

STOCHASTIC MODELS FOR CHEMICAL REACTIONS:
I. THEORY OF THE UNIMOLECULAR
REACTION PROCESS

ANTHONY F. BARTHOLOMAY*
HARVARD UNIVERSITY
BOSTON, MASSACHUSETTS

A stochastic model for the basic unimolecular chemical reaction $A \xrightarrow{\mu} B$ is derived. This model provides a mathematical basis, altogether missing in the current kinetic theory, for the analysis of inherent random fluctuations about the strict concentration-time course prescribed by the existing deterministic theory. Limits on the extent of the predicted inherent variability are obtained and compared with those usually expected purely on the basis of random experimental errors of extraneous origin (not associated with the mechanism of reaction). The results support the extrapolation to chemical systems of a principle of statistical inaccuracy for physical systems which has been called by E. Schroedinger "the \sqrt{n} Law of Physics."

1. *Introduction.* Mathematical models currently used in chemical reaction kinetics are "deterministic": once the initial values of concentrations of the reactant species are indicated, ordinary mathematical functions assign precise concentration values to be expected at subsequent times. Large scale fluctuations about these predicted values are ascribed usually to extraneous experimental errors. The statistical method of least squares fitting is employed as a smoothing device in such cases in order to obtain "rate constants."

Little attention has been paid to the study of fluctuations arising "inherently" in the reaction mechanism itself and not associated with random experimental errors. In the case of extremely slow reactions such as radioactive decompositions, which can be followed with great precision, the randomness of molecular decompositions

*Department of Biostatistics, The Harvard School of Public Health; and the Biophysics Research Laboratory of the Department of Medicine, Harvard Medical School and the Peter Bent Brigham Hospital.

cannot be ignored (See Rutherford, Chadwick, Ellis, 1931). L. Pauling (1949, p. 335), e.g., in referring to unimolecular reactions in general, conjectures that the numbers of molecules predicted for various times are not precise; that molecules decompose at random at an "average rate" given by the deterministic equation but that a "small" statistical fluctuation from this rate is to be expected.

How does one go about drawing a line between different orders of fluctuations to be expected? The current theory for such reactions cannot be used since it predicts no fluctuations whatever. The present paper suggests a mathematical rationale for answering this and related questions on a purely theoretical basis by constructing a stochastic or probabilistic model which may be applied to any unimolecular reaction. In constructing this model a correspondence principle is invoked which guarantees that the mean value of the resulting random function coincides with the ordinary time function currently used. The resulting theory may therefore be thought of as an extension of the present theories which agree with the current deterministic model "on the average,"—a situation which is reminiscent of the relationship of modern quantum mechanics to classical mechanics.

2. *The mathematical basis.* The basic notion of this paper is that of a Markoff chain with finitely many states, continuous time parameter, and stationary transition probabilities. This process is simpler in many ways than other processes because it makes the assumptions (1) that the future state of a system depends only on the present state, not on the entire past history; (2) that the probability of a transition from one arbitrarily defined "state" of a system to another depends on the length of time interval alone, and not on the particular instant at which the present state was reached; and (3) that transitions can take place at arbitrary times on a continuous time scale, so that the evolutionary progress of a system is non-uniform in time. Because of these characteristics, the stationary Markoff process seems to be well suited not only to chemical reaction kinetics but also to a great number of physical and biological kinetic systems (see, e.g., Furry, 1937; Feller, 1939, 1950; Arley, 1943; Kendall, 1951; and Bartholomay, 1957a,b, 1958a,b). The reader is referred to J. L. Doob (1953) for a complete mathematical analysis of such processes in general and to an earlier paper on biological growth processes published in this journal (Bartholomay, 1958b) for the complete mathematical treatment of equations similar to those which

arose in the present investigation. K. Singer (1953) has used processes with somewhat similar structure to explain macroscopic fluctuations arising in branching nucleation processes.

The starting point in the derivation of the "unimolecular reaction process" is the conception of concentration (expressed in integral units of number of molecules per constant volume of reaction mixture) as a time-varying random variable, the states or "values" at different times being influenced by probability effects. The central problem in the mathematical description of this kind of stochastic model is the calculation of the matrix of transition probabilities $\{p_{ik}(t)\}$, where $p_{ik}(t)$ represents the probability of going from a concentration value of i to the value of $k < i$ in time t . Once these are known, the complete basis for all probability statements about the process is at hand. The concentration values $0, 1, 2, \dots, n_0$, where n_0 is the initial concentration of A in the reaction $A \xrightarrow{\mu} B$, define the "Markoff states" S_0, S_1, \dots, S_{n_0} , in classical terminology.

It has been shown (See Doob, *loc. cit.*, p. 241) that the $p_{ik}(t)$'s are the solutions of the following "forward" system of differential equations:

$$p'_{ik}(t) = q_{kk} p_{ik}(t) + \sum_{j \neq k}^{n_0} q_{jk} p_{ij}(t), \quad (1)$$

where $i, k = 0, 1, 2, \dots, n_0$, subject to the initial conditions

$$p_{ij}(0) = \delta_{ij} = \begin{cases} 1 & (i = j), \\ 0 & (i \neq j). \end{cases} \quad (2)$$

In this system of equations $p'_{ik}(t)$ is the time derivative of $p_{ik}(t)$; the constant coefficients, the q 's, comprise a "Q-matrix" defined as follows:

$$q_{ij} = p'_{ij}(0) > 0 \quad (j \neq i), \quad (3)$$

$$q_{ii} = - \sum_{j \neq i} q_{ij} < 0. \quad (4)$$

In adapting these equations and relationships to the unimolecular reaction, the q 's have first been estimated from theoretical considerations stemming from the deterministic equations, using a method introduced by the author (Bartholomay, 1958b).

3. *Derivation of the unimolecular stochastic process.* The deterministic mathematical model for the unimolecular reaction $A \xrightarrow{\mu} B$ is completely given by the ordinary differential equation

$$\frac{dn}{dt} = -\mu n \quad (5)$$

in which the concentration n of A at any time t is treated as an ordinary continuous function of time and μ is the rate constant (assuming that reaction conditions are held constant). If $n_0 = n(0)$ is the initial concentration of A , then (5) may be integrated to give

$$n = n_0 e^{-\mu t}. \quad (6)$$

According to these assumptions then, the approximate number $|\Delta n|$ of individual molecular transformations taking place in the infinitesimal time from t to $(t + \Delta t)$ is given by

$$|\Delta n| = \mu n \Delta t + o(\Delta t), \quad (7)$$

in which $o(\Delta t)$ is a higher-order infinitesimal, vanishing in the limit as $\Delta t \rightarrow 0$. Thus, the ratio

$$\frac{|\Delta n|}{n} = \mu \Delta t + o(\Delta t) \quad (8)$$

may be interpreted as the probability of a single chemical transformation from n possible reactant molecules in the time interval $(t, t + \Delta t)$.

Assuming then that the transformations of individual molecules are randomly-occurring independent events, let

$$p = \frac{|\Delta n|}{n} = \mu \Delta t + o(\Delta t) \quad (9)$$

be taken as the probability of "success," and $q = 1 - p$ the probability of a failure in a conceptual sequence of Bernoulli trials repeated n times (once for each of the n molecules present at time t) during the interval $(t, t + \Delta t)$. These "trials" are independent of the location of t on the time scale. According to the binomial distribution which governs such events, the probability of k "successes" in n trials is given by

$$b(n, k; p) = \binom{n}{k} [\mu \Delta t + o(\Delta t)]^k [1 - \{\mu \Delta t + o(\Delta t)\}]^{n-k}, \quad (10)$$

where $\binom{n}{k}$ is the symbol for the combinations of n objects chosen k at a time. Expansion of the right-hand side of (10) gives

$$\binom{n}{k} \left\{ \mu^k \Delta t^k + k \mu^{k-1} \Delta t^{k-1} o(\Delta t) + \frac{k(k-1)}{2} \mu^{k-2} \Delta t^{k-2} [o(\Delta t)]^2 + \dots + [o(\Delta t)]^k \right\} [1 - f(\Delta t)], \quad (10')$$

where $f(\Delta t)$ is a polynomial in Δt with constant coefficients, so that

$$\left. \begin{array}{ll} \text{when } k = 1 & b(n, 1; p) = \mu n \Delta t + o(\Delta t), \\ \text{when } k \geq 2 & b(n, k; p) = 0 + o(\Delta t), \end{array} \right\} \quad (11)$$

where $o(\Delta t)$ represents a polynomial expression in Δt , each term of which contains a power of Δt greater than 1. Thus, whereas the probability of one and only one molecular transformation in this time is directly proportional to the concentration n and to the length of infinitesimal interval Δt , (except for terms of higher order in Δt), the probability of more than one such transformation in this time contains only higher-order powers of the infinitesimal Δt .

The preceding considerations may be treated as the inductive basis for the following set of axioms which describe the unimolecular reaction process in mathematical, probabilistic terms:

A1. If exactly $(n - n_0)$ decompositions have already occurred in the time interval $(0, t)$, in the subsequent interval $(t, t + \Delta t)$ the probability of a single decomposition of n molecules present at time t equals $[\mu n \Delta t + o(\Delta t)]$. The probability of more than one such collision in this time equals $o(\Delta t)$.

A2. The probability parameter μ has the dimensions and properties of the ordinary rate constant, the reaction proceeding under controlled constant conditions of chemical parameter (e.g., temperature, pressure, volume), rendering μ constant.

A3. Reactant species A and product species B coexist without significant interactions of any kind.

A4. The reverse reaction $B \rightarrow A$ has zero probability.

This set of axioms implies that the chemical process described may be represented mathematically as a Markoff chain of the kind discussed in section 2 and leads to the specification of the Q-matrix of constants $\{q_{ij}\}$ in the following way. Assuming, as usual, that

the functions $\{p_{ij}(t)\}$ satisfy the required differentiability properties, each function may be expanded about 0 time into a Taylor-Maclaurin series:

$$p_{ij}(\Delta t) = p_{ij}(0) + p'_{ij}(0) \Delta t + o(\Delta t) \quad (j \neq i) \quad (12)$$

The initial conditions (2) together with this formula imply that

$$p_{ij}(\Delta t) = q_{ij} \cdot \Delta t + o(\Delta t), \quad (13)$$

i.e., neglecting second-order and higher-order terms in Δt , the probability of a transition from i to j in the time interval of length Δt equals $q_{ij} \cdot \Delta t$. But from axioms A1 and A4, $p_{ij}(\Delta t)$ has the values:

$$p_{ij}(\Delta t) = \begin{cases} \mu i \Delta t + o(\Delta t) & (j = i - 1), \\ o(\Delta t) & (j < i - 1), \\ 0 & (j > i). \end{cases} \quad (14)$$

Comparison of (14) with (13) and the use of formula (4) show then that the Q-matrix elements are

$$q_{ij} = \begin{cases} 0 & (j < i - 1, j > i) \quad (i = 0, 1, 2, \dots, n_0), \\ \mu i & (j = i - 1), \\ -\mu i & (j = i). \end{cases} \quad (15)$$

Appropriate elements of the Q-matrix substituted into (1) yield the fundamental set of differential equations of the process:

$$p'_{ik}(t) = -\mu k p_{ik}(t) + \mu(k+1) p_{i,k+1}(t) \quad (16)$$

for given i , where $0 < i \leq n_0$, and for all k 's such that $k < i$.

This system of ordinary differential equations is transformed into a single partial differential equation by use of the "probability generating function" (See Bartholomay, 1958b):

$$\text{p.g.f. } \{p_{ik}(t)\} = \phi_i(s, t) = \sum_{k=0}^{n_0} s^k p_{ik}(t) \quad (17)$$

in which s is an arbitrary variable, with no physical significance, introduced purely for mathematical convenience. Differentiating $\phi_i(s, t)$ partially with respect to s and t separately and making appropriate substitutions from (16) gives the fundamental partial differential equation of the system:

$$\frac{\partial \phi_i}{\partial t} = \mu(1-s) \frac{\partial \phi_i}{\partial s}, \quad (18)$$

with boundary conditions

$$\left. \begin{aligned} \phi_i(s, 0) &= s^i, \\ \phi_i(1, t) &= 1. \end{aligned} \right\} \quad (19)$$

The boundary conditions follow, respectively, from initial conditions (3) and the constraint that for each i and all k such that $0 \leq k \leq n_0$, the $p_{ik}(t)$ define a probability distribution; i.e.,

$$\sum_{k=0}^{n_0} p_{ik}(t) = 1 \text{ for all } t \text{ where each } p_{ik}(t) \geq 0.$$

Equation (18) is a homogeneous, first order, partial differential equation to which Lagrange's method of auxiliary differential equations* may be applied, yielding the solution:

$$\phi_i(s, t) = [(-1 + e^{\mu t}) + s]^i e^{-\mu i t} \quad (i = 1, 2, \dots, n_0). \quad (20)$$

Expanding the right-hand side, collecting powers of s and comparing the result with the definition (17) of $\phi_i(s, t)$ gives finally

$$p_{ik}(t) = \binom{i}{k} e^{-\mu i t} (-1 + e^{\mu t})^{i-k} \quad (0 \leq k < i \leq n_0). \quad (21)$$

The mean value and variance functions of k , relative to the $p_{ij}(t)$'s defining the distribution of k , are easily calculated to be

$$E_{ik}(t) = i e^{-\mu t}, \quad (22)$$

$$\sigma_{ik}^2(t) = i(e^{-\mu t} - e^{-2\mu t}). \quad (23)$$

This completes the mathematical derivation of the fundamental properties of the Markoff model.

4. *Comparison of the stochastic theory with the deterministic theory.* The transition probability functions given by (21) make it possible, given a probable or an exact value of the concentration of A at an arbitrary time, to determine the probability that it will have a particular value t seconds later. For example, if it is known only that a given concentration, say, $n = N_0$ (where $0 \leq N_0 \leq n_0$)

has probability $p(N_0)$ where $\sum_{N_0=0}^{n_0} p(N_0) = 1$, then the probability t seconds later of finding a concentration value $n < N_0$ is given by

$$p_n(t) = p(N_0) \cdot \binom{N_0}{n} e^{-\mu N_0 t} (-1 + e^{\mu t})^{N_0-n}. \quad (24)$$

*See Bartholomay, 1958b, equations (25)–(32) and (60)–(68).

On the other hand, if N_0 is determined as the exact value, then $p(N_0) = 1$ and $p(n) = 0$ ($n \neq N_0$) so that in the preceding formula the coefficient $p(N_0)$ no longer appears. In particular, in the usual situation in which the precise value of n is determinable only initially as n_0 , then the probability subsequently of any value $n < n_0$ is

$$p_n(t) = \binom{n_0}{n} e^{-\mu n_0 t} (-1 + e^{\mu t})^{n_0 - n}, \quad (25)$$

in which case the mean value and standard deviation functions for n become, respectively

$$E(n, t) = n_0 e^{-\mu t}, \quad (26)$$

$$\sigma(n, t) = \sqrt{n_0 (e^{-\mu t} - e^{-2\mu t})}. \quad (27)$$

But $n_0 e^{-\mu t}$, with $\mu = k$, is the precise value of n predicated by the deterministic theory. The stochastic model is therefore "consistent in the mean" with that theory. The progress of a unimolecular reaction according to the non-random model is thus seen to furnish predictions which hold only "on the average" in a very definite sense, according to the stochastic theory (see the reference to Pauling's book quoted earlier) and with a measurable degree of dispersion given by (27).

For comparing the two theories further, it will be convenient to rewrite (25) in the form

$$p_n(t) = \binom{n_0}{n} (1 - e^{-\mu t})^{n_0 - n} (e^{-\mu t})^n; \quad (28)$$

the special case when $n = n_0$ becomes

$$p_{n_0}(t) = e^{-\mu n_0 t}, \quad (29)$$

this latter showing how the probability of remaining at n_0 decays exponentially with time. If in (28) $e^{-\mu t}$ is identified with the parameter p , and $(1 - e^{-\mu t})$ with $q = 1 - p$, then for fixed t , $p_n(t)$ can be interpreted as the binomial probability of n successes in n_0 conceptual Bernoulli trials, where p stands for the probability of "success" and q for that of "failure."

Particularly in the cases of large values for n_0 the binomial form (28) of the transition probability functions can be well approximated by the Gaussian probability density,

$$p(x, t) = (2\pi)^{-\frac{1}{2}} \sigma^{-1}(n, t) \exp \left\{ -\frac{1}{2} \sigma^{-2}(n, t) [x - E(n, t)]^2 \right\}, \quad (30)$$

in which the continuous variate x replaces the discrete valued n . Relative to this approximation, a region defined by the boundaries $[E(n,t) \pm 2\sigma(n,t)]$ may be constructed in which the points (n,t) obtained in repeated experimental runs would be most likely to fall (with roughly 95% probability). A typical sample curve, corresponding to a single run, is predicted to be a non-increasing stepwise constant curve fluctuating randomly about $E(n,t)$ and lying mostly inside the region. The regression curve of a large number of repeated sample curves should lie close to $E(n,t)$ which, because of the "mean consistency" property of the stochastic process, is also the unique time course expected always by the deterministic theory (in the absence of error).

The differences in approach between the deterministic and the stochastic theories are summarized in Figure 1 which shows geometrically the variation of the reaction probability "force" p as a function of n and t . The resulting 3-dimensional surface will be called the "stochastic reaction surface." Its chief characteristics may be seen in terms of families of planar profiles in the planes $t = \text{constant}$ [using the Gaussian approximation given by (30)] and $n = \text{constant}$.

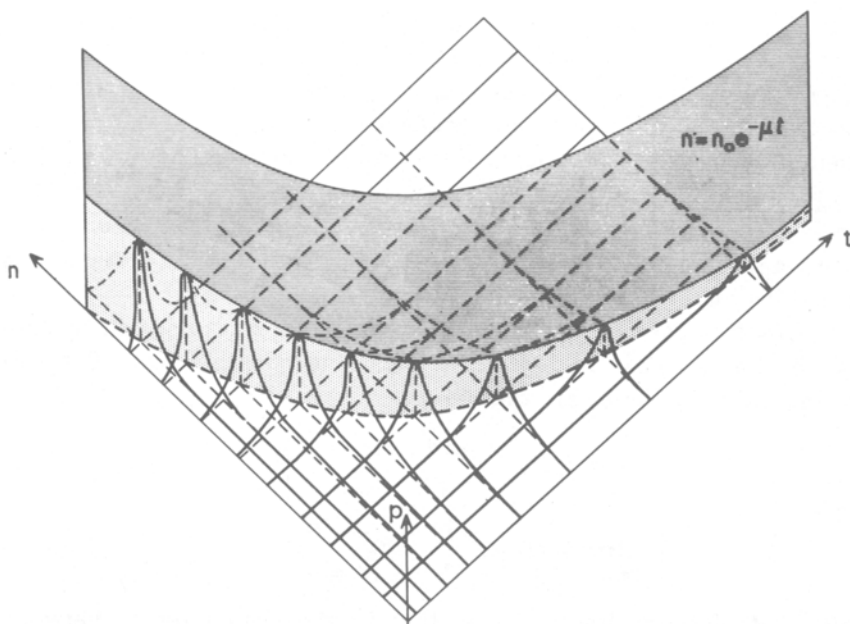


FIGURE 1. Stochastic reaction surface for the unimolecular reaction process.

The profiles in planes $t = K$ (constant) have been sketched from the continuous distribution in x given by equation (30) as a close approximation to equation (25). Since the maxima of this family of surface curves occur at the mean values of the approximating Gaussian probability densities $p_x(K)$, they correspond to points $(n_0 e^{-K\mu}, K)$ (where $0 \leq K < \infty$) in the (n, t) plane which lie on the deterministic curve $n = n_0 e^{-\mu t}$ shown in that plane.

In the planes $n = c$ (constant), using the actual probability function (25), the profiles have maxima located at points on this same curve; i.e., $t = -\frac{1}{\mu} \log_e \frac{c}{n_0}$ for $n = c$ (where $0 \leq c \leq n_0$). Consequently, while the surface has no relative maxima, there is a sharp crest, defined by the maxima of these two families of orthogonal profiles, the trace of this crest in the (n, t) plane being the ordinary deterministic concentration curve.

Thus, the contraction of the total unit area under each profile in the family of planes $t = K$ into a height of 1 probability unit above the intersection of its projection on the (n, t) plane with the curve (6) represents a mathematical transformation from the stochastic onto the deterministic theory.

5. *Some additional properties of the unimolecular stochastic process and the stochastic reaction surface.* In deriving additional properties of this process the transition functions $p_n(t)$ and the related mean value $E(n, t)$ and variance functions $\sigma^2(n, t)$ will be used in order to further emphasize the extension of the deterministic theory which has been obtained; i.e., n_0 will be taken as initial state i as in the latter theory. The results of this section, however, apply equally well to any element $p_{ij}(t)$ of the transition probability matrix.

a. *Limiting forms.* For fixed n the function $p_n(t)$ defined by (28) may be treated as a regular, analytical function of time. The following ordinary limits are obtained

$$\lim_{t \rightarrow 0} p_n(t) = \begin{cases} 1 & (n = n_0), \\ 0 & (n \neq n_0), \end{cases} \quad (31)$$

$$\lim_{t \rightarrow \infty} p_n(t) = \begin{cases} 0 & (n \neq 0), \\ 1 & (n = 0). \end{cases} \quad (32)$$

The first of these limits means that as the time interval between transitions grows smaller and smaller, the probability of transition

becomes negligibly small. The second of these indicates that when the allowable time of transition grows infinitely large the probability of any specific transition grows negligibly small also, except when $n = 0$ in which case there is certainty (in the sense of probability one). Both values then in (32) point to the ultimate extinction of the reactant species A as t grows infinitely large. These limiting values which have been checked in this particular model are quantities which would be expected a priori from purely mathematical considerations of the stationary Markoff Process.* All of the concentration "states" except the zero state are "transient states"; in such cases the probability of a particular transition is expected to approach 0 exponentially fast.

b. *Measure of inherent variability.* The expression for the standard deviation function $\sigma(n, t)$ [see formula (27)] is a natural one to suggest as a measure of the absolute amount of inherent variability of dispersion; particularly in view of the connection of the distribution function $p_n(t)$ with the normal distribution. The following values of this function are of particular interest:

$$(a) \lim_{t \rightarrow \infty} \sigma(n, t) = \lim_{t \rightarrow 0} \sigma(n, t) = 0, \quad (33)$$

$$(b) \sigma_{\max}(n, t) = \frac{1}{2} \sqrt{n_0}, \quad (34)$$

the maximal value occurring at time $t = \frac{1}{\mu} \log_e 2$ corresponding to the "half-time" of the reaction. Thus the inherent absolute variability grows steadily from 0 initially to a maximal value when n_0 has been reduced to $\frac{n_0}{2}$, receding gradually to zero as the reaction approaches completion. According to (34) then, a spread of experimental values about the mean or deterministic value can reach $\sqrt{n_0}$. This result appears to be consistent with what E. Schroedinger (1945) calls the " \sqrt{n} Rule" (see discussion, section 6).

The coefficient of variation $CV(t)$ is another statistical concept that can be applied to measure the variability, yielding the variation relative to the mean or expected value. Applied to the uni-molecular process it becomes

$$CV(t) = \sigma(t)/E(n, t) = \left[\frac{e^{\mu t} - 1}{n_0} \right]^{\frac{1}{2}}, \quad (35)$$

*See Doob, 1953, p. 236; and Feller, 1950, p. 324.

so that

$$\lim_{t \rightarrow 0} CV(t) = 0, \quad (36)$$

$$\lim_{t \rightarrow \infty} CV(t) = \infty, \quad (37)$$

the latter limit showing that while the absolute variability is declining this declines less rapidly than the mean value toward the end of reaction run. The relation between these two measures of dispersion is shown in Figure 2.

c. *Apparent completion time.* The "apparent" completions of unimolecular reactions which are generally observed in such cases can be interpreted as follows in terms of the stochastic theory. If $\varepsilon > 0$ is an arbitrarily small positive number, then the time T_ε required to obtain completion with probability equal to $(1 - \varepsilon)$ is determined as the solution of the equation

$$p_0(t) = (1 - e^{-\mu t})^{n_0} = 1 - \varepsilon; \quad (38)$$

$$\text{viz.,} \quad T_\varepsilon = -\frac{1}{\mu} \log_e [1 - (1 - \varepsilon)^{\frac{1}{n_0}}]. \quad (39)$$

in which $(1 - \varepsilon)^{\frac{1}{n_0}}$ is the principal n_0 th root of the probability number $(1 - \varepsilon)$.

6. *Discussion.* The principal purpose of the present paper has been to indicate a manner in which the effect of inherent random fluctuations in a chemical reaction can be treated theoretically.

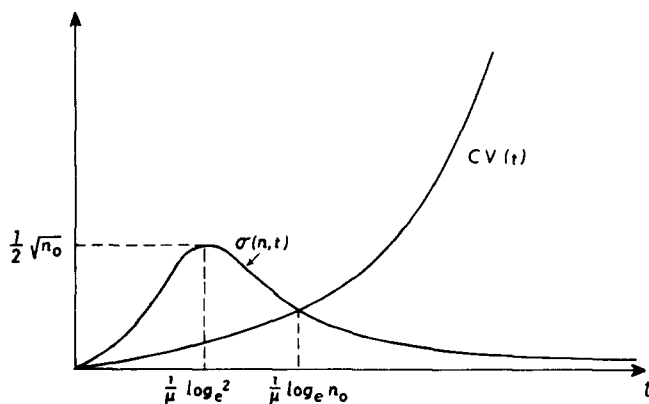


FIGURE 2. Standard deviation and coefficient of variation curves of the unimolecular reaction process.

The solution lies in the construction of a stochastic process showing how the probabilities of concentration values change with time. This contrasts with the deterministic theory which omits probability and gives the rule for the change in concentration as an ordinary non-random function of time.

The two approaches are not so far apart as might be imagined. The close relationship results from that property imposed on the stochastic model which makes the basic deterministic equation (6) the stochastic mean value. Random fluctuations about the mean value of the proposed stochastic process then amount to fluctuations about the values predicted by the current theory.

The magnitude of such deviations may be measured statistically in terms of the standard deviation or the coefficient of variation. The properties of these measures, as indicated in section 5b, show that often the magnitudes of the inherent variations are microscopic compared with random divergencies ascribable to extraneous factors such as experimental errors. They become macroscopic, however, in extremely dilute reactions and in certain epochs of the reaction time. The radioactive decomposition process (see Rutherford, Chadwick, Ellis, *loc. cit.*) is an example of a process in which the orders of magnitude of the predicted inherent fluctuations called for in this theory are realized under conditions controllable to the extent of almost completely eliminating experimental error. Another interesting feature is the agreement of the measure of dispersion (given by two standard deviations) with the order of divergency to be expected between predicted and (error-free) observed values in physical laws in general, as pointed out by E. Schroedinger *loc. cit.* He states that there is a " \sqrt{n} Rule of Physics": the laws of Physics may be inaccurate within a probable relative error of the order of $\frac{1}{\sqrt{n}}$, n being the number of molecules cooperating to bring about the law. In the present case the coefficient of variation as a measure of "relative" inherent, random dispersion would be $\frac{1}{\sqrt{n_0}}$ at maximal deviation, although in the later stages of the reaction a dispersion factor of many multiples of this quantity are possible.

It should be emphasized that other stochastic models are possible. There is, in general, no one-to-one correspondence between the class of all stochastic processes and a particular natural process. The one that has been used here has the advantage of great

simplicity, which derives from the stationary Markoff characteristic. In such processes the transition probabilities contain the answer to all probability questions which could be raised concerning the process. In addition to this advantage there are the consistencies pointed out above with what has already been assumed and tested in dealing with such reactions. The connection is close enough in fact to support the statement that the goodness of fit of this model to kinetic data is a natural corollary to the goodness of fit of the deterministic theory currently employed in the analysis of data.

The "stochastic reaction surface," which has been defined for such reactions, summarizes in geometric fashion the nature of the extension in kinetic theory introduced here. In this geometric context it may be seen to consist of the addition of a third dimension; viz., probability, to the concentration and time variables. The stochastic surface will in most instances be fairly flat, rising to a sharp crest along the intersecting negative exponential cylinder which corresponds to the non-random theory. Thus whereas in using the deterministic theory a unique value of concentration n is taken from a curve which sets up a correspondence between time and concentration, according to the stochastic theory, at a given time a whole curve of possible values (a section of the surface by the plane $t = \text{constant}$) exists. The height of the curve corresponds to the probability of occurrence of any concentration at that time. The microscopic character of dispersions about expected values in reactions with high concentrations n_0 shows that the existence of randomness does not correspond to complete lack of determinism; rather that the progress of the reaction is influenced by a "flow of probabilities" with time in a way which is pictured in this surface. The dominating "reaction force" of the deterministic theory is preserved as a strong trend, the crest of the surface.

Two additional advantages of the mathematical method given here are: (1) The generalizability of the method: it has already been applied to other types of reactions (Bartholomay, 1957a,b, 1958a), and (2) calculation of the rate constant: a new, simple method for calculating the rate constant k which results from its identification with the basic stochastic parameter μ will be presented in a subsequent paper.

The need for a kinetic analysis of this type will be more acutely felt when it becomes possible to analyze in detail some of the bio-

chemical reactions taking place in the individual cells of organisms. For example, the volume of a single gene has been calculated to be comparable to the volume of a cube the side of which is 300 angstroms. As E. Schroedinger (*loc. cit.*) points out, this can span only about 100 or 150 atomic distances in a liquid or solid—implying that a gene contains no more than a few million atoms which are not all similar nor similarly related. This number according to the stochastic model given here is much too small to expect anything like strict determinateness even if all molecules were involved in the same reaction (as in a unimolecular decomposition).

There is a universality to these considerations. For example, in discussing diffusion-like phenomena E. Schroedinger (*loc. cit.*) writes down the usual partial differential equation $\frac{\partial p}{\partial t} = D \nabla^2 p$, but hastens to point out that while this is a precise mathematical formulation its physical exactitude must be challenged in every physical application; it will be a very close approximation to reality only when an enormous number of molecules are involved in the diffusion process. As this number grows small, the random deviation increases, becoming highly significant when that number is very small—a characterization which applies also to the unimolecular reaction process established in the present paper.

The author is indebted to Dr. Hugo Muench, Chairman of the Department of Biostatistics, and to Dr. Bert L. Vallee, Director of the Biophysics Research Laboratory, for their interest and encouragement throughout the progress of this work. He also owes debts of gratitude to Dr. Frederic L. Hoch of the Biophysics Research Laboratory for his helpful suggestions concerning the organization of the present paper and to Dr. Ralph E. Thiers and Dr. Warren E. C. Wacker also of this laboratory for their comments.

This work was supported partially by the Rockefeller Foundation and by the Howard Hughes Medical Institute.

LITERATURE

- Arley, N. 1948. *On the Theory of Stochastic Processes and Their Application to the Theory of Cosmic Radiation*. New York: John Wiley and Sons, Inc.
- Bartholomay, A. F. 1957a. "A Stochastic Approach to Chemical Reaction Kinetics." (Abstract) *Proc. National Biophysics Conference*, Columbus, Ohio, p. 8.

- . 1957b. A Stochastic Approach to Chemical Reaction Kinetics. Unpublished S.D. Thesis. School of Public Health, Harvard University.
- . 1958a. "A Stochastic Model for the Reversible Reaction." (Abstract) *Proc. National Biophysics Soc.*, Cambridge, Mass. p. 48.
- . 1958b. "On the Linear Birth and Death Processes of Biology as Markoff Chains." *Bull. Math. Biophysics*, **20**, 97-118.
- Doob, J. L. 1953. *Stochastic Processes*. New York: John Wiley and Sons, Inc.
- Feller, W. 1939. "Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in Wahrscheinlichkeitstheoretischer Behandlung." *Acta Biotheoretica. Series A*, **5**, 11-39.
- Feller, W. 1950. *Probability Theory and Its Applications*. New York: John Wiley and Sons, Inc.
- Furry, W. H. 1937. "On Fluctuation Phenomena in the Passage of High Energy Electrons Thru Lead." *Phys. Rev.* **52**, 569-81.
- Kendall, D. G. 1948. "On the Generalized Birth and Death Process." *Ann. Math. Stat.* **19**, 1-16.
- Pauling, L. 1949. *General Chemistry*. San Francisco: W. H. Freeman and Co.
- Rutherford, E., Chadwick, J., Ellis, J. D. 1931. *Radiations from Radioactive Substances*. Cambridge: University Press.
- Schroedinger, E. 1945. *What is Life? The Physical Aspect of the Living Cell*. New York: The Macmillan Co.
- Singer, K. 1953. "Applications of the Theory of Stochastic Processes to the Study of Irreproducible Chemical Reactions and Nucleation Processes." *J. Roy. Stat. Soc., Ser. B*, **15**, 92-106.

RECEIVED 3-12-58