mitropolsky 1961; Sanders & Verhulst 1985). We could obtain equation (2.24) in the following, simpler, manner. Instead of (2.18), we could have made the transformation

$$f_1 = g_1 + \int \hat{X}_{-123}g_2A_3\,\mathrm{d}_{23}, \tag{2.25}$$

without the Y-term, and obtained the equation

$$\begin{aligned} \dot{g}_1 &+ \int \dot{\hat{X}}_{-123}g_2A_3\,\mathrm{d}_{23} + \int \hat{X}_{-123}\dot{g}_2A_3\,\mathrm{d}_{23} \\ &+ \int \hat{W}_{-123}\left\{g_2 + \int \hat{X}_{-256}g_5A_6\,\mathrm{d}_{56}\right\} A_3\,\mathrm{d}_{23} = 0. \end{aligned} \tag{2.26}$$

As previously, we choose the kernel X_{-123} according to (2.20), so that the linear terms in the field A_q disappear from equation (2.26). Then $\dot{g}_{\boldsymbol{k}} = O(\epsilon^2)$, and the integral with time derivative $\dot{g}$ are of order ϵ^3. Instead of equation (2.21), we have

$$\dot{g}_1 + \int \hat{W}_{-123}\hat{X}_{-256}g_5A_6A_3\,\mathrm{d}_{2356} + \epsilon^3\{\ldots\} = 0. \tag{2.27}$$

Now we average this equation with respect to the random variables A_q, assuming that g-variables are statistically independent of A-variables. As a result, we obtain equation (2.24).

This simplified (but less substantiated) derivation of the averaged equation (2.24) is actually similar to the averaging procedure (with respect to time) in the theory of dynamical systems. The same equation can be obtained in two ways, either we apply the near-identity transformation with all the terms, or we apply the near-identity transformation without the highest-order term and subsequently perform averaging.

We use only the simplified derivation, with the averaging, to obtain equations for the two-point correlation function (§ 4), and we do not substantiate them by the statistical near-identity transformation with all the terms.

2.2.6. *Time dependence*

Since equations (2.20) and (2.23) define only the time derivatives of the kernels X and Y in the near-identity transformation (2.18), we choose the constants of time integration so that the initial condition for the new variable $g_{\boldsymbol{k}}(t)$ would be the same as for the old variable $f_{\boldsymbol{k}}(t)$. Then, the near-identity transformation (2.18) is the identity initially.

Thus, according to (2.20),

$$X_{-123} = -U_{-123}\frac{\exp(-\mathrm{i}(-\sigma_1+\sigma_2+\omega_3)t)-1}{-\mathrm{i}(-\sigma_1+\sigma_2+\omega_3)}; \tag{2.28}$$

and according to (2.23),

$$\begin{aligned}\dot{Y}_{-1234} &= U_{-153}U_{-524}\exp(-\mathrm{i}(-\sigma_1+\sigma_5+\omega_3)t)\\ &\quad\times\left.\frac{\exp(-\mathrm{i}(-\sigma_5+\sigma_2+\omega_4)t)-1}{-\mathrm{i}(-\sigma_5+\sigma_2+\omega_4)}\right|_{\boldsymbol{k}_5=\boldsymbol{k}_1-\boldsymbol{k}_3=\boldsymbol{k}_2+\boldsymbol{k}_4}\\ &\quad+\{\text{same with exchange } 3\leftrightarrow 4\},\end{aligned}$$

so,

$$\begin{aligned}Y_{-1234} = &-U_{-153}U_{-524}\left\{\frac{\exp(-\mathrm{i}(-\sigma_1+\sigma_2+\omega_3+\omega_4)t)-1}{(-\sigma_1+\sigma_2+\omega_3+\omega_4)(-\sigma_5+\sigma_2+\omega_4)}\right.\\ &\left.-\frac{\exp(-\mathrm{i}(-\sigma_1+\sigma_5+\omega_3)t)-1}{(-\sigma_1+\sigma_5+\omega_3)(-\sigma_5+\sigma_2+\omega_4)}\right\}_{\boldsymbol{k}_5=\boldsymbol{k}_1-\boldsymbol{k}_3=\boldsymbol{k}_2+\boldsymbol{k}_4}\\ &-U_{-154}U_{-523}\left\{\frac{\exp(-\mathrm{i}(-\sigma_1+\sigma_2+\omega_3+\omega_4)t)-1}{(-\sigma_1+\sigma_2+\omega_3+\omega_4)(-\sigma_5+\sigma_2+\omega_3)}\right.\\ &\left.-\frac{\exp(-\mathrm{i}(-\sigma_1+\sigma_5+\omega_4)t)-1}{(-\sigma_1+\sigma_5+\omega_4)(-\sigma_5+\sigma_2+\omega_3)}\right\}_{\boldsymbol{k}_5=\boldsymbol{k}_1-\boldsymbol{k}_4=\boldsymbol{k}_2+\boldsymbol{k}_3}.\end{aligned} \tag{2.29}$$

Now, equation (2.24) can be rewritten in the form

$$\dot{g}_1 = g_1 \int U_{-1,2,3} U_{-2,1,-3} \frac{1 - \exp(-\mathrm{i}(-\sigma_1 + \sigma_2 + \omega_3)t)}{\mathrm{i}(-\sigma_1 + \sigma_2 + \omega_3)} E_3 \delta(-\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3)\, \mathrm{d}_{23} \quad (2.30)$$

(ϵ^3-terms are neglected).

2.3. *Green's function*

Since the initial conditions for the old variable $f_{\boldsymbol{k}}(t)$ and for the new variable $g_{\boldsymbol{k}}(t)$ coincide:

$$g_{\boldsymbol{k}}(0) = f_{\boldsymbol{k}}(0) = f^0_{\boldsymbol{k}},$$

equation (2.30) defines the solution

$$g_1(t) = f^0_1 \exp\left\{ \int U_{-1,2,3} U_{-2,1,-3} \frac{1 + \mathrm{i}(\sigma_1 - \sigma_2 - \omega_3)t - \exp(\mathrm{i}(\sigma_1 - \sigma_2 - \omega_3)t)}{(\sigma_1 - \sigma_2 - \omega_3)^2} \times E_3 \delta(\boldsymbol{k}_1 - \boldsymbol{k}_2 - \boldsymbol{k}_3)\, \mathrm{d}_{23} \right\}. \quad (2.31)$$

The solution $f_{\boldsymbol{k}}(t)$ of the original equation (2.12) is expressed through the function (2.31) by the near-identity transformation (2.18). Since (2.31) is not a random function, and the average of each Weiner–Hermite polynomial is zero, the ensemble average of the original solution $f_{\boldsymbol{k}}(t)$ is equal to the function (2.31):

$$\langle f_{\boldsymbol{k}}(t) \rangle = g_{\boldsymbol{k}}(t) \Rightarrow \langle \varphi(\boldsymbol{r}, t) \rangle = \int g_{\boldsymbol{k}}(t) \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r})\, \mathrm{d}\boldsymbol{k}, \quad (2.32)$$

where $\varphi(\boldsymbol{r}, t)$ is the solution of our equation in the real space representation. Equations (2.31) and (2.32) mean that

$$\langle \varphi(\boldsymbol{r}, t) \rangle = \int G(\boldsymbol{r} - \boldsymbol{r}_1, t) \varphi_0(\boldsymbol{r}_1)\, \mathrm{d}\boldsymbol{r}_1, \quad (2.33)$$

where $\varphi(\boldsymbol{r}, 0) = \varphi_0(\boldsymbol{r})$ is the initial condition, and the Green function $G(\boldsymbol{r}, t)$ is the Fourier transform of our solution (2.31) with the initial condition

$$f^0_{\boldsymbol{k}} = 1/(2\pi)^d \Leftrightarrow \varphi(\boldsymbol{r}, 0) = \delta(\boldsymbol{r}),$$

(d is the dimension of the medium). Thus,

$$G(\boldsymbol{r}, t) = \int g_{\boldsymbol{k}}(t) \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r})\, \mathrm{d}\boldsymbol{k} \quad \text{where} \quad g_1(t) = \frac{1}{(2\pi)^d} \times \exp\left\{ \int U_{-1,2,3} U_{-2,1,-3} \frac{1 + \mathrm{i}(\sigma_1 - \sigma_2 - \omega_3)t - \exp(\mathrm{i}(\sigma_1 - \sigma_2 - \omega_3)t)}{(\sigma_1 - \sigma_2 - \omega_3)^2}. \times E_3 \delta(\boldsymbol{k}_1 - \boldsymbol{k}_2 - \boldsymbol{k}_3)\, \mathrm{d}_{23} \right\} \quad (2.34)$$

(recall that $E_3 = E(\boldsymbol{k}_3, \omega_3)$, and so $\mathrm{d}_{23} = \mathrm{d}\boldsymbol{k}_2\, \mathrm{d}\boldsymbol{k}_3\, \mathrm{d}\omega_3$).

3. The evolution of the mean concentration

3.1. *Green's function for the passive tracer case*

Now we specify the general Green's function (2.34) for the passive tracer case (2.14)

$$g_1(t) = \frac{1}{(2\pi)^d} \exp\left\{ \int \mathrm{i}(\boldsymbol{k}_1 \cdot \boldsymbol{b}_3)\mathrm{i}(\{\boldsymbol{k}_1 - \boldsymbol{k}_3\} \cdot \boldsymbol{b}_{-3}) \frac{1 - \mathrm{i}\omega_3 t - \exp(-\mathrm{i}\omega_3 t)}{\omega_3^2} E_3 \, \mathrm{d}_3 \right\}.$$

This expression has the form

$$g_{\boldsymbol{k}}(t) = \frac{1}{(2\pi)^d} \exp(\tfrac{1}{2}\mathrm{i}\boldsymbol{k}'\boldsymbol{D}(t)\mathrm{i}\boldsymbol{k} - \mathrm{i}\boldsymbol{k}'\boldsymbol{R}(t)) \tag{3.1}$$

where $\boldsymbol{D}(t)$ is a symmetric matrix, and $\boldsymbol{R}(t)$ is a vector, defined by the following integrals

$$\boldsymbol{D} = 2\int \boldsymbol{B}_{\boldsymbol{k}} \frac{1-\cos\omega t}{\omega^2} E(\boldsymbol{k},\omega)\,\mathrm{d}\boldsymbol{k}\,\mathrm{d}\omega \qquad \text{where} \qquad \boldsymbol{B}_{\boldsymbol{k}} = \tfrac{1}{2}(\boldsymbol{b}_{\boldsymbol{k}}\boldsymbol{b}'_{-\boldsymbol{k}} + \boldsymbol{b}_{-\boldsymbol{k}}\boldsymbol{b}'_{\boldsymbol{k}}), \tag{3.2}$$

$$\boldsymbol{R} = \int \boldsymbol{B}_{\boldsymbol{k}}\boldsymbol{k} \frac{\omega t - \sin\omega t}{\omega^2} E(\boldsymbol{k},\omega)\,\mathrm{d}\boldsymbol{k}\,\mathrm{d}\omega + \frac{\mathrm{i}}{2}\int (\boldsymbol{b}_{\boldsymbol{k}}\boldsymbol{b}'_{-\boldsymbol{k}} - \boldsymbol{b}_{-\boldsymbol{k}}\boldsymbol{b}'_{\boldsymbol{k}})\boldsymbol{k} \frac{1-\cos\omega t}{\omega^2} E(\boldsymbol{k},\omega)\,\mathrm{d}\boldsymbol{k}\,\mathrm{d}\omega. \tag{3.3}$$

The prime denotes the transposition, so that $\boldsymbol{b}'_{\boldsymbol{k}}$ is a row, while $\boldsymbol{b}_{\boldsymbol{k}}$ is a column. The matrix $\boldsymbol{B}_{\boldsymbol{k}}$ (defined by the polarization vector $\boldsymbol{b}_{\boldsymbol{k}}$) can be called the polarization matrix. To obtain integrals (3.2)–(3.2), we used symmetrization $(\boldsymbol{k}_3, \omega_3) \leftrightarrow -(\boldsymbol{k}_3, \omega_3)$. We have obvious properties $\boldsymbol{b}_{\boldsymbol{k}} = \boldsymbol{b}^*_{-\boldsymbol{k}}$ and $\boldsymbol{B}_{\boldsymbol{k}} = \boldsymbol{B}^*_{-\boldsymbol{k}}$. If $\boldsymbol{b}_{\boldsymbol{k}} = \boldsymbol{b}_{-\boldsymbol{k}}$ or $\boldsymbol{b}_{\boldsymbol{k}} = -\boldsymbol{b}_{-\boldsymbol{k}}$, then the second integral in (3.2) vanishes.

If the waves are linear, their energy spectrum is proportional to a delta function of some surface in the $(\boldsymbol{k},\omega)$-space, and we can integrate in (3.2), (3.3) over $\mathrm{d}\omega$. In particular, in situations of both odd and even dispersion law (2.5)–(2.6) we find the same formulae

$$\boldsymbol{D} = 2\int \boldsymbol{B}_{\boldsymbol{k}} \frac{1-\cos\Omega_{\boldsymbol{k}t}}{\Omega_{\boldsymbol{k}^2}} \varepsilon_{\boldsymbol{k}}\,\mathrm{d}\boldsymbol{k}, \tag{3.4}$$

$$\boldsymbol{R} = \int \boldsymbol{B}_{\boldsymbol{k}}\boldsymbol{k} \frac{\Omega_{\boldsymbol{k}}t - \sin\Omega_{\boldsymbol{k}}t}{\Omega_{\boldsymbol{k}^2}} \varepsilon_{\boldsymbol{k}}\,\mathrm{d}\boldsymbol{k} + \frac{\mathrm{i}}{2}\int (\boldsymbol{b}_{\boldsymbol{k}}\boldsymbol{b}'_{-\boldsymbol{k}} - \boldsymbol{b}_{-\boldsymbol{k}}\boldsymbol{b}'_{\boldsymbol{k}})\boldsymbol{k} \frac{1-\cos\Omega_{\boldsymbol{k}t}}{\Omega_{\boldsymbol{k}^2}} \varepsilon_{\boldsymbol{k}}\,\mathrm{d}\boldsymbol{k}. \tag{3.5}$$

3.1.1. *Averaged equation*

Specifying (2.30) for the passive tracer case (2.14), we find

$$\dot{g}_1 = g_1 \int \mathrm{i}\boldsymbol{k}'_1\boldsymbol{b}_3\boldsymbol{b}'_{-3}\mathrm{i}(\boldsymbol{k}_1 - \boldsymbol{k}_3) \frac{1-\exp(-i\omega_3 t)}{i\omega_3} E_3\,\mathrm{d}_3.$$

This equation has the form

$$\dot{g}_1 = \mathrm{i}\boldsymbol{k}'_1\,\dot{\boldsymbol{D}}\,\mathrm{i}\boldsymbol{k}_1\,g_1 - \mathrm{i}\boldsymbol{k}'_1\dot{\boldsymbol{R}}\,g_1,$$

which in the $\boldsymbol{r}$-space is the convection–diffusion equation

$$\frac{\partial\Phi}{\partial t} + \dot{\boldsymbol{R}} \cdot \frac{\partial\Phi}{\partial\boldsymbol{r}} = \frac{\partial}{\partial\boldsymbol{r}}\left(\dot{\boldsymbol{D}}\frac{\partial\Phi}{\partial\boldsymbol{r}'}\right). \tag{3.6}$$

Here, $\Phi(\boldsymbol{r},t)$ is the Fourier transform of the function $g_{\boldsymbol{k}}(t)$ $(\partial/\partial\boldsymbol{r})$ is a row, while $(\partial/\partial\boldsymbol{r}')$ is a column). The function $\Phi(\boldsymbol{r},t)$ is not the original function $\varphi(\boldsymbol{r},t)$ of equation (2.1);

Φ is related to φ by the statistical near-identity transformation. The ensemble average of the original function $\varphi(\boldsymbol{r}, t)$ is equal to $\Phi(\boldsymbol{r}, t)$.

Note that the averaged equation (3.6) contains less information than Green's function; e.g. in the case of well-separated scales the matrix $\boldsymbol{D}$ approaches a constant value (see § 3.2.1), and $\dot{\boldsymbol{D}} = 0$ has no information about this constant value.

3.1.2. *The one-particle statistics*

The Green function $G(\boldsymbol{r}, t)$ has a direct physical meaning; it is the probability of finding a particle at point $\boldsymbol{r}$ at instant t, provided this particle was at the origin $\boldsymbol{r} = 0$ at instant $t = 0$. The function $(2\pi)^d g_{\boldsymbol{k}}$ is the characteristic function of the probability distribution $G(\boldsymbol{r}, t)$:

$$g_{\boldsymbol{k}} = \frac{1}{(2\pi)^d} \int G(\boldsymbol{r}, t) \exp(-\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r})\, \mathrm{d}\boldsymbol{r}$$

and therefore

$$\langle \boldsymbol{r} \rangle = \int G(\boldsymbol{r}, t)\boldsymbol{r}\, \mathrm{d}\boldsymbol{r} = \mathrm{i}\, \frac{\partial[(2\pi)^d g]}{\partial \boldsymbol{k}}\bigg|_{\boldsymbol{k}=0} = \boldsymbol{R},$$

$$\langle \boldsymbol{r}\, \boldsymbol{r}' \rangle = \int G(\boldsymbol{r}, t)\boldsymbol{r}\boldsymbol{r}'\, \mathrm{d}\boldsymbol{r} = -\, \frac{\partial[(2\pi)^d g]}{\partial \boldsymbol{k}^2}\bigg|_{\boldsymbol{k}=0} = \boldsymbol{D} + \boldsymbol{R}\boldsymbol{R}', \quad \langle \boldsymbol{r}\boldsymbol{r}' \rangle - \langle \boldsymbol{r} \rangle \langle \boldsymbol{r} \rangle' = \boldsymbol{D}.$$

According to (3.1), the mean displacement $\boldsymbol{R}(t)$ and the variance tensor $\boldsymbol{D}(t)$ completely define the probability distribution $G(\boldsymbol{r}, t)$.

3.2. *Anomalous transport*

In this paper, we are interested in the situation when there is no separation scales; so we would like to predict the evolution of the averaged concentration of the passive tracer during time t inside the range of time scales (1.17). However, we start with the asymptotics of large time.

3.2.1. *Trapping regime*

If $t \gg T_{max}$, the oscillating terms in the integrals (3.4) and (3.5) can be neglected, and so

$$\boldsymbol{D}(t) = \int (\boldsymbol{b}_{\boldsymbol{k}}\boldsymbol{b}'_{-\boldsymbol{k}} + \boldsymbol{b}_{-\boldsymbol{k}}\boldsymbol{b}'_{\boldsymbol{k}}) \frac{\varepsilon_{\boldsymbol{k}}}{\Omega_{\boldsymbol{k}^2}}\, \mathrm{d}\boldsymbol{k};$$

$$\boldsymbol{R}(t) = \boldsymbol{u}t, \quad \boldsymbol{u} = \int (\boldsymbol{b}_{\boldsymbol{k}}\boldsymbol{b}'_{-\boldsymbol{k}} + \boldsymbol{b}_{-\boldsymbol{k}}\boldsymbol{b}'_{\boldsymbol{k}})\boldsymbol{k} \frac{\varepsilon_{\boldsymbol{k}}}{\Omega_{\boldsymbol{k}}}\, \mathrm{d}\boldsymbol{k} + \frac{\mathrm{i}}{2} \int (\boldsymbol{b}_{\boldsymbol{k}}\boldsymbol{b}'_{-\boldsymbol{k}} - \boldsymbol{b}_{-\boldsymbol{k}}\boldsymbol{b}'_{\boldsymbol{k}})\boldsymbol{k} \frac{\varepsilon_{\boldsymbol{k}}}{\Omega^2_{\boldsymbol{k}}}\, \mathrm{d}\boldsymbol{k}.$$

Thus, as $t \to \infty$, the covariance matrix $\boldsymbol{D}(t)$ approaches a constant matrix, while the mean $\langle \boldsymbol{r} \rangle = \boldsymbol{R}$ becomes a linear function of time. The matrix $\dot{\boldsymbol{D}}$ approaches zero, and the terms with the second derivatives in the averaged equation (3.6) vanish. So the average evolution of the tracer is reduced to a drift with a constant speed $\boldsymbol{u}$. The tracer particles do not disperse, but remain trapped (cf. Kraichnan 1968; Majda & Kramer 1999) in a cluster, which drifts with the velocity $\boldsymbol{u}$. In this situation, the turbulent diffusion enters only at order ϵ^4 with respect to the small parameter (1.9) (see Herterich & Hasselmann 1982; Weichman & Glazman 1999, 2000; Balk & McLaughlin 1999).

3.2.2. *Vanishing of turbulent drift*

In general (at any observation time t), the turbulent drift $\boldsymbol{R}$ vanishes if the medium and the energy spectrum are isotropic. It also vanishes in an anisotropic situation if the velocity field is incompressible; according to (3.3), $\boldsymbol{R} = 0$ if $\boldsymbol{k} \cdot \boldsymbol{b_k} = 0$. So, in order to have a non-zero drift, we must assume both anisotropy and compressibility of the velocity field.

3.2.3. *The one-dimensional case*

In the one-dimensional case ($d = 1$) the velocity field is necessarily compressible, $\boldsymbol{b_k} = 1$. Consider a particular case of the general situation (1.10)–(1.11) when the energy spectrum is a power-law function (in the inertial interval), while the dispersion law is odd and has a power-law form for positive $\boldsymbol{k}$:

$$\varepsilon_k = \mathscr{C} k^{-\gamma} \ (K_{min} \ll k \ll K_{max}), \quad \Omega_k = \mathscr{A} k^{\alpha} \mathrm{sign}(\boldsymbol{k}) \ (\alpha > 0). \tag{3.7}$$

To calculate the integrals (3.4) and (3.5), we introduce a new integration variable $y = \mathscr{A} k^{\alpha} t$. So

$$\boldsymbol{D} = \xi^2 \left(\frac{t}{\tau}\right)^{\lambda} 4 \frac{I_1(\lambda)}{\alpha}, \tag{3.8}$$

where

$$\lambda = 2 + \frac{\gamma - d}{\alpha}, \quad I_1(\lambda) = \int_0^{\infty} \frac{1 - \cos y}{y^{\lambda}} \frac{\mathrm{d}y}{y}, \quad 0 < \lambda < 2 \Leftrightarrow d - 2\alpha < \gamma < d. \tag{3.9}$$

For reference, we present some values of the integral (3.9) (which are found with the aid of Maple): $I_1(\frac{1}{2}) = \sqrt{2\pi}$ (sub-diffusion), $I_1(1) = \frac{1}{2}\pi$ (normal diffusion), $I_1(\frac{3}{2}) = \frac{2}{3}\sqrt{2\pi}$ (super-diffusion). Also

$$\boldsymbol{R} = \xi \left(\frac{t}{\tau}\right)^{\mu} 2 \frac{I_2(\mu)}{\alpha}, \tag{3.10}$$

where

$$\mu = 2 + \frac{\gamma - d - 1}{\alpha} = \lambda - \frac{1}{\alpha}, \quad I_2(\mu) = \int_0^{\infty} \frac{y - \sin y}{y^{\mu}} \frac{\mathrm{d}y}{y} \ (1 < \mu < 3). \tag{3.11}$$

Although in this calculation we have considered only the case $d = 1$, we have written these formulae with dimension d as a parameter since the expressions for the exponents λ and μ are correct in any dimension d.

3.2.4. *The two-dimensional case*

As an example of the two-dimensional situation ($d = 2$), we consider isotropic medium (more precisely there is invariance with respect to rotations, but there is no mirror symmetry), but assume anisotropic energy spectrum. The dispersion law is just a power-law function, $\Omega_{\boldsymbol{k}} = \mathscr{A} k^{\alpha}$, and the polarization vector is

$$\boldsymbol{b_k} = \mathrm{i} p \frac{\boldsymbol{k}}{k} + \mathrm{i} q \frac{\tilde{\boldsymbol{k}}}{k},$$

where $\tilde{\boldsymbol{k}}$ is obtained from the vector $\boldsymbol{k} = (k_x, k_y)$ by 90° rotation: $\tilde{\boldsymbol{k}} = (k_y, -k_x)$; p and q are some scalar coefficients. To satisfy our normalization $|\boldsymbol{b_k}|^2 = l^2 + m^2 = 1$, we assume parameterization $p = \cos\theta, \ q = \sin\theta$. Introducing the polar coordinates $\boldsymbol{k} = (k\cos\phi, k\sin\phi)$, we have

$$\boldsymbol{b_k} = \mathrm{i}\cos\theta \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} + \mathrm{i}\sin\theta \begin{bmatrix} \sin\phi \\ -\cos\phi \end{bmatrix} = \mathrm{i} \begin{bmatrix} \cos(\phi - \theta) \\ \sin(\phi - \theta) \end{bmatrix}. \tag{3.12}$$

If the parameter $\theta = 0$, the polarization vector is parallel to the wave vector $\boldsymbol{k}$; if $\theta = \frac{1}{2}\pi$, the polarization vector is orthogonal to $\boldsymbol{k}$, i.e. the velocity field is incompressible.

We take the energy spectrum in the form

$$\varepsilon_{\boldsymbol{k}} = \mathscr{C}k^{-\gamma}\left\{1 + \sum_{n=1}^{\infty}(\Gamma_n \cos n\phi + \Upsilon_n \sin n\phi)\right\} \text{ in the inertial range } K_{min} \ll k \ll K_{max},$$

where Γ_n, Υ_n are some numbers, sufficiently small so that the spectrum $\varepsilon_{\boldsymbol{k}}$ stays positive for any ϕ. Then according to formulae (3.4)–(3.5), we find

$$\boldsymbol{D} = \boldsymbol{D}_0 \int_0^{\infty} \frac{1 - \cos \mathscr{A}k^{\alpha}t}{(\mathscr{A}k^{\alpha})^2}\mathscr{C}k^{-\gamma+d}\,\frac{\mathrm{d}k}{k}, \tag{3.13}$$

$$\boldsymbol{R} = \boldsymbol{R}_0 \int_0^{\infty} \frac{\mathscr{A}k^{\alpha}t - \sin \mathscr{A}k^{\alpha}t}{(\mathscr{A}k^{\alpha})^2}\mathscr{C}k^{1-\gamma+d}\,\frac{\mathrm{d}k}{k}, \tag{3.14}$$

where the matrix $\boldsymbol{D}_0$ and the vector $\boldsymbol{R}_0$ are the results of integration in (3.4) and (3.5) over the angle ϕ:

$$\boldsymbol{D}_0 = 2\int_0^{2\pi} \begin{bmatrix} \cos^2(\phi-\theta) & \cos(\phi-\theta)\sin(\phi-\theta) \\ \sin(\phi-\theta)\cos(\phi-\theta) & \sin^2(\phi-\theta) \end{bmatrix} \times \left\{1 + \sum_{n=1}^{\infty}(\Gamma_n \cos n\phi + \Upsilon_n \sin n\phi)\right\}\mathrm{d}\phi,$$

$$\boldsymbol{R}_0 = \int_0^{2\pi} \begin{bmatrix} \cos^2(\phi-\theta) & \cos(\phi-\theta)\sin(\phi-\theta) \\ \sin(\phi-\theta)\cos(\phi-\theta) & \sin^2(\phi-\theta) \end{bmatrix}\begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} \times \left\{1 + \sum_{n=1}^{\infty}(\Gamma_n \cos n\phi + \Upsilon_n \sin n\phi)\right\}\mathrm{d}\phi.$$

From the entire sum in brackets, the non-zero contribution to $\boldsymbol{D}_0$ comes only from the Γ_2- and Υ_2-terms, as well as from the unit. The non-zero contribution to $\boldsymbol{R}_0$ comes only from the Γ_1- and Υ_1-terms. So,

$$\boldsymbol{D}_0 = \pi \begin{bmatrix} 2 + \Gamma_2\cos 2\theta + \Upsilon_2 \sin 2\theta & -\Gamma_2 \sin 2\theta + \Upsilon_2 \cos 2\theta \\ -\Gamma_2 \sin 2\theta + \Upsilon_2 \cos 2\theta & 2 - \Gamma_2 \cos 2\theta - \Upsilon_2 \sin 2\theta \end{bmatrix}, \tag{3.15}$$

$$\boldsymbol{R}_0 = \frac{\pi}{2}\begin{bmatrix} \Gamma_1(1+\cos 2\theta) + \Upsilon_1 \sin 2\theta \\ -\Gamma_1 \sin 2\theta + \Upsilon_1(1+\cos 2\theta) \end{bmatrix}. \tag{3.16}$$

To calculate the integrals (3.13) and (3.14), we introduce the new integration variable $y = \mathscr{A}k^{\alpha}t$. So

$$\boldsymbol{D} = \xi^2 \left(\frac{t}{\tau}\right)^{\lambda} \boldsymbol{D}_0\,\frac{I_1(\lambda)}{\alpha}, \tag{3.17}$$

$$\boldsymbol{R} = \xi \left(\frac{t}{\tau}\right)^{\mu} \boldsymbol{R}_0\,\frac{I_2(\mu)}{\alpha}, \tag{3.18}$$

where the exponents λ and μ, as well as the functions $I_1(\lambda)$ and $I_2(\mu)$ are defined in (3.9) and (3.11).

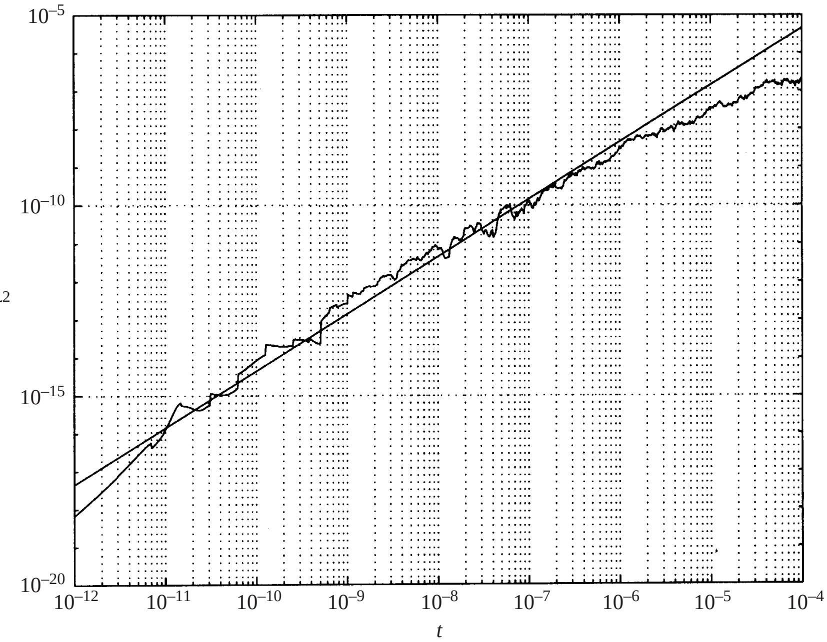

FIGURE 1. Super-diffusion in one dimension during a five-decade time interval. The velocity field exponents are $\alpha = 3/2$ and $\gamma = 1/4$ ($\delta = -1/4$); so, we have super-diffusion with the exponent $\lambda = 3/2$. The curve shows the numerical simulation, and the straight line represents the theoretical prediction. The graph is presented in non-dimensional form: time t is measured in units of time scale τ and the displacement σ is measured in units of length scale ξ. The non-dimensional cutoff parameters are $K_{min} = 10^4$, $K_{max} = 10^8$, $T_{min} = 10^{-12}$, $T_{max} = 10^{-6}$. (In dimensional units, we have this picture, e.g. if $\mathscr{A} = 1$ and $\mathscr{C} = 0.01$; then $\xi = 10^8$, $\tau = 10^{12}$, and therefore, $K_{min} = 10^{-4}$, $K_{max} = 10^4$, $T_{min} = 10$, $T_{max} = 10^6$). The value of the small parameter (1.15) is $\epsilon^2 = 0.1$. To obtain these results, we have used 10^4 grid points equally spaced between K_{min} and K_{max}; we averaged over 10 realizations of the velocity field.

In particular, if the velocity field is isotropic ($\Gamma_n = \Upsilon_n = 0,\ n = 1, 2, \ldots$), then $\boldsymbol{R} = 0$ and

$$\sigma^2 \equiv \mathrm{Trace}(\boldsymbol{D}) = \xi^2 \left(\frac{t}{\tau}\right)^{\lambda} 4\pi \frac{I_1(\lambda)}{\alpha}. \tag{3.19}$$

3.2.5. *Anomalous diffusion*

The exponent λ in (3.8) and (3.17) determines the anomalous diffusion. When $0 < \lambda < 1$, we have a sub-diffusive behaviour. When $1 < \lambda < 2$, we have a super-diffusion. At the lower limit, $\lambda = 0$, we have a trapping regime (see § 3.2.1), whereas the upper limit, $\lambda = 2$, corresponds to extremely large-scale velocity field, when the tracer moves with constant (but random) velocity. Balk (2000) presented the numerical confirmation of the anomalous diffusion.

Figure 1 presents super-diffusion observed in numerical simulation. The power law (3.8) is observed during time interval of more than a five-decade, during which the mean square displacement σ grows 10^4 times. Let us stress that not only the exponent λ, but also the factor in the power-law (3.8), are predicted.

3.2.6. *Anomalous drift*

From (3.10) and (3.18) we see that the passive tracer can exhibit not only the anomalous diffusion, but also an anomalous drift. When $\mu = 1$, we have normal drift, as in the case of well separated scales (§ 3.2.1). If $\mu > 1$, we have super-drift. The exponent μ cannot be less than 1, i.e. there is only normal drift or super-drift, but no slow-drift. Let us note that if $\alpha > 0$, the exponent μ cannot exceed 2, since $\mu = \lambda - 1/\alpha$ and $\lambda < 2$. See Balk (2001) for numerical confirmation of the super-drift.

4. The two-point correlation function

Now we would like to find the evolution of the two-point (same time) correlation function, i.e. the evolution of the ensemble average of the function

$$F_{12}(t) = f_1(t) f_2^*(t).$$

According to equation (2.12), the function $F_{12}(t)$ satisfies the following equation

$$\dot{F}_{12} + \int \hat{W}_{-134} F_{32} A_4 \, \mathrm{d}_{34} + \int \hat{W}^*_{-234} F_{13} A^*_4 \, \mathrm{d}_{34} = 0. \tag{4.1}$$

4.1. *Transformation*

In order to find the equation for the ensemble average of the function $F_{12}(t)$, we make the statistical near-identity transformation

$$F_{12} = G_{12} + \int \hat{X}_{-134} G_{32} A_4 \, \mathrm{d}_{34} + \int \hat{X}^*_{-234} G_{13} A^*_4 \, \mathrm{d}_{34}. \tag{4.2}$$

Substituting this into the equation (4.1), we find the equation for the new variable $G_{12}(t)$

$$\begin{aligned}
\dot{G}_{12} &+ \int \dot{\hat{X}}_{-134} G_{32} A_4 \, \mathrm{d}_{34} + \int \dot{\hat{X}}^*_{-234} G_{13} A^*_4 \, \mathrm{d}_{34} \\
&+ \int \hat{X}_{-134} \dot{G}_{32} A_4 \, \mathrm{d}_{34} + \int \hat{X}^*_{-234} \dot{G}_{13} A^*_4 \, \mathrm{d}_{34} \\
&+ \int \hat{W}_{-134} \left(G_{32} + \int \hat{X}_{-356} G_{52} A_6 \, \mathrm{d}_{56} + \int \hat{X}^*_{-256} G_{35} A^*_6 \, \mathrm{d}_{56} \right) A_4 \, \mathrm{d}_{34} \\
&+ \int \hat{W}^*_{-234} \left(G_{13} + \int \hat{X}_{-156} G_{53} A_6 \, \mathrm{d}_{56} + \int \hat{X}^*_{-356} G_{15} A^*_6 \, \mathrm{d}_{56} \right) A^*_4 \, \mathrm{d}_{34} = 0.
\end{aligned}$$

We choose the kernel X as in § 2.2.4 (see (2.28)), so that $\dot{X}_{-134} + W_{-134} = 0$, and the terms linear in A disappear from the equation.

4.1.1. *Eliminating divergence*

However, even after this elimination, $\dot{G}$ is not of the order ϵ^2, since now we allow the exponent γ to be greater than the dimension d. When $\gamma > d$, the mean square displacement $\boldsymbol{D}$ is infinite; and so, we assumed $\gamma < d$ in § 3. At the same time, the relative mean square displacement can be finite, since long waves transport a couple of sufficiently close particles in a similar way); and so, in this section, we consider a wider range of the exponent γ, in particular, $\gamma > d$ (cf. § 1.2). There is a problem with terms containing $\dot{G}$, e.g. term

$$Q_{12} = \int \hat{X}_{-134} \dot{G}_{32} A_4 \, \mathrm{d}_{34}.$$

Its mean square is

$$\langle |Q_{12}|^2\rangle = \int \hat{X}_{-134}\hat{X}^*_{-156}\dot{G}_{32}\dot{G}^*_{52}\langle A_4A^*_6\rangle\, \mathrm{d}_{3456}$$
$$= \int |X_{-134}|^2|\dot{G}_{32}|^2 E_4\delta(-\boldsymbol{k}_1+\boldsymbol{k}_3+\boldsymbol{k}_4)\,\mathrm{d}_{34}.$$

If $\gamma > d$, this integral diverges as $k_4 \to 0$. Indeed, $E_4 \propto k_4^{-\gamma}$ and $k_3 \to k_1$, as $k_4 \to 0$.

So, we write the new equation in the form

$$\dot{G}_{12}\left[1 + \int \hat{X}_{-134}A_4\,\mathrm{d}_{34} + \int \hat{X}^*_{-234}A^*_4\,\mathrm{d}_{34}\right]$$
$$+ \int \hat{X}_{-134}(\dot{G}_{32}-\dot{G}_{12})A_4\,\mathrm{d}_{34} + \int \hat{X}^*_{-234}(\dot{G}_{13}-\dot{G}_{12})A^*_4\,\mathrm{d}_{34}$$
$$+ \int \hat{W}_{-134}X_{-356}G_{52}A_4A_6\,\mathrm{d}_{3456} + \int \hat{W}_{-134}X^*_{-256}G_{35}A_4A^*_6\,\mathrm{d}_{3456}$$
$$+ \int \hat{W}^*_{-234}\hat{X}_{-156}G_{53}A^*_4A_6\,\mathrm{d}_{3456} + \int \hat{W}^*_{-234}\hat{X}^*_{-356}G_{15}A^*_4A^*_6\,\mathrm{d}_{3456} = 0.$$

The terms in the second line of this equation are of order ϵ^3, and we neglect them.

The first line in the passive scalar case has the form.

$$\dot{G}_{12}[1 + \mathrm{i}(\boldsymbol{k}_2-\boldsymbol{k}_1)\cdot \boldsymbol{Z}(t)]$$

where

$$\boldsymbol{Z}(t) = \int \boldsymbol{b}_{\boldsymbol{k}}\frac{1-\mathrm{e}^{-\mathrm{i}\omega t}}{\mathrm{i}\omega}A_q\,\mathrm{d}\boldsymbol{k}\,\mathrm{d}\omega.$$

We consider $\boldsymbol{k}_2 = \boldsymbol{k}_1$, so that the vector $\boldsymbol{Z}(t)$ disappears from the equation. Thus, we can find only the behaviour of the function

$$N_{\boldsymbol{k}} = \langle F_{\boldsymbol{kk}}\rangle = \langle |f_{\boldsymbol{k}}(t)|^2\rangle = \frac{1}{(2\pi)^{2d}}\int \langle \phi(\boldsymbol{r}_1,t)\phi^*(\boldsymbol{r}_2,t)\rangle \exp(-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2))\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2. \tag{4.3}$$

Moments $\langle F_{12}(t)\rangle$ for $\boldsymbol{k}_1 \neq \boldsymbol{k}_2$ seem to be non-universal (depending on the details at the ends of the inertial interval).

4.2. *The function $N_{\boldsymbol{k}}$ and the relative mean square displacement*

The calculation of the universal moments $\langle F_{12}(t)\rangle$ with $\boldsymbol{k}_1 = \boldsymbol{k}_2$ turns out to be sufficient to find the quantities (1.7) and (1.8). Differentiating (4.3) with respect to $\boldsymbol{k}$ and then assuming $\boldsymbol{k} = 0$, we find

$$\mathbf{S} = -(2\pi)^{2d}\left[\frac{\partial^2}{\partial \boldsymbol{k}'\partial \boldsymbol{k}}N_{\boldsymbol{k}}\right]_{\boldsymbol{k}=0}, \quad \rho^2 = -(2\pi)^{2d}\left[\frac{\partial^2}{\partial \boldsymbol{k}\,\partial \boldsymbol{k}'}N_{\boldsymbol{k}}\right]_{\boldsymbol{k}=0}, \tag{4.4}$$

($\partial/\partial\boldsymbol{k}$ is a row, while $\partial/\partial\boldsymbol{k}'$ is a column, so that $\partial^2N/\partial\boldsymbol{k}'\partial\boldsymbol{k}$ is a $d\times d$ matrix, and $\partial^2N/\partial\boldsymbol{k}\partial\boldsymbol{k}'$ is a scalar).

Consider, for example, the isotropic situation. Suppose the medium and the turbulence spectrum $E(\boldsymbol{k},\omega)$ are isotropic; suppose also that the initial condition for the function $N_{\boldsymbol{k}}$ is isotropic. Then, the function $N_{\boldsymbol{k}}$ remains isotropic for any $t > 0$, $N_{\boldsymbol{k}} = N(k,t)$, and the formula (4.4) gives

$$\rho^2 = -(2\pi)^{2d}\left[\frac{\partial^2 N}{\partial k^2} + \frac{d-1}{k}\frac{\partial N}{\partial k}\right]_{k=0} = -(2\pi)^{2d}d\left[\frac{\partial^2 N}{\partial k^2}\right]_{k=0}. \tag{4.5}$$

According to the definition (4.3), the function $N_{\boldsymbol{k}}(t)$ has the following properties:

(i) $N_{\boldsymbol{k}}(t)$ is dimensionless;
(ii) $N_{\boldsymbol{k}}(t)$ is positive;
(iii) If – as in the passive tracer case – $\phi(\boldsymbol{r}, t)$ is real, then $N_{\boldsymbol{k}} = N_{-\mathbf{k}}$;
(iv) If – as in the passive tracer case – $\phi(\boldsymbol{r}, t)$ is positive, then $N_{\boldsymbol{k}} \leqslant N|_{\boldsymbol{k}=0}$;
(v) The normalization (1.3) implies that $N|_{\boldsymbol{k}=0} = 1/(2\pi)^{2d}$.

4.2.1. *Examples*

Example 1. Isotropic passive tracer cloud of characteristic size s.

Consider a two-dimensional isotropic cloud with Gaussian distribution

$$\varphi(\boldsymbol{r}) = \frac{1}{2\pi s^2} \exp\left(-\frac{r^2}{2s^2}\right). \tag{4.6}$$

It satisfies the normalization (1.3); for this distribution, $\langle x^2 \rangle = \langle y^2 \rangle = s^2$. The variance of (4.6) is $\sigma^2 = \int r^2 \varphi(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} = 2s^2$, and so the mean squared distance between two particles in this distribution is $\rho^2 = 2\sigma^2 = 4s^2$.

By (4.3), we can find the corresponding function $N_{\boldsymbol{k}}$. Considering (4.6) as an initial distribution, we do not need averaging in (4.3) over the ensemble of the velocity fields, and we find

$$N_{\boldsymbol{k}} = \frac{1}{(2\pi)^4} \exp(-s^2 k^2). \tag{4.7}$$

This example illustrates the general rule that follows from the definition (4.3). If ρ is the characteristic size of the passive tracer cloud, then $1/\rho$ is the characteristic size of the support of the function $N_{\boldsymbol{k}}$, i.e. the characteristic size of the region in the $\boldsymbol{k}$-space where the function $N_{\boldsymbol{k}}$ is essentially non-zero.

Example 2. Two particles separated by distance s.

Consider a set of two particles, located at the points $\frac{1}{2}$ and $-\frac{1}{2}\boldsymbol{s}$ ($\boldsymbol{s}$ is the radius-vector from one particle to the other). In this case

$$\varphi(\boldsymbol{r}) = \tfrac{1}{2}[\delta(\boldsymbol{r} - \tfrac{1}{2}\boldsymbol{s}) + \delta(\boldsymbol{r} + \tfrac{1}{2}\boldsymbol{s})]. \tag{4.8}$$

The diameter of this two-particle cloud is $\rho = s/\sqrt{2}$ (which agrees with the fact that ρ^2 is the average of 4 possible squared distances in the set of 2 tracer particles: $\rho^2 = \frac{1}{4}(s^2 + s^2 + 0 + 0)$).

According to (4.3), we have

$$N_{\boldsymbol{k}} = \frac{1}{2(2\pi)^{2d}} [1 + \cos(\boldsymbol{k} \cdot \boldsymbol{s})]. \tag{4.9}$$

4.3. *Averaging*

Now we average our equation, assuming that the new variable G is statistically independent from the field A (see § 2.2.5):

$$\dot{N}_1 + \int \hat{W}_{-134} X_{-31-4} N_1 E_4 \,\mathrm{d}_{34} + \int \hat{W}_{-134} X^*_{-134} N_3 E_4 \,\mathrm{d}_{34} + \int \hat{W}^*_{-134} X_{-134} N_3 E_4 \,\mathrm{d}_{34} + \int \hat{W}^*_{-134} X^*_{-31-4} N_1 E_4 \,\mathrm{d}_{34} = 0.$$

It is a closed equation for the function $N_{\boldsymbol{k}} = G_{\boldsymbol{kk}} = \langle F_{\boldsymbol{kk}} \rangle$. Now we recall the formulae (2.13) and (2.28) for the kernels W and X, and find the following equation

$$\begin{aligned}\dot{N}_1 &= 2\int |U_{-134}|^2 \frac{\sin(\sigma_1-\sigma_3-\omega_4)t}{\sigma_1-\sigma_3-\omega_4}\delta(\boldsymbol{k}_1-\boldsymbol{k}_3-\boldsymbol{k}_4)N_3E_4\,\mathrm{d}_{34} \\ &+N_1\int \{U_{-134}U_{-31-4}+U^*_{-134}U^*_{-31-4}\}\,\frac{\sin(\sigma_1-\sigma_3-\omega_4)t}{\sigma_1-\sigma_3-\omega_4}\delta(\boldsymbol{k}_1-\boldsymbol{k}_3-\boldsymbol{k}_4)E_4\,\mathrm{d}_{34} \\ &+\mathrm{i}N_1\int\{U_{-134}U_{-31-4}-U^*_{-134}U^*_{-31-4}\}\,\frac{1-\cos(\sigma_1-\sigma_3-\omega_4)t}{\sigma_1-\sigma_3-\omega_4}\delta(\boldsymbol{k}_1-\boldsymbol{k}_3-\boldsymbol{k}_4)E_4\,\mathrm{d}_{34}.\end{aligned} \tag{4.10}$$

We could derive this equation without averaging if in the original equation (4.1) we made near-identity transformation up to order ϵ^2 (instead of the transformation (4.2) up to order ϵ). Then, we would push the randomness to order ϵ^3 and just neglect the ϵ^3-terms (cf. § 2.2.5). However, this leads to more cumbersome calculations.

4.3.1. *Remark about Hamiltonian symmetry*

Suppose that the kernel U satisfies the condition

$$U_{-31-4} = -U^*_{-134} \quad \text{provided} \quad \boldsymbol{k}_1-\boldsymbol{k}_3-\boldsymbol{k}_4 = 0. \tag{4.11}$$

Then equation (4.10) is reduced to the following form

$$\dot{N}_1 = 2\int |U_{-134}|^2(N_3-N_1)E_4\frac{\sin(\sigma_1-\sigma_3-\omega_4)t}{\sigma_1-\sigma_3-\omega_4}\,\delta(\boldsymbol{k}_1-\boldsymbol{k}_3-\boldsymbol{k}_4)\,\mathrm{d}_{34}. \tag{4.12}$$

The symmetry (4.11) means that the system has a Hamiltonian structure (cf. Zakharov 1974). In particular, this symmetry takes place for the Schrödinger equation with random potential, when $U = \mathrm{i}$ (see (2.15)). It also holds in the passive tracer case (2.14) if the velocity field is incompressible, $\boldsymbol{k}\cdot\boldsymbol{b}_{\boldsymbol{k}} = 0$. The symmetry (4.11) does not hold for compressible flows. However, even in the compressible case, the averaged equation still has the form (4.12) owing to certain cancellations (see § 5).

Equation (4.12) reminds us of the wave kinetic equation

$$\dot{N}_1 = 2\int |U_{-134}|^2(N_3-N_1)E_4\delta(\sigma_1-\sigma_3-\Omega_4)\delta(\boldsymbol{k}_1-\boldsymbol{k}_3-\boldsymbol{k}_4)\,\mathrm{d}_{34}$$

for the scattering of waves with dispersion law σ on waves with dispersion law Ω (see Zakharov, Musher & Rubenchik 1985; Ryzhik, Papanicolaou & Keller 1996). However, there are two important differences. First, instead of the usual frequency delta function, our equation contains a specific time-dependence. Secondly, in our equation, the function $N_{\boldsymbol{k}}$ is not the spectrum of the wave field with dispersion law $\sigma_{\boldsymbol{k}}$; in the latter case we would have $\langle f_1 f_2^* \rangle = N_1\delta(\boldsymbol{k}_1-\boldsymbol{k}_2)$. Instead, $N_{\boldsymbol{k}} = \langle f_{\boldsymbol{k}} f^*_{\boldsymbol{k}} \rangle$; the f-field is the Fourier image of the field φ which is not statistically homogeneous, but concentrated in a finite region of the $\boldsymbol{r}$-space.

The Hamiltonian symmetry (4.11) implies the conservation of 'the total number of quasi-particles'

$$\mathscr{N} = \int N_{\boldsymbol{k}}\,\mathrm{d}\boldsymbol{k} = \frac{1}{(2\pi)^d}\int |\varphi(\boldsymbol{r},t)|^2\,\mathrm{d}\boldsymbol{r}. \tag{4.13}$$

Indeed, the time derivative of the quantity $\mathscr{N}$ by virtue of equation (4.12) is equal zero.

The Hamiltonian symmetry (4.11) implies the existence of 'entropy'

$$\Theta = \tfrac{1}{2} \int N_{\boldsymbol{k}}^2 \, \mathrm{d}\boldsymbol{k}. \tag{4.14}$$

Indeed, by virtue of (4.12) we have

$$\dot{\Theta} = -\int |U_{-134}|^2 (N_3 - N_1)^2 E_4 \frac{\sin(\sigma_1 - \sigma_3 - \omega_4)t}{\sigma_1 - \sigma_3 - \omega_4} \delta(\boldsymbol{k}_1 - \boldsymbol{k}_3 - \boldsymbol{k}_4) \, \mathrm{d}_{134}.$$

Here, we have taken into account the symmetry (4.11) and the fact that $E_q = E_{-q}$. If the main lobe $|\sigma_1 - \sigma_3 - \omega_4| < \pi/2$ of the function $(\sin(\sigma_1 - \sigma_3 - \omega_4)t)/(\sigma_1 - \sigma_3 - \omega_4)$ dominates, then $\dot{\Theta} \leqslant 0$.

5. The spreading of the passive tracer cloud

5.1. *Equations*

Specifying equation (4.10) for the passive tracer case (2.14), we find (after some straightforward algebra) the following equation

$$\dot{N}_1 = 2 \int |\boldsymbol{k}_1 \cdot \boldsymbol{b}_4|^2 \frac{\sin \omega_4 t}{\omega_4} (N_3 - N_1) E_4 \delta(\boldsymbol{k}_1 - \boldsymbol{k}_3 - \boldsymbol{k}_4) \, \mathrm{d}_{34}. \tag{5.1}$$

We could expect to obtain an extra term, proportional to N_1, from the last two integrals in (4.10), besides that already present in (5.1); however, the extra term actually vanishes owing to the symmetry $E_q = E_{-q}$.

If the velocity field is due to linear waves with dispersion law $\omega = \Omega_{\boldsymbol{k}}$ (see (2.5) and (2.6)), then we can integrate over $\mathrm{d}\omega_4$ in (5.1) and find

$$\dot{N}_1 = 2 \int |\boldsymbol{k}_1 \cdot \boldsymbol{b}_4|^2 \frac{\sin \Omega_4 t}{\Omega_4} (N_3 - N_1) \varepsilon_4 \delta(\boldsymbol{k}_1 - \boldsymbol{k}_3 - \boldsymbol{k}_4) \, \mathrm{d}\boldsymbol{k}_3 \, \mathrm{d}\boldsymbol{k}_4. \tag{5.2}$$

This equation preserves the properties (i)–(v) described in § 4.3. In particular,

$$0 \leqslant (2\pi)^{2d} N_{\boldsymbol{k}}(t) \leqslant 1.$$

Let us show the validity of the first inequality. Initially, $N_{\boldsymbol{k}}$ is positive. Suppose, at some instant t_0, the quantity $N_{\boldsymbol{k}}$ turns into zero for the first time, at some $\boldsymbol{k}_0$, to become negative at k_0 for $t > t_0$; $N_{\boldsymbol{k}}(t_0)$ is still positive for all other $\boldsymbol{k} \neq \boldsymbol{k}_0$. Such a situation contradicts equation (5.2). $\dot{N}_{\boldsymbol{k}_0}(t_0)$ is negative (or zero), but the right-hand side of (5.2) is positive at $t = t_0, \boldsymbol{k} = \boldsymbol{k}_0$. To prove the second inequality, consider function $M_{\boldsymbol{k}}(t) = 1 - (2\pi)^{2d} N_{\boldsymbol{k}}(t)$.

5.1.1. *Convergence conditions*

The integral (5.2) has singularities at $k_4 \to 0$ and $k_4 \to \infty$. We will consider the convergence when $\varepsilon_{\boldsymbol{k}} = \varepsilon_{-\boldsymbol{k}}$ (actually we need this condition only near the origin $\boldsymbol{k} = 0$, namely $\varepsilon_{\boldsymbol{k}} - \varepsilon_{-\boldsymbol{k}} = O(k)$, $k \to 0$). To see the convergence conditions, we change the integration variable $\boldsymbol{k}_4 \to -\boldsymbol{k}_4$ and rewrite equation (5.2) in the form

$$\dot{N}_1 = \int |\boldsymbol{k}_1 \cdot \boldsymbol{b}_4|^2 \frac{\sin \Omega_4 t}{\Omega_4} \varepsilon(\boldsymbol{k}_4) \left[N(\boldsymbol{k}_1 - \boldsymbol{k}_4, t) - 2N(\boldsymbol{k}_1, t) + N(\boldsymbol{k}_1 + \boldsymbol{k}_4, t) \right] \mathrm{d}\boldsymbol{k}_4. \tag{5.3}$$

Here, we took into account that $|\boldsymbol{k}_1 \cdot \boldsymbol{b}_4| = |\boldsymbol{k}_1 \cdot \boldsymbol{b}_{-4}|$ (which follows from the property $\boldsymbol{b}_{-4} = \boldsymbol{b}_4^*$) and $\sin \Omega_4 t / \Omega_4 = \sin \Omega_{-4} t / \Omega_{-4}$ (which is valid for even and odd dispersion laws).

As $k_4 \to 0$, the integral (5.3) converges if $\gamma < d + 2$. As $k_4 \to \infty$, the integral

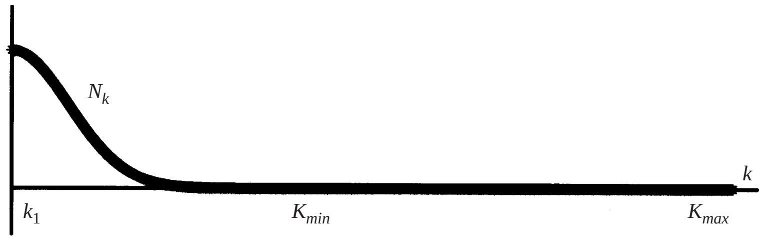

FIGURE 2. For a 'very large' passive tracer cloud, the characteristic size of the support of $N_{\boldsymbol{k}}$ in the $\boldsymbol{k}$-space is 'very small'.

converges if $\gamma > d - 2\alpha$. Thus, if $d - 2\alpha < \gamma < d + 2$, the spreading of the passive tracer cloud is universal, i.e. independent of the details of the energy spectrum $\varepsilon_{\boldsymbol{k}}$ at the ends of the inertial interval (1.11). This is similar to the universality of the Kolmogorov–Zakharov spectra of weak turbulence (see Zakharov 1985).

The convergence of the integral in (5.2) is necessary for the applicability of the perturbational approach. For the applicability of equation (5.2), we can require that the time derivative given by equation (5.2) be much less than the frequency of the travelling waves: $\dot{N} \ll \Omega$.

The behaviour of the passive tracer cloud crucially depends on the relation of the size ρ of the cloud and the observation time t to the length and time scales of the velocity field. A cloud with initial diameter ρ_0 starts spreading in four different regimes:

(i) 'very small' cloud, $\rho_0 \ll \Lambda_{min}$,
(ii) 'small' cloud, $\Lambda_{min} \ll \rho_0 \ll \xi$,
(iii) 'large' cloud, $\xi \ll \rho_0 \ll \Lambda_{max}$,
(iv) 'very large' cloud, $\Lambda_{max} \ll \rho_0$,

where the scale ξ is defined in (1.12) via the dimensional constants $\mathscr{A}$ and $\mathscr{C}$.

5.2. *The case of well-separated scales:* $\rho_0 \gg \Lambda_{max}$ *or* $\rho_0 \ll \Lambda_{min}$

5.2.1. *'Very large' cloud*

First, suppose that we have a 'very large' passive tracer cloud, $\rho_0 \gg \Lambda_{max}$. Then, the characteristic size of the support of $N_{\boldsymbol{k}}$ in the $\boldsymbol{k}$-space is 'very small': $1/\rho_0 \ll K_{min}$, which is shown schematically in figure 2.

Since we are interested in statistical moments, determined by the derivatives of the function $N_{\boldsymbol{k}}$ at $\boldsymbol{k} = 0$, such as (4.4), we can consider $\boldsymbol{k}_1$ in (5.2) being near zero ($k_1 \ll 1/\rho_0$). Then $\boldsymbol{k}_3 \approx -\boldsymbol{k}_4$, and we can neglect N_3 in (5.2):

$$\dot{N}_1 = -2N_1 \int |\boldsymbol{k}_1 \cdot \boldsymbol{b}_4|^2 \frac{\sin \Omega_4 t}{\Omega_4} \varepsilon_4 \, \mathrm{d}\boldsymbol{k}_4. \tag{5.4}$$

By virtue of (4.4), we can find the evolution of the relative mean square displacement; differentiate both parts of (5.4) with respect to $\boldsymbol{k}_1$ twice, and then put $\boldsymbol{k}_1 = 0$):

$$\dot{\mathbf{S}} = 2\dot{\mathbf{D}} \Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}\rho^2 = 2\frac{\mathrm{d}}{\mathrm{d}t}\sigma^2, \tag{5.5}$$

where the tensor of one-particle displacement $\mathbf{D}$ is given in (3.4). This agrees with the fact that two distant particles (separated by distance $\gg \Lambda_{max}$) are statistically independent. We can find the solution of (5.4) in the form

$$N_{\boldsymbol{k}}(t) = N^0_{\boldsymbol{k}}\,(2\pi)^{2d} g_{\boldsymbol{k}}(t) g^*_{\boldsymbol{k}}(t),$$

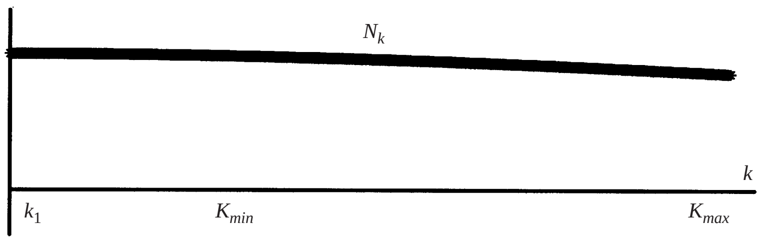

FIGURE 3. For a 'very small' passive tracer cloud, the characteristic size of the support of $N_{\boldsymbol{k}}$ in the $\boldsymbol{k}$-space is 'very large'.

where $g_{\boldsymbol{k}}(t)$ is Green's function in the Fourier representation, see § 2.3; $N^0_{\boldsymbol{k}} = N_{\boldsymbol{k}}(0)$ is the initial condition.

Suppose also that the observation time t is much larger than the largest time scale of the velocity field, $t \gg T_{max}$. Then $\dot{N}_{\boldsymbol{k}} = 0$ along with $\dot{\boldsymbol{S}} = 2\dot{\boldsymbol{D}} = 0$. As $t \to \infty$, the function $N_{\boldsymbol{k}}(t)$ approaches some constant function (independent of t, but depending on $\boldsymbol{k}$). This situation was studied by Weichman & Glazman 1999, 2002).

5.2.2. *'Very small' cloud*

Now suppose that we have a 'very small' passive tracer cloud, $\rho_0 \ll \Lambda_{min}$. Then, the characteristic size of the support of $N_{\boldsymbol{k}}$ in the $\boldsymbol{k}$-space is very large, $1/\rho_0 \gg K_{max}$, which is shown in figure 3.

For $k \ll 1/\rho_0$ we can approximate $N_{\boldsymbol{k}}$ by Taylor's formula:

$$N_{\boldsymbol{k}} = \frac{1}{(2\pi)^{2d}}(1 - \tfrac{1}{2}\boldsymbol{k}'\boldsymbol{S}\boldsymbol{k}).$$

We use this approximation in (5.2). Consider $\boldsymbol{k}_1$ near zero ($k_1 \ll K_{min}$), and so, $\boldsymbol{k}_3 \approx -\boldsymbol{k}_4$ (in particular, this implies that $\boldsymbol{k}_3 \ll 1/\rho_0$, and the approximation is valid for N_3 as well). Then

$$\dot{\boldsymbol{S}} = 2\int \boldsymbol{B}_{\boldsymbol{k}} \frac{\sin \Omega_{\boldsymbol{k}} t}{\Omega_{\boldsymbol{k}}} \varepsilon_{\boldsymbol{k}} (\boldsymbol{k}'\boldsymbol{S}\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}$$

($\boldsymbol{B}_{\boldsymbol{k}}$ is the polarization matrix defined in (3.2) via the polarization vector $\boldsymbol{b}_{\boldsymbol{k}}$).

If, in addition, the observation time is also 'very small': $t \ll T_{min}$, then $\sin \Omega_{\boldsymbol{k}} t = \Omega_{\boldsymbol{k}} t$, and

$$\frac{\mathrm{d}\boldsymbol{S}}{2t\mathrm{d}t} = \int \boldsymbol{B}_{\boldsymbol{k}} \varepsilon_{\boldsymbol{k}} (\boldsymbol{k}'\boldsymbol{S}\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}. \tag{5.6}$$

If we introduce new time $\eta = t^2$, the entries of the matrix $\boldsymbol{S}$ satisfy a system of linear homogeneous o.d.e. with constant coefficients, and therefore $\boldsymbol{S}$ is the sum of exponential functions of $\eta = t^2$. Of course, in each realization of the velocity field, two close particles diverge exponentially in time. However, when we average the Gaussian ensemble of the velocity fields, we find $\boldsymbol{S}$ to be the sum of exponential functions of $\eta = t^2$, according to (5.6). (This is easy to understand using the following simple model. Let $y(t) = y_0 \mathrm{e}^{ut}$, where u is a random variable with a Gaussian distribution $p(u) = (a/\sqrt{\pi})\exp(-a^2u^2)$ (y_0 and $a > 0$ are some constants). Then $\langle y(t)\rangle = y_0(a/\sqrt{\pi}) \int \exp(ut)\exp(-a^2u^2)\mathrm{d}u = y_0 \exp(t^2/(4a^2))$.)

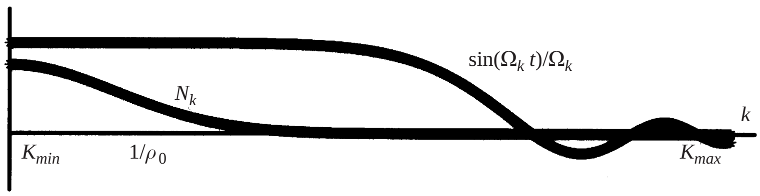

FIGURE 4. The support of both functions shrinks with time.

5.3. *Estimates showing two different anomalous regimes*

Now we start considering the anomalous situation when the initial diameter ρ_0 of the passive tracer cloud is well inside the range of the length scales of the velocity field, $\Lambda_{min} \ll \rho_0 \ll \Lambda_{max}$. We will see that the spreading of the cloud can occur in two different regimes, described in §§ 6 and 7. This depends on the relation of the diameter ρ_0 and the scale ξ (defined in § 1.2 via the dimensional factors $\mathscr{A}$ and $\mathscr{C}$). In order to see the existence of these two regimes, we make rough estimates in (5.3), which is equivalent to (5.2).

The passive tracer cloud is spreading in the $\boldsymbol{r}$-space, and, respectively, the 'support' of the function $N_{\boldsymbol{k}}$ (the area of essentially non-zero $N_{\boldsymbol{k}}$) is shrinking in the $\boldsymbol{k}$-space. In (5.2) or (5.3) we have two functions whose 'support' in the $\boldsymbol{k}$-space is shrinking with time

$$N_{\boldsymbol{k}}(t), \qquad \frac{\sin \Omega_{\boldsymbol{k}} t}{\Omega_{\boldsymbol{k}}}. \tag{5.7}$$

The 'support' of the second function is roughly $1/(\mathscr{A}t)^{1/\alpha}$. Initially, the support of the first function is $1/\rho_0$, and the support of the second function is infinite. As time progresses the 'supports' of both functions shrink (see figure 4).

At time $t = t_0 \sim \rho_0^\alpha/\mathscr{A}$, the 'support' of the second function reaches the value $1/\rho_0$, i.e. becomes comparable with the initial 'support' of the first function. Below, we estimate whether during this time the function $N_{\boldsymbol{k}}$ changes significantly, or changes little, and therefore its support stays almost the same, equal to $1/\rho_0$.

According to (5.3), the time derivative $N_{\boldsymbol{k}}(t)$ is less than or of the order of the following value

$$\frac{1}{\rho_0^2}\mathscr{C}\left(\frac{1}{\rho_0}\right)^\gamma t_0 \left(\frac{1}{\rho_0}\right)^d,$$

and so we have an estimate

$$|N_{\boldsymbol{k}}(t) - N_{\boldsymbol{k}}(0)| \sim \frac{\mathscr{C}t_0^2}{\rho_0^{2+d-\gamma}} \sim \frac{\mathscr{C}}{\mathscr{A}^2}\rho_0^{-2-d+\gamma+2\alpha} = \left(\frac{\rho_0}{\xi}\right)^{-\delta},$$

where ξ and δ are introduced in § 1.2; it is assumed throughout this paper that $\delta < 0$. From this estimate, we see that if $\rho_0 \ll \xi$, then we can neglect the change $N_{\boldsymbol{k}}(t)$ during the initial time t_0. After this time interval t_0, for some time (see § 6.3.1) we can assume that the support of the second function in (5.7) is much smaller than the support of the first. This situation is considered in § 6.

If, on the other hand, $\rho_0 \gg \xi$, then the function $N_{\boldsymbol{k}}(t)$ changes significantly during the initial time interval t_0, and the passive tracer cloud spreads significantly during t_0. The support of the first function in (5.7) shrinks and stays much smaller than the support of the second. We study this situation in § 7.

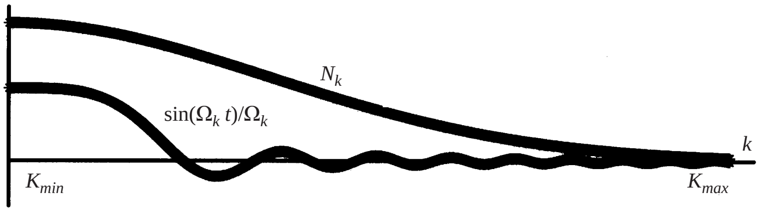

FIGURE 5. The situation where the 'support' of N_k is much larger than the 'support' of $\sin(\Omega_k t)/\Omega_k$.

6. 'Small' passive tracer cloud: differential approximation

In this section, we simplify the general equation (5.2) for the situation of the small cloud, when the 'support' of the second function in (5.7) is much smaller than the 'support' of the first. This situation is shown schematically in figure 5.

The essential contribution to the integral (6.1) comes only from the region of 'support' of the function $\sin \Omega_4 t/\Omega_4$, i.e. where $|\Omega_4|t < 1 \Leftrightarrow k_4 < (\pi/At)^{1/\alpha}$. Since $\boldsymbol{k}_1 - \boldsymbol{k}_3 = \boldsymbol{k}_4$, we can approximate the difference $N_3 - N_1$ by a differential expression and reduce (6.1) to a differential equation.

To be specific, let us consider the two-dimensional incompressible velocity field. Then in polar coordinates (k, ϕ)

$$\boldsymbol{k} = k \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix}, \quad \boldsymbol{b_k} = \mathrm{i} \begin{bmatrix} \sin\phi \\ -\cos\phi \end{bmatrix}, \quad \boldsymbol{k}_1 \cdot \boldsymbol{b}_4 = k_1 \sin(\phi_4 - \phi_1),$$

and (5.2) takes the form

$$\dot{N}_1 = 2k_1^2 \int \sin^2(\phi_4 - \phi_1) \frac{\sin \Omega_4 t}{\Omega_4} (N_3 - N_1)\varepsilon_4 \delta(\boldsymbol{k}_1 - \boldsymbol{k}_3 - \boldsymbol{k}_4)\, \mathrm{d}\boldsymbol{k}_3 \mathrm{d}\boldsymbol{k}_4. \tag{6.1}$$

In addition, we assume that the velocity field is statistically isotropic, i.e. the dispersion law and the energy spectrum are functions of the wavenumber only: $\Omega_{\boldsymbol{k}} = \Omega(k), \varepsilon_{\boldsymbol{k}} = \varepsilon(k)$. Note that the distribution of the passive tracer, i.e. the function $N_{\boldsymbol{k}}(t)$, can be anisotropic, depending on the polar angle ϕ, $N = N(k, \phi, t)$; this happens if the initial condition $N_{\boldsymbol{k}}(0)$ is anisotropic.

6.1. *Deriving the differential approximation via the weak form*

In order to derive the differential approximation, we rewrite equation (6.1) in the weak form: multiply (6.1) by a test function $\chi_1 = \chi(\boldsymbol{k}_1)$ and integrate over $\mathrm{d}\boldsymbol{k}_1$

$$\frac{\mathrm{d}}{\mathrm{d}t} \int N_1 \chi_1 \mathrm{d}\boldsymbol{k}_1 = -\int k_1^2 \sin^2(\phi_4 - \phi_1)(N_3 - N_1)(\chi_3 - \chi_1)\varepsilon_4 \frac{\sin \Omega_4 t}{\Omega_4} \times \delta(\boldsymbol{k}_1 - \boldsymbol{k}_3 - \boldsymbol{k}_4)\, \mathrm{d}_{134}. \tag{6.2}$$

Using the smallness of $\boldsymbol{k}_4$, we replace the finite differences by the corresponding differential expressions

$$N_3 - N_1 = -\frac{\partial N_1}{\partial \boldsymbol{k}_1} \boldsymbol{k}_4 = k_4 \left\{ \frac{\partial N_1}{\partial k_1} \cos(\phi_4 - \phi_1) + \frac{1}{k_1} \frac{\partial N_1}{\partial \phi_1} \sin(\phi_4 - \phi_1) \right\},$$

$$\chi_3 - \chi_1 = -\frac{\partial \chi_1}{\partial \boldsymbol{k}_1} \boldsymbol{k}_4 = k_4 \left\{ \frac{\partial \chi_1}{\partial k_1} \cos(\phi_4 - \phi_1) + \frac{1}{k_1} \frac{\partial \chi_1}{\partial \phi_1} \sin(\phi_4 - \phi_1) \right\}.$$

Then the integration over $\mathrm{d}\boldsymbol{k}_3$ in (6.2) just takes away the delta function. After that, we integrate over $\mathrm{d}\phi_4$, using the isotropy of the velocity field,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int N_1\chi_1 \mathrm{d}\boldsymbol{k}_1 = -\int \left\{ k_1^2 \frac{\partial N_1}{\partial k_1}\frac{\partial \chi_1}{\partial k_1} + 3\frac{\partial N_1}{\partial \phi_1}\frac{\partial \chi_1}{\partial \phi_1}\right\} \dot{Y}(t)\,\mathrm{d}\boldsymbol{k}_1,$$

where

$$Y(t) = \frac{\pi}{4}\int_0^\infty k^{2+d}\frac{1-\cos\Omega_k t}{\Omega_k^2}\varepsilon_k \frac{\mathrm{d}k}{k}. \tag{6.3}$$

We have defined the time-dependent coefficient in our equation as a time derivative $\dot{Y}(t)$ – instead of denoting it by some function of time, say $X(t)$ – in order to simplify further formulae; besides, $Y(t)$ (not $\dot{Y}(t)$) is dimensionless. Since the test function $\chi(\boldsymbol{k})$ is arbitrary, we have the following differential equation

$$\dot{N} = \dot{Y}(t)\frac{\partial}{\partial k}\left[k^2\frac{\partial N}{\partial k}\right] + 3\dot{Y}(t)\frac{\partial^2 N}{\partial \phi^2}. \tag{6.4}$$

6.1.1. *The power-law situation*

If the dispersion law and the energy spectrum are power-law functions of the wavenumber, $\Omega_k = \mathscr{A}k^\alpha$, $\varepsilon_k = \mathscr{C}k^{-\gamma}$, in the inertial range, then the quantity Y is a power-law function of time:

$$Y(t) = \mathscr{B}\left(\frac{t}{\tau}\right)^\beta \quad \text{where} \quad \beta = 2 + \frac{\gamma - d - 2}{\alpha} = \lambda - \frac{2}{\alpha}, \quad \mathscr{B} = \frac{\pi}{4}\frac{I_1(\beta)}{\alpha}, \tag{6.5}$$

(to derive this formula, we made the change of the integration variable in (6.3), $\mathscr{A}k^\alpha t = y$ and expressed the factors $\mathscr{A}$ and $\mathscr{C}$ via the scales ξ and τ, defined in § 1.2; the integral I_1 is defined in (3.9)).

The applicability of the differential approximation requires the convergence of the integral (6.3). This integral converges at $k \to 0$ if $2 + d - \gamma > 0$, and at $k \to \infty$ if $2 + d - \gamma - 2\alpha < 0$; the latter is the condition $\delta < 0$ assumed from the very beginning, in § 1.2.

6.2. *Mellin and Taylor expansions*

Equation (6.4) is easily solved by the Fourier transform in ϕ and Mellin transform in k. Our primary goal is the tensor of relative displacement (1.7), which is determined by the derivatives of the function $N_{\boldsymbol{k}}$ at the origin (see (4.4)). So, we find the function $N_{\boldsymbol{k}}$ only near the origin. Assuming that $N_{\boldsymbol{k}}$ is analytic at $\boldsymbol{k} = 0$, we expand it in the Taylor series

$$N = \sum_{u,v=0}^{\infty} h_{uv}(t)k_x^u k_y^v.$$

Since

$$k_x = k\frac{\mathrm{e}^{\mathrm{i}\phi} + \mathrm{e}^{-\mathrm{i}\phi}}{2}, \quad k_y = k\frac{\mathrm{e}^{\mathrm{i}\phi} - \mathrm{e}^{-\mathrm{i}\phi}}{2\mathrm{i}},$$

we have

$$N = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}{}' H_{lm}(t)k^l \mathrm{e}^{\mathrm{i}m\phi} \quad (H_{l,m}^* = H_{l,-m}), \tag{6.6}$$

where the second sum (marked by a prime) implies summation only over every other integer from $-l$ to l: $m = -l, -l+2, \ldots, l-2, l$. Substituting this expansion into (6.4) we find independent equations for the coefficients $H_{lm}(t)$

$$\dot{H}_{lm} = \dot{Y}(t)[l(l+1) - 3m^2]H_{lm}. \tag{6.7}$$

According to (6.6),

$$N = H_{00} + H_{11}(k_x + \mathrm{i}k_y) + H_{11}^*(k_x - \mathrm{i}k_y) + H_{20}(k_x^2 + k_y^2) \\ + H_{22}(k_x + \mathrm{i}k_y)^2 + H_{22}^*(k_x - \mathrm{i}k_y)^2 + O(k^3) \quad (k \to 0).$$

Using this in (4.4), we have

$$\mathbf{S} = -2(2\pi)^4 \begin{bmatrix} H_{20} + 2\mathrm{Re}(H_{22}) & -2\mathrm{Im}(H_{22}) \\ -2\mathrm{Im}(H_{22}) & H_{20} - 2\mathrm{Re}(H_{22}) \end{bmatrix}, \quad \rho^2 = -4(2\pi)^4 H_{20} \tag{6.8}$$

(Re and Im are real and imaginary parts). By virtue of (6.7) with $l = 2$ and $m = 0$, we find the evolution of the diameter of the passive tracer cloud

$$\rho(t) = \rho_0 \exp\{3Y(t)\}, \tag{6.9}$$

where ρ_0 is the initial diameter.

6.2.1. *Example: two particles* (*with fixed initial locations*)

Consider the evolution of two particles initially located at the two points $\frac{1}{2}\boldsymbol{s}$ and $-\frac{1}{2}\boldsymbol{s}$, so that the initial distance between the particles is s. Then initially,

$$\varphi(\boldsymbol{r}) = \tfrac{1}{2}[\delta(\boldsymbol{r} - \tfrac{1}{2}\boldsymbol{s}) + \delta(\boldsymbol{r} + \tfrac{1}{2}\boldsymbol{s})] \Rightarrow \\ N_{\boldsymbol{k}} = \frac{1}{2(2\pi)^4}[1 + \cos(\boldsymbol{k}\cdot\boldsymbol{s})] = \frac{1}{(2\pi)^4}\left[1 - \tfrac{1}{4}(\boldsymbol{k}\cdot\boldsymbol{s})^2 + O(k^4)\right] \quad (k \to 0).$$

If the polar axis is chosen in the direction of vector $\boldsymbol{s}$, then initially,

$$N_{\boldsymbol{k}} = \frac{1}{(2\pi)^4}\left[1 - \tfrac{1}{8}k^2s^2 - \tfrac{1}{16}k^2s^2\mathrm{e}^{2\mathrm{i}\phi} - \tfrac{1}{16}k^2s^2\mathrm{e}^{-2\mathrm{i}\phi} + O(k^4)\right] \quad (k \to 0),$$

$$H_{22} = \tfrac{1}{2}H_{20} = -\frac{s^2}{16(2\pi)^4}.$$

From these initial values by virtue of (6.7), we find the coefficients H_{20} and H_{22} at instant t, and then by (6.8) we obtain the tensor of relative displacement

$$\mathbf{S} = \frac{s^2}{4}\begin{bmatrix} \exp\{6Y(t)\} + \exp\{-6Y(t)\} & 0 \\ 0 & \exp\{6Y(t)\} - \exp\{-6Y(t)\} \end{bmatrix}. \tag{6.10}$$

This expression shows that, with time, the passive tracer cloud becomes more isotropic (on average): the distinction between the x-axis and y-axis disappears.

The mean diameter of the passive tracer cloud grows with time according to the law

$$\rho(t) = \frac{s}{\sqrt{2}}\exp\{3Y(t)\}.$$

Thus, the divergence of 'close' tracer particles is

$$\begin{array}{rlcl} \text{exponential} & \text{if} \quad \beta = 1 & \Leftrightarrow & \gamma = d + 2 - \alpha, \\ \text{sub-exponential} & \text{if} \quad 0 < \beta < 1 & \Leftrightarrow & d + 2 - 2\alpha < \gamma < d + 2 - \alpha, \\ \text{super-exponential} & \text{if} \quad 1 < \beta < 2 & \Leftrightarrow & d + 2 - \alpha < \gamma < d + 2. \end{array}$$

6.2.2. *Example: isotropic passive tracer cloud*

Consider the evolution of an ensemble of particles with an isotropic distribution $\varphi(\boldsymbol{r}, t)$ of initial characteristic size s, so that initially

$$\varphi(\boldsymbol{r}) = \frac{1}{2\pi s^2}\exp\left(-\frac{r^2}{2s^2}\right) \Rightarrow$$
$$N_{\boldsymbol{k}} = \frac{1}{(2\pi)^4}\exp(-s^2k^2) = \frac{1}{(2\pi)^4}[1 - s^2k^2 + O(k^4)] \quad (k \to 0).$$

Therefore,

$$\rho(t) = 2s\exp\{3Y(t)\}.$$

6.3. *The applicability of the differential approximation*

6.3.1. *Estimating the time during which the differential approximation is applicable*

The differential approximation can work only during some initial stage, when the size of the passive tracer cloud is still small enough. Suppose that, initially, the passive tracer is concentrated in some 'small' region of the $\boldsymbol{r}$-space, and, respectively, the function $N_{\boldsymbol{k}}$ has a 'wide' support in the $\boldsymbol{k}$-space. With time, the passive tracer cloud spreads in the $\boldsymbol{r}$-space, and, respectively, the region of essentially non-zero function $N_{\boldsymbol{k}}$ shrinks in the $\boldsymbol{k}$-space. However, according to (6.1), the value of $N_{\boldsymbol{k}}$ at the origin remains constant ($\dot{N}_{\boldsymbol{k}} = 0$ if $\boldsymbol{k} = 0$). So, the $\boldsymbol{k}$-derivatives of the function $N_{\boldsymbol{k}}$ become large, and the differential approximation can break down (see more accurate estimates in §§ 5.2.6 and 5.2.7).

Let us estimate the time during which the differential approximation (6.4) is applicable.

If ρ is the characteristic size of the passive tracer cloud in the $\boldsymbol{r}$-space, then $1/\rho$ is the characteristic size of the passive tracer cloud in the $\boldsymbol{k}$-space; more precisely, the function $N_{\boldsymbol{k}}$ changes from its constant value $N_0 = 1/(2\pi)^{2d}$ at the origin $\boldsymbol{k} = 0$ to essentially zero value at distance $1/\rho$ from the origin. The differential approximation requires that $k_4 = |\boldsymbol{k}_1 - \boldsymbol{k}_3|$ is small compared with $1/\rho$. The typical value of k_4 is obtained from the condition $\Omega_4 t = 1$, i.e. $k_4 = (1/At)^{1/\alpha}$. Thus, the differential approximation is applicable if

$$\left(\frac{1}{\mathscr{A}t}\right)^{1/\alpha} \ll \frac{1}{\rho} \quad \text{where} \quad \rho = \rho_0\exp\{3\mathscr{B}(t/\tau)^{\beta}\},$$

(parameters β and $\mathscr{B}$ are given in (6.5)). The smaller the initial diameter of the passive tracer cloud ρ_0, the longer the differential approximation is applicable. We can rewrite this estimate in the form

$$\frac{x}{h} \gg e^x, \quad \text{where} \quad x = 3\alpha\beta\mathscr{B}\left(\frac{t}{\tau}\right)^{\beta}, \quad h = 3\alpha\beta\mathscr{B}\left(\frac{\rho_0}{\xi}\right)^{\alpha\beta},$$

where x is a new variable, and h is a new parameter; they are both non-dimensional. For sufficiently small h, this condition defines an interval of possible x. This interval shrinks to a point $x = 1$ when $h = e^{-1}$. If $h > e^{-1}$, it is impossible to satisfy this condition, and the differential approximation does not take place at all. If $h \ll e^{-1}$ (i.e. $\rho_0 \ll \xi$), the estimate gives

$$h \ll x \ll \log(1/h).$$

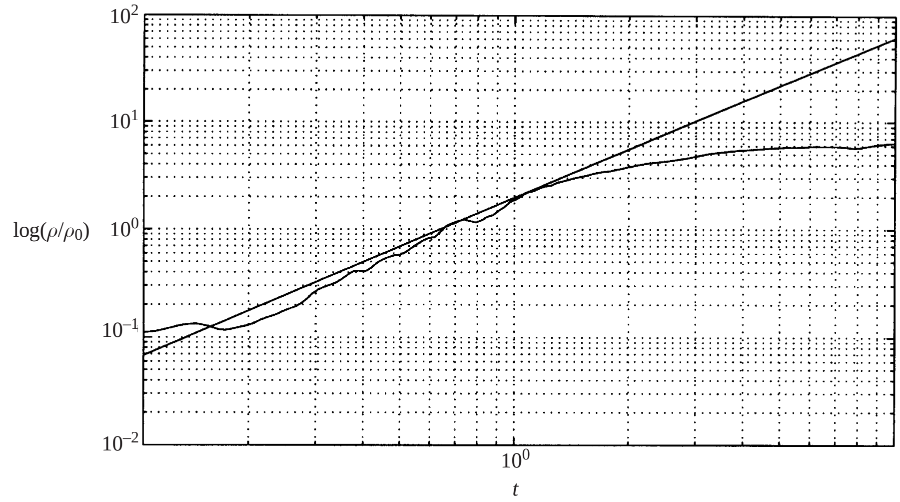

FIGURE 6. The super-exponential divergence of 'close' passive tracer particles. In non-dimensional units (when $\xi = \tau = 1$), $\Omega_k = k^2$ and $\varepsilon_k = k^{-3}$; so, according to the theoretical prediction, $\rho = \rho_0 \exp(2t^{3/2})$. The curve shows the numerical simulation, while the straight line represents the theoretical prediction. To obtain these results, we averaged over 10 realizations of the velocity field.

In terms of the original time t this means

$$\left(\frac{\rho_0}{\xi}\right)^{\alpha} \ll \frac{t}{\tau} \ll \left[\frac{1}{3\mathscr{B}} \log \frac{\xi}{\rho_0} - \frac{\log(3\alpha\beta\mathscr{B})}{3\alpha\beta\mathscr{B}}\right]^{1/\beta}. \tag{6.11}$$

The above estimate assumes that the passive tracer cloud is roughly isotropic; if a strong anisotropy is present, the differential approximation can break down even earlier.

6.3.2. *Saturation of the initial 'fast' divergence*

The initial regime of (sub- or super-) exponential divergence saturates for sufficiently large time. This regime can last till time t defined by the upper estimate in (6.11). At this time the average distance (6.9) between two tracer particles reaches the value

$$\bar{\rho} = \xi(3\alpha\beta\mathscr{B})^{-1/\alpha\beta}.$$

Thus, at the end of the initial regime the average distance between two tracer particles does not depend on the initial distance ρ_0; it is roughly equal to the scale ξ. If the initial distance ρ_0 exceeds the scale ξ, then the initial regime of 'fast' divergence and the differential approximation do not take place at all.

6.3.3. *Numerical simulation*

In order to check the prediction (6.9), we computed the evolution of two particles, initially located at points $(-\frac{1}{2}s, 0)$ and $(\frac{1}{2}s, 0)$, $s = 0.03$, see the first example in § 6.2.1. The results are presented in figure 6 in the log–log scale $\log(\rho/\rho_0)$ *vs.* time t.

According to the theory, the super-exponential divergence should saturate when $\rho \approx \xi = 1$, i.e. $\log(\rho/\rho_0) = \log(1/(s/\sqrt{2})) \approx 4$ (which agrees with figure 6). During the regime when the differential approximation is still applicable, the distance between the particles increases about 20 times.

Let us stress that the theory gives not only the slope of the straight line on figure 6, but also its location. In other words, the theory gives not only the exponent β, but also the pre-factor $\mathscr{B}$.

We simulated the velocity field following the procedure of modelling of Gaussian processes described in Kovalenko, Kuznetsov & Shurenkov (1996). We divide the domain

$$K_{min} \leqslant k \leqslant K_{max}, \quad 0 \leqslant \phi \leqslant 2\pi,$$

into many grid-cells Δ_{mn} in the $\boldsymbol{k}$-plane:

$$k_m \leqslant k \leqslant k_{m+1} \ (m = 1, 2, \ldots, \bar{m}), \quad \phi_n \leqslant \phi \leqslant \phi_{n+1} \ (n = 1, 2, \ldots, \bar{n}).$$

The grid should be sufficiently dense, i.e. the numbers $\bar{m}$ and $\bar{n}$ should be sufficiently large, to ensure that the discrete representation approximates the continuous Gaussian velocity field well. (In our calculation we took $\bar{m} = 300$, $\bar{n} = 200$.) In each of the cells we choose some 'central' value of the wave vector $\boldsymbol{K}_{mn} = (K_{mn}, \Phi_{mn}) \in \Delta_{mn}$. Then, the velocity field is approximated by the sum

$$\boldsymbol{v}(x, y, t) = \sum_{mn} \zeta_{mn} \boldsymbol{b}_{mn} \exp\{\mathrm{i}p_{mn}x + \mathrm{i}q_{mn}y - \mathrm{i}\Omega_{mn}t\}, \tag{6.12}$$

where

$$p_{mn} = K_{mn} \cos \Phi_{mn}, \quad q_{mn} = K_{mn} \sin \Phi_{mn}, \quad \Omega_{mn} = \Omega(\boldsymbol{K}_{mn}), \quad \boldsymbol{b}_{mn} = \boldsymbol{b}(\boldsymbol{K}_{mn})$$

and ζ_{mn} are independent Gaussian variables with zero mean and variance

$$\langle |\zeta_{mn}|^2 \rangle = I_{mn} = \int_{\Delta_{mn}} E_{\boldsymbol{k}} \, \mathrm{d}\boldsymbol{k}.$$

To generate the complex amplitude ζ_{mn}, we use the MATLAB random number generator to find (i) the random 'phase' θ_{mn} uniformly distributed in the interval $(0, 2\pi)$ and (ii) the random (real positive) 'amplitude' A_{mn} with distribution function $1 - \exp(-x^2/I_{mn})$; $\zeta_{mn} = A_{mn} \exp(\mathrm{i}\theta_{mn})$.

Once we have a realization, (6.12), of the velocity field, we use the standard MATLAB o.d.e.-solver to solve the system of two ordinary differential equations

$$\dot{x} = \boldsymbol{v}(x, y, t),$$
$$\dot{y} = \boldsymbol{v}(x, y, t),$$

with two different initial conditions: $(x(0), y(0)) = (-\tfrac{1}{2}s, 0)$ and $(x(0), y(0)) = (\tfrac{1}{2}s, 0)$.

The time step is chosen automatically by MATLAB to ensure relative accuracy 10^{-3} and absolute accuracy 10^{-6}. Then we repeat this calculation with another realization of the velocity field. Finally, we average over 10 realizations.

7. 'Large' passive tracer cloud: self-similar spreading

In this section we consider 'large' passive tracer clouds, when the 'support' of the first function in (5.7) is much smaller than the 'support' of the second: $1/\rho(t) \ll (1/\mathscr{A}t)^{1/\alpha}$; this situation is shown schematically in figure 7.

7.1. *Splitting the integration domain*

We can introduce some intermediate scale $\mathscr{K}(t)$ such that

$$\frac{1}{\rho(t)} \ll \mathscr{K}(t) \ll \frac{1}{(\mathscr{A}t)^{1/\alpha}}$$

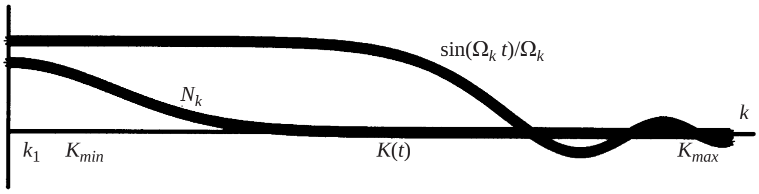

FIGURE 7. The situation where the 'support' of N_k is much smaller than the 'support' of $\sin(\Omega_k t)/\Omega_k$.

(figure 7) and split integration in (5.2) over $\boldsymbol{k}_4$ into two regions: $|\boldsymbol{k}_4| < \mathscr{K}(t)$ and $|\boldsymbol{k}_4| > \mathscr{K}(t)$. In the first region we can replace $\sin \Omega_4 t/\Omega_4$ by its limit at $k_4 \to 0$, which is equal to t. Since in order to find the mean relative displacement we require only the derivatives of $N_{\boldsymbol{k}}$ at $\boldsymbol{k} = 0$ (see (4.4)), we assume that in equation (5.2) the vector $\boldsymbol{k}_1$ is 'close to the origin' (k_1 is smaller than any scale). Then, in the second region we have $|\boldsymbol{k}_3| \approx |\boldsymbol{k}_4| > \mathscr{K}(t) \gg 1/\rho(t)$, and therefore, we can replace N_3 by zero. In the result, (5.2) takes the form

$$\dot{N}_1 = 2t \int_{|\boldsymbol{k}_4|<\mathscr{K}(t)} |\boldsymbol{k}_1 \cdot \boldsymbol{b}_4|^2 (N_3 - N_1)\varepsilon_4 \delta(\boldsymbol{k}_1 - \boldsymbol{k}_3 - \boldsymbol{k}_4)\, \mathrm{d}_{34}$$
$$-2N_1 \int_{|\boldsymbol{k}_4|>\mathscr{K}(t)} |\boldsymbol{k}_1 \cdot \boldsymbol{b}_4|^2 \frac{\sin \Omega_4 t}{\Omega_4} \varepsilon_4\, \mathrm{d}_4. \qquad (7.1)$$

We will see that the final result is independent of $\mathscr{K}(t)$.

Recalling the formula (4.4) for the tensor $\boldsymbol{S}$ of relative mean square displacement, we differentiate equation (7.1) with respect to $\boldsymbol{k}_1$ and then assume $\boldsymbol{k}_1 = 0 \Rightarrow \boldsymbol{k}_3 = \boldsymbol{k}_4$

$$\tfrac{1}{4}\dot{\boldsymbol{S}} = t \int_{|\boldsymbol{k}_4|<\mathscr{K}(t)} \boldsymbol{B}_4 [1 - (2\pi)^{2d} N_4] \varepsilon_4\, \mathrm{d}\boldsymbol{k}_4 - \int_{|\boldsymbol{k}_4|>\mathscr{K}(t)} \boldsymbol{B}_4 \frac{\sin \Omega_4 t}{\Omega_4} \varepsilon_4\, \mathrm{d}\boldsymbol{k}_4. \qquad (7.2)$$

7.2. *Assuming isotropy and self-similarity*

Now we make three isotropy assumptions:

(i) The medium is isotropic; this means that the dispersion law depends only on the wavenumber $|\boldsymbol{k}| = k$, $\Omega_{\boldsymbol{k}} = \Omega(k)$, and the integral of the polarization matrix $\boldsymbol{B}_{\boldsymbol{k}}$ over angular variables ζ of $\boldsymbol{k}$ is proportional to the identity matrix $\boldsymbol{I}$

$$\int \boldsymbol{B}_{\boldsymbol{k}} \mathrm{d}\zeta = \boldsymbol{I}\frac{1}{d} \int \mathrm{d}\zeta = \boldsymbol{I}\frac{1}{d} \times \begin{cases} 2\pi & \text{if } d = 2, \\ 4\pi & \text{if } d = 3 \end{cases} \qquad (\mathrm{d}\boldsymbol{k} = k^{d-1}\, \mathrm{d}k\, \mathrm{d}\zeta).$$

In the two-dimensional situation, which we consider below, the polarization vector $\boldsymbol{b}_{\boldsymbol{k}}$ is given by the formula (3.12) where the parameter θ can depend on k; and so

$$\int_0^{2\pi} \boldsymbol{B}_4\, \mathrm{d}\phi_4 = \pi \boldsymbol{I}.$$

(ii) The velocity field is statistically isotropic, i.e. $\varepsilon_{\boldsymbol{k}} = \varepsilon(k)$.

(iii) The distribution of the passive tracer is isotropic, i.e. $N_{\boldsymbol{k}}(t) = N(k,t)$, and therefore,

$$\boldsymbol{S} = \frac{\rho^2}{d}\boldsymbol{I}.$$

Property (iii) is implied by (i) and (ii) if the initial condition is isotropic, $N_{\boldsymbol{k}}(0) = N(k,0)$.

Integrating over $d\phi_4$ in (7.2), we find

$$\frac{1}{8\pi}\frac{d}{dt}\rho^2 = t\int_0^{\mathscr{K}(t)}[1-(2\pi)^{2d}N_4]\varepsilon_4 k_4^d\frac{dk_4}{k_4} + \int_{\mathscr{K}(t)}^{\infty}\frac{\sin\Omega_4 t}{\Omega_4}\varepsilon_4 k_4^d\frac{dk_4}{k_4}. \tag{7.3}$$

In other dimensions $d \neq 2$ we would have the same equation, but with a different factor only in the left-hand side, e.g. instead of 8π in two dimensions we would have 16π in three dimensions. Though we consider the two dimensional case, we keep general d in (7.3), because this allows us to obtain the converge conditions for an arbitrary d.

Let us further assume that the dispersion law and the energy spectrum are power-law functions of the wavenumber, $\Omega_k = \mathscr{A}k^\alpha$, $\varepsilon_k = \mathscr{C}k^{-\gamma}$ (in the inertial range), and the function $N_{\boldsymbol{k}}(t)$ has a self-similar form

$$N_{\boldsymbol{k}}(t) = N(p) \quad \text{where} \quad p = k\rho(t). \tag{7.4}$$

Now we transform integrals in (7.2); introducing new integration variables

$\boldsymbol{p} = \boldsymbol{k}_4\rho(t)$ (for the first integral), and $y = \mathscr{A}k_4^\alpha t$ (for the second integral),

we have

$$\frac{\rho\dot\rho}{4\pi\mathscr{C}} = t\rho^{\gamma-d}\int_0^{\mathscr{K}\rho}\frac{1-(2\pi)^{2d}N_p}{p^{\gamma-d}}\frac{dp}{p} + \frac{1}{\alpha}t(\mathscr{A}t)^{(\gamma-d)/\alpha}\int_{\mathscr{A}\mathscr{K}^\alpha t}^{\infty}\frac{\sin y}{y^{1+(\gamma-d)/\alpha}}\frac{dy}{y}.$$

Since $\mathscr{K}\rho \gg 1$ and $\mathscr{A}\mathscr{K}^\alpha t \ll 1$ (figure 7), we can rewrite these integrals in the form

$$\int_0^{\mathscr{K}\rho}\frac{1-(2\pi)^{2d}N_p}{p^{\gamma-d}}\frac{dp}{p} = \int_0^{L}\frac{1-(2\pi)^{2d}N_p}{p^{\gamma-d}}\frac{dp}{p} + \int_L^{\mathscr{K}\rho}p^{d-\gamma}\frac{dp}{p},$$

$$\int_{\mathscr{A}\mathscr{K}^\alpha t}^{\infty}\frac{\sin y}{y^{1+(\gamma-d)/\alpha}}\frac{dy}{y} = \int_l^{\infty}\frac{\sin y}{y^{1+(\gamma-d)/\alpha}}\frac{dy}{y} + \int_{\mathscr{A}\mathscr{K}^\alpha t}^{l}y^{(d-\gamma)/\alpha}\frac{dy}{y}.$$

where L is sufficiently large that N_p is practically zero for $p > L$, and l is sufficiently small number such that $\sin y$ is practically y for $y < l$. So

$$\frac{\rho\dot\rho}{4\pi\mathscr{C}} = t\rho^{\gamma-d}\left\{\int_0^{L}\frac{1-(2\pi)^{2d}N_p}{p^{\gamma-d}}\frac{dp}{p} + \frac{1}{d-\gamma}\left[(K\rho)^{d-\gamma} - L^{d-\gamma}\right]\right\}$$
$$+\frac{1}{\alpha}t(\mathscr{A}t)^{(\gamma-d)/\alpha}\left\{\int_l^{\infty}\frac{\sin y}{y^{1+(\gamma-d)/\alpha}}\frac{dy}{y} + \frac{\alpha}{d-\gamma}\left[l^{(d-\gamma)/\alpha} - (\mathscr{A}\mathscr{K}^\alpha t)^{(d-\gamma)/\alpha}\right]\right\}$$

We see that the scale $\mathscr{K}(t)$ drops out and we find the following equation

$$\rho\dot\rho = a\mathscr{C}t\rho^{\gamma-d} + b\mathscr{C}t(\mathscr{A}t)^{(\gamma-d)/\alpha} \tag{7.5}$$

where

$$a = 4\pi\lim_{L\to\infty}\left[\int_0^L\frac{1-(2\pi)^{2d}N_p}{p^{\gamma-d}}\frac{dp}{p} - \frac{L^{d-\gamma}}{d-\gamma}\right],$$

$$b = 4\pi\lim_{l\to 0}\left[\frac{1}{\alpha}\int_l^{\infty}\frac{\sin y}{y^{1+(\gamma-d)/\alpha}}\frac{dy}{y} + \frac{l^{(d-\gamma)/\alpha}}{d-\gamma}\right].$$

The factors a and b are dimensionless. The extra terms in these formulae (the finite terms, without integrals) 'help the integrals to converge', or, more precisely, they widen the region of the exponent γ for which the limits are finite. The factor a is finite if $\gamma < d+2$ (provided $N_p \to 0$ sufficiently fast as $p \to 0$). The factor b is finite if $d - 2\alpha < \gamma < d + 2\alpha$. Thus, the condition

$$d - 2\alpha < \gamma < \min(d + 2, d + 2\alpha)$$

is necessary for the applicability of equation (7.5).

In non-dimensional form, when length and time are measured in units ξ and τ, respectively, equation (7.5) is

$$\rho\dot{\rho} = at\rho^{\gamma-d} + bt^{1+(\gamma-d)/\alpha} \tag{7.6}$$

(see also § 1.3.4).

If $\gamma < d$, then

$$a = -4\pi \int_0^\infty \frac{(2\pi)^{2d} N_p}{p^{\gamma-d}} \frac{\mathrm{d}p}{p} < 0, \quad b = \frac{4\pi}{\alpha} \int_0^\infty \frac{\sin y}{y^\lambda} \mathrm{d}y = \lambda \frac{4\pi I_1(\lambda)}{\alpha} > 0;$$

the exponent λ and the integral $I_1(\lambda)$ are defined in (3.9); to establish the last equality, we integrated by parts.

If $\gamma > d$, then

$$a = 4\pi \int_0^\infty \frac{1 - (2\pi)^{2d} N_p}{p^{\gamma-d}} \frac{\mathrm{d}p}{p} > 0, \quad b = -\frac{4\pi}{\alpha} \int_0^\infty \frac{y - \sin y}{y^\lambda} \mathrm{d}y < 0;$$

here, the first inequality follows from the properties (iv), (v) in § 4.3.

7.3. *Evolution of the diameter of the passive tracer cloud*

If $\gamma < d$, then initially (for small t) the second term on the right-hand side of (7.6) dominates the first one; let us find the condition when this dominance is preserved as $t \to \infty$:

$$bt^{1+(\gamma-d)/\alpha} \gg |a| t\rho^{\gamma-d} \Rightarrow \frac{\rho}{t^{1/\alpha}} \gg \left(\frac{b}{|a|}\right)^{1/(\gamma-d)}. \tag{7.7}$$

If we neglect the first term on the right-hand side of (7.6), then

$$\rho\dot{\rho} = bt^{1+(\gamma-d)/\alpha} \Leftrightarrow \rho^2 - \rho_0^2 = \frac{2b}{\lambda} t^\lambda = 2\frac{4\pi I_1(\lambda)}{\alpha} t^\lambda. \tag{7.8}$$

The dominance (7.7) holds as $t \to \infty$ if $\frac{1}{2}\lambda > 1/\alpha \Leftrightarrow \delta < 0$; the latter inequality was assumed in the beginning, § 1.2. When $\gamma < d$, the mean square displacement σ^2 is finite, and equation (7.8) means that $\rho^2(t) - \rho_0^2 = 2\sigma^2(t)$ (compare (7.8) with (3.19)). This $\rho(t)$ corresponds to the average distance between two uncorrelated particles. Since the factor a is negative, the correlation slows down the growth of the relative displacement. On the other hand, if $\gamma > d$, then $a > 0$ and $b < 0$, which means that the relative displacement between two particles grows owing to the correlation.

If $\gamma > d$, then initially (for small t) the first term on the right-hand side of (7.6) dominates the second one; let us find the condition when this dominance is preserved as $t \to \infty$:

$$at\rho^{\gamma-d} \gg |b| t^{1+(\gamma-d)/\alpha} \Rightarrow \frac{\rho}{t^{1/\alpha}} \gg \left(\frac{b}{|a|}\right)^{1/(\gamma-d)}. \tag{7.9}$$

If we neglect the second term on the right-hand side of (7.6), then

$$\rho\dot{\rho} = at\rho^{\gamma-d} \Rightarrow \rho^{2/\nu} - \rho_0^{2/\nu} = \frac{at^2}{\nu} \quad \text{where} \quad \nu = \frac{2}{2 + d - \gamma} > 0. \tag{7.10}$$

As $t \to \infty$, $\rho(t)$ grows like t^{ν}, and the dominance (7.9) holds if $\nu > 1/\alpha$, which is again equivalent to the condition $\delta < 0$ assumed in § 1.2.

Thus,

$$\text{as } t \to \infty, \quad \rho(t) \sim \begin{cases} \sqrt{2b/\lambda}\, t^{\lambda/2} & \text{if} \quad \gamma < d, \\ (a/\nu)^{\nu/2} t^{\nu} & \text{if} \quad \gamma > d. \end{cases} \tag{7.11}$$

Let us stress, when we say $t \to \infty$, we mean $t \gg \tau$ (in non-dimensional form $\tau = 1$), but still $t \ll T_{max}$. While the constant b is easily calculated (see (7.8)), the calculation of the constant a requires the solution of the integral-differential equation for the function N_k.

The derivation of the diameter equation (7.6) requires that the passive tracer cloud be large enough, i.e. the 'support' of the first function in (5.7) is much smaller than the 'support' of the second: $1/\rho(t) \ll 1/t^{1/\alpha}$; see figure 7. Will this condition hold as $t \to \infty$ (when the initial diameter ρ_0 becomes negligible)? According to (7.11), the answer is 'Yes', provided $\delta < 0$, the condition assumed in § 1.2.

8. Conclusion

In this paper we have answered the two physical questions formulated in § 1.1.

The developed approach can be applied to several other problems. In particular, we intend to apply it to the problem of wave propagation through a random medium when there is a wide range of heterogeneity scales. The wavelength of the propagating signal is well inside this range. Such a problem arises, for example, in geophysics (see Sato & Fehler 1998).

We also intend to apply the developed approach to a few problems of wave turbulence in which the wave kinetic equation does not work.

The rigorous mathematical justification of this approach is interesting because the above predictions are observed only statistically, on average over many realizations; the individual trajectories of the tracer particles significantly deviate from each other.

I wish to thank R. Glazman for useful discussions. I am grateful to P. Milewski, L. Smith, E. Tabak and F. Waleffe, the organizers of the Conference on Dispersive Wave Turbulence (Massachusetts, June 2000), for the invitation to present my research at the Conference and for their support. I also wish to thank G. Papanicolaou for the invitation to participate in the Geophysical Mathematics Summer School (Stanford University, August 2000) and for the support provided. This work was supported by Grant DMS-9971640 of the National Science Foundation.

REFERENCES

Avellaneda, M. & Majda, A. J. 1990 Mathematical models with exact renormalization for turbulent transport. *Commun. Pure Appl. Maths* **131**, 381–429.

Avellaneda, M. & Majda, A. J. 1992 Mathematical models with exact renormalization for turbulent transport, II: fractal interfaces, non-Gaussian statistics and the sweeping effect. *Commun. Pure Appl. Maths* **146**, 139–204.

Balk, A. M. 2001*a* Anomalous diffusion of a tracer advected by wave turbulence. *Phys. Lett.* A **279**, 370–378.

Balk, A. M. 2001*b* Anomalous transport by wave turbulence. In *Advances in Wave Interaction and Turbulence* (ed. P. Milewski, L. Smith, E. Tabak & F. Waleffe), *Contemporary Mathematics*, vol. 283, pp.13–25. American Mathematical Society, Providence, RI.

Balk, A. M. & McLaughlin, R. M. 1999 Passive scalar in a random wave field: the weak turbulence approach. *Phys. Lett.* A **256**, 299–306.

Batchelor, G. K. & Townsend, A. A. 1956 Turbulent diffusion. In *Surveys in Mechanics* (ed. G. K. Batchelor & H. Bondi), pp. 352–399. Cambridge University Press.

Bensoussan, A., Lions, J. L. & Papanicolaou, G. 1978 *Asymptotic Analysis of Periodic Structures*. North-Holland.

Bogoliubov, N. N. & Mitropolsky, Y. A. 1961 *Asymptotic Methods in the Theory of Nonlinear Oscillations*. Gordon and Breach.

Cherkaev, A. 2000 *Variational Methods for Structural Optimization*. Springer.

Eftimiu, C. 1990 Wave propagation in random media: Wiener–Hermite expansion approach. *J. Electromag. Waves Applics* **4**, 847–864.

Gill, A. E. 1982 *Atmosphere–Ocean Dynamics*. Academic.

Herterich, K. & Hasselmann, K. 1982 The horizontal diffusion of tracers by surface waves. *J. Phys. Oceanogr.* **12**, 704–711.

Kovalenko, I. N., Kuznetsov, N. Yu. & Shurenkov, V. M. 1996 *Models of Random Processes: a Handbook for Mathematicians and Engineers*. CRC Press.

Kraichnan, R. H. 1968 Small-scale structure of a scalar field convected by turbulence. *Phys. Fluids* **11**, 945–953.

Majda, A. J. & Kramer, P. R. 1999 Simplified models for turbulent diffusion: theory, numerical modeling, and physical phenomena. *Phys. Rep.* **314**, 237–574.

Milton, G. W. 2000 *Effective Properties of Composites*. Cambridge University Press.

Ryzhik, L., Papanicolaou, G. & Keller, J. B. 1996 Transport equations for elastic and other waves in random media. *Wave Motion* **24**, 327–370.

Sanders, J. A. & Verhulst, F. 1985 *Averaging Methods in Nonlinear Dynamical Systems*. Springer.

Sato, H. & Fehler, C. F. 1998 *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*. Springer.

Sheng, P. 1995 *Introduction to Wave Scattering, Localization, and Mesoscopic Phenomena*. Academic.

Sobczyk, K. & Kirkner, D. J. 2001 *Stochastic Modeling of Microstructures*. Birkhäuser, Boston.

Weichman, P. B. & Glazman, R. E. 1999 Turbulent fluctuation and transport of passive scalars by random wave fields. *Phys. Rev. Lett.* **83**, 5011–5014.

Weichman, P. B. & Glazman, R. E. 2000 Passive scalar transport by travelling wave fields. *J. Fluid Mech.* **420**, 147–200.

Weichman, P. B. & Glazman, R. E. 2002 Spatial variations of a passive tracer in a random wave field. *J. Fluid Mech.* **453**, 263–287.

Wiener, N. 1958 *Nonlinear Problems in Random Theory*. MIT Press.

Zakharov, V. E. 1974 Hamiltonian formalism in the theory of waves in nonlinear media with dispersion. *Izv. VUZov, Radiofizika* **17**, 431–453.

Zakharov, V. E. 1985 Kolmogorov spectra in problems of weak turbulence. In *Basic Plasma Physics*, vol. 2 (ed. A. A. Galeev & R. N. Sudan), pp. 3–36. Elsevier.

Zakharov, V. E., Musher, S. L. & Rubenchik, A. M. 1985 Hamiltonian approach to the description of nonlinear plasma phenomena. *Phys. Rep.* **129**, 285–366.