

Large deviations in Taylor dispersion

Marcel Kahlen, Andreas Engel, and Christian Van den Broeck
Universität Oldenburg, Institut für Physik, 26111 Oldenburg, Germany
Hasselt University, Faculty of Sciences, B-3590 Diepenbeek, Belgium

We establish a link between the phenomenon of Taylor dispersion and the theory of empirical distributions. Using this connection, we derive, upon applying the theory of large deviations, an alternative and much more precise description of the long-time regime for Taylor dispersion.

I. INTRODUCTION

Around 1953, G. I. Taylor [1] published a series of papers in which he studies the dispersion of particles injected in a Poiseuille flow through a cylindrical tube. As the particles are carried along, their elution profile broadens along the flow axis x into a co-moving Gaussian shape, with a mean square displacement described by an effective diffusive law, $\langle [x(t) - \langle x(t) \rangle]^2 \rangle = 2Kt$. He calculated the value of the "Taylor dispersion coefficient": $K = \bar{u}^2 a^2 / (48D)$, a result also obtained by Westhaver in his study of electro-migration [2]. Here D is the molecular diffusion coefficient of the suspended particles, \bar{u} the average flow velocity and a the radius of the tube. As an application, he discusses how such a set-up allows to estimate D , which is typically very small, by measuring K , a method used routinely in biophysics by now [3]. The catch is that the Gaussian regime is attained for times larger than the typically very long time a^2/D for the particles to diffuse over the tube's radius. The dispersion of particles in a flow is one example of a more general phenomenon, referred to as generalized Taylor dispersion, which we briefly review below, and which includes applications in chromatography, the sedimentation of non-spherical particles, spin relaxation and phase diffusion in limit cycles, models for kinetic theory, random evolution, etc. [4, 5].

After pioneering work by H. Cramer and I. N. Sanov, the mathematical theory of large deviations was primarily shaped in a series of influential papers by M. D. Donsker and S. R. S. Varadhan [6]. Recent years have seen an ever increasing number of studies of large deviations in statistical mechanics, see, e.g., [7–9]. The Gaussian distribution as it appears in Taylor's problem corresponds to the quadratic approximation for the so-called large deviation function (ldf). The explicit evaluation of the ldf for Taylor dispersion was done in [10] and provides a more precise description of the large time regime. But, as we will show in the present paper, the connection between both problems runs deeper: the large deviation description for generalized Taylor dispersion can be mapped onto that for empirical distributions. The empirical distribution is the probability distribution observed in a sample realization of a stochastic process. It plays a crucial role in statistics and its properties form a cornerstone of large deviation theory proper. The connection between Taylor dispersion and empirical distributions allows to compare and exchange techniques and

results between both fields of research.

The paper is organized as follows. In section II we shortly review generalized Taylor dispersion and explicitly establish the relation to the theory of empirical distributions. We then turn in section III to the large deviation analysis of Taylor dispersion in layered systems for which the underlying stochastic process acts in a discrete state space. In section IV we discuss continuous flow fields including the Poiseuille flow studied originally by Taylor. Finally, section V contains some concluding remarks.

II. GENERAL THEORY

To set the scene, we start with a brief review of generalized Taylor dispersion. A variable x has a rate of change u , which is driven by a stochastic process \mathbf{y} , independent of x :

$$\dot{x}(t) = u(\mathbf{y}(t)). \quad (1)$$

In the example described in the introduction, x is the coordinate measured along the tube, $\mathbf{y} = \mathbf{r}$ is the radial position of the particle undergoing molecular diffusion, and

$$u(\mathbf{r}) = 2\bar{u}\left(1 - \frac{\mathbf{r}^2}{a^2}\right) \quad (2)$$

the flow profile, cf. Fig. 1a. By integrating the equation of motion, one finds the relation between the sample position $x = x(t)$ at time t , and the particular realisation $\mathbf{y}(\tau)$, $\tau \in [0, t]$, of the stochastic process:

$$x(t) = \int_0^t u(\mathbf{y}(\tau)) d\tau. \quad (3)$$

We will assume that the stochastic process \mathbf{y} is stationary, with stationary distribution $p(\mathbf{y})$, and does not display long-time correlations. The resulting displacement x will, for long times, be the sum of essentially uncorrelated and identically distributed contributions, corresponding to subsequent displacements during time intervals larger than the correlation time of \mathbf{y} . Hence one expects from the central limit theorem that the distribution for x approaches a Gaussian in the long time limit:

$$P(x, t) \sim \frac{1}{\sqrt{4\pi Kt}} \exp\left\{-\frac{(x - \bar{u}t)^2}{4Kt}\right\}, \quad (4)$$

with the average speed

$$\langle u(\mathbf{y}) \rangle = \int u(\mathbf{y})p(\mathbf{y})d\mathbf{y} = \bar{u}, \quad (5)$$

and the Taylor dispersion coefficient given by a Green-Kubo formula:

$$K = \int_0^\infty \langle [u(\mathbf{y}(\tau)) - \bar{u}][u(\mathbf{y}(0)) - \bar{u}] \rangle d\tau. \quad (6)$$

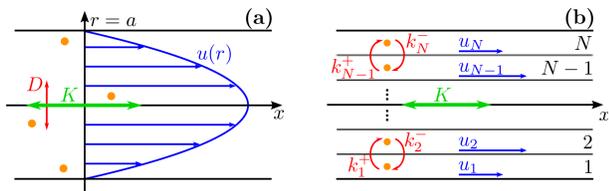


Figure 1. a) Taylor dispersion for a parabolic flow profile $u(\mathbf{y})$. Molecular diffusion with coefficient D perpendicular to the flow gives rise to a dispersion of particles parallel to the flow with effective diffusion constant K . b) In a discretized model the particles reside in N layers with drift velocities u_i and make random transitions between adjacent layers with rates k_i^+ and k_i^- .

The large deviation analysis of generalized Taylor dispersion builds on the well studied large deviation properties of the empirical density of the underlying stochastic process [6, 11]. To make the connection explicit let us consider the case in which the rate of change u_i of the coordinate x depends on a discrete variable $i \in \mathbf{I}$, rather than the continuous variable \mathbf{y} discussed before, cf. Fig. 1b. The equation of motion (1), is thus replaced by $\dot{x}(t) = u_{i(t)}$. The discrete state i undergoes an independent stationary stochastic process with steady state distribution p_i . During a realization of the stochastic process of duration t , the system will spend relative times

$$q_i(t) = \frac{1}{t} \int_0^t \delta_{i,i(\tau)} d\tau \quad (7)$$

in each of the states $i \in \mathbf{I}$. The object, $\mathbf{q} = \{q_i, i \in \mathbf{I}\}$, called the empirical density and itself a random variable, is to be distinguished from the "genuine" distribution $\mathbf{p} = \{p_i, i \in \mathbf{I}\}$. Nevertheless, for an ergodic process \mathbf{q} will converge to \mathbf{p} in the long-time limit. More precisely, one finds for large t

$$P(\mathbf{q}, t) \sim \exp\{-tJ(\mathbf{q})\}, \quad (8)$$

with the ldf $J(\mathbf{q})$ satisfying $J(\mathbf{p}) = 0$ and $J(\mathbf{q}) > 0$ for $\mathbf{q} \neq \mathbf{p}$. In other words, deviations of \mathbf{q} from \mathbf{p} are exponentially suppressed in the long-time limit.

Instead of considering the sample displacement at time t , $x = x(t)$, we now focus on the sample velocity $v = v(t) = x(t)/t$, which is clearly given by:

$$v = \sum_i u_i q_i = \mathbf{u} \cdot \mathbf{q}. \quad (9)$$

This expression links the stochastic properties of v to those of the empirical distribution \mathbf{q} . In particular, the large time properties of the sample speed are specified by a large deviation principle of the form

$$P(v, t) \sim \exp\{-tI(v)\}, \quad (10)$$

with the ldf $I(v)$ linked to $J(\mathbf{q})$ by the contraction principle of large deviation theory [7]:

$$I(v) = \inf_{\mathbf{u} \cdot \mathbf{q} = v, \mathbf{1} \cdot \mathbf{q} = 1} J(\mathbf{q}). \quad (11)$$

This property corresponds to the intuitively obvious observation that an exponentially rare sample speed v is realised by the least unlikely, exponentially rare realisation of an empirical distribution \mathbf{q} . We have explicitly included the constraint that \mathbf{q} is normalized, $\sum_i q_i = \mathbf{1} \cdot \mathbf{q} = 1$. Note that the central limit approximation, cf. Eq. (4), corresponds to the quadratic approximation of the ldf around its minimum,

$$I(v) \cong \frac{(v - \bar{u})^2}{4K}, \quad (12)$$

whereas the exact large deviation function will have contributions of higher order in v .

To proceed, we will assume that the stochastic process i is Markovian, and represents an equilibrium situation. This covers most applications in generalized Taylor dispersion. The Markov process is characterized by a transition matrix \mathbf{W} . Its elements are the rates $W_{i,j}$, corresponding to the probability per unit time to make a transition from state j to state i for $i \neq j$, and $W_{i,i} = -\sum_{j, j \neq i} W_{j,i}$. They obey the detailed balance property:

$$W_{j,i}p_i = W_{i,j}p_j. \quad (13)$$

In this case, the ldf of the empirical distribution is known explicitly [6]:

$$J(\mathbf{q}) = \frac{1}{2} \sum_{i,j} (\sqrt{W_{i,j}q_j} - \sqrt{W_{j,i}q_i})^2. \quad (14)$$

To account for the constraints in the contraction (11) we introduce two Lagrange multipliers λ_1 and λ_2 . The vector $\bar{\mathbf{q}}$ minimizing

$$J(\mathbf{q}) + \lambda_1(\mathbf{1} \cdot \mathbf{q} - 1) + \lambda_2(\mathbf{u} \cdot \mathbf{q} - v) \quad (15)$$

has to fulfill the system of linear equations:

$$(\lambda_1 + \lambda_2 u_i) \sqrt{q_i} - \sum_j \sqrt{W_{i,j}W_{j,i}} \sqrt{q_j} = 0 \quad \forall i \in \mathbf{I}. \quad (16)$$

Multiplying by $\sqrt{q_i}$ and summing over i yields:

$$I(v) = J(\bar{\mathbf{q}}) = -(\lambda_1 + \lambda_2 v). \quad (17)$$

So all that needs to be determined are the values of the Lagrange multipliers λ_1 and λ_2 . In general there will be

several solutions. To select the proper one it is instructive to write (16) in matrix form:

$$\mathbf{L} \begin{pmatrix} \sqrt{q_1} \\ \vdots \\ \sqrt{q_N} \end{pmatrix} = \lambda_1 \begin{pmatrix} \sqrt{q_1} \\ \vdots \\ \sqrt{q_N} \end{pmatrix}, \quad (18)$$

showing that λ_1 is an eigenvalue of the symmetric $N \times N$ matrix \mathbf{L} with elements $L_{i,j} = \sqrt{W_{i,j}W_{j,i}} - \lambda_2 u_i \delta_{i,j}$, with all entries of the corresponding eigenvector positive. Hence, by the Frobenius-Perron theorem, the largest eigenvalue of L is the relevant solution for the Lagrange multiplier λ_1 .

As a first application of the above formalism and to demonstrate its consistency with previous work, we re-derive the exact expression and bounds for the Taylor dispersion coefficient K . In order to capture the Gaussian regime we only need the quadratic expansion of the ldf $I(v)$ around the most probable sample value at $v = \bar{u}$, cf. Eq. (12). Starting from (14) we derive in appendix A the general expression [12, 13]

$$K = - \sum_{i,j} \tilde{W}_{i,j}^{-1} p_j (\bar{u} - u_i)(\bar{u} - u_j). \quad (19)$$

Here $\tilde{W}_{i,j}^{-1}$ denotes the inverse of $W_{i,j}$ on the subspace of non-zero eigenvalues. It is usually difficult if not impossible to evaluate this inverse exactly. As an alternative, the variational form of (11) invites for the derivation of bounds on K . Since $v = \bar{u}$ corresponds to $q_i = p_i$ it suffices to keep the first order deviations $q_i \cong p_i + f_i(v - \bar{u})$ in (14) before doing the contraction. In this way we show in appendix A

$$K \geq - \frac{1}{\langle \mathbf{f} | W | \mathbf{f} \rangle}, \quad (20)$$

where we used the scalar product $\langle \mathbf{f} | \mathbf{g} \rangle = \sum_i f_i g_i / p_i$ and \mathbf{f} is any vector perpendicular to \mathbf{p} and fulfilling $\langle \mathbf{f} | \mathbf{u} \mathbf{p} \rangle = 1$. This expression is identical to the lower bound for K derived previously in [13].

III. DISCRETE FLOW FIELDS

Turning to the large deviation properties, we first focus on the simplest model capturing the essential ingredients of Taylor dispersion. It consists of just two layers with velocities u_1 and u_2 , respectively. Because of Galilei invariance we may choose $u_2 = -u_1 =: u$ without loss of generality. We denote by $k_+ = W_{2,1}$ and $k_- = W_{1,2}$ the rates of transition from layer 1 to 2 and back, respectively. In this case, the exact elution profile is known for all times, and was derived in many different settings

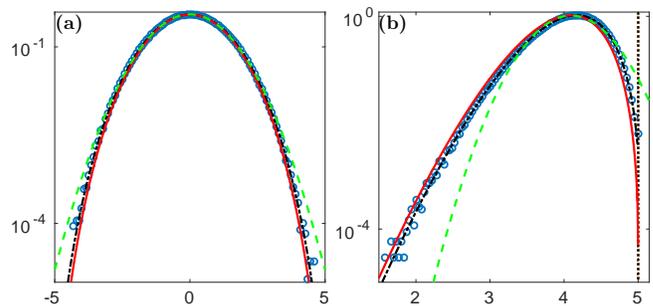


Figure 2. Logarithmic plot of the displacement distribution $P(x,t)$ at $t = 10$ for the discrete model of Fig. 1b with $N = 2$ layers and $u_2 = -u_1 = .5$. In a) the transition rates are symmetric, $k_+ = k_- = 2$, in b) they are biased, $k_+ = 10, k_- = 1$. Shown are the exact solution (21) (black), the large deviation result (red), the Gaussian approximation (dashed green), and results from numerical simulations using the Gillespie algorithm for 10^6 realisation (blue circles). The dotted vertical line in b) marks the cut-off of the distribution at $x = 5$.

[14–19].

$$P(x,t) = \frac{1}{4\pi} \left[\frac{\partial}{\partial t} + (k_+ + k_-) \right] \left[\frac{2\pi}{u} \exp \left(-\frac{k_+}{2} \left(t - \frac{x}{u} \right) - \frac{k_-}{2} \left(t + \frac{x}{u} \right) \right) I_0 \left(\frac{\sqrt{k_- k_+}}{u} \sqrt{(ut)^2 - x^2} \right) \theta(t) \theta(ut - |x|) \right] \quad (21)$$

Here $\theta(x)$ denotes the Heaviside step-function and $I_0(x)$ is the modified Bessel function of order zero. The ldf $I(v)$ follows directly from (11) because the two constraints $\bar{q}_1 + \bar{q}_2 = 1$ and $u(\bar{q}_1 - \bar{q}_2) = v$ already uniquely fix $\bar{\mathbf{q}}$ in terms of u and v . Hence, the contraction becomes trivial:

$$I(v) = \frac{k_+}{2} \left(1 - \frac{v}{u} \right) + \frac{k_-}{2} \left(1 + \frac{v}{u} \right) - \sqrt{k_+ k_-} \sqrt{1 - \frac{v^2}{u^2}}. \quad (22)$$

This result can, of course, also be obtained directly from the large-time limit of (21). In Fig. 2 we exemplarily compare the exact solution (21), with the Gaussian approximation (12), and the large-deviation result (22). For the latter we have determined the prefactor such that it coincides with the exact result for $v = \bar{u}$. Included are also results of numerical simulations using the Gillespie algorithm. As expected the large-deviation expression is superior to the Gaussian approximation for reproducing the tails of the distribution and follows the exact result down to rather low probabilities. Note that this is particularly striking for unbalanced transition rates as shown in Fig. 2b.

We next turn to a discrete model with N layers and nearest neighbour transitions, as sketched in Fig. 1b. No explicit analytic result for $I(v)$ is available for the case of general N , however, using standard numerical techniques the non-trivial solution $\sqrt{\bar{\mathbf{q}}}(\lambda_1, \lambda_2)$ of the homogeneous

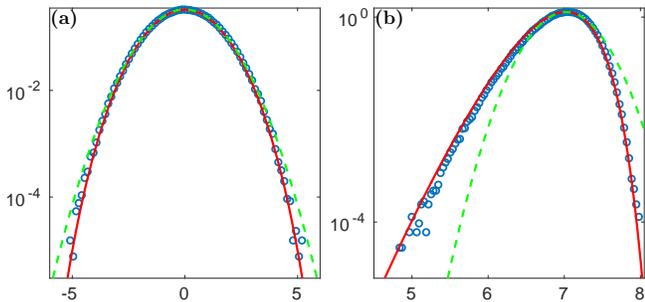


Figure 3. Same as figure 2 for a system of $N = 10$ layers. The symmetric situation a) is specified by $k_+ = k_- = 45$ the asymmetric one by $k_+ = 80, k_- = 40$. No exact solution is available for this case.

linear system (16) may be determined. Plugging the result into the two constraints results in two equations for λ_1 and λ_2 from which $I(v)$ is determined via (17). For the purpose of illustration, we focus on the simple case of constant upward and downward rates given by $W_{i+1,i} = k_+$ and $W_{i-1,i} = k_-$ for all i (with however boundary rates $W_{N+1,N} = W_{1,0} = 0$), and consider the simple linear shear flow profile

$$u_i = -1 + \frac{2}{N} \left(i - \frac{1}{2} \right). \quad (23)$$

Fig. 3 demonstrates the accuracy of the large deviation approach also for this system. As in Fig. 2 the ldf compares very well with the simulation results and is clearly superior to the Gaussian approximation. The failure of the latter is again particularly prominent for asymmetric transition rates $k_- \neq k_+$. This situation models a horizontal layered system in which Taylor dispersion is affected by sedimentation. Furthermore, the non-Gaussian deviations become stronger when the number of layers N increases. This can be explained by the fact that particles have more options to get stuck in slow or fast lanes.

IV. CONTINUOUS FLOW FIELDS

We now return to the case of a continuous variable \mathbf{y} instead of the discrete index i . If $\mathbf{y}(t)$ represents a stochastic process with detailed balance we may again build on explicit expressions for the ldf of the empirical density. More precisely, let the dynamics of \mathbf{y} be given by the Stratonovich Langevin equation

$$\partial_t \mathbf{y}(t) = -\nabla U(\mathbf{y}) + \sqrt{2D(\mathbf{y})} \boldsymbol{\xi}(t) \quad (24)$$

with a deterministic drift derived from a potential $U(\mathbf{y})$ and a Gaussian noise source with zero mean and correlations $\langle \xi_i(t) \xi_j(t') \rangle = \delta_{i,j} \delta(t - t')$. The large deviation functional for the empirical density (cf. (7))

$$q(\mathbf{y}) := \frac{1}{t} \int_0^t \delta(\mathbf{y} - \mathbf{y}(\tau)) d\tau \quad (25)$$

is then given by [6]:

$$J[q(\cdot)] = \int D(\mathbf{y}) p(\mathbf{y}) \left(\nabla \sqrt{\frac{q(\mathbf{y})}{p(\mathbf{y})}} \right)^2 d\mathbf{y}, \quad (26)$$

where, as before, $p(\mathbf{y})$ denotes the stationary distribution of the process (24). In appendix B we include a simple derivation of this expression from (14) for the special case of a biased random walk in one dimension.

Starting with (26) the contraction to $I(v)$ is now defined by the variational problem

$$I(v) = \inf_{q(\cdot)} \int p(\mathbf{y}) D(\mathbf{y}) \left(\nabla \sqrt{\frac{q(\mathbf{y})}{p(\mathbf{y})}} \right)^2 d\mathbf{y} \quad (27)$$

under the constraints

$$\int q(\mathbf{y}) d\mathbf{y} = 1 \quad \text{and} \quad \int q(\mathbf{y}) u(\mathbf{y}) d\mathbf{y} = v. \quad (28)$$

It is convenient to minimize in $\hat{q}(\mathbf{y}) := \sqrt{q(\mathbf{y})}$. Introducing as before the Lagrange multipliers λ_1 and λ_2 the corresponding Euler-Lagrange equation reads

$$0 = -\frac{1}{\sqrt{p}} \nabla \left(p D \nabla \left(\frac{\hat{q}}{\sqrt{p}} \right) \right) + (\lambda_1 + \lambda_2 u) \hat{q} \quad (29)$$

where at the boundary the normal component of

$$\sqrt{p} D \nabla \left(\frac{\hat{q}}{\sqrt{p}} \right) \quad (30)$$

has to vanish. Multiplying (29) by \hat{q} , integrating over \mathbf{y} , and using the boundary condition (30) gives back (17). Eqs. (29) and (30) form a Sturm-Liouville problem, its concrete solution depends on the flow profile $u(\mathbf{y})$.

Note that for $D(y)$ independent of y , which is the case in many physical applications, the large deviation function becomes proportional to D . For the original problem of Taylor dispersion in a Poiseuille flow (2) through a cylindrical tube one can, by a choice of the units of time and space, furthermore set $a = 1$ and $\bar{u} = 1$. Hence the corresponding large deviation function $I(v)$ divided by D reduces to a single master function, which we proceed to calculate below.

Starting with $p(\mathbf{y}) = \text{const.}$ and $u(\mathbf{y}) = 2(1 - y^2)$ (29) acquires the form

$$0 = -\Delta \hat{q} + \left(\lambda_1 + 2\lambda_2 (1 - y^2) \right) \hat{q}. \quad (31)$$

Due to rotational symmetry the solution of this equation will depend only on $r = |\mathbf{y}|$ and it is sufficient to keep the radial part of the Laplacian. We hence have to solve

$$0 = \partial_r^2 \hat{q}(r) + \frac{1}{r} \partial_r \hat{q}(r) - (\lambda_1 + 2\lambda_2 (1 - r^2)) \hat{q} \quad (32)$$

together with the boundary conditions

$$r \partial_r \hat{q}(r)|_{r=0} = 0 \quad \text{and} \quad \partial_r \hat{q}(r)|_{r=1} = 0. \quad (33)$$

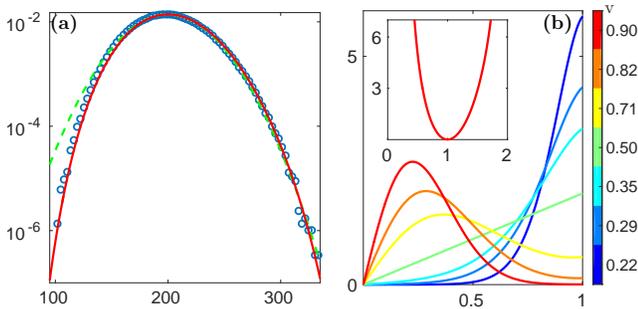


Figure 4. a) Logarithmic plot of the displacement distribution $P(x,t)$ at $t = 200$ for a Poiseuille flow (2) with $a = 1$ and $\bar{u} = 1$ and $D = 0.01$. Shown are the large deviation result (red), the Gaussian approximation (dashed green), and the outcome of numerical simulations (blue circles). b) Plot of $\tilde{q}(r)$ for different values of v identifying regions of the tube cross section that dominate the various values of v . The inset displays the function $I(v)/D$ that universally characterizes the large deviation properties of Taylor dispersion in Poiseuille flow.

This eigenvalue problem was also obtained in [10] (their Eq. (3.28)). The auxiliary function $\Phi(r)$ introduced in this paper is hence nothing but the square root $\tilde{q}(r)$ of the empirical density.

The general solution of (32) may be written as

$$\tilde{q}(r) = C_1 e^{-\frac{\sqrt{B}r^2}{2}} M\left(A, 1, \sqrt{B}r^2\right) + C_2 e^{-\frac{\sqrt{B}r^2}{2}} U\left(A, 1, \sqrt{B}r^2\right) \quad (34)$$

with the Kummer functions $M(a, b, z)$ and $U(a, b, z)$ [20], two constants C_1 and C_2 as well as

$$A = \frac{1}{2} \left(1 - \frac{i}{2} \frac{\lambda_1 + 2\lambda_2}{\sqrt{2\lambda_2}} \right) \quad \text{and} \quad B = -2\lambda_2. \quad (35)$$

The first part of (34) involving $M(a, b, z)$ obeys the first boundary condition, at $r = 0$, automatically. Because of the behaviour of $U(a, 1, z)$ for small z this boundary condition cannot be fulfilled by the second part of (34) implying $C_2 = 0$. The second boundary condition, at $r = 1$, and the two constraints (28) then form a system of equations to determine C_1 , λ_1 , and λ_2 . From the latter two $I(v)/D$ follows using (17).

In Fig. 4a we show the ldf obtained in this way together with the Gaussian approximation and numerical simulations. Again, the ldf compares very well with the numerical results. In particular the asymmetry of the distribution, which in this case is due to the circular geometry of the setup, and which is clearly beyond the capabilities of the Gaussian approximation is accurately reproduced. Fig. 4b displays the function $\tilde{q}(r)$ for different values of v that highlights which regions in the cross section of the pipe contribute most to the respective large deviations of the sample velocity. As intuitive small velocities require long residence of the particles near the wall whereas large values of v are realized by particles

staying most of the time near the centre of the tube. The inset shows the universal function $I(v)/D$ for Taylor dispersion in Poiseuille flow. Note that both curves are independent of t .

V. DISCUSSION

We discussed the connection between the large deviation properties of Taylor dispersion and the empirical measure of the underlying stochastic process. The large deviation function $I(v)$ provides an accurate long-time characterization of particle separation in shear flows. In particular for biased transport perpendicular to the flow direction the Gaussian approximation was found to yield a rather poor description only whereas the large deviation function reproduces the actual distribution down to very small probabilities. We conclude with a discussion about a converse issue, namely the interest and relevance of Taylor dispersion for the theory of large deviations. First, results from Taylor dispersion can feedback into the theory of empirical distributions. An example is the exact solution for the dispersion in a general nearest neighbour random walk [12], which can be translated back into an exact expression for the correlations of the empirical distribution. Second, we note that the ldf $I(v)$, albeit obtained via contraction from the ldf $J(\mathbf{q})$, is in fact a functional of the velocity field \mathbf{u} . Hence, if we know $I_{\mathbf{u}}(v)$ for a suitable chosen set of flow profiles \mathbf{u} we can in principle reconstruct $J(\mathbf{q})$. Third, as far as experiments are concerned, Taylor dispersion provides a rather direct way to measure the empirical distribution. If the transition matrix governing the \mathbf{y} -dynamics can be modified experimentally, e.g., by applying appropriate biasing fields between the states, the results of this "tilt" will be promptly seen in the particle dispersion along the flow.

Appendix A: Expressions and bounds for the Taylor dispersion coefficient K

Here we give a short account on how our results for the large deviation function $I(v)$ of Taylor dispersion relates to results for the Taylor dispersion coefficient K derived previously. Since K describes the Gaussian long time regime of the process it is determined by the second derivative of $I(v)$ at $v = \bar{u}$:

$$\frac{1}{2K} = \frac{d^2 I}{dv^2}(\bar{u}) = \frac{d^2 J}{d\mathbf{q}^2}(\mathbf{p}), \quad (A1)$$

where $J(\mathbf{q})$ is given by (cf. (14))

$$J(\mathbf{q}) = \frac{1}{2} \sum_{i,j} \left(\sqrt{W_{i,j} q_j} - \sqrt{W_{j,i} q_i} \right)^2, \quad (A2)$$

and $\bar{\mathbf{q}}(v)$ is determined by

$$\frac{\partial J}{\partial q_i}(\bar{\mathbf{q}}) = -(\lambda_1 + \lambda_2 u_i). \quad (A3)$$

Then, since $I(v) = J(\bar{\mathbf{q}}(v))$,

$$\frac{d^2 I}{dv^2} = \frac{d^2 J(\bar{\mathbf{q}})}{dv^2} = \sum_{i,j} \frac{\partial^2 J}{\partial q_i \partial q_j}(\bar{\mathbf{q}}) \frac{d\bar{q}_i}{dv} \frac{d\bar{q}_j}{dv} + \sum_i \frac{\partial J}{\partial q_i}(\bar{\mathbf{q}}) \frac{d^2 \bar{q}_i}{dv^2}. \quad (\text{A4})$$

From the constraints $\mathbf{1} \cdot \bar{\mathbf{q}} = 1$ and $\mathbf{u} \cdot \bar{\mathbf{q}} = v$ we find by differentiation with respect to v

$$\mathbf{1} \cdot \frac{d\bar{\mathbf{q}}}{dv} = 0, \quad \mathbf{u} \cdot \frac{d\bar{\mathbf{q}}}{dv} = 1, \quad \mathbf{1} \cdot \frac{d^2 \bar{\mathbf{q}}}{dv^2} = \mathbf{u} \cdot \frac{d^2 \bar{\mathbf{q}}}{dv^2} = 0. \quad (\text{A5})$$

Hence, together with (A3), it follows

$$\sum_i \frac{\partial J}{\partial q_i}(\bar{\mathbf{q}}) \frac{d^2 \bar{q}_i}{dv^2} = - \sum_i (\lambda_1 + \lambda_2 u_i) \frac{d^2 \bar{q}_i}{dv^2} = 0 \quad (\text{A6})$$

and the second term in (A4) vanishes. Therefore

$$\frac{1}{2K} = \sum_{i,j} \frac{\partial^2 J}{\partial q_i \partial q_j}(\mathbf{p}) \frac{d\bar{q}_i}{dv}(\bar{u}) \frac{d\bar{q}_j}{dv}(\bar{u}). \quad (\text{A7})$$

Using the detailed balance condition $W_{i,j} p_j = W_{j,i} p_i$ we find from (A2)

$$\frac{\partial^2 J}{\partial q_i \partial q_j}(\mathbf{p}) = -\frac{1}{2} \frac{W_{i,j}}{p_i} = -\frac{1}{2} \frac{W_{j,i}}{p_j}, \quad (\text{A8})$$

and therefore

$$\frac{1}{K} = - \sum_{i,j} \frac{W_{i,j}}{p_i} \frac{d\bar{q}_i}{dv}(\bar{u}) \frac{d\bar{q}_j}{dv}(\bar{u}). \quad (\text{A9})$$

To determine $d\bar{q}_i/dv$ at $v = \bar{u}$ we take the derivative of (A3) with respect to v to find

$$\sum_j \frac{\partial^2 J}{\partial q_i \partial q_j}(\bar{\mathbf{q}}) \frac{\partial \bar{q}_j}{dv} = -\frac{d\lambda_1}{dv} - \frac{d\lambda_2}{dv} u_i, \quad i \in \mathbf{I}. \quad (\text{A10})$$

Since J is a homogeneous function of \mathbf{q} , $J(\kappa \mathbf{q}) = \kappa J(\mathbf{q})$, we have

$$J(\mathbf{q}) = \sum_i \frac{\partial J}{\partial q_i} q_i \quad (\text{A11})$$

and, for $\mathbf{q} = \bar{\mathbf{q}}$,

$$J(\bar{\mathbf{q}}) = - \sum_i (\lambda_1 + \lambda_2 u_i) \bar{q}_i = -\lambda_1 - \lambda_2 v, \quad (\text{A12})$$

reproducing (17). This implies

$$\frac{dJ(\bar{\mathbf{q}})}{dv} = -\frac{d\lambda_1}{dv} - \frac{d\lambda_2}{dv} v - \lambda_2. \quad (\text{A13})$$

On the other hand, multiplying (A3) by $\partial \bar{q}_i / \partial v$ and summing over i gives

$$\frac{dJ(\bar{\mathbf{q}})}{dv} = -\lambda_2(v) \quad (\text{A14})$$

such that

$$\frac{d\lambda_1}{dv} = -\frac{d\lambda_2}{dv} v. \quad (\text{A15})$$

Therefore

$$-\frac{d\lambda_2}{dv}(\bar{u}) = \frac{1}{2K} \quad \text{and} \quad \frac{d\lambda_1}{dv}(\bar{u}) = \frac{1}{2K} \bar{u}. \quad (\text{A16})$$

From (A10) we hence obtain

$$\sum_j \frac{\partial^2 J}{\partial q_i \partial q_j}(\mathbf{p}) \frac{d\bar{q}_j}{dv} = \frac{\bar{u} - u_i}{2K}. \quad (\text{A17})$$

or, using (A8),

$$\sum_j W_{i,j} \frac{d\bar{q}_j}{dv}(\bar{u}) = p_i \frac{\bar{u} - u_i}{K}. \quad (\text{A18})$$

This is a system of linear inhomogeneous equations to determine $d\bar{q}_i/dv$ at $v = \bar{u}$ in terms of \mathbf{p} and \mathbf{u} . The matrix $W_{i,j}$ is singular, however, by definition of \bar{u} , the r.h.s of (A18) is orthogonal to the left eigenvector $(1, 1, \dots, 1)$ corresponding to the zero eigenvalue and the system is hence solvable. We denote by $\tilde{W}_{i,j}^{-1}$ the inverse of $W_{i,j}$ on the subspace orthogonal to the zero eigenvector to get

$$\frac{d\bar{q}_i}{dv}(\bar{u}) = \frac{1}{2K} \sum_k \tilde{W}_{i,k}^{-1} p_k (\bar{u} - u_k). \quad (\text{A19})$$

Plugging this result into (A9) we find the well-known expression [13]

$$K = - \sum_{i,j} \tilde{W}_{i,j}^{-1} p_j (\bar{u} - u_i)(\bar{u} - u_j). \quad (\text{A20})$$

In addition to this exact expression for K the variational form of the contraction (11) forms a convenient starting point to derive bounds on K . From (A20) we see that a Galilei transformation $u_i \rightarrow u_i + u$ does not change K . In the following we therefore assume for simplicity $\bar{u} = 0$. Small deviations v from $\bar{u} = 0$ result from small deviations of $\bar{\mathbf{q}}$ from \mathbf{p} . If we parametrize these deviations by $q_i = p_i(1 + \varepsilon_i v)$ and use the detailed balance condition (13) we find

$$J(\bar{\mathbf{q}}) \cong \frac{v^2}{8} \sum_{i,j} W_{i,j} p_j (\varepsilon_i - \varepsilon_j)^2, \quad (\text{A21})$$

which, using (A1) and (11), yields

$$\frac{1}{K} = \inf_{\boldsymbol{\varepsilon}} \frac{1}{2} \sum_{i,j} W_{i,j} p_j (\varepsilon_i - \varepsilon_j)^2, \quad (\text{A22})$$

Here the infimum has to be taken under the constraints $\sum_i \varepsilon_i q_i = 0$ and $\sum_i \varepsilon_i q_i u_i = 1$.

As an interesting trial ansatz, we set $\varepsilon_i = \alpha u_i$. Then the first constraint is automatically satisfied while the second implies

$$\frac{1}{\alpha} = \sum_i u_i^2 p_i = \bar{u}^2. \quad (\text{A23})$$

This gives

$$\frac{1}{K} \leq \frac{1}{2\bar{u}^2} \sum_{i,j} W_{i,j} p_j (u_i - u_j)^2. \quad (\text{A24})$$

The equality sign is realized for a Kangaroo process defined by $W_{i,j} = p_i/\tau$, where τ denotes the overall relaxation time. Indeed one finds in this case that the infimum in (A22) is realized for $\varepsilon_i = u_i/\bar{u}^2$. Hence for a Kangaroo process with $\bar{u} = 0$ we get the simple result

$$K = \frac{2\tau\bar{u}^2}{\sum_{i,j} p_i p_j (u_i - u_j)^2} = \tau\bar{u}^2. \quad (\text{A25})$$

Finally, using $\sum_i W_{i,j} = 0$ we may rewrite (A22) in the form

$$\begin{aligned} \frac{1}{K} &= -\inf_{\boldsymbol{\varepsilon}} \frac{1}{2} \sum_{i,j} W_{i,j} p_j \varepsilon_i \varepsilon_j \\ &= -\inf_{\boldsymbol{\varepsilon}} \frac{1}{2} \sum_{i,j} \varepsilon_i p_i W_{i,j} \varepsilon_j p_j / p_i \\ &= -\inf_{\mathbf{f}} \langle \mathbf{f} | W | \mathbf{f} \rangle, \end{aligned} \quad (\text{A26})$$

where the infimum is now over all vectors \mathbf{f} orthogonal to the stationary distribution $\langle \mathbf{f} | \mathbf{p} \rangle = 0$ and fulfilling $\langle \mathbf{f} | \mathbf{u} \rangle = 1$. Here, the scalar product is defined by

$$\langle \mathbf{f} | \mathbf{g} \rangle = \sum_i \frac{f_i g_i}{p_i}. \quad (\text{A27})$$

The bound

$$K \geq -\frac{1}{\langle \mathbf{f} | W | \mathbf{f} \rangle} \quad (\text{A28})$$

resulting from (A26) was already derived in [13].

Appendix B: Continuum limit of the ldf for a biased random walk

Consider a biased nearest neighbour random walk in one dimension with reflecting boundary conditions. The probability $P_i(t)$ of the walker being at site $i \in \mathbf{I}$ at time t obeys a Master equation

$$\partial_t P_i(t) = \sum_j \left(W_{i,j} P_j(t) - W_{j,i} P_i(t) \right) \quad (\text{B1})$$

with

$$W_{i,j} = \begin{cases} k_{i-1}^+ & \text{for } j = i - 1 \\ -(k_i^+ + k_i^-) & \text{for } j = i \\ k_{i+1}^- & \text{for } j = i + 1 \\ 0 & \text{else} \end{cases} \quad (\text{B2})$$

The stationary distribution p_i fulfills detailed balance $k_i^+ p_i = k_{i+1}^- p_{i+1}$ automatically. Using this condition we find from (A2)

$$J(\mathbf{q}) = \sum_i k_{i-1}^+ p_{i-1} \left(\sqrt{\frac{q_{i-1}}{p_{i-1}}} - \sqrt{\frac{q_i}{p_i}} \right)^2. \quad (\text{B3})$$

The continuum limit is obtained from this expression by introducing a small space interval η and using the replacements

$$\begin{aligned} i &\rightarrow \frac{y}{\eta}, \quad p_i \rightarrow \eta p(y), \quad q_i \rightarrow \eta q(y) \\ \sum_i &\rightarrow \int \frac{dy}{\eta}, \quad J(\mathbf{q}) \rightarrow J[q(\cdot)] \\ k_i^\pm &\rightarrow \pm \frac{w(y)}{2\eta} + \frac{D(y)}{\eta^2}. \end{aligned} \quad (\text{B4})$$

The probability density function $P(y, t)$ then fulfills the Fokker-Planck equation

$$\partial_t P(y, t) = -\partial_y \left((w(y) - \partial_y D(y)) P(y, t) \right) + \partial_y \left(D(y) \partial_y P(y, t) \right) \quad (\text{B5})$$

complemented by reflecting boundary conditions. The large deviation functional acquires the form

$$J[q(\cdot)] = \int p(y) D(y) \left(\partial_y \sqrt{\frac{q(y)}{p(y)}} \right)^2 dy \quad (\text{B6})$$

corresponding to (26).

[1] G. I. Taylor, Proc. R. Soc. A **219**, 168-203 (1953).

[2] J.W. Westhaver, Natl. Bur. Stand. **38**, 1766 (1947).

[3] <http://www.malvern.com/en/products/technology/taylor-dispersion-analysis/>

[4] H. Brenner and D. A. Edwards, *Macrotransport Processes*, (Butterworth-Heinemann, Stoneham, 1993)

[5] C. Van den Broeck, Physica A **168**, 677-696 (1990).

[6] M. D. Donsker, and S. R. S. Varadhan, Comm. Pure Appl. Math. **28**, 1-47 (1975); **28**, 279-301 (1975); **29**, 389-461 (1976).

[7] H. Touchette, Physics Reports **478**, 1-69 (2009).

[8] B. Derrida, JSTAT P01030 (2011)

[9] A. Vulpiani, F. Cecconi, M. Cencini, A. Puglisi, D. Vergni (eds.), *Large Deviations in Physics: The Legacy of the*

- Law of Large Numbers*, (Springer, Heidelberg, 2014)
- [10] P. H. Haynes and J. Vanneste *J. Fluid Mech.* **745**, 321 (2014)
- [11] G. R. Shorack and J. A. Wellner, *Empirical Processes with Applications to Statistics*, (Wiley, New York, 1986)
- [12] C. Van den Broeck and R. M. Mazo, *Phys. Rev. Lett.* **51**, 1309 (1983), *J. Chem. Phys.* **81**, 3624 (1984)
- [13] C. Van den Broeck, and D. Maes, and M. Bouten, *Phys. Rev. A* **36**, 5025–5030 (1987).
- [14] S. Goldstein, *Quart. J. Mech. Appl. Math.* **4**, 129–156 (1951).
- [15] J. C. Giddings, and H. Eyring, *J. Phys. Chem.* **59**, 416–421 (1955).
- [16] K. Mysels, *J. Phys. Chem.* **24**, 371 (1956).
- [17] H. P. McKean, *J. Math. Phys.* **8**, 547 (1967).
- [18] M. Kac, *Rocky Mountain J. Math.* **4**, 497 (1974).
- [19] W. C. Thacker, *J. Phys. Ocean.* **6**, 66 (1975).
- [20] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, (National Bureau of Standards, Washington D.C., 1972)