

Diffusion at finite speed and random walks

Joseph B. Keller*

Department of Mathematics, Stanford University, Stanford, CA 94305-2125

Contributed by Joseph B. Keller, November 12, 2003

Diffusion, which occurs with infinite speed, results from a random walk with steps of finite speed. We resolve this paradox and derive a modified diffusion equation with finite speed.

Particle diffusion results when a particle makes a random walk, i.e., takes a sequence of random steps. Even though each step is made at finite speed, the resulting diffusion occurs with infinite speed. We resolve this paradox by showing that the diffusion approximation to the probability density of the random walk holds only for a limited range of distance. Then, we derive a large deviation result, which holds for greater distances. Finally, we derive a partial differential equation for the probability density of the random walk in one and more dimensions, which shows that the probability density spreads at finite speed.

1. A Paradox and Its Resolution

Consider a particle moving along the x axis starting at $x = 0$ when $t = 0$. During the time interval of duration $\tau = 1$ from $t = i - 1$ to $t = i$, $i = 1, 2, \dots$, the particle moves with velocity $v = +1$ or $v = -1$, each with probability $1/2$. Thus, it takes a step $h_i = \pm 1$ of length 1, and the steps are assumed to be independent. Its position x_i at time t is $x_i = \sum_{i=1}^t h_i$. The expected value of x_i is $Ex_i = 0$ and the variance of x_i is $\text{Var } x_i = E(x_i - Ex_i)^2 = t$.

The central limit theorem of probability theory states that for large t , the probability density of the sum x_i is asymptotic to a normal density with mean 0 and variance t . Writing this density as $p(x, t)$ yields

$$p(x, t) \sim \frac{e^{-x^2/2t}}{(2\pi t)^{1/2}}, \quad |x| = O(\sqrt{t}). \quad [1.1]$$

The density on the right side of Eq. 1.1 satisfies the diffusion equation with diffusion coefficient $D = 1/2$ and initial value $\delta(x)$:

$$p_t(x, t) = Dp_{xx}(x, t), \quad p(x, 0) = \delta(x). \quad [1.2]$$

The result (Eq. 1.1) indicates that for each $t > 0$, there is a positive probability density for the particle to be at any position x , however large. This finding contradicts the fact that $|x_i|$ has a maximum value $|x_i| = t$, attained when all the steps are in the same direction. Thus $|x_i| \leq t$, so

$$p(x, t) = 0, \quad |x| > t. \quad [1.3]$$

To reconcile Eq. 1.1 with Eq. 1.3, we note that the central limit theorem result Eq. 1.1 holds only for $|x| = O(\sqrt{t})$. However, Eq. 1.3 applies when $|x| > t$. Thus, Eqs. 1.1 and 1.3 are valid in different ranges of x . No contradiction occurs provided that Eq. 1.1 is used only where $x = O(\sqrt{t})$.

2. Large Deviation Theory

We shall now determine $p(x, t)$ for values of $|x|$ larger than $O(\sqrt{t})$, where Eq. 1.1 is not valid. To do so we denote by j the number of steps to the right taken by the particle up to time t , and then $t - j$ is the number it has taken to the left. Therefore, its position is

$$x_{ij} = j - (t - j) = (2j - t). \quad [2.1]$$

The probability of this value x_{ij} is given by the binomial distribution

$$P(x_{ij}, t) = \frac{t!}{j!(t-j)!2^t}, \quad j = 0, 1, \dots, t, \quad [2.2]$$

$$= 0 \quad j < 0 \text{ or } j > t.$$

To simplify Eq. 2.2 for t large, we use Stirling's formula and use Eq. 2.1 to get $j = (x + t)/2$. To convert Eq. 2.2 to a density with respect to x , we divide it by 2. In this way we obtain

$$p(x, t) \sim \frac{1}{(2\pi t)^{1/2}} \left(1 + \frac{x}{t}\right)^{-(1/2)(t+x+1)} \left(1 - \frac{x}{t}\right)^{-(1/2)(t-x+1)},$$

$$= 0 \quad \begin{matrix} |x/t| < 1 & t \gg 1, \\ |x/t| > 1. \end{matrix} \quad [2.3]$$

This is called a large deviation result because it is valid for displacements $|x|$, which are large compared with those for which Eq. 1.1 holds. It is valid also for smaller displacements, and it yields Eq. 1.1 when it is expanded for x/t small and t large.

For either j or $t - j$ small, Stirling's formula does not apply, which explains why Eq. 2.3 does not hold at $x/t = \pm 1$. A result valid up to $x/t = +1$, but not near $x/t = -1$, follows from Eq. 2.2 when Stirling's formula is used for $t!$ and $j!$, but not for $(t - j)!$, as in Kuske and Keller (1), equation B9. After dividing the right side of Eq. 2.2 by 2 to make it a density, and applying Stirling's formula, we get

$$p(x, t) \sim \frac{1}{\sqrt{2}} \left(1 + \frac{x}{t}\right)^{-(1/2)(t+x+1)} \frac{\left(\frac{t}{2}\right)^{(1/2)(t-x)} e^{-(1/2)(t-x)}}{\left(\frac{t-x}{2}\right)!}, \quad [2.4]$$

$$= 0, \quad \begin{matrix} 1 \geq x/t > -1, & t \gg 1, \\ |x/t| > 1. \end{matrix}$$

At $x = t$ this yields $p(t, t) \approx 2^{-t-1}$. The factorial in Eq. 2.4 can be interpreted as a gamma function for noninteger values of its argument. When the argument is large, and Stirling's formula is used for the factorial, Eq. 2.4 reduces to Eq. 2.3. By changing x to $-x$, a result valid up to $x = -t$, but not at $x = +t$, is obtained. Therefore, with x replaced by $|x|$, Eq. 2.4 is valid for all x .

Fig. 1 shows the diffusion approximation (Eq. 1.1), the exact probability (Eq. 2.2), the large deviation result (Eq. 2.3), and the improved large deviation result (Eq. 2.4) at $t = 100$ for $-0.8 \leq x/t \leq 0.8$. The exact result (Eq. 2.2) holds only at the discrete values $|x|/t = 0, 0.2, 0.4, 0.6, 0.8, 1.0$, and $|x|/t > 1$. The diffusion approximation yields a positive value for all $|x|/t$, whereas Eq. 2.2 vanishes for $|x|/t > 1$.

Fig. 2 shows these same results at $t = 10$ for $0.8 \leq x/t \leq 1.05$ on an expanded scale, which shows up their differences. It shows that the large deviation result (Eq. 2.3) does not agree with the exact result (Eq. 2.2) at $x/t = 1$. However, the improved large deviation result (Eq. 2.4) does agree with the exact result (Eq. 2.2)

*E-mail: keller@math.stanford.edu.

© 2004 by The National Academy of Sciences of the USA

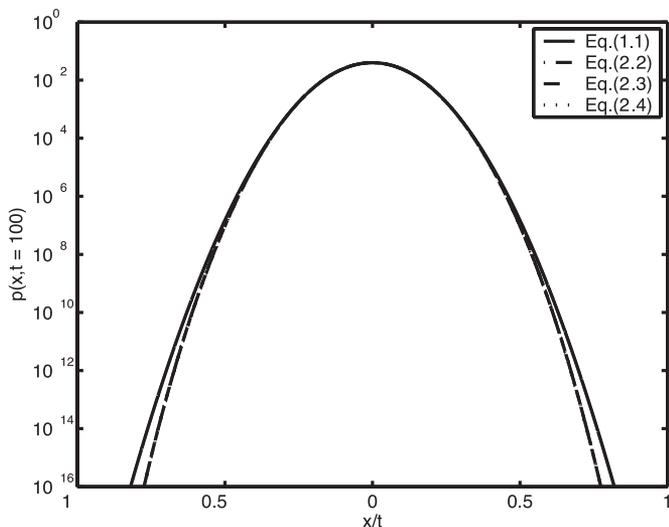


Fig. 1. The diffusion approximation (Eq. 1.1), the exact probability (Eq. 2.2), the large deviation result (Eq. 2.3), and the improved large deviation result (Eq. 2.4) at $t = 100$ are shown as functions of x/t . The latter three curves are indistinguishable on the scale of this figure.

3. A Differential Equation for $p(x,t)$

We shall now derive a differential equation for $p(x,t)$ which yields a better approximation to the exact probability (Eq. 2.2) than does the diffusion equation (Eq. 1.2). To derive it, we let $x_j = j$, $j = 0, \pm 1, \pm 2, \dots$, and $t = 0, 1, \dots$. Then, from the definition of the random walk, the probability $p(x_j, t + 1)$ is given by

$$p(x_j, t + 1) = \frac{1}{2} p(x_{j-1}, t) + \frac{1}{2} p(x_{j+1}, t). \quad [3.1]$$

This is an exact difference equation for p . It is convenient to subtract $p(x_j, t)$ from both sides of Eq. 3.1 to get

$$p(x_j, t + 1) - p(x_j, t) = \frac{1}{2} [p(x_{j-1}, t) - 2p(x_j, t) + p(x_{j+1}, t)]. \quad [3.2]$$

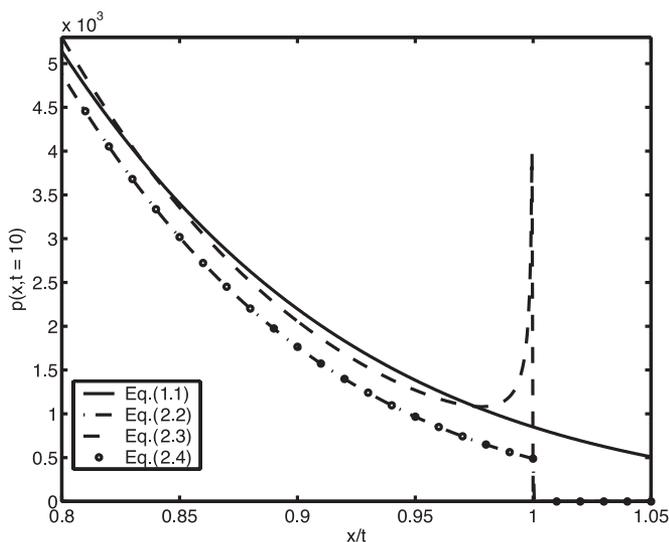


Fig. 2. The same four results as in Fig. 1 are shown on an expanded scale in the interval $0.8 \leq x/t \leq 1.05$ at $t = 10$. The differences between the various curves are apparent. The exact result (Eq. 2.2) and the improved large deviation result (Eq. 2.4) are indistinguishable, whereas the large deviation result (Eq. 2.3) and the Gaussian (Eq. 1.1) differ from them in this interval.

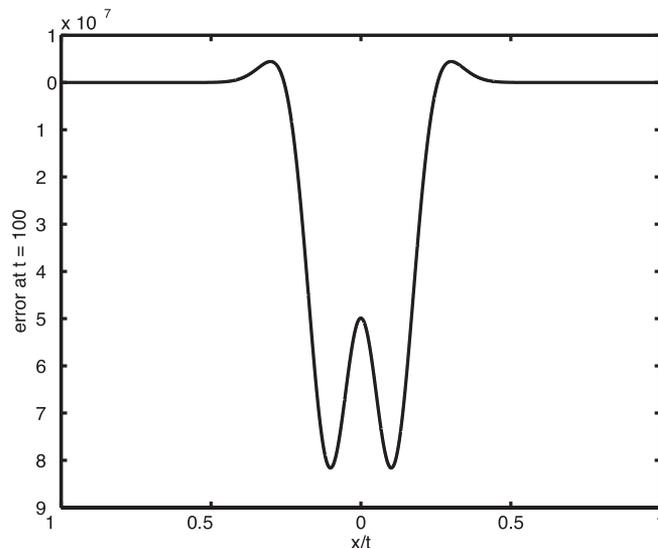


Fig. 3. The error at $t = 100$, as a function of x/t . It is obtained by using the large deviation result (Eq. 2.3) in the telegraph equation (Eq. 3.3), and subtracting the right side from the left side. Note that the vertical scale unit is 10^{-7} .

If $p(x, t)$ is a smooth function of its arguments, we can expand p in a Taylor series in x and t . Then, when we omit derivatives of an order higher than two, the left side is $p_t + \frac{1}{2} p_{tt}$ and the right side is $\frac{1}{2} p_{xx}$, so Eq. 3.2 becomes

$$p_t(x, t) + \frac{1}{2} p_{tt}(x, t) = \frac{1}{2} p_{xx}(x, t). \quad [3.3]$$

Eq. 3.3 is the telegraph equation for p . This is a hyperbolic equation with characteristic velocities $+1$ and -1 . Therefore, the solution p propagates with speed ≤ 1 , which is the same speed as the steps of the random walk.

To illustrate the accuracy of this equation, we have substituted the large deviation approximation (Eq. 2.3) into it. The error, defined to be the left side of Eq. 3.3 minus the right side, is plotted in Fig. 3 at $t = 100$. Its magnitude is $< 10^{-6}$ for $|x|/t < 1$.

To solve an initial value problem for Eq. 3.3, it is necessary to specify both $p(x, 0)$ and $p_t(x, 0)$. For the walk starting at $t = 0$, the initial value of p is

$$p(x, 0) = \delta(x). \quad [3.4]$$

We obtain p_t by taking the t derivative of Eq. 2.3 and evaluating it for $|x|/t$ small. The result is that $p_t \approx (-1/2t)p$. Then from Eq. 3.4, the asymptotic behavior of p_t is

$$p_t(x, t) \sim -\frac{\delta(x)}{2t} \text{ as } t \rightarrow 0. \quad [3.5]$$

This result also follows from Eq. 1.1.

Next we consider a general initial value of p , say,

$$p(x, 0) = f(x). \quad [3.6]$$

To obtain the initial behavior of p_t , we can use the linearity of the differential equation (Eq. 3.3). Therefore we replace $\delta(x)$ on the right side of Eq. 3.5 by $f(x)$ to obtain

$$p_t(x, t) \sim -\frac{f(x)}{2t} \text{ as } t \rightarrow 0. \quad [3.7]$$

4. More General Random Walks and Higher Dimensions

Now we shall extend the result of Section 3 to obtain a partial differential equation for more general random walks in one or more dimensions. We denote by $p(x, t)$ the probability density for the walk to be at x at time t , where x is the position vector in N dimensions with components x_α . Let τ be the duration of a step, and let the vector s , with components s_α , be the random step. We assume that the steps are independent identically distributed random variables with common density $\rho(s)$. Then $p(x, t + \tau)$ is given by

$$p(x, t + \tau) = \int p(x - s, t) \rho(s) ds. \quad [4.1]$$

We now expand p on the left side of Eq. 4.1 as a Taylor series in τ , and p on the right side as a Taylor series in s :

$$\begin{aligned} p(x, t) + \tau p_t(x, t) + \frac{\tau^2}{2} p_{tt}(x, t) + O(\tau^3) \\ = \int \left[p(x, t) - s_\alpha p_{x_\alpha}(x, t) \right. \\ \left. + \frac{s_\alpha s_\beta}{2} p_{x_\alpha x_\beta}(x, t) + O(|s|^3) \right] \cdot \rho(s) ds \\ = p(x, t) - (E s_\alpha) p_{x_\alpha}(x, t) + \frac{E(s_\alpha s_\beta)}{2} p_{x_\alpha x_\beta}(x, t) + E[O(|s|^3)]. \end{aligned} \quad [4.2]$$

Here each term with a repeated subscript α or β is summed over the values α or $\beta = 1, \dots, N$.

When the terms $O(\tau^3)$ and $E[O(|s|^3)]$ are omitted and the common term $p(x, t)$ is canceled, Eq. 4.2 becomes

$$\tau p_t + \frac{\tau^2}{2} p_{tt} = - (E s_\alpha) p_{x_\alpha} + \frac{1}{2} E(s_\alpha s_\beta) p_{x_\alpha x_\beta}. \quad [4.3]$$

Suppose that $E(s_\alpha) = 0$. We introduce the diffusion coefficient matrix $D_{\alpha\beta} = E(s_\alpha s_\beta)/2$. We also set $t' = t/\tau$ and write $p(x, t, \tau) = p'(x, t')$. Then Eq. 4.3 becomes the telegraph equation in N dimensions:

$$p'_{t'} + \frac{1}{2} p'_{t't'} = D_{\alpha\beta} p'_{x_\alpha x_\beta}. \quad [4.4]$$

For $n = 1$ and $D_{11} = 1/2$ this is exactly Eq. 3.3.

Suppose that the initial value of p is given by Eq. 3.6 where x is the N dimensional position vector. Then an analysis like that in Section 3 leads to the result

$$p_t(x, t) \sim -\frac{N}{2t} f(x) \quad \text{as } t \rightarrow 0. \quad [4.5]$$

Thus Eqs. 3.6 and 4.5 provide the initial data required by Eq. 4.4.

Conclusion

The considerations of Section 1 resolve the paradox of infinite speed of diffusion. Section 2 shows how to find the density $p(x, t)$ for values of x and t for which the diffusion approximation does not apply. From the random walk we find that $p(x, t)$ satisfies the telegraph equation (Eqs. 3.3 or 4.4), which has finite speed of propagation. Because this equation is of second order in t , it requires two initial conditions instead of the single one required by the diffusion equation. We have shown how to obtain these two conditions.

Many other authors have derived the telegraph equation, or some similar hyperbolic equation, for diffusion. An extensive review of this work was given by Joseph and Preziosi (2). The derivations sometimes invoke various physical hypotheses, and, in general, they do not obtain the second initial condition, which is needed to solve the initial value problem.

I thank Dr. Arnold Kim for producing the figures, Professor Amir Dembo for a helpful discussion, and the Air Force Office of Scientific Research for partial support.

1. Kuske, R. & Keller, J. B. (1997) *Eur. J. Appl. Math.* **8**, 567–580.

2. Joseph, D. D. & Preziosi, L. (1989) *Rev. Mod. Phys.* **61**, 41–73.