THE ROYAL SOCIETY

Received 8 March 2002
Accepted 5 June 2002
Published online 4 September 2002

From HIV infection to AIDS: a dynamically induced percolation transition?

Christel Kamp[1*] **and Stefan Bornholdt**[1,2]

[1]*Institut für Theoretische Physik, Universität Kiel, Leibnizstrasse 15, 24098 Kiel, Germany*
[2]*AG Statistische Biophysik, Interdisziplinäres Zentrum für Bioinformatik, Universität Leipzig, Kreuzstrasse 7b, 04103 Leipzig, Germany*

The origin of the unusual incubation period distribution in the development of AIDS is largely unresolved. A key factor in understanding the observed distribution of latency periods, as well as the occurrence of infected individuals not developing AIDS at all, is the dynamics of the long-lasting struggle between HIV and the immune system. Using a computer simulation, we study the diversification of viral genomes under mutation and the selective pressure of the immune system. In non-HIV infections, vast spreading of viral genomes in genome space usually does not take place. In the case of an HIV infection, this may occur, as the virus successively weakens the immune system by the depletion of $CD4^+$ cells. In a sequence space framework, this leads to a dynamically induced percolation transition, corresponding to the onset of AIDS. As a result, we obtain a prolonged shape of the incubation period distribution, as well as a finite fraction of non-progressors that do not develop AIDS, comparing well with results from recent clinical research.

Keywords: HIV; AIDS; incubation period distribution; percolation; theoretical model; computer simulation

1. INTRODUCTION

It is a well-known empirical fact that incubation times of most diseases obey a lognormal distribution that is often referred to as 'Sartwell's model' (Sartwell 1966). More recently, the underlying dynamics that generates the incubation period distribution, as well as mechanisms that lead to deviations from the common distribution, have gained attention (Philippe 2000). One of the most prominent examples of a deviation from the lognormal case is the distribution of waiting times between HIV infection (seroconversion) and the onset of AIDS, which is supported by datasets from various studies (CASCADE Collaboration 2000*a*,*b*; Robert Koch Institut 1998). The divergence from lognormality—extraordinarily long incubation times and the occurrence of non-progressors (patients not developing AIDS)—indicate more complex generating dynamics than observed in other infectious diseases. While much effort has been spent on parametric estimates of the incubation period distribution (Dangerfield & Roberts 1999), here we study possible mechanisms of the underlying dynamics. Any such attempt has to take into account the HIV-specific negative feedback to the host's immune system. While the immune system develops an ordinary epitope-specific answer to HIV, HIV targets the replication machinery of $CD4^+$ immune cells, which are depleted when viruses proliferate. As a result, the host's resistance against antigens is globally weakened.

In earlier differential equation approaches, the onset of AIDS has been associated with the passage of an antigenic diversity threshold (Nowak *et al.* 1991). More recently, progress has been made to overcome the limitations of analytical models with respect to topological effects in the shape space of receptors and in physical space. Inspired by early models (Pandey & Stauffer 1990), cellular automaton models have been defined and investigated that show the typical separation between the time-scales of primary infection and the onset of AIDS (Hershberg *et al.* 2001; Zorzenon dos Santos & Coutinho 2001).

In this article, we take an alternative approach and combine cellular automata with a sequence space framework in order to model typical characteristics of the time-course of HIV infection. In the following sections, we first define a framework to represent ordinary infections within the scope of percolation theory. From there we extend the model to describe the special case of HIV infection and discuss the distribution of incubation periods. Numerical simulations are complemented by a stochastic model for the origin of the variety in incubation period distributions. Finally, we discuss our findings in the context of empirical data on HIV survival. Here, we extend the scope of earlier models, making estimates on the behaviour of the incubation period distribution beyond the range of empirical data.

2. PERCOLATION MODEL OF INFECTION

In the course of an infection, one generally observes a diversification of viral genomes due to mutation and the selective pressure of the immune system. These coevolutionary dynamics can be modelled within a sequence space framework (Perelson & Weisbuch 1997; Kamp & Bornholdt 2002). Representing viral genomes by strings of length n, built up from an alphabet of length λ, we can describe their diversification as a spread in sequence space. Analogously, let us assign a sequence to the respective immune receptor matching the viral strain. As any string in sequence space is assumed to represent a viral epitope, as well as its complementary immune receptor, it is characterized by a viral and an immunological state variable. A mathematical framework to describe the

*Author for correspondence (kamp@theo-physik.uni-kiel.de).

dynamics in such a space can be found in percolation theory (Stauffer & Aharony 1992) and in theories for epidemic spreading, i.e. SIR models (Hethcote 2000; Moreno *et al.* 2001), which are equivalent to bond percolation (Grassberger 1983). However, while those models apply cellular automata to the interaction of organisms, we here apply the mathematical concept to modelling populations of immune cells and viruses within one organism. Adopting the notation of SIR models, we call a site in sequence space susceptible, if it in principle can harbour a virus. It is denoted as infected if the system contains a virus with an epitope motif represented by the site's string. If a viral sequence meets an immune response, it is removed and the system is immunized against it. In both this case and in the case in which a site is in principle inaccessible for a virus, it is called recovered (or removed). Apart from this, two immunological states are distinguished. An immune receptor shape may or may not be present within the immune repertoire. We set up a system in which a site is inaccessible for viral sequences with probability D_0 accounting for the fact that the viral genome is not arbitrary. In addition, we introduce a probability of immunological presence at a site in sequence space $\rho_{is}(t)$, with $\rho_{is}(0) = \rho_0$. This means that for sufficiently large systems the initial density of recovered sites is $R(0) = D_0 + \rho_0 - D_0\rho_0$. Also, taking into account the densities of susceptible sites $S(t)$ and of infected sites $\rho_v(t)$, one obtains the relationship

$$S(t) + \rho_v(t) + R(t) = 1 \quad \forall t. \tag{2.1}$$

The replication of viral and immunological entities is afflicted with copy fidelities $q_v < 1$ and $q_{is} < 1$, denoting the respective probability of correct duplication of a sequence's digit. As a result, the system shows viral (and in response immunological) spread in sequence space. Introducing some viral strains into a so far unaffected system leads to dynamics that is modelled within the cellular automaton approach by iterating the following steps.

(i) Choose a random site.
(ii) If the site represents an active immune receptor:
  (a) mutate any bit with probability $1 - q_{is}$,
  (b) if a new immunological strain is generated and the mutant matches an infected site, reset the site's viral status to recovered and occupy that site with an immune receptor.
(iii) If the site is infected:
  (a) mutate any bit with probability $1 - q_v$,
  (b) if a new strain is generated and corresponds to a susceptible site the site gets infected.

A viral strain generates a specific mutant strain at Hamming distance d (which is the number of differing digits) with probability $(1 - q_v)^d q_v^{n-d}/(\lambda - 1)^d$, which will survive as long as it meets a susceptible site. Equally, an immunological mutant strain is originated with probability $(1 - q_{is})^d q_{is}^{n-d}/(\lambda - 1)^d$ under the condition that it coincides with an infected site. Otherwise, we assume that the immunological mutant is not sufficiently amplified to establish a new strain.

This scenario corresponds to bond percolation in a fully connected graph with an occupation probability decaying exponentially with the Hamming distance from the source of infection. Such a system shows two regimes of qualitatively different behaviour. Below a percolation threshold, the source of infection will stay negligible in size compared with the system size, such that in the limit of infinite system size $R(\infty) = R(0)$. Above the percolation threshold, a virus will spread all over the system before it gets defeated. Accordingly $R(\infty) > R(0)$.

To determine the threshold conditions within a mean field approach ('fully mixed' approximation), we introduce the following system of differential equations

$$\frac{dS}{dt} = -(1 - q_v^n)\rho_v S, \tag{2.2}$$

$$\frac{d\rho_v}{dt} = -(1 - q_{is}^n)\rho_{is}\rho_v + (1 - q_v^n)S\rho_v, \tag{2.3}$$

$$\frac{d\rho_{is}}{dt} = (1 - q_{is}^n)\rho_v\rho_{is}, \tag{2.4}$$

$$\frac{dR}{dt} = (1 - q_{is}^n)\rho_{is}\rho_v. \tag{2.5}$$

Following arguments very similar to those in Moreno *et al.* (2001), one gets a fixed point equation for the fraction of recovered sites $R_\infty = R(\infty)$ in the stationary state:

$$R_\infty = 1 - (1 - D_0 - \rho_0 + D_0\rho_0)\left(\frac{\rho_0}{R_\infty - D_0 + D_0\rho_0}\right)^{\frac{1-q_v^n}{1-q_{is}^n}}. \tag{2.6}$$

Apart from the trivial solution $R_\infty = D_0 + \rho_0 - D_0\rho_0 = R(0)$, which means that no virus enters the system or at least cannot gain macroscopic areas in sequence space, above the percolation threshold a second solution of the fixed point equation can be found. This is equivalent to the fact that percolation occurs if the following inequality holds (Moreno *et al.* 2001):

$$\frac{1 - q_v^n}{1 - q_{is}^n} > \frac{\rho_0}{1 - R(0)}. \tag{2.7}$$

The theoretical findings have been confirmed by computer simulations with various sets of parameters. In the example of $D_0 = 0.5$, $q_v = q_{is} = 0.95$, $n = 15$, $\lambda = 2$ the simulations lead to a critical immunological density $\rho_0^c = 0.32$ comparing well with the theoretical value from equation (2.7) $\rho_0^c = 1/3$.

Obviously, the immune system generally operates below the percolation threshold, as an adequate immune response can defeat a viral attack before strains spread all over sequence space. Nonetheless, it is reasonable to assume that the immune system operates near the percolation threshold, as unnecessarily high immune receptor densities ρ_0 involve competitive disadvantages.

3. PERCOLATION TRANSITION FROM HIV INFECTION TO AIDS

We are now in the position to extend our model to include HIV dynamics. An HIV model has to take care of characteristic peculiarities of HIV infections, i.e. the destruction of the immune system by the virus. We consider this by extending the algorithm of § 2 by the

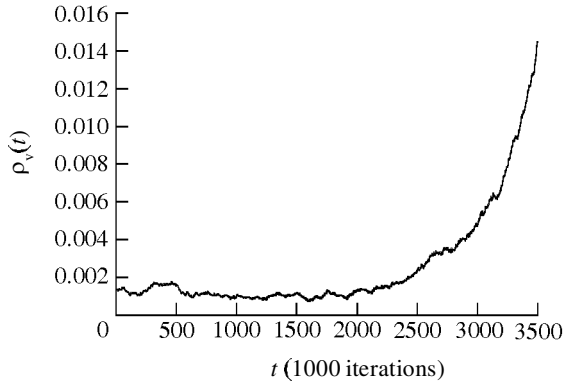

Figure 1. Density of viral strains $\rho_v(t)$ in sequence space under evolution of the system ($D_O = 0.5$, $\rho_O = 0.325$, $q_v = q_{is} = 0.95$, $n = 15$, $\lambda = 2$, $p = 0.0001$).

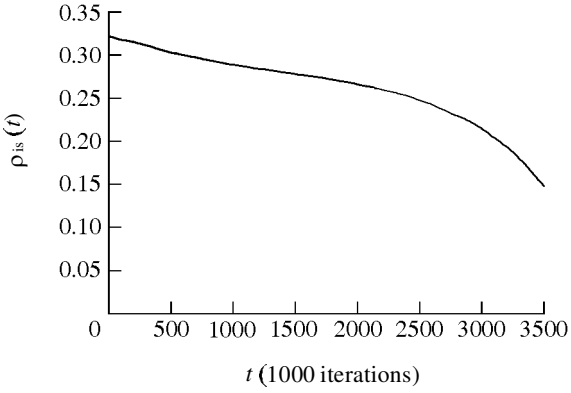

Figure 2. Density of immunologically active sites $\rho_{is}(t)$ in sequence space ($D_O = 0.5$, $\rho_O = 0.325$, $q_v = q_{is} = 0.95$, $n = 15$, $\lambda = 2$, $p = 0.0001$). Note the analogy to the decline in $CD4^+$ cells under HIV infection.

following rule: at any iteration step each viral strain is given a chance to meet a random immunological clone with probability $\rho_{is}(t)$, which thereafter is destroyed with probability p. If the affected site in principle is accessible for a viral strain the viral status changes back to susceptible. We initialize the system near, but below, the percolation threshold, which is the natural state of an immune system that has not yet suffered from an attack by HIV. As the system's qualitative behaviour shows insensitivity to the specific choice of parameters, we choose the parameter settings: $D_0 = 0.5$, $\rho_0 = 0.325$, $q_v = q_{is} = 0.95$, $n = 15$, $\lambda = 2$.

Figures 1 and 2 show simulation results for $p = 0.0001$, exhibiting characteristics typical to the course of disease from HIV infection to the onset of AIDS. One observes a drift of viral epitopes due to immune pressure as found in HIV-infected individuals (Wei *et al.* 1995; Barouch *et al.* 2002). Moreover, the simulations show fluctuations in the total number of actual strains, eventually sharply increasing, which corresponds to the onset of AIDS (Nowak *et al.* 1991). Likewise, it is an empirical fact that the disease progresses with a depletion of $CD4^+$ cells (Vergis & Mellors 2000), which can be assumed to be accompanied by a loss in diversity of the immune repertoire, as shown in figure 2. In this picture, the immune system is successively weakened while fighting the viral attack and ultimately breaks down when the virus begins to percolate in sequence space. The virus dynamically drives the system from a subcritical regime above the percolation threshold.

It will be interesting to investigate the distribution of waiting times until percolation among systems only differing in their random initialization, which corresponds to the incubation period distribution. To understand the generated distribution from a theoretical point of view, we have to take care of the stochastic nature of ρ_v as seen in figure 1. As described in Appendix A, we assume ρ_v to follow a generalized geometric Brownian motion with time-dependent growth rate $r(t)$. This process ρ_v has a lower absorbing boundary for $\rho_v(t) = 2^{-n}$ and converts into exponential growth after having passed an upper point of no return ρ_v^c. The first passage time distribution with respect to the upper boundary corresponds to the incubation period distributions under investigation. It is derived in Appendix A and is discussed in the context of simulation results and empirical HIV data in the following section.

4. RESULTS AND DISCUSSION

The virgin system is infected within a ball that includes, for $n = 15$, 15 one- and

$$\binom{15}{2}$$

two-bit mutants leading to

$$\rho_v(0) = 2^{-15}\left(1 + \binom{15}{1} + \binom{15}{2}\right)(1 - D_0 - \rho_0 + D_0\rho_0)$$
$$\approx 0.0012.$$

A lower absorbing boundary of ρ_v is given by 2^{-15} as less than one viral strain cannot exist. Further evaluation of the simulations yields estimates of $\rho_v^c = 0.002$ where the virus begins to percolate. Taking this together, we will be able to analyse the simulation results from the point of view of first passage time distributions (see Appendix A).

We have run simulations for various choices of p mimicking viruses with different aggressiveness towards the immune system. For p as large as 0.005, we hardly see any time-period of struggle between the immune system and the virus leading to an immediate exponential growth of ρ_v. The system shows very short incubation periods and a vanishing probability of viral defeat. The distribution of incubation periods can then be approximated by a simple inverse Gaussian distribution. Decreasing p leads to longer incubation periods that correspond to periods of combat between virus and immune system as observed in figure 1.

For further discussions, we focus on simulations with $p = 0.0001$ as they show a distribution of incubation periods that are in best accordance with real data on HIV incubation periods.

Figure 3 offers a comparison of a survival function generated by our cellular automaton model with the respective data describing the probability that an HIV-positive patient has not yet developed AIDS at time t after seroconversion (Robert Koch Institut 1998). Note that the survival curve fits for patients not treated with highly active antiretroviral therapy (CASCADE Collaboration 2000*a*,*b*). It shows that the model reproduces the main characteristics of the real system. In particular, the simula-

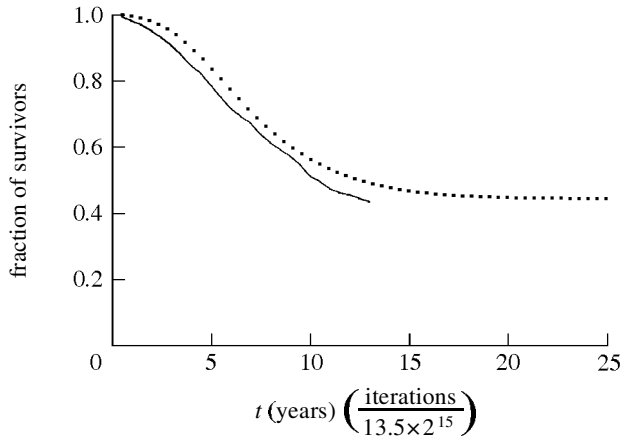

Figure 3. Comparison of the probability for HIV-positive patients of not yet having developed AIDS with a survival distribution generated by our simulations (after adequate renormalization of the time axis, $D_0 = 0.5$, $\rho_0 = 0.325$, $q_v = q_{is} = 0.95$, $n = 15$, $\lambda = 2$, $p = 0.0001$, $\rho_v(0) = 0.0012$, $\rho_v^c = 0.002$). Solid line, HIV survival data; dotted line, simulation results.

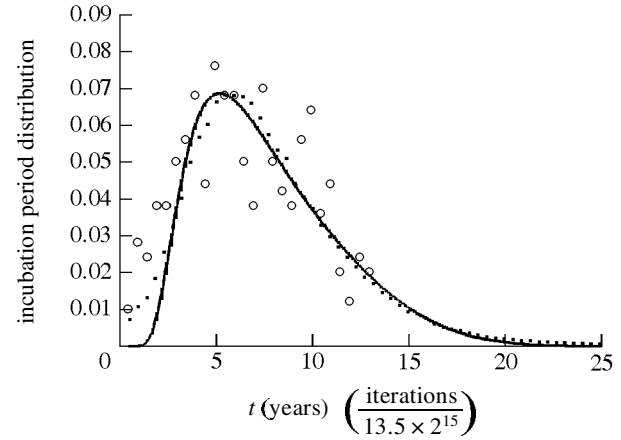

Figure 4. Comparison of the incubation period distributions corresponding to figure 3 with the theoretical model with $r(t) = 0.064 - 0.0092t$, $\sigma^2 = 0.0091$. Open circles, HIV incubation period distribution; dotted line, simulation results; solid line, fit of theoretical distribution.

tions predict the occurrence of long-term survivors as observed in reality and link it to a dynamical percolation mechanism. We would like to emphasize that, in this framework, a quantitative comparison of our model parameters with experimental data is not very meaningful. However, any parameter setting that corresponds to a system that is initially below the percolation threshold and that is attacked with moderate aggressiveness (moderate values of p) will show the same qualitative behaviour. This demonstrates the robustness of our model and ensures its applicability to larger sequence spaces than those simulated here.

To analyse the data in the light of a first passage time distribution, we have to specify the functional form of the viral growth rate $r(t)$. Unlike the case of a very aggressive virus (large p), a constant growth rate $r(t) = \mu > 0$ does not fit the simulation results for viruses that are only moderately destructive (small p). Therefore, let us approximate $r(t)$ underlying the simulations by an expansion in powers of t as $r(t) = \mu + \gamma t$.

This is exemplified by figure 4, where the incubation period distributions corresponding to the survival curves shown in figure 3 are approximated by a first passage time distribution with $\mu = 0.064$, $\gamma = -0.0092$ and $\sigma^2 = 0.0091$. It corresponds to the picture that the viral species is able initially to establish new strains, but that its opportunities for spreading in sequence space are successively diminished. In many cases, the virus nevertheless is able to percolate sequence space if its suppression takes effect too slowly. This happens in a non-deterministic manner due to stochastic fluctuations corresponding to $\sigma^2 > 0$ and generates the observed incubation period distribution. Limitations of the linear approximation become obvious with increasing t, but can easily be handled when considering further terms in the expansion of $r(t)$.

Describing the behaviour of incubation periods within our model, we can summarize that one observes an increase in waiting times before percolation and an enlarged fraction of cases where viral strains go extinct with decreasing p, i.e. less aggressive viral strains. This finds clear correspondence in real HIV statistics. p is a measure of the vulnerability of the immune system under the attack of HIV. This virus is capable of destructive penetration into T-helper cells ($CD4^+$ cells), not only by membrane fusion mediated by CD4, but also it generally needs an additional co-receptor, which is referred to as CCR5. As almost all HIV strains rely on this mechanism for replication in T cells, individuals who show a homozygous mutation leading to a non-expression of the CCR5 receptor have proven to be resistant against HIV infection (Lu *et al.* 1997). This is well in accordance with our model, which, for $p = 0$, predicts that no percolation will occur. More recently it has also been shown that, in individuals with heterozygous genotypes, a slower progression to AIDS can be observed. Moreover, those patients have a reduced risk in maintaining the HIV infection and developing AIDS (Marmor *et al.* 2001). Therefore, a reduction of CCR5 receptors on $CD4^+$ cells, making viral fusion more difficult, already improves the chance for prolonged or even total survival. This fits well with the predictions of the model for a decrease in p.

Recent progress in vaccine research (Amara *et al.* 2001; Lifson & Martin 2002; Shiver *et al.* 2002) further supports the model. From the model's point of view, vaccination corresponds to a local increase in the immune receptors' density ρ_0. This drives the system far below the percolation threshold and, accordingly, HIV will hardly manage to spread in sequence space.

In conclusion, non-trivial aspects of HIV/AIDS phenomenology are predicted within a dynamical cellular automaton model. Prolonged survival, as well as a finite fraction of non-progressors, can be traced back to an enhanced stability below the percolation transition in this framework. Consequently, from the percolation model's point of view, vaccination and receptor blocking are encouraged as efficient strategies to overcome an HIV infection.

C.K. thanks the Stiftung der Deutschen Wirtschaft for financial support.

APPENDIX A: FIRST PASSAGE TIME DISTRIBUTIONS FOR GEOMETRIC BROWNIAN MOTION BETWEEN TWO ABSORBING BOUNDARIES

Facing the stochastic nature of $\rho_v(t)$, we choose an ansatz in the regime before the percolation transition that expects a time-dependent viral growth rate $r(t)$ of $\rho_v(t)$ superposed by noise. Within the Stratonovich interpretation (Øksendal 1998), the underlying stochastic differential equation leads to

$$\rho_v(t) = \rho_v(0)\mathrm{e}^{R(t)+B_t(0,\sigma^2)}, \tag{A 1}$$

$$R(t) = \int_0^t r(t')\mathrm{d}t', \tag{A 2}$$

with $B_t(0,\sigma^2)$ denoting Brownian motion where mean is 0 and variance is $\sigma^2 t$. Accordingly, ρ_v is described by geometric Brownian motion that is locked between two absorbing boundaries at 2^{-n} (less than one strain cannot exist) and an upper critical concentration ρ_v^c that leads to percolation of the virus. This can be translated to Brownian motion $B_t(R(t),\sigma^2)$ (mean $R(t)$ and variance $\sigma^2 t$) with $B_0 = 0$ and limited by

$$-a = \ln\left(\frac{2^{-n}}{\rho_v(0)}\right) < 0,$$

$$b = \ln\left(\frac{\rho_v^c}{\rho_v(0)}\right) > 0.$$

The Fokker–Planck equation for the probability density $p(x,t)$ associated with $B_t(R(t),\sigma^2)$ is solved by an adequate superposition of Gaussians considering the boundary conditions $p(x,0) = \delta(0)$ for all $x \in [a,b]$, $p(-a,t) = 0$ and $p(b,t) = 0$ for all t (reflection principle, mirror charge solution) (Honerkamp 1993). From this, one can derive the probability flow $\mathcal{J}(x,t)$ corresponding to $B_t(R(t),\sigma^2)$ (Gardiner 1983; Honerkamp 1993). $\mathcal{J}(b,t)$ represents the contribution of the probability flow being absorbed at the boundary $b > 0$ at time t. Accordingly, it is the first passage time distribution of the process $\rho_v(t)$ with respect to the upper boundary ρ_v^c, requiring that it has not passed the lower absorbing boundary at 2^{-n}. To quantify $\mathcal{J}(b,t)$, one obtains, after lengthy but canonical calculations,

$$\mathcal{J}(b,t) = \frac{F(a,b,\sigma^2 t)}{\sqrt{2\pi\sigma^2 t^3}}\exp\left(-\frac{[b-R(t)]^2}{2\sigma^2 t}\right), \tag{A 3}$$

$$F(a,b,\sigma^2 t) = \exp\left(\frac{2b(a+b)}{\sigma^2 t}\right)\left\{-a\left[1-\exp\left(\frac{-2b(a+b)}{\sigma^2 t}\right)\right] + b\left[\exp\left(\frac{2a(a+b)}{\sigma^2 t}\right)-1\right]\right\}\Big/\left[\exp\left(\frac{2(a+b)^2}{\sigma^2 t}\right)-1\right] \overset{a\to\infty}{\to} b.$$

Obviously, in the case of only one absorbing boundary (and $r(t) = \mu$, $R(t) = \mu t$) we get the inverse Gaussian distribution as a well-known solution for this special problem (Feller 1968; Redner 2001). A parameter setting of $D_0 = 0.5$, $\rho_0 = 0.325$, $q_v = q_{is} = 0.95$, $n = 15$, $\lambda = 2$,

$$\rho_v(0) = 2^{-15}\left(1+\binom{15}{1}+\binom{15}{2}\right)(1-D_0-\rho_0+D_0\rho_0)$$

≈ 0.0012, as discussed in § 4, leads to $a = 3.7$ and $b = 0.51$. In the range of observation considered in § 4, one finds $F(a,b,\sigma^2 t) \approx b$ corresponding to the substitution of the lower boundary 2^{-n} by zero, justifying this semi-continuous approximation.

REFERENCES

Amara, R. (and 21 others) 2001 Control of mucosal challenge and prevention of AIDS by a multiprotein DNA/MVA vaccine. *Science* **292**, 69–74.

Barouch, D. (and 13 others) 2002 Eventual AIDS vaccine failure in a rhesus monkey by viral escape from cytotoxic T lymphocytes. *Nature* **415**, 335–339.

CASCADE Collaboration 2000*a* Survival after introduction of HAART in people with known duration of HIV-1 infection. *Lancet* **355**, 1158–1159.

CASCADE Collaboration 2000*b* Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. *Lancet* **355**, 1131–1137.

Dangerfield, B. & Roberts, C. 1999 Optimisation as a statistical estimation tool: an example in estimating the AIDS treatment-free incubation period distribution. *Syst. Dyn. Rev.* **15**, 273–291.

Feller, W. 1968 *An introduction to probability theory and its applications*, vol. 1. London: Wiley.

Gardiner, C. 1983 *Handbook of stochastic methods*. Berlin: Springer.

Grassberger, P. 1983 On the critical behaviour of the general epidemic process and dynamical percolation. *Math. Biosci.* **63**, 157–172.

Hershberg, U., Louzoun, Y., Atlan, H. & Solomon, S. 2001 HIV time hierarchy: winning the war while losing all the battles. *Physica* A **289**, 178–190.

Hethcote, H. 2000 The mathematics of infectious diseases. *SIAM Rev.* **42**, 599–653.

Honerkamp, J. 1993 *Stochastic dynamical systems*. New York: Wiley–VCH.

Kamp, C. & Bornholdt, S. 2002 Co-evolution of quasispecies: B-cell mutation rates maximize viral error catastrophes. *Phys. Rev. Lett.* **88**. (DOI 068104.)

Lifson, J. & Martin, M. 2002 One step forward, one step back. *Nature* **415**, 272–273.

Lu, Z.-H. (and 12 others) 1997 Evolution of HIV-1 coreceptor usage through interactions with distinct CCR5 and CXCR4 domains. *Proc. Natl Acad. Sci. USA* **94**, 6426–6431.

Marmor, M., Sheppard, H., Donnell, D., Bozeman, S., Celum, C., Buchbinder, S., Koblin, B. & Seage, G. R. 2001 Homozygous and heterozygous CCR5-Delta32 genotypes are associated with resistance to HIV infection. *J. Acquir. Immune Defic. Syndr.* **27**, 472–481.

Moreno, Y., Pastor-Satorras, R. & Vespignani, A. 2001 Epidemic outbreak in complex heterogeneous networks. LANL e-print archive: http://arxiv.org/abs/cond-mat/0107267.

Nowak, M., Anderson, R., McLean, A., Wolfs, T., Goudsmit, J. & May, R. 1991 Antigenic diversity thresholds and the development of AIDS. *Science* **254**, 963–969.

Øksendal, B. 1998 *Stochastic differential equations*. Heidelberg, Germany: Springer.

Pandey, R. B. & Stauffer, D. 1990 Metastability with probabilistic cellular automata in an HIV infection. *J. Stat. Phys.* **61**, 235–240.

Perelson, A. S. & Weisbuch, G. 1997 Immunology for physicists. *Rev. Mod. Phys.* **69**, 1219–1267.

Philippe, P. 2000 Epidemiology and self-organized critical systems: an analysis in waiting times and disease heterogeneity. *Nonlinear Dyn. Psychol. Life Sci.* **4**, 275–295.

Redner, S. 2001 *A guide to first-passage processes*. Cambridge University Press.

Robert Koch Institut 1998 See http://www.rki.de/infekt/aids_std/epidemio/konvert.htm.

Sartwell, P. 1966 The incubation period and the dynamics of infectious disease. *Am. J. Epidemiol.* **83**, 204–216.

Shiver, J. (and 51 others) 2002 Replication-incompetent adenoviral vaccine vector elicits effective anti-immunodeficiency-virus immunity. *Nature* **415**, 331–335.

Stauffer, D. & Aharony, A. 1992 *Introduction to percolation theory*. London: Taylor & Francis.

Vergis, E. & Mellors, J. 2000 Natural history of HIV-1 infection. *Infect. Dis. Clin. North Am.* **14**, 809–825.

Wei, X. (and 11 others) 1995 Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**, 117–122.

Zorzenon dos Santos, R. & Coutinho, S. 2001 Dynamics of HIV infection: a cellular automata approach. *Phys. Rev. Lett.* **87**. (DOI 168102.)