Normal and Anomalous Diffusion: A Tutorial

Loukas Vlahos, Heinz Isliker

Department of Physics, University of Thessaloniki, 54124 Thessaloniki, Greece

Yannis Kominis, and Kyriakos Hizanidis

School of Electrical and Computer Engineering, National Technical University of Athens, 15773 Zografou, Athens, Greece

Abstract

The purpose of this tutorial is to introduce the main concepts behind normal and anomalous diffusion. Starting from simple, but well known experiments, a series of mathematical modeling tools are introduced, and the relation between them is made clear. First, we show how Brownian motion can be understood in terms of a simple random walk model. Normal diffusion is then treated (i) through formalizing the random walk model and deriving a classical diffusion equation, (ii) by using Fick's law that leads again to the same diffusion equation, and (iii) by using a stochastic differential equation for the particle dynamics (the Langevin equation), which allows to determine the mean square displacement of particles. (iv) We discuss normal diffusion from the point of view of probability theory, applying the Central Limit Theorem to the random walk problem, and (v) we introduce the more general Fokker-Planck equation for diffusion that includes also advection. We turn then to anomalous diffusion, discussing first its formal characteristics, and proceeding to Continuous Time Random Walk (CTRW) as a model for anomalous diffusion. It is shown how CTRW can be treated formally, the importance of probability distributions of the Levy type is explained, and we discuss the relation of CTRW to fractional diffusion equations and show how the latter can be derived from the CTRW equations. Last, we demonstrate how a general diffusion equation can be derived for Hamiltonian systems, and we conclude this tutorial with a few recent applications of the above theories in laboratory and astrophysical plasmas.

Key words: Random Walk, Normal Diffusion, Anomalous Diffusion, Continuous Time Random Walk, Diffusion Equation

[1] In the memory of our friend and colleague Simos Ichtiaroglou

1 Introduction

The art of doing research in physics usually starts with the observation of a natural phenomenon. Then follows a qualitative idea on "How the phenomenon can be interpreted", and one proceeds with the construction of a model equation or a simulation, with the aim that it resembles very well the observed phenomenon. This progression from natural phenomena to models and mathematical prototypes and then back to many similar natural phenomena, is the methodological beauty of our research in physics.

Diffusion belongs to this class of phenomena. All started from the observations of several scientists on the irregular motion of dust, coal or pollen inside the air or a fluid. The roman Lucretius in his poem on the Nature of Things (60 BC) described with amazing details the motion of dust in the air, Jan Ingenhousz described the irregular motion of coal dust on the surface of alcohol in 1785, but Brownian motion is regarded as the discovery of the botanist Robert Brown in 1827, who observed pollen grains executing a jittery motion in a fluid. Brown initially thought that the pollen particles were "alive", but repeating the experiment with dust confirmed the idea that the jittery motion of the pollen grains was due to the irregular motion of the fluid particles.

The mathematics behind "Brownian motion" was first described by Thiele (1880), and then by Louis Bachelier in 1900 in his PhD thesis on "the theory of speculation", in which he presented a stochastic analysis of the stock and option market. Albert Einstein's independent research in 1905 brought to the attention of the physicists the main mathematical concepts behind Brownian motion and indirectly confirmed the existence of molecules and atoms (at that time the atomic nature of matter was still a controversial idea). As we will see below, the mathematical prototype behind Brownian motion became a very useful tool for the analysis of many natural phenomena.

Several articles and experiments followed Einstein's and Marian Smoluchowski's work and confirmed that the molecules of water move randomly, therefore a small particle suspended in the fluid experiences a random number of impacts of random strength and direction in any short time. So, after Brown's observations of the irregular motion of "pollen grains executing a jittery motion", and the idea of how to interpret it as "the random motion of particles suspended inside the fluid", the next step is to put all this together in a firm mathematical model, "the continuous time stochastic process." The end result is a convenient prototype for many phenomena, and today's research on "Brownian motion" is used widely for the interpretation of many phenomena.

This tutorial is organized as follows: In Sec. 2, we give an introduction to Brownian motion and classical random walk. Sec. 3 presents different models for classical diffusion, the Langevin equation, the approach through Fick's law, Einstein's approach, the Fokker-Planck equation, and the central limit theorem. In Sec. 4, the characteristics of anomalous diffusion are described, and a typical example, the rotating annulus, is presented. Sec. 5 introduces Continuous Time Random Walk, the waiting and the velocity model are explained, methods to solve the equations are discussed, and also the Levy distributions are introduced. In Sec. 6, it is shown how, starting from random walk models, fractional diffusion equations can be constructed. In Sec. 7 we show how a quasi-linear diffusion equation can be derived for Hamiltonian systems. Sec. 8 briefly comments on alternative ways to deal with anomalous diffusion, Sec. 9 contains applications to physics and astrophysics, and Sec. 10 presents the conclusions.

2 Brownian Motion and Random walks

2.1 Brownian Motion Interpreted as a Classical Random Walk

To build a firm base for the stochastic processes involved in Brownian motion, we may start with a very simple example.

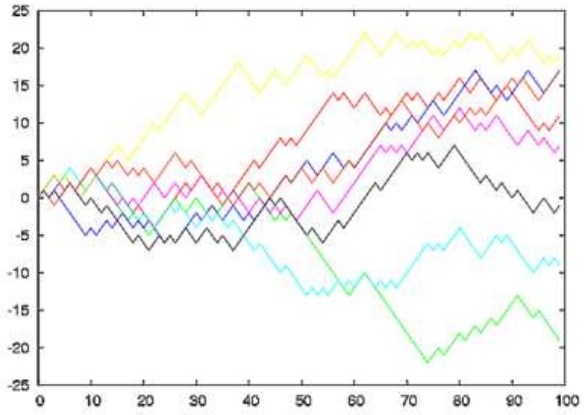

Fig. 1. Random walk in one dimension (along the vertical axis) as a function of time (to the right on the horizontal axis).

We consider a random walk in one dimension (1D) and assume that the particles' steps Δz are random and equally likely to either side, left or right, and of constant length ℓ (see Fig. 1). The position z_N of a particle starting at

$z_0 = 0$ after N steps is

$$z_N = \Delta z_N + \Delta z_{N-1} + ... + \Delta z_1 = \sum_{i=1}^{N} \Delta z_i, \tag{1}$$

so that the squared length of the path equals

$$\begin{aligned} z_N^2 &= \left(\sum_{j=1}^{N} \Delta z_j \right) \left(\sum_{k=1}^{N} \Delta z_k \right) = \sum_{j,k=1}^{N} \Delta z_j \Delta z_k \\ &= \sum_{j=1,\, k=j}^{N} \Delta z_j^2 + \sum_{j,k=1,\, k\neq j}^{N} \Delta z_j \Delta z_k = N\ell^2 + \sum_{j,k=1,\, j\neq k}^{N} \Delta z_j \Delta z_k. \end{aligned} \tag{2}$$

When averaging over a large number of particles, we find the mean squared path length as

$$< z_N^2 >= N\ell^2 + \left\langle \sum_{j,k=1,\, j\neq k}^{N} \Delta z_j \Delta z_k \right\rangle. \tag{3}$$

Each step of the walk is equally likely to the left or to the right, so that the displacements Δz_i are random variables with zero mean. The products $\Delta z_j \Delta z_k$ are also random variables, and, since we assume that Δz_j and Δz_k are independent of each other, the mean value of the products is zero, so that the expectation value of the mixed term in Eq. (3) is zero. We thus find

$$< z^2 >= N\ell^2. \tag{4}$$

The root-mean square displacement after N steps of constant length ℓ (mean free path) is

$$R := \sqrt{< z_N^2 >} = \ell\sqrt{N}. \tag{5}$$

We can now estimate the number of steps a photon starting from the Sun's core needs to reach the surface of the Sun. From Eq. (5), we have $N = (R/\ell)^2$ and since the Sun's radius is $\sim 10^{10}$ cm and the characteristic step (taken into account the density in the solar interior) is ~ 1 cm, we conclude that photons make 10^{20} steps before exiting from the Sun's surface (this can answer questions like: if the Sun's core stops producing energy, how long will it take until we feel the difference on Earth?).

The mean free path ℓ can be estimated with a simple model. By assuming that a particle is moving inside a gas with a mean speed $< v >$, the distance

traveled between two successive collisions is $\ell =< v > \tau$, where τ is called collision time. If the particle has radius a and travels a distance L inside the gas with density n, then it will suffer $4\pi a^2 Ln$ collisions, which is just the number of particles in the volume $4\pi a^2 L$ the particle sweeps through. The mean free path is then defined through the relation $4\pi a^2 \ell n = 1$, i.e. ℓ is the distance to travel and to make just one collision, so that

$$\ell = \frac{1}{4\pi a^2 n}. \tag{6}$$

We may thus conclude that the number of steps a particle executes inside a gas during a time t is $N = t/\tau$, and, with Eq. (4) and the above relation $\ell =< v > \tau$, the mean squared distances it travels is

$$< z^2 >= N\ell^2 = (t/\tau)(< v > \tau)\ell = (< v > \ell)t. \tag{7}$$

Assuming that the random walk takes place in 3 dimensions and that the gas is in equilibrium and isotropic, we expect that $< x^2 >=< y^2 >=< z^2 >= \frac{<r^2>}{3}$, and the mean square path length in 3 dimensions is

$$< r^2 >= 3 < v > \ell t = Dt, \tag{8}$$

where $D := 3 < v > \ell$ is called the diffusion coefficient, which is a useful parameter to characterize particle diffusion in the normal case (see Sec. 3). Important here is to note the linear scaling relation between $< r^2 >$ and time t.

2.2 Formal Description of the Classical Random Walk

More formally, we can define the classical random walk problem as follows. We consider the position $\vec{r}$ of a particle in 1, 2, or 3 dimensional space, and we assume that the position changes in repeated random steps $\Delta\vec{r}$. The time Δt elapsing between two subsequent steps is assumed to be constant, time plays thus a dummy role, it actually is a simple counter. The position $\vec{r}_n$ of a particle after n steps, corresponding to time $t_n = n\Delta t$, is

$$\vec{r}_n = \Delta\vec{r}_n + \Delta\vec{r}_{n-1} + \Delta\vec{r}_{n-2} + \ ... \ + \Delta\vec{r}_1 + \vec{r}_0 \tag{9}$$

where $\vec{r}_0$ is the initial position, and $\Delta\vec{r}_i$ is the ith step (or increment, or displacement). The position $\vec{r}_n$ as well as the increments $\Delta\vec{r}_i$ are all random variables. To specify the problem completely, we have to prescribe the probability distribution $q_{\Delta\vec{r}}(\Delta\vec{r})$ for the increments $\Delta\vec{r}_i$, which yields the

probability for the particle to make a certain step $\Delta\vec{r}_i$ (with given length and direction). Writing $q_{\Delta\vec{r}}(\Delta\vec{r})$ in this way, we have made the assumptions that all the increments have the same probability distribution and that the increments are independent of each other (the value $\Delta\vec{r}_i$ takes in any realization is completely independent of the value taken in the previous step by $\Delta\vec{r}_{i-1}$). Generalizations to time-dependent increment distributions or correlated increments are of course possible.

Since $\vec{r}_n$ is a random variable, the solution we are looking for in the random walk problem is in the form of the probability distribution $P(\vec{r}, t_n)$, which yields the probability for a particle to be at position $\vec{r}$ at time $t = t_n \equiv n\Delta t$.

If we are interested in the mean square displacement, we can just square Eq. (9), and, rearranging the terms in the same way as in Eq. (2), we find for $\vec{r}_0 = 0$

$$\langle \vec{r}_n^{\,2} \rangle = \sum_{j=1,\,(k=j)}^{n} \langle \Delta\vec{r}_j^{\,2} \rangle + \sum_{j,k=1,\,k\neq j}^{n} \langle \Delta\vec{r}_j \Delta\vec{r}_k \rangle. \tag{10}$$

The first term on the right hand side is just a sum over the variances $\sigma^2_{\Delta\vec{r},\,j}$ of $q_{\Delta\vec{r}}(\Delta\vec{r})$, since by definition $\sigma^2_{\Delta\vec{r}} := \langle \Delta\vec{r}_j^{\,2} \rangle = \int \Delta\vec{r}^{\,2} q_{\Delta\vec{r}}(\Delta\vec{r})\, d\Delta\vec{r}$ if the mean value of $q_{\Delta\vec{r}}(\Delta\vec{r})$ is zero. The second term on the right hand side is the covariance $\mathrm{cov}(\Delta\vec{r}_j, \Delta\vec{r}_k)$ of the random walk steps, and it is zero if the steps a particle takes are independent of each other. We thus can write Eq. (10) as

$$\langle \vec{r}_n^{\,2} \rangle = \sum_{j=1,\,(k=j)}^{n} \sigma^2_{\Delta\vec{r},\,j} + \sum_{j,k=1,\,k\neq j}^{n} \mathrm{cov}(\Delta\vec{r}_j, \Delta\vec{r}_k). \tag{11}$$

The particular random walk we considered in Sect. 2.1 can thus be understood on the base of Eq. (11) as a case with zero covariance and variance $\sigma^2_{\Delta\vec{r},\,j} = \ell^2$, due to the constant step length, which leads to the mean square displacement in Eq. (4).

3 Models for Normal Diffusion

3.1 Langevin's Equation

We turn now to a different way of treating Brownian motion. We again consider a particle with mass m performing a random walk inside a fluid due to the bombardment by the fluid molecules, which obey an equilibrium

(Maxwellian) distribution. Pierre Langevin (Langevin, 1908) described this motion with a simple but very interesting stochastic differential equation (let us work in one dimension for simplicity),

$$m\ddot{x} = -a\dot{x} + F(t), \tag{12}$$

where the term $a\dot{x}$ represents the friction force, $\dot{x}$ is the particle velocity, a is the damping rate and depends on the radius of the particle and the viscosity of the fluid, and $F(t)$ is a random fluctuating force due to the random bombardment of the particle by the fluid molecules. If the random fluctuating force were absent, the particle starting with an initial velocity v_0 would gradually slow down due to the friction term. Multiplying Eq. (12) with x, we have

$$mx\ddot{x} = m\left[\frac{d(x\dot{x})}{dt} - \dot{x}^2\right] = -ax\dot{x} + xF(t),$$

and after taking averages over a large number of particles we find, since $< xF(t) >= 0$ due to the irregular nature of the force $F(t)$,

$$m\frac{d < x\dot{x} >}{dt} = m < \dot{x}^2 > -a < x\dot{x} > . \tag{13}$$

Since the background gas is is in equilibrium, the kinetic energy of the particle is proportional to the gas temperature, $m < \dot{x}^2 > /2 = kT/2$, where k is the Boltzmann constant and T the temperature of the gas. Eq. (13) now takes the form

$$\left(\frac{d}{dt} + \gamma\right) < x\dot{x} >= \frac{kT}{m},$$

where $\gamma = a/m$, which has the solution

$$< x\dot{x} >= \frac{1}{2}\frac{d < x^2 >}{dt} = Ce^{-\gamma t} + \frac{kT}{a}. \tag{14}$$

At $t = 0$, the mean square displacement is zero, so that $0 = C + kT/a$, and Eq. (14) becomes

$$\frac{1}{2}\frac{d < x^2 >}{dt} = \frac{kT}{a}(1 - e^{-\gamma t}).$$

On integrating the above equation we find the solution

$$< x^2 >= \frac{2kT}{a}\left[t - \frac{1}{\gamma}\left(1 - e^{-\gamma t}\right)\right]. \tag{15}$$

In the limit $t << 1/\gamma$ (time much shorter than the collision time) the solution in Eq. (15) is of the form $< x^2 >\sim t^2$ (expanding the exponential up to second order), which is called "ballistic" diffusion and means that at small times particles are not hindered by collisions yet and diffuse very fast, see Sect. 4.2. In the other limit, $t >> 1/\gamma$, the solution has the form

$$< x^2 >\sim \frac{2kT}{a}t, \tag{16}$$

or, for the 3 dimensional case, if again the gas is in equilibrium and isotropic so that $< r^2 > /3 =< x^2 >$,

$$< r^2 >= \frac{6kT}{a}t = Dt, \tag{17}$$

where $D = 6kT/a$ is an expression for the diffusion constant in terms of particle and fluid characteristics (cf. Eq. (8)), and note that again $< r^2 >$ has a simple scaling relation with time, $< r^2 >= Dt$, as in Sect. 2.1.

3.2 Modeling Diffusion with Fick's Law

Diffusion usually occurs if there is a spatial difference in concentration of e.g. particles or heat etc., and it usually acts such as to reduce the spatial inhomogeneities in concentration.

Let us consider particle diffusion along the z-direction in 3-dimensional space, and let us assume that two elementary areas perpendicular to the flow (in the x-y-plane) are a distance Δz apart. Particle conservation implies that the time variation of the density $n(z,t)$ inside the elementary volume $\Delta x \Delta y \Delta z$ equals the inflow minus the outflow of particles, so that, if $J(z,t)$ denotes the particle flux,

$$\frac{\partial n(z,t)}{\partial t}\Delta x \Delta y \Delta z = J(z)\Delta x \Delta y - J(z+\Delta z)\Delta x \Delta y = -\frac{\partial J}{\partial z}\Delta x \Delta y \Delta z$$

which leads to the diffusion equation in its general form,

$$\frac{\partial n(z,t)}{\partial t} = -\frac{\partial J(z,t)}{\partial z}. \tag{18}$$

The problem that remains is to determine the particle flux J. From its physical meaning, it obviously holds that

$$J(z,t) = n(z,t)v(z,t), \tag{19}$$

where $v(z,t)$ is an average particle flow velocity. Using this expression in Eq. (18) leads to a closure problem, we would need to find ways to determine $v(z,t)$.

It is well documented experimentally that the flux of particles J crossing a certain area (again, say in the x-y-plain) is proportional to the density gradient along the z-axis (Fick's Law),

$$J_z = -D(z)\frac{\partial n}{\partial z}, \tag{20}$$

where D is the diffusion coefficient discussed already in the previous sections, and which generally may also depend on z. With Eq. (20), the diffusion equation takes the classical form

$$\frac{\partial n(z,t)}{\partial t} = \frac{\partial}{\partial z}D(z)\frac{\partial n(z,t)}{\partial z}, \tag{21}$$

or, for constant diffusion coefficient,

$$\frac{\partial n(z,t)}{\partial t} = D\frac{\partial^2 n(z,t)}{\partial z^2}. \tag{22}$$

In infinite space, and if all particles start initially from $z = 0$, the solution of Eq. (22) is

$$n(z,t) = \frac{N_0}{\sqrt{4\pi Dt}}e^{-z^2/4Dt}, \tag{23}$$

where N_0 is the total number of particles inside the volume under consideration. The solution obviously is identical to a Gaussian distribution with mean zero and variance $2Dt$. The variance is defined as

$$< z^2(t) >= \int z^2 n(z,t)\, dz = 2Dt, \tag{24}$$

which is just identical to the mean square displacement, so that the results obtained earlier, using the simple version of the random walk in Sect. 2.1 or the stochastic differential equation of Langevin in Sect. 3.1, are again confirmed.

Diffusion obeying Eq. (24) is called **normal diffusion** and is characteristic for the diffusion processes in systems that are in equilibrium or very close to equilibrium. **Generalizing the above results (Eqs. (21), (22), (23), (24) is simple and can be found in the literature (see references).**

3.3 Einstein's Formalism for the Classical Random Walk and the Diffusion Equation

A different approach to treat normal diffusion was introduced by Bachelier and by Einstein (see Einstein, 1905). Here, the starting point is the classical random walk as defined in Sec. 2.2, and we consider the 1-dimensional case. According to Sec. 2.2, the solution of the random walk problem is in the form of the probability distribution $P(z,t)$ for a particle at time t to be at position z. Assume that we would know the distribution $P(z, t-\Delta t)$ one time-step Δt earlier (remember Δt is assumed constant). If particles are conserved, the relation

$$P(z,t) = P(z-\Delta z, t-\Delta t)\, q_{\Delta z}(\Delta z). \tag{25}$$

must hold, with $q_{\Delta z}$ the distribution of random walk steps. Eq. (25) states that the probability to be at time t at position z equals the probability to have been at position $z-\Delta z$ at time $t-\Delta t$, and to have made a step of length Δz. We still have to sum over all possible Δz, which leads to the Einstein (or Bachelier) diffusion equation,

$$P(z,t) = \int_{\infty}^{\infty} P(z-\Delta z, t-\Delta t)\, q_{\Delta z}(\Delta z)\, d\Delta z \tag{26}$$

This is an integral equation that determines the solution $P(z,t)$ of the random walk problem as defined in Sec. 2.2. The power of this equation will become clear below when we will show ways to treat cases of anomalous diffusion. Here, we still focus on normal diffusion. As will become clear later, it is actually a characteristic of normal diffusion that the particles take only small steps Δz compared to the system size. This implies that $q_{\Delta z}(\Delta z)$ is non-zero only for small Δz, the integral in Eq. (26) is only over a small Δz-range, and we can expand $P(z-\Delta z, t-\Delta t)$ in z and t (also Δt is small),

$$\begin{aligned} P(z-\Delta z, t-\Delta t) = {} & P(z,t) - \Delta t\, \partial_t P(z,t) - \Delta z\, \partial_z P(z,t) \\ & + \frac{1}{2}\Delta z^2\, \partial_z^2 P(z,t) \end{aligned} \tag{27}$$

Inserting into Eq. (26), we find

$$P(z,t) = \int P(z,t)\, q_{\Delta z}(\Delta z)\, d\Delta z - \int \Delta t \partial_t P(z,t)\, q_{\Delta z}(\Delta z)\, d\Delta z$$
$$- \int \Delta z \partial_z P(z,t)\, q_{\Delta z}(\Delta z)\, d\Delta z$$
$$+ \int \frac{1}{2} \Delta z^2\, \partial_z^2 P(z,t)\, q_{\Delta z}(\Delta z)\, d\Delta z \tag{28}$$

Obviously, $P(z,t)$, its derivatives, and Δt are not affected by the integration, furthermore we use the normalization of $q_{\Delta z}$ ($\int q_{\Delta z}(\Delta z)\, d\Delta z = 1$), we assume $q_{\Delta z}$ to be symmetric (mean value zero, $\int \Delta z\, q_{\Delta z}(\Delta z)\, d\Delta z = 0$), and we use the definition of the variance $\sigma^2_{\Delta z}$ ($\int \Delta z^2\, q_{\Delta z}(\Delta z)\, d\Delta z = \sigma^2_{\Delta z}$), so that

$$P(z,t) = P(z,t) - \Delta t\, \partial_t P(z,t) + \frac{1}{2} \sigma^2_{\Delta z}\, \partial_z^2 P(z,t) \tag{29}$$

or

$$\partial_t P z, t = \frac{\sigma^2_{\Delta z}}{2\Delta t} \partial_z^2 P(z,t) \tag{30}$$

i.e. we again recover the simple diffusion equation, as in Sec. 3.2, with diffusion coefficient $D = \frac{\sigma^2_{\Delta z}}{2\Delta t}$, and the solution to it in infinite space is again the Gaussian of Eq. (23).

3.4 Fokker-Planck Equation

The Fokker-Planck (FP) equation (or Kolmogorov forward equation) is a more general diffusion equation than the simple equations introduced in Secs. 3.2 an 3.3. We again start from a description of diffusion in terms of a random walk, as in Sec. 3.3, we though relax two assumptions made there: (i) We assume now that the mean value $\mu_{\Delta z}$ of the random walk steps can be different from zero, which corresponds to a systematic motion of the particles in the direction of the sign of $\mu_{\Delta z}$, and (ii) we assume that both the mean and the variance can be spatially dependent, $\mu_{\Delta z} = \mu_{\Delta z}(z)$ and $\sigma^2_{\Delta z} = \sigma^2_{\Delta z}(z)$, which means that the distribution of increments depends on the spatial location, i.e. it is of the form $q_{\Delta z,z}(\Delta z, z)$. To be compatible with these assumptions, Eq. (26) must be rewritten in a slightly more general form,

$$P(z,t) = \int_{\infty}^{\infty} P(z - \Delta z, t - \Delta t) q_{\Delta z,z}(\Delta z, z - \Delta z)\, d\Delta z, \tag{31}$$

which is the *Chapman-Kolmogorov equation*, and where now $q_{\Delta z,z}(\Delta z, z)$ is the probability density for being at position z and making a step Δz in time

Δt. The FP equation can be derived in a way similar to the one presented in Sec. 3.3: We expand the integrand of Eq. (31) in a Taylor-series in terms of z, so that $P(z,t) = \int_\infty^\infty AB\, d\Delta z$, with

$$A = P(z,t) - \partial_t P(z,t)\Delta t - \partial_z P(z,t)\Delta z + \frac{1}{2}\partial_z^2 P(z,t)\Delta z^2 + ..., \tag{32}$$

where we have also expanded to first order in t, and which is of course the same as Eq. (27), and newly we have

$$B = q_{\Delta z,z}(\Delta z, z) - \partial_z q_{\Delta z,z}(\Delta z, z)\Delta z + \frac{1}{2}\partial_z^2 q_{\Delta z,z}(\Delta z, z)\Delta z^2 + ... \tag{33}$$

(note that the Taylor expansion is with respect to the second argument of $q_{\Delta z,z}$, we expand only with respect to z, not though with respect to Δz). In multiplying and evaluating the integrals, we use the normalization of $q_{\Delta z,z}$ ($\int q_{\Delta z,z}(\Delta z, z)\, d\Delta z) = 1$), the definition of the mean value ($\mu_{\Delta z}(z) := \int \Delta z q_{\Delta z,z}(\Delta z, z)\, d\Delta z$) and of the second moment ($\langle \Delta z^2 \rangle(z) := \int \Delta z^2 q_{\Delta z,z}(\Delta z, z)\, d\Delta z$), and expressions like $\int \Delta z\, \partial_z q_{\Delta z,z}(\Delta z, z)\, d\Delta z$ are considered to equal $\partial_z \int \Delta z q_{\Delta z,z}(\Delta z, z)\, d\Delta z \equiv \partial_z \mu_{\Delta z}(z)$, so that, on keeping all terms up to second order in Δz, we find the *Fokker-Planck equation*,

$$\partial_t P(z,t) = -\partial_z [V(z)P(z,t)] + \partial_z^2 [D(z)P(z,t)], \tag{34}$$

with $V(z) \equiv \mu_{\Delta z}(z)/\Delta t$ a drift velocity, and $D(z) \equiv \langle \Delta z^2 \rangle(z)/2\Delta t$ the diffusion coefficient (for a 3 dimensional formulation see e.g. Gardiner, 2004). The basic difference between the FP equation and the simple diffusion equation in Eq. (30) is the appearance of a drift term, and that both the drift velocity and the diffusion coefficient are allowed to be spatially dependent (Fick's law also allows a spatially dependent diffusion coefficient, see Eq. (21)). These differences allow the FP equation to model more complex diffusive behaviour.

The FP equation is also applied to velocity space, e.g. in plasma physics in order to treat collisional effects, or to position and velocity space together. It has the advantage of being a deterministic differential equations that allows to describe the evolution of stochastic systems, as long as the diffusivities and drift velocities are known, and as long as the conditions for its applicability are met, see the remarks below.

We can illustrate the typical use of the FP equation on the example of Brownian motion in the Langevin formalism in Sec. 3.1, which allowed us to calculate the diffusion coefficient in Eq. (17). If we are interested in the evolution of the distribution of particles P, then with the Langevin formalism we would have to follow a large number of individual particles over the times of

interest and then to construct the distribution function, which may become very intense in computing effort. Instead, one can use the diffusivity from Eq. (17), insert it into the FP equation, and solve the FP equation. Since in this example the diffusion coefficient is constant and there is no drift velocity, the FP equation reduces to the simple diffusion equation (22), whose solution for P is given in Eq. (23).

We just note that in the general case where the diffusion coefficient is z-dependent, $D = D(z)$, there is a small difference between the diffusive term in the Fokker-Planck equation and the Fickian diffusion equation, Eq. (21), in that the diffusivity is only once differentiated in the latter. This difference and its consequences are analyzed in van Milligen et al. (2005).

>From its derivation it is clear that the FP equation is suited only for systems close to equilibrium, with just small deviations of some particles from equilibrium, or, in the random walk sense, with just small steps of the particles performing the random walk, exactly as it holds for the simple diffusion equation in Sect. 3.3.

A further natural generalization for a diffusion equation in the approach followed here would be not to stop the Taylor expansion in Eqs. (32) and (33) at second order in z, but to keep all terms, which would lead to the so-called *Kramers-Moyal expansion*.

More details about the Fokker-Planck equation can be found in the literature (e.g. Gardiner, 2004).

3.5 Why Normal Diffusion Should Be the Usual Case

The appearance of normal diffusion in many natural phenomena close to equilibrium and the particular Gaussian form of the solution of the diffusion equation can also be understood from probability theory. The Central Limit Theorem (CLT) states that if a statistical quantity (random variable) is the sum of many other statistical quantities, such as the position of a random walker after n steps according to Eq. (9), and if (i) all the Δz_i have finite mean $\mu_{\Delta z}$ and variance $\sigma^2_{\Delta z}$, (ii) all the Δz_i are mutually independent, and (iii) the number n of the Δz is large, then, independent of the distribution of the Δz_i's, the distribution $P(z, t_n)$ of z_n is a Gaussian. In particular, if $\mu_{\Delta z} = 0$, $z_0 = 0$ and all the Δz_i have the same variance, then

$$P(z, t_n) = \frac{1}{2\pi n \sigma^2_{\Delta z}} e^{-\frac{z^2}{2n\sigma^2_{\Delta z}}} \tag{35}$$

with variance $\sigma^2_{z_n} = n\sigma^2_{\Delta z}$, or, if we set $n = t/\Delta t$, then $\sigma^2_{z_n} = t_n\sigma^2_{\Delta z}/\Delta t$.

The mean square displacement equals per definition the variance,

$$\langle z(t_n)^2\rangle = \int z^2 P(z,t_n)\,dz = \sigma^2_{z_n} = t_n\sigma^2_{\Delta z}/\Delta t \tag{36}$$

(for $z_0 = 0$), and diffusion is thus always normal in the cases where the CLT applies.

Moreover, the CLT predicts quantities that are the result of many small scale interactions to be distributed according to a Gaussian, and indeed this is what we found for the distribution of the classical random walker, see Eq. (23). Stated differently, we may say that the appearance of non-Gaussian distributions is something unexpected and unusual according to the CLT. We just mention that also the equilibrium velocity distributions of gas or fluid particles are in accordance with the CLT, the velocity components, say v_x, v_y, v_z, follow Gaussian distributions, and therewith the magnitude $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$ exhibits a Maxwellian distribution. Again then, the appearance of non-Maxwellian velocity distributions is unexpected on the base of the CLT.

4 Anomalous Diffusion

4.1 Systems Far from Equilibrium: The rotating Annulus

The simple experiment of the rotating annulus, shown in Fig. 2, allows to illustrate the differences between normal and anomalous diffusion (Solomon, Weeks & Swinney, 1994; Weeks, Urbach & Swinney, 1996). Water is pumped into the annulus through a ring of holes marked with I and pumped out through a second ring of holes marked with O. The annulus is completely filled with water and rotates as a rigid body (the inner and outer walls rotate together). The pumping of the fluid generates a turbulent flow in the annulus. A camera on top of the annulus records the formation of the turbulent eddies inside the rotating annulus and allows to track seeds of different tracer particles injected into the fluid and to monitor their orbits (see Fig. 3).

In the case of normal diffusion, which occurs mainly in fluids close to equilibrium, the particle trajectories are characterized by irregular, but small steps, which makes trajectories look irregular but still homogeneous (see Fig. 1). The trajectories shown in Fig. 3 for the highly turbulent rotating

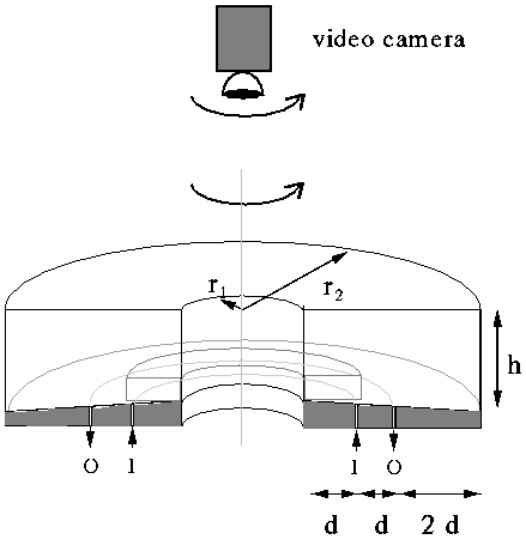

Fig. 2. Rotating annulus.

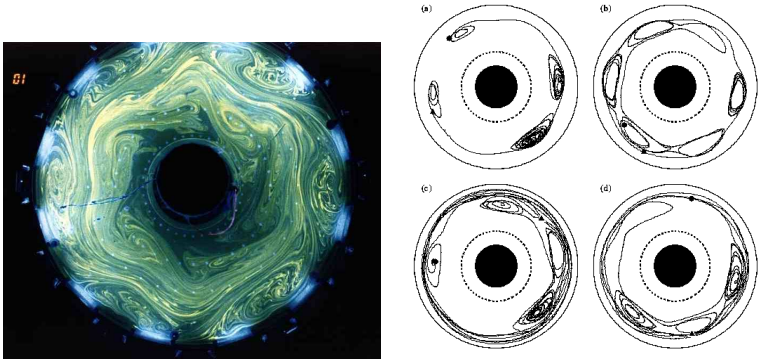

Fig. 3. (a) The formation of eddies inside the rotating annulus, as recorded by the camera (left panel), and (b) typical orbits of tracer particles inside the annulus (right panel).

annulus, which is far away from equilibrium, show different types of orbits, with two basic new characteristic, there is “trapping" of particles inside the eddies, where particles stay for ’unusually’ long times in a relatively small spatial area, and there are "long flights" of particles, where particles are carried in one step over large distances, in some cases almost through the entire system.

4.2 The Scaling of "Anomalous" Trajectories

Normal diffusion has as basic characteristic the linear scaling of the mean square displacement of the particles with time, $\langle r^2 \rangle \sim Dt$. Many different

experiments though, including the one shown in the previous section, reveal deviations from normal diffusion, in that diffusion is either faster or slower, and which is termed anomalous diffusion. A useful characterization of the diffusion process is again through the scaling of the mean square displacement with time, where though now we are looking for a more general scaling of the form

$$\langle r^2(t) \rangle \sim t^{\gamma}. \tag{37}$$

Diffusion is then classified through the scaling index γ. The case $\gamma = 1$ is normal diffusion, all other cases are termed anomalous. The cases $\gamma > 1$ form the family of super-diffusive processes, including the particular case $\gamma = 2$, which is called ballistic diffusion, and the cases $\gamma < 1$ are the sub-diffusive processes. If the trajectories of a sufficient number of particles inside a system are known, then plotting $\log < r^2 >$ vs $\log t$ is an experimental way to determine the type of diffusion occurring in a given system.

As an illustration, let us consider a particle that is moving with constant velocity v and undergoes no collisions and experiences no friction forces. It then obviously holds that $r = vt$, so that $\langle r^2(t) \rangle \sim t^2$. Free particles are thus super-diffusive in the terminology used here, which is also the origin of the name ballistic for the case $\gamma = 2$. Accelerated particles would even diffuse faster. The difference between normal and a anomalous diffusion is also illustrated in Fig. 4, where in the case of anomalous diffusion long "flights" are followed by efficient "trapping" of particles in localized spatial regions, in contrast to the more homogeneous picture of normal diffusion.

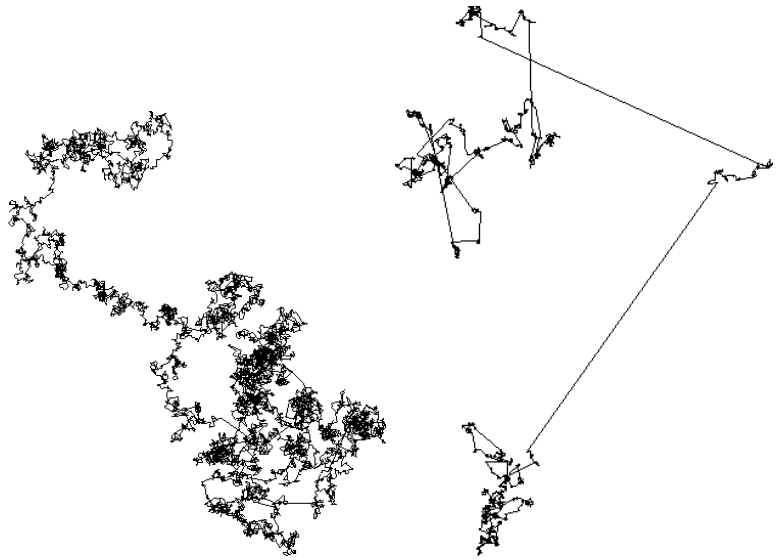

Fig. 4. (a) Random walk in dynamical systems close to equilibrium (normal diffusion; trajectory on the left), (b) random walk in dynamical systems far from equilibrium (anomalous diffusion; trajectory on the right).

It is to note that anomalous diffusion manifests itself not only in the scaling of Eq. (37) with $\gamma \neq 1$ (which experimentally may also be difficult to be measured), but also in 'strange' and 'anomalous' phenomena such as 'uphill' diffusion, where particles or heat diffuse in the direction of higher concentration, or the appearance of non-Maxwellian distributed particle velocities (see Sec. 3.5), very often of power-law shape, which is very common in high energy astrophysics (e.g. cosmic rays), etc.

5 Continuous Time Random Walk

5.1 Definition

Given the experimental ubiquity of anomalous diffusion phenomena, the question arises of how to model such phenomena. One way of tackling it is through the random walk formalism. So-far, we have used the random walk to model classical diffusion, and in Sec. 3.3 it had been shown how the random walk is related to a simple diffusion equation if the steps the particles take on their walk are small. One way to model anomalous diffusion is by relaxing the latter condition, and to allow the particles to also take large steps, where 'large' in a finite system means large up to system size, and in infinite systems it means that the steps are unbounded in length. Useful in this context is the family of Levy distributions as step-size distributions $q_{\Delta z}$. They are defined in closed form in Fourier space (see Sec. 5.3.3 below), and they have the property that

$$q^{L,\alpha}_{\Delta z}(\Delta z) \sim |\Delta z|^{-1-\alpha}, \quad \text{for } |\Delta z| \text{ large}, \quad 0 < \alpha < 2, \tag{38}$$

so that there is always a small, though finite probability for any arbitrarily large steps size. The Levy distributions all have an infinite variance, $\sigma^2_{L,\alpha} = \int \Delta z^2 q^{L,\alpha}_{\Delta z}(\Delta z)\, d\Delta z = \infty$, which makes their direct use as a step size distribution in the classical random walk of Sec. 2.2 and Eq. (9) impossible: Consider the case of a random walk in 1-D, with the position of the random walker after n steps given by the 1-D version of Eq. (9), and the mean square displacement (for $z_0 = 0$) given by Eq. (11). Let us assume that the steps are independent of each other, so that the covariances are zero and the mean square displacement is $\langle z_n^2 \rangle = n\sigma^2_{L,\alpha}$, which is infinite, already after the first step.

A way out of the problem is to release time from its dummy role and make it a variable that evolves dynamically, as the walker's position does. In this way, infinite steps in space can be accompanied by an infinite time for the

step to be completed, and the variance of the random walk, i.e. its mean square displacement, remains finite. The extension of the random walk to include the timing is called Continuous Time Random Walk (CTRW). Its formal definition consists again of Eq. (9), as described in Sec. 2.2, and, moreover, the time at which the nth step of the walk takes place is now also random (a random variable), and it evolves according to

$$t_n = \Delta t_n + \Delta t_{n-1} + \Delta t_{n-2} + \ ... \ + \Delta t_1 + t_0, \tag{39}$$

where t_0 is the initial time, and the Δt_i are random temporal increments. To complete the definition of the CTRW, we need also to give the probability distribution of the Δt_i, i.e. we must specify the probability for the ith step to last a time Δt_i.

Two case are usually considered (not least to keep the technical problems at a manageable level). (i) In the waiting model, the steps in position and time are independent, and one specifies two probabilities, one for $\Delta\vec{r}$, the $q_{\Delta\vec{r}}$ already introduced, and one for Δt, say $q_{\Delta t}$. Here then Δt is interpreted as a waiting time, the particle waits at its current position until the time Δt is elapsed, and then it performs a spatial step $\Delta\vec{r}$ during which no time is consumed (e.g. Montroll & Weiss, 1965). (ii) In the velocity model, the time Δt is interpreted as the traveling time of the particle, $\Delta t = |\Delta\vec{r}|/v$, where v is an assumed constant velocity (the velocity dynamics is not included, usually, see though Sec. 5.4), so that the distribution of increments is $q_{\Delta z,\Delta t} = \delta(\Delta t - |\Delta\vec{r}|/v)q_{\Delta\vec{r}}(\Delta\vec{r})$ (e.g. Shlesinger, West & Klafter, 1987). We just note that in the general case one would have to specify the joint probability distribution $q_{\Delta z,\Delta t}(\Delta z, \Delta t)$ for the spatial and temporal increments.

5.2 The CTRW Equations

The CTRW equations can be understood as a generalization of the Einstein equation, Eq. (26), or the Chapman-Kolmogorov equation, Eq. (31). It is useful to introduce the concept of the turning-points, which are the points at which a particle arrives at and starts a new random walk step. The evolution equation of the distribution of turning points $Q(z,t)$ (here in 1-D) follows basically from particle conservation,

$$\begin{aligned} Q(z,t) = &\int d\Delta z \int_0^t d\Delta t Q(z-\Delta z, t-\Delta t) q_{\Delta z,\Delta t}(\Delta z, \Delta t) \\ &+ \delta(t) P(z, t=0) + S(z,t), \end{aligned} \tag{40}$$

where the first term on the right side describes a completed random walk step, including stepping in space and in time, the second term takes the initial condition $P(z, t = 0)$ into account, and the third term S is a source term (see e.g. Zumofen & Klafter, 1993).

The expression for $P(z, t)$, the probability for the walker to be at position z at time t, is different for the waiting and for the velocity model, respectively. In case of the waiting model, where $q_{\Delta z,\Delta t}(\Delta z, \Delta t) = q_{\Delta t}(\Delta t)q_{\Delta z}(\Delta z)$, we have

$$P_W(z,t) = \int_0^t d\Delta t\, Q(z, t-\Delta t)\Phi_W(\Delta t), \tag{41}$$

with $\Phi_W(\Delta t) := \int_{\Delta t}^{\infty} dt' q_{\Delta t}(t')$ the probability to wait at least a time Δt (e.g. Zumofen & Klafter, 1993).

In the velocity model, where $q_{\Delta z,\Delta t}(\Delta z, \Delta t) = \delta(\Delta t - |\Delta z|/v)\, q_{\Delta z}(\Delta z)$, $P(z, t)$ takes the form

$$P_V(z,t) = \int_{-vt}^{vt} d\Delta z \int_0^t d\Delta t\, Q(z-\Delta z, t-\Delta t), \Phi_V(\Delta z, \Delta t) \tag{42}$$

with

$$\Phi_V(\Delta z, \Delta t) = \frac{1}{2}\delta(|\Delta z| - v\Delta t) \int_{|\Delta z|}^{\infty} dz' \int_{\Delta t}^{\infty} dt' \delta(t' - |z'|/v) q_{\Delta z}(z') \tag{43}$$

the probability to make a step of length at least $|\Delta z|$ and of duration at least Δt (e.g. Zumofen & Klafter, 1993; Shlesinger, West & Klafter, 1987).

Both, the expression for P_W and P_V determine the probability for seeing the particle when moving in-between two turning points, taking into account only the part of the random walk in which time is consumed by the particle.

The kind of diffusion that the CTRW formalism yields depends on the distribution of step increments. If the increments are small, then the treatment of Sec. 3.3 can be applied again, diffusion is normal, and again a simple diffusion equation can be derived. If the increments are not small, then super- as well as sub-diffusion can result, depending on the concrete choice of increment distributions. For instance, small spatial steps in combination with Levy-distributed, long waiting times will yield sub-diffusion in the waiting model. An important property of the CTRW equations is that they are non-local, both in space and time (which is also termed non-Markovian).

Anomalous diffusion phenomena in the CTRW approach are thus considered non-local processes, and with that they are far from equilibrium processes.

5.3 Treating the CTRW Equations

5.3.1 Remarks on the solution of the CTRW equations

A standard way to treat the CTRW equation is by transforming them to Fourier (F) and Laplace (L) space, whereby the convolution theorems of the two transforms are used. We will illustrate this procedure on the example of the waiting model below in Sec. 5.3.4.

The CTRW equations are though not always of a convolution type, e.g. the velocity model has not a convolution structure anymore, due to the appearance of time in the integration limits (see Eq. (42)), so that Fourier Laplace methods are not directly applicable anymore, and other methods are needed (actually also in the expression for Q, Eq. (40), time appears in the Δz-integration limits in case of the velocity model).

FL transforms, if applicable, usually do not allow to calculate the probability $P(z,t)$ in closed analytical form, but rather some asymptotic properties of it, such as the mean square displacement at large times (e.g. Klafter, Blumen & Shlesinger, 1987; Blumen, Zumofen & Klafter, 1989). On the other hand, FL transforms, if applicable, allow to transform the CTRW equations into other kinds of equations, e.g. in one instead of the two integral equations (Klafter, Blumen & Shlesinger, 1987; Blumen, Zumofen & Klafter, 1989, e.g.), into an integro-differential equation (or master equation, e.g. Klafter, Blumen & Shlesinger, 1987), or even into a fractional diffusion equation, which has the form of a diffusion equations, the fractional derivatives are though generalized, non-local differentiation operators. In Sec. 6.2, we will show how a fractional diffusion equation arises naturally in the context of the waiting model (see also e.g. Metzler & Klafter, 2000, 2004).

Another standard way of treating the CTRW equations is with Monte Carlo simulations (see e.g. Vlahos, Isliker & Lepreti, 2004), or else, the equations can be solved numerically, with an appropriate method (Isliker, 2008).

5.3.2 Fourier and Laplace transforms of probability densities

For any probability density function (pdf) such as $q_{\Delta z}$, we can define the Fourier transform as ($z \rightarrow k$)

$$\hat{q}_{\Delta z}(k) = \int e^{-ik\Delta z} q_{\Delta z}(\Delta z)\, d\Delta z, \tag{44}$$

which is often called the characteristic function of $q_{\Delta z}$. Considering then the expression

$$i^n \partial_k^n \hat{q}_{\Delta z}(k)\Big|_{k=0} = i^n \int (-i\Delta z)^n e^{-ik\Delta z} q_{\Delta z}(\Delta z)\, d\Delta z\Big|_{k=0}$$
$$= \int \Delta z^n q_{\Delta z}(\Delta z)\, d\Delta z, \tag{45}$$

for $n = 0, 1, 2, 3...$, we see that, because $q_{\Delta z}$ is a pdf, the last expression is the expectation value $\langle \Delta z^n \rangle$ of Δz^n, the so-called nth moment, and we have

$$i^n \partial_k^n \hat{q}_{\Delta z}(0) = \langle \Delta z^n \rangle. \tag{46}$$

In particular, it always holds that $\langle \Delta z^0 \rangle = 1$, since $q_{\Delta z}$ is a pdf that is normalized to one ($\langle \Delta z^0 \rangle = \int q_{\Delta z}(\Delta z)\, d\Delta z = 1$). Furthermore, $\langle \Delta z^1 \rangle = \int \Delta z q_{\Delta z}(\Delta z)\, d\Delta z$ is the mean value of $q_{\Delta z}$.

In the use of Fourier and Laplace transforms for solving the CTRW equations, we will concentrate on the asymptotic, large $|z|$ regime, which corresponds to small values of k (small wave-numbers correspond to large length-scales or wave-lengths). We thus can make a Taylor expansion of $\hat{q}_{\Delta z}(k)$ around $k = 0$ and keep only a few low order terms,

$$\hat{q}_{\Delta z}(k) = \hat{q}_{\Delta z}(0) + \partial_k \hat{q}_{\Delta z}(0) k + \frac{1}{2} \partial_k^2 \hat{q}_{\Delta z}(0) k^2 + ..., \tag{47}$$

With Eq. (46), the derivatives can be replaced with the moments,

$$\hat{q}_{\Delta z}(k) = 1 - i\langle \Delta z \rangle k - \frac{1}{2} \langle \Delta z^2 \rangle k^2 + ... \tag{48}$$

(with $\langle \Delta z^0 \rangle = 1$). The Taylor expansion of a pdf in terms of the moments is practical because the moments are the natural characteristics of a pdf. Often, the distribution of spatial increments is assumed to be symmetric around $z = 0$, so that $\langle \Delta z \rangle = 0$, and the Taylor expansion writes in this case as

$$\hat{q}_{\Delta z}(k) = 1 - \frac{1}{2} \langle \Delta z^2 \rangle k^2 + ... \tag{49}$$

Temporal distributions, such as the time-step distribution $q_{\Delta t}(\Delta t)$ in the waiting model, have the characteristic to be 'one-sided', i.e. they are defined and used only for $t \geq 0$, so that it is more appropriate to use Laplace

transforms in this case, defined as

$$\tilde{q}_{\Delta t}(s) = \int_0^\infty e^{-s\Delta t} q_{\Delta t}(\Delta t)\, d\Delta t.. \tag{50}$$

Straightforward calculations and the use of the respective definitions leads to the analogue of Eq. (46) for Laplace transforms,

$$(-\partial_s)^n \tilde{q}_{\Delta t}(s)\Big|_{s=0} = \int_0^\infty \Delta t^n q_{\Delta t}(\Delta t)\, d\Delta t = \langle \Delta t^n \rangle, \tag{51}$$

where the $\langle \Delta t^n \rangle$ are again the moments. As with respect to z, we will focus on the asymptotic, large t regime, which corresponds to small values of s, and we make a Taylor-expansion of the Laplace transform around $s = 0$, replacing through Eq. (51) the derivatives by the moments,

$$\tilde{q}_{\Delta t}(s) = 1 - \langle \Delta t \rangle s + ..., \tag{52}$$

in complete analogy to Eq. (48) ($\langle \Delta t^0 \rangle = 1$ is the normalization of $q_{\Delta t}$). The Taylor expansions in Eqs. (48), (49), and (52) can of course only be used if the involved moments are finite.

5.3.3 The symmetric and the one-sided Levy distributions

The *symmetric Levy distributions* are defined in Fourier space as

$$\hat{q}^{L,\alpha}_{\Delta z}(k) = \exp(-a|k|^\alpha), \tag{53}$$

with $0 < \alpha \leq 2$. It is not possible to express them in closed form in real space, with two exceptions, the case $\alpha = 2$ is the usual Gaussian distribution (the Fourier back-transform of a Gaussian is a Gaussian), and the case $\alpha = 1$ is known as the Cauchy distribution (see e.g. Hughes, 1995, Chap. 4.3). As mentioned in Sec. 5.1, the Levy distributions for $\alpha < 2$ all have infinite variance, and for $\alpha \leq 1$ they even have an infinite mean value, so that the expansion in the form of Eq. (49) is not applicable. We can though directly expand the exponential in Eq. (53) and find the small k expansion as

$$\hat{q}^{L,\alpha}_{\Delta z}(k) = 1 - a|k|^\alpha + ..., \tag{54}$$

The case $\alpha = 2$ corresponds obviously to the classical case of Eq. (49) with finite second moment ($a \equiv (1/2)\langle \Delta z^2 \rangle$) and zero mean (we are using only the symmetric Levy-distributions).

If one-sided distributions with infinite variance are needed, then they can also be defined via Fourier space as a specific asymmetric Levy distribution, or, more convenient for our purposes, as shown in Hughes (1995) (Chap. 4.3.2), they can be defined in Laplace space as

$$\tilde{q}^{L1,\beta}(s) = \exp(-bs^{\beta}), \tag{55}$$

with b strictly positive, and where now $0 < \beta \leq 1$. These one-sided Levy distributions decay as $t^{-1-\beta}$ for $t \to \infty$, and they have the small s expansion

$$\tilde{q}^{L1,\beta}(s) = 1 - bs^{\beta} + ... \tag{56}$$

Note that for $\beta = 1$ we recover the finite mean case of Eq. (52), with $b = \langle \Delta t \rangle$, and the back-transform of Eq. (55) in this case yields the distribution $\delta(\Delta t - b)$, i.e. the time-steps are constant and equal to b ($\Delta t \equiv \langle \Delta t \rangle \equiv b$).

In the following, we will use the (small k) Fourier expansion for $\hat{q}_{\Delta z}(k)$ in the form,

$$\hat{q}_{\Delta z}(k) = 1 - a|k|^{\alpha} + ..., \tag{57}$$

which for $\alpha < 2$ corresponds to the Levy case, Eq. (54), and for $\alpha = 2$ it recovers the normal, finite variance case of Eq. (49), with $a = (1/2)\langle \Delta z^2 \rangle$. Correspondingly, the (small s) Laplace expansion for $\tilde{q}_{\Delta t}(s)$ will be used in the form

$$\tilde{q}_{\Delta t}(s) = 1 - bs^{\beta} + ..., \tag{58}$$

which for $\beta < 1$ yields the Levy distribution, Eq. (56), and for $\beta = 1$ the normal, finite mean case of Eq. (52), with $b = \langle \Delta t \rangle$.

5.3.4 Solving the CTRW equations with Fourier and Laplace transforms

In this section, we will illustrate the use of the Fourier and Laplace (F-L) transform to solve the CTRW equation on the example of the waiting model, Eqs. (40) and (41). The use of the respective convolution theorems, namely

$$\int f(x-y)g(y)\,dy \to \hat{f}(k)\,\hat{g}(k) \tag{59}$$

for Fourier transforms, and

$$\int_0^t \phi(t-\tau)\psi(\tau)\,d\tau \rightarrow \tilde{\phi}(s)\,\tilde{\psi}(s) \tag{60}$$

for Laplace transforms (with f, g, ϕ, and ψ any transformable functions), allows in the case of the waiting model to determine the solution in Fourier Laplace space: For the initial condition $P(z,t=0)=\delta(t)\delta(z)$ and in the absence of any source ($S=0$), Eq. (40) turns into $\tilde{\hat{Q}}(k,s)=\tilde{\hat{Q}}(k,s)\,\hat{q}_{\Delta z}(k)\,\tilde{q}_{\Delta t}(s)+1$, and Eq. (41) takes the form $\tilde{\hat{P}}_W(k,s)=\tilde{\hat{Q}}(k,s)\tilde{\Phi}_W(s)$, and we can eliminate $\tilde{\hat{Q}}$ and solve for $\tilde{\hat{P}}$. Noting further that $\Phi_W(t)=\int_t^\infty q_{\Delta t}(\Delta t)\,d\Delta t = 1-\int_\infty^t q_{\Delta t}(\Delta t)\,d\Delta t$, so that $\tilde{\Phi}_W(s)=(1-\tilde{q}_{\Delta t}(s))/s$, we find

$$\tilde{\hat{P}}_W(k,s)=\frac{1-\tilde{q}_{\Delta t}(s)}{s\,[1-\hat{q}_{\Delta z}(k)\tilde{q}_{\Delta t}(s)]}, \tag{61}$$

which is known as the *Montroll-Weiss equation* (e.g. Montroll & Weiss, 1965; Zumofen & Klafter, 1993; Klafter, Blumen & Shlesinger, 1987).

Looking for asymptotic solutions, we insert the general form of the transformed temporal and spatial step distribution, Eqs. (57) and (58), respectively, into Eq. (61), which yields

$$\tilde{\hat{P}}_W(k,s)=\frac{bs^{\beta-1}}{bs^\beta+a|k|^\alpha}. \tag{62}$$

Unfortunately, it is not possible to Fourier and Laplace back-transform $\tilde{\hat{P}}_W(k,s)$ analytically. We can though use Eq. (46) for $n=2$, namely

$$\langle z^2(s)\rangle = -\partial_k^2\tilde{\hat{P}}_W(k=0,s), \tag{63}$$

to determine the mean square displacement in the asymptotic regime (note that we set $k=0$ at the end, which clearly is in the large $|z|$ regime). Inserting $\tilde{\hat{P}}_W(k,s)$ from Eq. (62) into Eq. (63), without yet setting $k=0$, we find

$$\langle z^2(s)\rangle = -\frac{2a^2\alpha^2|k|^{2\alpha-2}}{bs^{2\beta+1}}+\frac{a\alpha(\alpha-1)|k|^{\alpha-2}}{bs^{\beta+1}} \tag{64}$$

The first term on the right side diverges for $\alpha<1$, and the second term diverges for $\alpha<2$, so that $\langle z^2(s)\rangle$ is infinite in these cases. This divergence must be interpreted in the sense that the diffusion process is very efficient,

so that in the asymptotic regime, P_W has already developed so fat wings at large $|z|$ (power-law tails) that the variance, and with that the mean square displacement, of P_W is infinite, P_W has already become a Levy type distribution (see also Klafter, Blumen & Shlesinger, 1987; Balescu, 2007a). Of course, with the formalism we apply we cannot say anything about the transient phase, before the asymptotic regime is reached. We just note here that the velocity model (Eq. (42))in this regard is not so over-efficient, it allows super-diffusion with a more gradual build-up of the fat wings of the distribution P_V (Klafter, Blumen & Shlesinger, 1987).

Less efficient diffusion can only be achieved in the frame of the waiting model for $\alpha = 2$, i.e. for normal, Gaussian distributed spatial steps (see Eq. (57)). In this case, Eq. (64) takes the form

$$\langle z^2(s)\rangle = \frac{\langle \Delta z^2\rangle}{bs^{\beta+1}} \tag{65}$$

($a = (1/2)\langle \Delta z^2\rangle$ for $\alpha = 2$). This expression is valid for small s, and with the help of the Tauberian theorems, which relate the power-law scaling of a Laplace transform at small s to the scaling in original space for large t (see e.g. Hughes, 1995; Feller, 1971) it follows that

$$\langle z^2(s)\rangle\rangle \sim t^{\beta}. \tag{66}$$

With our restriction $0 < \beta \leq 1$, diffusion is always of sub-diffusive character, and for $\beta = 1$ it is normal, as expected, since we have in this case waiting times with finite mean and variance (see Eq. (58)).

5.4 Including Velocity Space Dynamics

Above all in applications to turbulent systems, and mainly to turbulent or driven plasma systems, it may not be enough to monitor the position and the timing of a particle, since its velocity may drastically change, e.g. if it interacts with a local electric field generated by turbulence. An interesting extension of the standard CTRW for these cases is to include, besides the position space and temporal dynamics, also the velocity space dynamics, which allows to study anomalous diffusive behaviour also in energy space.

To formally define the extended CTRW that also includes momentum space, we keep Eq. (9) and Eq. (39) for position and time evolution as they are, and newly the momentum (or velocity) also becomes a random, dynamic

variable, with temporal evolution of the form

$$\vec{p}_n = \Delta\vec{p}_n + \Delta\vec{p}_{n-1} + \Delta\vec{p}_{n-2} + \ ... \ + \Delta\vec{p}_1 + \vec{p}_0, \tag{67}$$

with p_0 the initial momentum, and the $\Delta\vec{p}_i$ the momentum increments. Again, one has to specify a functional form for the distribution of momentum increments $q_{\Delta\vec{p}}(\Delta\vec{p})$ in order to specify the random walk problem completely. The solution of the extended CTRW is in the form of the distribution $P(\vec{r},\vec{p},t)$ for a particle at time t to be at position $\vec{r}$ and to have momentum $\vec{p}$.

The extended CTRW can be treated by Monte-Carlo simulations, as done in Vlahos, Isliker & Lepreti (2004), or in Isliker (2008) a set of equations for the extended CTRW has been introduced, which basically is a generalization of Eq. (40) and Eq. (42), and a way to solve the equations numerically is presented.

6 From random walk to fractional diffusion equations

The purpose of this section is to show how fractional diffusion equations naturally arise in the context of random walk models. The starting point here are the CTRW equations for the waiting model, Eqs. (40) and (41), which in Fourier Laplace space take the form of Eq. (61), and on inserting the small k and small s expansion of the step-size and waiting-time distributions, Eqs. (57) and (58), respectively, the waiting CTRW equation takes the form of Eq. (62), with $\alpha \leq 2$ and $\beta \leq 1$. Multiplying Eq. (62) by the numerator on its right side,

$$\tilde{\hat{P}}_W(k,s)(bs^\beta + a|k|^\alpha) = bs^{\beta-1}, \tag{68}$$

and rearranging, we can bring the equation for $\tilde{\hat{P}}_W$ to the form

$$s^\beta \tilde{\hat{P}}_W(k,s) - s^{\beta-1} = -\frac{a}{b}|k|^\alpha \tilde{\hat{P}}_W(k,s). \tag{69}$$

It is illustrative to first consider the case of normal diffusion, with $\beta = 1$ and $\alpha = 2$, where according to Eqs. (57) and (58) we have $a = (1/2)\langle\Delta z^2\rangle$ and $b = \langle\Delta t\rangle$, so that

$$s\tilde{\hat{P}}_W(k,s) - s^0 = -\frac{\langle\Delta z^2\rangle}{2\langle\Delta t\rangle}|k|^2 \tilde{\hat{P}}_W(k,s) \tag{70}$$

Now recall how a first order temporal derivative is expressed in Laplace space,

$$\frac{d}{dt}\psi(z) \rightarrow s\tilde{\psi}(s) - s^0\psi(0), \tag{71}$$

and a how a spatial derivative translates to Fourier space,

$$\frac{d^n}{dz^n}f(z) \rightarrow (-ik)^n\hat{f}(k) \tag{72}$$

Obviously, for $P_W(z, t = 0) = \delta(z)$, Eq. (70) can be back-transformed as

$$\partial_t P_w(z,t) = \frac{\langle \Delta z^2 \rangle}{2\langle \Delta t \rangle}\partial_z^2 P_W(z,t), \tag{73}$$

so that we just recover the simple diffusion equation of the normal diffusive case.

6.1 Fractional derivatives

Fractional derivatives are a generalization of the usual derivatives of nth order to general non-integer orders. There exist several definitions, and in original space (z or t) they are a combination of usual derivatives of integer order and integrals over space (or time). The latter property makes them non-local operators, so that fractional differential equations are non-local equations, as are the CTRW integral equations. For the following, we need to define the Riemann-Liouville left-fractional derivative of order α,

$${}_aD_z^\alpha f(z) = \frac{1}{\Gamma(n-\alpha)}\frac{d^n}{dz^n}\int_a^z \frac{f(z')}{(z-z')^{\alpha+1-n}}\,dz', \tag{74}$$

with Γ the usual Gamma-function, a a constant, n an integer such that $n - 1 \leq \alpha < n$, α a positive real number, and f any suitable function. Correspondingly, the Riemann-Liouville right-fractional derivative of order α is defined as

$${}_zD_b^\alpha f(z) = \frac{(-1)^n}{\Gamma(n-\alpha)}\frac{d^n}{dz^n}\int_z^b \frac{f(z')}{(z'-z)^{\alpha+1-n}}\,dz'. \tag{75}$$

It is useful to combine these two asymmetric definitions into a a new, symmetric fractional derivative, the so-called Riesz fractional derivative,

$$D^{\alpha}_{|z|} f(z) = -\frac{1}{2\cos(\pi\alpha/2)} \left({}_{-\infty}D^{\alpha}_{z} + {}_{z}D^{\alpha}_{\infty} \right) f(z).. \tag{76}$$

The Riesz fractional derivative has the interesting property that its representation in Fourier space is

$${}^{(R)}D^{\alpha}_{|z|} f(z) \rightarrow -|k|^{\alpha} \hat{f}(k). \tag{77}$$

Comparison of this simple expression with Eq. (72) makes obvious that the Riesz derivative is a natural generalization of the usual derivative with now non-integer α.

To treat time, a different variant of fractional derivative is useful, the Caputo fractional derivative of order β,

$${}^{(C)}D^{\beta}_{t} \psi(t) = \frac{1}{\Gamma(n-\beta)} \int_{0}^{t} \frac{1}{(t-t')^{\beta+1-n}} \frac{d^n}{dt'^n} \psi(t')\, dt', \tag{78}$$

with n an integer such that $n-1 \leq \beta < n$, and ψ any appropriate function. The Caputo derivative translates to Laplace space as

$${}^{(C)}D^{\beta}_{t} \psi(t) \rightarrow s^{\beta} \tilde{\psi}(s) - s^{\beta-1} \psi(0), \tag{79}$$

for $0 < \beta \leq 1$, which is again a natural generalization of Eq. (71) for the usual derivatives for now non-integer β (the Caputo derivative is also defined for $\beta \geq 1$, with Eq. (79) taking a more general form).

Further details about fractional derivatives can be found e.g. in Podlubny (1999) or in the extended Appendix of Balescu (2007b).

6.2 Fractional diffusion equation

Turning now back to Eq. (69), we obviously can identify the fractional Riesz and Caputo derivatives in their simple Fourier and Laplace transformed form, Eqs. (77) and (79), respectively, and write

$${}^{(C)}D^{\beta}_{t} P_W(z,t) = \frac{a}{b}\ {}^{(R)}D^{\alpha}_{|z|} P_W(z,t) \tag{80}$$

>From this derivation it is clear that the order of the fractional derivatives, α and β, are determined by the index of the step-size ($q_{\Delta z}$) and the waiting time ($q_{\Delta t}$) Levy distributions, respectively. It is also clear that Eq. (80) is just an alternative way of writing Eq. (69) or (62), and as such it is the asymptotic, large $|z|$, large t version of the CTRW equations (40) and (41). It allows though to apply different mathematical tools for its analysis that have been developed specially for fractional differential equations.

As an example, we may consider the case $\beta = 1$ and $0 < \alpha \leq 2$, where the diffusion equation is fractional just in the spatial part,

$$\partial_t P_w(z,t) = \frac{a}{b} \ {}^{(R)}D^{\alpha}_{|z|} P_W(z,t) \tag{81}$$

In Fourier Laplace space, this equation takes the form

$$\hat{P}_W(k,s) = \frac{b}{bs + a|k|^{\alpha}}, \tag{82}$$

which, on applying the inverse Laplace transform, yields

$$\hat{P}_W(k,t) = \exp\left(-\frac{a}{b}\,|k|^{\alpha} t\right), \tag{83}$$

which is the Fourier-transform of a symmetric Levy-distribution with time as a parameter (see Eq. (53)), and with index α equal to the one of the spatial step distribution $q_{\Delta z}$. Thus, for $\alpha < 2$, the solution has power-law tails, and the mean square displacement (or variance or second moment) is infinite, as we had found it in Eq. (64). For $\alpha = 2$, the solution $P_W(z,t)$ is a Gaussian (the Fourier back-transform of a Gaussian is a Gaussian), and we have normal diffusion.

7 Action diffusion in Hamiltonian systems

So-far, our starting point for modeling diffusion was mostly the random walk approach and a probabilistic equation of the Chapman Kolmogorov type (Eq. (31)). Here now, we turn to Hamiltonian systems, and we will show how from Hamilton's equations a quasi-linear diffusion equation can be derived. This diffusion equation is of practical interest when the Hamiltonian system consists in a large number of particles, so that it becomes technically difficult to follow the individual evolution of all the particles.

Let us consider a generic $N-$degrees of freedom Hamiltonian system with Hamiltonian $H(\mathbf{q},\mathbf{p})$ and equations of motion given by

$$\frac{d\mathbf{q}}{dt} = \frac{\partial H}{\partial \mathbf{p}}, \tag{84}$$

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{q}}, \tag{85}$$

where $\mathbf{q} = (q_1, ..., q_N)$ and $\mathbf{p} = (p_1, ..., p_N)$ are the canonical coordinates and momenta, respectively. In order to be integrable such a system should have N independent invariants of the motion, corresponding to an equal number of symmetries of the system (Goldstein, 1980). The integrability of a system is a very strong condition which does not hold for most systems of physical interest. However, in most cases we can consider our system as a perturbation of an integrable one and split the Hamiltonian accordingly to an integrable part and a perturbation. Then the description of the system can be given in terms of the action-angle variables of the integrable part (Note that a periodic integrable system can always be transformed to action-angle variables (Goldstein, 1980)), so that we can write

$$H(\mathbf{J}, \boldsymbol{\theta}, t) = H_0(\mathbf{J}) + \epsilon H_1(\mathbf{J}, \boldsymbol{\theta}, t), \tag{86}$$

with $\mathbf{J} = (J_1, ..., J_N)$ and $\boldsymbol{\theta} = (\theta_1, ...\theta_N)$ being the action and angle variables, respectively. H_0 is the integrable part of the original Hamiltonian and H_1 is the perturbation. The parameter ϵ is dimensionless and will be used only for bookkeeping purposes in the perturbation theory; it can be set equal to unity, in the final results. The evolution of the integrable system H_0 is given by the following equations of motion

$$\dot{\mathbf{J}} = 0 \tag{87}$$

$$\dot{\boldsymbol{\theta}} = \boldsymbol{\omega_0} t + \boldsymbol{\theta_0}, \tag{88}$$

where $\boldsymbol{\omega_0} = \partial H_0 / \partial \mathbf{J}$ are the frequencies of the integrable system H_0. The N action variables correspond to the N invariants of the motion required for the integrability of the system.

The perturbation H_1 leads to the breaking of this invariance due to its $\boldsymbol{\theta}$ dependence. The derivation of a quasilinear diffusion equation in the action space is the subject of this section, and the method to be used is the canonical perturbation theory applied for finite time intervals. This method of derivation is based on first principles and does not imply any statistical assumptions for the dynamics of the system, such as the presence of strong chaos resulting in phase mixing or loss of memory for the system. Moreover, it is as systematic as the underlying perturbation scheme of the canonical perturbation theory, so it can be extended to higher order and provide results beyond the quasilinear approximation (Kominis, 2008). Also, it is important to note that the method makes quite clear what physical ef-

fects are taken into account in the quasilinear limit and what effects are actually omitted. It is worth mentioning that non-quasilinear diffusion has been studied both analytically and numerically for a variety of physical systems (Cary, Escande & Verga, 1990; Helander & Kjellberg, 1994; Benisti & Escande, 1998; Laval & Pesme, 1983, 1999).

The basic idea of the canonical perturbation theory is the search of canonical transformations for the perturbed system (i.e. transformations which preserve the Hamiltonian structure of the system) under which the new (transformed) Hamiltonian is a function of the action only. For a near-integrable system this can be done approximately, and the new actions correspond to approximate invariants of the motion which contain all the essential features of the phase space structure. The transformations involved in canonical perturbation theory are expressed in terms of the so-called mixed-variable generating functions. These can be functions of a subset of the old variables along with a subset of the new ones (Goldstein, 1980). Thus, the transformation from $(\mathbf{J}, \boldsymbol{\theta})$ to $(\bar{\mathbf{J}}, \bar{\boldsymbol{\theta}})$ can be expressed by a generating function of the form $S(\bar{\mathbf{J}}, \boldsymbol{\theta}, t)$. The transformation equations are given in the following implicit form:

$$\mathbf{J} = \bar{\mathbf{J}} + \epsilon \frac{\partial S(\bar{\mathbf{J}}, \boldsymbol{\theta}, t)}{\partial \boldsymbol{\theta}}, \tag{89}$$

$$\bar{\boldsymbol{\theta}} = \boldsymbol{\theta} + \epsilon \frac{\partial S(\bar{\mathbf{J}}, \boldsymbol{\theta}, t)}{\partial \bar{\mathbf{J}}}. \tag{90}$$

Following a standard procedure (Goldstein, 1980), we seek a transformation to new variables $(\bar{J}, \bar{\theta})$ for which the new Hamiltonian $\bar{H}$ is a function of the action $\bar{\mathbf{J}}$ alone. Expanding S and $\bar{H}$ in power series of a small parameter ϵ

$$S = \bar{\mathbf{J}}\boldsymbol{\theta} + \epsilon S_1 + \epsilon^2 S_2 + ... \tag{91}$$

$$\bar{H} = \bar{H}_0 + \epsilon \bar{H}_1 + \epsilon^2 \bar{H}_2 + ..., \tag{92}$$

where the lowest-order term has been chosen to generate the identity transformation $\mathbf{J} = \bar{\mathbf{J}}$ and $\bar{\boldsymbol{\theta}} = \boldsymbol{\theta}$. The old action and angle can be also expressed as power series in ϵ:

$$\mathbf{J} = \bar{\mathbf{J}} + \epsilon \frac{\partial S_1(\bar{\mathbf{J}}, \boldsymbol{\theta}, t)}{\partial \boldsymbol{\theta}} + \epsilon^2 \frac{\partial S_2(\bar{\mathbf{J}}, \boldsymbol{\theta}, t)}{\partial \boldsymbol{\theta}} + ... \tag{93}$$

$$\bar{\boldsymbol{\theta}} = \boldsymbol{\theta} + \epsilon \frac{\partial S_1(\bar{\mathbf{J}}, \boldsymbol{\theta}, t)}{\partial \bar{\mathbf{J}}} + \epsilon^2 \frac{\partial S_2(\bar{\mathbf{J}}, \boldsymbol{\theta}, t)}{\partial \bar{\mathbf{J}}} + ..., \tag{94}$$

and the new Hamiltonian is

$$\bar{H}(\bar{\mathbf{J}},\bar{\boldsymbol{\theta}},t)=H(\mathbf{J},\boldsymbol{\theta},t)+\frac{\partial S(\bar{\mathbf{J}},\boldsymbol{\theta},t)}{\partial t}. \tag{95}$$

By substituting the respective power series in Eq. (95) and equating like powers of ϵ for the zero order we have

$$\bar{H}_0=H_0, \tag{96}$$

while in the first and second order we have the equations

$$\frac{\partial S_i}{\partial t}+\boldsymbol{\omega_0}\frac{\partial S_i}{\partial \boldsymbol{\theta}}=\bar{H}_i-F_i(\mathbf{J},\boldsymbol{\theta},t), \qquad i=1,2, \tag{97}$$

with

$$F_1(\mathbf{J},\boldsymbol{\theta},t)=H_1, \tag{98}$$

$$F_2(\mathbf{J},\boldsymbol{\theta},t)=\frac{1}{2}\frac{\partial^2 H_0}{\partial \mathbf{J}^2}\left(\frac{\partial S_1}{\partial \boldsymbol{\theta}}\right)^2+\frac{\partial H_1}{\partial \mathbf{J}}\frac{\partial S_1}{\partial \boldsymbol{\theta}} \tag{99}$$

providing the first and second order generating function S_1 and S_2, respectively. The latter are linear partial equations which can be solved in a time interval of interest $[t_0,t]$ by the method of characteristics (i.e. integration along the unperturbed orbits). Note that $\bar{H}_1$ and $\bar{H}_2$ are arbitrary functions which can be set equal to zero (for the application of the canonical perturbation theory in infinite time intervals, these functions have to be chosen so that they cancel secular terms (Goldstein, 1980)). For a general perturbation of the form

$$H_1=\sum_{\mathbf{m}\neq 0} H_{\mathbf{m}}(\mathbf{J},t)e^{i\mathbf{m}\cdot\boldsymbol{\theta}}, \tag{100}$$

the solution for the first order generating function S_1 can be written as

$$S_1(\bar{\mathbf{J}},\boldsymbol{\theta},t;t_0)=-\sum_{\mathbf{m}\neq 0} e^{i\mathbf{m}\cdot(\boldsymbol{\theta}-\boldsymbol{\omega_0}t)}\int_{t_0}^{t} H_{\mathbf{m}}(\mathbf{J},s)e^{i\mathbf{m}\cdot\boldsymbol{\omega_0}s}ds. \tag{101}$$

Similarly, the solution for S_2 can be readily obtained. The resulting expression (too lengthy to be presented here) is a periodic function of $\boldsymbol{\theta}$. This is the only information we need for S_2, since its exact form will not be involved in our calculations.

In the following, we show that the results of first order perturbation theory can be utilized in order to provide an evolution equation for the angle-averaged distribution function, which is accurate up to second order with respect to the perturbation parameter ϵ, namely a quasilinear action diffusion equation. Therefore, we can relate results from perturbation theory applied for a single particle motion, to the distribution function, describing collective particle motion. The latter is of physical interest in all cases where a large number of particles is involved in collective phenomena so that a statistical approach is required.

The evolution of the phase space distribution function F is governed by Liouville's equation (Goldstein, 1980)

$$\frac{\partial F}{\partial t} + [F, H] = 0, \tag{102}$$

where $[.,.]$ denotes the Poisson bracket, defined as $[f_1, f_2] = \nabla_{\mathbf{q}} f_1 \cdot \nabla_{\mathbf{p}} f_2 - \nabla_{\mathbf{p}} f_1 \cdot \nabla_{\mathbf{q}} f_2$ with $\mathbf{q}$ and $\mathbf{p}$ being the canonical positions and momenta, respectively. This equation simply expresses the incompressibility of the Hamiltonian flow and the invariance of the number of particles. It is well-known that for an integrable system, any function of the invariants of the motion (actions) is a solution of the Liouville's equation. For the case of a near-integrable system, an approximate distribution function can be obtained as a function of the approximate invariant of the motion, namely the new actions $\bar{\mathbf{J}}$, so that we can write

$$F(\mathbf{J}, \boldsymbol{\theta}, t) = F(\bar{\mathbf{J}}), \tag{103}$$

with $\bar{\mathbf{J}}$ given implicitly by Eq. (93). To second order, with respect to ϵ we have

$$\bar{\mathbf{J}} = \mathbf{J} - \epsilon \Delta_1 \mathbf{J} + \frac{\epsilon^2}{2} \frac{\partial}{\partial \mathbf{J}} \cdot (\Delta_1 \mathbf{J} \Delta_1 \mathbf{J}) + \epsilon^2 \Delta_2 \mathbf{J}, \tag{104}$$

with

$$\Delta_i \mathbf{J}(\mathbf{J}, \boldsymbol{\theta}, t; t_0) = \frac{\partial S_i(\mathbf{J}, \boldsymbol{\theta}, t; t_0)}{\partial \boldsymbol{\theta}}, \qquad i = 1, 2. \tag{105}$$

Substituting (104) in (103) and utilizing a Taylor expansion with respect to ϵ we have

$$F(\mathbf{J}, \boldsymbol{\theta}, t) = F(\mathbf{J}) - \epsilon \frac{\partial F(\mathbf{J})}{\partial \mathbf{J}} \cdot (\Delta_1 \mathbf{J}) + \frac{\epsilon^2}{2} \frac{\partial}{\partial \mathbf{J}} \cdot \left[(\Delta_1 \mathbf{J} \Delta_1 \mathbf{J}) \cdot \frac{\partial F(\mathbf{J})}{\partial \mathbf{J}} \right]$$

$$+\epsilon^2 \frac{\partial F(\mathbf{J})}{\partial \mathbf{J}}(\Delta_2 \mathbf{J}). \tag{106}$$

Noting that $\Delta_i \mathbf{J}(\mathbf{J}, \boldsymbol{\theta}, t_0; t_0) = 0$ we have $F(\mathbf{J}, \boldsymbol{\theta}, t_0) = F(\mathbf{J})$, and by averaging over the angles, we obtain

$$f(\mathbf{J}, t) = f(\mathbf{J}, t_0) + \frac{\epsilon^2}{2} \frac{\partial}{\partial \mathbf{J}} \cdot \left[\langle (\Delta_1 \mathbf{J} \Delta_1 \mathbf{J}) \rangle_{\boldsymbol{\theta}} \cdot \frac{\partial f(\mathbf{J}, t_0)}{\partial \mathbf{J}} \right], \tag{107}$$

where f is the angle-averaged distribution function

$$f(\mathbf{J}, t) = \langle F(\mathbf{J}, \boldsymbol{\theta}, t) \rangle_{\boldsymbol{\theta}}\,, \tag{108}$$

and we have used the fact that $\langle \Delta_i \mathbf{J} \rangle_{\boldsymbol{\theta}} = 0$ as obtained from Eqs. (105) and the fact that S_i are sinusoidal functions of $\boldsymbol{\theta}$. Taking the limit $t \to t_0$ in Eq. (107), we finally obtain the quasilinear action diffusion equation

$$\frac{\partial f}{\partial t} = \epsilon^2 \frac{\partial}{\partial \mathbf{J}} \cdot \left[\mathbf{D}(\mathbf{J}, t) \cdot \frac{\partial f}{\partial \mathbf{J}} \right], \tag{109}$$

with

$$\mathbf{D}(\mathbf{J}, t) = \frac{1}{2} \frac{\partial \langle (\Delta_1 \mathbf{J} \Delta_1 \mathbf{J}) \rangle_{\boldsymbol{\theta}}}{\partial t} \tag{110}$$

being the corresponding quasilinear diffusion tensor. Identifying that $\Delta_1 \mathbf{J}$ corresponds to the first order action variation, we see that $\mathbf{D}$ in Eq. (110) corresponds to the common definition of the diffusion tensor, as provided by the statistical approach, which is based on the Kramers-Moyal expansion of the master equation (see Sec. 3.4, and van Kampen, 1981).

It is worth mentioning that this form [Eq. (109)] of the action diffusion equation is similar to the diffusion equation obtained with the utilization of the Fick's Law [Eq. (21)]. However, it has been shown that for any Hamiltonian system, this form of the action diffusion equation is equivalent to the Fokker-Planck equation [Eq. (34)], due to the fact that the corresponding drift velocity and diffusion terms are related through

$$\mathbf{V} = \frac{\partial \mathbf{D}}{\partial \mathbf{z}}, \tag{111}$$

where $\mathbf{z} = \mathbf{J}$ for the case of the action diffusion equation (see Chap. 4 in Lichtenberg & Lieberman, 1992, and references therein) . This property is

implied directly from the canonical form of the underlying equations of motion. More generally, it has been shown that this form of the diffusion equation has further relation with the more general class of the microscopically reversible systems (see Molvig & Hizanidis, 1984, and references therein).

It is important to emphasize that the derivation procedure described here does not prerequisite any statistical assumption and is only based on the underlying equations of motion. In contrast to other statistical approaches, it is not necessary to assume strong stochasticity related to the completely chaotic regime where loss of memory takes place and the orbits are completely decorrelated. Therefore, Eq. (109), is capable of describing not only diffusion in the almost homogeneous "chaotic sea" in the phase space, but also intermittent motion in an inhomogeneous phase space structure where resonant islands and chaotic areas are interlaced. It is worth mentioning that the diffusion tensor $\mathbf{D}$ is time-dependent. The dependence of $\mathbf{D}$ on the actions, for the case of time-periodic perturbations, is through smooth localized functions which are centered around the corresponding resonances (Kominis, 2008). In the limit $t \to \infty$ these localized functions tend to Dirac delta functions, which commonly appear in standard derivations of the quasilinear diffusion tensors (Zaslavski & Filonenko, 1968). However, the consideration of this limit in the derivation of $\mathbf{D}$ corresponds to an extension to infinity of the limits of integration in the derivation of S_1, in Eq. (101), which implies an ergodicity assumption and a steady-state approach. In the most general case, the diffusion tensor $\mathbf{D}$ [Eq. (110)] is capable of describing not only steady-state but also transient diffusion phenomena, through its time dependence. In this case the time scales of the action diffusion are determined by the time-dependence of $\mathbf{D}$ as well as by the factor ϵ^2 in the r.h.s. of Eq. (109), which implies an actually slow diffusion process. Note that these time scales come naturally into play and there is no need for a priori separation of the distribution function in slow and fast varying parts as in many heuristic derivations of the Fokker-Planck equation.

In the previous paragraphs we have derived a general quasilinear action diffusion equation for a Hamiltonian system, with a minimum of assumptions on the underlying dynamics. The fact that this method of derivation is closely related to a systematic and rigorous perturbation scheme allows for extending these results beyond the quasilinear approximation. Therefore we can carry our perturbation scheme to higher order in order to provide a hierarchy of diffusion equations having higher-order derivatives of the action distribution function with respect to the action, in direct analogy to statistical approaches where higher-order Kramers-Moyal expansions are considered (see Sec. 3.4, and van Kampen, 1981). Note that the most appropriate method for handling calculations involved in higher-order perturbation theory is the utilization of Lie transforms (Kominis, 2008). The hierarchy of higher-order diffusion equations does not only provide better accuracy with

respect to the perturbation parameter ϵ, but is also capable of describing non-Gaussian evolution of the distribution function and resonant processes between the particle and beats of multiple spectral components of the perturbation, known as nonlinear resonances.

8 Other ways to model anomalous diffusion

Escande & Sattin (2007) review and discuss under what circumstances the Fokker-Planck equation (Eq. (34)) is able to model anomalous diffusion. In summary, the FP equation, which is a local model, is able to model anomalous diffusive behaviour in cases where there is a non-zero drift velocity, $V(z) \neq 0$, anomalous diffusion is thus based on drift effects.

Klafter, Blumen & Shlesinger (1987) briefly review attempts of using the Fokker-Planck equation with zero drift velocity, but spatially or temporally dependent diffusion coefficient. In order to account for anomalous diffusion, the spatial or temporal dependence of the diffusion coefficient must though be chosen in very particular ways, which are difficult to interpret physically.

Lenzi, Mendes & Tsallis (2003) shortly discuss non-linear diffusion equations, which have the form of the simple diffusion equation as in Eq. (30), with P though raised on one side of the equation to some power γ.

9 Applications in Physics and Astrophysics

CTRW has successfully been applied to model various phenomena of anomalous diffusion, including sub- and super-diffusive phenomena, in the fields of physics, chemistry, astronomy, biology, and economics (see the references in Metzler & Klafter, 2000, 2004).

Laboratory plasma in fusion devices (tokamaks) show a variety of anomalous diffusion phenomena. Balescu (1995) was the first to apply CTRW to plasma physical problems. Later, van Milligen, Sánchez & Carreras (2004) and van Milligen, Carreras & Sánchez (2004) developed a CTRW model for confined plasma, the critical gradient model, which was able to explain observed anomalous diffusion phenomena such as 'up-hill' transport, where particles diffuse against the driving gradient. Isliker (2008) studied the same physical system, with the use though of the extended CTRW that includes momentum space dynamics, and they studied the evolution of the density and temperature distribution and the particle and heat diffusivities.

Also in astrophysical plasmas anomalous diffusion is ubiquitous, there are many astrophysical systems where non-thermal (i.e. not Maxwellian distributed) particles are directly or indirectly, through their emission, observed.

Dmitruk et al. (2003, 2004) analyzed the acceleration of particles inside 3-D MHD turbulence. The compressible MHD equations were solved numerically. In these simulations, the decay of large amplitude waves was studied. After a very short time (a few Alfvén times), a fully turbulent state with a broad range of scales has been developed (Fig. 5).

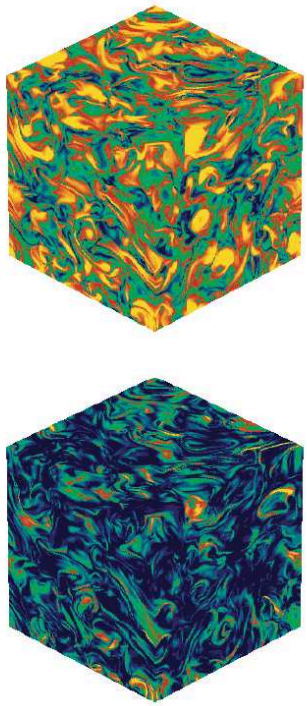

Fig. 5. Visualization of the turbulent magnetic field $\mid B \mid$ (top) and electric field $\mid E \mid$ (bottom) in the simulation box. High values are in yellow (light) and low values in blue (dark).

The magnetic field is directly obtained from the numerical solution of the MHD equations, with electric field derived from Ohm's law. It is obvious that the electric field is an intermittent quantity with the high values distributed in a less space filling way. Magnetic and electric fields show a broad range of scales and high degree of complexity. The energy spectrum of the MHD fields is consistent with a Kolmogorov-5/3 power law. The structure of the velocity field and the current density along the external magnetic field (J_z) can be seen in Fig. 6. The formation of strong anisotropies in the magnetic field, the fluid velocity and the associated electric field is observed. The overall picture is that current sheet structures along the DC field are formed as a natural evolution of the MHD fields.

Fig. 6. Cross section of the current density J_z along the external magnetic field in color tones. Yellow (light) is positive J_z, blue (dark) is negative, and the superposed arrows represent the velocity field.

Following thousands of particles particles inside the simulation box, we can learn many of the statistical properties of their evolution, e.g. the mean square displacements $\sqrt{<\Delta x^2>}$, $\sqrt{<\Delta v^2>}$, or the velocity distribution etc. can be determined. Electrons and ions are accelerated rapidly at the nonlinear small scale structures formed inside the turbulent volume, and non-thermal tails of power-law shape are formed in the velocity distributions. Most particles seem to escape the volume by crossing only a few of the randomly appearing current sheets. A few particles are trapped in these structures and accelerated to very high energies. The Fokker-Planck equation is not the appropriate tool to capture particle motion in the presence of the random appearance of coherent structures inside such a turbulent environment.

Vlahos, Isliker & Lepreti (2004) performed a Monte Carlo simulation of the extended CTRW in position and momentum space, in application to flares in the solar corona, with particular interest in the appearance of the non-thermal energy distributions of the so-called solar energetic particles.

10 Summary and Discussion

Brownian motion is a prototype of normal diffusion, and its analysis has brought forth a number of tools that today are very much in use for modeling a wide variety of phenomena. Normal diffusion occurs in systems which are close to equilibrium, like the water in Brown's experiment. It has now become evident that phenomena of anomalous diffusion are very frequent, because many systems of interest are far from equilibrium, such as turbulent systems, or because the space accessible to the diffusing particles has a strange, e.g. fractal structure. The tools to model these phenomena, continuous time random walk, stochastic differential equations, and fractional diffusion equations, are still active research topics.

References

Balescu, R., Phys. Rev. E **51**, 4807 (1995).

Balescu, R., Chaos Solitons & Fractals **34**, 62 (2007).

Balescu, R., arXiv:0704.2517v1, (2007).

Benisti D., and Escande, D.F., Phys. Rev. Lett. **80**, 4871 (1998).

Blumen, A., Zumofen, G., Klafter, J., Phys. Rev. A **40**, 3964 (1989).

Cary, J.R., Escande D.F. and Verga, A.D. , Phys. Rev. Lett. **65**, 3132 (1990).

P. Dmitruk, W.H. Matthaeus, N. Seenu, M.R. Brown: Astrophys. J. Lett., **597**, L81 (2003)

P. Dmitruk, W.H. Matthaeus, N. Seenu: Astrophys. J., **617**, 667 (2004)

Einstein, A., Ann. Phys. **17**, 549 (1905)

Escande, D.F & Sattin, F., Phys. Rev. Lett. **99**, 185005 (2007).

Feller, W., An Introduction to Probability Theory and its Applications, Vol. 2, 2nd edition, Wiley (New York), 1971.

Gardiner, C.W., Handbook of Stochastic Methods, 3rd ed. (Springer; Berlin & Heidelberg; 2004)

H. Goldstein, Classical Mechanics (2nd edition, Addison-Wesley, Massachusets, 1980).

Helander P., and Kjellberg, L., Phys. Plasmas **1**, 210 (1994).

Isliker, H., arXiv:0710:5152, and submitted to Phys. Rev. E (2008)

Hughes, B.D., Random Walks and Random Environments, Volume 1, Random Walks, Oxford (Clarendon Press), 1995.

N.G. van Kampen, Stochastic processes in physics and chemistry (North-Holland, Amsterdam, 1981).

Klafter, J., Blumen, A., Shlesinger, M.F., Phys. Rev. A **35**, 3081 (1987).

Kominis, Y., Phys. Rev. E **77**, 016404 (2008).

Langevin, P., C. R. Acad. Sci. (Paris) **146**, 530 (1908) [translated to English and commented in Lemons, D.S, Gythiel, A., Am. J. Phys. **65**, 1079 (1997)]

Laval, G. and Pesme, D., Phys. Fluids **26**, 66 (1983).

G. Laval and D. Pesme, Plasma Phys. Control. Fusion **41**, A239 (1999).

Lenzi, E.K., Mendes, R.S & Tsallis, C., Phys. Rev. E **67**, 031104 (2003).

Lichtenberg, A. J. and Lieberman, M. A., âœRegular and Chaotic Dynamicsâ (Springer-Verlag, New York, 1992).

Metzler, R., Klafter, J., Physics Reports **339**, 1 (2000).

Metzler, R., Klafter, J., Journ. of Phys. A **37**, R161 (2004).

van Milligen, B. Ph., Sánchez, R., Carreras, B.A., Phys. of Plasmas **11**, 2272 (2004)

van Milligen, B. Ph., Carreras, B.A., Sánchez, R., Physics of Plasmas **11**, 3787 (2004).

van Milligen, B. Ph., Bons, P.D., Carreras, B.A., Sánchez, R., Eur. J. Phys. **26**, 913 (2005).

Molvig, K., and Hizanidis, K., 'Transport theory: Microscopic reversibility and symmetry', Phys. Fluids **27**, 2847-2858 (1984).

E.W. Montroll, G.H. Weiss, J. Math. Phys. **6**, 167 (1965).

Podlubny, I., Fractional differential equations, New York (Academic press), 1999.
Shlesinger, M.F., West, B., Klafter, J., Phys. Rev. Lett. **58**, 1100 (1987).
Solomon, T.H. Weeks, E.R., Swinney, H.L., Physica D **76**, 70 (1994).
Vlahos, L., Isliker, H., Lepreti, F., Astrophys. Journ. **608**, 540 (2004).
Weeks,E.R., Urbach, J.S., Swinney, H.L., Physica D **97**, 291 (1996).
G.M. Zaslavski and N.N. Filonenko, Stochastic instability of trapped particles and conditions of applicability of the quasi-linear approximation, Sov. Phys. JETP **25**, 851 (1968).
Zumofen, G., Klafter, J., Phys. Rev. E **47** 851 (1993).