Homework 1

Robert Picardo

BIOENG 104 Biological Transport Phenomena | Aaron Streets

University of California, Berkeley

1 Problem 1

We showed that if N particles all start at $x = 0$ at $t = 0$, after a random walk of time t, then they have mean position $\langle x \rangle = 0$, and mean-square position $\langle x^2 \rangle = 2Dt$. Now assume that all the particles start at a position $x = L$ at $t = 0$. The number of time steps is n, the time for a single step is τ, and the total elapsed time is $t = n\tau$.

(a) Following the original derivation, derive the new mean position $\langle x(n) \rangle$.

By definition, the mean position of N particles along the x-axis at step n is

$$\langle x(n) \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i(n) \tag{1.1}$$

We assume that, for each step n, each particle changes its position x_i by distance $\pm\delta$ over time τ. The sign of $\pm\delta$ is to be completely random. Thus, we have $x_i(n) = x_i(n-1) \pm \delta$, and

$$\langle x(n) \rangle = \frac{1}{N} \sum_{i=1}^{N} \left[x_i(n-1) \pm \delta \right] = \frac{1}{N} \sum_{i=1}^{N} x_i(n-1) + \frac{1}{N} \sum_{i=1}^{N} (\pm\delta) \tag{1.2}$$

If we backtracked the path of all particles n steps back to time $t = 0$, each particle would have moved n iterations of $\pm\delta$ before the nth step.

$$\langle x(n) \rangle = \frac{1}{N} \sum_{i=1}^{N} x_i(n-n) + \frac{n}{N} \sum_{i=1}^{N} (\pm\delta) = \frac{1}{N} \sum_{i=1}^{N} x_i(0) + \frac{n}{N} \sum_{i=1}^{N} (\pm\delta) \tag{1.3}$$

We note that all particles start at position L at the zeroth step (i.e. $x_i(0) = L$ for all i). The randomness of the sign of $\pm\delta$ also implies that the particles are distributed around L with no favor to the right or left of L. In other words, for large enough N, $\langle \pm\delta \rangle = 0$.

$$\langle x(n) \rangle = \frac{1}{N} \sum_{i=1}^{N} L + \frac{n}{N} \sum_{i=1}^{N} (\pm\delta) = L + \frac{n}{N} \langle \pm\delta \rangle = L \tag{1.4}$$

(b) Now show that the variance σ^2 of this new particle distribution is the same as the original situation where $x = 0$ at $t = 0$. In your proof, use the equation $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \langle x\rangle\right)^2$.

Noting that the new particle distribution has $\langle x\rangle = L$ for large enough N and t, the variance is given by

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \langle x\rangle\right)^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - L)^2 \tag{1.5}$$

and expanding out, we get

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}x_i^2 - 2L\frac{1}{N}\sum_{i=1}^{N}x_i + \frac{1}{N}\sum_{i=1}^{N}L^2 = \langle x^2\rangle - 2L\langle x\rangle + L^2 = \langle x^2\rangle - L^2 \tag{1.6}$$

We then separately calculate $\langle x^2\rangle$ before plugging it back into σ^2. If we backtrack to the $(n-1)$th step, the positions of each ith particle would have differed from those at the nth step by $\pm\delta$. So, $x_i(n) = x_i(n-1) \pm \delta$ for every ith particle, and

$$\begin{aligned}
\langle x^2\rangle &= \frac{1}{N}\sum_{i=1}^{N}[x(n)]^2 = \frac{1}{N}\sum_{i=1}^{N}[x_i(n-1) \pm \delta]^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}[x_i(n-1)]^2 + \cancelto{0}{\frac{2}{N}\sum_{i=1}^{N}[x_i(n-1)(\pm\delta)]} + \frac{1}{N}\sum_{i=1}^{N}\delta^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}[x_i(n-1)]^2 + \frac{1}{N}\sum_{i=1}^{N}\delta^2
\end{aligned}$$

Backtracking one step again, we have

$$\langle x^2\rangle = \frac{1}{N}\sum_{i=1}^{N}[x_i(n-2) \pm \delta]^2 = \frac{1}{N}\sum_{i=1}^{N}[x_i(n-2)]^2 + \frac{2}{N}\sum_{i=1}^{N}\delta^2$$

and if we backtrack all the way to the zeroth step at $t = 0$, we have

$$\langle x^2\rangle = \frac{1}{N}\sum_{i=1}^{N}[x_i(0)]^2 + \frac{n}{N}\sum_{i=1}^{N}\delta^2 = \frac{1}{N}\sum_{i=1}^{N}L^2 + \frac{n}{N}\sum_{i=1}^{N}\delta^2 = L^2 + n\delta^2 \tag{1.7}$$

Since $\sigma^2 = \langle x^2\rangle - L^2$, then

$$\sigma^2 = (L^2 + n\delta^2) - L^2 = n\delta^2 \tag{1.8}$$

which is the variance of the original particle distribution with particles all at $x = 0$ at $t = 0$ (as given by the Book 1: Diffusion, Chapter 1).

2 Problem 2

Suppose you're told that, for N particles all at $x = 0$ when $t = 0$, the concentration of particles at a given point on the x-axis at a given time can be calculated using the equation

$$c(x,t) = \frac{N}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}$$

Since the equation has one length dimension x, the concentration has units of number of particles/meter. Verify that this equation agrees with our random walk analysis by showing that $\langle x^2 \rangle = 2Dt$ for $c(x,t)$ above.

By definition, the expected value for x^2 is

$$\langle x^2 \rangle = \frac{1}{N} \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{4Dt}}\,\mathrm{d}x = \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} x \left(x e^{-\frac{x^2}{4Dt}} \right) \mathrm{d}x \tag{2.9}$$

where f is the probability density function—in this case, f may be taken as the concentration function c. Before performing integration by parts, we determine the antiderivative of $x \exp\bigl(-x^2/(4Dt)\bigr)$ via integration by substitution.

$$\begin{aligned}
\int x e^{-\frac{x^2}{4Dt}}\,\mathrm{d}x &= -2Dt \int e^{\frac{-x^2}{4Dt}} \left(-\frac{2x}{4Dt} \right) \mathrm{d}x \\
&= -2Dt \int e^{\frac{-x^2}{4Dt}}\,\mathrm{d}\left(-\frac{x^2}{4Dt} \right) \\
&= -2Dt e^{\frac{-x^2}{4Dt}} + \text{constant}
\end{aligned}$$

Therefore, by integration by parts, the integral in (2.9) evaluates to be

$$\begin{aligned}
\langle x^2 \rangle &= \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} x \frac{\mathrm{d}}{\mathrm{d}x} \left(-2Dt e^{\frac{-x^2}{4Dt}} \right) \mathrm{d}x = \frac{-2Dt}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} x \frac{\mathrm{d}}{\mathrm{d}x} \left(e^{\frac{-x^2}{4Dt}} \right) \mathrm{d}x && (2.10) \\
&= \frac{-2Dt}{\sqrt{4\pi Dt}} \left(\left[x e^{\frac{-x^2}{4Dt}} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}(x)\, e^{\frac{-x^2}{4Dt}}\,\mathrm{d}x \right) && (2.11) \\
&= \frac{-2Dt}{\sqrt{4\pi Dt}} \left(\left[x e^{\frac{-x^2}{4Dt}} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{\frac{-x^2}{4Dt}}\,\mathrm{d}x \right) = \frac{2Dt}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{4Dt}}\,\mathrm{d}x && (2.12)
\end{aligned}$$

The eliminative step in (2.12), where the left term in parentheses equals zero, comes from the fact that, by L'Hôpital's rule,

$$\left[x e^{\frac{-x^2}{4Dt}} \right]_{-\infty}^{\infty} = \lim_{a \to -\infty} a e^{\frac{-a^2}{4Dt}} + \lim_{b \to \infty} b e^{\frac{-b^2}{4Dt}} = \lim_{a \to -\infty} \left[-\frac{4Dt}{2a} e^{\frac{-a^2}{4Dt}} \right] + \lim_{b \to \infty} \left[-\frac{4Dt}{2b} e^{\frac{-b^2}{4Dt}} \right] = 0$$

The integral in (2.12) can be evaluated by applying the Leibniz integral rule to the Gaussian integral:

$$\int_{-\infty}^{\infty} e^{\frac{-x^2}{4Dt}}\,\mathrm{d}x = \int_{-\infty}^{\infty} e^{-ax^2}\,\mathrm{d}x = \sqrt{\frac{\pi}{a}} = \sqrt{4\pi Dt} \tag{2.13}$$

for $a = 1/(4Dt)$. Inserting the result of (2.13) to the integral of (2.12) simplifies the expectation of x^2 into the desired result, which verifies the random walk analysis approach.

$$\langle x^2 \rangle = \frac{2Dt}{\sqrt{4\pi Dt}}\sqrt{4\pi Dt} = 2Dt \tag{2.14}$$

3 Problem 3

The experiment below has been collected from analysis of microscope images. Use it to calculate the diffusion coefficient, D, for the cell surface-bound photoactivatable protein (in µm^2 s^{-1}). You should use all of the data in your calculation by performing a least-squares fit and state the confidence of your measurement, using, for example, the coefficient of determination or R^2 value.

t (s)	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
$\sqrt{\langle r^2 \rangle}$ (µm)	0.59	0.95	1.01	1.26	1.39	1.44	1.64	1.68	1.78

The mean square radial position resulting from diffusion in two dimensions is

$$\langle r^2 \rangle = 4Dt \tag{3.15}$$

for $r^2 = x^2 + y^2$, where x and y are Cartesian coordinates, and time t.

We perform a linear regression on the experimental data above via the least-squares method. Given a collection of data points (t_i, y_i), each ith data point can be roughly approximated by an equation $y_i \approx b + mt_i$ for constants m and b which represent the slope and intercept, respectively, of the line of best fit. Here, since we assume that all particles are positioned at $r = 0$ at $t = 0$, then we set $b = 0$.

If we denote the vector containing all values of y_i as $\boldsymbol{y}$ and the vector containing all values of t as $\boldsymbol{t}$, then the equations $y_i \approx mt_i$ can be re-expressed as

$$\boldsymbol{y} \approx m\boldsymbol{t} \implies \operatorname{proj}_{\boldsymbol{t}} \boldsymbol{y} = m\boldsymbol{t}$$

where $\operatorname{proj}_{\boldsymbol{t}} \boldsymbol{y}$ is the orthogonal projection of $\boldsymbol{y}$ onto the linear span of $\boldsymbol{t}$.

$$\operatorname{proj}_{\boldsymbol{t}} \boldsymbol{y} = \frac{\boldsymbol{y}\cdot\boldsymbol{t}}{\boldsymbol{t}\cdot\boldsymbol{t}}\boldsymbol{t} \implies m = \frac{\boldsymbol{y}\cdot\boldsymbol{t}}{\boldsymbol{t}\cdot\boldsymbol{t}}$$

We can solve for the slope m by multiplying $\operatorname{proj}_{\boldsymbol{t}} \boldsymbol{y}$ by the inverse of $\boldsymbol{t}$. We can implement this easily using Python. Here, our values of y_i are obtained by squaring the values of $\sqrt{\langle r^2 \rangle}$, thereby matching Equation (3.15). This yields the slope $m = 4D = 36.35$ µm^2 s^{-1}.

```python
import numpy as np

# Datasets (times t and sqrt(<r^2>) values y)
t = np.array([0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09])
sqrtr2 = np.array([0.59, 0.95, 1.01, 1.26, 1.39, 1.44, 1.64, 1.68, 1.78])
y = sqrtr2**2

# Linear regression
m = np.dot(y,t)/np.dot(t,t)
print(m) # output: 36.34663157894737
```

An additional calculation shows that the experimental measurements can be modeled with relatively high goodness of fit, having a coefficient of determination $R^2 = 0.984$. We determine $R^2 = 1 - SS_{\text{res}}/SS_{\text{tot}}$ by finding the sum of squared differences from the regression model ($SS_{\text{res}} = \sum_i (y_i - L)^2$) and the sum of squared differences from the mean ($SS_{\text{tot}} = \sum_i (y_i - \mu)^2$).

```python
# Coefficient of determination
SSres = np.sum((y - m*t)**2)
SStot = np.sum((y - np.mean(y))**2)
R2 = 1 - SSres/SStot
print(R2) # output: 0.9840220692965168
```

Dividing $m = 4D = 36.35\,\mu\text{m}^2\,\text{s}^{-1}$ by four gives

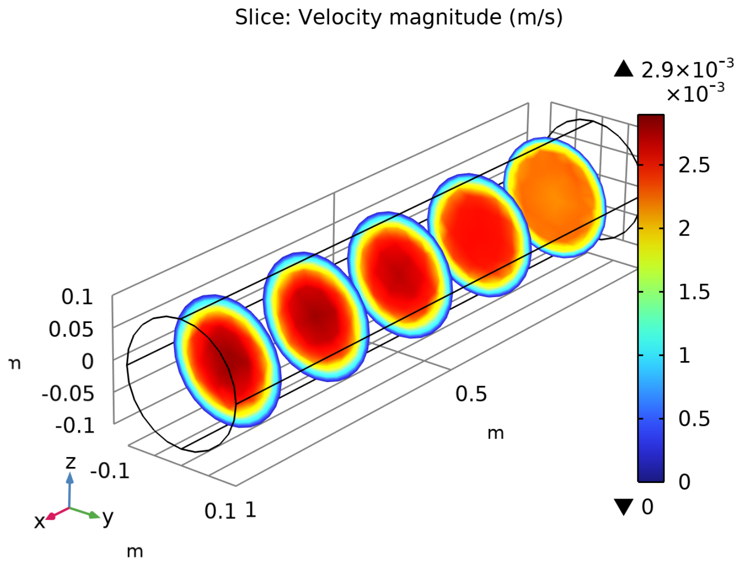
$$\boxed{D = 9.09\,\mu\text{m}^2\,\text{s}^{-1}}$$

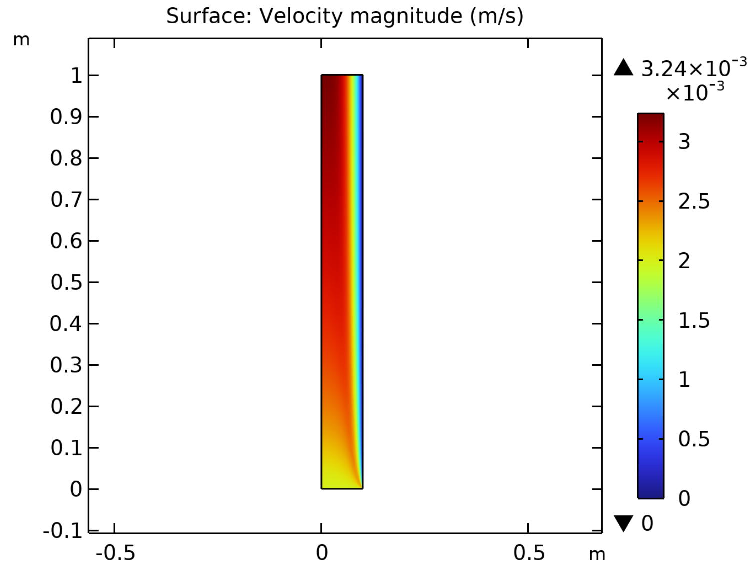
with coefficient of determination $R^2 = 0.984$.

4 Lab Questions

4.1 Problem 4

Complete the Lab 1 protocol. Show the flow profile for your model with printouts of `Lab01_3D.jpg` and `Lab01_2D.jpg`.





4.2 Problem 5

> The flow rate is the integral of velocity over any cross section of the flow. For example, in the 3D case, flow rate is the integral of $\int v \cdot \mathrm{d}A$ where v is the axial velocity, $\mathrm{d}A$ is area element and the integral is over the circular cross-section. Verify that the flow rate at the inlet is approximately equal to the flow rate at the outlet for your 3D model, and is also equal to the theoretical flow rate $v_{\text{in}}A$, where A is the area of the inlet, or the outlet.

The flow rate at the inlet and outlet, according to COMSOL, is

$$v_{\text{in}} = 6.0312 \times 10^{-5}\,\mathrm{m^3\,s^{-1}}; \quad v_{\text{out}} = 6.0228 \times 10^{-5}\,\mathrm{m^3\,s^{-1}}$$

The theoretical flow rate at the inlet or outlet is

$$v_{\text{in}}A = v_{\text{in}}\pi r^2 = \pi(0.02\,\mathrm{m\,s^{-1}})(0.1\,\mathrm{m})^2 = \boxed{6.0283 \times 10^{-5}\,\mathrm{m^3\,s^{-1}}}$$

where $A = \pi r^2$, $v_{\text{in}} = 0.02\,\mathrm{m\,s^{-1}}$, and $r = 0.1\,\mathrm{m}$.