Vol. 29 (1998) *ACTA PHYSICA POLONICA B* No 6

STATISTICAL PROPERTIES OF DYNAMICAL SYSTEMS WITH DISTURBANCES: VARIATION IN PARAMETERS*

A. Kleczkowski

King's College
Cambridge CB2 1ST, England
e-mail: ak133@cus.cam.ac.uk

(Received March 4, 1997)

Statistical properties of kinetic equations are studied for reactions in which the (effective) rate decays to zero with time. For such systems the final state depends on initial condition and on the parameters. Time evolution of the probability distribution associated with a concentration of one of the reagents is studied, and analytical formulas are obtained for the case when the parameters are drawn from a random sample, but remain constant for a particular realization. Even if the underlying distribution of the parameters is symmetrical, the resulting distribution of the concentration is highly skewed. This results in a magnification of variability as small differences in the parameters lead to high levels of variability in the outcome of the reaction. The magnification of the variability is also quantified using a concept analogous to the Lyapunov exponent in chaos theory.

PACS numbers: 02.30. q, 02.50. Ey, 02.50. Cw

1. Introduction

The theory of stochastic processes aims at a prediction of the behaviour of dynamical systems, from initial conditions until the final state [1]. However, the resulting equations are usually too difficult to solve, and one is left with the information about a final steady state. It is also commonly assumed that the initial state of the system is not relevant for the long-term dynamics as the system settles on its (widely understood) attractor determined by the dynamics and stochastic perturbations. Yet, there are often situations when

* Submitted for the Proceedings of the Marian Smoluchowski Symposium on Statistical Physics, Zakopane, Poland, September 1–10, 1997.

the transient behaviour is more interesting than the final state. This is especially true for applications of stochastic systems in biochemistry and biology. For such systems, characteristic time scales for observation and/or changes in parameters are comparable with relaxation times needed to equilibrate the system. In another interesting class of systems, the final state depends on the history of the dynamics and in particular on the initial state [2]. We shall consider some examples of such systems below.

In the paper we study some simple chemical reactions for which the probability of one or more reactions is a decreasing function of time. The question originally arose in the context of population dynamics [3–5] where the ability of species to interact with each other declines in time due to *e.g.* environmental forcing or an internal clock. However, many processes in biology and medicine can be characterized by a similar 'race against time'. Other examples include chemical reactions limited by an availability of catalytic substances which become inactive or exhausted over time, crystallization processes and crystal growth in the presence of evaporation.

In studying the evolution of chemical and biological systems, the traditional approach often concentrates on the fluctuations resulting from the finite size of the system and/or small numbers of reaction agents. This mainly represents the within-sample variability, understood here in terms of factors limiting our ability to predict the behaviour of the system at later times based on our knowledge of the past for the same replicate. Fluctuations in the dynamical behaviour of the system also affect the between-sample variability, reducing a predictability of the behaviour of one replicated system from the past records for another replicate. However, there is another component of the between-sample variability, caused by variations in initial conditions, reaction rates and environmental factors. In this paper we concentrate on the latter aspect, assuming that the within-sample part is negligibly small so that once the parameters are drawn from a random sample, the dynamics is deterministic. Such assumptions are equivalent, *e.g.* to the approximation of a large size of an individual replicate.

We start by presenting a suite of models of increasing complexity, characterized by a decay in the effective rate of the reaction. We study the evolution of a probability distribution associated with the concentration of one of the reagents. We assume that the evolution is deterministic but the parameters are drawn from a random sample, so that the model is described in terms of probabilistic equations. Finally, we use a concept analogous to the Lyapunov exponent to quantify the variability. Appendix A contains some mathematical details related to the analytical solutions of one of the examples, whereas Appendix B contains general results for the error propagation.

2. Models

The basic model describes the reaction of the type

$$A \quad + \quad B \quad \xrightarrow{k(t)} \quad 2A \;, \tag{1}$$

with $k(t) \to 0$ when $t \to \infty$. Note that $a + b = n = \text{const}$, where a and b denote concentrations of A and B. Two cases can be distinguished here. One (Example 1) has $n \gg a$ so that B is always in abundance throughout the period when $k(t) \gg 0$. In this case, the reaction is limited by the time decay of the rate. In another case (Example 2), the abundance of b limits the reaction in addition to the decay of $k(t)$.

The decay in the rate can be caused by an external forcing or by a limited supply of an additional catalytic reagent as in

$$\begin{aligned} A \quad + \quad B \quad + \quad X \quad &\xrightarrow{k} \quad 2A \quad + \quad X \\ X \quad &\xrightarrow{g} \quad 0 \qquad\quad . \end{aligned} \tag{2}$$

This leads to a special form of $k(t)$ in (1), $k(t) = k_0 \exp(-gt)$.

Another example of this type of reactions occurs for a simple catalytic reaction, as in Example 3

$$B \quad + \quad Y \quad \xrightarrow{k(t)} \quad A \quad + \quad Y \;, \tag{3}$$

where $k(t) \to 0$ with $t \to \infty$. Alternatively,

$$\begin{aligned} B \quad + \quad Y \quad &\xrightarrow{k} \quad A \quad + \quad Y \\ B \quad &\xrightarrow{g} \quad 0 \qquad\quad . \end{aligned} \tag{4}$$

The order of the reaction (1) can be higher than 1, with individual reactions summarizing more complicated processes, as in Example 4,

$$pA \quad + \quad qB \quad \xrightarrow{k(t)} \quad (p+q)A \;, \tag{5}$$

where $p \geq 0$ and $q \geq 0$. p and q do not need to be integer.

Another situation arises when one of the reagents is depleted by a reaction (or series of reactions), as in Example 5:

$$\begin{aligned} A \quad + \quad B \quad &\xrightarrow{k} \quad 2A \\ B \quad &\xrightarrow{g} \quad 0 \,, \end{aligned} \tag{6}$$

where the second reaction can summarize several reactions. More complicated reaction schemes can be also analysed using methods discussed below.

3. Results

In the following we are mainly interested in the dynamics of the reagent A. We begin by writing deterministic equations describing the kinetics represented by Equations (1)–(6). Exact analytical solutions are obtained whenever possible, and approximate solutions derived when no exact solution is available. We then use the formulas to obtain a general form of the probability distribution associated with a given distribution of the parameters and initial conditions. The variability can be associated with the between-samples variability. This section is split into two parts, the first discussing the deterministic dynamics with fixed parameters. The second part concentrates on the error propagation due to uncertainty in the parameters and/or initial conditions. We also discuss one of the examples in detail.

3.1. Deterministic dynamics

General results In the following, the concentrations of A, B, *etc.* are denoted by small letters a, b, *etc.* Apart from Example 5, all other equations can be written in a general form

$$\frac{da}{dt} = k(t) f(a)\,. \tag{7}$$

The solution can be obtained in an implicit form

$$\mathcal{F}(a) \equiv \int\limits_{a_0}^{a} \frac{da}{f(a)} = \int\limits_{0}^{t} k(\tau)\, d\tau\,, \tag{8}$$

where the initial conditions in the form $a(t) = a_0$ for $t = 0$ are assumed. Equation (8) also defines the functional $\mathcal{F}(a)$.

Note that the transformation of the independent variable t into $k(t)$ removes the time dependence in Eq. (7), but does not provide any equilibrium (long-term) values. The transformation becomes singular at $t \to \infty$, so that it is often easier to analyse equations in original variables.

Example 1 If the rates in Eq. (1) satisfy the inequality $-dk/dt \gg g$, the dynamics are dominated by the decay in the rates, and we can write the equation as

$$\frac{da}{dt} \simeq k(t) n a\,. \tag{9}$$

The solution can be obtained for the initial value problem $a(t=0)=a_0$ and is listed in Table I. A special case of $k(t)=k_0\exp(-gt)$ is given in Table I as well.

TABLE I

Analytical solutions for Example 1. General solution is shown together with a special case for $k(t)=k_0\exp(-gt)$, and the asymptotic value.

Example	1
Reaction	$A+B\to 2A$
General solution	$a_0\exp\left(n\int\limits_0^t k(\tau)\,d\tau\right)$
Special case	$a_0\exp\left[\frac{nk_0}{g}\left(1-\exp(-gt)\right)\right]$
Asymptotic value	$a_0\exp\left(\frac{nk_0}{g}\right)$

The solution to Eq. (9) is known as a Gompertz function and is often used in biology to describe population and organ growth. The shape is sigmoidal with a single inflection point and asymptotic value depending on the rates and the initial condition, $a^*=a_0\exp\left(\frac{nk_0}{g}\right)$ (a star here and below denotes the deterministic equilibrium value). Figure 1(a) gives some examples of the solution for k_0 varying according to a Gaussian distribution with the mean $\mu=1/n$ and the standard deviation $\sigma=0.1/n$. Note that although the rate varies by $\pm 10\%$, the equilibrium value a^* differs by an order of magnitude.

Example 2 If the concentrations of A described by (1) reach levels comparable with concentration of B during the period when $k(t)\gg 0$, we need to use the full kinetic equation

$$\frac{da}{dt}\simeq k(t)(n-a)a\,. \tag{10}$$

with $a+b=n$ and therefore $b=n-a$. The equation can be solved analytically [4]. See Table II for the general form and a special case of exponential decay in the rates. Figure 1(b) shows examples of the solutions $a(t)$. The equilibrium levels of a are reduced as compared with example 1, and they are reached earlier, but there is little change in the qualitative dynamics. A solution to Eq. (10) for $k(t)=\text{const}$ is called a logistic function.

Example 3 The equation for $k(t)=\text{const}$ is known as a monomolecular equation, and can be generalized as follows

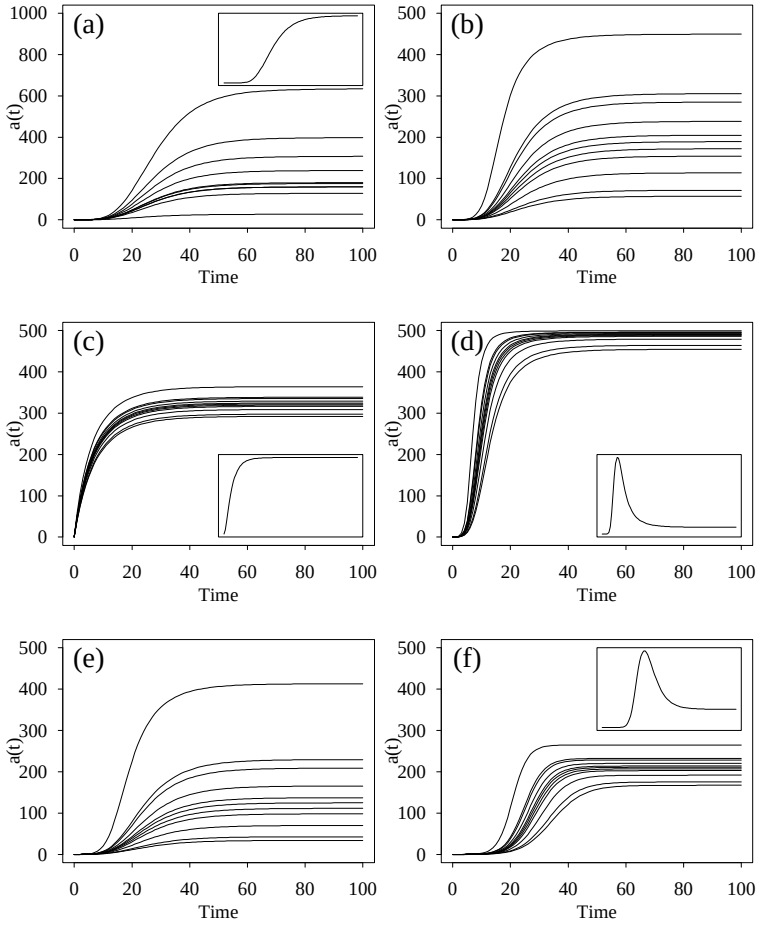


Fig. 1. Solutions for $a(t)$ for examples considered in the text. (a) corresponds to example 1 (exponential dynamics), (b) to example 2 (logistic dynamics), (c) to example 3 (monomolecular dynamics). (d) and (e) show the sensitivity of the dynamics to changes in p in example 4 (a generalized logistic dynamics). (f) represents example 5 (system of equations). Insets show the dynamics of the variance for (a), (c), (d) and (f). For (b) and (e) the dynamics are qualitatively the same as in (a). 10 randomly chosen solutions starting from the same initial condition $a_0 = 0.01$ are shown, with $k(t) = k_0 \exp(-gt)$ and k_0 varying according to a Gaussian distribution with the mean $\mu = 1/n$ and the standard deviation $\sigma = 0.1\mu$. Other parameters are $g = 0.1$ and $n = 500$, apart from (c) where $\mu = 0.1/n$ and (f) where $\mu = 0.5/n$ and $g = 0.03$. In (d), $p = 0.9$, whereas in (e) $p = 1.01$; $q = 1$. Time is in arbitrary units

TABLE II

The same as Table I, but for the Example 2. $F(t) \equiv \exp\left(n \int\limits_0^t k(\tau)\, d\tau\right)$

Example	2
Reaction	$A + B \to 2A$
General solution	$\frac{nF(t)}{\frac{n-a_0}{a_0}+F(t)}$
Special case	$n\left\{1+\frac{n-a_0}{a_0}\exp\left[-\frac{nk_0}{g}\left(1-\exp(-gt)\right)\right]\right\}^{-1}$
Asymptotic value	$n\left\{1+\frac{n-a_0}{a_0}\exp\left(-\frac{nk_0}{g}\right)\right\}^{-1}$

TABLE III

The same as Table I, but for the Example 3. $F(t) \equiv \exp\left(n \int\limits_0^t k(\tau)\, d\tau\right)$.

Example	3
Reaction	$B + Y \to A + Y$
General solution	$n - (n - a_0)\left\{F(t)\right\}^{-1}$
Special case	$n - (n - a_0)\exp\left[-\frac{nk_0}{g}\left(1-\exp(-gt)\right)\right]$
Asymptotic value	$n - (n - a_0)\exp\left(\frac{nk_0}{g}\right)$

$$\frac{da}{dt} \simeq k(t)(n-a)\,. \tag{11}$$

Eq. (11) can be easily solved to yield the results listed in Table III. The initial slope of $a(t)$ as a function of t, $k_0 n$ is larger compared to previous examples ($k_0 n a_0$), if a_0 is small, and the individual curves as well as the mean and variance do not exhibit any inflection point, figure 1(c). Note also a significantly reduced variability in the equilibrium values, caused by a linear rather than exponential character of the growth for small t.

Example 4 The nonlinear equation describing kinetics of Example 4

$$\frac{da}{dt} \simeq k(t)a^p(n-a)^q \tag{12}$$

has no analytical solution in a general case. For positive integer p (and any q) a recursive formula can be used in order to evaluate $\mathcal{F}(a)$ [6].

$$\int \frac{da}{a^p(n-a)^q} = \frac{-1}{(p-1)na^{p-1}(n-a)^q} - \frac{2-p-q}{n(p-1)} \int \frac{da}{a^{p-1}(n-a)^q} \quad (13)$$

$$\int \frac{da}{(n-a)^q} = \frac{1}{(q-1)(n-a)^{q-1}}\,. \quad (14)$$

Change of the variable from a to $n-a$ can be also used if q is a positive integer number, for any p.

Although the implicit equation (8) cannot be solved in this case for most values of p and q, the evolution of the probability distribution can be obtained if $\mathcal{F}(a)$ can be found; see below and Appendix B.

The solution $a(t)$ follows a sigmoidal curve as long as $p > 0$. The dynamics are very sensitive to changes in p, and less to changes in q. Figure 1(d) and 1(e) compare the results for $q = 1$ and $p = 0.9$ and $p = 1.01$; see also figure 1(b) corresponding to $p = 1.0$. Lowering p results in a behaviour dominated by the availability of B (equilibrium close to n and little variability), whereas increasing p above 1 results in a rapid drop of the average equilibrium concentration of A. Variability first increases, but then decreases again as $a^* \to 0$ with p increasing (figure 1(d)).

Example 5 In this example, B is depleted by two processes, a reaction with A and a separate decay. As a result, the dynamics is governed by a set of nonlinear equations with constant coefficients,

$$\frac{da}{dt} = kab\,, \quad (15)$$

$$\frac{db}{dt} = -kab - gb \quad (16)$$

with initial conditions $a(0) = a_0$ and $b(0) = b_0$. No general analytical solutions can be obtained. However, $\mathcal{F}(a)$ can be found in two special cases, when a is close to and when it is large compared with a_0. We show in Appendix A that a satisfies an equation

$$\frac{da}{dt} = ka\left(a_0 + b_0 - \frac{g}{k}\ln\left(\frac{a}{a_0}\right) - a\right). \quad (17)$$

The equilibrium value is given by a non-zero solution to the equation

$$a = a_0 + b_0 - \frac{g}{k}\ln\left(\frac{a}{a_0}\right). \quad (18)$$

Although no analytical formula can be found, it is clear that the equilibrium value depends on initial conditions as well as on the parameters. The evolution of the probability distribution can be obtained in this case, see below and Appendix B.

Some solutions of Eq. (16) are shown in figure 1(f), for different values of $k = k_0$. Note that although the overall shape of a single replicate is similar to other Examples (1, 2 and 4), the variability among the replicates follows a different pattern. There is a well pronounced increase in the variability around $t = 30$ (arbitrary units) followed by a decrease to a non-zero value at the equilibrium. The decrease is caused by the dynamics of b rather than by an approach to the maximum concentration as in figure 1(d) and Example 4.

3.2. Uncertainty due to parameter change

In the following we assume that the parameters (and/or initial conditions) are drawn from a certain random distribution, but they remain fixed throughout the course of the reaction. In particular, we assume that the rate $k_0 \equiv k(t = 0)$ is the main determining factor, and is distributed with a probability distribution $\rho(k_0)$. Alternatively, the initial condition for a (and/or b) as well as the decay rate g can vary between replicates.

Examples of the distributions include a uniform distribution on an interval $(\mu - \xi, \mu + \xi)$, $U(\mu - \xi, \mu + \xi)$ and a Gaussian distribution, $G(k_0; \mu, \sigma) = (2\pi\sigma^2)^{-1/2} \exp(-(k_0 - \mu)^2/(2\sigma^2))$. We denote the average value of k_0 by μ and the variance by σ^2.

In Appendix B we show how the uncertainty in the parameters propagates through the solution of the differential equation. For the varying rate, k_0, we show that

$$P\{a(t)\} = \left(\frac{\partial a(t)}{\partial k_0}\right)^{-1} \rho\left(k_0\left(a(t)\right)\right) , \tag{19}$$

where $P\{a(t)\}$ is the distribution of $a(t)$ at a given time. $\rho\left(k_0\left(a(t)\right)\right)$ denotes a function $\rho(k_0)$ where k_0 is substituted by an appropriate expression in terms of a. If k_0 is varying in the rate function $k(t) = k_0 \exp(-gt)$, then

$$P(a) = \frac{g}{(1 - \exp(-gt))\ f(a)}\ \rho\left(\frac{g\mathcal{F}(a)}{1 - \exp(-gt)}\right) . \tag{20}$$

This equation holds for Examples 1 through 4. Table IV lists $f(a)$ and $\mathcal{F}(a)$ so that the appropriate formulas for $P(a)$ can be obtained.

TABLE IV

Functionals $f(a)$ and $\mathcal{F}(a)$ for examples 1–4.

Example	$f(a)$	$\mathcal{F}(a)$
1	na	$n \ln\left(\frac{a}{a_0}\right)$
2	$(n-a)a$	$n \ln\left(\frac{a}{n-a} \,/\, \frac{a_0}{n-a_0}\right)$
3	$n-a$	$n \ln\left(\frac{n-a_0}{n-a}\right)$
4	$a^p(n-a)^q$	Recursive formula, (13) and (14)
5	$ka\left(b_0 - \frac{g}{k} \ln\left(\frac{a}{a_0}\right) - (a-a_0)\right)$	Available in special cases, see Appendix A

3.3. Particular example

We use Example 1 in order to illustrate the results obtained above. Figures 2(a) and 2(b) show sample solutions of Eq. (9) for the same initial conditions and the parameter k_0 varying between replicates (compare figure 2(a) and 2(b) with figure 1(a)). Other parameters are fixed, $g = 0.1$ and $n = 500$. k_0 was drawn from a uniform distribution on the interval $(0.9, 1.1)/n$ (figure 2(a)) or from a Gaussian distribution with the mean $1/n$ and the standard deviation $0.1/n$ (figure 2(b)). There is a considerable variation in the equilibrium values, much more pronounced for the Gaussian distribution which is unlimited at its upper tail. Although k_0 should always be strictly positive, we did not consider a cut-off in the Gaussian distribution. Rare events of $k_0 < 0$ can be interpreted as a reverse reaction, $A \longleftarrow 2A$.

Figures 2(c) and 2(d) compare the distribution $P(a)$ for $t \to \infty$ (equilibrium value). The distributions are well approximated by appropriate analytical formulas (see Table IV and Eq. (20)). The distributions are highly skewed, with a long tail for large values of a. Note the differences in limits for the horizontal axis in figures 2(c) and 2(d).

The availability of analytical formulas makes it possible to study the evolution of probability distribution $P(a)$ in time. Figure 3 compares the results of numerical solution of Eq. (9) for 500 values of k_0 with predictions of analytical formulas from Table IV and Eq. (20), for some values of t (arbitrary units). The agreement between numerical solutions and analytical prediction is very good, although the size of the sample was relatively small (500) as compared with the variability between replicates. The shape of

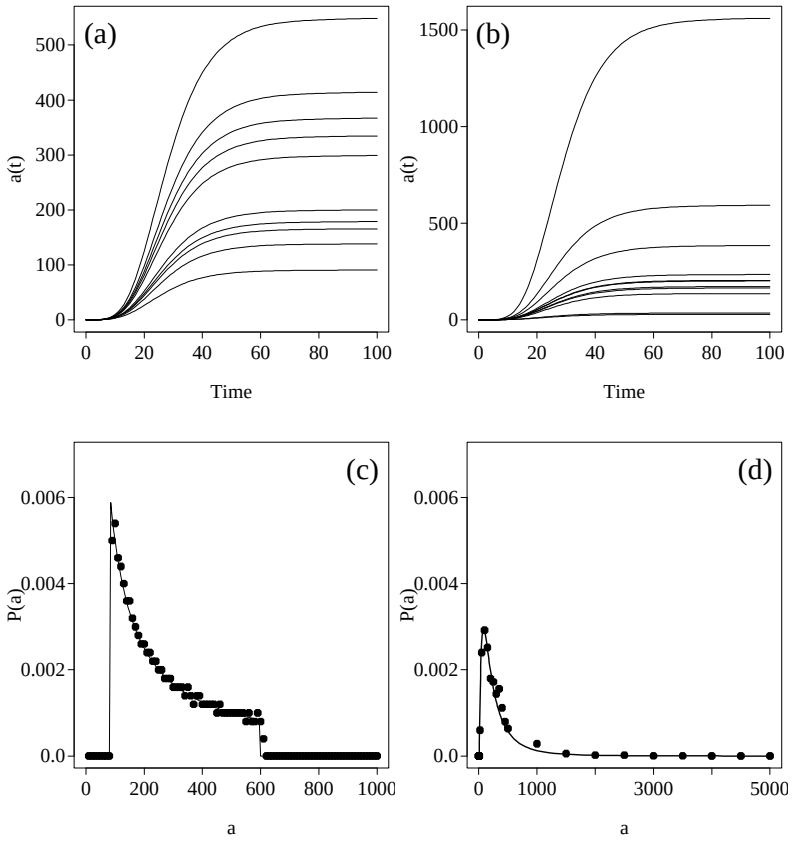


Fig. 2. Example 1: (a) and (b) show 10 random replicates of the solution of Eq. (9) for identical initial conditions $a_0 = 0.01$ and parameters $g = 0.1$ and $n = 500$, differing by the parameter k_0 drawn from a uniform distribution on the interval $(0.9, 1, 1)/n$ (a) and a Gaussian distribution with the mean $1/n$ and the standard deviation $0.1/n$ (b). Probability distributions $P(a)$ for $t \to \infty$ are shown in (c) (underlying uniform distribution) and in (d) (underlying Gaussian distribution), estimated by numerical solution of Eq. (9) for 500 values of k_0 (points) and predicted by analytical formulas in Table IV and Eq. (20) (lines), cf.Eq. (**??**)

$P(a)$ is Gaussian if plotted as a function of $\ln(a)$, reflecting the formula

$$\begin{aligned} P(a) &= \frac{g}{(1-\exp(-gt))\,a}\,\rho\left(\frac{g\ln(a/a_0)}{1-\exp(-gt)}\right) \\ &\sim \exp\left\{-\frac{\left(\ln(a/a_0)-\frac{1-\exp(-gt)}{g}\mu\right)^2}{2\sigma^2\left(\frac{1-\exp(-gt)}{g}\right)^2}\right\} \end{aligned} \tag{21}$$

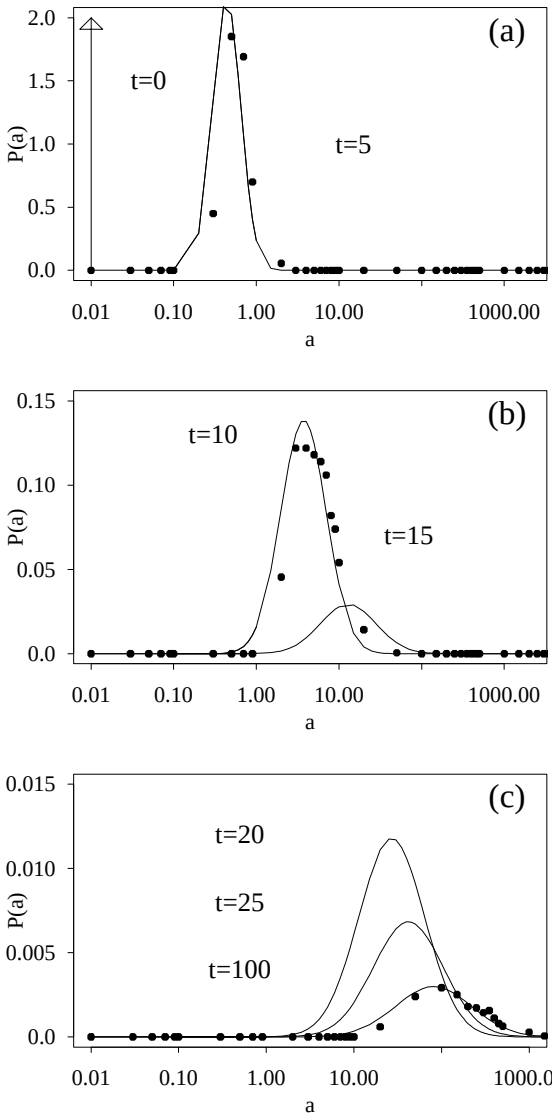


Fig. 3. Example 1: Evolution of the probability distribution $P(a)$ obtained by solving Eq. (9) with 500 different values of k_0 (points) and by applying analytical formulas, Eq. (21). $t = 0$ and 5 in (a), 10 and 15 in (b), and 20, 25 and 100 in (c) (time in arbitrary units). Solutions of Eq. (9) are shown for $t = 5$, 10 and 100 only. Other parameters as in figure 2. Note that the horizontal axis is logarithmic in this figure, but linear in figure 2(d)

The maximum of the distribution (the mode) moves according to the formula

$$a_{\max} = a_0 \exp\left(\frac{1-\exp(-gt)}{g}\,\mu\right)$$

and follows the deterministic dynamics with the average parameter $k_0 = \mu$. Note that the distribution $P(a)$ is not symmetric, although the distribution of k_0 and of $P(\ln(a))$ are symmetric. As a result of this, the mode and median of the distribution of $P(a)$ coincide, but the mean can be very different. We explore this idea further elsewhere.

3.4. Characterizing the variability

The magnification of differences in initial conditions is often characterized by the Lyapunov exponent [7, 8]. Since the solutions of the equations considered here always settle on the equilibrium, the concept of the Lyapunov exponent is not applicable. Nevertheless, we can discuss the dependence of equilibrium values on the parameters. An analogue of the Lyapunov exponent has been introduced in [3]. Define χ_j in terms of a partial derivative of the final state of one of the reactants (*e.g.* $a(t)$) with respect to a given parameter x_j (*e.g.* k_0)

$$\chi_j = \frac{\partial\,(a^\infty)}{\partial x_j}\,\left(\frac{a^\infty}{x_j}\right)^{-1}. \tag{22}$$

The magnification factor, χ_j, is defined in terms of *relative* changes, so that if x_j is changed into $x_j(1+\epsilon)$, a^∞ becomes $a^\infty(1+\chi\epsilon)$ (with ϵ being a small number). Here we assume that the parameters form a vector, $\vec{x}$, and x_j is its j-th component. Then

$$\chi_j^2 = \frac{V(a^\infty)}{V(x_j)}\,\left(\frac{<a^\infty>_{x_j}}{<x_j>}\right)^{-2}. \tag{23}$$

$<x_j>$ and $V(x_j)$ represent the mean and the variance of x_j, and $<a>_{x_j}$ and $V(a)$ represent the mean and the variance of a generated by an uncertainty in x_j. The analogue of the Lyapunov exponent can then be defined as [3]

$$\lambda_j = \ln(\chi_j)$$

for any of the parameters. High and positive values of λ suggest a strong dependence of the final state of the reaction on the particular parameter.

The above definitions can be generalized for $a(t)$ at any time t. This allows us to study the evolution of the uncertainty in time and to find parts

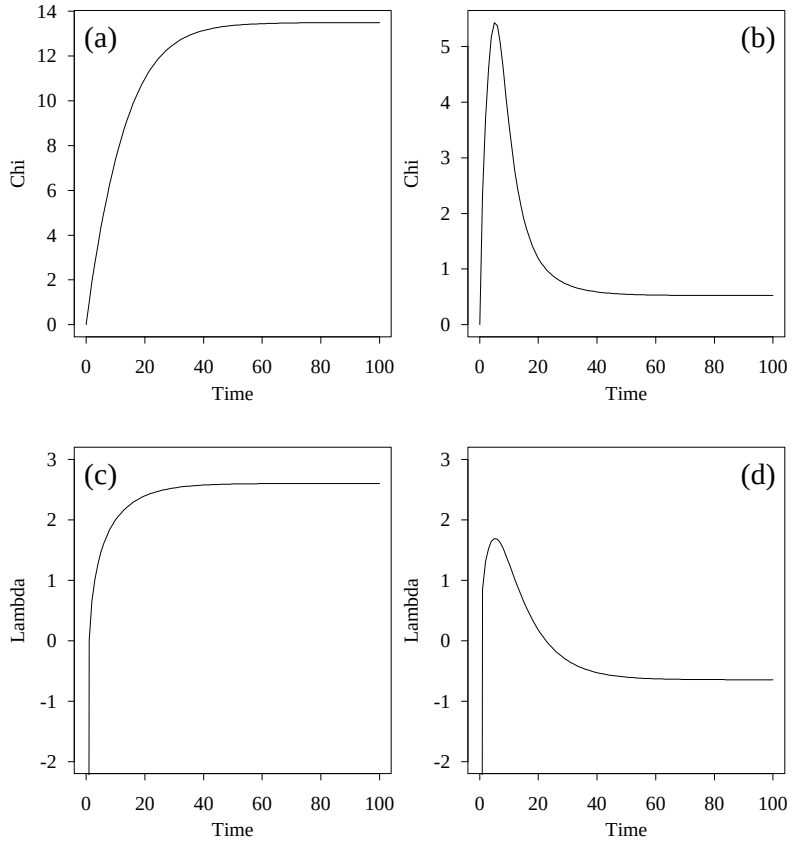


Fig. 4. Example 1: (a) and (c): Time evolution of the evolution of the magnification factor, χ_j, and (b) and (d) of the analogue of the Lyapunov exponent, λ. (a) and (c) show the results for the Example 1 (dynamics limited by the decaying rate) and (b) and (d) for the Example 5 (dynamics limited by the availability of one if the reagents). All parameters as in previous figures

of the dynamics which are mostly influenced by changes in the particular parameter. Although for a given distribution and a model it is possible to obtain an analytical formula for λ_j, we give numerical solutions for the Example 1, see figure 4(a)–(b), and compare with the results for the Example 5, figure 4(c)–(d). For Example 1, λ quickly reaches positive values, suggesting the particular sensitivity of the solution to changes in k_0, both at the equilibrium and in the transient. On the contrary, changes in k_0 affect strongly the transient dynamics of the Example 5, but the equilibrium seems to be rather unaffected ($\lambda < 0$).

4. Discussion

In the paper we have discussed the evolution of a dynamical variable describing the concentration of a substrate in one or more chemical reactions for which the probability of the reaction decays to zero with time. For such systems transients are reflected in the final state which varies as the parameters and/or initial conditions change. We analysed the evolution of the associated probability distribution of the concentration, with the stochastic element introduced by differences in one or more parameters. The variability is caused by imperfections in preparing chemical systems and environmental factors. Once the parameters are fixed, the evolution is considered to be deterministic. In fact, additional sources of uncertainty must be also considered. Amongst them the demographic variability caused by small numbers of reacting molecules or growing foci is most important. This effect depends on the size of the system and can be reduced by increasing the size of an individual replicate. In order to reduce the between-sample variability we need to consider dividing the system into more independent components. In many situations this leads to a trade-off effect, and an optimal size of the sample can be found.

The results presented here are important from the point of view of parameter estimation and fitting. Very often the exact parameters and sometimes even the exact mechanisms are not known, and the model must be identified from the data. The estimation procedure depends on the shape of the probability distribution, which in turn depends on unknown distributions of parameters and the model. However, some simple assumptions (like the Gaussian shape of the parameter k_0 considered here) can be made. The existence of analytical solutions and predictions about the evolution of the probability distribution can be crucial in determining a fitting procedure.

The work presented here can be extended in two directions. First, more complicated and more realistic models can be discussed. For some of them, analytical deterministic solutions can be found, and for some the evolution of the probability distribution can be obtained without solving the equations; see Example 4 and 5. Second, other sources of variability, and especially time-dependent additive and/or multiplicative noise, can be considered. Some previous work [9,10] suggests that analytical solutions for the transient behaviour can be obtained in such a case, and we plan to include this type of analysis in the second part of the paper.

Finally, a spatial aspect of the reactions can be included in the analysis. Many biochemical and biological processes occur in the presence of low or no mixing, whereas the above formalism assumes perfect mixing and/or homogenous initial state. If the mixing is low, complicated spatial patterns can result and affect the time evolution (Kleczkowski, unpubl.), both in the

deterministic and stochastic case. The spatial structure can be confined to a plane or extend into three dimensions.

I am very grateful to the organizers of the Marian Smoluchowski Symposium on Statistical Physics, Zakopane, 1997, for allowing me to take part in the conference. Discussions with C.A. Gilligan and K.M. Briggs (Cambridge) were also extremely useful.

Appendix A

Analytical solution of model 3

Consider the following set of equations

$$\frac{da}{dt} = kab\,, \tag{24}$$

$$\frac{db}{dt} = -kab - gb\,, \tag{25}$$

$$\frac{dz}{dt} = gb\,, \tag{26}$$

where we introduced a variable z representing an 'inactive' state of substrate B, so that $a + b + z = \text{const} = n$. Then,

$$\frac{da}{dz} = \frac{k}{g}\,a\,, \tag{27}$$

$$\ln\left(\frac{a}{a_0}\right) = \frac{k}{g}\,(z - z_0)\,, \tag{28}$$

$$b = n - z - a = n - z_0 - \frac{g}{k}\,\ln\left(\frac{a}{a_0}\right) - a\,, \tag{29}$$

where $z_0 = n - a_0 - b_0$. Using the above formula we can obtain a closed equation for a, cf. Eq. (16).

Special cases The integral

$$\int \frac{da}{a\left(b_0 - \frac{g}{k}\,\ln\left(\frac{a}{a_0}\right) - (a - a_0)\right)} \equiv \int \frac{da}{f(a)} \tag{30}$$

cannot be evaluated analytically for any value of a. Two special cases can be, however, distinguished. If $a = a_0(1 + \delta)$ with δ small, then

$$f(a) = a\left(b_0 - a_0\delta - \frac{g}{k}\,\ln(1 + \delta)\right)\,, \tag{31}$$

$$= a_0(1+\delta)\left(b_0 - a_0\delta - \frac{g}{k}\left(\delta - \delta^2 + O(\delta^3)\right)\right), \tag{32}$$

$$= a_0\left(b_0 + \left(b_0 - \frac{g}{k} - a_0\right)\delta - \left(a_0 - 2\,\frac{g}{k}\right)\delta^2 + O(\delta^3)\right). \tag{33}$$

Alternatively, if $a - a_0 \gg g/k\,\ln(a/a_0)$, *i.e.* for $a \gg a_0$, the logarithm in (30) can be neglected and

$$\int \frac{da}{a\left(b_0 - \frac{g}{k}\,\ln\left(\frac{a}{a_0}\right) - (a-a_0)\right)} \simeq \int \frac{da}{a\,(b_0 + a_0 - a)}$$

$$= \frac{1}{a_0 + b_0}\,\ln\left(\frac{a}{a_0 + b_0 - a}\right). \tag{34}$$

Equations (33) and (34) can now be evaluated analytically.

Appendix B

General properties of the error propagation

Let us assume that a is a solution of the equation

$$\frac{da}{dt} = k(t)f(a)\,. \tag{35}$$

In the following we shall consider a particular form of $k(t) = k_0\exp(-gt)$. Then, the solution can be obtained by separating the variables, and given in an implicit form

$$\mathcal{F}(a) \equiv \int_{a_0}^{a} \frac{da}{f(a)} = \int_0^t k(\tau)\,d\tau = \frac{k_0}{g}\,(1 - \exp(-gt)) \equiv \mathcal{G}(k_0)\,. \tag{36}$$

Then, the conversion factor can be obtained by an implicit differentiation of $F(a)$,

$$\frac{\partial a}{\partial k_0} = \left(\frac{\partial \mathcal{F}(a)}{\partial a}\right)^{-1}\left(\frac{d\mathcal{G}(k_0)}{dk_0}\right) = f(a)\,\frac{1-\exp(-gt)}{g}\,. \tag{37}$$

Note that there is no need of solving Eq. (35) explicitly or even evaluating the integral in Eq. (36).

In order to express the probability distribution of the parameter(s) in terms of $a(t)$ we need to change the variables in the form of the underlying probability distribution. From (36) we obtain

$$k_0 = \frac{g\mathcal{F}\,(a)}{1-\exp(-gt)}\,. \tag{38}$$

Combining Eq. (37) and Eq. (38) we obtain a general form

$$P(a) = \frac{g}{(1-\exp(-gt))\ f(a)}\ \rho\left(\frac{g\mathcal{F}(a)}{1-\exp(-gt)}\right)\ . \tag{39}$$

Similar expressions can be obtained if g or a_0 is changing. For the latter, if the full solution $a(t)$ is known, conditioned on the initial condition $a(t=0)=a_0$, we can express a_0 in terms of $a(t)$ by changing the independent variable t into $-t$.

For a particular choice of k_0 distributed according to a Gaussian distribution with a mean μ and a variance σ^2 we obtain

$$P(a) = \frac{g}{\sqrt{2\pi}\,\sigma\, f(a)\ (1-\exp(-gt))}\ \exp\left(-\frac{(\mathcal{F}(a)-\mu\,(1-\exp(-gt))/g)^2}{2\,(\sigma\,(1-\exp(-gt))/g)^2}\right)\ . \tag{40}$$

Note that if $g \to 0$ in Eq. (39), $k(t) \to \mu$ and $(1-\exp(-gt))/g \to t$, so that

$$P(a) = \frac{1}{tf(a)}\,\rho\left(\frac{\mathcal{F}(a)}{t}\right)\ , \tag{41}$$

and, for the Gaussian shape,

$$P(a) = \frac{1}{\sqrt{2\pi}\,\sigma\, f(a)\, t}\ \exp\left(-\frac{(\mathcal{F}(a)-\mu\, t)^2}{2\,(\sigma\, t)^2}\right) \tag{42}$$

This represents a function, centred at the deterministic solution determined by an implicit equation $\mathcal{F}(a) = \mu\, t$, with a width $\sigma\, f(a)\, t$. If the solution, $a(t)$, of the deterministic equation (35) settles on an equilibrium a^*, then $f(a) \to 0$, and usually $f(a) \sim \exp(-\mu t)$ so that $f(a)\, t \to 0$ for $t \to 0$. The probability distribution in (40) becomes a δ-function reflecting the deterministic character of the motion.

REFERENCES

[1] C.W. Gardiner, *Handbook of Stochastic Methods*, second edition Springer Verlag, Berlin, Heidelberg, New York 1985.

[2] I.R. Epstein, *Nature* **374**, 321 (1995).

[3] A. Kleczkowski, D. Bailey, C.A. Gilligan, *Proc. Roy. Soc.* **B263**(1371), 777 (1996).

[4] C.A. Gilligan, A. Kleczkowski, *Philos. Trans. R. Soc. London* **B352**, 591 (1997).

[5] A. Kleczkowski, C.A. Gilligan,D.J. Bailey, *Proc. Roy. Soc. London* **B264**, 979 (1997).

[6] I.S. Gradshteyn, I.M. Ryzhik, A. Jeffrey (Editor), *Table of Integrals, Series, and Products*. Academic Press, fifth edition, 1994. Translated from the Russian.

[7] Q.W. Yao, H. Tong, *J. Roy. Stat. Soc. Series B- Methodological* **56(4)**, 701 (1994).

[8] R.J. Deissler, J.D. Farmer, *Physica*, **D55**, 155 (1992).

[9] R. Zygadło, *Phys. Rev.* **E47**, 106 (1993).

[10] R. Zygadło, *Phys. Rev.*, **E47**, 4067 (1993).