Determination of Boltzmann's Constant through Statistical Analysis of 2D Brownian Motion

Sara Parvaresh Rizi (1010913451), Aarya Shah (1010871157)

March 31, 2026

Abstract

This experiment aimed to determine Boltzmann's constant, k, by observing the Brownian motion of microscopic beads suspended in water using video microscopy. The particle trajectories were tracked and analyzed with two distinct statistical models: the 1D mean squared displacement (MSD) as a function of time lag, and the probability distribution of 2D radial step lengths over a constant time interval of 0.5 s.

The MSD method yielded a highly inaccurate value of $k = (7 \pm 2) \times 10^{-22} \frac{J}{K}$, resulting in a 4711.22% difference from the accepted value. Conversely, analyzing the Rayleigh distribution of individual short-term step lengths was significantly closer to k. Extracting the diffusion coefficient through a curve fit of the step distribution yielded a 18.16%error with $k = (1.13 \pm 0.09) \times 10^{-23} \frac{J}{K}$, while the maximum likelihood estimate (MLE) provided the most accurate experimental result of $k = (1.2 \pm 0.1) \times 10^{-23} \frac{J}{K}$, with a percent difference of only 13.47%. The results confirm that the MLE method avoids histogram binning biases and is the most reliable statistical approach for extracting thermodynamic constants from discrete, noise-prone experimental tracking data.

1 Introduction

Brownian motion is the random motion of microscopic particles from thermal fluctuations, caused by collisions with fast-moving molecules in a fluid (Elsevier 2026). In this experiment, video microscopy is used to track the Brownian motion of microscopic beads in water to experimentally determine the value of Boltzmann's constant, k.

For an ensemble of particles stepping randomly left and right in 1 dimension every τ second, each particle steps a distance δ in either direction with equal probability. The position of the *i-th* particle, x_i after n steps, is

$$x_i(n) = x_i(n-1) \pm \delta$$

For N particles, the average position of the ensemble of particles is:

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

The mean displacement for these particles is zero, as $\delta \to 0$, whereas the mean squared displacement (MSD) is given by Equation 1 (*Physics 127b: Statistical Mechanics Brownian Motion* n.d.).

$$\langle x^2 \rangle = \frac{2kT}{\gamma} t \tag{1}$$

Here, k is Boltzmann's constant, T is the absolute temperature of the fluid and γ is the viscous drag coefficient of the particle.

In Equation 1, $\frac{kT}{\gamma}$ can be written as in Equation 2, by the Stokes-Einstein relation (Srivastava and Khanna 2009).

$$D = \frac{kT}{\gamma} \tag{2}$$

For a spherical particle of radius a moving through a fluid with dynamic viscosity η, the drag coefficient is given by Stokes' law as $\gamma = 6\pi\eta a$. Substituting this into Equation 2 yields $D = \frac{kT}{6\pi\eta a}$ (Serbanescu, Ladan, and Ryu 2025). Plotting ($\langle x^2 \rangle$) versus time lag (τ) gives a straight line whose slope can be used to find D.

The probability distribution for the displacement of a particle in 2D (or 3D) can be shown by a 2D Gaussian, with Equation 3 (*Mean Squared Displacement* 2018).

$$p(x, y; t) = \frac{1}{4\pi Dt} e^{-\frac{x^2+y^2}{4Dt}} \tag{3}$$

This differs from the 1D case, being $P(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}$, in the scaling factor and that the exponent changed to x^2+y^2. This gives two independent variables, causing the distribution to be a surface, not a curve. To convert to a curve, note that the distribution is isotropic and therefore depends only on radial distance, not direction; thus, the substitution $r^2 = x^2 + y^2$ can be made. To fully make this an integral substitution, for $a < r < b$, Equation 4 can be used.

$$P(r < R; t) = \int_0^{2\pi} \int_0^R \frac{1}{4\pi Dt} r e^{-\frac{r^2}{4Dt}} dr d\theta = \int_0^R \frac{r}{2Dt} e^{-\frac{r^2}{4Dt}} dr d\theta \tag{4}$$

The probability density function for the step size in a 2D diffusion process is therefore given by Equation 5.

$$p(r; t) = \frac{r}{2Dt} e^{\frac{-r^2}{4Dt}} \tag{5}$$

This is a Rayleigh distribution, where t is the 'step' size (not total time) and r is the distance moved in 1 'step,' *not* total distance (Nema 2022). The same derivation in 3 dimensions provides the Maxwell-Boltzmann equation (Yirsaw and Goshu 2024). Thus, if a position (or momentum) follows a

Gaussian distribution, its magnitude follows the Maxwell-Boltzmann distribution (Siegrist 2026). Looking at individual step sizes over a constant time interval (Δt) maps data to the Rayleigh distribution (Equation 5) to find D a second way.

2 Materials and Methods

The methods for this lab report adhere to the guidelines in the 'Thermal Motion' manual (Serbanescu, Ladan, and Ryu 2025), with the setup shown in Figure 1.

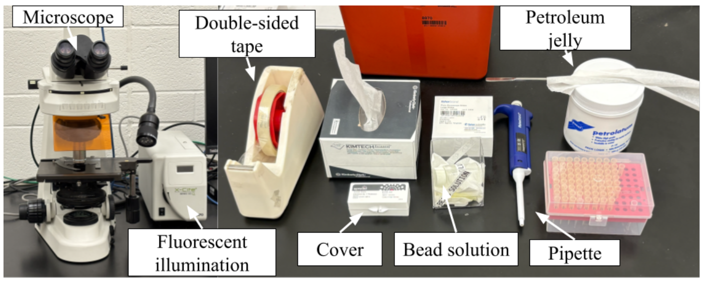

Figure 1: A photo of the equipment used throughout this lab to create and analyze the motion of the beads in solution. Included is 50 µL of bead solution, petroleum jelly, double-sided tape, a pipette, glass covers, fluorescence illumination (X-Cite Box), and a microscope.

A sealed sample chamber was constructed to observe the Brownian motion of fluorescent polystyrene beads (1.9 μm diameter) suspended in water. To prevent bulk fluid flow and evaporation, a 1.2 cm square wide, 3–4 cm long rectangular enclosure was created on a microscope slide using double-sided tape and petroleum jelly. 50 µL of the dilute bead solution was pipetted into the enclosure and sealed with a glass cover slip, pressing down gently to eliminate air bubbles (Figure 2).

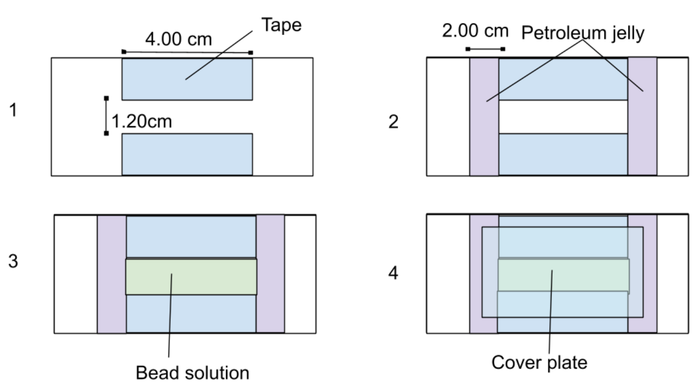

Figure 2: A diagram showing the progression of steps (1 to 4) to apply the bead solution to the microscope plate, alongside the dimensions used within this lab. The pictures of the real-life setup are shown in (Serbanescu, Ladan, and Ryu 2025).

The slide was placed on the stage of a fluorescence microscope equipped with an X-Cite Box illumination source and a 40x phase-contrast objective. To ensure accurate thermodynamic calculations, the ambient room temperature was recorded, being $296.5 \pm 0.5K$. Using the 'Microscope Camera Controller' software, the camera's gain and brightness were optimized to clearly isolate the fluorescent beads against a dark background. Video data was recorded using the Multiple Image Capture feature at a frame rate of 2 frames per second for 120 frames. This process was repeated across 10 different regions of the slide to capture a robust statistical sample. Following image acquisition, a stage micrometre (calibration slide) was imaged under the same 40x magnification to determine the pixel-to-meter conversion factor. The 'Image Object Tracker' software was then used to extract the 2D Cartesian coordinates of 2 distinct beads per video, yielding 20 independent particle trajectories. These raw pixel coordinates were subsequently exported to a Python script (see Section 6.1) to compute the MSDs and radial step distributions.

3 Data and Analysis

For this section, the microscope had a calibration conversion of 1 pixel equivalent to 0.12048 ± 0.003 μm. The bead diameter was stated by the manufacturer as 1.9 ± 0.1 μm. The viscosity of water was taken as 1.00 ± 0.05 cP (where 1 cP $= 10^{-3}$ Pa$\cdot$s) at 20°C, adjusted for a recorded ambient room temperature of 296.5 ± 0.5 K.

3.1 Mean-Squared Distance vs. Time

In 1 dimension, plotting the MSD against time lag (τ), using Equation 1, should yield a straight line with a slope equal to $2D$, where D is given by Equation 2. The diffusion coefficient D can then be used to calculate the Boltzmann constant by rearranging the Stokes-Einstein relation to $k = \frac{\gamma D}{T}$.

Instead of measuring the absolute distance a particle travels from its initial starting point at $t = 0$, the MSD is calculated using a rolling time lag (τ). A time lag represents the time interval between any two frames in the particle's trajectory. By calculating and averaging the squared displacements for all possible overlapping intervals of length τ within a single continuous track, the statistical power of the dataset is vastly increased (Spagnolo and Luin 2024). This rolling average smoothens the inherent random noise of thermal fluctuations, yielding a highly reliable ensemble average for the linear fit.

For each tracked particle, the squared distance travelled was averaged over all trajectories for each time lag and fitted linearly to extract D. This plot is shown in Figure 3.

From the linear fit in Figure 3, a diffusion coefficient $D = (1.1 \pm 0.3) \times 10^{-11}$ m^2/s was extracted. Using Equation 2, this yields a Boltzmann constant of $k = (7 \pm 2) \times 10^{-22}$ J/K. Comparing this to the accepted value of $k = 1.38 \times 10^{-23}$ J/K, there is an extremely large difference of 4711.22%, indicating a failure of the model assumptions (Serbanescu, Ladan, and Ryu 2025).

The fit in Figure 3 has a reduced chi-squared of 0.01. This exceptionally low value suggests the model fits the data "too well." This implies that the assigned statistical error bars on

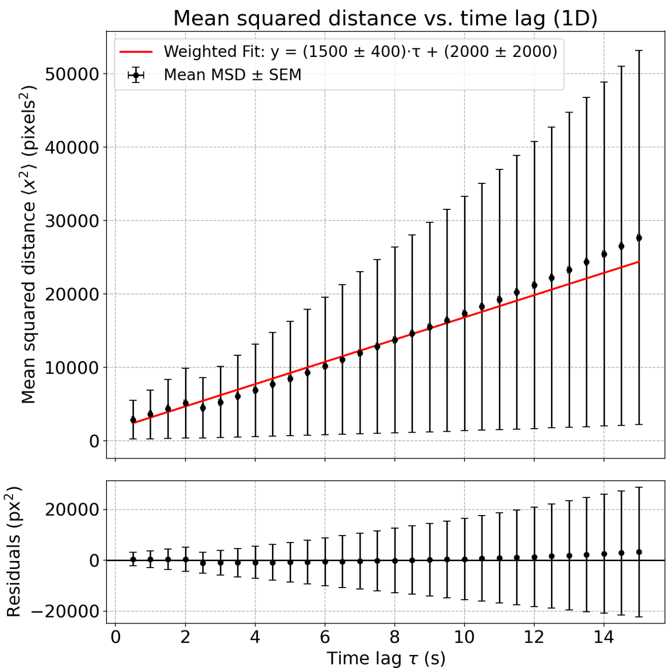

Figure 3: The 1D MSD of 20 tracked particles plotted against time-lag (τ). The linear fit, shown in red, has the equation $\langle x^2 \rangle = (1500 \pm 400)\tau + (2000 \pm 2000)$ px^2. The non-zero intercept suggests a static positional uncertainty or uncorrected drift, consistent with the discussion in 4.1. Vertical error bars representing the Standard Error of the Mean (SEM) are shown, while horizontal time uncertainties (0.03s) are too small to be visible. The mean residual is 4.81×10^2 px^2 with a standard deviation of 1.18×10^3 px^2.

the data points are overestimated, which is expected because MSD data points at sequential time lags are calculated from the same underlying trajectory and are therefore highly correlated, violating the assumption of independent, identically distributed (i.i.d.) errors required for standard chi-squared analysis (Shih and Fay 2017).

3.2 Probability Distribution of Step Lengths

In 2 dimensions, the theoretical expectation for the mean squared step length over a specific time interval is $\langle r^2 \rangle = \frac{4kT}{\gamma}\Delta t$. When fitting a probability distribution function to experimental step lengths, using an MLE is preferred to determine parameters, as changing histogram bin sizes can artificially skew calculations. Binning violates the least-squares assumptions of standard curve-fitting modules (Soto 2025). For the Rayleigh distribution, the maximum-likelihood estimate for the parameter $2D\Delta t$ is given by Equation 6 (Soch 2021).

$$(2D\Delta t)_{est} = \frac{1}{2N}\sum_{i=1}^{N} r_i^2 \tag{6}$$

From the bead position data, the 2D radial distance travelled ($r = \sqrt{\Delta x^2 + \Delta y^2}$) between consecutive frames ($\Delta t =$ 0.5 s) was calculated. To account for software tracking artifacts where the program temporarily lost the bead and snapped to a distant particle, unphysical step lengths larger than 30 pixels were filtered out of the dataset (see Section 6.1). This cleaned step size data was compiled into a single 1D array and plotted as a normalized histogram using the `hist` module from SciPy.

The probability density function (Equation 5) was fitted to the histogram to determine D_{fit}, and k_{fit} was calculated using Equation 2. Simultaneously, the MLE (Equation 6) was utilized directly on the unbinned step array to calculate D_{MLE} and subsequently k_{MLE}. The Rayleigh distribution using both estimates was plotted on the same axes over the empirical data, as shown in Figure 4.

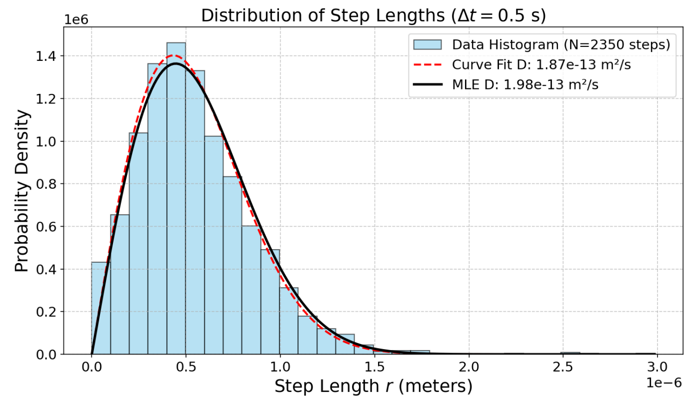

Figure 4: Normalized histogram of 2D step lengths (r) over a constant time interval ($\Delta t = 0.5$ s) with fitted Rayleigh PDFs (Equation 5). The MLE method, shown in black, is the statistically preferred model, as it directly utilizes the data's second moment, while the red dashed curve fit provides visual confirmation of the distribution shape.

From the standard curve fit, $D_{fit} = (1.9 \pm 0.1) \times 10^{-13}$ m^2/s, yielding $k_{fit} = (1.13 \pm 0.09) \times 10^{-23}$ J/K, which represents an 18.16% difference from the accepted value of $k = 1.38 \times 10^{-23}$ J/K (Serbanescu, Ladan, and Ryu 2025). Conversely, the preferred MLE method yielded $D_{MLE} = (2.0 \pm 0.1) \times 10^{-13}$ m^2/s and $k_{MLE} = (1.2 \pm 0.1) \times 10^{-23}$ J/K. This represents a percent difference of 13.47%, indicating strong agreement with the accepted value.

4 Discussion

The objective of this experiment was to determine Boltzmann's constant, k, by analyzing the Brownian motion of suspended fluorescent beads. Two distinct statistical approaches were used: the time-averaged mean squared distance (MSD) and the distribution of radial step lengths over a constant time interval. The results reveal a significant difference in how these two methods respond to systematic experimental errors.

4.1 Comparison of Methods to Calculate k

The MSD method yielded a significantly overestimated value for k with a percent difference of 4711.22%. This deviation indicates the presence of severe uncompensated systematic errors in the time-averaged tracking data. The most probable cause for this is bulk fluid drift. Over long tracking periods, slight evaporation, convection currents, or an uneven microscope stage can cause the fluid sample to slowly drift in one direction (Rodal 2019). While pure Brownian motion says that root-mean-square displacement scales as $x_{\text{rms}} \propto \sqrt{t}$, directed bulk flow causes distance to scale linearly with time ($x \propto t$), resulting in a parabolic, rather than linear, MSD curve (Maris et al. 2022). The MSD acquires an additional quadratic-in-time component due to superimposed drift, deviating from the expected linear relationship, which artificially inflates the slope of the linear fit, leading to a massive overestimation of the diffusion coefficient D and, consequently, k.

Furthermore, the reduced chi-squared value for the MSD linear fit was approximately 0.01. A value $\chi^2_\nu \ll 1$ indicates that the uncertainties on the data points were significantly overestimated (OriginLab n.d.). This occurs because the MSD values for sequential time lags are calculated from the same underlying continuous trajectory; therefore, the data points are highly correlated and violate the assumption of statistical independence required for standard chi-squared analysis. Increasing the total number of independent trajectories (sample size) would give the statistical test more power and provide a more accurate representation of variance (T 2019).

Conversely, analyzing the Rayleigh distribution of individual step lengths over a short, constant time interval ($\Delta t = 0.5$ s) proved to be more accurate. This method only looks at frame-to-frame jumps; thus, it successfully isolates the high-frequency random thermal motion while still being mostly immune to the slow and long-term effects of bulk fluid drift. Within this approach, the MLE had a percent difference of 13.47%, outperforming the standard curve fit (18.15%). This confirms that the MLE is the superior statistical tool, as it evaluates the raw second moment of the data ($\langle r^2 \rangle$) directly, avoiding biases and loss of information introduced by sorting data into histogram bins.

This highlights the importance of selecting statistical methods that respect the independence and timescale assumptions inherent in the underlying diffusion model.

4.2 Sources of Error

Several other sources of error (SOE) impacted the precision and accuracy of these calculations. For instance, the beads are approximately 2 μm in size, but the tracking software follows a rectangular bounding box that may not remain perfectly centered on the bead. Combined with fluctuations in the microscope's focus, this introduces a static positional uncertainty likely exceeding 0.1 μm. For example, a systematic drift velocity of even $0.1\mu m/s$ over a 60 s recording would contribute an additional displacement of 6μm, comparable to the expected Brownian displacement (around 5–7μm), thereby substantially inflating the MSD (Barbosa et al. 2018). This explains the large overestimation of the MSD slope and resulting diffusion coefficient. While a static error simply shifts the y-intercept of the MSD graph, severe tracking failures—such as the software snapping to a different bead—drastically skew the step distributions.

Secondly, to seal the slide chamber, petroleum jelly was used. If a small amount of this grease seeped into the interior bead solution, it would create local regions of higher viscosity. Because the calculations assumed the viscosity of pure water ($\eta = 1.00$ cP), an artificially higher viscosity in the actual sample would result in an inaccurate extraction of k.

Furthermore, the fluorescence illumination (X-Cite Box) acts as a powerful heat source. Over the course of the 10 tracking trials, the light likely heated the slide above the recorded ambient room temperature (296.5 ± 0.5 K). Because fluid viscosity decreases with temperature, even slight localized heating would cause the beads to diffuse faster than anticipated, skewing the final calculations (Wang 2025).

Lastly, the equations used in this lab assume purely 2D diffusion within the focal plane. However, the beads are suspended in a 3D fluid volume. If a bead diffuses upward or downward along the Z-axis, its apparent 2D projection on the camera slows down, or it blurs out of focus entirely, leading to an underestimation of its true 3D spatial step size.

5 Conclusion

This experiment successfully captured the random thermal motion of microscopic particles and demonstrated the mathematical relationship between microscopic diffusion and macroscopic thermodynamic properties. While the time-averaged mean squared distance approach failed to provide a reasonable estimate due to its extreme susceptibility to bulk fluid drift, analyzing the step-length distribution over short time intervals effectively mitigated this systematic error. By applying a maximum likelihood estimate to the Rayleigh distribution of 0.5 s step lengths, an experimental value of $k_{MLE} = (1.2 \pm 0.1) \times 10^{-23}$ J/K was determined. Therefore, the MLE result of $k = (1.2 \pm 0.1) \times 10^{-23}$ J/K is taken as the final experimental value. This result, with a percent difference of only 13.47% from the accepted value, confirms that short-interval statistical distributions are a highly robust method for extracting fundamental constants from noise-prone experimental tracking data.

References

Barbosa, Hugo et al. (Mar. 2018). "Human Mobility: Models and Applications". In: *Physics Reports* 734, pp. 1–74. DOI: 10.1016/j.physrep.2018.01.001. URL: https://www.sciencedirect.com/science/article/pii/S037015731830022X.

Elsevier (2026). *Brownian Movement - an overview.* ScienceDirect Topics. URL: https://www.sciencedirect.com/topics/engineering/brownian-movement (visited on 04/02/2026).

Maris, Erik et al. (June 2022). “Classification-based motion analysis of single-molecule trajectories using DiffusionLab”. In: *Scientific Reports* 12. DOI: 10.1038/s41598-022-13446-0. URL: https://www.nature.com/articles/s41598-022-13446-0#Fig2 (visited on 10/28/2024).

Mean Squared Displacement (2018). Fiveable.me. URL: https://fiveable.me/physical-chemistry-i/key-terms/mean-squared-displacement (visited on 04/03/2026).

Nema, Pranjul (Dec. 2022). *RPubs - RAY*. Rpubs.com. URL: https://rpubs.com/pranjul/985994 (visited on 04/03/2026).

OriginLab (n.d.). *Interpreting Regression Results*. www.originlab.com. URL: https://www.originlab.com/doc/Origin-Help/Interpret-Regression-Result.

Physics 127b: Statistical Mechanics Brownian Motion (n.d.). URL: http://mcc.caltech.edu/Ph127/b/Lecture15.pdf.

Rodal, Michael (Dec. 2019). *How to optimize the QCM baseline stability*. Biolinscientific.com. URL: https://www.biolinscientific.com/blog/how-to-optimize-the-qcm-baseline-stability (visited on 04/03/2026).

Serbanescu, Ruxandra, John Ladan, and W. Ryu (Mar. 2025). *Thermal Motion*. (Visited on 04/01/2026).

Shih, Joanna H. and Michael P. Fay (Feb. 2017). “Pearson’s chi-square test and rank correlation inferences for clustered data”. In: *Biometrics* 73, pp. 822–834. DOI: 10.1111/biom.12653.

Siegrist, Kyle (2026). *The Maxwell Distribution*. Randomservices.org. URL: https://www.randomservices.org/random//special/Maxwell.html (visited on 04/03/2026).

Soch, Joram (Nov. 2021). *Maximum likelihood estimation for simple linear regression*. The Book of Statistical Proofs. URL: https://statproofbook.github.io/P/slr-mle.html (visited on 04/03/2026).

Soto, Eddie (Feb. 2025). *The Art of Histograms: Finding the Perfect Bin Size for Data Visualization*. Medium. URL: https://medium.com/@eddie.sotocastillo/the-art-of-histograms-finding-the-perfect-bin-size-for-data-visualization-b2264aa3f103.

Spagnolo, Chiara Schirripa and Stefano Luin (Aug. 2024). “Trajectory Analysis in Single-Particle Tracking: From Mean Squared Displacement to Machine Learning Approaches”. In: *International Journal of Molecular Sciences* 25.16, p. 8660. DOI: 10.3390/ijms25168660. URL: https://www.mdpi.com/1422-0067/25/16/8660.

Srivastava, Rajat M and K.N Khanna (Apr. 2009). “StokesEinstein Relation in Two- and Three-Dimensional Fluids”. In: *ACS Publications* 54, pp. 1452–1456. DOI: 10.1021/je800698t. (Visited on 07/09/2023).

T, Dennis (July 2019). *Sample Size’s impact on Effect Size and Power*. Medium. URL: https://medium.com/@dtuk81/sample-sizes-impact-on-effect-size-and-power-fbd5084c7c47.

Wang, Stefan (2025). *How Does Viscosity Change with Temperature? - Martests Instrument — A Leading Rotational Viscometer Manufacturer in China*. Martests Instrument — A Leading Rotational Viscometer Manufacturer in China. URL: https://martests.com/how-does-viscosity-change-with-temperature-and-why-is-this-critical-for-your-business/.

Yirsaw, Adugna Gelaw and Ayele Taye Goshu (Jan. 2024). “Extended Rayleigh Probability Distribution to Higher Dimensions”. In: *Journal of Probability and Statistics* 2024. Ed. by Ramón M. Rodríguez-Dagnino. DOI: 10.1155/2024/7677855.

6 Appendix

6.1 Experimental Code

To maintain the brevity and readability of this report, the complete Python scripts used for the data analysis (including the mean squared distance linear regression, the Rayleigh maximum likelihood estimate, and all error propagation calculations) have been hosted externally. The custom scripts, along with the raw ‘.txt‘ tracking coordinates and generated output files, can be accessed and reviewed in full at the following GitHub repository:

https://github.com/sarapr06/thermal-motion

6.2 Mean Squared Distance

To convert tracking data (with units of pixels squared) into physical area (meters squared), Equation 7 was used.

$$\langle x^2\rangle_{true} = MSD_{px} \times c^2 \quad (7)$$

Because the calibration factor is squared, its relative uncertainty is doubled inside the sum due to differentiation, given below.

$$\delta\langle x^2\rangle_{true} = \langle x^2\rangle_{true} \times \sqrt{(\frac{\delta MSD_{px}}{MSD_{px}})^2 + (2\frac{\delta c}{c})^2} \quad (8)$$

The variables used in these errors are defined as follows:

- MSD_{px}: Raw 1D MSD for specific time lag
- c: Calibration factor, taken from the microscope in the lab having 1 px = 0.12048 ± 0.003 µm, thus, this value is 0.12048×10^{-6}m/px
- δc: Uncertainty in calibration factor, being 0.003×10^{-6} m/px (or 3.0×10^{-9} m/px)

One other variable was δMSD_{px}, the statistical uncertainty of the MSD at a specific time lag (Equation 10). This was calculated as the standard error of the mean (SEM) across all N tracked particles. First, the sample standard deviation (σ_{MSD}) of the squared displacements at that time lag was calculated:

$$\sigma_{MSD} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i^2 - \langle x^2\rangle\right)^2} \quad (9)$$

The variables are as follows:

- N = The total number of particle trajectories analyzed. Here, this was 20.
- x_i^2 = The squared displacement of the i-th individual particle at that specific time lag.
- $\langle x^2 \rangle$ = The average squared displacement of all particles at that time lag (MSD_{px} value).

The standard error (δMSD_{px}) is then found by dividing the standard deviation by the square root of the number of trajectories:

$$\delta MSD_{px} = \frac{\sigma_{MSD}}{\sqrt{N}} \tag{10}$$

6.3 Error for Boltzmann's Constant k

Using the 1D equation (Equation 1) and Stokes-Einstein relation (Equation 2), substitute the true slope of the graph, $m_{true} = m_{px} \times c^2$. Then, if $m_{true} = 2D$ (Equation 1 and 2), we get:

$$k = \frac{3\pi\eta r m_{px} c^2}{T} \tag{11}$$

Constant factors (e.g., 3 and π) do not contribute to relative uncertainty and are omitted; thus, the relative uncertainty is the sum of the squares of relative uncertainties of all measured variables, being:

$$\delta k = k\sqrt{(\frac{\delta\eta}{\eta})^2 + (\frac{\delta r}{r})^2 + (\frac{\delta m_{px}}{m_{px}})^2 + (2\frac{\delta c}{c})^2 + (\frac{\delta T}{T})^2} \tag{12}$$

The variables used are defined below:

- m_{px}: Slope of linear fit (see script) in $pixels^2/s$
- δm_{px}: Standard error of slope, being the statistical uncertainty of fit, usually given by linear regression.
- c and δc: Same calibration factor and error as in Section 6.2
- r: Radius of bead, being 1.9 µm (Serbanescu, Ladan, and Ryu 2025)
- δr: Uncertainty of bead radius, being 0.1 µm (Serbanescu, Ladan, and Ryu 2025)
- T: Absolute room temperature in Kelvin, being 296.5 K (Serbanescu, Ladan, and Ryu 2025)
- δT: Reading error of thermometer, being 0.5 K (Serbanescu, Ladan, and Ryu 2025)
- η: Dynamic viscosity of water at specific temperature, T, being 1.00 centipoise ($\frac{g}{cms}$) $\delta\eta$: Uncertainty in viscosity, being 0.05 ($\frac{g}{cms}$) (Serbanescu, Ladan, and Ryu 2025)

Using the MLE equation (Equation 6), the error for k will remain the same; however, D will be changed to its corresponding D_{MLE}.

6.4 MLE Error

From the maximum-likelihood estimate of the Rayleigh Distribution, given in Equation 6, the diffusion coefficient is calculated directly from the mean of step sizes, given in Equation 13.

$$D_{MLE} = \frac{\langle r^2 \rangle_{true}}{4\Delta t} = \frac{c^2 \langle r_{px}^2 \rangle}{4\Delta t} \tag{13}$$

The constants are as follows:

- c: Calibration factor (m/pixel), given in Section 6.3
- $\langle r_{px}^2 \rangle$: The average of all squared step lengths in pixels.
- Δt: The constant time interval (0.5 s)

The time interval Δt and the constant 4 have zero uncertainty in this theoretical limit. The relative uncertainty stems from the statistical spread of the jump sizes ($\delta\langle r_{px}^2 \rangle$) and the calibration factor (δc), given in Equation 14.

$$\frac{\delta D}{D} = \sqrt{\left(\frac{\delta\langle r_{px}^2 \rangle}{\langle r_{px}^2 \rangle}\right)^2 + \left(2\frac{\delta c}{c}\right)^2} \tag{14}$$

Where $\delta\langle r_{px}^2 \rangle$ is the Standard Error of the Mean (SEM) of the squared step array, calculated as $\frac{\sigma_{r_{px}^2}}{\sqrt{N}}$ (N being the total number of individual steps analyzed).