# 1. A Historical Introduction

## 1.1 Motivation

Theoretical science up to the end of the nineteenth century can be viewed as the study of solutions of differential equations and the modelling of natural phenomena by deterministic solutions of these differential equations. It was at that time commonly thought that if all initial data could only be collected, one would be able to predict the future with certainty.

We now know this is not so, in at least two ways. Firstly, the advent of quantum mechanics within a quarter of a century gave rise to a new physics, and hence a new theoretical basis for all science, which had as an essential basis a purely statistical element. Secondly, more recently, the concept of chaos has arisen, in which even quite simple differential equation systems have the rather alarming property of giving rise to essentially unpredictable behaviour. To be sure, one can predict the future of such a system given its initial conditions, but any error in the initial conditions is so rapidly magnified that no practical predictability is left. In fact, the existence of chaos is really not surprising, since it agrees with more of our everyday experience than does pure predictability—but it is surprising perhaps that it has taken so long for the point to be made.

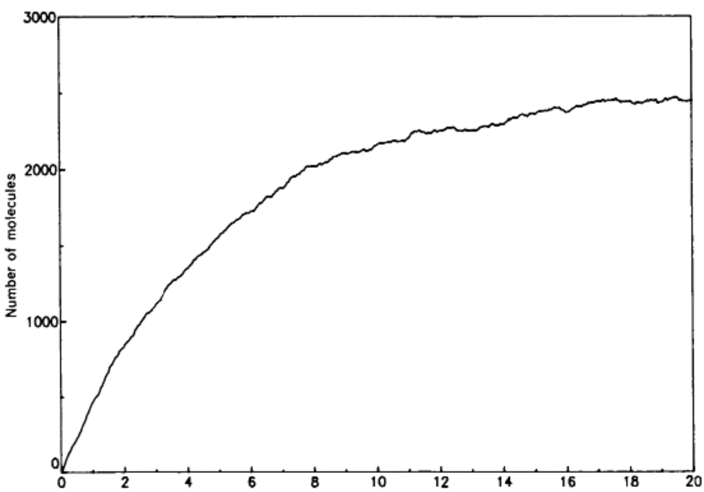


**Fig. 1.1.** Stochastic simulation of an isomerisation reaction $X \rightleftharpoons A$

Chaos and quantum mechanics are not the subject of this chapter. Here I wish to give a semihistorical outline of how a phenomenological theory of fluctuating phenomena arose and what its essential points are. The very usefulness of predictable models indicates that life is not entirely chaos. But there is a limit to predictability, and what we shall be most concerned with in this book are models of limited predictability. The experience of careful measurements in science normally gives us data like that of Fig. 1.1, representing the growth of the number of molecules of a substance $X$ formed by a chemical reaction of the form $X \rightleftharpoons A$. A quite well defined deterministic motion is evident, and this is reproducible, unlike the fluctuations around this motion, which are not.

## 1.2 Some Historical Examples

### 1.2.1 Brownian Motion

The observation that, when suspended in water, small pollen grains are found to be in a very animated and irregular state of motion, was first systematically investigated by Robert Brown in 1827, and the observed phenomenon took the name *Brownian Motion* because of his fundamental pioneering work. Brown was a botanist—indeed a very famous botanist—and of course tested whether this motion was in some way a manifestation of life. By showing that the motion was present in any suspension of fine particles—glass, minerals and even a fragment of the sphinx—he ruled out any specifically organic origin of this motion. The motion is illustrated in Fig. 1.2.

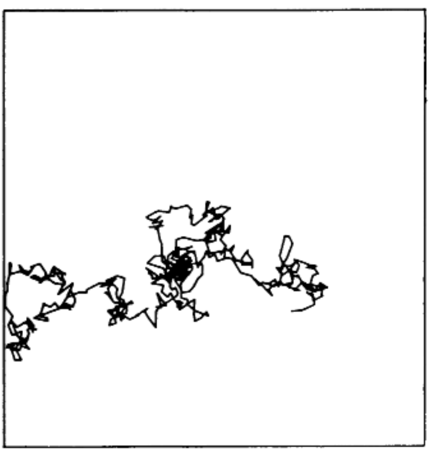

**Fig. 1.2.** Motion of a point undergoing Brownian motion

The riddle of Brownian motion was not quickly solved, and a satisfactory explanation did not come until 1905, when *Einstein* published an explanation under the rather modest title "über die von der molekular-kinetischen Theorie der

Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen" (concerning the motion, as required by the molecular-kinetic theory of heat, of particles suspended in liquids at rest) [1.2]. The same explanation was independently developed by *Smoluchowski* [1.3], who was responsible for much of the later systematic development and for much of the experimental verification of Brownian motion theory.

There were two major points in Einstein's solution to the problem of Brownian motion.

(i) The motion is caused by the exceedingly frequent impacts on the pollen grain of the incessantly moving molecules of liquid in which it is suspended.
(ii) The motion of these molecules is so complicated that its effect on the pollen grain can only be described probabilistically in terms of exceedingly frequent statistically independent impacts.

The existence of fluctuations like these ones calls out for a statistical explanation of this kind of phenomenon. Statistics had already been used by Maxwell and Boltzmann in their famous gas theories, but only as a description of possible states and the likelihood of their achievement and not as an intrinsic part of the time evolution of the system. *Rayleigh* [1.1] was in fact the first to consider a statistical description in this context, but for one reason or another, very little arose out of his work. For practical purposes, Einstein's explanation of the nature of Brownian motion must be regarded as the beginning of stochastic modelling of natural phenomena.

Einstein's reasoning is very clear and elegant. It contains all the basic concepts which will make up the subject matter of this book. Rather than paraphrase a classic piece of work, I shall simply give an extended excerpt from Einstein's paper (author's translation):

"It must clearly be assumed that each individual particle executes a motion which is independent of the motions of all other particles; it will also be considered that the movements of one and the same particle in different time intervals are independent processes, as long as these time intervals are not chosen too small.

"We introduce a time interval $\tau$ into consideration, which is very small compared to the observable time intervals, but nevertheless so large that in two successive time intervals $\tau$, the motions executed by the particle can be thought of as events which are independent of each other.

"Now let there be a total of $n$ particles suspended in a liquid. In a time interval $\tau$, the $X$-coordinates of the individual particles will increase by an amount $\Delta$, where for each particle $\Delta$ has a different (positive or negative) value. There will be a certain *frequency law* for $\Delta$; the number $dn$ of the particles which experience a shift which is between $\Delta$ and $\Delta + d\Delta$ will be expressible by an equation of the form

$$dn = n\phi(\Delta)d\Delta, \tag{1.2.1}$$

where

$$\int\limits_{-\infty}^{\infty} \phi(\Delta)d\Delta = 1 \tag{1.2.2}$$

and $\phi$ is only different from zero for very small values of $\Delta$, and satisifes the condition

$$\phi(\Delta) = \phi(-\Delta). \tag{1.2.3}$$

"We now investigate how the diffusion coefficient depends on $\phi$. We shall once more restrict ourselves to the case where the number $\nu$ of particles per unit volume depends only on $x$ and $t$.

"Let $\nu = f(x, t)$ be the number of particles per unit volume. We compute the distribution of particles at the time $t + \tau$ from the distribution at time $t$. From the definition of the function $\phi(\Delta)$, it is easy to find the number of particles which at time $t + \tau$ are found between two planes perpendicular to the $x$-axis and passing through points $x$ and $x + dx$. One obtains

$$f(x, t + \tau)dx = dx \int_{-\infty}^{\infty} f(x + \Delta, t)\phi(\Delta)d\Delta \,. \tag{1.2.4}$$

But since $\tau$ is very small, we can set

$$f(x, t + \tau) = f(x, t) + \tau \frac{\partial f}{\partial t} \,. \tag{1.2.5}$$

Furthermore, we develop $f(x + \Delta, t)$ in powers of $\Delta$:

$$f(x + \Delta, t) = f(x, t) + \Delta \frac{\partial f(x, t)}{\partial x} + \frac{\Delta^2}{2!} \frac{\partial^2 f(x, t)}{\partial x^2} + \cdots \,. \tag{1.2.6}$$

We can use this series under the integral, because only small values of $\Delta$ contribute to this equation. We obtain

$$f + \frac{\partial f}{\partial \tau}\tau = f \int_{-\infty}^{\infty} \phi(\Delta)d\Delta + \frac{\partial f}{\partial x} \int_{-\infty}^{\infty} \Delta\phi(\Delta)d\Delta + \frac{\partial^2 f}{\partial x^2} \int_{-\infty}^{\infty} \frac{\Delta^2}{2} \phi(\Delta)d\Delta \,. \tag{1.2.7}$$

Because $\phi(x) = \phi(-x)$, the second, fourth, etc., terms on the right-hand side vanish, while out of the 1st, 3rd, 5th, etc., terms, each one is very small compared with the previous. We obtain from this equation, by taking into consideration

$$\int_{-\infty}^{\infty} \phi(\Delta)d\Delta = 1 \tag{1.2.8}$$

and setting

$$\frac{1}{\tau} \int_{-\infty}^{\infty} \frac{\Delta^2}{2} \phi(\Delta)d\Delta = D \,, \tag{1.2.9}$$

and keeping only the 1st and third terms of the right-hand side,

$$\frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial x^2} \cdots \,. \tag{1.2.10}$$

This is already known as the differential equation of diffusion and it can be seen that $D$ is the diffusion coefficient. ...

"The problem, which corresponds to the problem of diffusion from a single point (neglecting the interaction between the diffusing particles), is now completely determined mathematically: its solution is

$$f(x, t) = \frac{n}{\sqrt{4\pi D}} \frac{e^{-x^2/4Dt}}{\sqrt{t}} \ldots \tag{1.2.11}$$

"We now calculate, with the help of this equation, the displacement $\lambda_x$ in the direction of the $X$-axis that a particle experiences on the average or, more exactly, the square root of the arithmetic mean of the square of the displacement in the direction of the $X$-axis; it is

$$\lambda_x = \sqrt{\bar{x}^2} = \sqrt{2Dt}\,.\text{"} \tag{1.2.12}$$

Einstein's derivation is really based on a discrete time assumption, that impacts happen only at times $0, \tau, 2\tau, 3\tau \ldots$, and his resulting equation (1.2.10) for the distribution function $f(x, t)$ and its solution (1.2.11) are to be regarded as approximations, in which $\tau$ is considered so small that $t$ may be considered as being continuous. Nevertheless, his description contains very many of the major concepts which have been developed more and more generally and rigorously since then, and which will be central to this book. For example:

i) *The Chapman-Kolmogorov Equation* occurs as Einstein's equation (1.2.4). It states that the probability of the particle being at point $x$ at time $t + \tau$ is given by the sum of the probability of all possible "pushes" $\Delta$ from positions $x + \Delta$, multiplied by the probability of being at $x + \Delta$ at time $t$. This assumption is based on the independence of the push $\Delta$ of any previous history of the motion: it is only necessary to know the initial position of the particle at time $t$—not at any previous time. This is the *Markov postulate* and the Chapman Kolmogorov equation, of which (1.2.4) is a special form, is the central dynamical equation to all Markov processes. These will be studied in detail in Chap. 3.

ii) *The Fokker-Planck Equation:* Eq. (1.2.10) is the diffusion equation, a special case of the Fokker-Planck equation, which describes a large class of very interesting stochastic processes in which the system has a continuous sample path. In this case, that means that the pollen grain's position, if thought of as obeying a probabilistic law given by solving the diffusion equation (1.2.10), in which time $t$ is continuous (not discrete, as assumed by Einstein), can be written $x(t)$, where $x(t)$ is a *continuous function of time*-but a random function. This leads us to consider the possibility of describing the dynamics of the system in some direct probabilistic way, so that we would have a *random* or *stochastic differential equation for the path*. This procedure was initiated by Langevin with the famous equation that to this day bears his name. We will discuss this in detail in Chap. 4.

iii) *The Kramers-Moyal* and similar expansions are essentially the same as that used by Einstein to go from (1.2.4) (the Chapman-Kolmogorov equation) to the diffusion equation (1.2.10). The use of this type of approximation, which effectively replaces a process whose sample paths need not be continuous with one whose paths are continuous, has been a topic of discussion in the last decade. Its use and validity will be discussed in Chap. 7.

### 1.2.2 Langevin's Equation

Some time after Einstein's original derivation, *Langevin* [1.4] presented a new method which was quite different from Einstein's and, according to him, "infinitely more simple." His reasoning was as follows.

From statistical mechanics, it was known that the mean kinetic energy of the Brownian particle should, in equilibrium, reach a value

$$\langle \tfrac{1}{2}mv^2 \rangle = \tfrac{1}{2}kT \tag{1.2.13}$$

($T$; absolute temperature, $k$; Boltzmann's constant). (Both Einstein and Smoluchowski had used this fact). Acting on the particle, of mass $m$ there should be two forces:

i) a viscous drag: assuming this is given by the same formula as in macroscopic hydrodynamics, this is $-6\pi\eta a\, dx/dt$, $\eta$ being the viscosity and $a$ the diameter of the particle, assumed spherical.

ii) another *fluctuating force* $X$ which represents the incessant impacts of the molecules of the liquid on the Brownian particle. All that is known about it is that fact, and that it should be positive and negative with equal probability. Thus, the equation of motion for the position of the particle is given by Newton's law as

$$m\frac{d^2x}{dt^2} = -6\pi\eta a\frac{dx}{dt} + X \tag{1.2.14}$$

and multiplying by $x$, this can be written

$$\frac{m}{2}\frac{d^2}{dt^2}(x^2) - mv^2 = -3\pi\eta a\frac{d(x^2)}{dt} + Xx\,, \tag{1.2.15}$$

where $v = dx/dt$. We now average over a large number of different particles and use (1.2.13) to obtain an equation for $\langle x^2 \rangle$:

$$\frac{m}{2}\frac{d^2\langle x^2\rangle}{dt^2} + 3\pi\eta a\frac{d\langle x^2\rangle}{dt} = kT\,, \tag{1.2.16}$$

where the term $\langle xX \rangle$ has been set equal to zero because (to quote Langevin) "of the irregularity of the quantity $X$". One then finds the general solution

$$\frac{d\langle x^2\rangle}{dt} = kT/(3\pi\eta a) + C\exp(-6\pi\eta at/m)\,, \tag{1.2.17}$$

where $C$ is an arbitrary constant. Langevin estimated that the decaying exponential approaches zero with a time constant of the order of $10^{-8}$ s, which for any practical observation at that time, was essentially immediately. Thus, for practical purposes, we can neglect this term and integrate once more to get

$$\langle x^2 \rangle - \langle x_0^2 \rangle = [kT/(3\pi\eta a)]t\,. \tag{1.2.18}$$

This corresponds to (1.2.12) as deduced by Einstein, provided we identify

$$D = kT/(6\pi\eta a)\,, \tag{1.2.19}$$

a result which Einstein derived in the same paper but by independent means.

Langevin's equation was the first example of the *stochastic differential equation*—a differential equation with a random term $X$ and hence whose solution is, in some sense, a random function. Each solution of Langevin's equation represents a different random trajectory and, using only rather simple properties of $X$ (his fluctuating force), measurable results can be derived.

One question arises: Einstein explicitly required that (on a sufficiently large time scale) the change $\Delta$ be completely independent of the preceding value of $\Delta$. Langevin did not mention such a concept explicitly, but it is there, implicitly, when one sets $\langle Xx \rangle$ equal to zero. The concept that $X$ is extremely irregular *and* (which is not mentioned by Langevin, but is implicit) that $X$ and $x$ are *independent* of each other—that the irregularities in $x$ as a function of time, do not somehow conspire to be always in the same direction as those of $X$, so that the product could possibly not be set equal to zero; these are really equivalent to Einstein's independence assumption. The method of Langevin equations is clearly very much more direct, at least at first glance, and gives a very natural way of generalising a dynamical equation to a probabilistic equation. An adequate mathematical grounding for the approach of Langevin, however, was not available until more than 40 years later, when Ito formulated his concepts of stochastic differential equations. And in this formulation, a precise statement of the independence of $X$ and $x$ led to the calculus of stochastic differentials, which now bears his name, and which will be fully developed in Chap. 4.

As a physical subject, Brownian motion had its heyday in the first two decades of this century, when Smoluchowski in particular, and many others carried out extensive theoretical and experimental investigations, which showed complete agreement with the original formulation of the subject as initiated by himself and *Einstein*, see [1.5]. More recently, with the development of laser light scattering spectroscopy, Brownian motion has become very much more quantitatively measurable. The technique is to shine intense, coherent laser light into a small volume of liquid containing Brownian particles, and to study the fluctuations in the intensity of the scattered light, which are directly related to the motions of the Brownian particles. By these means it is possible to observe Brownian motion of much smaller particles than the traditional pollen, and to derive useful data about the sizes of viruses and macromolecules. With the preparation of more concentrated suspensions, interactions between the particles appear, generating interesting and quite complex problems related to macromolecular suspensions and colloids [1.6].

The general concept of fluctuations describable by such equations has developed very extensively in a very wide range of situations. The advantages of a continuous description turn out to be very significant, since only a very few parameters are required, i.e., essentially the coefficients of the derivatives in (1.2.7):

$$\int_{-\infty}^{\infty} \Delta\phi(\Delta)d\Delta, \quad \text{and} \quad \int_{\infty}^{\infty} \Delta^2\phi(\Delta)d\Delta\,. \tag{1.2.20}$$

It is rare to find a problem which cannot be specified, in at least some degree of approximation, by such a system, and for qualitative simple analysis of problems it is normally quite sufficient to consider an appropriate Fokker-Planck equation, of a form obtained by allowing both coefficients (1.2.20) to depend on $x$, and in a space of an appropriate number of dimensions.

## 1.3 Birth-Death Processes

A wide variety of phenomena can be modelled by a particular class of process called a birth-death process. The name obviously stems from the modelling of human or animal populations in which individuals are born, or die. One of the most entertaining models is that of the prey-predator system consisting of two kinds of animal, one of which preys on the other, which is itself supplied with an inexhaustible food supply. Thus letting $X$ symbolise the prey, $Y$ the predator, and $A$ the food of the prey, the process under consideration might be

$$X + A \rightarrow 2X \tag{1.3.1a}$$

$$X + Y \rightarrow 2Y \tag{1.3.1b}$$

$$Y \rightarrow B \tag{1.3.1c}$$

which have the following naive, but charming interpretation. The first equation symbolises the prey eating one unit of food, and reproducing immediately. The second equation symbolises a predator consuming a prey (who thereby dies—this is the only death mechanism considered for the prey) and immediately reproducing. The final equation symbolises the death of the predator by natural causes. It is easy to guess model differential equations for $x$ and $y$, the numbers of $X$ and $Y$. One might assume that the first reaction symbolises a rate of production of $X$ proportional to the product of $x$ and the amount of food; the second equation a production of $Y$ (and an equal rate of consumption of $X$) proportional to $xy$, and the last equation a death rate of $Y$, in which the rate of death of $Y$ is simply proportional to $y$; thus we might write

$$\frac{dx}{dt} = k_1ax - k_2xy \tag{1.3.2a}$$

$$\frac{dy}{dt} = k_2xy - k_3y\,. \tag{1.3.2b}$$