

## APPROXIMATE SOLUTION METHODS FOR LINEAR STOCHASTIC DIFFERENCE EQUATIONS

### I. MOMENTS

J.B.T.M. ROERDINK

*Institute for Theoretical Physics, Princetonplein 5, Utrecht, The Netherlands*

Received 9 December 1982

The cumulant expansion for linear stochastic differential equations is extended to the case of linear stochastic difference equations. We consider a vector difference equation, which contains a deterministic matrix  $A_0$  and a random perturbation matrix  $A_1(t)$ . The expansion proceeds in powers of  $\alpha\tau_c$ , where  $\tau_c$  is the correlation time of the fluctuations in  $A_1(t)$  and  $\alpha$  a measure for their strength. Compared to the differential case, additional cumulants occur in the expansion. Moreover one has to distinguish between a nonsingular and a singular  $A_0$ . We also discuss a limiting situation in which the stochastic difference equation can be replaced by a stochastic differential equation. The derivation is not restricted to the case where in the limit the stochastic parameters in the difference equation are replaced by white noise.

### 1. Introduction

In this and a subsequent paper we study the stochastic difference equation

$$u(t) = A(t, \omega)u(t-1) + f(t, \omega) \quad (t = t_0 + 1, t_0 + 2, \dots), \quad (1.1a)$$

$$u(t_0) = u_0(\omega), \quad (1.1b)$$

where  $u(t)$  is an  $n$ -dimensional vector.  $A(t, \omega)$  and  $f(t, \omega)$  are a random time-dependent coefficient matrix and inhomogeneous vector, respectively\*. The initial condition  $u_0(\omega)$  may in general be taken random too. The random nature of all these quantities is indicated by the parameter  $\omega$ , which is an element of a set  $\Omega$  which, together with a  $\sigma$ -algebra  $\Sigma$  on  $\Omega$  and a probability measure  $P$  on  $\Sigma$ , constitutes a probability space. In the following we will usually omit the parameter  $\omega$ .

We are interested in the statistical properties of the solution  $u(t, \omega)$  of (1.1), such as its moments, correlation function (or covariance), distribution function etc., when the statistical properties of  $A$ ,  $f$  and  $u_0$  are prescribed (by giving their

\*We use the words "random" and "stochastic" interchangeably. A non-negative matrix, whose columns sum to one will be called a "Markov matrix"<sup>1)</sup>, instead of "stochastic matrix"<sup>2)</sup>.

joint probability distribution or joint moments, cumulants etc.). Averages with respect to  $P$  are indicated by angular brackets, e.g.

$$\langle u(t) \rangle = \int_{\Omega} dP(\omega) u(t, \omega). \quad (1.2)$$

Probably the first studied equations of this type are the autoregressive equations, which are used in time series analysis and economy<sup>3-7</sup>). In this case the inhomogeneous term  $f$  in (1.1) is random, but not  $A$ . Such equations are called additive.

Equations where also the matrix  $A$  is random are called multiplicative. A one-dimensional example was discussed by Paulson and Uppuluri<sup>8</sup>) in their study of compartment problems. For an application to random walks in random environments see ref. 9, section 4. A summary of known results for the one dimensional case is given by Vervaat<sup>10</sup>), who also provides an extensive bibliography. The  $n$ -dimensional case was studied by Kesten<sup>11</sup>).

All authors mentioned so far consider the situation where the sequence of pairs  $\{A(t), f(t)\}$  is independent identically distributed (i.i.d), although  $A$  and  $f$  may be statistically dependent. In this special case it is often possible to obtain exact results concerning the limiting behaviour of the probability distribution of  $u(t)$  as  $t \rightarrow \infty$ .

In this paper we will be concerned with eq. (1.1) where  $\{A(t)\}$  is an autocorrelated stochastic matrix sequence,  $f(t)$  is identically zero and the initial condition is non-random. In a second paper we treat the general case where both  $A$  and  $f$  are autocorrelated and possibly crosscorrelated with each other and with the random initial condition  $u_0(\omega)$ .

Because of the much higher complexity of the problem it will in general be impossible to obtain exact results<sup>†</sup>. So our goal will be to obtain a systematic perturbation expansion for the case of small and rapid fluctuations in  $A$ . That is we assume  $A(t)$  to be of the form  $A(t) = A_0 + \alpha A_1(t, \omega)$  where  $A_0$  is a deterministic matrix,  $A_1$  random and  $\alpha$  a small parameter measuring the magnitude of  $A_1$ . An important property of the expansion is that it is valid for large  $t$ , i.e. it avoids secular terms.

The method we use is a straightforward generalization of the cumulant expansion for linear stochastic differential equations<sup>13-16</sup>). The basic assumption is that the fluctuations in  $A_1(t, \omega)$  (as measured by the parameter  $\alpha$ ) are small and that the cumulants of  $A_1$  decay rapidly. In this case one can define an autocorrelation time  $\tau_c$  which is a measure for the time interval within which the

<sup>†</sup>For some limit theorems on products of random (possibly correlated) matrices, see Furstenberg and Kesten<sup>12</sup>).

autocorrelation is still appreciable. The expansion then proceeds in powers of  $\alpha\tau_c$ . The result is that the average of the solution  $u(t, \omega)$  of (1.1) with  $f \equiv 0$  and a fixed initial condition  $u_0$  obeys itself a first order difference equation

$$\langle u(t) \rangle = \{1 + K(t/t_0)\} \langle u(t-1) \rangle \quad (t = t_0 + 1, t_0 + 2, \dots), \quad (1.3)$$

where  $K(t/t_0)$  is a deterministic matrix, which is obtained in the form of an expansion involving the correlation functions of  $A_1(t, \omega)$ . Moreover,  $K(t/t_0)$  is independent of  $t_0$  when  $t - t_0 \gg \tau_c$ . Higher moments of  $u$ , such as  $\langle u(t) \otimes u(t) \rangle$  where  $\otimes$  denotes a Kronecker product, can be obtained in the same way, because  $u(t) \otimes u(t)$  obeys an equation of the same type as  $u(t)$  itself. Multi-time averages, such as the correlation functions  $\langle u(t) \otimes u(t') \rangle$  will be considered in part II.

Equations of type (1.1) with autocorrelated noise have recently been studied in population biology as a model for age-structured populations in random environments, see for example Cohen<sup>17</sup>), Tuljapurkar and Orzack<sup>18</sup>) and Tuljapurkar<sup>19</sup>). These authors assume that the matrix  $A_1(t, \omega)$  in (1.1) is an  $n$ -dimensional Markov chain\* endowed with certain ergodic or mixing properties which are mathematically more precise formulations of our assumption of a small correlation time  $\tau_c$ . In this case the bivariate process  $\{u(t), A(t)\}$  constitutes again a Markov chain, which makes it possible to evaluate various quantities exactly, such as moments or the invariant long run distribution (see ref. 17, III). In part II we will show that the moments satisfy an exact difference equation from which they can easily be calculated<sup>26</sup>).

It should be emphasized that we do not make these assumptions in the present treatment. Our only assumptions concern the parameters  $\alpha$  and  $\tau_c$  as explained above. It will turn out that if  $\alpha\tau_c$  is small enough the lowest order moments of  $A_1(t)$  determine the value of  $\langle u(t) \rangle$ , whatever the precise nature of  $A_1(t)$  may be. Moreover, even if  $\{A_1(t)\}$  is a Markov chain and therefore  $\langle u(t) \rangle$  can be exactly calculated, it may well be hard to do so explicitly (even numerically) if the dimension of  $A_1$  is large. For this reason Tuljapurkar<sup>19</sup>) also resorted to a perturbation expansion, but he did this only for the sum of the (in his case nonnegative) components of  $u(t)$  and again assumed a Markov chain  $A_1(t)$ .

The paper is organized as follows: section 2 is devoted to the derivation of the expansion in the case that the unperturbed matrix  $A_0$  is nonsingular. (Since we will repeatedly use the analogy with the expansion for a s.d.e.†, we summarize the results for that case as far as necessary in an appendix.) The reader who wants

\* By a Markov chain we mean a Markov process with discrete time parameter and discrete or continuous state space.

† Henceforth we use the abbreviations: r.d.e. = random difference equation; s.d.e. = stochastic differential equation.

to skip the details of the derivation finds a summary of the expansion to order  $\alpha^2\tau_c$  in subsection 2.5.

Section 3 considers different limiting cases. In particular we show that in a suitable limit we can replace (1.1) by a s.d.e. in the sense that the result of the cumulant expansion for the r.d.e. approaches that for the s.d.e. Also the first correction to this limit is calculated for the case in which the discreteness of time steps constitutes only a small deviation.

Section 4 illustrates the results with a simple example for which exact results are available for comparison. Finally we consider in section 5 the case that the deterministic  $A_0$  is singular, which necessitates some modifications in the results of section 2.

## 2. The cumulant expansion

### 2.1. Definition of the interaction representation

Consider the vector difference equation

$$u(t) = A(t, \omega)u(t-1) \quad (t \in I_{t_0+1}), \quad (2.1a)$$

$$u(t_0) = u_0, \quad (2.1b)$$

where  $A(t) = A_0(t) + \alpha A_1(t, \omega)$ ,  $A_0(t)$  being a deterministic matrix,  $A_1(t, \omega)$  random,  $u(t)$  an  $n$ -vector and  $\alpha$  a parameter. We conform here to the notation of Miller<sup>4)</sup> and denote by  $I = \{\dots, -1, 0, 1, \dots\}$  the set of all integers and  $I_a = \{a, a+1, a+2, \dots\}$ , where  $a \in I$ . The initial condition  $u_0$  is taken to be non-random.

The essential assumption of the present method is that  $A_1(t)$  possesses a short correlation time  $\tau_c$ , in the sense that  $A_1(t)$  and  $A_1(t')$  are statistically independent when  $|t - t'| \gg \tau_c$ . E.g. if  $A_1$  is generated by a Markov matrix  $p_{ij}$  we can define  $\tau_c$  by  $e^{-1/\tau_c} = |\lambda_1|$ , where  $\lambda_1$  is the eigenvalue of  $p_{ij}$  with largest modulus smaller than 1.

For notational convenience we first define the analog of the matrizant<sup>2)</sup> for the case of a matrix differential equation (denoted by  $Q_A$  in the case of a s.d.e.<sup>16)</sup>). That is, the unique matrix solution of\*

$$X(t) = A(t)X(t-1) \quad (t \in I_{t_0+1}),$$

$$X(t_0) = 1,$$

\*We denote the unit matrix just by 1 if no confusion is possible.

is written as

$$\mathcal{X}_A(t/t_0) = \begin{cases} \tilde{T} \prod_{s=t_0+1}^t A(s), & t \in I_{t_0+1}, \\ 1, & t = t_0. \end{cases} \quad (2.2)$$

The symbol  $\tilde{T}$  (in 2.2) stands for the time-ordering operator which orders the matrices in decreasing timeorder from left to right. The inverse of  $\mathcal{X}_A$  exists if  $A$  is nonsingular and is given by

$$\mathcal{X}_A^{-1}(t/t_0) = \begin{cases} \tilde{T} \prod_{s=t_0+1}^t A^{-1}(s), & t \in I_{t_0+1}, \\ 1, & t = t_0, \end{cases} \quad (2.3)$$

where  $\tilde{T}$  is the anti-time-ordering operator (times increasing from left to right).

With these definitions we now introduce the interaction representation (denoted by a superscript <sup>(1)</sup>) via

$$u(t) = \mathcal{X}_{A_0}(t/t_0)u^{(1)}(t) \quad (t \in I_{t_0}) \quad (2.4a)$$

and

$$A_1^{(1)}(t) = \mathcal{X}_{A_0}^{-1}(t/t_0)A_1(t)\mathcal{X}_{A_0}(t-1/t_0) \quad (t \in I_{t_0+1}) \quad (2.4b)$$

under the assumption that  $A_0(t)$  is nonsingular on  $I_{t_0+1}$ . Then one finds from (2.1) the transformed equation

$$u^{(1)}(t) = \{1 + \alpha A_1^{(1)}(t, \omega)\}u^{(1)}(t-1) \quad (t \in I_{t_0+1}), \quad (2.5a)$$

$$u^{(1)}(t_0) = u_0. \quad (2.5b)$$

## 2.2. The expansion

From now on we will only consider the equation in the interaction representation, eq. (2.5). Therefore we drop superfluous super- and subscripts and write (2.5) as

$$\Delta u(t-1) = \alpha A(t, \omega)u(t-1) \quad (t \in I_{t_0+1}), \quad (2.6a)$$

$$u(t_0) = u_0. \quad (2.6b)$$

In the following we assume that the norm of  $A$  is of order one and  $\alpha < 1$ . By the introduction of the difference operator  $\Delta$ , defined by  $\Delta f(t) = f(t+1) - f(t)$  for arbitrary  $f$ , we want to emphasize the analogy with the case of a linear s.d.e.. In fact, most of the theory for the latter case can be straightforwardly generalized

to the present situation of difference equations (a noteworthy exception being the case of a singular  $A_0$ , which poses no difficulty for differential equations). Hence we will not repeat the arguments for expanding in cumulants rather than in moments, nor consider in detail the question how to justify the expansion, but refer to the previous literature<sup>13-15</sup>). A summary of the most relevant results is given in the appendix.

The formal solution of (2.6) is

$$u(t) = \left[ \tilde{T} \prod_{s=t_0+1}^t \{1 + \alpha A(s)\} \right] u_0 \quad (t \in I_{t_0+1}), \quad (2.7)$$

from which one derives the moment expansion for the average of  $u$ :

$$\begin{aligned} \langle u(t) \rangle = & \left[ 1 + \alpha \sum_{t \geq t_1 > t_0} \langle A(t_1) \rangle + \alpha^2 \sum_{t \geq t_1 > t_2 > t_0} \langle A(t_1)A(t_2) \rangle \right. \\ & \left. + \dots + \alpha^{t-t_0} \langle A(t)A(t-1) \dots A(t_0+1) \rangle \right] u_0. \end{aligned} \quad (2.8)$$

All summations run over all discrete values between the indicated limits. Now, in contrast to the expansion for a s.d.e., the series (2.8) terminates. We can of course formally write (2.8) as ( $t \geq t_i > t_0$ ,  $i = 1, 2, \dots, l$ )

$$\langle u(t) \rangle = \left[ 1 + \sum_{l=1}^{\infty} \alpha^l \sum_{t_1 > t_2 > \dots > t_l} \langle A(t_1)A(t_2) \dots A(t_l) \rangle \right] u_0, \quad (2.9)$$

because there are no sets  $\{t_1, t_2, \dots, t_l\}$  satisfying  $t \geq t_1 > t_2 > \dots > t_l > t_0$  when  $t - t_0 < l$ . In the form (2.9) the analogy with the case of a s.d.e. is clearly displayed (compare eq. (A.3) of the appendix).

As is well known<sup>13-15</sup>), a series of this kind cannot be truncated to obtain valid approximations for large  $t$ . For this reason we use the following method†. Write (2.8) as

$$\Delta \langle u(t-1) \rangle = \left\langle \alpha A(t) \left[ \tilde{T} \prod_{s=t_0+1}^{t-1} \{1 + \alpha A(s)\} \right] \right\rangle u_0. \quad (2.10)$$

Express  $u_0$  in  $\langle u(t-1) \rangle$  by inverting (2.8) (with  $t$  replaced by  $t-1$ ), which will be possible if  $\alpha$  is small enough, and substitute the result in (2.10) to find

$$\Delta \langle u(t-1) \rangle = K(t/t_0) \langle u(t-1) \rangle, \quad (2.11a)$$

where

$$K(t/t_0) = \left\langle \alpha A(t) \left[ \tilde{T} \prod_{s=t_0+1}^{t-1} \{1 + \alpha A(s)\} \right] \right\rangle \left\langle \tilde{T} \prod_{s=t_0+1}^{t-1} \{1 + \alpha A(s)\} \right\rangle^{-1}. \quad (2.11b)$$

†For the justification of such a procedure, see ref. 13, 14.

The eqs. (2.11) constitute the central result upon which our subsequent considerations will be based.

The matrix  $K(t/t_0)$  is expanded in powers of  $\alpha$

$$K(t/t_0) = \sum_{m=1}^{\infty} K_m(t/t_0), \quad (2.12)$$

where  $K_m$  is of order  $\alpha^m$ . For example

$$K_1 = \alpha \langle A(t) \rangle, \quad (2.13a)$$

$$K_2 = \alpha^2 \sum_{t > t_1 > t_0} \{ \langle A(t)A(t_1) \rangle - \langle A(t) \rangle \langle A(t_1) \rangle \}, \quad (2.13b)$$

$$\begin{aligned} K_3 = & \alpha^3 \sum_{t > t_1 > t_2 > t_0} \{ \langle A(t)A(t_1)A(t_2) \rangle - \langle A(t) \rangle \langle A(t_1)A(t_2) \rangle \} \\ & + \alpha^3 \sum_{t > t_1 > t_0} \sum_{t > t_2 > t_0} \{ - \langle A(t)A(t_1) \rangle \langle A(t_2) \rangle + \langle A(t) \rangle \langle A(t_1) \rangle \langle A(t_2) \rangle \}. \end{aligned} \quad (2.13c)$$

The general expression for  $K_m$  can be written in terms of the averaging operator  $\mathcal{P}$ , defined by

$$\mathcal{P} \cdots = \langle \cdots \rangle. \quad (2.14a)$$

For later use we also introduce here the averaging operator  $\mathcal{Q}$  by

$$\mathcal{Q} = 1 - \mathcal{P}. \quad (2.14b)$$

Now all terms of order  $\alpha^m$  in (2.11b) can be represented by the formula

$$K_m(t/t_0) = \alpha^m \sum_{\{t_1, \dots, t_{m-1}\}}^* \langle A(t)(1 - \mathcal{P})A(t_1)(1 - \mathcal{P})A(t_2) \dots (1 - \mathcal{P})A(t_{m-1}) \rangle, \quad (2.15)$$

where the asterisk means the following: after expressing the summand of (2.15) in the moments of  $A$ , one has to sum each term over all  $t_i$ ,  $t > t_i > t_0$ , with the restriction that successive time variables within the same moment are always decreasing (see for example (2.13c)). The summand of (2.15) has previously been called a "totally time ordered cumulant" ( $t$ -cumulant for short) and is denoted as (see also the appendix)

$$\langle A(t)A(t_1) \dots A(t_{m-1}) \rangle_t.$$

To save writing we will often use the abbreviated notation  $\alpha A(t_i) \equiv i$  ( $i = 1, 2, \dots, m-1$ ),  $\alpha A(t) \equiv 0$ , so (2.15) becomes

$$K_m(t/t_0) = \sum_{\{t_1, \dots, t_m\}}^* \langle 01 \dots (m-1) \rangle_t. \quad (2.16)$$

### 2.3. The $p$ -cumulants

Now we are going to express the quantity (2.15) as a summation over the so-called "partially time-ordered cumulants"<sup>15)</sup> (abbreviated as  $p$ -cumulants). This is analogous to the case of a s.d.e. (eq. (A.6) of the appendix), but now additional cumulants emerge.

As the first step we will collect all terms in (2.15) for which all time variables  $t_1, t_2, \dots, t_{m-1}$  take on different values (i.e. no two equal). It will be shown that after a relabeling of summation variables all these terms can be written as an ordered summation over  $t_1, t_2, \dots, t_{m-1}$ , denoted as

$$K_m^{(m-1)} = \sum_{[t_1, \dots, t_{m-1}]} \dots, \quad (2.17)$$

where  $[t_1, t_2, \dots, t_{m-1}]$  stands for the ordered sequence  $t_1 > t_2 > \dots > t_{m-1}$ , and the summation extends over all integer values between  $t_0$  and  $t$  compatible with this ordering. The superscript  $^{(m-1)}$  in (2.17) indicates that all  $m-1$  summation variables are different.

One term of (2.15) which contributes to (2.17) is that for which  $t_1 > t_2 > \dots > t_{m-1}$  in the original time variables

$$\sum_{[t_1, \dots, t_{m-1}]} \langle 012 \dots (m-1) \rangle_t.$$

Secondly one might take all terms of (2.15) with  $t_2 > t_1 > t_3 > t_4 > \dots > t_{m-1}$ . This means that  $A(t_1)$  and  $A(t_2)$  necessarily belong to different moments. So by a relabeling of the summation variables  $t_1$  and  $t_2$  we can write this term as

$$\sum_{[t_1, \dots, t_{m-1}]} -\langle 02 \rangle_t \langle 134 \dots (m-1) \rangle_t.$$

In fact, in the case of a s.d.e.  $K_m$  is also of the form (2.15) with all summations replaced by integrations, but with the same restrictions on the integration variables that are indicated by the asterisk. However, in the latter case there are no separate terms for equal integration variables\* (see ref. 15, subsection 3.3). So we can immediately conclude that all terms in (2.15) with distinct time variables  $t_1, t_2, \dots, t_{m-1}$  can be written in the form (A.6) of the appendix with integrations replaced by summations

$$K_m^{(m-1)} = \sum_{[t_1, \dots, t_{m-1}]} \langle 012 \dots (m-1) \rangle_p. \quad (2.18)$$

The symbol  $\langle \dots \rangle_p$  in (2.18) denotes the  $p$ -cumulant of  $m$  matrices  $A$ . It is a sum of  $(m-1)!$  terms (one term for each permutation of  $1, 2, \dots, m-1$ ), each of

\* These are included as upper limits of the integration domains, see (A.6).



which is a product of  $t$ -cumulants. For the precise definition we refer to the appendix. The most important feature of this decomposition in  $t$ -cumulants is that two successive matrices  $A(t_i)$  and  $A(t_j)$  occurring in a permutation belong to the same  $t$ -cumulant if  $t_i > t_j$  (so  $i < j$ ), but to different  $t$ -cumulants if  $t_i < t_j$  ( $i > j$ ) (the reason is again that these matrices  $A(t_i)$  and  $A(t_j)$  originally (i.e. in (2.15)) must have belonged to the same or to different moments, respectively).

Now we are left with the terms in (2.15) where two or more time variables are equal. Again these terms can be written as ordered summations over  $m - 2$ ,  $m - 3, \dots$  distinct time variables. The method is the same as above, the only difference being that successive matrices  $A(t_i)$  and  $A(t_j)$  belong to the same, resp. different  $t$ -cumulants if  $t_i > t_j$  ( $i < j$ ), resp.  $t_i \leq t_j$  ( $i \geq j$ ) (matrices with equal time variables can never have belonged to the same moment). For example all terms in (2.15) with  $t_1 = t_2 > t_3 > \dots > t_{m-1}$  can after relabeling of time variables be written as

$$\sum_{[t_1, \dots, t_{m-2}]} -\langle 01 \rangle_t \langle 12 \dots (m-2) \rangle_t \quad (2.19)$$

or all terms with  $t_2 = t_4 > t_3 > t_1 > t_5 > t_6 > \dots > t_{m-1}$  yield

$$\sum_{[t_1, \dots, t_{m-2}]} \langle 03 \rangle_t \langle 12 \rangle_t \langle 145 \dots (m-2) \rangle_t. \quad (2.20)$$

In this way one can collect all terms in (2.15) with two equal time variables which are larger than the remaining ones and we write them as

$$\sum_{[t_1, \dots, t_{m-2}]} \langle 01123 \dots (m-2) \rangle_p. \quad (2.21)$$

The summand in (2.21) is a  $p$ -cumulant with equal numerals (corresponding to matrices with equal time variables). Its decomposition into  $\frac{1}{2}(m-1)!$  products of  $t$ -cumulants (of which the summands of (2.19) and (2.20) are two examples) is given in the appendix. Of course there are also terms with two equal summation variables where of the  $(m-3)$  other ones there is one larger and  $(m-4)$  smaller than these two. They can be written as

$$\sum_{[t_1, \dots, t_{m-2}]} \langle 012234 \dots (m-2) \rangle_p. \quad (2.22)$$

In this way one obtains all terms in (2.15) with precisely  $m - 2$  distinct time variables in the set  $\{t_1, t_2, \dots, t_{m-1}\}$ . The collection of all these terms is denoted by  $K_m^{(m-2)}$ .

In general we denote the collection of all terms with  $i$  distinct time variables out of the set  $\{t_1, t_2, \dots, t_{m-1}\}$  by  $K_m^{(i)}$ ,  $i = 1, 2, \dots, m - 1$ . Since there can at most be

$t - t_0 - 1$  different integers between  $t_0$  and  $t$  we have that  $K_m^{(i)} = 0$  for  $i \geq t - t_0$  if  $m > t - t_0$ .

Summarizing the above considerations we get the following rules for obtaining  $K_m(t/t_0)$  in (2.12):

- i) Choose a number  $i \in \{1, 2, \dots, m-1\}$ .
- ii) Choose  $n_1$  times a number 1,  $n_2$  times a 2,  $\dots$ ,  $n_i$  times a  $i$ , so that  $n_1 + n_2 + \dots + n_i = m-1$ ;  $n_k = 1, 2, \dots$  (this can be achieved in  $\binom{m-1}{i}$  possible ways); call the sets of numbers thus obtained  $S_l(i, m-1)$ ,  $l = 1, 2, \dots, \binom{m-1}{i-1}$ .
- iii) For each  $l$ , write a sequence of  $m$  dots. Write a zero on the first dot and place the numerals in  $S_l$  on the remaining dots in nondecreasing order.
- iv) Put brackets  $\langle \dots \rangle_p$  around each sequence of numerals.
- v) Replace each numeral  $k$  by  $\alpha A(t_k)$  (and 0 by  $\alpha A(t)$ ) and sum the resulting expression over  $t_1, t_2, \dots, t_i$  where each time variable runs through all integer values between  $t_0$  and  $t$  under the restriction  $t_1 > t_2 > \dots > t_i$ .
- vi) Sum all contributions from the sets  $S_l(i, m-1)$  for different  $l$ . This yields  $K_m^{(i)}$ .
- vii) Finally sum over all values of  $i$  to obtain

$$K_m(t/t_0) = \sum_{i=1}^{m-1} K_m^{(i)}(t/t_0).$$

Examples (all primed summations  $\Sigma'$  are between  $t_0$  and  $t$ ):

$$K_1 \equiv K_1^{(0)} = \alpha \langle 0 \rangle_p,$$

$$K_2 = K_2^{(1)} = \alpha^2 \sum'_{t_1} \langle 01 \rangle_p,$$

$$K_3^{(2)} = \alpha^3 \sum'_{t_1 > t_2} \langle 012 \rangle_p; K_3^{(1)} = \alpha^3 \sum'_{t_1} \langle 011 \rangle_p, \quad (2.23)$$

$$K_4^{(3)} = \alpha^4 \sum'_{t_1 > t_2 > t_3} \langle 0123 \rangle_p; K_4^{(2)} = \alpha^4 \sum'_{t_1 > t_2} \{ \langle 0112 \rangle_p + \langle 0122 \rangle_p \},$$

$$K_4^{(1)} = \sum'_{t_1} \langle 0111 \rangle_p.$$

A general definition of the p-cumulant is given in the appendix. The lowest order ones are

$$\langle 0 \rangle_p = \langle 0 \rangle; \langle 01 \rangle_p = \langle 01 \rangle - \langle 0 \rangle \langle 1 \rangle = \langle\langle 01 \rangle\rangle.$$

With double brackets  $\langle\langle \dots \rangle\rangle$  we denote ordinary cumulants<sup>14</sup>). If the matrices  $A(t)$  and  $A(t')$  commute for all  $t$  and  $t'$  the p-cumulants can be expressed in ordinary cumulants<sup>13</sup>).

Finally we want to give a very formal expression for  $K(t/t_0)$  from which one can obtain all the terms  $K_m$  by expansion with respect to  $\alpha$ . Let us first formally sum all terms  $K_m^{(1)}$

$$\begin{aligned} \sum_{m=2}^{\infty} K_m^{(1)} &= \sum_{t_1}' \langle 01 \rangle_p + \langle 011 \rangle_p + \langle 0111 \rangle_p + \dots \\ &= \alpha^2 \sum_{t_1}' \langle A(t):A(t_1)(1 - \alpha A(t_1))^{-1}: \rangle_p. \end{aligned} \quad (2.24)$$

The colons in (2.24) indicate that to evaluate the p-cumulant one should first expand the expression between the colons in powers of  $\alpha$  and subsequently take the p-cumulant of each term. In the same way one finds

$$\sum_{m=3}^{\infty} K_m^{(2)} = \sum_{t_1 > t_2}' \alpha^3 \langle A(t):\hat{A}(t_1)\hat{A}(t_2): \rangle_p, \quad (2.25)$$

where  $\alpha \hat{A}(t) = \alpha A(t)(1 - \alpha A(t))^{-1}$ . Continuing in this way we arrive at

$$\begin{aligned} K(t/t_0) &= \alpha \langle A(t) \rangle + \sum_{m=1}^{t-t_0-1} \alpha^{m+1} \sum_{t_1 > t_2 > \dots > t_m}' \langle A(t):\hat{A}(t_1)\hat{A}(t_2)\dots\hat{A}(t_m): \rangle_p \\ &= \left\langle \alpha A(t): \left[ \tilde{T} \prod_{s=t_0+1}^{t-1} (1 + \alpha \hat{A}(s)) \right]: \right\rangle_p. \end{aligned} \quad (2.26)$$

This compact but formal representation of  $K$  will be useful when studying the correlation functions of  $u(t)$ . This will be taken up in part II.

#### 2.4. Estimates for large time

It is known that the p-cumulants have the cluster property<sup>14,15</sup>), i.e. if the moments  $\langle A(t)A(t_1)\dots A(t_{m-1}) \rangle$  factorize, the corresponding p-cumulant  $\langle A(t)A(t_1)\dots A(t_{m-1}) \rangle_p$  vanishes. Since we assume that the stochastic matrix-sequence  $\{A(t)\}$  has a finite correlation "time"  $\tau_c$ , this will be the case if there is a gap between two successive times  $t_i$  and  $t_{i+1}$  large compared to  $\tau_c$ . If two or more time variables in a p-cumulant are equal, the same is true if there is a gap of order  $\tau_c$  between two successive distinct time variables (this can be shown via the decomposition in  $t$ -cumulants along the same lines as in ref. 15, appendix A).

As a consequence the summations in the cumulant functions  $K_m(t/t_0)$  virtually extend only over intervals of order  $\tau_c$ , and in particular we may replace the initial

time  $t_0$  in  $K(t/t_0)$  by  $-\infty$  if  $t - t_0 \gg \tau_c$ . This enables us to deduce the following estimates for the successive  $K_m^{(i)}$ 's (assuming that  $\|A(t)\|$  is of order one\*), which are valid after a transient time much larger than  $\tau_c$  (compare the discussion for a s.d.e. in ref. 15)

$$\begin{aligned} K_1 &\sim \alpha, \\ K_2^{(1)} &\sim \alpha(\alpha\tau_c), \\ K_3^{(2)} &\sim \alpha(\alpha\tau_c)^2, K_3^{(1)} \sim \alpha^2(\alpha\tau_c), \\ K_4^{(3)} &\sim \alpha(\alpha\tau_c)^3, K_4^{(2)} \sim \alpha^2(\alpha\tau_c)^2, K_4^{(1)} \sim \alpha^3(\alpha\tau_c), \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \end{aligned} \quad (2.27)$$

And generally

$$K_m^{(i)} \sim \alpha^{m-i}(\alpha\tau_c)^i. \quad (2.28)$$

These estimates can be sharpened in many cases. To give a concrete example, let us assume that

$$\|\langle A(t)A(t_1) \dots A(t_m) \rangle_p\| \leq C_m |\rho|^{t-t_m} \quad (|\rho| < 1), \quad (2.29)$$

where  $C_m$  is a constant†. This condition is met in several important cases<sup>20</sup>. Then we have the estimate

$$K_m^{(i)} \sim \alpha^{m-1} \gamma^i, \quad \gamma = \alpha \frac{|\rho|}{1-|\rho|}. \quad (2.30)$$

If we now define  $\tau_c$  by  $|\rho| = e^{-1/\tau_c}$ , then  $\gamma \leq \alpha\tau_c$ .

The most important result of this section is that after a transient time of order  $\tau_c$ , the average of the solution  $u(t)$  of (2.6) obeys itself a first order difference equation with coefficients which are independent of the initial time  $t_0$ . If  $\{A(t)\}$  is a stationary sequence the cumulant functions  $K_m$  are independent of  $t_0$  and  $t$  for  $t - t_0 \gg \tau_c$ , e.g.

$$\alpha^{-2} K_2 = \sum_{t_1=t_0+1}^{t-1} \langle\langle A(t)A(t_1) \rangle\rangle \rightarrow \sum_{\tau=1}^{\infty} C(\tau), \quad (2.31)$$

where  $C(\tau) = \langle\langle A(t)A(t-\tau) \rangle\rangle$  is the autocorrelation matrix (also called covariance matrix) of  $A(t)$ .

\* $\|\dots\|$  denotes a matrix norm.

†The dependence of the constants  $C_m$  on  $m$  determines whether the expansion (2.12) is convergent or asymptotic.

*Comment*

The above considerations apply to the eq. (2.5) in the interaction representation. In particular, the autocorrelation time  $\tau_c$  is that of  $A_1^{(i)}(t)$  and is not the same as that of  $A_1(t)$ , but is modified by the deterministic matrix  $A_0$ . (See also ref. 15, appendix B.) An example will be given in part II, see also Tuljapurkar<sup>19</sup>).

2.5. Result to second order in  $\alpha$

We have found that if both  $\alpha \ll 1$  and  $\alpha\tau_c \ll 1$ , the average of the solution  $u^{(i)}(t)$  of (2.5) obeys the first order difference equation

$$\langle u^{(i)}(t) \rangle = \left[ 1 + \alpha \langle A_1^{(i)}(t) \rangle + \alpha^2 \sum_{t=t_0+1}^{t-1} \langle\langle A_1^{(i)}(t) A_1^{(i)}(t_1) \rangle\rangle \right] \langle u^{(i)}(t-1) \rangle, \quad (2.32)$$

where  $\langle\langle \dots \rangle\rangle$  denotes an ordinary second order cumulant. In the original representation (2.1) we have in the case of a time independent deterministic matrix  $A_0$  and after a transient period of order  $\tau_c$

$$\begin{aligned} \langle u(t) \rangle = & \left[ A_0 + \alpha \langle A_1(t) \rangle + \alpha^2 \sum_{\tau=1}^{\infty} \right. \\ & \times \langle\langle A_1(t) [A_0^{-1} A_1(t-\tau)] \rangle\rangle A_0^{-\tau} \left. \right] \langle u(t-1) \rangle. \end{aligned} \quad (2.33)$$

For many practical applications this approximation will suffice. Note however, that even if  $A_1(t)$  is a scalar Gaussian process, the approximation (2.33) is not exact (in contrast with the case of a s.d.e.). The reason is that although in this case  $K_m^{(m-1)} = 0$  for  $m > 2$ , the second order cumulant of  $A_1$  also occurs in the cumulant functions  $K_m^{(i)}$ ,  $i < m-1$  for  $m > 2$ .

3. Limiting cases

3.1. Comparison of time scales

In the case of a s.d.e. we had only one dimensionless expansion parameter, viz.  $\alpha\tau_c$ <sup>13-16</sup>). To obtain a valid expansion we had to assume that  $\alpha\tau_c \ll 1$ . In the present situation the "time" variables are in fact just integers, so we have two dimensionless expansion parameters:  $\alpha$  and  $\tau_c$  (or alternatively  $\alpha$  and  $\alpha\tau_c$ ). Accordingly we may distinguish between different limiting cases. From (2.27) we deduce that at any rate both  $\alpha$  and  $\alpha\tau_c$  have to be small.

a)  $\tau_c = 0$  (no correlations)

Only  $K_1$  contributes.  $\{A(t)\}$  is a sequence of independently distributed random matrices, i.e. there are no correlations at all. In this case the solution of (2.6) constitutes a Markov chain.

b)  $\tau_c \simeq 1$  (*correlations last only a few time steps*)

All  $K_m^{(i)}$ ,  $i = 1, 2, \dots, m-1$  are of order  $\alpha^m$ , so in this case one can just expand in  $\alpha$  itself.

c)  $\tau_c \gg 1$  (*correlations over many time steps*)

Since  $\alpha\tau_c$  must be smaller than one, we also require that  $\alpha \ll 1$ . This implies that the evolution of  $u(t)$  is very slow on the original time scale (see (2.6)), but the correlations are relatively long-lived. Now the most important contribution to  $K_m$  is  $K_m^{(m-1)}$ . But the terms  $K_m^{(m-1)}$  are precisely those which are present in the case of a s.d.e. (with summations replaced by integrations). And indeed we show below for a one dimensional example that in a suitable limit we can replace (2.6) by a s.d.e., in the sense that the results of the cumulant expansions for both cases become identical in the limit. The effects of the autocorrelation are retained in the limiting procedure, so the resulting s.d.e. is in general not a diffusion equation<sup>\*26</sup>).

In section 3.3 we then calculate the first correction to this limit, which describes in lowest order the effect of the discreteness of time steps.

### 3.2. The limit $\alpha \rightarrow 0$ , $\tau_c \rightarrow \infty$

In this subsection we investigate whether we can replace the r.d.e. (2.6) by a s.d.e. in the limit in which the change of  $u(t)$  per time step becomes smaller and smaller, i.e. when  $\alpha \rightarrow 0$ . To have still a nonvanishing effect in this limit, we also have to rescale the independent variable  $t$  in such a way that one step on the new time scale corresponds to many steps on the old one. Moreover we also want to retain the effect of the finite autocorrelation time  $\tau_c$  of the noise  $A(t)$  in (2.6). Accordingly we let  $\tau_c$  grow indefinitely at the same rate with which the time  $t$  is rescaled as  $\alpha \rightarrow 0$ .

To see what is involved let us consider the following one dimensional case

$$\begin{aligned} \Delta u(t) &= \alpha \xi(t+1)u(t), \\ u(t_0) &= u_0, \end{aligned} \quad (3.1)$$

where  $\{\xi(t)\}_{t=-\infty}^{\infty}$  is a stationary random sequence with auto-correlation time  $\tau_c$ . Then we know from the results of section 2 that, if  $\alpha\tau_c \ll 1$ , the average of  $u$  obeys

$$\Delta \langle u(t) \rangle = [\alpha K^{(1)}(t+1/t_0) + \alpha^2 K^{(2)}(t+1/t_0) + \dots] \langle u(t) \rangle, \quad (3.2)$$

where

$$\alpha K^{(1)} = \sum_{m=1}^{\infty} K_m^{(m-1)} \quad (3.3a)$$

and in general

$$\alpha^i K^{(i)} = \sum_{m=i+1}^{\infty} K_m^{(m-i)}, \quad i = 2, 3, \dots \quad (3.3b)$$

\*For a direct diffusion approximation to stochastic difference equations, see refs. 21 and 22.

The order of magnitude of  $K_m^{(j)}$  is given by  $(t - t_0 \gg \tau_c)$

$$K_m^{(j)} \sim \alpha^{m-j} (\alpha \tau_c)^j, \quad (3.4)$$

so after a transient period all functions  $K^{(1)}, K^{(2)} \dots$  are expansions in powers of  $\alpha \tau_c$  (i.e. the parameter  $\alpha$  does not appear separately). From this one sees that as  $\alpha$  becomes small, while  $\alpha \tau_c$  stays constant (so  $\tau_c$  becomes large), the most important contribution in (3.2) comes from  $\alpha K^{(1)}$ . Its terms of order  $\alpha$  and  $\alpha^2$  are

$$K_1^{(0)}(t + 1/t_0) = \alpha \langle \xi \rangle, \quad (3.5a)$$

$$K_2^{(1)}(t + 1/t_0) = \alpha^2 \sum_{s=1}^{t-t_0} C(s; \tau_c), \quad (3.5b)$$

where  $\langle \xi \rangle$  is the constant average of  $\xi(t)$  and  $C(s; \tau_c)$  is the autocorrelation function

$$C(s; \tau_c) = \langle \xi(t) \xi(t-s) \rangle. \quad (3.6)$$

We explicitly indicate the dependence of  $C$  on the autocorrelation time  $\tau_c$ . In (3.5b) only the values of  $C$  at discrete time points occur, but we assume that  $C$  is given by an expression which is defined for all real  $s$ , that  $C$  is Riemann-integrable and  $\int_0^\infty ds |C(s)| < \infty$ . To be more specific we assume that  $C$  depends only on the quotient  $s/\tau_c$ , so that

$$C(\epsilon s; \epsilon \tau_c) = C(s; \tau_c), \quad \forall \epsilon > 0. \quad (3.7)$$

In particular this implies that

$$\left| \int_0^\infty ds C(s; \tau_c) \right| \rightarrow \infty, \quad \text{as } \tau_c \rightarrow \infty \quad (3.8)$$

at least if the integral in (3.8) is nonzero for some finite value of  $\tau_c$ . And the condition (3.8) is needed if we want to obtain a nonzero value for  $K_2^{(1)}/\alpha$  in (3.5b) as  $\alpha \rightarrow 0$ .

Now consider the limit

$$\alpha \rightarrow 0, \quad \alpha \tau_c = \text{constant} = \gamma, \quad (3.9)$$

where  $\gamma$  denotes the value of the constant. Also define a new time scale  $\tau$  and a new function  $v(\tau)$  by

$$t = [\tau/\alpha]; \quad v(\tau) = u([\tau/\alpha]), \quad (3.10)$$

where  $\tau$  can now be any real number, and  $[x]$  denotes the entier of  $x$ . In this limit  $\alpha^{-1} K_2^{(1)}$  will approach a limiting value of order  $\gamma$ , and we expect that also the higher contributions  $\alpha^{-1} K_m^{(m-1)}$ ,  $m = 3, 4, \dots$  to  $K^{(1)}$  approach a value of order  $\gamma^{m-1}$ . But  $K^{(2)}, K^{(3)} \dots$  will not occur because they are multiplied by additional factors  $\alpha$ .

So let us consider (3.2) to  $\mathcal{O}(\alpha^2)$  (take  $t_0 = 0$  for convenience)

$$\Delta \langle u(t) \rangle = \left[ \alpha \langle \xi \rangle + \alpha^2 \sum_{s=1}^t C(s; \tau_c) \right] \langle u(t) \rangle. \quad (3.11)$$

In terms of  $v(\tau)$  this yields by using (3.9)

$$\frac{\langle v(\tau + \alpha) \rangle - \langle v(\tau) \rangle}{\alpha} = \left[ \langle \xi \rangle + \alpha \sum_{s=1}^{[\tau/\alpha]} C\left(s; \frac{\gamma}{\alpha}\right) \right] \langle v(\tau) \rangle. \quad (3.12)$$

Then take the limit  $\alpha \rightarrow 0$ . For the expression involving  $C$  we find

$$\alpha \sum_{s=1}^{[\tau/\alpha]} C\left(s; \frac{\gamma}{\alpha}\right) = \alpha \sum_{s=1}^{[\tau/\alpha]} C(\alpha s; \gamma) \rightarrow \int_0^\tau d\tau' C(\tau'; \gamma), \quad \text{as } \alpha \rightarrow 0, \quad (3.13)$$

where we used (3.7) and the integrability of  $C$  in order to replace the Riemann sum by the integral. Thus (3.11) becomes

$$\partial_\tau \langle v(\tau) \rangle = \left[ \langle \xi \rangle + \int_0^\tau d\tau' C(\tau'; \gamma) \right] \langle v(\tau) \rangle. \quad (3.14)$$

Now we ask the question whether we can reproduce the result (3.14) by starting directly from a stochastic differential equation instead of (3.1). This will indeed be possible if the function  $C(\tau'; \gamma)$  in (3.14) is itself again a correlation function of some continuous time stationary random process, which is the case if  $C(\tau'; \gamma)$  is a nonnegative-definite function of  $\tau'$  on  $\mathbb{R}^{23}$ ). Under this assumption there exists\* a random process  $\bar{\xi}(t)$  with

$$\langle \bar{\xi}(t) \rangle = \langle \xi \rangle; \quad \langle \bar{\xi}(t) \bar{\xi}(t') \rangle = C(|t - t'|; \tau_c), \quad (3.15)$$

where  $\tau_c$  is now the correlation time of the process  $\bar{\xi}$ . So if one starts with the s.d.e.

$$\begin{aligned} \partial_t u(t) &= \alpha \bar{\xi}(t) u(t), \\ u(0) &= u_0 \end{aligned} \quad (3.16)$$

and applies the cumulant expansion (A.4)–(A.6) to it, using (3.15), one finds back the eq. (3.14) by putting  $t = (\tau/\alpha)$ ,  $u(t) = v(\tau)$  and  $\tau_c = (\gamma/\alpha)$  in the result. So our conclusion is that in the limit (3.9) combined with an appropriate time-rescaling, the r.d.e. (3.1) and the s.d.e. (3.16) are equivalent in the sense that their respective cumulant expansions give the same answer.

\*E.g. one could take the Gaussian process  $\bar{\xi}$  satisfying (3.15).



### Example

Consider (3.1) where  $\xi(t)$  is a stationary two-state Markov chain (values  $\pm 1$ ) with transition matrix

$$T = \begin{pmatrix} 1-v & v \\ v & 1-v \end{pmatrix}, \quad 0 < v < 1. \quad (3.17)$$

This implies for every sequence  $t_1 > t_2 > \dots > t_m$

$$\langle \xi(t_1)\xi(t_2) \dots \xi(t_m) \rangle = \begin{cases} 0, & m \text{ odd}, \\ \rho^{t_1-t_2} \rho^{t_3-t_4} \dots \rho^{t_{m-1}-t_m}, & m \text{ even}, \end{cases} \quad (3.18)$$

where

$$\rho = 1 - 2v, \quad -1 < \rho < 1. \quad (3.19)$$

With the help of (3.18) we can calculate all cumulant functions  $K_m$  of (2.12), for example

$$K_2(t/-\infty) = \alpha^2 \sum_{s=1}^{\infty} \rho^s = \alpha \delta \quad (3.20)$$

with

$$\delta = \frac{\alpha \rho}{1 - \rho}. \quad (3.21)$$

If we define  $\tau_c$  by

$$|\rho| = e^{-1/\tau_c}, \quad (3.22)$$

we see that the limit  $\tau_c \rightarrow \infty$  corresponds to  $|\rho| \rightarrow 1$ . Now as  $\rho \rightarrow 1$  it is true that  $\delta/\alpha \rightarrow \infty$  in accordance with (3.8), but if  $\rho \rightarrow -1$  we find that  $\delta/\alpha \rightarrow -\frac{1}{2}$ . And indeed, (3.7) is satisfied if  $0 < \rho < 1$ , but not if  $-1 < \rho < 0$ . So in the latter case the general estimate (2.28) gives no indication at all of the real behaviour of the  $K$ -functions as  $\tau_c \rightarrow \infty$ . The underlying reason is of course the increasingly rapid oscillations of the correlation function of  $\xi$  as  $\rho \rightarrow -1$  ( $\alpha \rightarrow 0$ ):

$$\alpha \sum_{s=1}^{[\tau/\alpha]} C\left(s, \frac{\gamma}{\alpha}\right) = \alpha \sum_{s=1}^{[\tau/\alpha]} (-)^s e^{-\alpha s/\gamma}. \quad (3.23)$$

In other words, if  $\rho \rightarrow 1$ , we approximate the area under the curve  $e^{-\tau/\gamma}$  by finer and finer partitions of the interval  $[0, \tau]$  (fig. 1), while as  $\rho \rightarrow -1$  the function  $C(s, (\gamma/\alpha))$  (which is now only defined for integer  $s$ ) becomes more and more oscillating as the partitions become finer, due to the factor  $(-)^s$  in (3.23) (fig. 2), and as  $\alpha \rightarrow 0$  the resulting sum in (3.23) (i.e. the shaded area in fig. 2) goes to zero.

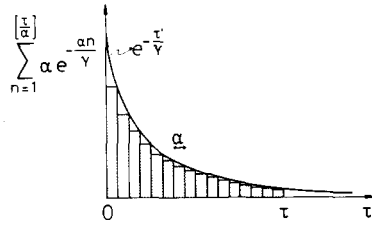


Fig. 1. Approximation of the sum (3.13) by an integral.

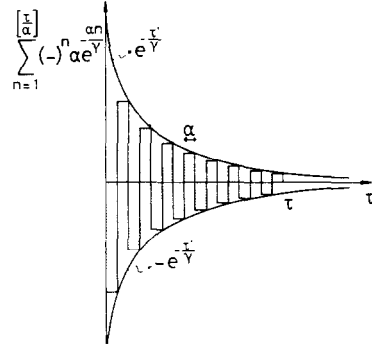


Fig. 2. Approximation of the sum (3.23) by an integral.

So we consider the limit (3.9) corresponding to  $\rho \rightarrow 1$ . Then the parameter  $\delta$  in (3.21) approaches

$$\delta = \frac{\alpha e^{-1/\tau_c}}{1 - e^{-1/\tau_c}} \rightarrow \alpha \tau_c = \gamma \quad (3.24)$$

and  $\bar{\xi}(t)$  is now a process with zero mean and covariance

$$\langle\langle \bar{\xi}(t) \bar{\xi}(t') \rangle\rangle = e^{-|t-t'|/\tau_c}. \quad (3.25)$$

It is not difficult to show that in the limit (3.9) also the higher corrections to (3.14) (i.e. the remaining contributions to  $K^{(1)}$ ) approach those in the expansion\* for (3.16), if one takes for  $\bar{\xi}(t)$  the stationary dichotomic Markov process<sup>14)</sup> (values  $\pm 1$ ) with generator  $W$  and transition matrix  $T_\tau$  given by

$$W = \begin{pmatrix} -v' & v' \\ v' & -v' \end{pmatrix}; \quad T_\tau = e^{\tau W} = \frac{1}{2} \begin{pmatrix} 1 + e^{-2v'\tau} & 1 - e^{-2v'\tau} \\ 1 - e^{-2v'\tau} & 1 + e^{-2v'\tau} \end{pmatrix}. \quad (3.26)$$

where  $v'$  is related to the parameter  $v$  in (3.17) by

$$v = \frac{1}{2}(1 - e^{-2v'}) \quad (3.27)$$

\* That means use (A.6) or alternatively the terms  $K_m^{(m-1)}$  of (2.23) with summations replaced by integrations.

and  $\tau_c$  in (3.25) is given by

$$\tau_c = (2\nu')^{-1}. \quad (3.28)$$

This process  $\bar{\xi}(t)$  is the continuous time embedding<sup>24</sup> of the discrete process  $\xi(t)$  generated by (3.17). This embedding is only possible if  $0 < \nu < \frac{1}{2}$ , i.e. if  $0 < \rho < 1$ , which again explains why we had to take  $\rho \rightarrow 1$  and not  $\rho \rightarrow -1$ .

It seems plausible that these results can be extended to the case of a vector equation (2.6), as long as only a single autocorrelation time is involved, so that one can make an assumption like (3.7). But if different correlation times occur, there may be different ways to obtain limiting equations, depending on which  $\tau_c$  one chooses in (3.9). Therefore we will not discuss this point further here<sup>26</sup>.

### 3.3. First correction to eq. (3.14)

Now we want to calculate the first correction to the result (3.14) by taking the discreteness of time steps into account in lowest order in  $\alpha$ . We illustrate the method again by the example of the previous subsection.

Again we have to specify the way in which  $\tau_c$  depends on  $\alpha$  as  $\alpha$  becomes small. It turns out to be convenient if instead of (3.9) we take the combination (3.21) to be constant, i.e.

$$\delta = \frac{\alpha e^{-1/\tau_c}}{1 - e^{-1/\tau_c}} = \text{constant} \equiv \gamma', \quad (3.29)$$

where  $\gamma'$  denotes the value of the constant. As  $\alpha \rightarrow 0$ ,  $\tau_c$  varies with  $\alpha$  as  $(\gamma'/\alpha)$ , so in zeroth order we will find back the result (3.14) if we choose equal numerical values for the constants  $\gamma$  in (3.9) and  $\gamma'$  in (3.29). Now the matrices  $K_m^{(i)}$  in (3.3) are of order

$$K_m^{(i)} \sim \alpha^{m-i} (\delta)^i \quad (t - t_0 \gg \tau_c). \quad (3.30)$$

This estimate is valid after a transient time. This implies that the matrices  $K^{(1)}$ ,  $K^{(2)}$ , ... in (3.2) depend on  $\delta$  only (for small times they depend on  $\delta$  and  $\alpha$  separately). The same holds on the new time scale  $\tau$  (defined in (3.10)): if we take  $\tau - \tau_0 \gg \delta$  in the final result, the functions  $K^{(i)}$  depend only on  $\delta$  (as we will show at the end of this subsection). So for the rest of the derivation we assume that the  $K^{(i)}$  are equal to their asymptotic values for  $t_0 \rightarrow -\infty$ . Introduce  $v(\tau)$  by (3.10) and put

$$\langle v(\tau + \alpha) \rangle = \langle v(\tau) \rangle + \alpha \langle \dot{v}(\tau) \rangle + \frac{1}{2} \alpha^2 \langle \ddot{v}(\tau) \rangle + \mathcal{O}(\alpha^3). \quad (3.31)$$

Inserting this in (3.2) we get

$$\begin{aligned} \frac{\langle v(\tau + \alpha) \rangle - \langle v(\tau) \rangle}{\alpha} &= \langle \dot{v}(\tau) \rangle + \frac{1}{2}\alpha \langle \ddot{v}(\tau) \rangle + \mathcal{O}(\alpha^2) \\ &= [K^{(1)} + \alpha K^{(2)} + \mathcal{O}(\alpha^2)] \langle v(\tau) \rangle. \end{aligned} \quad (3.32)$$

Now also expand  $\langle v(\tau) \rangle$  itself in powers of  $\alpha$

$$\langle v(\tau) \rangle = \langle v^{(0)}(\tau) \rangle + \alpha \langle v^{(1)}(\tau) \rangle + \mathcal{O}(\alpha^2). \quad (3.33)$$

Substitute this in (3.32) and collect terms with the same power of  $\alpha$

$$\mathcal{O}(\alpha^0): \quad \langle \dot{v}^{(0)}(\tau) \rangle = K^{(1)} \langle v^{(0)}(\tau) \rangle. \quad (3.34)$$

This is our previous result (subsection 3.2).

$$\mathcal{O}(\alpha^1): \quad \langle \dot{v}^{(1)}(\tau) \rangle + \frac{1}{2} \langle \ddot{v}^{(0)}(\tau) \rangle = K^{(1)} \langle v^{(1)}(\tau) \rangle + K^{(2)} \langle v^{(0)}(\tau) \rangle.$$

Using (3.34) we obtain

$$\langle \dot{v}^{(1)}(\tau) \rangle = K^{(1)} \langle v^{(1)}(\tau) \rangle + \{K^{(2)} - \frac{1}{2}K^{2(1)}\} \langle v^{(0)}(\tau) \rangle. \quad (3.35)$$

Combining (3.34) and (3.35) we find for  $\langle v(\tau) \rangle$  itself

$$\langle \dot{v}(\tau) \rangle = K^{(1)} \langle v(\tau) \rangle + \alpha \{K^{(2)} - \frac{1}{2}K^{2(1)}\} \langle v(\tau) \rangle + \mathcal{O}(\alpha^2), \quad (3.36)$$

where a term of  $\mathcal{O}(\alpha^2)$  was included in the second term on the r.h.s. of (3.36) to get a closed expression in  $\langle v(\tau) \rangle$ . Of course one can obtain higher corrections by taking  $K^{(3)}$ ,  $K^{(4)}$ , ... into account, but we will not go into this here.

If  $\gamma'$  is small, so that we can neglect terms of order  $(\gamma')^2$  and higher, the result (3.36) for the example of the previous subsection yields for  $K^{(1)}$  and  $K^{(2)}$

$$\begin{aligned} K^{(1)} &= \langle \xi(t) \rangle + \alpha \sum_{s=1}^{\infty} \langle \langle \xi(t) \xi(t-s) \rangle \rangle = \gamma', \\ K^{(2)} &= -\alpha \sum_{s=1}^{\infty} \langle \langle \xi(t) \xi(t-s) \rangle \rangle \langle \xi(t-s) \rangle = 0. \end{aligned}$$

This we can again reproduce by the introduction of the continuous time Markov process  $\bar{\xi}(t)$  as follows

$$K^{(1)} = \langle \bar{\xi}(t) \rangle + \alpha \int_0^{\infty} d\tau_1 \langle \langle \bar{\xi}(t) \bar{\xi}(t - \tau_1) \rangle \rangle, \quad (3.37)$$

$$K^{(2)} = -\alpha \int_0^{\infty} d\tau_1 \langle \langle \bar{\xi}(t) \bar{\xi}(t - \tau_1) \rangle \rangle \langle \bar{\xi}(t - \tau_1) \rangle, \quad (3.38)$$

provided we make the identification

$$\bar{\tau}_c = (\gamma' / \alpha), \quad (3.39)$$

where  $\bar{\tau}_c$  is the correlation time of  $\bar{\xi}(t)$ . The lowest order result (3.34) coincides with (3.14) if we choose the same numerical values for  $\gamma$  in (3.9) and  $\gamma'$  in (3.29).

It remains to be shown that the matrices  $K^{(i)}$  contain no terms of order  $\alpha$  if  $\tau - \tau_0 \gg \gamma'$ . Now (take  $\tau_0 = 0$ )

$$K_2^{(1)} = \alpha^2 \sum_{\tau_1=1}^{\tau} \rho^{\tau_1} = \alpha \gamma' (1 - \rho^{[\tau/\alpha]}) = \alpha \gamma' \left[ 1 - e^{-\tau/\gamma'} \left\{ 1 + \frac{1}{2} \frac{\alpha \tau}{\gamma'^2} \right\} + \mathcal{O}(\alpha^2) \right].$$

So if  $\tau \gg \gamma'$ ,  $K_2^{(1)} = \alpha \gamma' + \mathcal{O}(\alpha^3)$ . Higher orders are checked similarly, so  $K^{(1)} = \bar{K}^{(1)}(\gamma') + \mathcal{O}(\alpha^2)$ . We don't have to do it for  $K^{(2)}$  because (3.32) shows that only the part of order  $\alpha^0$  contributes to (3.35).

#### 4. Example

In this section we illustrate the results of section 2 by a simple example which has an exact solution for comparison.

We start with a definition<sup>20)</sup>:

##### Definition

An  $m$ -dependent stochastic sequence  $\{\xi_n\}$  is a sequence of stochastic variables which satisfies

$$\langle \xi_i \dots \xi_k \xi_{k+n} \dots \xi_j \rangle = \langle \xi_i \dots \xi_k \rangle \langle \xi_{k+n} \dots \xi_j \rangle \quad \text{if } n > m.$$

The simplest case is a 0-dependent process (no correlations at all) for which the cumulant functions of section 2 are

$$K_1 = \langle \xi_i \rangle; \quad K_m = 0, \quad m \geq 2. \quad (4.1)$$

Let us now consider the scalar difference equation

$$\Delta u(t-1) = \alpha \xi(t) u(t-1) \quad (t \in I_1), \quad (4.2)$$

$$u(0) = u_0,$$

where  $\xi(t) = a_0 \zeta(t) + a_1 \zeta(t-1)$ , and  $\{\zeta(t)\}_{t=-\infty}^{\infty}$  is a i.i.d. random sequence with mean zero and variance  $\sigma^2$ . Then  $\{\zeta(t)\}$  is a 0-dependent process, and  $\{\xi(t)\}$  is 1-dependent. The only non-zero  $t$ -cumulant of  $\xi$  is:

$$\langle \xi(t) \xi(t-1) \rangle_t = a_1 a_0 \sigma^2 \quad (t = 2, 3, \dots). \quad (4.3)$$

The  $K$ -functions are:

$$\begin{aligned}
 K_{2m+1} &= 0 \quad (m = 0, 1, 2, \dots), \\
 K_2 &= \alpha^2 \langle \xi(t) \xi(t-1) \rangle_p = \beta, \\
 K_4 &= K_4^{(2)} = \alpha^4 \langle \xi(t) \xi(t-1) \xi(t-1) \xi(t-2) \rangle_p \\
 &= -\alpha^4 \langle \xi(t) \xi(t-1) \rangle_i \langle \xi(t-1) \xi(t-2) \rangle_i = -\beta^2, \\
 K_6 &= K_6^{(3)} + K_6^{(2)} \\
 &= \alpha^6 \{ \langle \xi(t) \xi(t-1) \xi(t-1) \xi(t-2) \xi(t-2) \xi(t-3) \rangle_p \\
 &\quad + \langle \xi(t) \xi(t-1) \xi(t-1) \xi(t-1) \xi(t-2) \xi(t-2) \rangle_p \} \\
 &= \alpha^6 \{ \langle \xi(t) \xi(t-1) \rangle_i \langle \xi(t-1) \xi(t-2) \rangle_i \langle \xi(t-2) \xi(t-3) \rangle_i \\
 &\quad + \langle \xi(t) \xi(t-1) \rangle_i \langle \xi(t-1) \xi(t-2) \rangle_i \langle \xi(t-1) \xi(t-2) \rangle_i \} = 2\beta^3,
 \end{aligned} \tag{4.4}$$

where

$$\beta = \alpha^2 a_1 a_0 \sigma^2. \tag{4.5}$$

These results are valid if  $t \geq 4$ . In the same way one finds

$$K_8 = \begin{cases} -5\beta^4, & t > 4, \\ -4\beta^4, & t = 4, \end{cases} \tag{4.6}$$

because the term

$$\langle \xi(t) \xi(t-1) \rangle_i \langle \xi(t-1) \xi(t-2) \rangle_i \langle \xi(t-2) \xi(t-3) \rangle_i \langle \xi(t-3) \xi(t-4) \rangle_i$$

is absent when  $t = 4$  (subsection 2.3).

As a check of these results, we transfer the scalar equation (4.2) with autocorrelated noise to a vector equation with uncorrelated noise which can be solved exactly. That is, define the vector

$$X(t) = \begin{pmatrix} u(t) \\ \zeta(t)u(t) \end{pmatrix} \quad (t \in I_1); \quad X_0 = \begin{pmatrix} u_0 \\ \zeta(0)u_0 \end{pmatrix}, \tag{4.7}$$

which obeys the equation

$$\Delta X(t-1) = M(\zeta(t))X(t-1) \quad (t \in I_1), \tag{4.8}$$

$$X(0) = X_0,$$

where  $M(\zeta)$  is the matrix

$$M(\zeta) = \begin{pmatrix} \alpha a_0 \zeta & \alpha a_1 \\ (1 + \alpha a_0 \zeta) \zeta & \alpha a_1 \zeta - 1 \end{pmatrix}; \quad \langle M \rangle = \begin{pmatrix} 0 & \alpha a_1 \\ \alpha a_0 \sigma^2 & -1 \end{pmatrix}. \tag{4.9}$$

Since  $\{\zeta(t)\}$  is a i.d.d. sequence, we have

$$\Delta \langle X(t-1) \rangle = \langle M(\zeta(t)) \rangle \langle X(t-1) \rangle. \quad (4.10)$$

Because the initial condition  $X_0$  in (4.8) is random (it contains  $\zeta(0)$ ), there will in general be inhomogeneous terms in (4.10). In this case they vanish since  $X_0$  is statistically independent of  $M(\zeta(t))$ ,  $t = 1, 2, \dots$  (see part II). From (4.9) and (4.10) we find for  $\langle u(t) \rangle$  the second order difference equation

$$\langle u(t+2) \rangle - \langle u(t+1) \rangle - \beta \langle u(t) \rangle = 0 \quad (t \in I_2), \quad (4.11)$$

$$\langle u(0) \rangle = u_0, \quad \langle u(1) \rangle = u_0 \quad (4.12)$$

with solution

$$\langle u(t) \rangle = (1 + 4\beta)^{-1/2} (\lambda_+^{t+1} - \lambda_-^{t+1}) u_0, \quad (4.13)$$

where

$$\lambda_{\pm} = \frac{1}{2} \{1 \pm (1 + 4\beta)^{1/2}\}. \quad (4.14)$$

From (4.13) we can derive a difference equation for  $\langle u(t) \rangle$

$$\Delta \langle u(t-1) \rangle = K(t) \langle u(t-1) \rangle \quad (4.15)$$

with

$$K(t) = -\lambda_+ \lambda_- (\lambda_+^t - \lambda_-^t)^{-1} (\lambda_+^{t-1} - \lambda_-^{t-1}). \quad (4.16)$$

Expanding  $K(t)$  in powers of  $\alpha^2$  (i.e. in powers of  $\beta$ ) we find

$$K(t) = \begin{cases} \beta(1 - \beta + 2\beta^2 - 5\beta^3 \dots), & t > 4, \\ \beta(1 - \beta + 2\beta^2 - 4\beta^3 \dots), & t = 4, \end{cases}$$

in agreement with the previous results (4.4) and (4.6).

## 5. The singular case

In this final section we discuss the necessary modifications of the theory in the preceding sections when the difference equation is of the form

$$u(t) = \{A_0 + A_1(t, \omega)\} u(t-1), \quad t \in I_{t_0+1}, \quad (5.1)$$

$$u(t_0) = u_0,$$

where the deterministic matrix  $A_0$  is singular. We restrict ourselves here to a re-derivation of the result up to  $\mathcal{O}(\alpha^2)$ . As a first step we apply a variant of Terwiel's method<sup>25)</sup> to derive an exact difference equation of order  $t - t_0$  for the average  $\langle u(t) \rangle$ .

The method runs as follows: apply the averaging operators  $\mathcal{P}$  and  $\mathcal{Q}$ , as defined in (2.14), to both sides of (5.1) and express all occurring quantities in  $\mathcal{P}u$  and  $\mathcal{Q}u$ :

$$\mathcal{P}u(t) = \{A_0 + \alpha \mathcal{P}A_1(t)\} \mathcal{P}u(t-1) + \alpha \mathcal{P}A_1(t) \mathcal{Q}u(t-1), \quad (5.2a)$$

$$\mathcal{Q}u(t) = \{A_0 + \alpha \mathcal{Q}A_1(t)\} \mathcal{Q}u(t-1) + \alpha \mathcal{Q}A_1(t) \mathcal{P}u(t-1). \quad (5.2b)$$

Now (5.2b) is an inhomogeneous difference equation for  $\mathcal{Q}u(t)$  with initial condition  $\mathcal{Q}u(0) = 0$ . Its formal solution is

$$\begin{aligned} \mathcal{Q}u(t) &= \alpha \mathcal{Q}A_1(t) \mathcal{P}u(t-1) \\ &+ \sum_{s=t_0+1}^{t-1} \left[ \tilde{T} \prod_{s'=s+1}^t \{A_0 + \alpha \mathcal{Q}A_1(s')\} \right] \alpha \mathcal{Q}A_1(s) \mathcal{P}u(s-1). \end{aligned} \quad (5.3)$$

Inserting this in (5.2a) and using the definition of  $\mathcal{P}$  we find

$$\begin{aligned} \langle u(t) \rangle &= \{A_0 + \alpha \langle A_1(t) \rangle\} \langle u(t-1) \rangle + \alpha^2 \langle\langle A_1(t) A_1(t-1) \rangle\rangle \langle u(t-2) \rangle \\ &+ \alpha^2 \sum_{s=t_0+1}^{t-2} \left\langle A_1(t) \left[ \tilde{T} \prod_{s'=s+1}^{t-1} \{A_0 + \alpha \mathcal{Q}A_1(s')\} \right] \mathcal{Q}A_1(s) \right\rangle \langle u(s-1) \rangle, \end{aligned} \quad (5.4)$$

which is an exact equation for the average of  $u(t)$ . Neglecting all terms of order  $\alpha^3$  and higher, (5.4) reduces to

$$\begin{aligned} \langle u(t) \rangle &= \{A_0 + \alpha \langle A_1(t) \rangle\} \langle u(t-1) \rangle \\ &+ \alpha^2 \sum_{s=t_0+1}^{t-1} \langle\langle A_1(t) A_0^{t-1-s} A_1(s) \rangle\rangle \langle u(s-1) \rangle. \end{aligned} \quad (5.5)$$

If  $A_0$  is nonsingular we can put

$$\langle u(s-1) \rangle = A_0^{s-t} \langle u(t-1) \rangle + \mathcal{O}(\alpha) \quad (5.6)$$

in (5.5) because we neglected terms of  $\mathcal{O}(\alpha^3)$  already. This gives back our previous approximation (2.33).

Now consider the case that  $A_0$  is singular. Then one or more eigenvalues of  $A_0$  are zero. It turns out to be convenient if we consider the Jordan canonical form<sup>1)</sup>  $A_0^J$  of the matrix  $A_0$ :

$$A_0^J = S A_0 S^{-1}, \quad (5.7a)$$

where  $S$  is a nonsingular matrix. We can choose  $S$  in such a way that  $A_0^J$  takes the form ( $\emptyset$  denotes a null-matrix)

$$A_0^J = \left( \begin{array}{c|c} A_0^{(n)} & \emptyset \\ \hline \emptyset & A_0^{(s)} \end{array} \right), \quad (5.7b)$$



where  $A_0^{(n)}$  (dimension  $n_1 \times n_1$ ) is nonsingular and  $A_0^{(s)}$  (dimension  $n_2 \times n_2$ ,  $n_2 \geq 1$ ) contains all blocks with zero eigenvalues on their diagonal and is therefore singular. Now we distinguish between the following cases:

i) All eigenvalues of  $A_0$  are zero ( $n_2 = n$ , where  $n = \dim A_0$ )

Then  $(A_0^1)^n = 0$  and also  $A_0^n = 0$ . From (5.1) follows

$$u(t) = \left[ \tilde{T} \prod_{s=t-n+1}^t \{A_0 + \alpha A_1(s)\} \right] u(t-n),$$

where the matrix between brackets  $[\dots]$  is of order  $\alpha$  and higher, since the term of order  $\alpha^0$  is  $A_0^n = 0$ . So if we define  $v(t)$  by  $v(t) = u(tn)$ ,  $v(t)$  obeys

$$v(t) = A_1'(t)v(t-1),$$

where the matrix  $A_1'$  contains only matrices of order  $\alpha$  or higher. Clearly in this case no perturbation expansion in  $\alpha$  is possible.

ii) At least one of the eigenvalues of  $A_0$  is nonzero

In this case there is a block matrix  $A_0^{(n)}$  in (5.7b) which is non-singular. Now it is possible to derive a difference equation of the first order for  $\langle u(t) \rangle$ , which is valid after a transient time of order  $\tau_c$ . To derive the result, we first put (by using the matrix  $S$  in (5.7a))

$$Su(t) = \begin{pmatrix} u^{(n)}(t) \\ u^{(s)}(t) \end{pmatrix},$$

where  $u^{(n)}$  and  $u^{(s)}$  have dimension  $n_1$  and  $n_2$  and correspond to nonsingular and singular components of  $A_0$  respectively. The matrix  $A_1$  is transformed by  $S$  to a partitioned matrix which we denote by

$$SA_1S^{-1} = \begin{pmatrix} A_1^{(1)} & A_1^{(2)} \\ A_1^{(3)} & A_1^{(4)} \end{pmatrix},$$

where  $A_1^{(1)}$ ,  $A_1^{(2)}$ ,  $A_1^{(3)}$  and  $A_1^{(4)}$  are matrices of dimension  $n_1 \times n_1$ ,  $n_1 \times n_2$ ,  $n_2 \times n_1$  and  $n_2 \times n_2$ , respectively\*. In the same way we put

$$SA_1(t)A_0^{t-1-s}A_1(s)S^{-1} = \begin{pmatrix} B^{(1)} & B^{(2)} \\ B^{(3)} & B^{(4)} \end{pmatrix}.$$

Then by eq. (5.5) we find for the components  $u^{(n)}$  and  $u^{(s)}$  of  $u$  (which are

\*If  $A_1^{(2)}$  happens to be zero,  $u^{(n)}(t)$  obeys a closed equation which can be treated by the method of section 2.

themselves again vectors in general)

$$\begin{aligned} \langle u^{(n)}(t) \rangle &= \{A_0^{(n)} + \alpha \langle A_1^{(1)}(t) \rangle\} \langle u^{(n)}(t-1) \rangle + \alpha \langle A_1^{(2)}(t) \rangle \langle u^{(s)}(t-1) \rangle \\ &\quad + \alpha^2 \sum_{s=t_0+1}^{t-1} \{ \langle B^{(1)} \rangle \langle u^{(n)}(s-1) \rangle + \langle B^{(2)} \rangle \langle u^{(s)}(s-1) \rangle \}, \end{aligned} \quad (5.8a)$$

$$\begin{aligned} \langle u^{(s)}(t) \rangle &= \{A_0^{(s)} + \alpha \langle A_1^{(4)}(t) \rangle\} \langle u^{(s)}(t-1) \rangle + \alpha \langle A_1^{(3)}(t) \rangle \langle u^{(n)}(t-1) \rangle \\ &\quad + \alpha^2 \sum_{s=t_0+1}^{t-1} \{ \langle B^{(3)} \rangle \langle u^{(n)}(s-1) \rangle + \langle B^{(4)} \rangle \langle u^{(s)}(s-1) \rangle \}. \end{aligned} \quad (5.8b)$$

Now

$$\langle u(t-1) \rangle = A_0^{t-s} \langle u(s-1) \rangle + \mathcal{O}(\alpha),$$

so

$$\langle u^{(n)}(t-1) \rangle = (A_0^{(n)})^{t-s} \langle u^{(n)}(s-1) \rangle,$$

$$\langle u^{(n)}(s-1) \rangle = (A_0^{(n)})^{s-t} \langle u^{(n)}(t-1) \rangle$$

and

$$\langle u^{(s)}(t-1) \rangle = (A_0^{(s)})^{t-s} \langle u^{(s)}(s-1) \rangle.$$

From the last equality it follows that in lowest order

$$\langle u^{(s)}(t) \rangle = (A_0^{(s)})^{t-t_0} u^{(s)}(t_0),$$

so

$$\langle u^{(s)}(t) \rangle = 0 \quad \text{if } t - t_0 \geq n_2. \quad (5.9)$$

Since  $A_1(t)$  has a finite autocorrelation time  $\tau_c^\dagger$ , the sums in (5.8) virtually extend only over those values of  $s$  such that  $|t-1-s| \lesssim \tau_c$ . So if  $t-t_0-1 \gtrsim n_2 + \tau_c$  we have only a contribution if  $s-1-t_0 \gtrsim n_2$ :  $\langle u^{(s)}(s-1) \rangle \simeq \mathcal{O}(\alpha)$ . Therefore after a transient period the terms  $\langle u^{(s)}(s-1) \rangle$  under the sums in (5.8) can be put equal to zero, since they only contribute to order  $\alpha^3$ .

Summarizing, we can effectively make the replacement

$$\left( \begin{array}{c} \langle u^{(n)}(s-1) \rangle \\ \langle u^{(s)}(s-1) \rangle \end{array} \right) = \left( \begin{array}{c} (A_0^{(n)})^{s-t} \\ \frac{\emptyset}{\emptyset} \end{array} \right) \left( \begin{array}{c} \emptyset \\ \langle u^{(n)}(t-1) \rangle \\ \langle u^{(s)}(t-1) \rangle \end{array} \right)$$

in (5.8).

<sup>†</sup>See however the comment at the end of section 2.4.

### Conclusion

If  $A_0$  is singular, the result of the cumulant expansion for eq. (5.1) to order  $\alpha^2$  for times exceeding the transient time can be written as

$$\langle u(t) \rangle = \left[ A_0 + \alpha \langle A_1(t) \rangle + \alpha^2 \sum_{\tau=1}^{\infty} \langle\langle A_1(t) [A_0^{\tau-1} A_1(t-\tau)] \rangle\rangle (A_{0,N}^{-1})^{\tau} \right] \times \langle u(t-1) \rangle, \quad (5.10)$$

where  $A_{0,N}^{-1}$  ( $N$  for "nonsingular") is obtained via the following prescription: transform  $A_0$  to Jordan canonical form (5.7). Put  $A_0^{(s)} = 0$ , take the inverse of  $A_0^{(n)}$  and transform back to get  $A_{0,N}^{-1}$ :

$$A_{0,N}^{-1} = S^{-1} \left( \begin{array}{c|c} (A_0^{(n)})^{-1} & \emptyset \\ \hline -\frac{1}{\emptyset} & \emptyset \end{array} \right) S. \quad (5.11)$$

### Remark

If  $A_0^{(s)} \equiv 0$ , and  $[A_0, A_1] = 0$ , then  $\langle u^{(s)}(t) \rangle = \mathcal{O}(\alpha^{t-t_0})$  and therefore one can put  $\langle u^{(s)} \rangle \equiv 0$  in (5.8). So in this case only the equation for  $\langle u^{(n)} \rangle$  remains.

### Example

$$u(t) = \left[ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \alpha \xi(t) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] u(t-1) \quad (t \in I_1), \quad (5.12)$$

$$u(0) = u_0$$

or

$$u^{(n)}(t) = (1 + \alpha \xi(t)) u^{(n)}(t-1) + u^{(s)}(t-1), \quad (5.13a)$$

$$u^{(s)}(t) = \alpha \xi(t) u^{(s)}(t-1). \quad (5.13b)$$

In this case we can first solve (5.13b)

$$u^{(s)}(t) = \alpha^t \xi(t) \xi(t-1) \dots \xi(1) u_0^{(s)}$$

and substitute it in (5.13a). This gives an inhomogeneous equation for  $u^{(n)}$ :

$$\Delta u^{(n)}(t-1) = \alpha \xi(t) u^{(n)}(t-1) + \alpha^{t-1} \xi(t-1) \dots \xi(1) u_0^{(s)}, \quad (5.14)$$

where the inhomogeneous term is correlated with the multiplicative term  $\alpha \xi(t)^*$ . But since the additive term is of order  $\alpha^{t-1}$  we can ignore it after a few time steps. The remaining equation can be treated with the method of section 2, and the result coincides with that of (5.10).

\* In part II we will derive an exact expansion for the case of mutually correlated multiplicative and additive noise.

## Acknowledgements

The author is greatly indebted to Professor N.G. van Kampen for numerous critical discussions and valuable advice.

This investigation is part of the research program of the Stichting voor Fundamenteel Onderzoek der Materie (F.O.M.), which is financially supported by the Nederlandse Organisatie voor Zuiver Wetenschappelijk Onderzoek (Z.W.O.).

## Appendix

In this appendix we summarize the most important facts in relation to the cumulant expansion for linear stochastic differential equations.

Consider the linear s.d.e. with sure initial condition

$$\dot{u}(t) = \alpha A(t)u(t), \quad (\text{A.1})$$

$$u(t_0) = u_0. \quad (\text{A.2})$$

It has the formal solution

$$u(t) = \left[ \tilde{T} \exp \int_{t_0}^t ds \alpha A(s) \right] u_0$$

with moment expansion

$$\langle u(t) \rangle = \left[ 1 + \sum_{l=1}^{\infty} \alpha^l \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{l-1}} dt_l \langle A(t_1) A(t_2) \dots A(t_l) \rangle \right] u_0. \quad (\text{A.3})$$

The cumulant expansion yields for the average the following result<sup>13-15)</sup>

$$\langle \dot{u}(t) \rangle = K(t/t_0) \langle u(t) \rangle, \quad (\text{A.4})$$

where the matrix  $K(t/t_0)$  is a sure matrix with expansion

$$K(t/t_0) = \sum_{m=1}^{\infty} K_m(t/t_0), \quad (\text{A.5})$$

$$K_m(t/t_0) = \alpha^m \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{m-2}} dt_{m-1} \langle A(t) A(t_1) \dots A(t_{m-1}) \rangle_p. \quad (\text{A.6})$$

By definition  $K_1(t/t_0) = \alpha \langle A(t) \rangle$ . The cumulant functions  $K_m(t/t_0)$  are independent of  $t_0$  if  $t - t_0 \gg \tau_c$ , where  $\tau_c$  is the autocorrelation time of  $A(t)$ . In general  $\alpha \tau_c \ll 1$  is necessary.

The integrand of (A.6) is the so called "partially time-ordered cumulant" ( $p$ -cumulant), which is a certain combination of moments of  $A(\cdot)$  with a specific ordering of the time variables (see ref. 15, section 2.1, for the precise definition). These  $p$ -cumulants can be expressed in another type of cumulant, the so called "totally time-ordered cumulant" ( $t$ -cumulant), denoted as  $\langle \cdots \rangle_t$ . For  $m$  non-commuting quantities  $A_0(t), A_1(t_1), \dots, A_{m-1}(t_{m-1})$ , where  $t > t_1 > \dots > t_{m-1}$ , the  $t$ -cumulant  $\langle A_0(t)A_1(t_1) \dots A_{m-1}(t_{m-1}) \rangle_t$  is defined in terms of the moments ( $A_k(t_k)$  abbreviated as  $k$ )

$$\begin{aligned} \langle 0 \rangle_t &= \langle 0 \rangle, \\ \langle 012 \dots (m-1) \rangle_t &= \langle (0)\mathcal{Q}(1)\mathcal{Q}(2) \dots \mathcal{Q}(m-1) \rangle, \end{aligned} \quad (\text{A.7})$$

where  $\mathcal{Q}$  is the averaging operator defined in (2.14). The connection between  $p$ - and  $t$ -cumulants was given in appendix A of ref. 15.

To take into account the possibility that in the  $p$ -cumulant some of the time variables  $t_i$  may be identical we generalize the rules given in ref. 15 as follows. To obtain  $\langle A_0(t)A_1(t_1) \dots A_{m-1}(t_i) \rangle_p$ ; where the  $m-1$  numerals  $\{1, 2, 2, \dots, \dots, i, \dots, i\}$  are chosen from a set  $S(i, m-1)$  as defined in section 2.3, use the following prescription:

- i) Write a sequence of  $m$  dots.
- ii) Write a zero on the first dot, and any distinct permutation of the  $m-1$  numerals (some of which may be equal) in  $S(i, m-1)$  on the remaining dots.
- iii) Partition each of the distinct permutations of numerals into subsequences by inserting angular brackets  $\langle \cdots \rangle_t$  (denoting  $t$ -cumulants) in such a way that two successive numerals belong to the same subsequence if and only if the first one is smaller than the second.
- iv) For each partition consisting of  $p$  subsequences supply a factor  $(-)^{p-1}$ .
- v) Replace each numeral  $n$  on the  $k$ th dot by  $A_{k-1}(t_n)$ .

As an illustration we give here all the  $p$ -cumulants occurring in (2.23) in terms of the  $t$ -cumulants (which can in their turn be expressed in the moments via (A.7)):

$$\begin{aligned} \langle 0 \rangle_p &= \langle 0 \rangle_t, \quad \langle 01 \rangle_p = \langle 01 \rangle_t, \\ \langle 012 \rangle_p &= \langle 012 \rangle_t - \langle 02 \rangle_t \langle 1 \rangle_t; \quad \langle 011 \rangle_p = -\langle 01 \rangle_t \langle 1 \rangle_t, \\ \langle 0123 \rangle_p &= \langle 0123 \rangle_t - \langle 013 \rangle_t \langle 2 \rangle_t - \langle 02 \rangle_t \langle 13 \rangle_t - \langle 023 \rangle_t \langle 1 \rangle_t \\ &\quad - \langle 03 \rangle_t \langle 12 \rangle_t + \langle 03 \rangle_t \langle 2 \rangle_t \langle 1 \rangle_t, \\ \langle 0122 \rangle_p &= -\langle 012 \rangle_t \langle 2 \rangle_t - \langle 02 \rangle_t \langle 12 \rangle_t + \langle 02 \rangle_t \langle 1 \rangle_t \langle 1 \rangle_t, \end{aligned}$$

$$\begin{aligned}\langle 0112 \rangle_p &= -\langle 012 \rangle_t \langle 1 \rangle_t - \langle 01 \rangle_t \langle 12 \rangle_t + \langle 02 \rangle_t \langle 1 \rangle_t \langle 1 \rangle_t, \\ \langle 0111 \rangle_p &= \langle 01 \rangle_t \langle 1 \rangle_t \langle 1 \rangle_t.\end{aligned}$$

## References

- 1) R. Bellman, *Introduction to Matrix Analysis* (McGraw Hill, New York, 1970).
- 2) F.R. Gantmacher, *Applications of the Theory of Matrices* (Intersc. Publ., New York, 1959).
- 3) M.S. Bartlett, *An Introduction to Stochastic Processes, Methods and Applications* (Cambr. Univ. Press, Cambridge, 1956).
- 4) K.S. Miller, *Linear Difference Equations* (Benjamin, New York, 1968).
- 5) T.W. Anderson, *An Introduction to Statistical Analysis of Time Series* (Wiley, New York, 1969).
- 6) H.B. Mann and A. Wald, *Econometrika* **11** (1943) 173.
- 7) E. Malinvaud, *Statistical Methods of Econometrics* (North-Holland, Amsterdam, 1970).
- 8) A.S. Paulson and V.R.R. Uppuluri, *Math. Biosc.* **13** (1972) 325.
- 9) F. Solomon, *Ann. Prob.* **3** (1975) 1.
- 10) W. Vervaat, *Adv. Appl. Prob.* **11** (1979) 750.
- 11) H. Kesten, *Acta Math.* **131** (1973) 207.
- 12) H. Furstenberg and H. Kesten, *Ann. Math. Statist.* **31** (1960) 457.
- 13) N.G. van Kampen, *Physica* **74** (1974) 215, 239; *Phys. Rep.* **C24** (1976) 173.
- 14) N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam 1981), ch. XIV.
- 15) J.B.T.M. Roerdink, *Physica* **109A** (1981) 23.
- 16) J.B.T.M. Roerdink, *Physica* **112A** (1982) 557.
- 17) J.E. Cohen, I. J. Amer. Stat. Ass. **71** (1976) 335; II: *Adv. Appl. Prob.* **9** (1977) 18; III: *Adv. Appl. Prob.* **9** (1977) 462; *J. Math. Anal. Appl.* **69** (1979) 243.
- 18) S. D. Tuljapurkar and S. H. Orzack, I: *Theor. Pop. Biol.* **18** (1980) 314 and references quoted therein.
- 19) S.D. Tuljapurkar, II: *Theor. Pop. Biol.* **21** (1982) 114; III: *Theor. Pop. Biol.* **21** (1982) 141.
- 20) P. Billingsley, *Convergence of Probability Measures* (Wiley, New York, 1968).
- 21) H. A. Guess and J. H. Gillespie, *J. Appl. Prob.* **14** (1977) 58.
- 22) M. Turelli, *Theor. Pop. Biol.* **12** (1977) 140.
- 23) M. Loève, *Probability Theory II* (Springer, New York, 1978), pp. 130–135.
- 24) J. Lamperti, *Stochastic Processes* (*Appl. Math. Sc.*, vol. 23; Springer, New York, 1977) pp. 118–121.
- 25) R. H. Terwiel, *Physica* **74** (1974) 248.
- 26) See also: R. Bellman, T.T. Soong and R. Vasudevan, *J. Math. Anal. Appl.* **40** (1972) 286.