Random Walks on a Complete Graph: A Model for Infection

Nilanjana Datta
Statistical Laboratory,
Centre for Mathematical Sciences,
Wilberforce Road, Cambridge CB3 0WB, United Kingdom.
Email: n.datta@statslab.cam.ac.uk

Tony C. Dorlas
Dublin Institute for Advanced Studies
School of Theoretical Physics
10 Burlington Road, Dublin 4, Ireland.
Email: dorlas@stp.dias.ie

December 22, 2003

Abstract

We introduce a new model for the infection of one or more subjects by a single agent and calculate the probability of infection after a fixed length of time. We model the agent and subjects as random walkers on a complete graph of N sites, jumping with equal rates from site to site. When one of the walkers is at the same site as the agent for a length of time τ, we assume that the infection probability is given by an exponential law with parameter γ, i.e. $q(\tau) = 1 - e^{-\gamma\tau}$. We introduce the boundary condition that all walkers return to their initial site ('home') at the end of a fixed period T. We also assume that the incubation period is longer than T so that there is no immediate propagation of the infection.

In this model, we find that for short periods T, i.e. $\gamma T \ll 1$ and $T \ll 1$, the infection probability is remarkably small and behaves like T^3. On the other hand, for large T, the probability tends to 1 (as might be expected) exponentially. However, the dominant exponential rate is given approximately by $\frac{2\gamma}{(2+\gamma)N}$ and is therefore small for large N.

1 Introduction

Models for the spread of an infection are important for studying their asymptotic behaviour, especially for long time. Numerous models have already been studied [1, 2, 3]. Here we introduce a particularly simple model which contains a new element. Namely, we take into account the fact that contact with an infectious agent generally takes place during the day (i.e. a set period) after which people return to their respective homes. This introduces a particular boundary condition into the problem.

We model the infectious agent and the initially uninfected subjects as random walkers on a complete graph of N sites. For simplicity we assume that all walkers jump from site to site with the same rate, which we take to be 1. (This obviously means that the length of the day, or appropriate period, is measured in terms of this rate.) We then calculate the probability distribution of the time a second walker coincides with the infected walker, starting and ending after a period of time T at a different site. This allows us to compute in particular the probability of infection of the second walker if we assume that during contact with the agent the probability of infection is given by an exponential law: $q(\tau) = 1 - e^{-\gamma\tau}$ with some constant $\gamma > 0$. The graph of Figure 1 below shows the behaviour of this infection probability.

Asymptotically for small T, the infection probability behaves like

$$P_{\text{infection}}(T) \sim \frac{\gamma T^3}{3(1 + (N-1)e^{-T})^2}. \tag{1.1}$$

(For $T \ll 1$ this reduces to $\frac{1}{3N^2}T^3$ but the above formula (1.1) is more accurate for large N and moderately small T.) This is seen more clearly in Figure 1, which also shows the graph for the approximation as a dadot line.

On the other hand, for large T, the infection probability tends to 1 at an exponential rate: $P_{\text{infection}}(T) \sim$ constant $\times\, e^{-r_N(\gamma)T}$, the dominant rate being given by

$$r_N(\gamma) = \frac{1}{2}\left(2 + \gamma - \sqrt{(2+\gamma)^2 - \frac{8}{N}\gamma}\right) \sim \frac{2\gamma}{(2+\gamma)N} \tag{1.2}$$

which is clearly small for large N (as well as small γ).

Random walks on a complete graph have been studied before, especially by Winkler [5, 4] and Aldous [6, 7], but the problem studied in this paper has not been addressed before. Random walks on graphs have various other applications, in particular in physics. The relation with the physics of polymers [11, 12, 13] and Feynman-Kac representation of quantum many-body systems [9, 10] is well-known. A particular application of random walks on

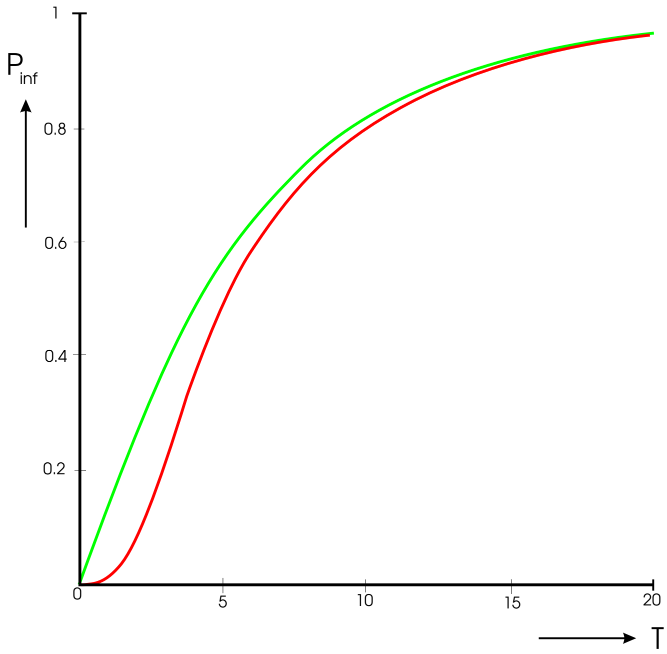

Figure 1: *Infection probability of a single walker as a function of the period T for $N = 10$. The solid line corresponds to the case where the walkers are constrained to return to their starting points; the dashed line corresponds to unconstrained walkers.*

a complete graph to a lattice model of a boson gas was elaborated by Toth [8].

2 The time of coincidence of two walkers

We label the infected agent as walker ξ_0 and the initially uninfected walker as walker ξ_1. We want to compute the probability distribution of the total length of time that the two walkers coincide, i.e. are at the same site. Similar problems were considered by Winkler [5, 4] and Aldous [6, 7]. We can subdivide the total time of coincidence into a number k of intervals. Between the intervals, the walks must avoid one another. We start by computing the transition probabilities for two random walks between two intervals of coincidence.

Denote by $P_t(A)$ the probability of the event A for walks during a time interval of length t and let $\bar{P}_t(A) = P_t(A \text{ and the walks do not coincide})$.

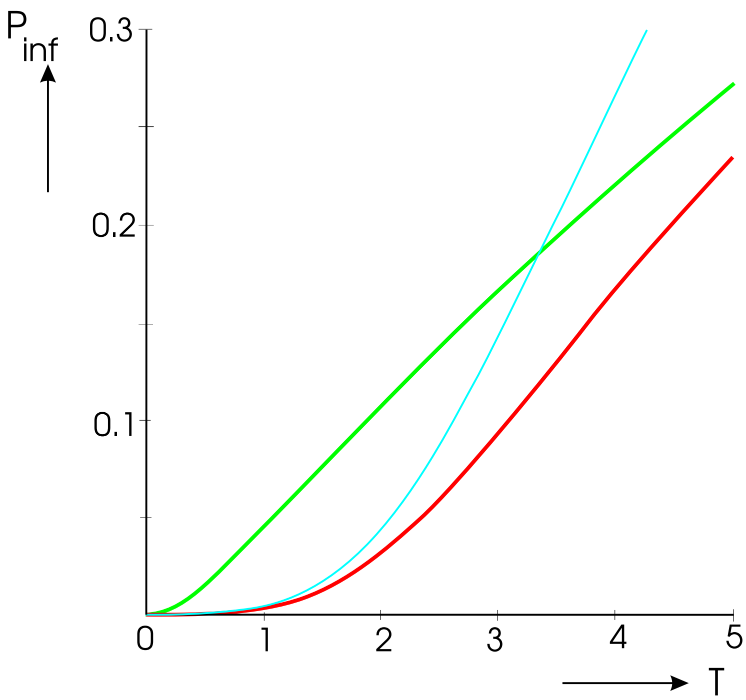

Figure 2: *Infection probability as a function of the period T for small T. The drawn line corresponds to the case where the walkers are constrained to return to their starting points; the dashed line corresponds to unconstrained walkers, and the dadot line is the approximation* (1.1).

Assuming unit jump rate, we have, for a single walker ξ,

$$\mathbb{P}\left(\xi(t+\delta t)=x' \,|\, \xi(t)=x\right)=\begin{cases} \frac{1}{N}\delta t & \text{if} \quad x'\neq x \\ 1-\frac{N-1}{N}\delta t & \text{if} \quad x'=x. \end{cases} \tag{2.1}$$

and

$$P_t(\xi : x\to x')=e^{-t}\delta_{x,x'}+\frac{1}{N}(1-e^{-t}). \tag{2.2}$$

Now define

$$g_1(t)=\bar{P}_t\begin{pmatrix}\xi_0 : x\to x \\ \xi_1 : x'\to \cdot\end{pmatrix}:=\sum_{x''\neq x}\bar{P}_t\begin{pmatrix}\xi_0 : x\to x \\ \xi_1 : x'\to x''\end{pmatrix}. \tag{2.3}$$

and

$$g_2(t)=\bar{P}_t\begin{pmatrix}\xi_0 : x\to x' \\ \xi_1 : x'\to \cdot\end{pmatrix}:=\sum_{x''\neq x'}\bar{P}_t\begin{pmatrix}\xi_0 : x\to x' \\ \xi_1 : x'\to x''\end{pmatrix}. \tag{2.4}$$

We can compute $g_1(t)$ and $g_2(t)$ as follows. Let t' denote the last time at which ξ_0 jumps. Then

$$\begin{aligned} g_1(t) &= \bar{P}_t \begin{pmatrix} \xi_0 : x \to x \text{ constant} \\ \xi_1 : x' \to \cdot \end{pmatrix} \\ &\quad + \sum_{x'' \neq x} \sum_{x''' \neq x, x''} \int_0^t \bar{P}_{t'} \begin{pmatrix} \xi_0 : x \to x''' \\ \xi_1 : x' \to x'' \end{pmatrix} \frac{dt'}{N} e^{-(t-t')}, \end{aligned} \tag{2.5}$$

where we use the fact that

$$\bar{P}_t \begin{pmatrix} \xi_0 : x \to x \text{ constant} \\ \xi_1 : x' \to \cdot \end{pmatrix} = e^{-t}. \tag{2.6}$$

This follows from

$$\begin{aligned} \bar{P}_{\delta t} \begin{pmatrix} \xi_0 : x \to x \text{ constant} \\ \xi_1 : x' \to \cdot \end{pmatrix} &= \left(1 - \frac{N-1}{N}\delta t\right)^2 \\ &\quad + \left(1 - \frac{N-1}{N}\delta t\right) \frac{N-2}{N}\delta t \\ &\sim 1 - \delta t. \end{aligned} \tag{2.7}$$

We can express the probability in the integrand of (2.5) as follows:

$$\begin{aligned} \sum_{x'' \neq x} \sum_{x''' \neq x, x''} \bar{P}_t \begin{pmatrix} \xi_0 : x \to x''' \\ \xi_1 : x' \to x'' \end{pmatrix} &= \bar{P}_t \begin{pmatrix} \xi_0 : x \to \cdot \\ \xi_1 : x' \to \cdot \end{pmatrix} - \bar{P}_t \begin{pmatrix} \xi_0 : x \to \cdot \\ \xi_1 : x' \to x \end{pmatrix} \\ &\quad - \sum_{x'' \neq x} \bar{P}_t \begin{pmatrix} \xi_0 : x \to x \\ \xi_1 : x' \to x'' \end{pmatrix} \\ &= e^{-(2/N)t} - g_2(t) - g_1(t) \end{aligned} \tag{2.8}$$

The expression $e^{-\frac{2}{N}t}$ is derived in the same way as (2.6): either both walks stay where they are, or one walk jumps to one of $N-2$ other positions.

$$\begin{aligned} \bar{P}_{\delta t} \begin{pmatrix} \xi_0 : x \to \cdot \\ \xi_1 : x' \to \cdot \end{pmatrix} &= \left(1 - \frac{N-1}{N}\delta t\right)^2 + 2\left(1 - \frac{N-1}{N}\delta t\right) \frac{N-2}{N}\delta t \\ &\sim 1 - \frac{2}{N}\delta t. \end{aligned} \tag{2.9}$$

Analogous to (2.5) we also have

$$\begin{aligned} g_2(t) &= \sum_{x'' \neq x} \sum_{x''' \neq x'', x} \int_0^t \bar{P}_{t'} \begin{pmatrix} \xi_0 : x \to x''' \\ \xi_1 : x' \to x'' \end{pmatrix} \frac{dt'}{N} e^{-t+t'} \\ &= g_1(t) - e^{-t}. \end{aligned} \tag{2.10}$$

Inserting this into (2.5) we obtain the following integral equation for $g_1(t)$:

$$g_1(t) = e^{-t} + \int_0^t \left[e^{-\frac{2}{N}t'} + e^{-t'} - 2g_1(t') \right] e^{-(t-t')} \frac{dt'}{N}. \tag{2.11}$$

With the obvious initial condition $g_1(0) = 1$ the solution is

$$g_1(t) = \frac{N-2}{2N} e^{-\frac{N+2}{N}t} + \frac{1}{N} e^{-\frac{2}{N}t} + \frac{1}{2} e^{-t}. \tag{2.12}$$

Hence, by (2.10),

$$g_2(t) = \frac{N-2}{2N} e^{-\frac{N+2}{N}t} + \frac{1}{N} e^{-\frac{2}{N}t} - \frac{1}{2} e^{-t}. \tag{2.13}$$

We will also need

$$f(t) = \bar{P}_t \begin{pmatrix} \xi_0 : x \longrightarrow x'' \\ \xi_1 : x' \longrightarrow \cdot \end{pmatrix}, \tag{2.14}$$

where $x'' \neq x, x'$. (Clearly, if $x'' = x$ then the right-hand side is just $g_1(t)$ and if $x'' = x'$ then it is $g_2(t)$ by symmetry.) If $x'' \neq x, x'$ then

$$\begin{aligned} f(t) = \bar{P}_t \begin{pmatrix} x \to x'' \\ x' \to \cdot \end{pmatrix} &= \frac{1}{N-2} \left[\bar{P}_t \begin{pmatrix} x \to \cdot \\ x' \to \cdot \end{pmatrix} \right. \\ &\qquad \left. - \bar{P}_t \begin{pmatrix} x \to x' \\ x' \to \cdot \end{pmatrix} - \bar{P}_t \begin{pmatrix} x \to x \\ x' \to \cdot \end{pmatrix} \right] \\ &= \frac{1}{N-2} \left[e^{-\frac{2}{N}t} - g_2(t) - g_1(t) \right] \\ &= \frac{1}{N} e^{-\frac{2}{N}t} (1 - e^{-t}). \end{aligned} \tag{2.15}$$

(Here and in the following we omit the explicit mention of the walks ξ_0 and ξ_1.)

Now consider first two walks conditioned to start and end after time T at the same point x. We compute the probability distribution of the total time τ of coincidence of the two walks. Clearly, there is a finite probability p_T that the walks coincide over the entire time interval, i.e. if they do not jump. This probability is

$$p_T = \exp \left[-2 \frac{N-1}{N} T \right]. \tag{2.16}$$

For $\tau < T$, there will in general be a number k of intervals where the walks *do not* coincide. We denote $\rho_k^{(=)}(T, \tau)$ the probability density for k intervals of non-coincidence with equal initial and final points of coincidence, and similarly, $\rho_k^{(\neq)}(T,\tau)$ for unequal initial and final points. Let the intervals of non-coincidence be $(t_1,t_2),(t_3,t_4),\ldots,(t_{2k-1},t_{2k})$ where $0<t_1<t_2<\cdots<t_{2k}<T$. Considering the last interval of non-coincidence (t_{2k-1},t_{2k}), one of the walks has to jump in a infinitesimal time dt_{2k-1} and one has to jump in an infinitesimal time dt_{2k}. These jumps happen with probability $\frac{dt_{2k-1}\,dt_{2k}}{N^2}$. We have to distinguish two cases: either the two points of coincidence before t_{2k-1} and after t_{2k} are the same, or they are different. In the first case, there are two possibilities (see Figure 2): if the walk that jumps first also jumps last, the transition probability is $g_1(t)$, if the walk that jumps first does not jump last, it is $g_2(t)$, where $t=t_{2k}-t_{2k-1}$.

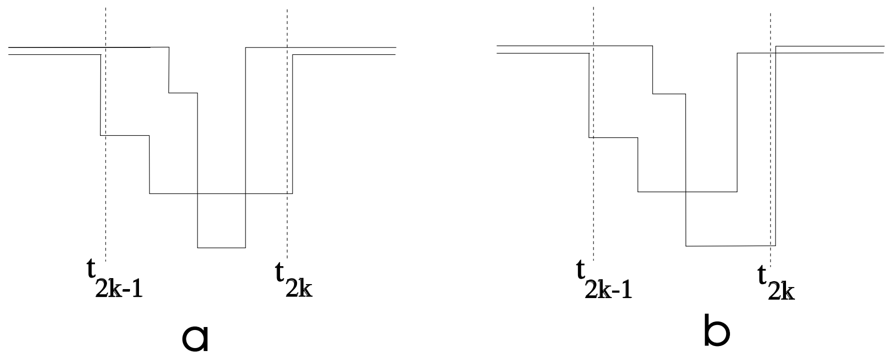

Figure 3: *Example trajectories of walks starting and ending at the same point. (For clarity the initial and final trajectories of the two walks have been separated by a small distance; they represent in fact the same point.) In* (a) *the first walk to jump away is the last to return to the initial point; in* (b) *it returns to the initial point before the other walk.*

The total transition probability in the case in which the points of coincidence immediately before and after the time interval of length t are the same, is therefore

$$\eta^{(=)}(t)=g_1(t)+g_2(t)=\frac{N-2}{N}e^{-\frac{N+2}{N}t}+\frac{2}{N}e^{-\frac{2}{N}t}. \tag{2.17}$$

Similarly, the total transition probability in the case in which the points of coincidence correspond to different levels is

$$\begin{aligned}\eta^{(\neq)}(t) &= [g_1(t)+(N-2)f(t)]+[g_2(t)+(N-2)f(t)]\\ &= 2e^{-\frac{2}{N}t}-\eta^{(=)}(t)\\ &= \left(2\frac{N-1}{N}-\frac{N-2}{N}e^{-t}\right)e^{-\frac{2}{N}t}.\end{aligned} \tag{2.18}$$

Finally, we remark that the total time of coincidence τ in case of k intervals of non-coincidence equals

$$\tau = \tau' + T - t_{2k},$$

where τ' is the time of coincidence of walks in the time interval $[0, T - t_{2k-1}]$. We can therefore write the following recursive equations:

$$\begin{aligned}\rho_k^{(=)}(T,\tau) &= \frac{2}{N^2}\int_0^\tau d\tau' \int_{\tau'}^{T-\tau+\tau'} dt_{2k-1}\, e^{-2\frac{N-1}{N}(T-t_{2k})} \\ &\quad \Big\{ (N-1)\rho_{k-1}^{(=)}(t_{2k-1},\tau')\eta^{(=)}(T-\tau+\tau'-t_{2k-1}) \\ &\quad +(N-1)\rho_{k-1}^{(\neq)}(t_{2k-1},\tau')\eta^{(\neq)}(T-\tau+\tau'-t_{2k-1}) \Big\} \end{aligned} \tag{2.19}$$

and

$$\begin{aligned}\rho_k^{(\neq)}(T,\tau) &= \frac{2}{N^2}\int_0^\tau d\tau' \int_{\tau'}^{T-\tau+\tau'} dt_{2k-1}\, e^{-2\frac{N-1}{N}(T-t_{2k})} \\ &\quad \Big\{ \rho_{k-1}^{(=)}(t_{2k-1},\tau')\eta^{(\neq)}(T-\tau+\tau'-t_{2k-1}) \\ &\quad +(N-2)\rho_{k-1}^{(\neq)}(t_{2k-1},\tau')\eta^{(\neq)}(T-\tau+\tau'-t_{2k-1}) \\ &\quad +(N-1)\rho_{k-1}^{(\neq)}(t_{2k-1},\tau')\eta^{(=)}(T-\tau+\tau'-t_{2k-1}) \Big\} \end{aligned} \tag{2.20}$$

Note that the factor of 2 arises because either of the two walkers can jump first. The factor $\exp\left[-2\frac{N-1}{N}(T-t_{2k})\right]$ corresponds to the final interval where neither walker jumps.

The solutions to these equations are, as can be easily verified,

$$\begin{aligned}\rho_k^{(=)}(T,\tau) &= \frac{2^k(N-1)}{N^{2k+1}} e^{-2\frac{N-1}{N}\tau} e^{-\frac{2}{N}(T-\tau)} \\ &\quad \times \frac{\tau^k}{k!}\frac{(T-\tau)^{k-1}}{(k-1)!}\left[2^k(N-1)^{k-1} + (N-2)^k e^{-T+\tau}\right] \end{aligned} \tag{2.21}$$

and

$$\begin{aligned}\rho_k^{(\neq)}(T,\tau) &= \frac{2^k}{N^{2k+1}} e^{-2\frac{N-1}{N}\tau} e^{-\frac{2}{N}(T-\tau)} \\ &\quad \times \frac{\tau^k}{k!}\frac{(T-\tau)^{k-1}}{(k-1)!}\left[2^k(N-1)^k - (N-2)^k e^{-T+\tau}\right]. \end{aligned} \tag{2.22}$$

In the Appendix, we verify that (2.21) satisfies the consistency condition

$$p_T + \sum_{k=1}^{\infty} \int_0^T \rho_k^{(=)}(T,\tau) d\tau = P_T(x \to x)^2 = \frac{1}{N^2}\left(1 + (N-1)e^{-T}\right)^2, \quad (2.23)$$

where the right-hand side follows from (2.2).

Next consider two walkers returning to their starting points in the case in which the two starting points are different. Let $p_k(\tau)$ denote the probability density for time of coincidence τ in this case, with k intervals of *coincidence.* We can reduce this to the previous case. Assume first $k > 1$. There are then at least two intervals of coincidence. In between the first and last interval, the distribution is given by $\rho_{k-1}^{(=)}$ or $\rho_{k-1}^{(\neq)}$, depending on whether the corresponding points of coincidence are equal or not. Consider the first case. Then the points of coincidence can either be one of the initial points of the walks, or a different point. If they are one of the initial points, we have, as before, a contribution $\eta^{(=)}$ for the interval of time before the first coincidence, and by symmetry also for the interval after the last coincidence. If the points of coincidence are different from the initial points, we get a contribution $2f(t_0)$ for the first interval (i.e. the time interval before the first coincidence) and $2f(t_l)$ for the last interval, where t_0 and t_l are the lengths of these intervals. (Notice that this is the contribution for a *specific* point of coincidence. We then have to multiply by a factor $N-2$ for the number of such points.) The total contribution to $p_k(\tau)$ for equal first and last points of coincidence is thus

$$2\int_0^{T-\tau} \frac{dt_0}{N} \int_{t_0+\tau}^{T} \frac{dt_{2k-1}}{N} \eta^{(=)}(t_0)\rho_{k-1}^{(=)}(t_{2k-1}-t_0,\tau)\eta^{(=)}(T-t_{2k-1})$$
$$+4(N-2)\int_0^{T-\tau} \frac{dt_0}{N} \int_{t_0+\tau}^{T} \frac{dt_{2k-1}}{N} f(t_0)\rho_{k-1}^{(=)}(t_{2k-1}-t_0,\tau)f(T-t_{2k-1}). \quad (2.24)$$

In the case that the first and last point of coincidence are unequal, there are 4 possibilities: both are equal to an initial point, the first is equal to an initial point and the last is not, or vice versa, or both are different from the initial points. As is easily seen , this yields the contributions

$$
\begin{aligned}
&2\int_0^{T-\tau}\frac{dt_0}{N}\int_{t_0+\tau}^{T}\frac{dt_{2k-1}}{N}\eta^{(=)}(t_0)\rho^{(\neq)}_{k-1}(t_{2k-1}-t_0,\tau)\eta^{(=)}(T-t_{2k-1})\\
&+4(N-2)\int_0^{T-\tau}\frac{dt_0}{N}\int_{t_0+\tau}^{T}\frac{dt_{2k-1}}{N}\eta^{(=)}(t_0)\rho^{(\neq)}_{k-1}(t_{2k-1}-t_0,\tau)f(T-t_{2k-1})\\
&+4(N-2)\int_0^{T-\tau}\frac{dt_0}{N}\int_{t_0+\tau}^{T}\frac{dt_{2k-1}}{N}f(t_0)\rho^{(\neq)}_{k-1}(t_{2k-1}-t_0,\tau)\eta^{(=)}(T-t_{2k-1})\\
&+4(N-2)(N-3)\int_0^{T-\tau}\frac{dt_0}{N}\int_{t_0+\tau}^{T}\frac{dt_{2k-1}}{N}f(t_0)\\
&\qquad\qquad\rho^{(\neq)}_{k-1}(t_{2k-1}-t_0,\tau)f(T-t_{2k-1}).
\end{aligned}
\tag{2.25}
$$

Some algebra, using (2.17) and (2.15), reduces the sum of all contributions to

$$
\begin{aligned}
p_k(\tau) &= \frac{4}{N^2}\int_0^{T-\tau}dt_0\int_{t_0+\tau}^{T}dt_{2k-1}e^{-\frac{2}{N}t_0}\rho^{(\neq)}_{k-1}(t_{2k-1}-t_0,\tau)e^{-\frac{2}{N}(T-t_{2k-1})}\\
&+\frac{2}{N^2}\int_0^{T-\tau}dt_0\int_{t_0+\tau}^{T}dt_{2k-1}\eta^{(=)}(t_0)\\
&\quad\left(\rho^{(=)}_{k-1}(t_{2k-1}-t_0)-\rho^{(\neq)}_{k-1}(t_{2k-1}-t_0,\tau)\right)\eta^{(=)}(T-t_{2k-1})\\
&+\frac{4(N-2)}{N^2}\int_0^{T-\tau}dt_0\int_{t_0+\tau}^{T}dt_{2k-1}f(t_0)\\
&\quad\left(\rho^{(=)}_{k-1}(t_{2k-1}-t_0)-\rho^{(\neq)}_{k-1}(t_{2k-1}-t_0,\tau)\right)f(T-t_{2k-1}),
\end{aligned}
\tag{2.26}
$$

which reduces further to

$$
\begin{aligned}
p_k(\tau) &= \frac{2(N-2)}{N^3}\int_0^{T-\tau}dt_0\int_{t_0+\tau}^{T}dt_{2k-1}e^{-\frac{N+2}{N}t_0}\\
&\quad\left(\rho^{(=)}_{k-1}(t_{2k-1}-t_0)-\rho^{(\neq)}_{k-1}(t_{2k-1}-t_0,\tau)\right)e^{-\frac{N+2}{N}(T-t_{2k-1})}\\
&+\frac{4}{N^3}\int_0^{T-\tau}dt_0\int_{t_0+\tau}^{T}dt_{2k-1}e^{-\frac{2}{N}t_0}\\
&\quad\left(\rho^{(=)}_{k-1}(t_{2k-1}-t_0)+(N-1)\rho^{(\neq)}_{k-1}(t_{2k-1}-t_0,\tau)\right)e^{-\frac{2}{N}(T-t_{2k-1})}.
\end{aligned}
\tag{2.27}
$$

Inserting the formulas for $\rho^{(=)}$ and $\rho^{(\neq)}$, the integrals are trivial, and the result is:

$$\begin{aligned} p_k(\tau) &= \frac{2^k}{N^{2k+1}} e^{-2\frac{N-1}{N}\tau} e^{-\frac{2}{N}(T-\tau)} \\ &\quad \times \frac{\tau^{k-1}}{(k-1)!} \frac{(T-\tau)^k}{k!} \left[2^k (N-1)^{k-1} + (N-2)^k e^{-T+\tau}\right]. \end{aligned} \tag{2.28}$$

This formula in fact also holds for $k = 1$. Indeed, it is easily seen that

$$\begin{aligned} p_1(\tau) &= \frac{2}{N^2} \int_0^{T-\tau} dt_0 \eta^{(=)}(t_0) e^{-2\frac{N-1}{N}\tau} \eta^{(=)}(T - t_0 - \tau) \\ &\quad + \frac{4(N-2)}{N^2} \int_0^{T-\tau} dt_0 f(t_0) e^{-2\frac{N-1}{N}\tau} f(T - t_0 - \tau) \\ &= \frac{2}{N^3} e^{-2\frac{N-1}{N}\tau} e^{-\frac{2}{N}(T-\tau)} (T-\tau) \left[2 + (N-2)e^{-T+\tau}\right]. \end{aligned} \tag{2.29}$$

Finally, there is a finite probability $p_0 = p_0(T)$ that the paths do not meet at all. This can be computed as follows. We have

$$\begin{aligned} \bar{P}_{t+\delta t} \begin{pmatrix} x \to x \\ x' \to x' \end{pmatrix} &= \bar{P}_t \begin{pmatrix} x \to x \\ x' \to x' \end{pmatrix} P_{\delta t}(\xi_0, \xi_1 \text{ constant}) + \\ &\quad +2 \left[\bar{P}_t \begin{pmatrix} x \to x \\ x' \to \cdot \end{pmatrix} - \bar{P}_t \begin{pmatrix} x \to x \\ x' \to x' \end{pmatrix} \right] \frac{\delta t}{N} \end{aligned} \tag{2.30}$$

and hence

$$p_0(t + \delta t) = p_0(t) \left(1 - 2\frac{N-1}{N}\delta t\right) + 2[g_1(t) - p_0(t)]\frac{\delta t}{N}. \tag{2.31}$$

Thus we have the differential equation

$$p_0'(t) = -2p_0(t) + \frac{2}{N} g_1(t) \tag{2.32}$$

with solution

$$\begin{aligned} p_0(t) &= e^{-2t} + \frac{2}{N} e^{-2t} \int_0^t g_2(t') e^{2t'} dt' \\ &= \frac{1}{N} \left[e^{-\frac{N+2}{N}t} + e^{-t} + (N-2)e^{-2t} \right] + \frac{1}{N(N-1)} \left[e^{-\frac{2}{N}t} - e^{-2t} \right]. \end{aligned} \tag{2.33}$$

Analogous to (2.23) we now have the consistency check (see the Appendix):

$$p_0 + \sum_{k=1}^{\infty} \int_0^T d\tau p_k(\tau) = P_T \begin{pmatrix} x \to x \\ x' \to x' \end{pmatrix} = \frac{1}{N^2}(1 + (N-1)e^{-T})^2. \tag{2.34}$$

3 The infection probability

We first remark that if there are n initially uninfected subjects then, provided the incubation time is longer than T, the expected number of infections after time T is just n times the probability of infection of a single walker. It would of course be interesting to investigate the case where cross-infection can occur, but that case is considerably more difficult.

In order to compute the probability of transfer of infection from walker ξ_0 to ξ_1 we want to compute the *conditional* probability of the second walker being infected after time T given that both walkers return home after time T. Assuming that the probability of infection upon contact for a period τ is given by

$$q(\tau) = 1 - e^{-\gamma\tau}, \tag{3.1}$$

this probability is given by

$$\mathrm{P}\ (\text{infection in time } T \,|\, \text{both walkers return home}) = \frac{\sum_{k=1}^{\infty} \int_0^T q(\tau) p_k(\tau) d\tau}{K_N(T)}, \tag{3.2}$$

where $K_N(T)$ is the probability of both walkers returning home:

$$K_N(T) = \frac{1}{N^2} \left[1 + (N-1)e^{-T}\right]^2. \tag{3.3}$$

Notice that the exponential law (3.1) is natural in that the probability that the infection is not transferred upon contact over a time τ, $\bar{q}(\tau) = 1 - q(\tau)$ satisfies $\bar{q}(\tau_1 + \tau_2) = \bar{q}(\tau_1)\bar{q}(\tau_2)$. This also means that we only need the probability distribution of the total time τ of contact in (3.2).

The numerator in (3.2) can be calculated exactly (see (A.9)), but the result is complicated. Some insight can already be obtained (for small γ) by approximating it with the first order contribution in γ:

$$\gamma I_T^{(1)} = \gamma \sum_{k=1}^{\infty} \int_0^T \tau p_k(\tau) d\tau. \tag{3.4}$$

Obviously, this approximation is appropriate only for values of γ which are small compared to $1/T$. (Notice that the jump rate of each walker has been set equal to 1, so T has to be measured in these terms.)

We now proceed by computing the quantity $I_T^{(1)}$. Using the formula

$$\begin{aligned} \int_0^\beta \tau^n (\beta - \tau)^m e^{-\alpha\tau} d\tau &= \sum_{p=0}^{m} \binom{m}{p} \frac{(n+p)!}{\alpha^{n+p+1}} (-1)^p \beta^{m-p} \\ &- \sum_{p=0}^{n} \binom{n}{p} \frac{(p+m)!}{\alpha^{m+p+1}} (-1)^m \beta^{n-p} e^{-\alpha\beta} \end{aligned} \tag{3.5}$$

the integrals can be evaluated:

$$
\begin{aligned}
I_T^{(1)} \;=\; & \sum_{k=1}^{\infty} \frac{2^k}{N^{2k+1}} e^{-\frac{2}{N}T} \frac{1}{k!(k-1)!} \Bigg\{ \\
& 2^k (N-1)^{k-1} \left[\sum_{p=0}^{k} \binom{k}{p} \frac{(k+p)!}{2^{k+p+1}} \left(\frac{N}{N-2} \right)^{k+p+1} (-1)^p \, T^{k-p} \right. \\
& \left. - \sum_{p=0}^{k} \binom{k}{p} \frac{(k+p)!}{2^{k+p+1}} \left(\frac{N}{N-2} \right)^{k+p+1} (-1)^k \, T^{k-p} e^{-2\frac{N-2}{N}T} \right] \\
& +(N-2)^k e^{-T} \left[\sum_{p=0}^{k} \binom{k}{p} (k+p)! \left(\frac{N}{N-4} \right)^{k+p+1} (-1)^p \, T^{k-p} \right. \\
& \left. - \sum_{p=0}^{k} \binom{k}{p} (k+p)! \left(\frac{N}{N-4} \right)^{k+p+1} (-1)^k \, T^{k-p} e^{-\frac{N-4}{N}T} \right] \Bigg\}
\end{aligned}
\tag{3.6}
$$

To evaluate the sums, we use identity (A.2) from the Appendix, and differentiate to get

$$
\begin{aligned}
& \sum_{k=1}^{\infty} \sum_{p=0}^{k} \frac{(k+p)!}{p!(k-1)!(k-p)!} x^{k+p-1} y^p = \\
&= \frac{d}{dx} \left\{ \sum_{k=1}^{\infty} \sum_{p=0}^{k} \frac{(k+p-1)!}{p!(k-1)!(k-p)!} x^{k+p} y^p \right\} \\
&= \frac{d}{dx} \left\{ \frac{2x^2 y}{\sqrt{1-4x^2y}(1-\sqrt{1-4x^2y})} \exp \left[\frac{1-\sqrt{1-4x^2y}}{2xy} \right] \right\} \\
&= 2 \frac{4x^3y^2 - 4x^2y - 2xy + 1 + (2x^2y + 2xy - 1)\sqrt{1-4x^2y}}{(1-4x^2y)\sqrt{1-4x^2y}(1-\sqrt{1-4x^2y})^2}.
\end{aligned}
\tag{3.7}
$$

The appropriate insertions for x and y we need to make in the four terms of $I_T^{(1)}$ are:

$$
\text{in the first term: } x = \frac{2(N-1)T}{N(N-2)} \text{ and } y = -\frac{N^2}{4(N-1)T^2}; \tag{3.8}
$$

$$
\text{in the second term: } x = -\frac{2(N-1)T}{N(N-2)} \text{ and } y = -\frac{N^2}{4(N-1)T^2}; \tag{3.9}
$$

$$
\text{in the third term: } x = \frac{2(N-2)T}{N(N-4)} \text{ and } y = -\frac{N^2}{2(N-2)T^2}; \tag{3.10}
$$

and

$$\text{in the fourth term: } x = -\frac{2(N-2)T}{N(N-4)} \text{ and } y = -\frac{N^2}{2(N-2)T^2}. \tag{3.11}$$

This yields, after some algebra,

$$I_T^{(1)} = \frac{T-1}{N^3} + \frac{T+1}{N^3}e^{-2T} + \frac{2(N-2)(T-2)}{N^3}e^{-T} + \frac{2(N-2)(T+2)}{N^3}e^{-2T}. \tag{3.12}$$

A simple expansion shows that for small T this behaves like $\frac{1}{3N^2}T^3$. Dividing by $K_N(T)$ results in the formula (1.1) quoted in the Introduction. The large T behaviour cannot be read off from $I_T^{(1)}$ as this approximation is only valid for $\gamma T \ll 1$. The dadot curve in Figure 1 shows how $\gamma I_T^{(1)}$ deviates from $I_T(\gamma)$ for large T. The exact solution for $I_T(\gamma)$ is derived in the Appendix.

4 Unconstrained walks

The above model made the reasonable assumption that the walkers return home after a fixed time T. If we discard this assumption, the problem becomes much easier to analyse. In that case, we can represent the process by a flow diagram as in Figure 4.

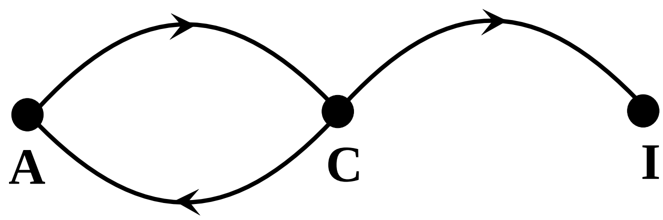

Figure 4: *Flow diagram for two walkers on a complete graph in the case in which one only distinguishes the states where the two walkers coincide or not and the second walker is not yet infected, and that where the second walker is infected.*

Here the vertices correspond to the cases that the walks are at different sites, i.e. apart (A), and that where they coincide (C), and (I) where the second walker has been infected. The transition probabilities are:

$$P_{\delta t}(A \to C) = \frac{2}{N}\delta t, \tag{4.1}$$

$$P_{\delta t}(C \to A) = 2\frac{N-1}{N}\delta t, \tag{4.2}$$

and

$$P_{\delta t}(C \to I) = \gamma \delta t. \tag{4.3}$$

The transition matrix P_t is therefore given by $P_t = e^{-tQ}$ with

$$Q = \begin{pmatrix} \frac{2}{N} & -\frac{2}{N} & 0 \\ -2\frac{N-1}{N} & 2\frac{N-1}{N} + \gamma & -\gamma \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.4}$$

The eigenvalues are 0 and

$$\lambda_{\pm} = \frac{1}{2}\left(2 + \gamma \pm \sqrt{(2+\gamma)^2 - \frac{8}{N}\gamma}\right). \tag{4.5}$$

The corresponding eigenvectors are

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ 1 - \frac{N}{2}\lambda_{\pm} \\ 0 \end{pmatrix}. \tag{4.6}$$

The infection probability after time T is given by

$$P_{\text{inf}}(T) = (1, 0, 0)\, e^{-TQ} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{4.7}$$

A simple calculation now shows that

$$P_{\text{inf}}(T) = 1 + \frac{\lambda_-}{\lambda_+ - \lambda_-} e^{-T\lambda_+} - \frac{\lambda_+}{\lambda_+ - \lambda_-} e^{-T\lambda_-}. \tag{4.8}$$

Clearly, for large T, the asymptotic behaviour is the same as in the case of walks restrained to return to their starting points: it is given by the smallest exponent, i.e. λ_-. This is just (1.2). This is also clearly visible in Figure 1. On the other hand, for small T, $P_{\text{inf}}(T)$ behaves like

$$P_{\text{inf}}(T) \sim \frac{1}{2}\lambda_+ \lambda_- T^2 = \frac{1}{N}\gamma T^2. \tag{4.9}$$

The dashed graph in Figure 1 shows the behaviour of this infection probability as a function of T, for small T, $N = 10$ and $\gamma = 1$. Notice that for large N the prefactor of the asymptotics is more important than the power of T.

Appendix

Here we derive an explicit formula for the expectation of $e^{-\gamma\tau}$ and hence the infection probability. Evaluating the integrals we have, using (3.5),

$$\begin{aligned}
\int_0^T e^{-\gamma\tau} p_k(\tau)d\tau \;=\; & \frac{4^k(N-1)^{k-1}}{N^{2k+1}} \frac{e^{-\frac{2}{N}T}}{k!(k-1)!} \\
& \times\Bigg\{ \sum_{p=0}^{k} \binom{k}{p} \frac{(k+p-1)!}{\left(\frac{2}{N}(N-2)+\gamma\right)^{k+p}} (-1)^p \, T^{k-p} \\
& - \sum_{p=0}^{k-1} \binom{k-1}{p} \frac{(k+p)!}{\left(\frac{2}{N}(N-2)+\gamma\right)^{k+p+1}} (-1)^k \, T^{k-p-1} \\
& \times e^{-\left(\frac{2}{N}(N-2)+\gamma\right)T} \Bigg\} \\
& + \frac{2^k(N-2)^k}{N^{2k+1}} \frac{e^{-\left(\frac{2}{N}+1\right)T}}{k!(k-1)!} \\
& \times\Bigg\{ \sum_{p=0}^{k} \binom{k}{p} \frac{(k+p-1)!}{\left(\frac{N-4}{N}+\gamma\right)^{k+p}} (-1)^p \, T^{k-p} \\
& - \sum_{p=0}^{k-1} \binom{k-1}{p} \frac{(k+p)!}{\left(\frac{N-4}{N}+\gamma\right)^{k+p+1}} (-1)^k \, T^{k-p-1} \\
& \times e^{-\left(\frac{N-4}{N}+\gamma\right)T} \Bigg\}
\end{aligned} \tag{A.1}$$

To perform the sum over k we make use of the identities

$$\sum_{k=1}^{\infty} \sum_{p=0}^{k-1} \frac{(k+p)!}{p!(k-p-1)!k!} x^{k-p} y^p = \frac{x(1-\sqrt{1-4y})}{2y\sqrt{1-4y}} \exp\left[\frac{x(1-\sqrt{1-4y})}{2y}\right] \tag{A.2}$$

and

$$\sum_{k=1}^{\infty} \sum_{p=0}^{k} \frac{(k+p-1)!}{p!(k-p)!(k-1)!} x^{k-p} y^p = \frac{1+\sqrt{1-4y}}{2\sqrt{1-4y}} \exp\left[\frac{x(1-\sqrt{1-4y})}{2y}\right] - 1. \tag{A.3}$$

The first formula applies to the second and fourth terms, the second formula to the first and third terms. The resulting contributions are, for the first term:

$$T_1 = \frac{1}{2N(N-1)} \left(1 + \frac{2+\gamma-\frac{4}{N}}{\sqrt{(2+\gamma)^2 - \frac{8}{N}\gamma}}\right) \times \exp\left[-\frac{1}{2}\left(2+\gamma-\sqrt{(2+\gamma)^2-\frac{8}{N}\gamma}\right)T\right] - \frac{e^{-\frac{2}{N}T}}{N(N-1)}; \tag{A.4}$$

and for the third term,

$$T_3 = \frac{1}{2N} \left(1 + \frac{1+\gamma-\frac{4}{N}}{\sqrt{(1+\gamma)^2 - \frac{8}{N}\gamma}}\right) \times \exp\left[-\frac{1}{2}\left(3+\gamma-\sqrt{(1+\gamma)^2-\frac{8}{N}\gamma}\right)T\right] - \frac{1}{N}e^{-\frac{N+2}{N}T}; \tag{A.5}$$

and for the second term,

$$T_2 = \frac{1}{2N(N-1)} \left(1 - \frac{2+\gamma-\frac{4}{N}}{\sqrt{(2+\gamma)^2 - \frac{8}{N}\gamma}}\right) \times \exp\left[-\frac{1}{2}\left(2+\gamma+\sqrt{(2+\gamma)^2-\frac{8}{N}\gamma}\right)T\right]; \tag{A.6}$$

and, finally, for the fourth term,

$$T_4 = \frac{1}{2N} \left(1 - \frac{1+\gamma-\frac{4}{N}}{\sqrt{(1+\gamma)^2 - \frac{8}{N}\gamma}}\right) \exp\left[-\frac{1}{2}\left(3+\gamma+\sqrt{(1+\gamma)^2-\frac{8}{N}\gamma}\right)T\right]. \tag{A.7}$$

Notice that, in particular, at $\gamma = 0$ we obtain

$$\sum_{k=1}^{\infty} \int_0^T p_k(\tau)d\tau = \frac{1}{N^2} - \frac{1}{N(N-1)}e^{-\frac{2}{N}T} + \frac{1}{N^2(N-1)}e^{-2T} + \frac{N-2}{N^2}e^{-T} - \frac{1}{N}e^{-\frac{N+2}{N}T} + \frac{2}{N^2}e^{-2T}. \tag{A.8}$$

Adding this to $p_0(T)$ given by (2.28) yields $K_N(T)$ as stated in (2.33). To obtain the infection probability, we have to subtract the contributions T_1,

T_2, T_3 and T_4 from the right-hand side of (A.8) and divide by $K_N(T)$. The numerator is

$$
\begin{aligned}
I_T(\gamma) &= \sum_{k=1}^{\infty}\int_0^T q(\tau)p_k(\tau) \\
&= \frac{1}{N^2} + \frac{N-2}{N^2}e^{-T} + \frac{2N-1}{N^2(N-1)}e^{-2T} \\
&\quad -\frac{1}{2N(N-1)}\left(1+\frac{2+\gamma-\frac{4}{N}}{\sqrt{(2+\gamma)^2-\frac{8}{N}\gamma}}\right) \\
&\quad\quad \times\exp\left[-\frac{1}{2}\left(2+\gamma-\sqrt{(2+\gamma)^2-\frac{8}{N}\gamma}\right)T\right] \\
&\quad -\frac{1}{2N(N-1)}\left(1-\frac{2+\gamma-\frac{4}{N}}{\sqrt{(2+\gamma)^2-\frac{8}{N}\gamma}}\right) \\
&\quad\quad \times\exp\left[-\frac{1}{2}\left(2+\gamma+\sqrt{(2+\gamma)^2-\frac{8}{N}\gamma}\right)T\right] \\
&\quad -\frac{1}{2N}\left(1+\frac{1+\gamma-\frac{4}{N}}{\sqrt{(1+\gamma)^2-\frac{8}{N}\gamma}}\right) \\
&\quad\quad \times\exp\left[-\frac{1}{2}\left(3+\gamma-\sqrt{(1+\gamma)^2-\frac{8}{N}\gamma}\right)T\right] \\
&\quad -\frac{1}{2N}\left(1-\frac{1+\gamma-\frac{4}{N}}{\sqrt{(1+\gamma)^2-\frac{8}{N}\gamma}}\right) \\
&\quad\quad \times\exp\left[-\frac{1}{2}\left(3+\gamma+\sqrt{(1+\gamma)^2-\frac{8}{N}\gamma}\right)T\right]. \qquad \text{(A.9)}
\end{aligned}
$$

A simple but tedious differentiation shows that $I_T^{(1)} = \frac{d}{d\gamma}I_T(\gamma)\Big|_{\gamma=0}$ as it should.

Acknowledgements

The authors would like to thank Philippe Martin for inviting them to Lausanne, where this work was started. We are also grateful to Sir John Kingman and James Norris for useful discussions and suggestions.

References

[1] N. T. J. Bailey, *The mathematical theory of infectious diseases and its applications.* (2nd Edn.) Charles Griffin & Comp. Ltd., London and High Wycombe, 1975.

[2] D. J. Daley & J. Gani, *Epidemic Modelling: An Introduction.* Cambridge Studies in Mathematical Biology No. 15. Cambridge Univ. Press, 1999.

[3] N. G. Becker, *Analysis of Infectious Disease Data.* Monographs on Statistics and Applied Probability. Chapman & Hall; London, New York 1989.

[4] D. Coppersmith, P. Tetali & P. Winkler, *Collisions among random walks on a graph.* SIAM J. Discrete Math. **6**, 363–374 (1993).

[5] P. Winkler, *Dependent percolation and colliding random walks.* Random Struct. & Algorithms **16**, 58–84 (2000).

[6] D. Aldous, *Meeting times for independent Markov chains.* Stoch. Proc. and their Appl. **38**, 185–193 (1991).

[7] D. Aldous, *Applications of random walks on finite graphs.* Selected Proceedings of the Sheffield Symposium on Applied Probability (Sheffield, 1989), pp. 12–26, IMS Lecture Notes Monogr. Ser. **18**.

[8] B. Toth, *Phase transition in an interacting Bose system. An application of the theory of Ventsel' and Freidlin.* J. Stat. Phys. **61**, 749–764 (1990).

[9] J. Ginibre, *Some applications of functional integration in statistical mechanics.* In: *Statistical Mechanics and Quantum Field Theory.* Les Houches summer school 1970; C. De Witt & R./ Stora (eds.). New York, London, Paris: Gordon & Breach Science Publ. 1971, pp. 327–427.

[10] D. Brydges & Ph-A. Martin, *Coulomb Systems at Low Density: A Review.* J. Stat. Phys. **96**, 1163–1330 (1999).

[11] P.-G. de Gennes, *Scaling Concepts in Polymer Physics.* Cornell Univ. Press, 1979.

[12] M. Doi, *Introduction to Polymer Physics.* Clarendon Press, Oxford, 1996.

[13] J. M. Ziman, *Models of Disorder.* Cambridge Univ. Press, 1979.