The Principle of the Diffusion of Arbitrary Constants

Andrew D. Barbour

Journal of Applied Probability, Vol. 9, No. 3. (Sep., 1972), pp. 519-541.

Stable URL:
http://links.jstor.org/sici?sici=0021-9002%28197209%299%3A3%3C519%3ATPOTDO%3E2.0.CO%3B2-6

Journal of Applied Probability is currently published by Applied Probability Trust.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/apt.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Fri Nov 24 12:46:52 2006

J. Appl. Prob. **9**, 519–541 (1972)
Printed in Israel
© *Applied Probability Trust* 1972

THE PRINCIPLE OF THE DIFFUSION OF ARBITRARY CONSTANTS

ANDREW D. BARBOUR, *University of Cambridge*

Abstract

Equations are derived describing a central limit type large population approximation for continuous time Markov lattice processes in one or more dimensions, such as are commonly encountered in biological models. A method of solving the equations using only the deterministic solution of the process is explained, and it is extended by the use of a martingale argument to provide more detailed information about the process.

MARKOV POPULATION PROCESSES; EPIDEMIC; RUMOUR; MARTINGALE; COMPETITION PROCESS

1. Introduction

The most elementary stochastic models of biological phenomena are those which can be represented as a Markov random walk in continuous time on a multi-dimensional integer lattice, with position (and possibly time) dependent transition rates: the birth and death process is a typical one-dimensional example. The general solution of such a process, when it can be evaluated, is usually tedious, and affords little insight into the qualitative dependence of the behaviour of the process on the parameters: Monte-Carlo estimates of the evolution of the process are often extremely expensive in computer time. This paper derives a technique for obtaining analytic approximations to the first few cumulants of the process, when the 'typical population size' becomes large, in cases where each transition rate is a multinomial expression in the lattice coordinates. The approach is essentially heuristic, being more concerned about the practical effectiveness of the ideas than about their rigorous establishment, and is susceptible of considerable generalization.

Sections 2–6 lay the foundations for the construction of the approximations, and establish ordinary differential equations for a Normal approximation to the time dependent solution, which can easily be solved by computer. Sections 7 and 8 derive the leading terms in the asymptotic expansions of the mean lattice position and of the corresponding covariance matrix entirely as functions of the deterministic solution of the process, and indicate how further approx-

Received in revised form 24 July 1971.

imations can be found. Section 9 illustrates the uses of the techniques in the case of the stochastic rumour in a closed population.

2. Rewriting the forward equation

Consider a continuous time discrete step Markov random walk on an n-dimensional integer lattice. Denote by $X(t) = (X_1(t), \cdots, X_n(t))$ the position at time t, and by $f^{\boldsymbol{J}}(\boldsymbol{X}, t)$ the transition rate at time t for jumps between $\boldsymbol{X}$ and $\boldsymbol{X} + \boldsymbol{J}$. Then (Bailey [1]) the Kolmogorov forward equation for the characteristic function of the process is

$$\frac{\partial F}{\partial t} = \sum_{\boldsymbol{J} \in \mathscr{J}} \left(e^{i\boldsymbol{J}.\boldsymbol{\theta}} - 1 \right) f^{\boldsymbol{J}} \left(\frac{1}{i}, \frac{\partial}{\partial \boldsymbol{\theta}}, t \right) F \tag{2.1}$$

where $F \equiv E\{e^{i\boldsymbol{\theta}.\boldsymbol{X}(t)}\}$, $\mathscr{J}$ is the set of all possible jumps with, for some $(\boldsymbol{X}, t)$, non-zero transition rates, and $f^{\boldsymbol{J}}((1/i)\partial/\partial\boldsymbol{\theta}, t)$ represents formally the operator derived by substituting $(1/i)\partial/\partial\theta_j$ for X_j in the (finite) series expansion of $f^{\boldsymbol{J}}(\boldsymbol{X}, t)$. The argument t is included in $f^{\boldsymbol{J}}(\boldsymbol{X}, t)$ to incorporate cases, such as an epidemic with active public health authorities, when the transition rates may be functions of time: in many applications there is no such time-dependence, and the transition rates can be taken as constant in time throughout this paper without impairing the essential argument.

Suppose that $\mathscr{J}$ is a fixed finite set with $|\boldsymbol{J}| \leqq B$ (for all $\boldsymbol{J} \in \mathscr{J}$), that each $f^{\boldsymbol{J}}(\boldsymbol{X}, t)$ admits of a multinomial expansion in terms of the variables $\{X_1, \cdots, X_n\}$, and that the highest degree of any term in any of these expansions is d; consider the behaviour of the process with initial condition $\boldsymbol{X}(0) = N\boldsymbol{x}(0)$, for a fixed choice $\boldsymbol{x}(0) = (x_1(0), \cdots, x_n(0))$, as N becomes large compared to B. N thus defined is a scale variable describing a typical size for $\boldsymbol{X}$, and so terms of degree p $(0 \leqq p \leqq d)$ in the multinomial expansions of the $f^{\boldsymbol{J}}(\boldsymbol{X}, t)$ will be of order N^p as $N \uparrow \infty$. In particular, any terms of degree strictly smaller than d will be of small order compared with the terms in N^d; in order to retain their contribution to the process, it will in general be necessary to allow the coefficients in the expansions of $f^{\boldsymbol{J}}(\boldsymbol{X}, t)$ to vary with N as well, and there will usually, in practical applications, be at least one natural way of adjusting the parameters influencing the process in order to achieve this. For $\boldsymbol{x} = N^{-1}\boldsymbol{X}$ and $u = N^{d-1}t$, let $g^{\boldsymbol{J}}(\boldsymbol{x}, u) \equiv N^{-d} f^{\boldsymbol{J}}(\boldsymbol{X}, t)$ be the N-independent scaled version of $f^{\boldsymbol{J}}(\boldsymbol{X}, t)$. Then from (2.1) we derive the following equation for $G \equiv E\{e^{i\boldsymbol{\theta}.\boldsymbol{x}}\} \equiv E\{e^{i\boldsymbol{\theta}N^{-1}.\boldsymbol{X}}\}$, the characteristic function for $\boldsymbol{x} = \boldsymbol{X}N^{-1}$:

$$\frac{\partial G}{\partial t} = \sum_{\boldsymbol{J} \in \mathscr{J}} \left(e^{i\boldsymbol{J}.\boldsymbol{\theta}N^{-1}} - 1 \right) f^{\boldsymbol{J}} \left(N, \frac{1}{i}, \frac{\partial}{\partial \boldsymbol{\theta}}, t \right) G, \tag{2.2}$$

and putting $u = N^{d-1}t$ we obtain, if $H(\boldsymbol{\theta}, u) \equiv E\{e^{i\boldsymbol{\theta}.\boldsymbol{x}(u)}\}$,

$$\frac{\partial H}{\partial u} = \sum_{\boldsymbol{J} \in \mathscr{J}} \{N(e^{i\boldsymbol{J}.\boldsymbol{\theta}N^{-1}} - 1)\} g^{\boldsymbol{J}}\left(\frac{1}{i}\frac{\partial}{\partial \boldsymbol{\theta}}, u\right) H. \tag{2.3}$$

The initial value of H is $\exp\{i\boldsymbol{\theta}.\boldsymbol{x}(0)\}$, independent of N by choice, and so the N-dependence in the solution of (2.3) is confined entirely to the term in braces, which, by expanding the exponential, can be expressed as a power series $\sum_{k=0}^{\infty} b_k N^{-k}$ in inverse powers of N. This, together with the fact that $|H| \leqq 1$, enables successive approximations to H in a fixed interval $u \in [0, U]$ to be derived by solving the equations obtained from (2.3) by neglecting high powers of N^{-1}.

If, as in the stochastic rumour investigated in Section 9, there is no natural way in which the $f^{\boldsymbol{J}}$ can be scaled to make the $g^{\boldsymbol{J}}$ entirely N-independent, we can at least assert that, since d is the highest power in the multinomial expansions of the $f^{\boldsymbol{J}}$, any N-dependent term in $g^{\boldsymbol{J}} \equiv N^{-d} f^{\boldsymbol{J}}$ must be of order N^{-r} for $r \geqq 1$, and so small compared with 1. Thus, although the N-dependence in (2.3) will no longer be confined to the term in braces, expansions of (2.3) to the relevant orders in N^{-1} can still be obtained, and the essential reasoning will remain unchanged.

3. The deterministic approximation

As a first approximation to the solution of (2.3), as $N \uparrow \infty$, take only the N-independent terms, and solve the equation

$$\frac{\partial H^{(0)}}{\partial u} = \sum_{\boldsymbol{J} \in \mathscr{J}} i\boldsymbol{\theta}.\boldsymbol{J} g^{\boldsymbol{J}}\left(\frac{1}{i}\frac{\partial}{\partial \boldsymbol{\theta}}, u\right) H^{(0)}. \tag{3.1}$$

The solution is of deterministic form $H^{(0)} = \exp\{i\boldsymbol{\theta}.\boldsymbol{\xi}(u)\}$, where $\boldsymbol{\xi}(u)$ is the (non random) vector function of u satisfying, for each component $\xi_\alpha(u)$ $(\alpha = 1, \cdots, n)$,

$$\frac{d\xi_\alpha(u)}{du} = \sum_{\boldsymbol{J} \in \mathscr{J}} J_\alpha g^{\boldsymbol{J}}(\boldsymbol{\xi}(u), u). \tag{3.2}$$

These differential equations for $\boldsymbol{\xi}(u)$ are obtained by substituting the desired form for $H^{(0)}$ in (3.1), and equating coefficients of θ_α, and are the usual deterministic equations for the process; they are to be solved with the initial condition $\boldsymbol{\xi}(0) = \boldsymbol{x}_0$.

4. The diffusion approximation; a central-limit type solution

To obtain an approximation to the stochastic fluctuation about the deterministic limiting path, as $N \uparrow \infty$, we now take also the terms of order N^{-1} in (2.3), and consider the equation

$$\frac{\partial H^{(1)}}{\partial u} = \sum_{\mathbf{J}\in\mathscr{J}} \left\{ i\boldsymbol{\theta}.\mathbf{J} - \frac{1}{2N}(\boldsymbol{\theta}.\mathbf{J})^2 \right\} g^{\mathbf{J}}\left(\frac{1}{i}\frac{\partial}{\partial\boldsymbol{\theta}}, u\right) H^{(1)}. \tag{4.1}$$

This can be shown to be the (Fokker-Planck) equation for the inhomogenous diffusion approximation to the process, expressed in characteristic function form. The Fokker-Planck equation is usually written as

$$\frac{\partial p(\mathbf{x},u)}{\partial u} = -\sum_{\alpha}\frac{\partial}{\partial x_\alpha}\left(\sum_{\mathbf{J}} J_\alpha g^{\mathbf{J}}(\mathbf{x},u)p(\mathbf{x},u)\right) + \frac{1}{2N}\sum_{\alpha,\beta}\frac{\partial^2}{\partial x_\alpha \partial x_\beta}\left(\sum_{\mathbf{J}} J_\alpha J_\beta g^{\mathbf{J}}(\mathbf{x},u)\,p(\mathbf{x},u)\right)$$

and, taking Fourier transforms of both sides with respect to $\mathbf{x}$, observing that

$$\mathfrak{F}\left(\frac{\partial}{\partial x_\alpha}(gp)\right) = -i\theta_\alpha \mathfrak{F}(gp) = -i\theta_\alpha g\left(\frac{1}{i}\frac{\partial}{\partial\boldsymbol{\theta}}\right)\mathfrak{F}(p)$$

and that

$$\mathfrak{F}\left(\frac{\partial^2}{\partial x_\alpha \partial x_\beta}(gp)\right) = -\theta_\alpha\theta_\beta g\left(\frac{1}{i}\frac{\partial}{\partial\boldsymbol{\theta}}\right)\mathfrak{F}(p),$$

Equation (4.1) follows.

Equation (4.1) has the solution, to order N^{-1}, of normal form,

$$H^{(1)}_{*} = \exp\left\{ i\boldsymbol{\theta}.\boldsymbol{\xi}(u) - \frac{1}{2N}\sum_{\alpha,\beta}\theta_\alpha\theta_\beta\sigma_{\alpha\beta}(u)\right\}$$

for, observing that

$$H^{(1)}_{*} = \exp\{i\boldsymbol{\theta}.\boldsymbol{\xi}(u)\}\left(1 - \frac{1}{2N}\sum_{\alpha,\beta}\sigma_{\alpha\beta}(u)\theta_\alpha\theta_\beta\right) + O(N^{-2})$$

and substituting this for $H^{(1)}$ in (4.1), we obtain, after cancelling $\exp\{i\boldsymbol{\theta}.\boldsymbol{\xi}\}$ on either side,

$$\left(1 - \frac{1}{2N}\sum_{\alpha,\beta}\sigma_{\alpha\beta}\theta_\alpha\theta_\beta\right)\left(i\sum_{\alpha}\dot{\xi}_\alpha\theta_\alpha - \frac{1}{2N}\sum_{\alpha,\beta}\dot{\sigma}_{\alpha\beta}\theta_\alpha\theta_\beta\right) \tag{4.2}$$

$$= \sum_{\mathbf{J}\in\mathscr{J}}\left\{i\mathbf{J}.\boldsymbol{\theta} - (\mathbf{J}.\boldsymbol{\theta})^2\frac{1}{2N}\right\}\left\{g + \frac{1}{2N}\sum_{\alpha,\beta}\sigma_{\alpha\beta}(g_{\alpha\beta} + i\theta_\alpha g_\beta + i\theta_\beta g_\alpha - \theta_\alpha\theta_\beta g)\right\} + O(N^{-2}),$$

where, for example, $g_{\alpha\beta} \equiv \partial^2 g^{\mathbf{J}}(\mathbf{x},u)/\partial x_\alpha \partial x_\beta$ evaluated at

$$\mathbf{x} = \boldsymbol{\xi}(u)\,;\; g \equiv g^{\mathbf{J}}(\boldsymbol{\xi}(u),u)\,;\; \dot{\xi}_\alpha = d\xi_\alpha/du\,.$$

Equating coefficients, we find that $H_*^{(1)}$satisfies (4.1) to terms in N^{-1}, provided that $\xi(u)$ and $(\sigma_{\alpha\beta}(u))$ satisfy the following *Normal Equations*,

$$(\theta_\alpha)\colon\ \dot{\xi}_\alpha = \sum_{J \in \mathscr{J}} J_\alpha \left(g + \frac{1}{2N} \sum_{\gamma,\beta} \sigma_{\gamma\beta} g_{\gamma\beta} \right) + O(N^{-2}), \tag{4.3}$$

$$(\theta_\alpha\theta_\beta)\colon \dot{\sigma}_{\alpha\beta} = \sum_{J \in \mathscr{J}} \left(J_\alpha \sum_\gamma g_\gamma \sigma_{\beta\gamma} + J_\beta \sum_\gamma g_\gamma \sigma_{\alpha\gamma} + J_\alpha J_\beta g \right) + O(N^{-1}), \tag{4.4}$$

for the terms then remaining of large order compared to N^{-2} are

$$\text{L.H.S.} \qquad -\frac{1}{2N} \left(\sum_{\alpha,\beta} \sigma_{\alpha\beta}\theta_\alpha\theta_\beta \right) i \sum_\gamma \dot{\xi}_\gamma \theta_\gamma$$

$$\text{R.H.S.} \qquad \sum_J \left\{ -\frac{1}{2N} \left(\sum_{\alpha,\beta} \sigma_{\alpha\beta}\theta_\alpha\theta_\beta \right) i \sum_\gamma J_\gamma g \theta_\gamma \right\}$$

and, by virtue of (4.3), these are equal (to order N^{-1}). Thus $H_*^{(1)}$ satisfies (4.1) up to terms in N^{-1}.

The initial value $H = \exp\{i\boldsymbol{\theta}.\boldsymbol{x}(0)\}$ implies that $\xi(0) = \boldsymbol{x}(0)$, $(\sigma_{\alpha\beta}(0)) = (0)$, furnishing the initial conditions together with which the Normal Equations are to be solved.

5. About the Normal Equations

The most striking feature of the Normal Equations is the "inevitable" emergence of a normal limiting form, even when there are cases, such as the simple stochastic closed epidemic, satisfying all the assumptions made, in which absorbing boundaries make such a limit seem unlikely. The explanation lies in the choice of the initial conditions $\boldsymbol{X}(0) = N\boldsymbol{x}(0)$ and in the stipulation that the limit holds only for $u \in [0, U]$, where $\boldsymbol{x}(0)$ and U are fixed as $N \uparrow \infty$: these together ensure that N can be chosen so large that the probability of the process hitting a boundary in the fixed u-interval is asymptotically small, unless the process actually started on the boundary, when the correct trivial result is obtained. This reveals the particular weakness of the approach; it cannot directly provide any quantitative information for situations where the boundaries are important, and these are the cases which are stochastically most interesting. It yields, however, a normal approximation to add to the deterministic solution in cases where deterministic solutions are appropriate. For contrasting descriptions of this type of large population approximation, see Whittle [4] and Bailey [2].

It is also possible to deduce, from the solution of the deterministic equations, when the approximation will break down. Define $\xi(0)$ to be a stability point for the deterministic solution if and only if, given any $\varepsilon > 0$, there exists $\delta > 0$ such that any deterministic solution $\xi'(u)$ satisfying $|\xi'(0) - \xi(0)| < \delta$ also satisfies

$|\xi'(u) - \xi(u)| < \varepsilon$ for all $u \geqq 0$, where $\xi(u)$ is the deterministic solution with initial condition $\xi(0)$. Then, if a sequence of initial conditions $x_N(0)$ is chosen for x such that $\lim_{N\uparrow\infty}\{x_N(0)\} = x_\infty(0)$, $x_N(0) \neq x_\infty(0)$, and $x_\infty(0)$ is not a stability point for the deterministic solution, it may be expected that the process will ultimately adopt one or other of the possible characters suggested by the local deterministic solutions with asymptotic probabilities depending in detail on the sequence $x_N(0)$. In many such cases, it will be possible to approximate the behaviour in the neighbourhood of $x_\infty(0)$ in other ways, using the local properties of the transition functions, in order to obtain these asymptotic probabilities, and then use the normal approximation to continue the process until it reaches the neighbourhood of other instability points. This is done for the stochastic rumour in Section 9 in the case $I = 1$; further illustrations of how to extend the use of the technique to accommodate situations in which boundaries are important—in particular, the stochastic epidemic—form the content of subsequent papers, and the martingale treatment in Sections 7 and 8 following will be extensively required therein.

6. Solving the Normal Equations

The Normal Equations (4.3) and (4.4) are of forward predictor form, and as such are particularly amenable to solution by computer: (4.3) gives the next term in the asymptotic expansion of the mean of the process after the deterministic estimate, and (4.4) the leading term in the covariance matrix. Further terms in these asymptotic expansions, and further cumulants of the process, can be evaluated by substituting an appropriate Gram-Charlier expansion for H in an approximation to (2.3) to terms of the required order in N^{-1}: the algebra becomes progressively heavier, and, since all the differential equations derived are coupled, computer solution becomes all that is feasible.

There is in general much more to be gained from an approximate analytic solution of the problem than there is from computer solution, and to this end consider the simplest form of the Normal Equations, sufficient for obtaining the leading terms in the expansions for means and coaariances, in which (4.3) is solved, ignoring the terms in N^{-1}, as the deterministic approximation, and in which (4.4) is subsequently solved using the deterministic values for $\xi(u)$ in the functions $g(\xi(u),u)$. In typical applications, the deterministic equations are usually relatively easily solved, but the equations (4.4), although linear in $(\sigma_{\alpha\beta})$, are coupled, and involve the functions $g(\xi(u),u)$ of u; a direct solution to them appears unlikely. However, there exists a special technique for finding analytically some or all of this solution using only integrals of the deterministic equations; this was first used by Daley and Kendall [3] in 1965, and was called by them the Principle of the Diffusion of Arbitrary Constants.

7. The principle of the diffusion of arbitrary constants

Consider the martingale (backward) equation for the x process: because of the Markov property assumed for the process, this can be expressed in the form

$$\begin{aligned}\phi(x,u+du)\left\{1-\sum_J f^J(X,uN^{1-d})N^{1-d}du\right\}\\ +\sum_J f^J(X,uN^{1-d})N^{1-d}du\,\phi(x+JN^{-1},u+du)\\ =\phi(x,u)+o(du),\end{aligned} \tag{7.1}$$

i.e.,

$$\frac{\partial\phi(x,u)}{\partial u}-N\sum_J\phi(x,u)g^J(x,u)+N\sum_J g^J(x,u)\phi(x+JN^{-1},u)=0. \tag{7.2}$$

Expanding the last term in (7.2) in inverse powers of N, using Taylor's expansion, we obtain

$$\frac{\partial\phi(x,u)}{\partial u}+\sum_J g^J(x,u)\sum_\alpha J_\alpha\frac{\partial\phi(x,u)}{\partial x_\alpha}=O(N^{-1}) \tag{7.3}$$

and so we are led to consider the solutions $h(x,u)$ of the *linearised martingale equation*

$$\frac{\partial h(x,u)}{\partial u}+\sum_J g^J(x,u)\sum_\alpha J_\alpha\frac{\partial h(x,u)}{\partial x_\alpha}=0 \tag{7.4}$$

in the expectation that they will have, asymptotically at least, some properties of a martingale in relation to our stochastic process. We now prove the following theorem.

Theorem. If $h^{(1)}(x,u)$ and $h^{(2)}(x,u)$ are any two solutions of the linearised martingale equation (7.4) (not necessarily distinct), then

$$\begin{aligned}&\mathrm{Cov}(h^{(1)}(x,u),h^{(2)}(x,u))\\ &=\frac{1}{N}\int_0^u\sum_J g^J(\xi(u),u)\left\{\sum_\alpha\frac{\partial h^{(1)}}{\partial\xi_\alpha}J_\alpha\right\}\left\{\sum_\beta\frac{\partial h^{(2)}}{\partial\xi_\beta}J_\beta\right\}du+O(N^{-2}),\end{aligned} \tag{7.5}$$

where $\xi(u)$ represents the deterministic solution to the problem, and is a function of u alone, and where expressions such as $\partial h^{(1)}/\partial\xi_\alpha$ denote the function of u defined by $\partial h(x,u)/\partial x_\alpha$ evaluated at $x=\xi(u)$.

Proof.

$$h^{(i)}(x,u)=h^{(i)}(\xi,u)+(x-\xi).\nabla h^{(i)}(\xi,u)+O(x-\xi)^2, \tag{7.6}$$

by Taylor's expansion, implying that

$$\text{(7.7)}\qquad \operatorname{Cov}\{h^{(1)}(x,u),h^{(2)}(x,u)\} = \frac{1}{N}\sum_{\alpha\beta}\sigma_{\alpha\beta}\frac{\partial h^{(1)}}{\partial\xi_\alpha}\frac{\partial h^{(2)}}{\partial\xi_\beta}+O(N^{-2}).$$

So we have

$$\text{(7.8)}\qquad \frac{d}{du}\operatorname{Cov}\{h^{(1)},h^{(2)}\}$$
$$= \frac{1}{N}\sum_{\alpha,\beta}\left\{\dot\sigma_{\alpha\beta}\frac{\partial h^{(1)}}{\partial\xi_\alpha}\frac{\partial h^{(2)}}{\partial\xi_\beta}+\sigma_{\alpha\beta}\left\{\frac{d}{du}\left(\frac{\partial h^{(1)}}{\partial\xi_\alpha}\right)\frac{\partial h^{(2)}}{\partial\xi_\beta}+\frac{\partial h^{(1)}}{\partial\xi_\alpha}\frac{d}{du}\left(\frac{\partial h^{(2)}}{\partial\xi_\beta}\right)\right\}\right\}+O(N^{-2}),$$

in which the coefficient of $\sigma_{\alpha\beta}/N$ is, to the first order,

$$\text{(7.9)}\qquad \begin{aligned}&\Big\{\sum_J\sum_\gamma J_\gamma\Big(\frac{\partial g^J}{\partial\xi_\alpha}\frac{\partial h^{(2)}}{\partial\xi_\beta}\frac{\partial h^{(1)}}{\partial\xi_\gamma}+\frac{\partial g^J}{\partial\xi_\beta}\frac{\partial h^{(1)}}{\partial\xi_\alpha}\frac{\partial h^{(2)}}{\partial\xi_\gamma}\Big)\\&\quad+\sum_J\sum_\gamma J_\gamma g^J\Big(\frac{\partial h^{(2)}}{\partial\xi_\beta}\frac{\partial^2 h^{(1)}}{\partial\xi_\alpha\partial\xi_\gamma}+\frac{\partial h^{(1)}}{\partial\xi_\alpha}\frac{\partial^2 h^{(2)}}{\partial\xi_\beta\partial\xi_\gamma}\Big)\\&\quad+\frac{\partial h^{(2)}}{\partial\xi_\beta}\frac{\partial^2 h^{(1)}}{\partial\xi_\alpha\partial u}+\frac{\partial h^{(1)}}{\partial\xi_\alpha}\frac{\partial^2 h^{(2)}}{\partial\xi_\beta\partial u}\Big\}+\left\{\begin{array}{l}\text{same expression with (1)}\\ \text{and (2) interchanged}\end{array}\right\},\end{aligned}$$

(expanding $\dot\sigma_{\alpha\beta}$ according to Equation (4.4); and setting

$$\frac{d}{du}\Big(\frac{\partial h^{(i)}}{\partial\xi_\alpha}\Big)=\sum_\gamma\frac{\partial^2 h^{(i)}}{\partial\xi_\alpha\partial\xi_\gamma}\dot\xi_\gamma+\frac{\partial^2 h^{(i)}}{\partial\xi_\alpha\partial u},\ \text{using (4.3) to expand } \dot\xi_\gamma)$$

which can be regrouped as

$$\text{(7.10)}\qquad \begin{aligned}&\Big\{\Big(\frac{\partial h^{(1)}}{\partial\xi_\alpha}\frac{\partial}{\partial\xi_\beta}+\frac{\partial h^{(1)}}{\partial\xi_\beta}\frac{\partial}{\partial\xi_\alpha}\Big)\Big(\frac{\partial h^{(2)}}{\partial u}+\sum_\gamma\frac{\partial h^{(2)}}{\partial\xi_\gamma}\sum_J J_\gamma g^J(\xi(u),u)\Big)\Big\}\\&\qquad+\{\text{same expression with (1) and (2) interchanged}\},\end{aligned}$$

which is zero, from (7.4), since $h^{(i)}(x,u)$ were chosen to satisfy the linearised martingale equation.

Thus

$$\text{(7.11)}\qquad \frac{d}{du}\operatorname{Cov}\{h^{(1)},h^{(2)}\}=\frac{1}{N}\sum_J g^J(\xi(u),u)\sum_{\alpha\beta}J_\alpha J_\beta\frac{\partial h^{(1)}}{\partial\xi_\alpha}\frac{\partial h^{(2)}}{\partial\xi_\beta}+O(N^{-2}),$$

the terms in (7.8) independent of $(\sigma_{\alpha\beta})$, arising from the expansion of $\dot\sigma_{\alpha\beta}$ according to Equation (4.4), and the theorem follows by integration.

Expression (7.5) provides a means of evaluating the covariance of any pair of random functions $h^{(1)}(x,u)$ and $h^{(2)}(x,u)$, where each one satisfies (7.4) using only known functions of the deterministic solution $\xi(u)$: the disappearance of the terms in $\sigma_{\alpha\beta}$, a consequence of the fact that $h^{(i)}(x,u)$ is asymptotically martingale, means that the Equations (4.4) *need not be solved.*

Furthermore, if r functionally independent solutions to (7.4) are known, an $(r \times r)$ matrix V_{ij} can be evaluated without reference to (4.4), where

$$V_{ij} = \mathrm{Cov}\{h^{(i)}(\boldsymbol{x},u), h^{(j)}(\boldsymbol{x},u)\}. \tag{7.12}$$

But, from Taylor's expansion, we have

$$V_{ij} = \frac{1}{N} \sum_{\alpha,\beta} \sigma_{\alpha\beta} \frac{\partial h^{(i)}}{\partial \xi_\alpha} \frac{\partial h^{(j)}}{\partial \xi_\beta} + O(N^{-2}), \tag{7.13}$$

and so, if we can find n functionally independent solutions to (7.4), we can invert this set of linear equations to obtain $\sigma_{\alpha\beta}$ in terms of the (already evaluated) V_{ij}, without directly solving (4.4), provided merely that the $(\frac{1}{2}n(n+1), \frac{1}{2}n(n+1))$ matrix of coefficients of $(\sigma_{\alpha\beta})$ in (7.13) is invertible; i.e., provided that the functions h have a non-zero jacobian determinant along the deterministic trajectory.

Equation (7.4) is a first order homogeneous linear partial differential equation in $(n+1)$ independent variables $\{x_1, \cdots, x_n; u\}$, and so will in general possess n functionally independent particular solutions, of which the general solution is an arbitrary function; from standard theory, the solutions of (7.4) are the integrals of the auxiliary equations associated with (7.4), and these are

$$du = dx_\alpha \Big(\sum_J J_\alpha g^J(\boldsymbol{x},u) \Big)^{-1} \qquad (\alpha = 1, \cdots, n) \tag{7.14}$$

which are precisely the deterministic equations. Thus, if n functionally independent integrals of the deterministic equations can be found explicitly, we can use the matrix V_{ij} to derive the $\sigma_{\alpha\beta}$ without ever solving (4.4) directly, using only integrals and solutions of the deterministic equations, provided Equations (7.13) are invertible.

The consideration of the invertibility of (7.13) elicits some important aspects of the character of the solutions to (7.4). If ϕ_0 is an integral of (7.4), then so is any function of ϕ_0, and in particular ϕ_0^2. Now every solution to (7.4) is constant on the deterministic path, and we can freely add a constant to it. Suppose ϕ_0 is chosen to be zero on the deterministic path; then, from expression (7.5), the variance of ϕ_0^2 has leading term $4\phi_0^2 \operatorname{var} \phi_0$ which is zero, since $\phi_0 = 0$ on the deterministic path. Although ϕ_0^2 is an integral of (7.4), it does not serve our purpose for finding the $\sigma_{\alpha\beta}$, as its every partial differential is zero on the deterministic path, and so (7.13) is not invertible (the jacobian has a row of zeros). This leads to the question of how we can be sure to discover suitable solutions to (7.4). Considering the solution of the partial differential equation (7.4) in terms of characteristic coordinates, the solutions of (7.4) may be represented by an arbitrary function, defined on an initial condition n-manifold which is nowhere characteristic, which is extended throughout $(n+1)$ space by the characteristics; the condition that the jacobian in (7.13) does not have a row of

zeros is then equivalent to demanding that the arbitrary function does not have a stationary value in the initial condition manifold at the point at which the deterministic path (a particular characteristic) intersects it, and it is clearly possible (except perhaps in degenerate cases) to choose the arbitrary function so that this condition is satisfied. Call the solution corresponding to an arbitrary function satisfying this condition *standard* if, in addition, it is chosen to be zero on the deterministic path: the importance of the powers and products of standard integrals will emerge in later sections.

The title "The Principle of the Diffusion of Arbitrary Constants" derives from the representation of the deterministic solution as

$$h^{(i)} = c_i, \qquad (i = 1,2,\cdots,n)$$

where the $h^{(i)}$ are any n functionally independent integrals of the deterministic equations, and the 'arbitrary' constants c_i are fixed by the initial conditions. In the stochastic formulation, the $h^{(i)}$, instead of remaining constant, are (asymptotically) martingale, and the process may be described by the stochastic variation of (as it were) the c_i, which will have an asymptotically diffusive behaviour for large N.

8. An alternative approach using martingales

Another approach to the problem, which extends the possibilities for approximation, is to solve the martingale equation (7.2) by successive approximation. We have the exact equation

$$\frac{\partial\phi(\boldsymbol{x},u)}{\partial u} - N \sum_{\boldsymbol{J}} g^{\boldsymbol{J}}(\boldsymbol{x},u)\phi(\boldsymbol{x},u) + N \sum_{\boldsymbol{J}} g^{\boldsymbol{J}}(\boldsymbol{x},u)\phi(\boldsymbol{x}+\boldsymbol{J}N^{-1},u) = 0 \tag{8.1}$$

whence, expanding by Taylor's series in inverse powers of N,

$$\frac{\partial\phi}{\partial u} + \sum_{\boldsymbol{J}} \sum_{\alpha} J_\alpha g^{\boldsymbol{J}} \frac{\partial\phi}{\partial x_\alpha} + \frac{1}{2N} \sum_{\boldsymbol{J}} \sum_{\alpha,\beta} J_\alpha J_\beta g^{\boldsymbol{J}} \frac{\partial^2\phi}{\partial x_\alpha \partial x_\beta} + \frac{1}{6N^2} \sum_{\boldsymbol{J}} \sum_{\alpha,\beta,\gamma} J_\alpha J_\beta J_\gamma g^{\boldsymbol{J}} \frac{\partial^3\phi}{\partial x_\alpha \partial x_\beta \partial x_\gamma} + \cdots = 0. \tag{8.2}$$

Writing the solution ϕ in the form $\phi = \phi_0 + N^{-1}\phi_1 + N^{-2}\phi_2 + \cdots$, we obtain a solution if

$$\text{(8.3) (terms independent of } N\text{):} \quad \frac{\partial\phi_0}{\partial u} + \sum_{\boldsymbol{J}} \sum_{\alpha} J_\alpha g^{\boldsymbol{J}} \frac{\partial\phi_0}{\partial x_\alpha} = 0,$$

$$\text{(8.4) (terms in } N^{-1}\text{):} \quad \frac{\partial\phi_1}{\partial u} + \sum_{\boldsymbol{J}} \sum_{\alpha} J_\alpha g^{\boldsymbol{J}} \frac{\partial\phi_1}{\partial x_\alpha} = -\tfrac{1}{2} \sum_{\boldsymbol{J}} \sum_{\alpha,\beta} J_\alpha J_\beta g^{\boldsymbol{J}} \frac{\partial^2\phi_0}{\partial x_\alpha \partial x_\beta},$$

$$\text{(terms in } N^{-2}\text{):}\quad \frac{\partial \phi_2}{\partial u} + \sum_{\boldsymbol{J}} \sum_{\alpha} J_\alpha g^{\boldsymbol{J}} \frac{\partial \phi_2}{\partial x_\alpha} = -\tfrac{1}{6} \sum_{\boldsymbol{J}} \sum_{\alpha,\beta,\gamma} J_\alpha J_\beta J_\gamma g^{\boldsymbol{J}} \frac{\partial^3 \phi_0}{\partial x_\alpha \partial x_\beta \partial x_\gamma}$$

$$- \tfrac{1}{2} \sum_{\boldsymbol{J}} \sum_{\alpha,\beta} J_\alpha J_\beta g^{\boldsymbol{J}} \frac{\partial^2 \phi_1}{\partial x_\alpha \partial x_\beta} \tag{8.5}$$

and so on, where the equations can be solved in turn for $\phi_0, \phi_1, \phi_2, \cdots$ each of them either the homogeneous or an inhomogeneous version of the linear equation (7.4); i.e., Lagrangian in form. Thus the auxiliary equations for the solution are the deterministic equations, together with an auxiliary equation involving ϕ_i which may be written as

$$\frac{d\phi_i}{du} = \{\text{R.H.S.}\}, \tag{8.6}$$

where the term in braces is the right-hand side of the equation for ϕ_i, and is evaluated in terms of $\boldsymbol{x}$ and u from the previous solution of the equations for $\phi_0, \phi_1, \cdots, \phi_{i-1}$.

Using this method of solution for (8.2), we may start with any integral ϕ_0 of (8.3) and by successive approximation through Equations (8.4), (8.5),$\cdots$ obtain a corresponding solution of (8.2): similarly, we may take any solution of (8.2) and, by expanding it in inverse powers of N, obtain an integral of (8.3) as leading term. This establishes a 1-1 correspondence between the solutions of (8.2) and the solutions of (8.3), and since the solution space of (8.3) may be parameterised by an n-dimensional manifold of initial conditions, it follows that there must exist an equivalent parameterisation for the solution space of (8.2).

The martingales of the process need only satisfy (8.2) when $N\boldsymbol{x} \in \boldsymbol{Z}^n$, since they are not defined for states which the process cannot assume, and it is clear from the unexpanded form (8.1) of (8.2) that its general solution can be written as

$$\phi(\boldsymbol{x}, u) = \phi_{\boldsymbol{\omega}}(\boldsymbol{r}N^{-1}, u), \tag{8.7}$$

where $\boldsymbol{\omega}(\boldsymbol{x})$ is uniquely determined by

$$\{\boldsymbol{x} = N^{-1}(\boldsymbol{r} + \boldsymbol{\omega});\ \boldsymbol{\omega} \in [0,1]^n;\ \boldsymbol{r} \in \boldsymbol{Z}^n\}, \tag{8.8}$$

and where, for each distinct $\boldsymbol{\omega}, \phi_{\boldsymbol{\omega}}$ is an arbitrary martingale of the process. Thus many of the different solutions of (8.2) are identical when restricted to the feasible states $(\boldsymbol{x}, u) \in N^{-1}\boldsymbol{Z} \times \boldsymbol{R}^+$, since the restriction identifies all solutions of (8.2) with common $\phi_{\boldsymbol{o}}$, but each different solution of (8.2) is martingale under this restriction. In the rest of the section, we shall pass from solutions of (8.3) to closer approximate solutions of (8.2), refer to them as *approximately martingale*, and treat them as functions of a continuous variable $\boldsymbol{x}$: the underlying justification for this procedure is inherent in the preceding paragraphs.

Before proceeding further, we pause to note that, if $\phi(\boldsymbol{x},u)$ is a function satisfying suitable regularity conditions, then

$$E\{\phi(\boldsymbol{x}(u),u)\} = E\{\phi(\xi(u),u) + \sum_{\alpha}(x_\alpha - \xi_\alpha)\partial\phi/\partial\xi_\alpha + \sum_{\alpha,\beta}(x_\alpha-\xi_\alpha)(x_\beta-\xi_\beta)\partial^2\phi/\partial\xi_\alpha\partial\xi_\beta + \cdots\} \tag{8.9}$$

by Taylor's expansion, and so, taking the expectation term by term,

$$E\{\phi(\boldsymbol{x}(u),u)\} = \phi(\xi(u),u) + O(N^{-1}) \tag{8.10}$$

from (3.2). This expansion is used extensively in what follows, and it is a measure of the heuristic nature of the paper that in Section 9 the conclusion is successfully drawn when, strictly speaking, the function ϕ thereby derived has infinite expectation, although $\phi(\xi(u),u)$ is everywhere finite on the deterministic path. We leave this without any detailed justification, remarking only that if, for a given sequence of initial conditions, the deterministic path approaches a point where one of the functions ϕ used becomes infinite, the method may be expected to break down, whereas, if this is not the case, the difficulty, as in Section 9, can probably be circumvented by means specific to the particular problem.

The optional stopping theorem for martingales tells us that, if ϕ is a martingale, then

$$E\{\phi(\boldsymbol{x},u)\,|\,u = \text{Markov time}\} = \phi(\boldsymbol{x}(0),0). \tag{8.11}$$

If we write ϕ in terms of its asymptotic expansion in inverse powers of N, and substitute in either side, we obtain

$$\sum_{r=0}^{R} E\{\phi_r(\boldsymbol{x},u)\,|\,u = \text{Markov time}\}N^{-r} = \sum_{r=0}^{R}\phi_r(\boldsymbol{x}(0),0)N^{-r} + O(N^{-R-1}) \text{ for any } R\in\boldsymbol{Z}^+; \tag{8.12}$$

we shall sketch an extension of Section 7 from this standpoint.

Suppose that ϕ_0 is a standard integral of (7.4): then applying (8.12) with $R = 0$ for the stopping rule $u = u_0$, we obtain

$$E\{\phi_0(\boldsymbol{x}(u_0),u_0)\} = \phi_0(\boldsymbol{x}(0),0) + O(N^{-1}) = O(N^{-1}). \tag{8.13}$$

Since ϕ_0 is an integral of (7.4), so also is ϕ_0^2: construct from (8.4) a function ϕ_1, such that $\{\phi_0 + N^{-1}\phi_1\}$ is the two-term expansion of the martingale corresponding to ϕ_0^2. Applying (8.12) with $R = 1$ for the stopping rule $u = u_0$, we obtain

$$E\{\phi_0^2(\boldsymbol{x}(u_0),u_0) + \phi_1(\boldsymbol{x}(u_0),u_0)N^{-1}\} = N^{-1}\phi_1(\boldsymbol{x}(0),0) + O(N^{-2}), \tag{8.14}$$

and so

$$\text{var}\,\phi_0 = N^{-1}\{\phi_1(\boldsymbol{x}(0),0) - E\phi_1(\boldsymbol{x}(u_0),u_0)\} + O(N^{-2}), \tag{8.15}$$

because, from (8.13),

$$(E\{\phi_0(\boldsymbol{x},u_0)\})^2 = O(N^{-2});$$

thus, using (8.10),

$$\begin{aligned}\text{var}\,\phi_0 &= N^{-1}\{\phi_1(\boldsymbol{x}(0),0) - \phi_1(\xi(u_0),u_0)\} + O(N^{-2}) \\ &= -\frac{1}{N}\int_0^{u_0}\frac{d\phi_1(\xi(u),u)}{du}du + O(N^{-2}).\end{aligned} \tag{8.16}$$

But from (8.4), and its auxiliary equation (8.6) connecting ϕ_1 and u

$$\begin{aligned}\frac{d\phi_1}{du} &= -\tfrac{1}{2}\sum_J\sum_{\alpha,\beta}J_\alpha J_\beta g^J\frac{\partial^2\{\phi_0{}^2\}}{\partial x_\alpha\partial x_\beta} \\ &= -\sum_J\sum_{\alpha,\beta}J_\alpha J_\beta g^J\left\{\frac{\partial\phi_0}{\partial x_\alpha}\frac{\partial\phi_0}{\partial x_\beta} + \phi_0\frac{\partial^2\phi_0}{\partial x_\alpha\partial x_\beta}\right\},\end{aligned}$$

and so

$$\frac{d\phi_1(\xi(u),u)}{du} = -\sum_J g^J(\xi(u),u)\left\{\sum_\alpha J_\alpha\frac{\partial\phi_0}{\partial\xi_\alpha}\right\}^2 \tag{8.17}$$

since

$$\phi_0(\xi(u),u) = 0 \qquad (\phi_0 \text{ standard}).$$

If we substitute this in (8.16), we recover the variance case of Formula (7.5): an equivalent calculation with $\{\phi_0^{(1)}\phi_0^{(2)}\}$ for $\phi_0^{(1)},\phi_0^{(2)}$ distinct standard integrals would establish the rest of (7.5).

We can now proceed to closer approximations, using the results already established. If ϕ_0 is a standard integral of (7.4), construct from (8.4) a function ϕ_1 such that $\{\phi_0 + N^{-1}\phi_1\}$ is the two-term expansion of the martingale corresponding to ϕ_0. Applying (8.12) with $R = 1$ for the stopping rule $u = u_0$, we obtain

$$E\{\phi_0(\boldsymbol{x}(u_0),u_0) + N^{-1}\phi_1(\boldsymbol{x}(u_0),u_0)\} = N^{-1}\phi_1(\boldsymbol{x}(0),0) + O(N^{-2}),$$

and so

$$\begin{aligned}E\{\phi_0(\boldsymbol{x}(u_0),u_0)\} &= -\frac{1}{N}\int_0^{u_0}\frac{d\phi_1(\xi(u),u)}{du}du + O(N^{-2}) \\ &= \frac{1}{2N}\int_0^{u_0}\sum_J\sum_{\alpha,\beta}J_\alpha J_\beta g^J(\xi(u),u)\frac{\partial^2\phi_0}{\partial\xi_\alpha\partial\xi_\beta}du + O(N^{-2}),\end{aligned} \tag{8.18}$$

from (8.10) and (8.6).

This provides the next term in the expectation of ϕ_0, and can be converted to provide the N^{-1} terms in $E\{\boldsymbol{x}(u_0)\}$ by observing that

$$E\{\phi_0(\boldsymbol{x}(u_0),u_0)\} = \sum_\alpha E(x_\alpha(u_0) - \xi_\alpha(u_0))\partial\phi_0/\partial\xi_\alpha + \tfrac{1}{2}\sum_{\alpha,\beta} E\{(x_\alpha - \xi_\alpha)(x_\beta - \xi_\beta)\}\partial^2\phi_0/\partial\xi_\alpha\partial\xi_\beta + O(N^{-2}). \tag{8.19}$$

Note that the L.H.S. and the expectation product in the second term of the R.H.S. have already been evaluated to the required accuracy. Here again, a matrix equation must be inverted in order to get the more precise estimates of $E\{\boldsymbol{x}(u_0)\}$.

If we now start with ϕ_0^3, where ϕ_0 is a standard integral of (7.4), evaluate the corresponding three-term martingale approximation $\{\phi_0^3 + N^{-1}\phi_1 + N^{-2}\phi_2\}$, and apply (8.12) with $R = 2$ for the stopping rule $u = u_0$, we obtain, using (8.10)

$$E\{\phi_0^3(\boldsymbol{x}(u_0),u_0)\} = -\frac{1}{N^2}\int_0^{u_0}\frac{d\phi_2(\xi(u),u)}{du}du + \frac{1}{N}\{\phi_1(\boldsymbol{x}(0),0) - E\{\phi_1(\boldsymbol{x}(u_0),u_0)\}\} + O(N^{-3}). \tag{8.20}$$

Now, from (8.6) and because ϕ_0 is standard,

$$\frac{d\phi_2(\xi(u),u)}{du} = -\tfrac{1}{6}\sum_J\sum_{\alpha,\beta,\gamma} J_\alpha J_\beta J_\gamma g^J \frac{\partial^3\{\phi_0^3\}}{\partial\xi_\alpha\partial\xi_\beta\partial\xi_\gamma} - \tfrac{1}{2}\sum_J\sum_{\alpha,\beta} J_\alpha J_\beta g^J \frac{\partial^2\phi_1}{\partial\xi_\alpha\partial\xi_\beta}$$

$$= -\sum_J g^J\Big(\sum_\alpha J_\alpha\frac{\partial\phi_0}{\partial\xi_\alpha}\Big)^3 - \tfrac{1}{2}\sum_J\sum_{\alpha,\beta} J_\alpha J_\beta g^J\frac{\partial^2\phi_1}{\partial\xi_\alpha\partial\xi_\beta} + O(N^{-1}), \tag{8.21}$$

where ϕ_1 must now be found explicitly as an integral of (8.4) (with ϕ_0^3 in the place of ϕ_0), and to find $E\{\phi_1(\boldsymbol{x}(u_0),u_0)\}$ to the required accuracy we must use

$$E\{\phi_1(\boldsymbol{x}(u_0),u_0)\} = \phi_1(\xi(u_0),u_0) + \sum_\alpha E\{x_\alpha(u_0) - \xi_\alpha(u_0)\}\partial\phi_1/\partial\xi_\alpha + \tfrac{1}{2}\sum_{\alpha,\beta} E\{(x_\alpha - \xi_\alpha)(x_\beta - \xi_\beta)\}\partial^2\phi_1/\partial\xi_\alpha\partial\xi_\beta + O(N^{-2}), \tag{8.22}$$

with $E(x_\alpha(u_0) - \xi_\alpha(u_0))$ evaluated to terms of order N^{-1} using (8.18) and (8.19). We can then estimate the third cumulants of the stochastic path.

The process can be continued, in principle, using higher powers of standard integrals and closer martingale approximations, to obtain further asymptotic terms in the expansions of the cumulants of $\boldsymbol{x}(u_0)$; and stopping rules other than $u = u_0$ can be employed, providing they do not distort the distribution to the extent that the error in replacing $E\{\phi(\boldsymbol{x}(S),S)\}$ by $\phi(\xi(s),s)$ is no longer $O(N^{-1})$, where S is the Markov time defined by the stopping rule and s the time at which

the deterministic process stops under this rule. The example in the next section illustrates the use of the technique in the case of the stochastic rumour; important information is derived, *even though only one of a possible two independent standard integrals can be found explicitly.*

9. Example: the stochastic rumour in a closed population

For a detailed description of the model, see Daley and Kendall [3]: a summary sufficient for this calculation is, in the notation of Section 2: $n = 2$; $(X, Y) =$ (number of ignorant, number of spreaders) with initial condition (N, I); $\mathscr{J} = \{(-1,1),(0,-1),(0,-2)\}$ with $f^{\mathbf{J}}(X,Y) = \{XY, Y(N+I-X-Y), \frac{1}{2}Y(Y-1)\}$, and so $d = 2$.

Consider the situation in which $I = kN$ and $N \uparrow \infty$: from Section 2 we are to take $u = N^{d-1}t = Nt$. To find the leading terms in the expansions of means and variances, we can write

$$g^{\mathbf{J}}(x,y,u) = g^{\mathbf{J}}(x,y) = \{xy, y(1+k-x-y), \tfrac{1}{2}y^2\}$$

neglecting the small order term $-y/2N$ in the third transition (which will be required later when a more precise approximation to $Ex(\infty)$ is derived). We then obtain the deterministic equations

$$\dot{\xi} = -\xi\eta, \tag{9.1}$$

$$\dot{\eta} = \xi\eta - \eta(1+k-\xi-\eta) - \eta^2 = \eta(2\xi-1-k), \tag{9.2}$$

which have only one integral expressible in elementary functions, namely $h(\xi,\eta) = \eta + 2\xi - (1+k)\log\xi$. The absence of the second independent integral prevents a full asymptotic solution by this method, but we can use the one we have to approximate the distributions of the proportion of people who, originally ignorant, remain so for ever more.

From expression (7.5), we have

$$\operatorname{var}\{h(x(U),y(U))\} = \frac{1}{N}\int_0^U \{\xi\eta(\xi^{-1}(1+k)-1)^2 + \eta(1+k-\xi-\eta)(-1)^2 + \tfrac{1}{2}\eta^2(-2)^2\}du + O(N^{-2}),$$

where $\xi(u)$, $\eta(u)$ are the deterministic solutions of (9.1) and (9.2). Using $\eta du = -d\xi/\xi$ from (9.1), we obtain

$$\begin{aligned}\operatorname{var} h &= \frac{1}{N}\int_{\xi(U)}^1 \left\{\left(\frac{1+k}{\xi}-1\right)^2 - 3 + \frac{3+2k}{\xi} + \frac{(1+k)\log\xi}{\xi}\right\}d\xi + O(N^{-2}) \\ &= \frac{1}{N}\left[-\xi^{-1}(1+k)^2 - 2\xi + \log\xi + \tfrac{1}{2}(1+k)\log^2\xi\right]_{\xi(U)}^1 + O(N^{-2}).\end{aligned} \tag{9.3}$$

In the case where $U \uparrow \infty$, $\eta(U) \downarrow 0$; and since $h(\xi,\eta) = \eta + 2\xi - (1+k)\log\xi = 2 + k$ everywhere on the deterministic path, we have $\xi(U) \downarrow \theta$, where θ is the root in $[0,1)$ of

$$2\theta - (1+k)\log\theta = 2 + k. \tag{9.4}$$

Furthermore, $y(\infty) = 0$ with probability 1, so that

$$\operatorname{var} h(x(\infty),y(\infty)) = (2 - \theta^{-1}(1+k))^2 \operatorname{var}(x(\infty)) + O(N^{-2});$$

hence, using (9.3) and (9.4) for $U \uparrow \infty$,

$$\operatorname{var}(x(\infty)) = \frac{1}{N}\left(\frac{\theta}{2\theta-1-k}\right)^2\left\{2\theta + \frac{(1+k)^2}{\theta} + \frac{2+k-2\theta}{1+k} - \frac{1+k}{2}\left(\frac{2+k-2\theta}{1+k}\right)^2 - 2 - (1+k)^2\right\} + O(N^{-2}),$$

and so

$$\operatorname{var}(x(\infty)) = \frac{1}{N}\left(\frac{\theta}{2\theta-1-k}\right)^2\left\{\frac{(1+k)^2}{\theta} - \frac{k(k+2)}{2(k+1)} - 2 - (k+1)^2 + 4\theta - \frac{2\theta^2}{1+k}\right\} + O(N^{-2}). \tag{9.5}$$

Of particular interest is the case when k is fixed $\ll 1$: the proportion of the initially ignorant populace which never learns of the rumour is then approximately distributed as

$$\mathcal{N}\left(\theta, \frac{1}{N}\frac{\theta}{(2\theta-1)^2}\{1 - 3\theta + 4\theta^2 - 2\theta^3\}\right) \tag{9.6}$$

where θ is the smaller positive root of $\log\theta = 2(\theta - 1)$. Since $k = 0$ is an instability point for the solution of the deterministic equations, we cannot assert directly that this approximation is valid for, say, $k = N^{-1}$ (i.e., $I = 1$) as $N \uparrow \infty$: in fact, a more detailed examination of the early behaviour of the process indicates that the limit in distribution holds in this case, even though the actual variance of the process does not converge to the suggested limiting form because of the probability $(2N-1)^{-1}$ of the rumour finishing with $X = N-1$.

In order to adopt the method of Section 8, to find a closer approximation to the mean of the ultimate distribution, we must include the neglected term in the third transition rate, and solve initially the auxiliary equations

$$\dot{x} = -xy, \tag{9.7}$$

$$\dot{y} = xy - y(1+k-x-y) - y^2 + yN^{-1} = y(2x - \{1 + k - N^{-1}\}): \tag{9.8}$$

again, there is only one suitable integral,

$$\phi_0(x,y) \equiv y + 2x - \{1 + k - N^{-1}\}\log x.$$

Applying the asymptotic form of the optional stopping theorem to ϕ_0 on the boundary $y = 0$, we obtain

$$E\{y + 2x - (1 + k - N^{-1})\log x \mid y = 0\} = 2 + k + O(N^{-1}), \tag{9.9}$$

and so

$$Ex = \theta + O(N^{-1}), \text{ where } 2\theta - (1 + k)\log\theta = 2 + k. \tag{9.10}$$

This is the usual deterministic approximation to the mean, and to obtain a more precise result we require a closer approximation to a martingale than that afforded by ϕ_0.

Accordingly, following Section 8, we take as a next approximation $\phi_0 + N^{-1}\phi_1$, where $\phi_1(x, y)$ satisfies the inhomogeneous linear partial differential equation with auxiliary equations (9.7) and (9.8) together with

$$\frac{d\phi_1}{du} = -\tfrac{1}{2}\sum_J \sum_{\alpha,\beta} J_\alpha J_\beta g^J \frac{\partial^2\phi_0}{\partial x_\alpha \partial x_\beta} = -\tfrac{1}{2}xy(1 + k - N^{-1})x^{-2}. \tag{9.11}$$

Using (9.7) we can write this as

$$d\phi_1 = dx(1 + k - N^{-1})\tfrac{1}{2}x^{-2},$$

and obtain an integral

$$\phi_1 = -(1 + k - N^{-1})\tfrac{1}{2}x^{-1}. \tag{9.12}$$

Because of the small but finite probability that, for any given N, $x(\infty)=0$, the expectation of $\phi_1(\infty)$ will, strictly speaking, be infinite. However, the addition of an absorbing barrier along $x = \varepsilon$, for ε fixed and strictly smaller than θ, will asymptotically in N have only exponentially small effect on the process, and will not alter the calculations that follow, except in justifying the use of (8.10) implicit in the derivation of ϕ_1 and in the expansion of the stopping expectations.

Applying the asymptotic form of the optional stopping theorem to $\phi_0 + N^{-1}\phi_1$, on the boundary $y = 0$, we obtain

$$\begin{aligned} &E\{y + 2x - (1 + k - N^{-1})\log x - (1 + k - N^{-1})/2Nx \mid y = 0\} \\ &\qquad = 2 + k - (1 + k - N^{-1})\tfrac{1}{2}N^{-1} + O(N^{-2}) \end{aligned} \tag{9.13}$$

and hence, putting $x = \theta + z$,

$$\begin{aligned} &2\theta + 2Ez - (1 + k - N^{-1})(\log\theta + Ez\theta^{-1} - Ez^2\tfrac{1}{2}\theta^{-2} + \cdots) - \frac{1}{2N}\left(\frac{1 + k - N^{-1}}{\theta} + \cdots\right) \\ &\qquad = 2 + k - (1 + k - N^{-1})\tfrac{1}{2}N^{-1} + O(N^{-2}), \end{aligned} \tag{9.14}$$

where the expectation has been expanded using Taylor's series, and where the neglected moments of z would introduce only terms of order N^{-2} or less (cf. Section 4). Thus, regrouping, we have

$$Ez(2-\theta^{-1}(1+k)) = \frac{1}{N}\{\tfrac{1}{2}(1+k)(\theta^{-1}-1)-\log\theta\} - \tfrac{1}{2}(1+k)\theta^{-2}Ez^2 + O(N^{-2}). \tag{9.15}$$

Since $Ez = O(N^{-1})$, we have $Ez^2 = \operatorname{var} x + O(N^{-2})$, and, substituting the leading term already derived for $\operatorname{var} x$, we deduce that

$$\begin{aligned} Ex(\infty) &= \theta + Ez \\ &= \theta + \frac{\theta}{N(2\theta-1-k)}\left\{\tfrac{1}{2}(1+k)(\theta^{-1}-1)-\log\theta - \frac{(1+k)}{2(2\theta-1-k)^2}\right. \\ &\qquad \left.\times\,[(1+k)^2(\theta^{-1}-1)+2(\theta-1)-\log\theta-\tfrac{1}{2}(1+k)\log^2\theta]\right\} + O(N^{-2}). \end{aligned} \tag{9.16}$$

As an example of how further cumulants can be estimated using Section 8, it was intended to exhibit also an explicit expression for the leading (N^{-2}) term in the third cumulant. The algebra, however, became so heavy that the formula obtained was of no intuitive value, and we omit it here, although it is used to provide the estimate of the third cumulant compared with the true value by computer evaluation.

A computer evaluation of the effectiveness of these approximations yielded the results shown in Tables 1 and 2 for the distribution of $x(\infty)$.

The differential equation estimates in Table 1 were obtained by iterating the Normal Equations, and true values for $Ex(\infty)$, $N\operatorname{var} x(\infty)$ and $N^2\kappa_3(x(\infty))$ were found exactly by solving the backward recurrence equations for the corresponding discrete time random walk. All estimates converge as predicted.

The results in Table 2, where the chosen sequence of initial conditions approaches the instability point $k = 0$ of the deterministic equations, exhibit the convergence in distribution to (9.6), as illustrated by the progressing accuracy of the estimates of 5% and 95% points (despite the irregularity caused by the true discrete nature of the distribution, which dies out only as $N^{-\frac{1}{2}}$), together with the failure of the variance estimate: the 'improved martingale estimate' (9.16) also fails to converge more rapidly than the deterministic estimate, as is to be expected.

In order to recover estimates for $Ex(\infty)$ and $N\operatorname{var}x(\infty)$ to the required N^{-1} accuracy, in the case $I = 1$, we can follow the procedure suggested in Section 5, and under these special circumstances split the distribution of $x(\infty)$ into two parts

$$\begin{aligned} &1)\quad x_1(\infty) = 1 - N^{-1} \quad \text{with probability } \tfrac{1}{2}N^{-1} + O(N^{-2}), \\ &2)\quad x_2(\infty) = \theta + z \qquad \text{with probability } 1 - \tfrac{1}{2}N^{-1} + O(N^{-2}), \end{aligned} \tag{9.17}$$

where $Ez = O(N^{-1})$ and θ is the deterministic solution.

We will now use the approach of Section 8 through the optional stopping theorem, observing that the asymptotic expansions of the martingales are valid

TABLE 1

$k = 0.1$

N	20	40	80	160	320	640
$Ex(\infty)$	0.258160	0.227280	0.223545	0.222618	0.222222	0.222037
% errors in estimates						
Deterministic, (9.4)	−14.06	−2.38	−0.75	−0.34	−0.16	−0.08
Improved M. G. (9.16)	−11.92	−1.17	−0.13	−0.03	−0.007	−0.002
Differential Eqn. (4.3)	−10.85	−0.96	−0.09	−0.02	−0.005	−0.001
N var $x(\infty)$	0.785835	0.412457	0.352738	0.343706	0.340034	0.338336
% errors in estimates						
Martingale, (9.5)	−57.15	−18.36	−4.54	−2.03	−0.97	−0.48
Differential Eqn. (4.4)	−54.30	−16.50	−3.61	−1.59	−0.76	−0.37
$N^2\kappa_3(x(\infty))$	5.87324	2.12131	0.940546	0.850194	0.819865	0.806400
% error in estimate						
Martingale	−85.23	−61.06	−13.99	−5.77	−2.74	−1.34

TABLE 2

$I = 1$ (Here $k = N^{-1} \downarrow 0$ as $N \uparrow \infty$, and $k = 0$ is an instability point of the deterministic equations)

N	10	20	40	80	160	320	640
$E\,x\,(\infty)$	0.270175	0.247843	0.223752	0.212310	0.207550	0.205328	0.204248
% errors in estimates							
Deterministic, (9.4)	−17.88	−14.09	−6.97	−3.11	−1.49	−0.73	−0.36
Improved M. G. (9.16)	−13.78	−12.19	−6.04	−2.66	−1.27	−0.62	−0.31
N var $x\,(\infty)$	0.590886	0.744968	0.715967	0.661485	0.642781	0.635102	0.631537
% error in estimate							
Martingale, (9.5)	−43.01	−56.38	−55.56	−52.46	−51.37	−50.93	−50.73
Estimated 5% point, P_5	−0.07997	0.003272	0.061475	0.102573	0.131747	0.152490	0.167237
$100\,P\,[x\,(\infty) \leqq P_5]$	0.00	3.00	2.57	3.92	4.48	3.89	4.79
Estimated 95% point, P_{95}	0.523689	0.422582	0.345848	0.308829	0.277155	0.255152	0.239773
$100\,P\,[x\,(\infty) \leqq P_{95}]$	86.42	87.89	91.82	92.33	93.80	93.92	94.41

irrespective of the initial conditions, and that we can use (8.12) effectively with stopping rule $u = \infty$, provided we split the distribution of x thereby obtained into its two components x_1 and x_2, and write the expectation as the probability average of its two parts, respectively the expectations conditional on x being sampled from x_1 and x_2.

Take $\phi_0 \equiv y + 2x - \log x - 2 - N^{-1}$ (since $k = N^{-1}$ for $I = 1$): then

$$E\{\phi_0^2(\boldsymbol{x}(\infty)) + \phi_1(\boldsymbol{x}(\infty))N^{-1}\} = \phi_1(\boldsymbol{x}(0))N^{-1} + O(N^{-2}),$$

where ϕ_1 is the next term in the expansion of the martingale corresponding to ϕ_0^2, by virtue of (8.12). Splitting the expectation into two parts corresponding to the possibilities $x_1(\infty)$ and $x_2(\infty)$, the contribution from $x_1(\infty)$ is of order N^{-2}, and so that from $x_2(\infty)$ satisfies, to order N^{-1}, the equation leading to the previous estimate of $\mathrm{var}\, x(\infty)$. Denote this estimate by $\mathrm{var}\,\theta$. Then we have

$$\begin{aligned} \mathrm{var}\, x(\infty) &= (1 - \tfrac{1}{2}N^{-1})\mathrm{var}\,\theta + \tfrac{1}{2}(1-\theta)^2N^{-1} + O(N^{-2}) \\ &= \mathrm{var}\,\theta + \tfrac{1}{2}(1-\theta)^2N^{-1} + O(N^{-2}). \end{aligned} \tag{9.18}$$

Thus we must correct our previous estimate by adding the term $\frac{1}{2}(1-\theta)^2N^{-1}$, the effective contribution from the rumour which terminates prematurely.

The two term martingale expansion corresponding to ϕ_0 is, from our earlier working, with $k = N^{-1}$,

$$y + 2x - \log x - 2 - N^{-1} - \tfrac{1}{2}(Nx)^{-1}.$$

Applying (8.12) with stopping rule $u = \infty$, and splitting the expectation as before, the contribution from x_1 is again of order N^{-2}, and, to order N^{-1}, the expectation of $x_2(\infty)$ is that calculated by the 'improved martingale estimate' (9.16). Denote this estimate by $\theta + e$. Then we have

$$\begin{aligned} Ex(\infty) &= (1 - \tfrac{1}{2}N^{-1})(\theta + e) + \tfrac{1}{2}N^{-1} + O(N^{-2}) \\ &= (\theta + e) + \tfrac{1}{2}(1-\theta)N^{-1} + O(N^{-2}). \end{aligned} \tag{9.19}$$

Thus we must correct our previous 'improved' estimate by adding the term $\frac{1}{2}(1-\theta)N^{-1}$.

The effect of these revised estimates is reflected in Table 3, which replaces two similar rows in Table 2. The new estimates clearly converge to the true values at the required rates, in marked contrast with the corresponding estimates used in Table 2, where the 'improved martingale estimate' converged no faster than the deterministic estimate, and where the variance was consistently underestimated by about 50%.

The work of Daley and Kendall [3] in deducing the leading terms in the

TABLE 3

$I = 1$

N	10	20	40	80	160	320	640
% errors in estimates							
Improved M. G. ($E\,x\,(\infty)$), (9.19)	0.62	−4.26	−1.61	−0.32	−0.07	−0.017	−0.004
Martingale (N var $x\,(\infty)$), (9.18)	8.39	−14.81	−11.78	−4.77	−2.14	−1.03	−0.50

asymptotic expansions of the mean and variance of the ultimate distribution of ignorants in the case $I = 1$ is very close to that of this section, and embodies the fundamental idea which this paper develops. Their results are numerically and algebraically consistent with the equivalent results in this paper, and they use the same techniques for explaining the discrepancy in the estimated variance. The extension of their ideas through the alternative approach in Section 8 enables one to go beyond their diffusion approximation, and to obtain in principle more detailed information about the mean and variance of the distribution as well as estimates of the higher cumulants. The results derived for a refined approximation to the expectation and for the third cumulant exhibit the practical relevance of this technique.

References

[1] BAILEY, N. T. J. (1964) *The Elements of Stochastic Processes.* (Chapter 10). John Wiley, New York.

[2] BAILEY, N. T. J. (1968) A perturbation approximation to the simple stochastic epidemic in a large population. *Biometrika* **55,** 199.

[3] DAILEY, D. J. AND KENDALL, D. G. (1965) Stochastic rumours. *J. Inst. Math. Appl.* **1,** 42–55.

[4] WHITTLE, P. (1957) On the use of the normal approximation in the treatment of stochastic processes. *J. Roy. Statist. Soc. Ser. B.* **19,** 268.