Simulation of One-Dimensional Brownian Motion by Stochastic Differential Equations

Ken Muranaka

K’s Garden Nishioji, Suite 401, Kasuga Hachijo Sagaru, Minami-ku, Kyoto 601-8312, Japan; kenneth@mbox.kyoto-inet.or.jp

Brownian motion is an important concept in chemistry and in other fields of science and engineering, and has several times been a topic in this Journal (1–8). Because the most popular application of Brownian motion is concerned with diffusion or transport phenomena, the theory of Brownian motion along with the relevant mathematics has usually been discussed in the context of deterministic models involving analytical solutions of ordinary or partial differential equations (ODEs or PDEs) such as the Fick’s laws.

However, a more appropriate model to describe a random process like Brownian motion is stochastic (or probabilistic), not deterministic. To be more precise, models formulated by stochastic differential equations (SDEs) are more versatile when simulating random processes. Diffusion and transport are not the only area where Brownian dynamics plays an important role. SDEs are being applied in molecular dynamics (9). In the near future various applications of SD E s in chemistry will become more popular as the computational speed of personal computers increases.

In this article, I suggest a simple, economical, and effective way to help students grasp what the diffusion process really is in terms of a modern stochastic simulation technique. A simulation program has been written in Microsoft Quick Basic, and can also be run by Microsoft QBASIC. To analyze and visualize the results of the stochastic simulations, spreadsheet software such as Lotus 1-2-3 or Excel is needed. This also gives students an opportunity to learn how to perform basic statistical analyses with a spreadsheet. All the exercises presented here can be done in one normal laboratory session, and should be suitable for inclusion in an introductory physical chemistry laboratory.

Theory

Brownian motion is a random walk, where the current position is not dependent upon the past positions or the paths undertaken. It resembles a drunk student wandering around in a campus quad. In mathematical language, random walk is called a stochastic process, or random process.

A foundation for the modern theory of stochastic processes was prepared in the 1940s by the Japanese mathematician Kiyosi Ito (15, 16). The theory of SDEs is today a hot area among probabilists, and its applications are being explored (9–11). Some rules in the stochastic calculus are quite different from those in regular calculus, and its rigorous treatment requires a background in measure theory (19, 20) far beyond the scope of undergraduate chemistry. However, a simple SDE can still demonstrate its versatility and power when applied for simulating random processes such as a one-dimensional diffusion process, and its connection to ODEs and PDEs can intuitively be seen as well.

As an example, the following SDE is considered throughout this article:

\[ dx = \mu dt + \sigma dz \]  

where \( x \) is position of, say, a particle, and \( t \) is time. Thus, \( dx \) is a small change in position; \( dt \), a small change in time. A change in position is described by the two terms on the right-hand side (\( \mu dt \), and \( \sigma dz \)); \( \mu \) is the expected or mean drift rate (the drift per unit of time), and can be understood easily if the second term \( \sigma dz \) is ignored. Then, the eq 1 becomes just

\[ dx = \mu dt \]  

which will give the following solution with the initial position \( x_0 \):

\[ x = x_0 + \mu t \]  

Without the \( \sigma dz \) term in the SDE, the linear eq 3 implies that \( x \), the position of a particle, has an expected or mean drift rate of \( \mu \) per unit time.

Now, added to the drift term \( \mu dt \) of \( x \) is the noise, or the random shock term \( \sigma dz \), where \( \sigma \) is a standard deviation of the noise and thus determines the extent of an uncertainty in the position of a particle in a small time interval. Sometimes \( \sigma \) is referred to as a diffusion coefficient; \( z \) is a random variable whose value changes over time in an uncertain way, and is termed a stochastic or random process.

A stochastic process can be either continuous or discrete. In a discrete process, the values of a random variable can be observed only at certain fixed points in time. Equation 1 represents a continuous-time stochastic process.

A random variable \( z \) has a property that the previous paths followed by a particle are all irrelevant to the present position. This is known as the Markov property, and it implies that predictions of future positions are completely uncertain and must therefore be based upon a probability distribution. Which probability distribution should we use?

In physical science, when a normal or Gaussian distribution is used, the stochastic process \( z \) is commonly called the Brownian motion. A random variable \( z \) changes over time in an uncertain way, and its increments having a normal distribution are uncorrelated over time by the Markov property. The probability distribution can now be incorporated into the random shock term \( \sigma dz \) in the following manner:

\[ dz = \varepsilon \sqrt{dt} \]  

in which \( \varepsilon \) is a normally distributed random variable with its mean 0 and standard deviation 1; in other words, the random variable \( \varepsilon \) is the standardized Gaussian noise. Let’s not think about a constant \( \sigma \) for a moment. Equation 4 says that
the variability of the position \( x \) taken randomly by a particle is proportional to the square root of a small time interval \( \Delta t \). Why do we have to define the \( dz \) in such a way? The reason will become clear if we consider eq 4 in terms of a larger time interval \( \Delta t \), that is, in discrete time.

Imagine that a particle with the initial position \( x_0 \) at the time \( t = 0 \) "randomly walks" to a final position \( x_f \) observed at a time \( t = f \) as depicted in Figure 1. Divide the entire time interval \( f \) into \( n \) smaller intervals of the equal length \( \Delta t \) so that

\[
f = n \Delta t
\]

For each \( t \), a small change in \( z \) is related to \( \Delta t \) by

\[
\Delta z = \varepsilon \sqrt{\Delta t}
\]

Then, the variance of \( \Delta z \) is just \( \Delta t \) as the variance of the standardized normal distribution \( \varepsilon \) is 1; that is, if \( \text{Var} \) is the variance operator, we see

\[
\text{Var}(\Delta z) = \text{Var}(\varepsilon \sqrt{\Delta t}) = \varepsilon^2 \text{Var}(\sqrt{\Delta t}) = \varepsilon^2 \Delta t
\]

This is due to a well-known property of random variables (13, 18). For the random variables \( X \) and \( Y \), if \( Y = cX \) where \( c \) is a constant, then the variance of \( Y \) is \( c^2 \text{Var}(X) \).

Because the stochastic process \( z \) is independent of previous paths, and the increments \( \Delta z \) are also independent, variances can be summed up to give the result

\[
\text{Variance of } (z_i - z_{i-1}) = n \Delta t = f
\]

Hence, the standard deviation of \( \Delta z \) is \( \sqrt{f} \). In a discrete time interval \( \Delta t \), a change in the position of a particle will be normally distributed with a mean of 0 and a standard deviation of \( \sqrt{f} \). The degree of uncertainty about the future position of a particle is proportional to the square root of how far in time one is looking ahead.

As \( \Delta t \to 0 \), \( \Delta z \) may be thought of as its limit \( dz \). But \( dz \) is really an unpredictable, infinitesimal, random shock, displaying extreme variation even in a very, very small time interval. For this reason, \( z \) is continuous but not differentiable everywhere. There lies a reason why ordinary calculus cannot be applied for SDEs. As can be seen in the simulation exercises, the random shock term \( dz \) not only adds a Gaussian noise to the drift term but also gives "kinks" to a path generated by eq 1, no matter how small the \( \Delta t \) may be. A constant parameter \( \sigma \) defines the extent of an uncertainty regarding an unpredictable, infinitesimal variation.

The whole random shock given by the \( \sigma \Delta z \) term can then be "drifted" up or down by a constant parameter \( \mu \). Therefore, an SDE as presented in the form of the eq 1 is very simple but contains all the essential components to model the dynamics of Brownian motion. The paths generated by SDEs are called sample paths or trajectories. Figure 1 is an example.

**Simulations and Results**

The program to simulate one-dimensional Brownian motion is named RWALK, which stands for "random walk". It can be downloaded from JCE Online. RWALK is written in Microsoft Quick Basic language. Both the source code and compiled version are available. RWALK can simulate positions of a particle starting at an initial position \( x_0 \) over some time periods, both the initial position and the periods to be simulated need to be entered as input values. A new period corresponds to a \( \Delta t \), which may be nanosecond, second, minute, hour, day, month, year, or any regular time interval for which an observation is made. A drift \( \mu \) and a standard deviation \( \sigma \) must be provided as well. It is convenient to set the initial time \( t \) to be 0, as stochastic processes are usually defined over a time course \( \geq 0 \).

RWALK is an interactive program. The first thing it will ask for is the kind of simulation to be done. RWALK can (i) simulate one sample path and (ii) iteratively simulate as many sample paths as needed to generate a data set so that confidence intervals and frequency distributions of the positions at 6 different periods can be obtained with the aid of a spreadsheet or statistical package. All the results can be written to a designated file on a disk.

To generate Figure 2, I used initial value = 0, drift = 0.1, and standard deviation = 1 over a time course of 1000 periods; this simulation was repeated 5 times to obtain 5 different sample paths. Lotus 1-2-3 was used to plot the paths. A straight line in this figure represents the constant drift without the random shock term. Note that all the paths are drifting around this straight line ("a trend line") with many tiny twists or kinks, but deviating farther and farther away from it as time passes. This is exactly what we are expecting to see: our uncertainty should increase as the square root of time in Brownian motion. As an exercise, generate several sample paths with different \( \mu \) and \( \sigma \), compute the mean and sample standard deviation for each of these paths and compare the estimates with the inputs (theoretical values). To compute the means and the sample standard deviations, first use spreadsheet software to compute the changes in position of a particle for each period; that is, compute for each period

\[
\Delta x = x_i - x_{i-1}
\]

where \( i = 1, 2, \ldots \). These changes are approximations of instantaneous changes in position, \( dx \), as modeled by eq 1. Then, simply use the statistical functions or a statistical package to compute the mean and standard deviation for each sample path generated. The mean and the standard deviation to be estimated here are the constant parameters \( \mu \) and \( \sigma \) in eq 1. All standard deviations from the sample paths should be very close to the input value for \( \sigma \). For \( \mu \), this is not the case. But the mean or the median of all the means computed from the sample paths should be close enough to the input drift pa...
the displacement because

When the standard deviation of distributions is the same as the root-mean-square (rms) of displacements for a large number of observations. To compute the sample standard deviation of displacements at a certain point in time $t$, the following formula can be used.

$$s = \sqrt{\frac{(\Delta x_1 - m)^2 + (\Delta x_2 - m)^2 + \ldots + (\Delta x_n - m)^2}{n - 1}} \quad (10)$$

When $m = 0$ and $n$ is large, $s$ is simply equal to the rms of the displacements because $n$ is approximately equal to $n - 1$.

For this particular case, the rms of the displacements is $\sqrt{2D \Delta t}$, where $D$ is the famous diffusion coefficient found in textbooks (21–24). The rms of the displacements is also the standard deviation of Gaussian (normal) distribution at a certain time $t$. This is the probability distribution of finding particles at different positions $x$ at times $t$ if the random walk can be modeled by Brownian motion without drift. Students can now go beyond this level because eq 1 is a model for diffusion with drift.

In eq 1, a constant noise measured by standard deviation $\sigma$ is assumed. If experimental data do not show a constant $\sigma$, Brownian motion cannot be modeled by a simple SDE such as eq 1. In such cases, $\sigma$ may be expressed as a function of time. The drift term can also be a function of time. Added to the random shock term $dz$ may be the Poisson jump diffusion process $dq$, where a finite jump can occur due to an abrupt change in the system and $q$ is a Poisson process (whereas the random variable $z$ in $dz$ is a Gaussian process). Unlike ODEs and PDEs, SDEs require no analytical solutions and can model a variety of systems by computer simulation.

The data analysis capability of Lotus 1-2-3 or Excel can be used to get a frequency distribution of simulated values for each period. Figure 4 is a three-dimensional plot summarizing all the frequency distributions at the six periods specified. Note how the spread of each bell-shaped frequency distribution gets wider as time passes. This graph is visually demonstrating that our uncertainty about the future position of a particle indeed increases as time progresses, and the centers of the bell-shaped curves are shifted upward owing to the drift. This is a one-dimensional Brownian motion seen in three dimensions by frequency distributions (which can be thought of as probability distributions for the random walks of a particle) that get wider and wider over time. As we increase the number of simulation runs, the frequency distribution curves will approach a theoretical limit of smooth, bell-shaped curves of Gaussian probability distribution, the distribution governing Brownian motions.
need not be constant and can be expressed as a function of time. A Poisson process can be added as well for sudden changes. SDEs are indeed very flexible and versatile when modeling dynamic and unstable systems.

As a laboratory exercise, students can model some of their own vital signs—pulse, blood pressure, breathing rate, or any other signs that randomly fluctuate over time. In this exercise, the model to be used is eq 1, a Gaussian process, and stochastic modeling involving a Poisson process should be reserved for more advanced students.

The following is a time-series record of my pulse taken just after I woke up around 6 a.m.:

67, 68, 70, 72, 69, 68, 68 (counts/minute)

Each pulse was recorded every minute as continuously as possible. We are interested in modeling the random fluctuations of the wrist pulse. Use eq 9 to compute the changes of the pulse ($\Delta x$), and estimate the mean and standard deviation for this stochastic component. We are approximating eq 1, a continuous-time process, by a discrete-time model. So, $\Delta t = 1$ minute, and $\Delta x$ is assumed to be normally distributed with mean $\mu \Delta t$ and standard deviation $\sigma \sqrt{\Delta t}$; therefore, the variance of this stochastic component $\Delta x$ is $\sigma^2 \Delta t$. For this data set, $\mu = 0.167$ min, and $\sigma = 1.94 \sqrt{\text{min}}$. Then, simply enter these parameters into eq 1 to get the SDE

$$dx = 0.167 dt + 1.94 dz$$ (11)

Run the program RWALK to generate a frequency distribution, which can be thought of as a probability distribution of the pulse, say, 30 minutes later. Then, record the wrist pulse to see if the count is actually within a predicted range given by this frequency distribution. (If a student's pulse rate is above the upper limit, it must be just before taking a final exam.)

For my pulse data, after 100 simulation runs (or by generating 100 sample paths), the 95% confidence interval at the 30th time period was between 57 and 92. For the initial value, I used the first pulse count taken when I had decided to make a prediction by simulation, and it was 70 counts per minute. Note that this initial value differs from one simulation to another. To obtain a 95% level of confidence for the distribution (or 99% for a higher level of confidence), sort all the values at the 30th time period in ascending order, and take the average of the 97th and 98th values for the upper limit, and the average of the 2nd and 3rd values for the lower limit. The upper and lower limits are just the percentile points of the 100 pulses generated, which correspond to the 100 possible outcomes at the 30th time period. My pulse taken 30 minutes later was 75 counts per minute—within the 95% confidence interval as predicted. Here lies the beauty of stochastic modeling: that any confidence interval can be obtained even in the case where the probability density function for the stochastic system is not available in analytic form.

This exercise using a student's own vital sign is simple and economical, but presents a real-life example of how a random process can be modeled by SDEs. In addition, this exercise raises a problem with regard to eq 1, the simplest model for Brownian motion. If a proposed SDE like eq 11 is used to model a situation in which a total simulation horizon (length of the simulation) is much longer than 30 minutes (such as days or weeks), negative values and unrealistically

Suggested Exercises

Although Brownian motion is a useful and important topic to chemists, its implementation in terms of SDEs is not as easily found in the chemical literature as in the literature of engineering and econometrics (14, 17). A good example in chemistry would be the molecular dynamics of polymers (9). I am hoping that teaching SDEs at the undergraduate level will motivate chemistry students, the future chemists, to apply stochastic models in various fields of chemistry. This article, per se should be useful because Brownian motion has never been presented with SDEs in physical chemistry for undergraduate students.

The question to be addressed in class is, “How can we have a control over a random process so that we can predict an unexpected event ahead of time in an unstable system?” A stochastic or random process is a real-valued process, a function of time, and can model a process largely governed by chance. The soybean price listed at the Chicago Board of Trade is a stochastic process, and can be modeled by an SDE (17). In fact, any process that randomly fluctuates or randomly vibrates over time can be modeled by SDEs. So, as a homework exercise, students can propose a chemical system that is governed by random fluctuations due to chance, and therefore can be modeled by an SDE such as eq 1. Which variable in a system is the random variable to be modeled, and what are the justifications for it? Also ask students to propose an experiment to estimate reliable values for the two important parameters $\mu$ and $\sigma$, which in eq 1 are assumed to be constant over time. To what extent is the assumption of constant $\mu$ and $\sigma$ reasonable? In more complex SDEs, these parameters

$\mu$ and $\sigma$ are the justifications for it? Also ask students to propose an experiment to estimate reliable values for the two important parameters $\mu$ and $\sigma$, which in eq 1 are assumed to be constant over time. To what extent is the assumption of constant $\mu$ and $\sigma$ reasonable? In more complex SDEs, these parameters

Figure 4. A three-dimensional graph done by Excel to show how the frequency distribution changes over time. Note that the spread or standard deviation as a measure of uncertainty gets wider and wider as time passes, and the center of each frequency distribution is shifted or “drifted” upward by a constant parameter called the drift rate in the SDE (eq 1). As the number of simulation runs (or sample paths generated by RWALK) increases, each frequency distribution will approach a bell-shaped normal curve because the SDE is a Gaussian stochastic process governed by the random shock term $dz$. For this reason, the frequency distribution obtained by computer simulations can be thought of as a resulting probability distribution for the Brownian dynamics under consideration.

$\text{JChemEd.chem.wisc.edu} \quad \text{Vol. 76} \quad \text{No. 7} \quad \text{July 1999} \quad \text{Journal of Chemical Education}$
large values for the pulse will be generated by RWALK. Why? Because eq 11 is a diffusion equation that does not take homeostasis into account. Our body has an inherent feedback mechanism to keep the vital signs within a certain range or around an equilibrium level. In other words, a feedback mechanism works to revert an extreme vital sign to a trend line defined by nature. Random processes like this can still be modeled, and an SDE may be what is known as a mean-reverting process, where the random variables fluctuate while reverting to the mean in a long run. Occasional jumps (sudden changes in temperature or heart rate) can also be modeled by incorporating a Poisson jump process.

The exercise to simulate the pulse can still surprise students with the great predictive power of SDEs as long as the simulation horizon is kept adequately short. It would also be a good exercise to suggest a maximum length of time for which the simulation model is reliable and appropriate for this exercise.

Conclusion

A simple stochastic differential equation (eq 1) was used as a motivation to learn that important phenomena such as diffusion and transport in chemistry can be explained in terms of Brownian dynamics. Unlike the classical approach to teach Brownian motion with analytical solutions of ODEs or PDEs, computer simulations that allow one to “see” the paths (trajectories) followed by a particle should give students a better intuition regarding random processes governing such physical phenomena as diffusion and should also be inspiring for further work in physical science. The Basic program RWALK is an easy-to-use tool for teaching not only Brownian motion but also some basic statistical concepts essential to physical scientists and engineers.

In this article, I have suggested some exercises for students, hoping that such laboratory experience will motivate them to further studies of stochastic processes and to find chemical applications of SDEs.

Note

*Supplementary materials for this article are available on JCE Online at http://jchemed.chem.wisc.edu/journal/issues/1999/jul/ab94.html.*

**Literature Cited**