Flooding dynamics of diffusive dispersion in a random potential

Michael Wilkinson[1,2] · Marc Pradas[2] · Gerhard Kling[3]

Abstract

We discuss the combined effects of overdamped motion in a quenched random potential and diffusion, in one dimension, in the limit where the diffusion coefficient is small. Our analysis considers the statistics of the mean first-passage time $T(x)$ to reach position x, arising from different realisations of the random potential. Specifically, we contrast the median $\bar{T}(x)$, which is an informative description of the typical course of the motion, with the expectation value $\langle T(x) \rangle$, which is dominated by rare events where there is an exceptionally high barrier to diffusion. We show that at relatively short times the median $\bar{T}(x)$ is explained by a 'flooding' model, where $T(x)$ is predominantly determined by the highest barriers which are encountered before reaching position x. These highest barriers are quantified using methods of extreme value statistics.

Keywords Diffusion · Ornstein-Uhlenbeck process

1 Introduction

There are many situations where particles move under the combined influence of thermal diffusion and a static (or quenched) random potential [1]. The particles might be electrons, holes or excitons diffusing in a disordered metallic or semiconductor sample [2], or molecules diffusing in a complex environment such as the cytoplasm of a eukaryotic cell [3]. Some

✉ Michael Wilkinson
   michael.wilkinson@czbiohub.org

   Marc Pradas
   marc.pradas@open.ac.uk

   Gerhard Kling
   gerhard.kling@abdn.ac.uk

1 Chan Zuckerberg Biohub, 499 Illinois Street, San Francisco, CA 94158, USA

2 School of Mathematics and Statistics, The Open University, Walton Hall, Milton Keynes MK7 6AA, England

3 Business School, King's College, University of Aberdeen, Aberdeen AB24 3FX, Scotland

significant aspects of the state of knowledge of this problem are surprisingly under-developed, and in this work we present new results on the simplest version of this problem, in one dimension, where the equation of motion is

$$\dot{x} = -\frac{dV}{dx} + \sqrt{2D}\,\eta(t)\,. \tag{1}$$

Here $V(x)$ is a random potential, D is the diffusion coefficient, and $\eta(t)$ is a white noise signal with statistics defined by

$$\langle \eta(t) \rangle = 0\,, \quad \langle \eta(t)\eta(t') \rangle = \delta(t - t') \tag{2}$$

($\langle \cdot \rangle$ denotes expectation value throughout). We assume that $V(x)$ is a smooth random function, defined by its statistical properties, which are stationary in x, and independent of the temporal white noise $\eta(t)$. The one and two-point statistics of this potential are

$$\langle V(x) \rangle = 0\,, \quad \langle V(x)V(x') \rangle = C(x - x')\,. \tag{3}$$

The correlation function $C(\Delta x)$ is assumed to be smooth and to decay rapidly as $|\Delta x| \to \infty$. We also assume that the tails of the distribution of V are characterised by a large-deviation 'rate' (or 'entropy') function $J(V)$, so that when $|V|$ is large, the probability density function of V is approximated by

$$P_V \sim \exp[-J(V)] \tag{4}$$

where throughout we shall use P_X to denote the probability density function (PDF) of a random variable X. If P_V is a Gaussian distribution, then the entropy function is quadratic, $J(V) \sim V^2/2C(0)$.

A cluster of particles which is released from $x = 0$ at time $t = 0$ are considered to evolve under different realisations of the temporal noise $\eta(t)$, with the potential $V(x)$ frozen. We use the term *dispersion* of the particles in a statistical sense, which may be described by specifying the probability density $p(x, t)$ of x at a given time t. This probability density function (PDF) is a functional of the potential $V(x)$, and we wish to understand averages over the function $V(x)$. It has been proposed that the behaviour of this system is diffusive (in a sense which will be made well-defined in Sect. 2.1 below). When $V(x)$ has a Gaussian distribution, Zwanzig [4] gave an elegant argument which implies that the effective diffusion coefficient vanishes very rapidly as $D \to 0$:

$$D_{\text{eff}} \sim D \exp\left[-\frac{C(0)}{D^2}\right]\,. \tag{5}$$

An earlier work by De Gennes [5] proposes a similar expression. We discuss the origin of this result, and present a generalisation of it to non-Gaussian distributions, in Sect. 2. When D is small, this estimate for the diffusion coefficient depends upon rare events where the potential is unusually large, and it is very difficult to verify equation (5) numerically. In addition, numerical experiments show that at relatively early times, the model exhibits subdiffusive behaviour and it has been suggested that there is a regime of anomalous diffusion, in the sense that $\langle x^2 \rangle \sim t^\alpha$, with $0 < \alpha < 1$ [6–8]. It is desirable to achieve an analytical understanding of the sub-diffusive behaviour which is observed in numerical simulations of (1).

We should mention that there are also exact results [9–11] on a closely related model (motion in a quenched velocity field, which is *not* the derivative of a potential with a well-defined probability distribution) showing that $\langle x^2 \rangle \sim (\ln t)^4$. This 'Sinai diffusion' process is fundamentally different, because the particle becomes trapped in successively deeper minima of the potential, from which it takes ever increasing time intervals to escape. The Sinai model

has been extended to consider a periodised random potential, so that the long-time dynamics is diffusive: in this case it is found [12] that the diffusion coefficient has a very broad probability distribution, and these studies have inspired other works on dispersion in the presence of traps, of which [13] is a recent example. Another related problem is the motion of a *quantum* particle in a statistically homogeneous random potential which exhibits even slower spatial dispersion: it may be completely localised [14], and in one or two dimensions localisation is almost certain [15,16]: see [17] for a review.

Very little can be deduced from the equation of motion (1) directly. Because the equation of motion contains random elements, we must characterise the solution statistically. We could try to estimate the probability distribution $p(x, t)$ of position x at a given value of time t, and hence compute the variance $\langle x^2 \rangle$. We shall adopt an alternative approach, however, considering the statistics of the mean time T taken for first passage of a marker at distance x from the starting point (averaging with respect to the noise, $\eta(t)$). We will argue that, while equation (5) and its generalisation to non-Gaussian distributions describes the long-time asymptote of the spatial dispersion of particles, the diffusive behaviour only emerges at very long times. At intermediate times, the dynamics of typical realisations is not diffusive. We show that it is determined by the time taken to diffuse across the largest potential barrier which must be traversed to reach position x. The diffusion process is able to traverse a barrier of height ΔV after a characteristic time $T \sim \exp(\Delta V / D)$ [18], and as time increases the height of the barriers which can be breached, leading to 'flooding' of the region beyond, increases. According to this picture, the dispersion distance x is determined by a problem in extreme value statistics: how large must x be before we reach a barrier of height $\Delta V \approx D \ln t$? Figure 4(b) from our numerical results in Sect. 5 gives a graphic illustration of this flooding process in action. The figure plots $T(x)$ against x on a logarithmic scale, for a fixed realisation of the potential. As T increases, the region which is accessible increases in steps, corresponding to potential barriers becoming penetrable as the time is increased.

Zwanzig's calculation [4] considers a double-average of the first-passage time: if $T(x)$ is the time for first passage of x, averaged over the noise η, Zwanzig considers the mean value $\langle T(x) \rangle$ averaged over the realisations of the potential. We argue that this approach does not give a reliable indication of the typical motion because the mean is dominated by realisations where $T(x)$ is unusually large, as a consequence of an unusually large potential barrier. Hence, we propose to characterise the dispersion by evaluating the median $\bar{T}(x)$ (that is, the value of $T(x)$ which is exceeded for exactly half of the realisations of $V(x)$). The median is, by definition, a good description of a typical realisation. We show that the median $\bar{T}$ (with respect to different realisations of the potential $V(x)$) of the mean-first-passage-time (averaged over $\eta(t)$) satisfies

$$\ln \left(\frac{x}{\tilde{x}} \right) \sim J(D \ln \bar{T}/2) \tag{6}$$

where $\tilde{x}$ is a lengthscale which characterises the typical distance between extrema of the potential. In the case where the potential has a Gaussian distribution, this implies that the dispersion is subdiffusive, satisfying

$$\ln \left(\frac{x}{\tilde{x}} \right) \sim \frac{D^2 (\ln \bar{T})^2}{8C(0)} \tag{7}$$

which is quite distinct from the usual anomalous diffusion behaviour, characterised by power-laws such as $\langle x^2 \rangle \sim t^\alpha$. After a sufficiently long time, the dynamics becomes diffusive, with a diffusion coefficient given by (5).

Our arguments will depend upon making estimates of sums of the form

$$S_N = \sum_{j=1}^{N} \exp(f_j/\epsilon) \tag{8}$$

where f_j are independent identically distributed (i.i.d.) random variables, and ϵ is a small parameter, which we identify with the diffusion coefficient D. We refer to S_N as 'extreme-weighted sums', because the largest values of f_j make a dominant contribution to S_N as $\epsilon \to 0$. In Sect. 2 we show how the mean-first-passage time is related to sums like (8), and in Sect. 3 we analyse some of their statistics, which are used in Sect. 4 to justify our principle result, equation (6). Section 5 describes our numerical investigations, and Sect. 6 summarises our findings.

2 The Mean First Passage Time

Our discussion of the dynamics of (1) will focus on the mean first passage problem: what is the mean time $T(x)$ at which a particle released from the origin reaches position x. First passage problems are discussed comprehensively in the book by Redner [19]. The result that we require can be found in multiple sources: it is based on the method introduced by Pontryagin *et al.* [20] and has been applied in several contexts, for example in [21] to describe self-diffusion of ions in a polyelectrolyte solution. The key formula, equation (9) below, was already applied to equation (1) by Zwanzig [4].

In this section we first quote the general formula for the mean first passage time $T(x)$, as a functional of the potential $V(x)$. In order to simplify the discussion, we consider the case where there is a reflecting barrier at the origin, $x = 0$. If particles are released at $x_0 = 0$, the mean first passage time to reach position x is given by (see, e.g., [22])

$$T(x) = \frac{1}{D} \int_0^x dy \ \exp[V(y)/D] \int_0^y dz \ \exp[-V(z)/D] \tag{9}$$

where the averaging is with respect to realisations of the noise $\eta(t)$ in the equation of motion (1), with $V(x)$ frozen, so that $T(x)$ is a functional of $V(x)$.

We then (Sect. 2.1) discuss the result obtained by Zwanzig [4] for the expectation value $\langle T(x) \rangle$ (averaged with respect to realisations of $V(x)$). Zwanzig gave the result for a potential with Gaussian fluctuations, which we extend to the case of a general form for the large-deviation entropy function (as defined by equation (4)). The result obtained by Zwanzig suggests that the dispersion is diffusive, with a diffusion coefficient D_{eff} which vanishes in a highly singular fashion as $D \to 0$. We shall argue that this result is a consequence of the expectation value of $T(x)$ being dominated by very rare large excursions of the potential $V(x)$, and that for typical realisations of the potential the dispersion is much more rapid than the value of $\langle T(x) \rangle$ suggests. This requires a more delicate analysis of the structure of the integrals in the expression for $T(x)$, equation (9). In Sect. 2.2, we discuss how these integrals may be approximated by sums, analogous to (8), in the limit as $D \to 0$.

2.1 Expression for Expectation Value of Mean First Passage Time

We can make an additional average of (9), with respect to different realisations of the potential, which leads to

$$\langle T(x) \rangle = \frac{1}{D} \int_0^x dy \int_0^y dz \left\langle \exp\left[\frac{V(y) - V(z)}{D} \right] \right\rangle$$
$$\sim \frac{x^2}{2D} \langle \exp(-V/D) \rangle \langle \exp(V/D) \rangle \qquad (10)$$

where in the second line we consider the leading order behaviour as $x \to \infty$ (where most of the contribution to the double integral is from values of x and y which are separated by a distance which is much larger than the correlation length of the potential, so that $V(x)$ and $V(y)$ can be treated as independent random variables). If the motion were simple diffusion, with $V = 0$, equation (9) would evaluate immediately to $\langle T \rangle = x^2/2D$, so that it is reasonable to identify $x^2/2\langle T \rangle$ as the effective diffusion coefficient. Hence, assuming that the PDF of $V(x)$ is symmetric between V and $-V$, we have

$$D_{\text{eff}} = \frac{D}{\left[\langle \exp(V/D) \rangle \right]^2} \ . \qquad (11)$$

The mean first passage time $T(x)$ must have a very broad distribution. It might be difficult to extract the diffusion coefficient from time series (see discussion in [23]) and it might be questioned whether the motion really is diffusive. In this regard, we note that the central limit theorem is applicable to the quantity $T(x)$ obtained by equation (9) in the limit as $x \to \infty$, so that $x^2/2T(x)$ does approach a limit, which we identify as the effective diffusion coefficient.

When D is small, $\langle \exp(V/D) \rangle$ is dominated by the tail of the PDF of V, and this expectation value is obtained using the Laplace method [24] which utilises the fact that as $\epsilon \to 0$ the integral of $\exp[f(x)/\epsilon]$ is well approximated using the Taylor series of f about its maximum. Using this approach, we obtain

$$\langle \exp(V/D) \rangle = \int_{-\infty}^{\infty} dV \, P_V \exp(V/D)$$
$$\sim \int_{-\infty}^{\infty} dV \exp[V/D - J(V)]$$
$$\sim \sqrt{\frac{2\pi}{J''(V^*)}} \exp[V^*/D - J(V^*)] \qquad (12)$$

where V^* is the stationary point of the exponent, satisfying

$$DJ'(V^*) = 1 \ . \qquad (13)$$

(We assume that P_V has a faster than exponential decay, so that $V - DJ(V)$ has a maximum.) From this we obtain

$$D_{\text{eff}} \sim \frac{DJ''(V^*)}{2\pi} \exp\left[2J(V^*) - \frac{2V^*}{D} \right] \ . \qquad (14)$$

In the Gaussian case, where

$$J(V) = \frac{V^2}{2C(0)} + \frac{1}{2} \ln(2\pi C(0)) \qquad (15)$$

equation (14) agrees with (5).

2.2 Summation Approximations

In order to understand the implications of equation (9), we should consider the behaviour of the integral

$$S(x) = \int_0^x \mathrm{d}y \ \exp[-V(y)/D] \tag{16}$$

in the limit as $D \to 0$. When D is small this quantity may be estimated (using the Laplace method to capture the contribution from each of the minima of the potential):

$$S(x) \sim \sum_{j=1}^N \sqrt{\frac{2\pi D}{V_j''^-}} \exp(-V_j^-/D) \equiv \sum_{j=1}^N \exp[-\tilde{V}_j^-/D] \tag{17}$$

where V_j^- are the values of the N minima between 0 and x, occurring at positions x_j^-, and where we have defined

$$\tilde{V}_j^- = V_j^- - \frac{D}{2} \ln\left(\frac{2\pi D}{|V_j''^-|}\right) . \tag{18}$$

Note that

$$T(x) = \frac{1}{D} \int_0^x \mathrm{d}y \ \exp[V(y)/D]S(y) \tag{19}$$

and consider how to estimate $T(x)$ in the limit as $D \to 0$. Note that $S(y)$ is determined by the values of the minima of $V(y)$ in the interval $[0, y]$, jumping by an amount $\exp[-\tilde{V}_j^-/D]$ at x_j^-. Similarly, if V_j^+ are local maxima of $V(x)$, occurring at positions x_j^+, then $T(x)$ jumps at local maxima. The evolution of $S(x)$ and $T(x)$ are therefore determined by a pair of coupled maps:

$$\begin{aligned}
S \to S' &= S + \exp[-\tilde{V}_j^-/D] &&\text{(at minima } x_j^-) \\
T \to T' &= T + \tfrac{1}{D} \exp[\tilde{V}_j^+/D]S &&\text{(at succeeding maximum } x_j^+)
\end{aligned} \tag{20}$$

where we have defined again $\tilde{V}_j^+ = V_j^+ + \frac{D}{2} \ln\left(\frac{2\pi D}{|V''^+|}\right)$. These equations are difficult to analyse in the general case, but in the next section we discuss an approach which can be used to treat the limit where D is small.

3 Statistics of Extreme-Weighted Sums

This section draws on extreme value statistics. We refer the reader to [25–27] for an introduction to extreme value statistics and applications in statistical mechanics, and we mention [28] as an example of recent work which uses extreme value statistics as a tool to analyse stochastic models for dispersion. We have seen that when D is small the integrals defining the mean first passage time are approximated by sums over extrema of the potential, as described by equation (17). Accordingly, we study properties of random sums of the form

$$S_N = \sum_{j=1}^N \exp(f_j/\epsilon) \tag{21}$$

where ϵ is a small parameter and where the f_j are random numbers drawn from a probability distribution function P_f. We also define the complementary cumulative distribution function (tail distribution) describing the probability for f_j being greater than f as $Q(f)$. In the case where f has a Gaussian distribution, Eq. (21) is a sum of log-normal distributed random variables. There is some earlier literature on these sums which shows very little overlap with our results, see [29] and references therein, also [30], which discusses a phase transition that arises in a limiting case.

We also consider sums of the form

$$T_N = \sum_{n=1}^{N} \exp(g_n/\epsilon) S_n \tag{22}$$

where g_j are i.i.d. random numbers drawn from the same distribution as the f_j. Equations (21) and (22) are models for the summation which approximate the integrals $S(x)$ and $T(x)$ defined by equations (16) and (19), respectively, where we have assumed that the potential V has a symmetric distribution. In the context of our analysis of $T(x)$, the random variables f_j and g_j are identified with extrema of the potential, $-\hat{V}_j^-$ and $+\hat{V}_j^+$ respectively. The sum (21) is, however, also an object which is of interest in its own right (with broader applications, such as those considered in [29,30]). When ϵ is sufficiently small, these sums are determined by the largest values of f_j and g_j, and for this reason we shall refer to S_N and T_N as extreme-weighted sums.

In this section we find it convenient to write the tail distribution function for f in the form

$$Q(f) = \exp[-\mathcal{J}(f)] \tag{23}$$

where $\mathcal{J}(f)$ is a large deviation rate function (note that this is defined in terms of the cumulative probability function, as opposed to the definition based upon probability density used in equation (4)).

We are interested in the asymptotic behaviour of statistics of the sums S_N and T_N for small ϵ and large N. The sums vary wildly in magnitude and the mean is dominated by the tail of the distribution of f. Unless N is sufficiently large, values of f_j which determine the mean are unlikely to be sampled. This suggests that it will be useful to characterise the distribution of the S_N by the median, rather than the mean. We denote the median of a quantity X by $\bar{X}$ and its expectation by $\langle X \rangle$.

We consider an estimate of the median of S_N in some detail in Sect. 3.1, leading to a quite precise formula, equation (42). However, the sum T_N is a model for the mean first passage time. Extending the approach, which yields equation (42), to encompass the more complex structure of the T_N sum would be very cumbersome. Accordingly, in Sect. 3.2 we simplify equation (42) by confining our attention to its leading-order behaviour as $\epsilon \to 0$ and $N \to \infty$, and extending these asymptotic formulae to describe the median of the T_N. Our estimate for the median of T_N is contained in equations (50) and (51), which are expressed in terms of logarithmic variables defined in (43).

3.1 Estimate of Median of S_N

The sum S_N may be well approximated by its largest term, which is

$$s_N = \exp(\hat{f}/\epsilon) \tag{24}$$

where $\hat{f}$ is the largest of the N realisations, f_j, with index $j = \hat{j}$. We write

$$S_N \equiv \exp[\hat{f}/\epsilon]F \equiv s_N F \tag{25}$$

where

$$F = 1 + \sum_{\substack{j=1 \\ j \neq \hat{j}}}^{N} \exp[-(\hat{f} - f_j)/\epsilon] . \tag{26}$$

If F is close to unity, we can estimate $\bar{S}_N$ by $\bar{s}_N$. Let us first estimate $\bar{s}_N$ and return to consider F later. Note that

$$\bar{s}_N = \exp(\bar{\hat{f}}/\epsilon) \equiv \exp(f^\dagger/\epsilon) \tag{27}$$

where we have defined $f^\dagger = \bar{\hat{f}}$ as the median of the largest value of N samples from the distribution of f. The value of $f^\dagger$ is determined by requiring that the probability that N samples of f are all less than $f^\dagger$ is equal to one half. This condition is:

$$\left[1 - Q(f^\dagger)\right]^N = \frac{1}{2} . \tag{28}$$

When $N \gg 1$, this is determined by the tails of the distribution, where $Q(f)$ is approximated using (23):

$$\exp\left[-N\exp[-\mathcal{J}(f^\dagger)]\right] = \frac{1}{2} \tag{29}$$

so that $f^\dagger$ satisfies

$$\ln N - \ln \ln 2 = \mathcal{J}(f^\dagger) . \tag{30}$$

An important special case is where f has a Gaussian distribution, so that in the case where $\langle f \rangle = 0$ and $\langle f^2 \rangle = 1$,

$$Q(f) = \frac{1}{\sqrt{2\pi}} \int_f^\infty dx \, \exp(-x^2/2) \sim \frac{1}{\sqrt{2\pi} f} \exp(-f^2/2) \tag{31}$$

implying that

$$\mathcal{J}(f) = \frac{f^2}{2} + \ln f + \frac{\ln(2\pi)}{2} \tag{32}$$

so that

$$\frac{f^{\dagger 2}}{2} + \ln f^\dagger = \ln N - \ln \ln 2 - \frac{\ln(2\pi)}{2} . \tag{33}$$

In the limit where N is extremely large, we can approximate $f^\dagger$ by

$$f^\dagger \sim \sqrt{2 \ln N} \tag{34}$$

and consequently the median of $\bar{s}_N$ is approximated by

$$\bar{s}_N \sim \exp\left(\frac{\sqrt{2 \ln N}}{\epsilon}\right) . \tag{35}$$

Next consider how to estimate the quantity F in equation (25), when $\epsilon \ll 1$. When $N \gg 1$, either F is close to unity or else it is the sum of a large number of terms which make a comparable contribution. The value of F depends upon $\hat{f}$. The f_j which contribute to F are i.i.d. random variables, each with a PDF which is the same as that of the general f_j, except that there is an upper cutoff at $\hat{f}$: the adjustment of the normalisation due to the loss of the tail, $f > \hat{f}$, can be neglected. Consider the PDF of f written as

$$P_f = \exp[-J(f)] \tag{36}$$

where $J(f)$ is a large deviation rate function. The expectation value of F is then obtained as follows

$$\exp(\hat{f}/\epsilon)[\langle F \rangle - 1] = (N-1) \int_{-\infty}^{\hat{f}} \mathrm{d}f \; P_f \exp(f/\epsilon)$$

$$\sim N \int_{-\infty}^{\hat{f}} \mathrm{d}f \; \exp[f/\epsilon - J(f)]$$

$$\sim \frac{N \exp[f^*/\epsilon - J(f^*)]}{\sqrt{J''(f^*)/2}} \int_{-\infty}^{\sqrt{J''(f^*)/2}(\hat{f}-f^*)} \mathrm{d}y \; \exp(-y^2) \tag{37}$$

where f^* satisfies

$$\epsilon J'(f^*) = 1 \; . \tag{38}$$

Noting that

$$\langle S_N \rangle = N \langle \exp(f/\epsilon) \rangle \sim \sqrt{2\pi} \frac{N \exp[f^*/\epsilon - J(f^*)]}{\sqrt{J''(f^*)}} \tag{39}$$

we have

$$\exp(\hat{f}/\epsilon)[\langle F \rangle - 1] \sim \langle S_N \rangle \frac{1}{2} \left[1 + \mathrm{erf} \left(\sqrt{\frac{J''(f^*)}{2}} (\hat{f} - f^*) \right) \right] \; . \tag{40}$$

Hence, we obtain a rather simple approximation for S_N, depending upon the extreme value $\hat{f}$ of the sample of N realisations of the f_j:

$$S_N \sim \exp(\hat{f}/\epsilon) + \langle S_N \rangle \frac{1}{2} \left[1 + \mathrm{erf} \left(\sqrt{\frac{J''(f^*)}{2}} (\hat{f} - f^*) \right) \right] \; . \tag{41}$$

The median of S_N is therefore approximated by

$$\bar{S}_N \sim \exp(f^\dagger/\epsilon) + \langle S_N \rangle \frac{1}{2} \left[1 + \mathrm{erf} \left(\sqrt{\frac{J''(f^*)}{2}} (f^\dagger - f^*) \right) \right] \tag{42}$$

where $f^\dagger$ is the solution of equation (33).

3.2 Interpretation and Generalisation to T_N

We shall see that equation (42) gives a quite precise approximation for the median $\bar{S}_N$, but it is not immediately clear when either of the two terms is dominant. In order to clarify the structure of equation (42), we consider an approximate form of the equation determining $f^\dagger$, and transform to logarithmic variables. As well as leading to a transparent understanding

of equation (42), this facilitates making an estimate for $\bar{T}_N$ in the limit where $\epsilon \ll 1$ and $N \gg 1$. We define

$$\eta = \ln N , \quad \sigma = \ln \bar{S}_N , \quad \tau = \ln \bar{T}_N . \tag{43}$$

Note that η and τ are logarithmic measures of, respectively, distance and time, so that a plot of η versus τ gives information about the dispersion due to the dynamics.

Let us consider the limit where the first term in (42), $\exp[f^\dagger/\epsilon]$, is dominant. Note that the condition (33), determining the extreme value of a set of N samples, can be approximated by the requirement that the PDF of f is approximately equal to $1/N$. That is $1 \sim N \exp[-J(f^\dagger)]$. For the purposes of considering the $N \to \infty$ and $\epsilon \to 0$ limit, we can therefore approximate $f^\dagger$ by a solution of the equation

$$J(f^\dagger) = \eta . \tag{44}$$

If the second term in equation (42) is negligible, as might be expected when $f^\dagger - f^* \ll 0$, equations (43) and (44) then yield a simple implicit equation for σ:

$$\eta = J(\epsilon\sigma) . \tag{45}$$

If $f^\dagger - f^* \gg 0$, and if the second term in (42) is dominant, then $\bar{S}_N \sim \langle S_N \rangle = N \langle \exp(f/\epsilon) \rangle$, and using the Laplace method [24], following the same approach as for equation (12) we find

$$\sigma = \eta + \frac{f^*}{\epsilon} - J(f^*) . \tag{46}$$

Note the (46) indicates that $\frac{\mathrm{d}\sigma}{\mathrm{d}\eta} = 1$. Let us compare this with the value of $\frac{\mathrm{d}\sigma}{\mathrm{d}\eta}$ obtained from (45), which predicts $\frac{\mathrm{d}\eta}{\mathrm{d}\sigma} = \epsilon J'(f)$. The approximation (45) therefore becomes sub-dominant when $1 = \epsilon J'(f)$, which is precisely the equation for f^*, equation (38). If we define η^* and σ^* by writing

$$\sigma^* = \frac{f^*}{\epsilon} , \quad \eta^* = J(\epsilon\sigma^*) \tag{47}$$

then assembling these results and definitions, the relationship between η and σ can be summarised in the following equation

$$\eta = \begin{cases} J(\epsilon\sigma) & 0 < \eta < \eta^* \\ \sigma - \sigma^* + \eta^* & \eta \geq \eta^* \end{cases} . \tag{48}$$

Note that $\eta(\sigma)$ and its first derivative are continuous functions. In the foregoing we defined $\bar{x}$ as the median of x, but it should be noted that our arguments will lead to equations (46) and (47) as $N \to \infty$ if $\bar{S}_N$ denotes any fixed percentile of S_N.

Thus far we have considered the behaviour of η as a function of σ rather than of τ, but it is the function $\eta(\tau)$ which describes the dynamics of the dispersion. Consider the form of the sum T_N defined in equation (22). When $\sigma < \sigma^*$, the value of S_N is almost always determined by $\hat{f}$, the largest value of f_j, and similarly, one of the factors $\exp(g_j/\epsilon)$ corresponding to $\hat{g}$, the largest of the g_j, will predominate over the others. In one half of realisations, those where $\hat{k}(N) > \hat{j}(N)$, the largest value of f_j contributes to the sum which is multiplied by $\exp(\hat{g}/\epsilon)$, and we have $T_N \sim \exp(\hat{g}/\epsilon)\exp(\hat{f}/\epsilon)$. In cases where $\hat{j} > \hat{k}$, T_N is expected to be small in comparison to this estimate. We conclude that T_N has a probability of approximately one quarter to exceed the median of $\exp[(\hat{f} + \hat{g})/\epsilon]$. Note that because $\hat{f}$ and $\hat{g}$ are extreme values of samples on N i.i.d. random variables, their probability distribution

has a very narrow support [26]. The median of $\hat{f} + \hat{g}$ is therefore well approximated by the sums of the medians, that is by $f^\dagger + g^\dagger$. Correspondingly, the median of $\exp[(\hat{f} + \hat{g})]/\epsilon$ is approximated by $\exp[(f^\dagger + g^\dagger)/\epsilon]$. If we now use the overbar to represent the upper *quartile* of the distribution of T_N, rather than the median, we have

$$\bar{T}_N \sim \exp(g^\dagger/\epsilon)\exp(f^\dagger/\epsilon) . \tag{49}$$

Using the assumption that the f_j and g_j have the same PDF, and noting that $\bar{S}_N \sim \exp(f^\dagger/\epsilon)$, we can conclude that $\bar{T}_N \sim \bar{S}_N^2$ and hence that $\tau = 2\sigma$. The equation describing the dispersion as a function of time is therefore

$$\eta = J(\epsilon\tau/2) , \quad \tau < \tau^* \tag{50}$$

where τ^* is determined by the condition that $\mathrm{d}\eta/\mathrm{d}\tau = \frac{1}{2}$ when $\tau = \tau^*$. When $\tau > \tau^*$, we have $\bar{T}_N \sim \langle T \rangle = N\langle\exp(f/\epsilon)\rangle^2$, implying that

$$\eta = J(\epsilon\tau^*/2) + \frac{\tau - \tau^*}{2} , \quad \tau > \tau^* . \tag{51}$$

Equations (50) and (51) are a description of the logarithm of the typical dispersion η as a function of the logarithm of the time, τ. Usually the function $J(V)$ has a quadratic behaviour for small values of V, so that the initial dispersion, described by (50), is sub-diffusive. The factor of one half in (51) indicates that the long-time limit is diffusive. Writing $\eta^2 \sim \tau + \ln D_{\mathrm{eff}}$, we see that the effective diffusion coefficient is

$$D_{\mathrm{eff}} \sim \exp\left[2J(\epsilon\tau^*/2) - \tau^*\right] \tag{52}$$

which is consistent with (14).

Note that while our arguments considered the medians of the S_N and the upper quartile of the T_N, the same reasoning applies for any percentiles of S_N or T_N.

4 Flooding Dynamics Model for Dispersion

In Sect. 2, we showed that the integrals which are used to compute the mean-first-passage time may be approximated by sums when D is small. In Sect. 3, we considered the statistics of these sums, S_N and T_N, defined by equations (21) and (22) respectively. In terms of the calculation discussed in Sect. 2, our estimate of $\bar{T}_N$ corresponds, for $N < N^*$, to the value of $\bar{T}(x)$ being determined by the difference between the lowest minimum of the potential and its highest maximum, provided the minimum occurs before the maximum. We can therefore think of $\bar{T}(x)$ being determined by a 'flooding' model, according to which the probability density for locating the particle occupies a region which is constrained by a potential barrier which can trap a particle for time $\bar{T}$. As $\bar{T}$ increases, higher barriers are required.

In terms of the original problem, discussed in Sect. 2, N is the number of extrema of the potential before we reach position x. The arguments of Sect. 3 imply that the upper quartile of the mean-first-passage time, $\bar{T}(x)$, satisfies an equation similar to (49). We define logarithmic variables

$$\eta = \ln\left(\frac{x}{\tilde{x}}\right) , \quad \tau = \ln \bar{T}(x) \tag{53}$$

where $\tilde{x}$ is the mean separation of minima of $V(x)$. In terms of these logarithmic variables, the dispersion is described by

$$\eta = J(D\tau/2) . \tag{54}$$
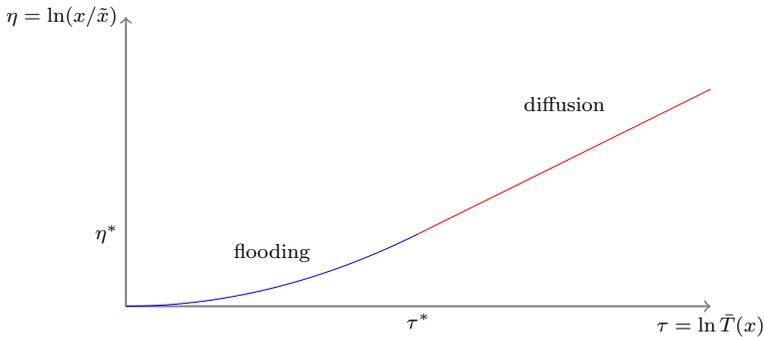
Fig. 1 The dynamics of a typical realisation, characterised by the median $\bar{T}$ of the mean-first-passage time, shows a crossover from subdiffusive 'flooding' dynamics to slow diffusion

which is valid up to τ^*, which is defined by the condition

$$\left.\frac{d\eta}{d\tau}\right|_{\tau^*} = \frac{1}{2} \ . \tag{55}$$

Equation (54) is our principal result. It applies to any percentile of the distribution which remains fixed when we take the limits $N \to \infty$ and $\epsilon \to 0$. When τ is large compared to τ^*, equation (54) is replaced by a linear relation, with an effective diffusion coefficient D_{eff}

$$2\eta \sim \tau + \ln(D_{\text{eff}}/D) \ , \quad D_{\text{eff}} \sim D \exp\left[2J(D\tau^*/2) - \tau^*\right] \ . \tag{56}$$

An important example is the case where V has a Gaussian distribution, so that $J \sim V^2/2C(0)$. In terms of the diffusion coefficient D, equations (54)-(56) give

$$\eta \sim \frac{D^2\tau^2}{8C(0)} \quad \tau < \tau^* = \frac{2C(0)}{D^2}$$

$$\eta \sim \frac{\tau}{2} - \frac{C(0)}{2D^2} \quad \tau > \tau^* \tag{57}$$

and using (56) we find $D_{\text{eff}} \sim D \exp[-C(0)/D^2]$, in agreement with (5). A sketch of the dependence of η upon τ for the Gaussian case is shown in Fig. 1.

5 Numerical Studies

We performed a variety of numerical investigations, using Gaussian distributed random variables f_j to test the theory of extreme-weighted sums, and a Gaussian random function $V(x)$ to test the analysis of continuous potentials. In both cases the Gaussian variables had zero mean and unit variance. In the case of the random potential, we also used a Gaussian for the correlation function, with a correlation length of order unity:

$$\langle V(x)V(x')\rangle = \exp[-(x - x')^2/2] \ . \tag{58}$$

5.1 Discrete Sums

We characterised the statistics of the discrete sum (8) by making a careful estimate of its median, equation (42). In order to evaluate equation (42), we need a solution of the implicit

equation (33), which determines $f^\dagger$. By substituting (34) into (33), we find

$$f^\dagger \approx \sqrt{2\ln N - \ln \ln N - \ln(2\pi) - 2\ln \ln 2} \,. \tag{59}$$

The expression for the median approaches that for the mean

$$S_N \sim N \exp\left(\frac{1}{2\epsilon^2}\right) \tag{60}$$

at large values of N when $f^\dagger$ exceeds $f^* = 1/\epsilon$.

For very large N and very small ϵ, the medians of S_N and T_N are estimated by simplified expressions, relating $\sigma = \ln \bar{S}_N$ and $\tau = \ln \bar{T}_N$ to $\eta = \ln N$. In the Gaussian case, these equations (48), (50) and (51) give

$$\sigma = \begin{cases} \dfrac{\sqrt{2\eta}}{\epsilon} & 0 < \eta < \eta^* \\ \eta - \eta^* + \sigma^* & \eta \geq \eta^* \end{cases} \tag{61}$$

and

$$\tau = \begin{cases} \dfrac{\sqrt{8\eta}}{\epsilon} & 0 < \eta < \eta^* \\ 2(\eta - \eta^*) + \tau^* & \eta > \eta^* \end{cases} \tag{62}$$

where

$$\eta^* = \frac{1}{2\epsilon^2} \,, \quad \sigma^* = \frac{1}{\epsilon^2} \,, \quad \tau^* = \frac{2}{\epsilon^2} \,. \tag{63}$$

These equations imply that, in the limit as $\epsilon \to 0$, if we plot $y = \sigma/\eta^*$ as function of $x = \eta/\eta^*$, the numerical data for $\bar{S}_N$ should collapse onto the function

$$y = f(x) = \begin{cases} 2\sqrt{x} & 0 < x < 1 \\ x + 1 & x > 1 \end{cases} \,. \tag{64}$$

Similarly, $y' = \tau/\eta^*$ plotted as a function of $x = \eta/\eta^*$ should collapse to $y' = 2f(x)$.

We computed $M \in \{10, 100, 1000\}$ realisations of the sums S_N and T_N, for $\epsilon \in \{1/3, 1/4, 1/6, 1/8\}$ and $N \leq 10^5$ (except for $M = 1000$, in which case $N \leq 5 \times 10^4$). We evaluated the sample average, sample median, and the sample upper octile as a function of N for each value of the sample size M. We also computed the same statistics for T_N.

Figure 2 plots $\ln \bar{S}_N$, and $\ln\langle S_N \rangle$ as a function of $\eta = \ln N$, for different sample sizes, for $\epsilon = 1/4$ (**a**) and $\epsilon = 1/6$ (**b**). We compare with the theoretical prediction, obtained from (42) and (59) (for the median) and (60) (for the mean). The agreement with the theory for the median is excellent. Note that the convergence of the mean value for different sample sizes is very poor when $\eta < \eta* = 1/2\epsilon^2$ (this is especially apparent for smaller values of ϵ).

In figure 3, we plot $y = \sigma/\sigma^*$ (**a**) and $y' = \tau/\tau^*$ (**b**) as a function of $x = \eta/\eta^*$, for all of the values of ϵ in our data set, using the largest sample size ($M = 1000$) in each case, comparing with the theoretical scaling function (64). We see convergence towards the function (64) as $\epsilon \to 0$. In panels (**c**) and (**d**), we make the same comparison using the upper octile rather than the median.

5.2 Continuous Potentials

In the case of a continuous potential, we require the mean separation of maxima or minima, $\tilde{x}$, in order to make a comparison with theory. The density $\mathcal{D}$ of zeros of $V'(x)$ may be
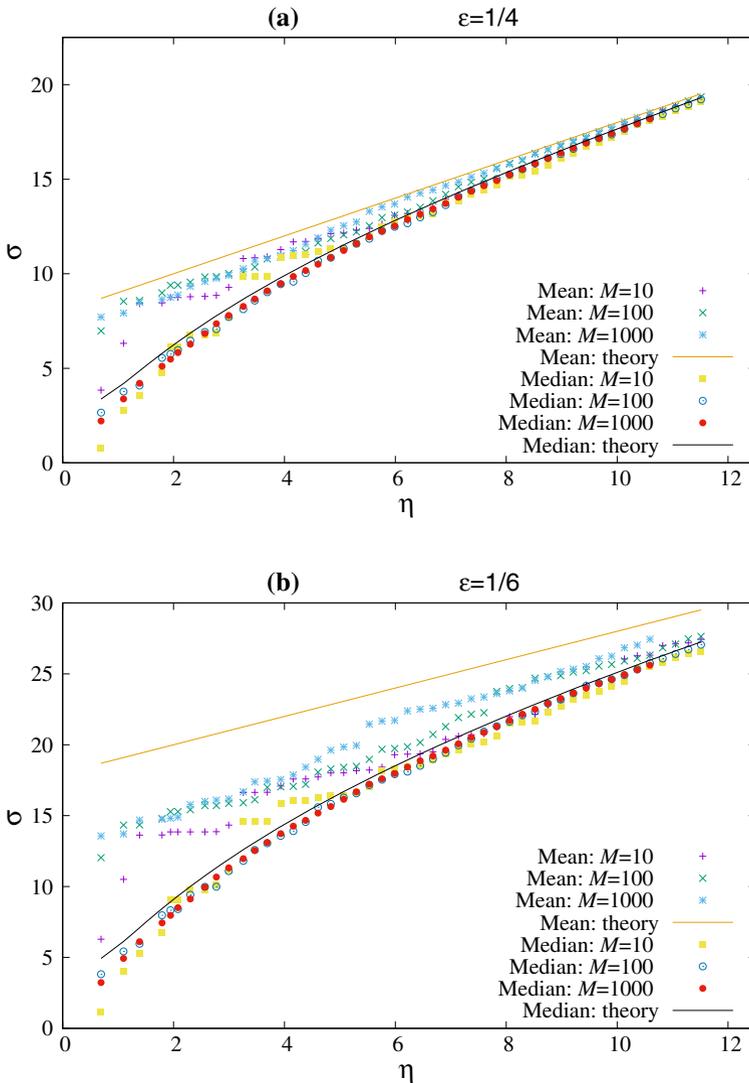
Fig. 2 Plot of $\ln \bar{S}_N$ and $\ln \langle S_N \rangle$, as a function of $\eta = \ln N$, for $\epsilon = 1/4$ (**a**) and $\epsilon = 1/6$ (**b**)

determined by the approach developed by Kac [31] and Rice [32]. If $\mathcal{P}(V, V', V'')$ is the joint PDF of $V(x)$ and its first two derivatives, evaluated at the same point, we find that

$$\mathcal{D} = \int_{-\infty}^{\infty} \mathrm{d}V \int_{-\infty}^{\infty} \mathrm{d}V'' \, \mathcal{P}(V, 0, V'') |V''| \, . \tag{65}$$

By noting that the vector (V, V', V'') has a multivariate Gaussian distribution, and expressing $\mathcal{P}(V, V', V'')$ in terms of the correlation function of the elements of this vector, we obtain $\mathcal{D}$ and hence the separation of minima $\tilde{x}$ for the potential satisfying (58):

$$\frac{2}{\tilde{x}} \equiv \mathcal{D} = \frac{\sqrt{3}}{2\pi} \, . \tag{66}$$
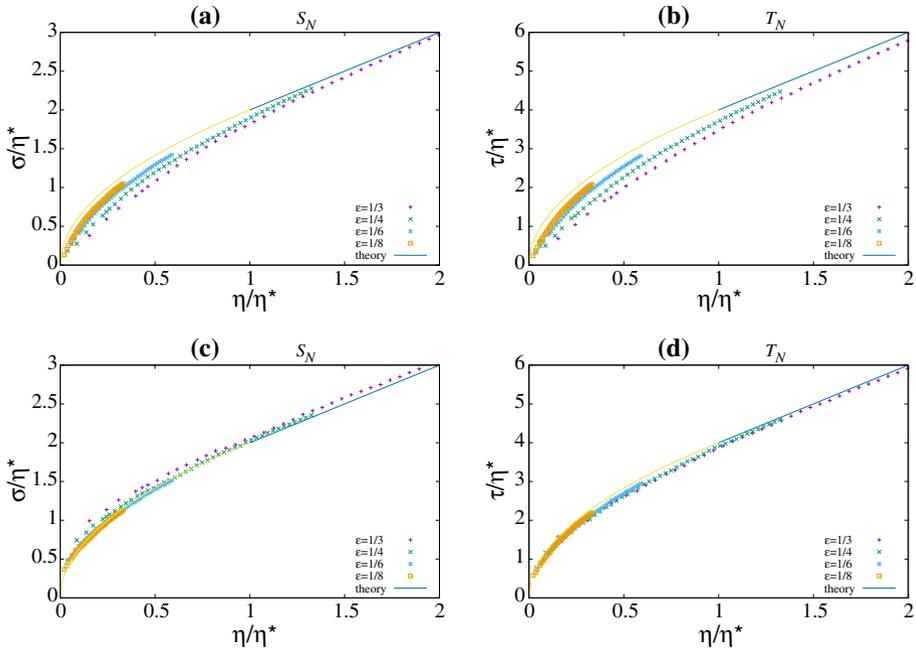
Fig. 3 Plot of σ/η^* (**a**) and τ/τ^* (**b**) based upon median values, as a function of η/η^*, compared with the theoretical prediction for the $\epsilon \to 0$ limit, equation (64). In (**c**) and (**d**) we show similar plots for the upper octile
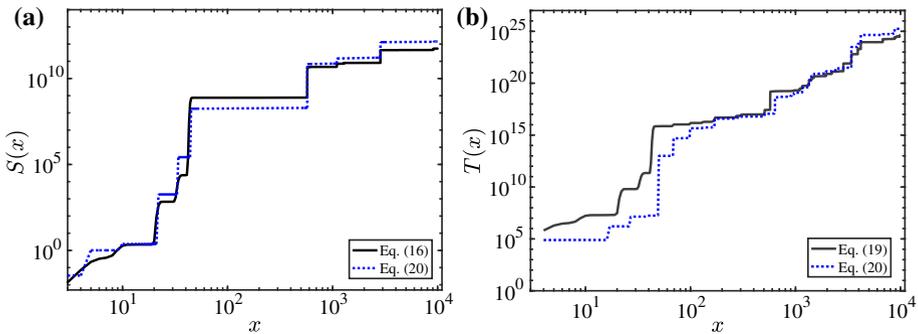
Fig. 4 Plots of $S(x)$ (**a**) and $T(x)$ (**b**), for one realisation of the smooth random potential, with $D = 1/8$. The numerically evaluated integrals (solid curves) are in good agreement with approximations based on sums over maxima and minima (dashed lines)

First we investigated whether the mean-first passage time can be accurately represented by sums over maxima and minima of the potential. In figure 4, we compare the numerical evaluation of the integrals $S(x)$ (**a**) and $T(x)$ (**b**), given by Eq. (16) and Eq. (19), respectively, with the approximations which estimate the integrals using maxima and minima, Eq. (20).

We evaluated the median $\bar{T}(x)$ and mean $\langle T(x) \rangle$ of the mean first passage time $T(x)$ for 1000 realisations of the potential $V(x)$, up to $x_{\max} = 10^4$, for $D \in \{1/3, 1/4, 1/5, 1/6, 1/7\}$. According to the discussion in Sect. 4, we expect that $\tau = \ln \bar{T}(x)$ and $\eta = \ln(x/\tilde{x})$ are related by $\eta = J(D\tau/2) = D^2\tau^2/8$, up to a maximum value of η, given by $\eta^* = 1/2D^2$.

Springer
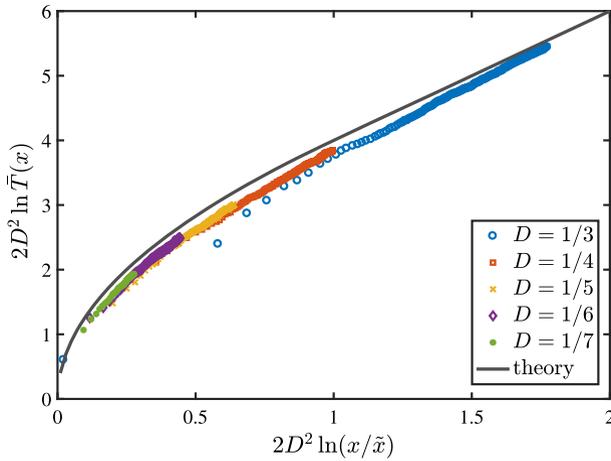
Fig. 5 Numerical results on the median $\bar{T}(x)$ of the mean first-passage time to reach x, for different values of the diffusion coefficient D. We plot $D^2 \ln \bar{T}(x)/2$ as a function of $X = 2D^2 \ln(x/\tilde{x})$ (where $\tilde{x}$ is the mean distance between maxima of $V(x)$). The numerical results converge towards the theoretically predicted scaling function, equation (64), as $D \to 0$

In figure 5 we plot $2D^2 \ln \bar{T}(x)$ as a function of $2D^2 \ln(x/\tilde{x})$ for different values of D, and compare with the theoretical scaling function, given by equation (64).

6 Conclusions

In his analysis of equation (1), Zwanzig considered the mean-first-passage time $T(x)$, which is averaged over different noise realisations, to reach displacement x. Computing the averaged value $\langle T(x) \rangle$ over different realisations of the random potential, he showed [4] that $\langle T(x) \rangle \sim x^2$, which is consistent with a diffusive dispersion, with an effective diffusion coefficient D_{eff}. The effective diffusion coefficient vanishes in a highly singular manner as $D \to 0$, and numerical studies have suggested that equation (1) exhibits anomalous diffusion [6–8]. It seems evident that the discrepancy between these two pictures of the dynamics results from the expectation value $\langle T(x) \rangle$ being dominated by rare events, where an unusually large fluctuation of the potential $V(x)$ acts as a barrier to dispersion. The central limit theorem is applicable to this problem, and at sufficiently large values of x the ratio $T(x)/\langle T(x) \rangle$ is expected to approach unity, for almost all realisations of $V(x)$. However, at values of x which are of practical relevance, most realisations $T(x)$ will be much smaller than $\langle T(x) \rangle$.

In order to give a description of the dynamics of (1) which is both empirically usable and analytically tractable, we considered the median (with respect to different realisations of the potential) of the mean-first-passage time. In the limit where the diffusion coefficient D is small, the integrals which appear in the expression for the first passage time, equation (9), are dominated by maxima and minima of the potential, described by equations (17) and (20). This observation led us to consider the statistics of sums of exponentials of random variables, equations (8) and (22). We gave a quite precise estimate, equation (42), for the median of (8) and also derived simple relations describing the asymptotic behaviour of these sums, equations (48) and (50).

It is these expressions which enable us to formulate a concise asymptotic description of the dynamics of (1) in the limit as $D \to 0$, in terms of the large deviation rate function of the potential, $J(V)$. We argued that at very long length scales $\bar{T}(x)$ approaches the expectation value $\langle T(x) \rangle$, and that the dispersion is diffusive, in accord with the theory of Zwanzig [4]. On shorter timescales $\bar{T}(x)$ is determined by a 'flooding' model, according to which the probability density for locating the particle occupies a region which is constrained by a potential barrier which can trap a particle for time $\bar{T}$. As $\bar{T}$ increases, higher and higher barriers are required. For a Gaussian distribution of barrier heights, equation (57) implies that the dispersion is described as subdiffusive, of the form

$$x \sim \tilde{x} \exp\left(\frac{D^2 (\ln \bar{T})^2}{8 C(0)} \right) \tag{67}$$

which is distinctively different from the power-law anomalous diffusion which has been reported by some authors [6–8]. Our numerical investigations of the dynamics of equation (1) for different values of D, illustrated in figure 5, show a data collapse which is in excellent agreement with equation (64), verifying (67).

References

1. Havlin, S., Ben-Avraham, D.: Diffusion in disordered media. Adv. Phys. **51**(1), 187–292 (2002)
2. Chaikin, P., Lubensky, T.: Principles of Condensed Matter Physics. University Press, Cambridge (1995)
3. Saxton, M.J.: A biological interpretation of transient anomalous subdiffusion. I. Qual. Model Biophys. J. **92**(4), 1178–1191 (2007)
4. Zwanzig, R.: Diffusion in a rough potential. PNAS **85**, 2029–30 (1988)
5. De Gennes, P.G.: Brownian motion of a classical particle through potential barriers. Application to the helix-coil transitions of heteropolymers. J. Stat. Phys. **12**, 463–481 (1975)
6. Khoury, M., Lacasta, A.M., Sancho, J.M., Lindenberg, K.: Weak Disorder: Anomalous transport and diffusion are normal yet again. Phys. Rev. Lett. **106**, 090602 (2011)
7. Simon, M.S., Sancho, J.M., Lindenberg, K.: Transport and diffusion of overdamped Brownian particles in random potentials. Phys. Rev. E **88**, 062105 (2013)
8. Goychuk, I., Kharchenko, V.O., Metzler, R.: Persistent Sinai-type diffusion in Gaussian random potentials with decaying spatial correlations. Phys. Rev. E **96**, 052134 (2017)
9. Sinai, G.Ya.: Theor. Prob. Appl. **27**, 256 (1982)
10. Comtet, A., Dean, D.S.: Exact results on Sinai's diffusion. J. Phys. A **31**, 8595–8605 (1998)
11. Le Doussal, P., Monthus, C., Fisher, D.S.: Random walkers in one-dimensional random environments: exact renormalization group analysis. Phys. Rev. E **5**, 4795–4840 (1999)
12. Dean, D.S., Gupta, S., Oshanin, G., Rosso, A., Schehr, G.: Diffusion in periodic, correlated random forcing landscapes. J. Phys. A **47**, 372001 (2014)
13. Akimoto, T., Saito, K.: Trace of anomalous diffusion in a biased quenched trap model. Phys. Rev. E **101**, 042133 (2020)
14. Anderson, P.W.: Absence of Diffusion in Certain Random Lattices. Phys. Rev. **109**, 1492–1505 (1958)
15. Mott, N.F., Twose, W.D.: The theory of impurity conduction. Adv. Phys. **10**, 107–163 (1961)

16. Abrahams, E., Anderson, P.W., Licciardello, D.C., Ramakrishnan, T.V.: Scaling theory of localization: absence of quantum diffusion in two dimensions. Phys. Rev. Lett. **42**, 673–676 (1979)
17. Akkermans, E., Montambaux, G.: Mesoscopic Physics of Electrons and Photons. University Press, Cambridge (2007) 9780511618833
18. Kramers, H.A.: Brownian motion in a field of force and the diffusion model of chemical reactions. Physica **7**, 284–304 (1940)
19. Redner, S.: A Guide to First-Passage Processes. Cambridge University Press, ISBN 0-521-65248-0 (2001)
20. Pontryagin, L., Andronov, A., Vitt, A.: On the statistical treatment of dynamical systems. *Zh. Eksp. Teor. Fiz.*, **3**, 165-80. [Reprinted in *Noise in Nonlinear Dynamical Systems*, 1989, ed. by F. Moss and P. V. E. McClintock (Cambridge University Press), Vol. 1, p. 329] (1933)
21. Lifson, S., Jackson, J.L.: On self-diffusion of ions in a polyelectrolyte solution. J. Chem. Phys. **36**, 2410–14 (1962)
22. Gardiner, C.W.: Handbook of Stochastic Methods for Physics. Springer-Verlag, Chemistry and the Natural Sciences (1983)3-540-20882-8
23. Boyer, D., Dean, D.S., Mejía-Monasterio, C., Oshanin, G.: Optimal fits of diffusion constants from single-time data points of Brownian trajectories. Phys. Rev. E **86**, 060101(R) (2012)
24. Erdelyi, A.: Asymptotic expansions. Dover Publications, New York (1956) 0486603180
25. Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.L.: Statistics of Extremes: Theory and Applications. Wiley, Chichester (2006) 0471976474
26. Gumbel, E.J.: Statistics of Extremes. Dover Publications, New York (2009) 0486436047
27. Touchette, H.: The large deviation approach to statistical mechanics. Phys. Rep. **478**, 1–69 (2009)
28. Wang, W., Vezzani, A., Burioni, R., Barkai, E.: Transport in disorderd systems: the single big jump approach. Phys. Rev. Res. **1**, 033172 (2019)
29. Romeo, M., Da Costa, V., Bardou, F.: Broad distribution effects in sums of lognormal random variables. Eur. Phys. J. B **32**, 513–525 (2003)
30. Pradas, M., Pumir, A., Wilkinson, M.: Uniformity transition for ray intensities in random media. J. Phys. A **51**, 155002 (2018)
31. Kac, M.: On the average number of real roots of a random algebraic equation. Bull. Am. Math. Soc. **49**, 314–20 (1943)
32. Rice, S.O.: Mathematical analysis of random noise. Bell Syst. Tech. J. **23**, 283–332 (1945)