Journal of Statistical Physics, Vol. 6, No. 2/3, 1972

On the Solutions and the Steady States of a Master Equation

Joel Keizer[1]

Received November 2, 1971

A complete characterization of the time behavior of the means and variance of a stochastic process which is generated by a finite number of independent systems is presented based on the master equation for the conditional probability. It is found that the means and variance relax to a steady state and that the steady state will be independent of the initial state if and only if a matrix related to the transition matrix is nonsingular. Finally, the result that the variance approaches its steady-state form at twice the rate of the means is shown to depend on the nonsingularity of the same matrix.

KEY WORDS: Master equation; stochastic process; steady states.

In a recent paper,[1] the approach to steady state of a collection of independent subsystems was discussed based on a master equation for the conditional probability. The transition probabilities in this equation were assumed linear in the populations of the subsystem states, although otherwise the form of the equation was quite general. It was shown that the column vector of means $\langle \mathbf{N} \rangle(t)$ with components $\langle N_i \rangle(t)$ satisfies the equation

$$d\langle \mathbf{N} \rangle / dt = -\Lambda \langle \mathbf{N} \rangle \tag{1}$$

and that the modified variance matrix $\xi_{ij}(t) = \sigma_{ij}(t) - f_{ij}(t)$, where σ is the usual variance matrix defined by

$$\sigma_{ij} = \langle N_i N_j \rangle - \langle N_i \rangle \langle N_j \rangle$$

[1] Department of Chemistry, University of California, Davis, California.

and

$$f_{ij} = \langle N_i \rangle \delta_{ij} - \langle N_i \rangle \langle N_j \rangle / B$$

satisfies the equation

$$d\xi/dt = -\Lambda\xi - \xi\Lambda^T \tag{2}$$

In these equations, $\sum_{j=0}^{r} N_j(t) = B$ (a fixed positive number), Λ is an $(r+1) \times (r+1)$ matrix with $-\Lambda_{ij} \geqslant 0$, and $\sum_{j \neq i} \Lambda_{ji} = -\Lambda_{ii}$, and the superscript T means the transpose. It is the purpose of this note to clarify and extend remarks made in several papers[1-4] concerning the nature of the solutions to Eqs. (1) and (2).

To provide a basis for discussion, the following statements concerning the matrix Λ and the properties of the equations are proven: (a) there is at least one eigenvector of $-\Lambda$ with eigenvalue zero, and all of the nonzero eigenvalues have negative real parts of magnitude less than 2 $\max_i \{\Lambda_{ii}\}$. Furthermore, the zero eigenvector is unique if the graph of Λ is strongly connected, that is, if all the states are accessible. (b) The truncated means and variance, $\langle N_i \rangle - \langle N_i \rangle^{ss}$, $i \neq 0$, and ξ_{ij}, $i, j \neq 0$, where ss denotes the steady state, satisfy linear equations similar to Eqs. (1) and (2) [see Eqs. (3) and (4) below] with an $r \times r$ matrix Λ_r replacing Λ. With the exception of an extra zero eigenvalue for Λ, Λ and Λ_r have the same eigenvalues and there is a simple correspondence between the nonzero eigenvectors of the two matrices. The zero eigenvalue of Λ is unique if and only if Λ_r is nonsingular. (c) If the means $\langle N_i \rangle$ are initially nonnegative, then they remain nonnegative and bounded for all future times.

Before proving statement (a), note that in case some of the components Λ_{ij} are zero, the validity of the first part of the statement does not follow directly from the Routh–Hurwitz criterion.[4] The proof given here—which includes this case—uses the elegant theorem of Perron and Frobenius.[5], [2] Thus let $-\Lambda$ be a matrix of the form given above and let $q = \max_i \{\Lambda_{ii}\} + \delta$, where $\delta > 0$. Then, the matrix $Q = -\Lambda + qI$, where I is the identity, is a nonnegative matrix and so by the Perron–Frobenius theorem,[3] Q has a positive eigenvalue r equal to the spectral radius of Q *and* there exists a nonnegative right eigenvector associated with r. But since $-\Lambda = Q - qI$, this eigenvector is also an eigenvector of $-\Lambda$ with eigenvalue $r - q$. However notice that if $-\Lambda\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha}$ for $\lambda \neq 0$, then the condition $\sum_i \Lambda_{ij} = 0$ implies that $\sum_i \alpha_i = 0$. Thus only the zero eigenvectors of $-\Lambda$ can have nonnegative components and so $r - q = 0$, and at least one zero eigenvector (with nonnegative components) exists for $-\Lambda$. Since the eigenvalues of $-\Lambda$ are trans-

[2] Although a proof of the first part of statement (a) is given in Ref. 2, the accessibility criterion for unique steady states does not follow from that proof.

[3] See Ref. 5, Theorem 9.3.1.

lations by $-q$ of the eigenvalues of Q and since the spectral radius of Q was shown to equal q, it follows that all the eigenvalues of $-\Lambda$ are in the closed disk of radius q centered at $-q$. However, $q = \max_i \{\Lambda_{ii}\} + \delta$ for arbitrary $\delta > 0$, so that the nonzero eigenvalues of $-\Lambda$ must have negative real parts of magnitude less than or equal to $2 \max_i \{\Lambda_{ii}\}$.[4] It should be noted that if $-\Lambda$ is irreducible or alternately if its graph is strongly connected,[5] then the same argument implies that the zero eigenvalue of Λ is *unique*. Since strong connectivity implies that there are nonzero matrix elements connecting all states together either directly or by multiple transitions, this provides a generalization of the result proven by Siegert[(6)] in case Λ satisfies detailed balance, namely, that if all states are accessible, the steady state is unique. In particular if the matrix elements of $-\Lambda$ are all nonzero, then the graph of $-\Lambda$ will be strongly connected and so a matrix with positive transition rates has a unique zero eigenvalue.

To verify statement (b), let $\langle \mathbf{N} \rangle_r$ be the column vector with components $\langle N_i \rangle - \langle N_i \rangle^{ss}$, $i \neq 0$, and let the matrix ξ_r be defined by $(\xi_r)_{ij} = \xi_{ij}$, $i, j \neq 0$. Then, it follows by direct substitution of the conservation relation $N_0 = B - \sum_{i=1}^{r} N_i$ in Eqs. (1) and (2) that

$$d\langle \mathbf{N} \rangle_r/dt = -\Lambda_r \langle \mathbf{N} \rangle_r \tag{3}$$

$$d\xi_r/dt = -\Lambda_r \xi_r - \xi_r \Lambda_r^T \tag{4}$$

where Λ_r is the $r \times r$ matrix whose components are $(\Lambda_r)_{ij} = \Lambda_{ij} - \Lambda_{i0}$ for $i, j \neq 0$ and where $\langle N_i \rangle^{ss}$ solves $d\langle N_i \rangle/dt = 0$. To see the relationship between the nonzero eigenvectors of Λ and Λ_r, let

$$P = \begin{pmatrix} 1 & 1 \cdots 1 \\ 0 & \\ \vdots & I_r \\ 0 & \end{pmatrix}$$

an $(r + 1) \times (r + 1)$ matrix where I_r is the $r \times r$ identity, so that

$$P^{-1} = \begin{pmatrix} 1 & -1 \cdots -1 \\ 0 & \\ \vdots & I_r \\ 0 & \end{pmatrix}.$$

[4] It is not hard to construct matrices whose eigenvalues are actually complex. For example,

$$-\Lambda = \begin{pmatrix} -2.1 & 0.9 & 0.65 \\ 0.1 & -1.9 & 1.35 \\ 2 & 1 & -2 \end{pmatrix}$$

has $x = 0, -3 \pm i/2$ as eigenvalues.

[5] See Ref. 5, p. 281.

Then, it is easily seen that

$$\hat{\Lambda} = P\Lambda P^{-1} = \begin{pmatrix} 0 & 0 \cdots 0 \\ \Lambda_{01} & \\ \vdots & \Lambda_r \\ \Lambda_{0r} & \end{pmatrix}$$

and that

$$\hat{\boldsymbol{\alpha}} = P\boldsymbol{\alpha} = \begin{pmatrix} \sum_i \alpha_i \\ \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix}$$

Since Λ and $\hat{\Lambda}$ are similar, it follows that $\Lambda\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha}$ implies $\hat{\Lambda}\hat{\boldsymbol{\alpha}} = \lambda\hat{\boldsymbol{\alpha}}$ or

$$\begin{pmatrix} 0 & 0 \cdots 0 \\ \Lambda_{01} & \\ \vdots & \Lambda_r \\ \Lambda_{0r} & \end{pmatrix} \begin{pmatrix} \sum_i \alpha_i \\ \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix} = \lambda \begin{pmatrix} \sum_i \alpha_i \\ \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix}. \tag{5}$$

But if $\lambda \neq 0$, then $\sum_i \alpha_i$ must vanish and so Eq. (5) implies that $\Lambda_r\boldsymbol{\alpha}_r = \lambda\boldsymbol{\alpha}_r$, where $\boldsymbol{\alpha}_r = (\alpha_1, ..., \alpha_r)^T$. Thus the components $\alpha_1, ..., \alpha_r$ of the nonzero eigenvectors of Λ are identical to the corresponding components of the nonzero eigenvectors of Λ_r and α_0 is determined by $\sum_i \alpha_i = 0$. Moreover, since $\hat{\Lambda}$ and Λ are similar, all of their eigenvalues are the same. But the eigenvalues of $\hat{\Lambda}$ are determined by $\det(\hat{\Lambda} - \lambda I) = -\lambda \det(\Lambda_r - \lambda I_r) = 0$ and so with the exception of an extra zero eigenvalue for Λ, Λ and Λ_r have identical eigenvalues. Thus Λ will have a unique zero eigenvalue if and only if Λ_r is nonsingular.[6]

To prove statement (c), the invariance of the nonnegativity of the means, recall that the solution to Eq. (1) is $\langle \mathbf{N}\rangle(t) = \exp(-\Lambda t)\langle \mathbf{N}\rangle(0)$. Now, it is known[7] that $-\Lambda_{ij} \geqslant 0$ implies that the matrix $\exp(-\Lambda t)$ is nonnegative. Thus if $\langle \mathbf{N}\rangle(0)$ is nonnegative, then so is $\langle \mathbf{N}\rangle(t)$. Because of the sum rule $\sum_j \langle N_j\rangle(t) = B$, the means will be bounded above by B if they are initially nonnegative.

Using these results, it is possible to obtain a fairly complete picture of the solutions to Eqs. (1) and (2). This can be done most easily using Eqs. (1) and (4). In particular, these equations are solved by $\langle \mathbf{N}\rangle(t) = \exp(-\Lambda t)\langle \mathbf{N}\rangle(0)$ and $\xi_r = \exp(-\Lambda_r t)\, \xi_r(0) \exp(-\Lambda_r^T t)$, as is verified by differentiation. More explicit expressions can be gotten from the spectral resolutions of the exponentials, a result which known for arbitrary complex-valued matrices.[8] Indeed,

[6] An example of a matrix with two zero eigenvalues is the 3×3 matrix with components $\Lambda_{12} = \Lambda_{32} = -1$, $\Lambda_{22} = 2$, and all other $\Lambda_{ij} = 0$.

[7] See Bellman.[7] Notice that the result is somewhat stronger than quoted in Ref. 3.

[8] See Ref. 5, pp. 189–191.

$\langle N_i \rangle(t)$ will be a sum of exponentials in the time whose time constants are the (possibly complex) eigenvalues of Λ and whose coefficients are polynomials in t of degree less than the index of the eigenvalue. In particular, the polynomial coefficient of the zero eigenvalues must be of degree zero since by the result in (c), $\langle N_i \rangle(t)$ is bounded if all $\langle N_i \rangle(0) \geqslant 0$. Also, if Λ is simple, that is, diagonalizable, then all the polynomial coefficients will be constants. Since the nonzero eigenvalues of Λ have negative real parts, the means will always relax to some steady-state values. The relaxation will involve, in the general case, polynomials as well as sines and cosines in the time with each term multiplied by a decaying exponential whose time constant is no greater than $2 \max_i \{\Lambda_{ii}\}$. The steady state to which the means decay is not necessarily unique and is unique if and only if the matrix Λ_r is nonsingular. For example, the matrix given in footnote 6 is simple and has two linearly independent zero eigenvectors. Thus the asymptotic expression for

$$\langle \mathbf{N} \rangle(t) = \exp(-\Lambda t) \langle \mathbf{N} \rangle(0)$$

is just the projection of $\langle \mathbf{N} \rangle(0)$ onto the space of zero eigenvalues. Since this space is two-dimensional, the steady state will not be unique. This occurs because the multiple zero eigenvalues of Λ have very different physical origins. The zero eigenvalue that results from the sum conditions $\sum_j N_j = B$ is actually a holonomic constraint on the variables[8] and so is "removeable" as indicated by (b) above. The domain of definition of the variables is not effected, however, by the "accidental" zero eigenvalues of Λ_r and so these zeros have the purely dynamical effect of leading to multiple steady states.

The form of the solution to the variance equation, Eq. (4), will be similar to that of the means except that the product of the spectral resolutions must be taken. This implies that the exponentials in the solution will be of the form $\exp[-(\lambda_k + \lambda_j)\, t]$, where λ_k and λ_j are the eigenvalues of Λ_r, and that the polynomial coefficients will be of degree less than the product of the indices of the two eigenvalues. If Λ_r can be diagonalized, then all the polynomial coefficients will be constants. While it is clear that the variance will relax to a steady state, it is again true that the steady state may or may not be unique. In fact, the steady state is determined by setting $d\xi/dt = 0$, that is, by the solutions to the equation

$$\Lambda_r \xi_r + \xi_r \Lambda_r^T = 0$$

Now, it is known that this equation will have a unique solution if and only if Λ_r and $-\Lambda_r^T$ have no eigenvalues in common.[9] Since the eigenvalues of Λ_r and Λ_r^T are identical and their nonzero eigenvalues have negative real parts, the solution will be unique if and only if Λ_r has no zero eigenvalues, that is,

[9] See Ref. 5, Theorem 8.5.1.

if Λ_r is nonsingular. In that case, the unique steady-state solution is $\xi_r = 0$, which implies that the steady-state form of the variance is

$$\sigma_{ij}^{ss} = \delta_{ij}\langle N_i\rangle^{ss} - \langle N_i\rangle^{ss}\langle N_j\rangle^{ss}/B$$

Also, if Λ_r is nonsingular, then all the exponentials occurring in $\xi(t)$ will contain sums of nonzero eigenvalues, and so the approach to the steady-state form of the variance will occur asymptotically at twice the rate of the approach of the mean.[1]

On the other hand, if Λ_r is not invertible, then multiple solutions to the steady-state variance equation exist. In this case, the approach to the steady state will not be asymptotically twice as fast as that of the means. This occurs because when Λ_r has a zero eigenvalue, exponentials of the form $\exp[-(0 + \lambda_j)\, t]$ can appear in $\xi(t)$ and these slowly varying components cannot be removed by simply subtracting out the asymptotic steady state. Moreover, if multiple steady states are possible, the steady-state value of the variance will depend on the initial value of ξ_r, which is arbitrary. Thus for singular Λ_r, it is no longer be possible to define the "form" of the variance at steady state,[10] except to say that $\sigma_{ij}^{ss} = f_{ij}^{ss} + \xi_{ij}(\infty)$. It is worth repeating that the existence of these pathological solutions is a result of the different physical origins of multiple zero eigenvalues for the matrix Λ, and in case all states are accessible, such things will not occur.[11]

ACKNOWLEDGMENT

The author would like to thank Professor Terrell Hill for a number of stimulating conversations and correspondence which led to many of the results in this note.

REFERENCES

1. T. L. Hill, *J. Chem. Phys.* **54**:34 (1971); T. L. Hill and I. W. Plesner, *J. Chem. Phys.* **43**:267 (1965).
2. P. J. Gans, *J. Chem. Phys.* **33**:691 (1960); I. M. Krieger and P. J. Gans, *J. Chem. Phys.* **32**:247 (1960).
3. I. Oppenheim, K. E. Shuler, and G. Weiss, *Advan. Mol. Relaxation Processes* **1**:13 (1967).
4. T. A. Bak, *Contributions to the Theory of Chemical Kinetics* (Benjamin, New York, 1963).
5. P. Lancaster, *Theory of Matrices* (Academic Press, New York, 1969).
6. A. J. F. Siegert, *Phys. Rev.* **76**:1708 (1949).
7. R. Bellman, *Introduction to Matrix Analysis*, 2nd Ed. (McGraw-Hill, New York, 1970), p. 176.
8. H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 1959), pp. 10–14.

[10] For example, this is the case for the matrix in footnote 6.

[11] Note added in proof: The author would like to point out an article by J. Z. Hearon [*Ann. N.Y. Acad. Sci.* **108**:36 (1963)] which has come to his attention and which also uses the Perron–Frobenius theorem to analyze Eq. (1).