STEVEN

CHAPTER 8

Analysis of Experimental Observables and Oscillations in Single-Molecule Kinetics

Marcel O. Vlad*,† and John Ross*

*Department of Chemistry, Stanford University
Stanford CA 94305-5080 USA

†Institute of Mathematical Statistics and Applied Mathematics
Casa Academiei Romane, Calea 13 Septembrie 13, Bucharest
050711 Romania

1. Introduction

This chapter deals with the theoretical analysis of different types of oscillations which may occur in single-molecule kinetics and their influence on different experimental observables. The main difference between a single-molecule system and a macroscopic kinetic system is that for a macroscopic system the details of intramolecular dynamics are lost due to the overlapping of a large number of signals produced by the different molecules present in the system, whereas for a single-molecule system these microscopic signals have a direct influence on the experimental observables. In particular, oscillations of observables due to intramolecular dynamics, which in macroscopic experiments are smoothed out and do not show up in the observed data, can be observed directly in single-molecule experiments.

In this chapter we do not intend to develop a general theory of oscillations in single-molecule kinetics; instead we consider a few models, which can be investigated in detail. We intend to show the reader how to build different stochastic models for single-molecule kinetics, with special reference to oscillations. Experiments in single-molecule kinetics[1–7] consist in studying the chemical changes of one large molecule, such as a protein or an enzyme, immobilized on a support; the process may involve either a large molecule alone or a large molecule interacting with smaller molecules; intramolecular and molecular fluctuations are large and thus the deterministic mass action laws of chemical kinetics do not hold and are replaced by probabilistic laws. Although it would be desirable to develop a microscopic description based on nonequilibrium statistical mechanics, this is an extremely difficult task and thus the approaches in single-molecule kinetics are based on a stochastic, mesoscopic description involving two different types of stochastic processes. In the following we use variations of the basic model: (a) A single-molecule can exist in different chemical states $u = 1, 2, \ldots$ and the random transitions from one chemical state to another can be described by a local, Markovian master equation with time-dependent transition rates, $k_{uu'} = k_{uu'}(t)$. (b) Due to the conformational and other (energy) fluctuations in the single molecule, the rate coefficients $k_{uu'}(t)$ themselves are random functions of time. Some approaches consider directly the fluctuations of the rate coefficients, whereas other approaches assume that the stochastic properties of the rate coefficients can be represented in terms of a set of control parameters,[3,7–9] such as total energy of the molecule or the energy corresponding to a given degree of freedom; in this chapter we use both approaches, (a) and (b). In both cases we can write a Markovian master equation with random rate coefficients for the probability $P_u(t)$ that the molecule is in the chemical state u at time t

$$\frac{\partial}{\partial t} P_u(t) = \sum_{u' \neq u} P_{u'}(t) k_{u'u}(t) - P_u(t) \sum_{u'' \neq u} k_{uu''}(t), \tag{1}$$

where $k_{uu'}(t)$ is the rate of transition (rate coefficient) from the state u to the state u' at time t. The rate coefficients are random functions of time; thus, we need additional stochastic equations for describing the fluctuations of the rate coefficients $k_{uu'}(t)$. The simplest approximation is to neglect the fluctuations of the rate coefficients altogether and describe the kinetics of the process in terms of a master equation with constant coefficients; such an approach is similar to traditional chemical kinetics and is used as a first approximation. More sophisticated approaches make certain assumptions regarding the fluctuations of the rate coefficients based on theoretical models or experimental observations. In this chapter, we assume that the fluctuations of the rate coefficients can be described in terms of known characteristic functionals.

Among the different states $u = 1, 2, \ldots$ of the molecule studied, some are fluorescent and some are not. The molecule undergoes a random walk among these states, resulting in random variations of the fluorescent signal. The direct, raw experimental observable in a single-molecule experiment is the fluorescent signal $I(t)$ as a function of time, collected from the molecule studied; since the experiments are usually carried out in a time-independent regime, the time series describing the evolution of $I(t)$ is stationary. The most commonly used approach of data analysis is based on the computation of the correlation functions[10] of the fluorescence signal at times $t_1, \ldots, t_m$

$$C_m = \langle (I(t_1) - \langle I(t_1) \rangle) \cdots (I(t_m) - \langle I(t_m) \rangle) \rangle. \qquad (2)$$

The second type of observables include the on/off time distributions,[6,7] that is, the distributions of the time intervals for which the fluorescent signal is on or off, respectively.

The third type of observables which can be extracted from the experimental data include the statistical properties of the numbers of reaction events,[11] that is, the numbers of occurrences of different reactions occurring in a given time interval. The models used in this paper are variations of the basic model based on the master equation with random coefficients (1) supplemented by suitable descriptions for the fluctuations of the rate coefficients. In the following sections we focus on using these models for investigating the connections among the three types of experimental observables mentioned earlier, and the possible occurrence of oscillations in single-molecule kinetics.

2. Correlation functions and oscillations

Following Ref. 9 we start out by considering a special class of systems for which the fluctuating rate coefficients obey a separability condition $k_{uu'} = k^0_{uu'}\chi(\mathbf{s})$, that is, they are made up of the multiplicative contributions of two factors: (a) a universal factor, $\chi(\mathbf{s})$ which is fluctuating and is the same for all interaction processes and (b) process-dependent factors, $k^0_{uu'}$ which depend on the initial and final chemical states of the molecule but are not random. This separability condition makes it possible to introduce an intrinsic timescale and use the method of characteristic functionals for computing the correlation functions of the fluorescent signal. The separability condition is consistent with the condition of detailed balance for a system with a unique equilibrium state and is automatically fulfilled by a system with two chemical states and a unique equilibrium.

In Ref. 9, the theory was developed for the general case when the single molecule has an arbitrary number of chemical states and general expressions were derived for correlation functions of all orders. However, for simplicity, we begin by considering a system with two different chemical states. By applying the general theory developed in the study of Ref. 9, we obtain the following formula for the second-order correlation function:

$$\langle \Delta I(t) \Delta I(t+\tau) \rangle = \frac{K}{(K+1)^2} \mathscr{I}(\tau), \tag{3}$$

where

$$K = k_+(t)/k_+(t) \text{ independent of } t \tag{4}$$

is the equilibrium constant of the process, which, for a single chemical equilibrium is time-independent and not random. The term

$$\mathscr{I}(\Delta t) = \left\langle \exp\left(-\int_0^{\Delta t} k_\Sigma(t')dt'\right)\right\rangle \tag{5}$$

is a dynamic damping factor and

$$k_\Sigma(t) = k_+(t) + k_-(t) \tag{6}$$

is a total fluctuating rate coefficient, which is the sum of forward and backward reaction rates $k_+(t)$ and $k_-(t)$, respectively, and $\langle ... \rangle$ denotes a dynamic average over all possible values of the total fluctuating rate coefficient $k_\Sigma(t)$. For a system with two chemical states the separability condition mentioned earlier is automatically fulfilled and Eq. (3) is not subjected to any restriction.

If we assume that the cumulants $\langle\langle k_\Sigma(t)\rangle\rangle$, $\langle\langle k_\Sigma(t_1)k_\Sigma(t_2)\rangle\rangle \ldots$ of the total rate coefficient exist and are finite, the damping factor can be expressed by a cumulant expansion. We arrive at

$$\mathscr{I}(\Delta t) = \exp\left\{\sum_{m=1}^{\infty}\frac{(-1)^m}{m!}\right. \\ \left.\times \int_0^{\Delta t}\cdots\int_0^{\Delta t}\langle\langle k_\Sigma(t_1)\cdots k_\Sigma(t_m)\rangle\rangle dt_1\cdots dt_m\right\}. \tag{7}$$

The data can be analyzed in terms of the effective decay rate

$$\begin{aligned} k_{\mathrm{eff}}(\Delta t) &= -\frac{\partial}{\partial \Delta t}\ln \mathscr{I}(\Delta t) \\ &= \langle k_\Sigma\rangle + \frac{\partial}{\partial \Delta t}\sum_{m=2}^{\infty}\frac{(-1)^{m-1}}{m!} \\ &\quad\times \int_0^{\Delta t}\cdots\int_0^{\Delta t}\langle\langle k_\Sigma(t_1)\cdots k_\Sigma(t_m)\rangle\rangle\, dt_1\cdots dt_m. \end{aligned} \tag{8}$$

The effective decay rate bears information about the nature of intramolecular fluctuations. If the fluctuations of the rate of change are of short range in time (Markovian or independent fluctuations), then in the long run, the effective decay rate is independent of the time difference Δt

$$k_{\text{eff}}(\Delta t) = \text{independent of } \Delta t \text{ as } \Delta t \to \infty. \tag{9}$$

If condition (9) is fulfilled by the experimental data, then the intramolecular fluctuations are of short range in time. If $k_{\text{eff}}(\Delta t)$ varies with Δt for large time differences Δt, then the intramolecular fluctuations are of long range. In addition, we notice that the effective rate is a better function for identifying the existence of oscillations in single-molecule kinetics than the correlation function of the fluorescent signal. In order to evaluate $k_{\text{eff}}(\Delta t)$, however, accurate measurements are necessary.

In order to express the contribution of intramolecular fluctuations to the effective decay rate, we evaluate the difference

$$\begin{aligned}\Delta k_{\text{eff}}(\Delta t) &= k_{\text{eff}}(\Delta t) - \langle k_\Sigma(t)\rangle \\ &= \frac{\partial}{\partial \Delta t}\sum_{m=2}^{\infty}\frac{(-1)^{m-1}}{m!} \\ &\quad \times \int_0^{\Delta t}\cdots\int_0^{\Delta t}\langle\langle k_\Sigma(t_1)\cdots k_\Sigma(t_m)\rangle\rangle dt_1\cdots dt_m.\end{aligned} \tag{10}$$

In the particular case of Gaussian fluctuations of the total rate coefficient, all cumulants of order higher than 2 vanish and the difference $\Delta k_{\text{eff}}(\Delta t)$ is simply given by

$$\begin{aligned}\Delta k_{\text{eff}}(\Delta t) &= -\frac{1}{2}\frac{\partial}{\partial \Delta t}\int_0^{\Delta t}\int_0^{\Delta t}\langle\langle k_\Sigma(t_1)k_\Sigma(t_2)\rangle\rangle dt_1 dt_2 \\ &= -k_\Sigma\frac{\partial}{\partial \Delta t}\int_0^{\Delta t}(\Delta t - x)g(x)dx,\end{aligned} \tag{11}$$

where

$$g(|t_2 - t_1|) = \frac{1}{\langle k_\Sigma\rangle^2}\langle\langle k_\Sigma(t_1)k_\Sigma(t_2)\rangle\rangle \tag{12}$$

is the relative value of the correlation function of the total rate coefficient. The relative correlation function can be evaluated from experimental data by solving Eq. (11) for $g(\Delta t)$, resulting in

$$g(\Delta t) = -\frac{1}{k_\Sigma}\frac{\partial}{\partial \Delta t}\Delta k_{\text{eff}}(\Delta t). \tag{13}$$

This simplified model makes it possible to discuss the possible existence of damped oscillations in single-molecule kinetics. We start out by analyzing the oscillations due to the intramolecular fluctuations. For simplicity we limit ourselves to the case of Gaussian fluctuations. In this case it is easy to show that if the relative correlation function of the rate of change, $g(\Delta t)$, which expresses the intramolecular fluctuations, displays damped oscillations, then damped oscillations may also occur in the correlation functions of the fluorescent signal. According to the normal mode theory,[12] the function $g(\Delta t)$ can be expressed as

$$g(|t_1 - t_2|) = \sum_q c_q \exp[-\varepsilon_q |t_1 - t_2|] + \int_q c(q)\exp[-\varepsilon(q)|t_1 - t_2|]dq. \tag{14}$$

In this equation, c_q, $c(q)$, ε_q, $\text{Re}(\varepsilon_q) > 0$ and $\varepsilon(q)$, $\text{Re}(\varepsilon(q)) > 0$ are amplitude and frequency factors attached to the different normal modes. Since, in general, both c_q and ε_q are complex, their values must be chosen in such a way that the corresponding Gaussian process is physically consistent. For a purely discrete mode spectrum in Eq. (14) the integral term is missing, and the stochastic process, even though generally nonMarkovian, has short memory. The Markovian memory corresponds to a single exponential, that is, to a single mode. For a discrete spectrum, the Markovian approximation is accurate for large time differences, because in this case the main contribution to the sum in Eq. (14) is given by a single exponential which corresponds to the frequency with the smallest absolute value. If the mode spectrum has a continuum branch, then the tail of the correlation function may obey a scaling law of the inverse power type and the system may display long memory. Our analysis in this chapter is limited to the case of short-range fluctuations, for which Eq. (14) contains only the contribution of the discrete spectrum.

We consider the following physical constraints:

(1) For a time difference equal to zero, the autocorrelation function is equal to the dispersion of the total relative rate, $\nu(t) = k_\Sigma(t)/\langle k_\Sigma(t)\rangle$ at time t, $\langle\langle \nu^2(t)\rangle\rangle$, which, by definition, must be non-negative.
(2) Since the characteristic frequency is a real function of time, the modes with complex frequencies ε_q must occur in conjugated pairs.
(3) For large times the autocorrelation function of the relative total rate must decay to zero.

We keep in Eq. (14) only the contribution of the discrete spectrum and express the contribution of real eigenvalues $\varepsilon_q^{(\mathrm{real})}$ and of complex eigenvalues $\varepsilon_q^{(\mathrm{compl})} = \mu_q \pm i\sigma_q$. After some calculations, we obtain

$$
\begin{aligned}
g(|t_1 - t_2|) &= \sum_{\text{real values}} c_q^{(\mathrm{real})} \exp[-\varepsilon_q^{(\mathrm{real})}|t_1 - t_2|] \\
&\quad + \sum_{\text{complex values}} c_q^{(\mathrm{compl})} \exp[-\varepsilon_q^{(\mathrm{compl})}|t_1 - t_2|] \\
&= \sum_{\text{real values}} c_q^{(\mathrm{real})} \exp[-\varepsilon_q^{(\mathrm{real})}|t_1 - t_2|] \\
&\quad + \sum_{\text{complex values}} 2\{a_q \cos[\sigma_q|t_1 - t_2|] \\
&\quad + b_q \sin[\sigma_q|t_1 - t_2|]\} \exp[-\mu_q|t_1 - t_2|], \qquad (15)
\end{aligned}
$$

where a_q and b_q are the real and imaginary parts of the complex amplitude factors, and $c_q^{(\mathrm{compl})} = a_q \pm ib_q$. In order that the constraints (1)–(3) be valid we introduce the following restrictions for the parameters in Eq. (15):

$$c_q^{(\text{real})},\ a_q > 0, \quad b_q > 0 \tag{16}$$

$$\varepsilon_q^{(\text{real})},\ \mu_q > 0. \tag{17}$$

The restrictions (16) ensure that the dispersion of the characteristic frequency is non-negative, whereas the restrictions (17) are necessary in order that the autocorrelation function tends to zero for large time differences. The damping factor of the correlation function $\mathscr{I}(\Delta t)$ can be expressed as

$$\mathscr{I}(\Delta t) = \exp[-k_\Sigma \Delta t + k_\Sigma \Theta(\Delta t)], \tag{18}$$

where $\Theta(\Delta t)$ is a phase factor given by

$$\begin{aligned}
\Theta(\Delta t) = \Delta t &\left[\sum_{\text{real values}} \frac{c_q^{(\text{real})}}{\varepsilon_q} + \sum_{\text{complex values}} \frac{2(a_q\mu_q + b_q\sigma_q)}{(\mu_q)^2 + (\sigma_q)^2} \right] \\
&- \sum_{\text{real values}} \frac{c_q^{(\text{real})}}{(\varepsilon_q)^2} \\
&+ 2 \sum_{\text{complex values}} \left[\frac{a_q[(\sigma_q)^2 - (\mu_q)^2] - 2b_q\mu_q\sigma_q}{[(\mu_q)^2 + (\sigma_q)^2]^2} \right] \\
&+ \sum_{\text{real values}} \frac{c_q^{(\text{real})}}{(\nu_q)^2} \exp(-\varepsilon_q \Delta t) + \frac{2}{[(\mu_q)^2 + (\sigma_q)^2]^2} \\
&\times \sum_{\text{complex values}} \{a_q[(\mu_q)^2 - (\sigma_q)^2] \cos(\sigma_q \Delta t) + 2b_q\mu_q\sigma_q\} \\
&\times \exp(-\mu_q \Delta t) + \frac{2}{[(\mu_q)^2 + (\sigma_q)^2]^2} \\
&\times \sum_{\text{complex values}} \{b_q[(\mu_q)^2 - (\sigma_q)^2] \sin(\sigma_q \Delta t) - 2a_q\mu_q\sigma_q\} \\
&\times \exp(-\mu_q \Delta t).
\end{aligned} \tag{19}$$

The phase factor $\Theta(\Delta t)$ has the following asymptotic behavior:

$$\Theta(\Delta t) \sim \begin{cases} \mathscr{M}\Delta t^2 & \text{as } \Delta t \to 0 \\ \mathfrak{S}\Delta t & \text{as } \Delta t \to \infty \end{cases}, \tag{20}$$

where the proportionality factors $\mathscr{M}$ and $\mathfrak{S}$ are given by

$$\mathscr{M} = \frac{1}{2} \sum_{\text{real values}} c_q^{(\text{real})} + \sum_{\text{complex values}} a_q > 0, \tag{21}$$

$$\mathfrak{S} = \sum_{\text{real values}} \frac{c_q^{(\text{real})}}{\varepsilon_q} + \sum_{\text{complex values}} \frac{2(a_q\mu_q + b_q\sigma_q)}{(\mu_q)^2 + (\sigma_q)^2} > 0. \tag{22}$$

In this case the variation of the effective rate coefficient has the asymptotic behavior

$$\Delta k_{\text{eff}}(\Delta t) \sim k_\Sigma \begin{cases} -2\mathscr{M}\Delta t & \text{as } \Delta t \to 0 \\ -\mathfrak{S} & \text{as } \Delta t \to \infty \end{cases}. \tag{23}$$

For large as well as short time differences, the variation $\Delta k_{\text{eff}}(\Delta t)$ is negative: as expected, for large time differences the variation is constant, as it should be for short-range intramolecular fluctuations.

According to Eq. (19) the phase factor $\Theta(\Delta t)$ may display damped oscillations in the time difference Δt. From Eq. (18) it follows that the same type of damped oscillation must be displayed by the damping factor $\mathscr{I}(\Delta t)$. We did a numerical study of the possible occurrence of damped oscillations in the correlation functions of the fluorescent signal, due to the presence of damped oscillations in intramolecular dynamics, represented by the complex eigenmodes in Eq. (19). In order for the damped oscillations to show up in the correlation functions on the fluorescent signal, it is necessary that the timescale of the chemical process be of the same order of magnitude as the timescale of intramolecular dynamics. Figure 1 shows such a damped oscillating behavior for the second-order correlation function of the fluorescent signal $C_2(\tau) = \langle \Delta I(t+\tau)\Delta I(t)\rangle$, which is similar to the oscillations observed in the experiments of Edman and Rigler.[10] Similar behavior is displayed by the correlation functions of higher order.

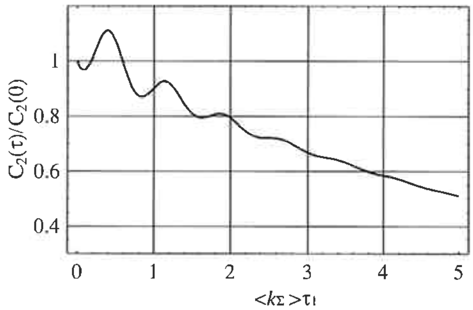

Fig. 1. Graphical representation of the absolute value of the second-order correlation function $C_2(\tau)$ versus the time difference $\tau = |t_2 - t_1|$ for a two-state Gaussian model. (The relative correlation function $g(\Delta t)$ is represented by a linear combination of complex exponential terms.)

It has been suggested that the damped oscillations of the correlation functions of high order can be used for the characterization of the nonMarkovian nature of the two-state fluorescent process.[10] A non-Markovian function (NMF) has been defined in terms of the second- and third-order correlation functions of the fluorescent signal:

$$\mathrm{NMF}(\tau_1, \tau_2) = p_f \left[\frac{C_3(\tau_1, \tau_2)}{C_2(\tau_2)} - C_2(\tau_1) \right], \tag{24}$$

where p_f is the stationary probability that the molecule is in a fluorescent state. Figure 2 shows a typical memory landscape for the NMF computed by applying our approach from Eqs. (3)–(5), (18), and (19), and the expression

$$\langle \Delta I(t) \Delta I(t + \tau_1) \Delta I(t + \tau_1 + \tau_2) \rangle = \mathscr{I}(\tau_1 + \tau_2) \frac{K(K-1)}{(K+1)^3}, \tag{25}$$

for the third-order correlation function, computed by applying our theory presented in Ref. 9. The computed landscape displays the same type of damped oscillations as the ones observed in the experiments of Edman and Rigler.[10]

Another possible cause for the occurrence of the damped oscillations of the correlation functions is the interaction between chemical kinetics and intramolecular dynamics. This cause was suggested by

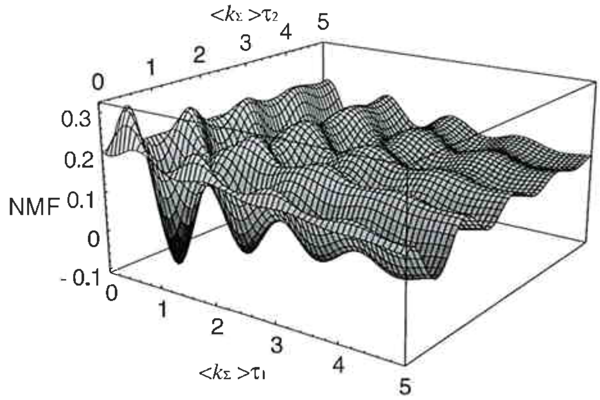

Fig. 2. Graphical representation of the non-Makovian function NMF(τ_1, τ_2) of Edman and Rigler[10] versus the time differences τ_1 and τ_2 for a two-state Gaussian model. (The relative correlation function $g(\Delta t)$ is represented by a linear combination of complex exponential terms.)

Edman and Rigler[10] in order to explain their experimental data on the oxidation reaction involving a single molecule of immobilized horseradish peroxidase. These authors neglected the random fluctuations of the rate coefficients and assumed that the kinetics of the process can be described by a simplified form of the master equation (1), where the rate coefficients $k_{u'u}$ are constant. They have chosen sets of rates $k_{u'u}$, which correspond to closed loops of states and violate detailed balance. It has been theoretically proven that a master equation with rates $k_{u'u}$, which violate the detailed balance, are capable of producing damped oscillations.[9] The model of Edman and Rigler[10] seems to contradict the principles of statistical mechanics, because they evaluate time invariant, equilibrium correlation functions by using a model which violates the detailed balance. However, we show shortly that this is not necessarily the case. A single molecule is not a macroscopic system; therefore, it does not have to obey equilibrium statistical mechanics; however, a single molecule is not isolated, but connected to its environment, and in most experiments, the ensemble molecule plus environment are at statistical equilibrium. Actually, most experimental studies of single-molecule kinetics involve the measurement of the regression of the equilibrium fluctuations of the fluorescent signal.

We consider a different approach, which shows that the model of Edman and Rigler[10] may be correct and do not violate detailed balance. We assume that the intramolecular dynamics, expressed in terms of the control parameters $\mathbf{s}(t) = (s_1(t), s_2(t), \ldots)$, can be described by a Markovian stochastic process. We denote by $\mathscr{R}(\mathbf{s}; t)d\mathbf{s}$ the probability that at time t the vector of control parameters is between $\mathbf{s}$ and $\mathbf{s} + d\mathbf{s}$ and assume that its time evolution is described by a linear evolution equation:

$$\frac{\partial}{\partial t}\mathscr{R}(\mathbf{s}; t) = \mathbb{L}\mathscr{R}(\mathbf{s}; t), \tag{26}$$

where $\mathbb{L}$ is a Markovian operator of the Fokker–Planck, master or Liouville type. We introduce the joint probability density $B_u(\mathbf{s}; t)$ for the chemical state of the molecule u and the control vector $\mathbf{s}$. This joint probability density is the solution of a compound stochastic Liouville equation[13,14]:

$$\begin{aligned}\frac{\partial}{\partial t}B_u(\mathbf{s}; t) = \mathbb{L}B_u(\mathbf{s}; t) &+ \sum_{u' \neq u} B_{u'}(\mathbf{s}; t)k_{u'u}(\mathbf{s}) \\ &- B_u(\mathbf{s}; t)\sum_{u' \neq u} k_{uu'}(\mathbf{s}).\end{aligned} \tag{27}$$

We are interested in the evaluation of the marginal probability

$$P_u(t) = \int B_u(\mathbf{s}; t)d\mathbf{s}, \tag{28}$$

in terms of which we can compute the experimental observables, the correlation functions of the fluorescence signal. A simple way would be to derive an approximate equation for the marginal probability $P_u(t)$ by eliminating the stochastic vector $\mathbf{s}$ from Eq. (27). This is a standard topic in statistical physics,[13,14] which is usually referred to as the "renormalization of stochastic evolution equations." In quantum field theory, a "bare" particle interacting with a field is replaced by a "dressed" particle with renormalized parameters, which take into account the contribution of the field. By analogy with quantum field theory, we start out with a set of "bare" rate coefficients, $k_{u'u}(\mathbf{s})$, which depend on the fluctuations of the control parameters and then introduce a set of "dressed", renormalized rate coefficients $\check{k}_{u'u}$, which express the contribution of the fluctuations of the control parameters. If the intramolecular fluctuations, expressed in terms of the control parameters, have a short correlation time compared to chemical dynamics, the renormalized rate coefficients can be evaluated by using Van Kampen's renormalization method based on a cumulant expansion.[13,14] We get a "dressed" master equation for the marginal probability $P_u(t)$

$$\frac{\partial}{\partial t}P_u(t) = \sum_{u' \neq u} P_{u'}(t)\check{k}_{u'u} - P_u(t)\sum_{u' \neq u} \check{k}_{uu'}, \tag{29}$$

where the renormalized rate coefficients $\check{k}_{u'u}$ are expressed by fluctuation–dissipation relations as integrals of functional transformations of the correlation functions of the control variables. These expressions can be derived by using a cumulant expansion; in order to save space they are not given here. The important thing is that Eq. (29) is of the type used by Edman and Rigler[10] in their analysis. By starting out from a set of "bare" rate coefficients $k_{u'u}(\mathbf{s})$, which obey detailed balance, we end up with a set of renormalized, "dressed" rate coefficients $\check{k}_{u'u}$, which do not have to obey a similar condition of detailed balance. In general the renormalized rate coefficients $\check{k}_{u'u}$ are different from the average values of the "bare" rate coefficients, $\langle k_{u'u}(\mathbf{s})\rangle$. The differences

$$\Delta\check{k}_{u'u} = \check{k}_{u'u} - \langle k_{u'u}(\mathbf{s})\rangle, \tag{30}$$

measure the contribution of intramolecular fluctuations to the values of the renormalized rate coefficients.

The model of Edman and Rigler is valid if the following conditions are fulfilled:

(a) The intramolecular fluctuations have a short correlation time compared to the chemical timescale. If this condition is fulfilled, then the stochastic Liouville equation (27) can be replaced by the time homogeneous Markovian equation (29).
(b) Although short, the correlation time of intramolecular fluctuations is long enough so that the renormalized rate coefficients $\check{k}_{u'u}$ are different from the average values of the "bare" rate coefficients, $\langle k_{u'u}(\mathbf{s})\rangle$, that is, $\Delta\check{k}_{u'u} \neq 0$. If this constraint is fulfilled, it is possible that the renormalized rate coefficients do not obey detailed balance even though the bare coefficients do.
(c) The chemical states of the system are connected with at least a loop[15]; for this condition to be fulfilled there must be at least three chemical states.

An interesting problem pointed out by a referee is to identify the conditions for which "bare" kinetic equations obeying detailed balance would lead to "renormalized" kinetic equations which violate detailed balance. In general, this is still an open problem. Based on the method of projection operators, recently we have identified a few particular cases for which "bare" detailed balance leads to "renormalized" detailed balance.

It is interesting to compare the two mechanisms for damped chemical oscillations of correlation functions discussed in this paper. Although both mechanisms involve the intramolecular fluctuations expressed in terms of the random variations of control variables, their role is different in the two cases. In the first case, the timescale of the intramolecular fluctuations must be of the same order of magnitude as the chemical dynamics, and the damped oscillations of the correlation functions of the fluorescent signal are a direct result of damped oscillations at the intramolecular level. Here chemical dynamics plays a marginal role in the occurrence of damped oscillations; the oscillations may emerge even if there are only two chemical states. In the second case the intramolecular fluctuations play an indirect role: their job is to produce "dressed" rate coefficients which may violate detailed balance. In this case the intramolecular fluctuations are characterized by a characteristic time, which is smaller than the timescale of chemical kinetics. The chemistry plays a major role in the generation of the oscillations: they are produced by feedback loops involving at least three chemical states, which may violate detailed balance. In contrast to the first case, damped oscillations may exist even if the intramolecular dynamics do not have an oscillatory component.

There is a unified approach which can be used for describing both types of oscillations, intramolecular and chemical, respectively. The starting point is the stochastic Liouville equation (27). Instead of using approximate renormalization group methods for the elimination of the control parameters, we use an exact method, based on the use of an additional variable, the age a[16,17] of a given chemical state of the single molecule. We introduce a joint probability density $\mathscr{B}_u(a, \mathbf{s}; t)$ for the state u of the molecule, the age a of the state u, and the vector $\mathbf{s}$ of the control parameters, all evaluated at time t and the marginal probability density:

$$\mathscr{R}_u(a; t) = \int_{\mathbf{s}} \mathscr{B}_u(a, \mathbf{s}; t) d\mathbf{s} \tag{31}$$

of the state u of the molecule and the age a of the state u at time t. $\mathscr{B}_u(a, \mathbf{s}; t)$ is the solution of a system of age-dependent stochastic Liouville equation:

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) \mathscr{B}_u(a, \mathbf{s}; t) = \mathbb{L}\mathscr{B}_u(a, \mathbf{s}; t) - \mathscr{B}_u(a, \mathbf{s}; t) \sum_{u' \neq u} k_{uu'}(\mathbf{s}), \tag{32}$$

$$\mathscr{B}_u(0, \mathbf{s}; t) = \sum_{u' \neq u} k_{u'u}(\mathbf{s}) \int_0^\infty \mathscr{B}_{u'}(a', \mathbf{s}; t) da'. \tag{33}$$

Equations (32) and (33) have a special structure which makes it possible to eliminate formally the vector $\mathbf{s}$ of control variables; such elimination methods were developed in connection with the description of heavy-ion collisions in nuclear physics.[17] The elimination leads to closed systems of equations for the marginal probability $\mathscr{R}_u(a; t)$:

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) \mathscr{R}_u(a; t) = -\mathscr{R}_u(a; t) \sum_{u' \neq u} \kappa_{uu'}(a, t), \tag{34}$$

$$\mathscr{R}_u(0; t) = \sum_{u' \neq u} \kappa_{u'u}(a, t) \int_0^\infty \mathscr{R}_{u'}(a'; t) da', \tag{35}$$

where $\kappa_{u'u}(a, t)$ are renormalized rate coefficients which are complicated functionals of the Green functions $\mathfrak{G}_u(\mathbf{s}, t|\mathbf{s}_0, t_0)$ attached to the operator $\mathbb{L}$, which are the solutions of the equation $(\partial_t - \mathbb{L})\mathfrak{G}_u(\mathbf{s}, t|\mathbf{s}_0, t_0) = \delta(\mathbf{s} - \mathbf{s}_0)\delta(t - t_0)$. The evaluation of $\kappa_{u'u}(a, t)$ requires the numerical evaluation of the Green functions $\mathfrak{G}_u(\mathbf{s}, t|\mathbf{s}_0, t_0)$. In general the renormalized rates $\kappa_{uu'}(a, t)$ depend both on the age a of the state u and on the current time t. If the fluctuations of the control parameters are stationary, then the renormalized rate coefficients depend only on age, not on time, $\kappa_{uu'}(a, t) = \kappa_{uu'}(a)$ independent of t. Unlike the renormalized master equations (29), Eqs. (34) and (35) are exact; they describe both intramolecular as well as chemical fluctuations. If the renormalized rate coefficients are independent of time, then Eqs. (34) and (35) can be transformed into a generalized master equation (GME)[16,18]:

$$\frac{\partial}{\partial t} P_u(t) = \int_0^t \sum_{u'} [P_{u'}(t - \Delta t)\omega_{u'u}(\Delta t) - P_u(t - \Delta t)\omega_{uu'}(\Delta t)] d\Delta t, \tag{36}$$

where $\omega_{u'u}(\Delta t)$ are acceleration coefficients which can be computed in terms of the renormalized rate coefficients. We have

$$\omega_{uu'}(\Delta t) = \mathscr{L}^{-1}_{(\Delta t, s)} \times \left\{ \frac{\mathscr{L}_{(s,a)}\left[\kappa_{uu'}(a) \exp\left(-\sum_{u''} \int_0^a \kappa_{uu''}(a'') da''\right)\right]}{\mathscr{L}_{(s,a)} \exp\left(-\sum_{u''} \int_0^a \kappa_{uu''}(a'') da''\right)} \right\}, \tag{37}$$

where $\mathscr{L}_{(s,a)}$ and $\mathscr{L}^{-1}_{(\Delta t,s)}$ denote the direct and inverse Laplace transformations, respectively.

Among the three sets of renormalized rate parameters, $\check{k}_{uu'}$, $\kappa_{uu'}(a)$, and $\omega_{uu'}(\Delta t)$, the following asymptotic relations are obtained:

$$\check{k}_{uu'} = \lim_{a\to\infty} \kappa_{uu'}(a) = \int_0^\infty \omega_{uu'}(\Delta t)d\Delta t. \tag{38}$$

The dependence of $\kappa_{uu'}(a)$ and $\omega_{uu'}(\Delta t)$ on a and Δt is due to intramolecular fluctuations. The approximate renormalized equations (29) do not contain detailed information about intramolecular dynamics, and thus they are able to describe only chemical oscillations. In contrast, the exact renormalized equations (34) and (35) and (36) are formally equivalent to the stochastic Liouville equations (32) and (33), and describe both intramolecular as well as chemical kinetics, therefore, they are able to describe both types of oscillations.

In conclusion, in single-molecule kinetics, both intramolecular as well as chemical processes can produce oscillations. There are no simple recipes to distinguish between these two types of oscillations; intramolecular oscillations tend to be faster than the chemical oscillations but there can be situations for which this is not true. Regarding the description of oscillation displayed by correlation functions, there are different types of models available. (1) Global techniques based on the use of characteristic functionals; (2) Local methods based on the use of stochastic Liouville equations, combined with the use of age-dependent and generalized master equations and the renormalization group approach.

3. On–off time distributions and oscillations

The analysis of on–off time distributions in single-molecule kinetics and their connections with chemical and intramolecular oscillations can be carried out by using an approach based on age-dependent master equations,[19] similar to Eqs. (32)–(35) combined with the method of characteristic functionals.[9] We consider a given realization of the random rate coefficients $\mathbf{k} = \mathbf{k}(t)$ with $\mathbf{k} = [k_{uu'}]$, $k_{uu} = 0$; for this realization we introduce the joint probability

$$\mathscr{P}_u(a; t, \mathbf{k}(t'))d\tau = \mathscr{P}_u(a; t)d\tau, \tag{39}$$

for the chemical state u of the molecule and for the age a of the state u at time t. In the following considerations we describe the fluctuations of the rates directly, without the use of any control parameters. This joint probability obeys the normalization condition $\sum_u \int_0^\infty \mathscr{P}_u(a; t)da = 1$. In terms of $\mathscr{P}_u(a; t)da$ we can introduce the probability density $\gamma(a|u; t)$ of the age (lifetime) of a given chemical state u, which obeys the normalization condition, $\int_0^\infty \gamma(a|u; t)da = 1$, and can be expressed in terms of the joint probability $\mathscr{P}_u(a; t)da$ as

$$\gamma(a|u; t)da = \frac{\mathscr{P}_u(a; t)da}{\int_0^\infty \mathscr{P}_u(a; t)da} = \frac{\mathscr{P}_u(a; t)da}{P_u(t)}, \tag{40}$$

where $P_u(t) = \int_0^\infty \mathscr{P}_u(a; t)da$ is a state probability which is the solution of the master equation (1). The conditional probability $\gamma(\tau|u; t)$ is the solution of a system of modified age-dependent master equations[19]:

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right)\gamma(a|u; t) = -\gamma(a|u; t)\tilde{\Theta}_u(t), \tag{41}$$

$$\gamma(0|u; t) = \sum_{u' \neq u} \tilde{k}_{u'u}(t) \int_0^\infty \gamma(a|u'; t)d\tau = \sum_{u' \neq u} \tilde{k}_{u'u}(t) = \tilde{\Theta}_u(t), \tag{42}$$

where

$$\tilde{k}_{u'u}(t) = k_{u'u}(t)P_{u'}(t)/P_u(t) \tag{43}$$

are adjoint rate coefficients and

$$\tilde{\Theta}_u(t) = \sum_{u' \neq u} \tilde{k}_{u'u}(t) \tag{44}$$

is a total (fluctuating) adjoint transition (decay) rate attached to the chemical state u. Equation (41) has an exact formal solution. We assume that the fluctuations of the rate coefficients are stationary, a hypothesis which is justified by the fact that usually single-molecule kinetic experiments are carried out at statistical equilibrium. For this reason, without loss of generality, we can push the initial condition to minus infinity, $t_0 \to -\infty$, and can represent this formal solution as

$$\gamma(a|u;t) = \tilde{\Theta}_u(t-a)\exp\left[-\int_{t-a}^{t}\tilde{\Theta}_u(t')dt'\right]. \qquad (45)$$

An experimental observable is the averaged distribution of the lifetime over all possible values of the rate coefficients $\varphi(a|u) = \langle\gamma(a|u;t)\rangle$. We have

$$\varphi(a|u) = \left\langle\tilde{\Theta}_u(t-a)\exp\left[-\int_{t-a}^{t}\tilde{\Theta}_u(t')dt'\right]\right\rangle, \qquad (46)$$

where the dynamic average $\langle\ldots\rangle$ is taken over all possible fluctuations of the rate coefficients. In Eq. (46) we take into account that the fluctuations of the rate coefficients are stationary, and as a result the average probability density of the lifetimes, $\varphi(a|u)$, depends on the lifetime a and is independent of the current time, t. For a given chemical state u, this dynamical average can be expressed in terms of a single random variable, the total adjoint rate $\tilde{\Theta}_u(t)$. By assuming that the cumulants of the total adjoint rate $\tilde{\Theta}_u(t)$ exist and are finite, the experimental observable $\varphi(a|u)$ can be easily evaluated. We have

$$\varphi(a|u) = \tilde{\Theta}_u^{\mathrm{eff}}(a)\exp\left[-\int_0^a\tilde{\Theta}_u^{\mathrm{eff}}(a')da'\right], \qquad (47)$$

where $\tilde{\Theta}_u^{\mathrm{eff}}(a)$ is the total adjoint effective decay rate attached to the state u, which is given by a cumulant expansion similar to Eq. (8)

from Sec. 2:

$$\tilde{\Theta}_u^{\text{eff}}(a) = \langle \tilde{\Theta}_u \rangle + \frac{\partial}{\partial a} \sum_{m=2}^{\infty} \frac{(-1)^{m-1}}{m!}$$
$$\times \int_0^a \cdots \int_0^a \langle\langle \tilde{\Theta}_u(t_1) \cdots \tilde{\Theta}_u(t_m) \rangle\rangle dt_1 \cdots dt_m. \qquad (48)$$

The identical structures of Eqs. (8) and (48) for $k_{\text{eff}}(\Delta t)$ and $\tilde{\Theta}_u^{\text{eff}}(a)$, respectively, make it possible to bring many of the results about correlation functions to the study of on/off time distributions, and vice versa. In particular, the survival function of a state u,

$$\ell(a|u) = \int_a^{\infty} \varphi(a|u) da = \exp\left[-\int_0^a \tilde{\Theta}_u^{\text{eff}}(a') da' \right], \qquad (49)$$

plays the same role as the two-point correlation function of the fluorescent signal. We notice however that the rate coefficients involved in the two cases are different. They are direct rates for correlation functions and adjoint rates for on/off time distributions. In addition, we notice that the results about correlation functions are valid only for separable models, whereas the results for on/off times are more general; they apply to any master equation (1) with fluctuating rate coefficients. As far as we know, no experimental data about oscillations and their possible connections with on/off statistics have been published in the literature. Anyway, the same approaches as the ones used for correlation functions can be applied for the analysis of the oscillations. In particular, the analysis of the effective rates $\tilde{\Theta}_u^{\text{eff}}(a)$ is more suitable for the identification and analysis of oscillation, rather than the distribution $\varphi(a|u)$ or the survival function $\ell(a|u)$.

4. Reaction event statistics and oscillations

Of all observables in single-molecule kinetics, the reaction events have the longest history; they were studied decades before single-molecule kinetic experiments were possible. The numbers of reaction events as stochastic variables in time were first discussed over

80 years ago in Jean Perrin's celebrated book, *The Atoms*.[20] In 1974, Milan Solc[21–23] computed the probability distribution of the reaction events for a monomolecular reaction. In connection with a problem of nuclear physics, Vlad and Pop[24] developed a systematic Lippmann–Schwinger expansion for the solution of a master equation for the joint probability density of the state and the number of transition events for Markovian process in continuous time and discrete state space.[24] A more general expansion approach, which in particular can be applied to chemical reactions, was developed by Vlad and Ross.[25] Two different Lippmann–Schwinger expansions are developed: the first one produces exact expressions for the probabilities of the reaction events, and the second one produces exact expressions for the factorial moments and cumulants of the reaction events.

Regarding the applications to single-molecule kinetics, a serious mathematical difficulty arises: the average over the fluctuations of the rate coefficients is very hard to compute analytically or numerically. The simplest approach is to ignore the fluctuations of rate coefficients.[26,27] Although the method of Lippmann–Schwinger expansions can be extended to fluctuating rate coefficients, its application is numerically intensive; it requires repeated numerical integration of the master equations for different realizations of the rate coefficients, storing the resulting Green functions in a database, followed by the evaluation of dynamic averages.

The basic theory starts from a master equation for the joint probability $F_u(\mathbf{q}; t)$ of the state u and of the matrix $\mathbf{q} = [q_{uu'}]$ of the numbers $q_{uu'}$ of reaction events $u \to u'$:

$$\frac{\partial}{\partial t} F_u(\mathbf{q}; t) = \sum_{u' \neq u} F_u(\cdots q_{uu'} - 1; \ldots; t) k_{u'u}(t) - F_u(\mathbf{q}; t) \times \sum_{u'' \neq u} k_{uu''}(t). \tag{50}$$

If the fluctuations of the rate coefficients are neglected, $k_{u'u}(t) = k_{u'u}$, independent of t, and $F_u(\mathbf{q}; t)$ as well as the moments and cumulants of $q_{uu'}$ can be easily evaluated by deriving from Eq. (50) an equation for the partial generating function of $F_u(\mathbf{q}; t)$, $G_u(s; t) = \sum_{\mathbf{q}} \prod_{uu'} (s_{uu'})^{q_{uu'}} F_u(\mathbf{q}; t)$ with $|s_{uu'}| < 1$, followed by the use of Lippmann–Schwinger expansions. (For details, see Ref. 25). In order to take the flucuations of the rate coefficients into account, the method should be applied repeatedly for different realizations of the rate coefficients $k_{u'u}(t)$, followed by the evaluation of dynamic averages.

For fluctuating rate coefficients, a simpler procedure is based on the use of age-dependent renormalization technique mentioned in Sec. 2. For simplicity we take the fluctuating rates themselves as control parameters, $\mathbf{k} = \mathbf{s}$, and write a system of stochastic Liouville, age-dependent equations for the joint probability density $Q_u(a, \mathbf{k}, \mathbf{q}; t)$ of the state u of the molecule, its age a, its matrix $\mathbf{k} = \mathbf{s}$ of the rate coefficients (control parameters), and its matrix $\mathbf{q}$ of the numbers of reaction events

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) Q_u(a, \mathbf{k}, \mathbf{q}; t) = \mathbb{L} Q_u(a, \mathbf{k}, \mathbf{q}; t) - Q_u(a, \mathbf{k}, \mathbf{q}; t) \sum_{u' \neq u} k_{uu'}, \tag{51}$$

$$Q_u(0, \mathbf{k}, \mathbf{q}; t) = \sum_{u' \neq u} k_{u'u} \int_0^\infty Q_u(a', \mathbf{k}, \ldots, q_{u'u} - 1; \ldots; t) da'. \tag{52}$$

The elimination of the fluctuation rate coefficients in Eqs. (51) and (52) leads to a system of renormalized age-dependent master equation for the joint probability density $\mathscr{F}_u(a, \mathbf{q}; t)$ of the state u of the molecule, its age a, and its matrix $\mathbf{q}$ of the numbers of reaction events

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) \mathscr{F}_u(a, \mathbf{q}; t) = -\mathscr{F}_u(a, \mathbf{q}; t) \sum_{u' \neq u} \kappa_{uu'}(a, t), \tag{53}$$

$$\mathscr{F}_u(0, \mathbf{q}; t) = \sum_{u' \neq u} \kappa_{u'u}(a, t) \int_0^\infty \mathscr{F}_{u'}(a', \ldots, q_{u'u} - 1; \ldots; t) da', \tag{54}$$

where $\kappa_{u'u}(a, t)$ are the same renormalized rate coefficients as the ones from Eqs. (34) and (35). In particular, if the fluctuations of the rate coefficients (control parameters) are stationary, then the renormalized rate coefficients depend only on age, and not on time $\kappa_{uu'}(a, t) = \kappa_{uu'}(a)$, independent of t. In this case too, it is possible to use a generating function transformation, $\mathscr{G}_u(a, s; t) = \sum_{\mathbf{q}} \prod_{uu'} (s_{uu'})^{q_{uu'}} \mathscr{F}_u(a, \mathbf{q}; t)$ with $|s_{uu'}| < 1$, followed by the use of Lippmann–Schwinger expansions.

The application of these approaches in the general case to experimental data is rather complicated and involves the use of advanced numerical techniques. For separable systems, the average over the fluctuations of the rate coefficients can be carried out analytically. Although complicated, the theory leads to analytical results, which are similar to the ones obtained for the correlation functions of the fluorescent signal. In particular, for a system with two states, one fluorescent and the other one non-fluorescent, the second-order correlation function of the "out" reaction events, attached to each state, can be expressed in terms of the damping factor $\mathscr{I}(\Delta t)$, given by Eq. (7). It follows that the considerations account for the relations between oscillations and correlation functions can be easily adapted to the study of the relations between oscillations and the statistics of the reaction events.

5. Conclusions

In this chapter we tried to get the reader acquainted with the process of building stochastic models for single-molecule kinetics, with special reference to oscillations. The emphasis was on model building, rather than on model classification or model solving. We have illustrated model building based on various techniques from nonequilibrum statistical physics, including both global as well as local methods: global characteristic functionals, stochastic Liouville equations, age-dependent and generalized master equations, and different stochastic renormalization techniques.

We tried to show how the models can be used for describing intramolecular as well as chemical oscillations in single-molecule kinetics. Based on theoretical considerations, we also tried to develop criteria for distinguishing between these two types of oscillations. The difficulty of the analysis of experimental data depends on the type of experimental observable used. On/off time data statistics are the easiest to process. Moreover, the theory imposes few restrictions for the models, which describe on/off time distributions; in particular, no separability conditions are needed. Building models for the other two observables, the correlation functions of the fluorescent signal and the statistics of the reaction events, is more complicated. These models can be solved analytically only in a few cases (separable models, deterministic rate coefficients). In general the use of these models require advanced numerical computations.

Acknowledgments

This research has been supported in part by the National Science Foundation and by the CEEX-M1-C2-3004/2006 Grant of the Romanian Ministry of Research and Education.

References

1. R. M. Dickson, A. B. Cubitt, R. Y. Tsien and E. W. Moerner, *Nature* **388** (1997) 355.
2. P. H. Lu and X. S. Xie, *Journal of Physical Chemistry B* **10** (1997) 2753.
3. G. K. Schenter, H. P. Lu and X. S. Xie, *Journal of Physical Chemistry A* **103** (1999) 10477.
4. E. J. Peterman, S. Brasselet and E. W. Moerner, *Journal of Physical Chemistry A* **103** (1999) 10553.
5. A. D. Mehta, M. Rief, J. A. Spudich, D. A. Smith and R. M. Simmons, *Science* **283** (1999) 1689.
6. W. E. Morner and M. Orrit, *Science* **283** (1999) 1670.
7. W. E. Moerner, *Journal of Physical Chemistry B* **106** (2002) 910.
8. J. Wang and P. Wolynes, *Journal of Chemical Physics A* **110** (1999) 4812.