Numerical analysis of the master equation

Ronald Dickman*

Departamento de Física, ICEx, Universidade Federal de Minas Gerais, Caixa Postal 702, 30161-970, Belo Horizonte - MG, Brazil

Applied to the master equation, the usual numerical integration methods, such as Runge-Kutta method, become inefficient when the rates associated with various transitions differ by several orders of magnitude. We introduce an integration scheme that remains stable with much larger time increments than can be used in standard methods. When only the stationary distribution is required, a direct iteration method is even more rapid; this method may be extended to construct the *quasistationary* distribution of a process with an absorbing state. Applications to birth-and-death processes reveal gains in efficiency of two or more orders of magnitude.

The master equation is the basic tool for describing Markovian stochastic processes on a discrete state space, in continuous time [1,2]. Despite its central role in stochastic analysis in the physical and biological sciences, only a handful of exactly soluble examples are known. Thus the need to integrate the master equation numerically arises frequently, and with it the issue of computational efficiency.

Consider a Markov process with transition rates $W_{n,m}$ from state m to state n. In the master equation (ME),

$$\dot{p}_n = \sum_{n'} W_{n,n'} p_{n'} - p_n \sum_{n'} W_{n',n} \quad (n=0,\dots,N), \quad (1)$$

the factor $\sum_{n'} W_{n',n} \equiv w_n$ multiplying p_n can become large in a typical birth-and-death process, where $W_{n',n} \propto n$ or some higher power of n. This poses a problem for numerical integration via the usual discretization schemes, such as the Runge-Kutta method (RKM). Since instabilities appear when $|w_n \Delta t| \gtrsim 1$, we must use a small time increment, $\Delta t \sim 1/(\max_n w_n)$, and convergence is slow. With some (but not all) of the transition rates large, the ME is, in effect, a *stiff* system of differential equations, requiring special numerical treatment [3]. In this note I introduce numerical integration schemes for the ME that are efficient when the transition rates vary over a wide range, and an iteration method that eliminates the need for step-by-step integration, when only the stationary (or quasi-stationary) distribution is required. Simple but detailed examples are used to illustrate the methods; further applications will be reported elsewhere [4]. (We consider *stationary* Markov processes, i.e., time-independent rates, although the method is not limited to this class of problem.)

We begin by writing the ME in the form

$$\dot{p}_n = -w_n p_n + r_n, \quad (2)$$

where

$$r_n(t) = \sum_{n'} W_{n,n'} p_{n'}(t). \quad (3)$$

Note that p_n does not appear in the sum for r_n, since $W_{n,n} \equiv 0$. Both w_n and r_n are non-negative. (In fact, w_n is zero only if state n is absorbing.) Integrating Eq. (2) we have

$$p_n(t) = e^{-w_n t} p_n(0) + \int_0^t dt' e^{-w_n(t-t')} r_n(t'). \quad (4)$$

This is only a formal solution, since we need the $p_n(t)$ to evaluate r_n, but it is a useful starting point for approximations. If we adopt a time increment Δt such that $r_n(t)$ is approximately constant over this interval, then we have

$$p_n(\Delta t) \simeq e^{-w_n \Delta t} p_n(0) + [1 - e^{-w_n \Delta t}] \frac{r_n(0)}{w_n}. \quad (5)$$

This relation can then be iterated: we use $p_n(\Delta t)$ to evaluate r_n at the start of the second interval, and thereby find $p_n(2\Delta t)$, and so on. This simple formal integration (FI) scheme, analogous to Euler's method for direct numerical integration, already represents a significant advantage over the usual approaches when some of the w_n are large. The reason is that the exponential factor is already included in the solution, whereas in the usual discretizion methods it has to be constructed term by term, in powers of $w_n \Delta t$.

Suppose that the ME of interest possesses a unique, stable stationary solution $\bar{p}_n$. For a stationary solution, $\bar{p}_n = \bar{r}_n / w_n$, where $\bar{r}_n$ is given by Eq. (3) with $p_n = \bar{p}_n$. Setting $p_n = \bar{p}_n$ on the right-hand side (rhs) of Eq. (5) we immediately see that $\bar{p}_n$ is a stationary solution of our iteration method. To investigate the linear stability of the stationary solution, let $p_n = \bar{p}_n + \delta_n$. According to the ME,

$$\dot{\delta}_n = \sum_{n'} W_{n,n'} \delta_{n'} - \delta_n \sum_{n'} W_{n',n}, \quad (6)$$

or, in matrix notation,

$$\frac{d}{dt} \delta = S \delta, \quad (7)$$

where $S_{n,m} = W_{n,m}$ for $n \neq m$ and $S_{n,n} = -w_n$. Since, by hypothesis, the stationary solution $\delta = 0$ is stable, S has one zero eigenvalue and all others negative.

*Electronic address: dickman@fisica.ufmg.br

If we insert $p_n = \bar{p}_n + \delta_n$ in the rhs of Eq. (5) (so that $r_n = \bar{r}_n + \Sigma_{n'} W_{n,n'} \delta_{n'}$), we find

$$\delta(t+\Delta t) = T \delta(t), \tag{8}$$

where $T_{n,m} = (1 - e^{-w_n \Delta t}) W_{n,m}/w_n$ for $n \neq m$, and $T_{n,n} = e^{-w_n \Delta t}$. Expanding to first order in Δt, we see that Eq. (7) applies in this case as well, so the stationary solution is stable if it is so for the original ME.

We refer to the integration method embodied in relation Eq. (5) as a "first-order FI scheme" since the error in the solution $p_n(t)$ is proportional (for fixed t) to Δt. (In each step, the error incurred in treating r_n as constant is $O(\Delta t)^2$ [since in fact $r_n(\Delta t) = r_n(0) + r'_n(0)\Delta t + \cdots$], and we require $N = t/\Delta t$ steps.) An obvious way to improve accuracy, analogous to going from Euler's method to the midpoint (or second-order Runge-Kutta) method, is to replace the assumption of a constant r_n on the interval $[t, t+\Delta t]$ with a linear approximation. To do this we first use Eq. (5) to find $p_n(t+\Delta t)$, and then estimate the $r_n(t+\Delta t)$ using these values. Then we form the linear approximation

$$r_n(s) \simeq r_n(t) + (s-t) r'_n \quad (t \leq s \leq t + \Delta t), \tag{9}$$

where

$$r'_n = \frac{r_n(t+\Delta t) - r_n(t)}{\Delta t}. \tag{10}$$

Inserting this in the formal solution, Eq. (4), we find

$$p_n(t+\Delta t) \simeq e^{-w_n \Delta t} p_n(t) + [1 - e^{-w_n \Delta t}] \frac{r_n(t)}{w_n}$$
$$+ \left[\Delta t - \frac{1 - e^{-w_n \Delta t}}{w_n} \right] \frac{r'_n}{w_n}. \tag{11}$$

(Note that we do not use the expression for dr_n/dt that follows from the master equation in place of r'_n, since it does not, in general, represent the behavior of r_n over a substantial time interval.) The total error associated with the second-order FI procedure is $\propto (\Delta t)^2$. Further improvement is clearly possible, by introducing higher order polynomial approximations to the r_n, but the second-order scheme is quite adequate for the applications considered here. The error estimates apply for a fixed, finite time; the error for $t \to \infty$ is *zero*, if the system possesses a unique stationary distribution.

A tremendous simplification and speedup is possible if one only requires the *stationary* probability distribution. The relation $\bar{p}_n = \bar{r}_n/w_n$ suggests an iterative procedure of the form

$$p'_n = a p_n + (1-a) \frac{r_n}{w_n}, \tag{12}$$

where $0 < a < 1$ is a parameter. We expect p'_n to converge to $A\bar{p}_n$, where the factor A depends on a and on the initial distribution.

Some insight into the choice of a is afforded by the simple example of a two-state system with transition rates $W_{1,0} = \lambda$ and $W_{0,1} = \mu$. The evolution of the probability distribution is given (in matrix notation) by $p' = Mp$, where

$$M = \begin{pmatrix} a & (1-a)\dfrac{\mu}{\lambda} \\ (1-a)\dfrac{\lambda}{\mu} & a \end{pmatrix}. \tag{13}$$

Matrix M has eigenvalues 1 and $2a-1$. Thus the iterative procedure converges to a multiple of the stationary distribution, (μ, λ) for $0 < a < 1$, and is instantaneous for $a = 1/2$.

A similar analysis of the three-state processes ($n = 0,1,2$), with rates (a) $W_{n+1,n} = \gamma$, $W_{n-1,n} = 1$, and (b) $W_{n+1,n} = \gamma$, $W_{2,0} = \gamma^2$, and $W_{n,m} = 1$ for $n < m$, yields $a = 1/3$ as the optimal choice. (By optimal choice we mean the value that minimizes $\max_i |\omega_i|$, the ω_i being the eigenvalues of M, excluding, of course $\omega_0 \equiv 1$, associated with the stationary distribution.) These results suggest that for processes with a large number of states, even smaller values of a may be advantageous. The numerical examples discussed below support this notion.

A simple modification of the iterative scheme, Eq. (12), generates the *quasistationary* distribution (if such exists) of a Markov processs with an absorbing state. Consider a process on the states $n = 0,1,2,\ldots,N$ with $n = 0$ absorbing, i.e., $W_{n,0} \equiv 0$ for all n, while $W_{0,n}$ for at least one $n > 0$. (It seems reasonable to exclude manifestly transient states, in other words, we assume that for each n, $W_{n,m} > 0$ for at least one m.) In this case the stationary state is $\bar{p}_n = \delta_{n,0}$, but it is possible that the probability distribution, *conditioned on survival*, attains a stationary form, that is, for long times $p_n(t) \to C(t) q_n$ $(n > 0)$, where the q_n are time independent. Such quasistationary distributions [5] arise in birth-and-death processes with saturation, for example the Malthus-Verhulst process or the contact process [4].

The defining feature of the quasistationary distribution is that relative probabilities $p_n(t)/p_m(t)$ $(n, m > 0)$ are constant, or equivalently, $\dot{p}_n/p_n = \kappa$, constant and independent of n for $n \geq 1$. Suppose the probability distribution has attained a quasistationary form at some time t, and that at this instant the distribution is normalized so: $\Sigma_{n \geq 1} p_n(t) = 1$. Using Eq. (2) we have $r_n = (\kappa + w_n) p_n$, and summing on $n \geq 1$ yields

$$\kappa = \sum_{n \geq 1} \dot{p}_n = - \sum_{n \geq 1} W_{0,n} p_n \equiv -r_0. \tag{14}$$

(r_0 is the decay rate of the survival probability in the quasistationary regime.) Thus we find that in the quasistationary state,

$$p_n = \frac{r_n}{w_n - r_0} \quad (n \geq 1). \tag{15}$$

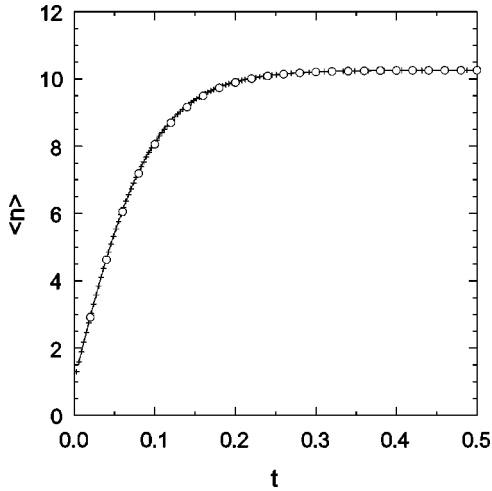This relation suggests that we iterate as follows:

FIG. 1. Mean population $\langle n \rangle$ versus time in the coagulation process, Eq. (17), with $\lambda = 100$, using Runge-Kutta integration (solid line), the second-order FI scheme with a fixed time step (circles), and with a variable time step, as described in the text (+).

$$p'_n = a p_n + (1-a) \frac{r_n}{w_n - r_0}. \tag{16}$$

The new distribution p' should be normalized after each iteration, since Eq. (15) assumes this property. [If $r_0 = 0$ the process of course possesses a true stationary distribution and Eq. (16) reduces to Eq. (12).] We have verified that this scheme converges to the quasistationary distribution much more rapidly than via integration of the ME.

As a first example we consider the nonlinear one-step process (on states $n = 1, 2, 3, \ldots$), with nonzero transition rates,

$$W_{n-1,n} = n(n-1), \quad W_{n+1,n} = \lambda. \tag{17}$$

This represents the coagulation process $A + A \rightarrow A$ with addition of species A at rate λ, in a well-stirred system. In Fig. 1 we compare the mean population size $\langle n \rangle_t$ as furnished by integration of the ME via a fourth-order RKM (solid line) with the second-order FI scheme, Eq. (11), (open circles). [Here $\lambda = 100$, $p_n(0) = \delta_{n,1}$, and we have set $p_n(t) \equiv 0$ for $n > 100$.] The results are nearly identical, but the RKM requires an integration step of $\Delta t \leq 2 \times 10^{-4}$ for stability, whereas the FI integration uses $\Delta t = 0.002$. The latter yields $\langle n \rangle_t$ with a relative error of less than 0.8%, the largest deviation occurring at relatively short times. (As noted above, the stationary values are identical in all cases.) To maintain fidelity to the solution at short, as well as long times, one may use the FI with an adjustable time step (+ in Fig. 1), such that $r'^* \Delta t$ is small, where $r'^* = \max_n \{r'_n\}$. For example, using the second-order FI with a time step of $\Delta t = 0.0005 + 0.02/(1 + r'^*)$, the relative error is reduced to $\leq 0.06\%$, while the average time step over the region of interest is 0.0024.

Comparison of the stationary distributions for $\lambda = 400$, obtained via the RKM, and by iteration of Eq. (12), reveals
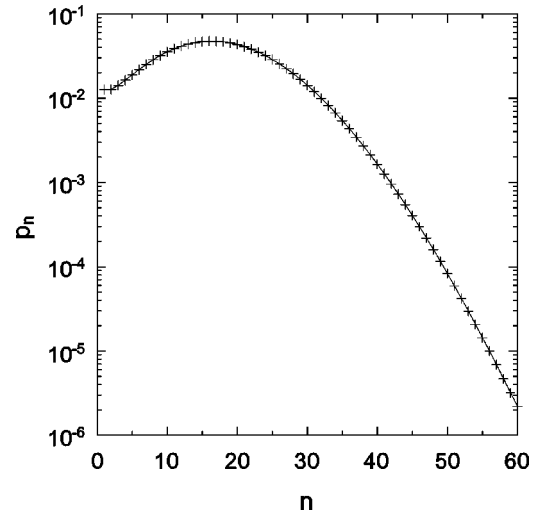

FIG. 2. Quasistationary probability distribution in the generalized Malthus-Verhulst process, Eq. (18), with $\lambda = 3$, $\mu = 0.1$, and $\gamma = 1$. Solid line, RKM; +, direct iteration of Eq. (16).

that they are identical to within one part in 10^5. In this case, the distribution converges (such that the error in $\langle n \rangle_t$ is $< 10^{-10}$), after about 900 iterations, when we use $a = 1/2$. As a is reduced, the number of iterations required falls steadily; in fact, the procedure converges even for $a = 0$, after only 460 steps.

Our second example is a multistep generalization of the Malthus-Verhulst process, with transition rates

$$W_{m,n} = \begin{cases} n[1 + \mu(n-1)]e^{\gamma(m-n+1)}, & m < n, \\ \lambda n e^{\gamma(n-m+1)}, & m > n, \\ 0, & m = n, \end{cases} \tag{18}$$

for $n, m = 0, 1, 2, \ldots$. (Thus transition rates between states n and m fall off $\propto \exp[-\gamma|m-n|]$.) We choose this example to show that the methods proposed here are not limited to one-step processes, which admit a relatively simple analysis [1,4].) In this case $n = 0$ is absorbing. Figure 2 compares the quasistationary distribution for $\lambda = 3$, $\mu = 0.1$, and $\gamma = 1$, obtained via the RKM (with $\Delta t = 10^{-4}$; larger values produce instability), with iteration of Eq. (16). As is clear from the plot, the results are identical in every detail. The RKM requires about 1270 s of CPU time on a DEC-alpha workstation to converge to the quasistationary distribution; the FI method requires about 50 s. Iteration of Eq. (16), (with $a = 0.5$), by contrast, converges in 9 s (about 1600 steps). Reducing parameter a to 0.1 yields convergence in just 900 steps (5 s CPU time), and in 815 steps for $a = 0$.

The steady decrease in computation time as we reduce a leads one to ask whether it is possible to use a *negative*. Iteration of Eq. (16) with $a < 0$ does not work, as negative probabilities are generated, but we can exclude these by writing

$$p'_n = \max\left[0, \; a p_n + (1-a) \frac{r_n}{w_n - r_0}\right], \tag{19}$$

This functions even for negative a, converging, in the present case, after about 630 steps when $a = -0.3$. For $a \leqslant -0.4$, however, the scheme does not converge. On the other hand, in the first example [i.e., rates given by Eq. (17)], using $a < 0$ does not offer any advantage over $a = 0$.

We conclude from these and other examples with large numbers of states [4], that $a \simeq 0$ generally yields rapid convergence, but that some amount of experimenting (for example, with negative a values) may reduce computation time further, in particular cases. But even without such optimization, the iterative schemes of Eqs. (12) and (16) yield economies in computation time of two or more orders of magnitude, compared to conventional integration methods. As the number of variables and/or the disparity in the magnitudes of the transition rates increases, direct iteration and the FI scheme become ever more valuable.

[1] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).

[2] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1990).

[3] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1996).

[4] R. Dickman and R. R. Vidigal, J. Phys. A **35**, 1147 (2002).

[5] P. A. Ferrari, H. Kestin, S. Martinez, and P. Picco, Ann. Prob. **23**, 501 (1995); P. A. Ferrari, H. Kestin, and S. Martinez, Ann. Appl. Probab. **6**, 577 (1996), and references therein.