A COMPREHENSIVE SPATIAL-TEMPORAL INFECTION MODEL

Harisankar Ramaswamy
Aerospace and Mechanical Engineering
Viterbi School of Engineering
University of Southern California
hramaswa@usc.edu

Assad A. Oberai
Aerospace and Mechanical Engineering
Viterbi School of Engineering
University of Southern California
aoberai@usc.edu

Yannis C. Yortsos [*]
Mork Family Department of Chemical Engineering and Materials Science
Viterbi School of Engineering
University of Southern California
yortsos@usc.edu

August 28, 2020

ABSTRACT

Motivated by analogies between the spreading of human-to-human infections and of chemical processes, we develop a comprehensive model that accounts both for infection (reaction) and for transport (mobility, advection and diffusion). In this analogy, the three different populations (susceptible, infected and recovered) of infection models correspond to three "chemical species". Areal densities (people/area), rather than populations, emerge as the key variables, thus capturing the effect of spatial density, widely considered important, but ignored or under-represented in existing models. We derive expressions for the kinetics of the infection rates and for the important parameter R_0, that include areal density and its spatial distribution. Coupled with mobility (through diffusion) the model allows the study of various effects.

We first present results for a "batch reactor", the chemical process equivalent of the SIR model. Because density makes R_0 a decreasing function of the process extent, the infection curves are different and smaller than for the standard SIR model, the difference increasing with R_0. We show that the effect of the initial conditions (density of infected individuals) is limited to the onset of the epidemic, everything else being equal. The same invariance is obtained for infection imported into initially non-infected regions. We derive effective infection curves for a number of cases, including a back-and-forth "commute" between regions of low (e.g. "home") and high (e.g. "work") R_0 environments.

We then consider spatially distributed systems. We show that diffusion leads to traveling waves, which in 1-D geometries (rectilinear or radial) propagate at a constant speed and with a constant shape, both of which are sole functions of R_0. The infection curves are slightly different than for the batch problem, as diffusion mitigates the infection intensity, thus leading to an effective lower R_0. The dimensional wave speed is found to be proportional to the product of the square root of the diffusivity and of an increasing function of R_0, confirming the importance of restricting mobility in arresting the propagation of infection. We examine the interaction of infection waves under various conditions and scenarios, and extend the wave propagation analysis to 2-D heterogeneous systems.

Keywords Infection Models · Chemical Processes · Wave Propagation · COVID-19

---

[*]Corresponding author

1 Introduction

Understanding the spread of infectious diseases, where infection is from human to human, is of significant interest, greatly relevant to epidemics, such as the one the world is experiencing today with COVID-19. A plethora of epidemiologic models have been proposed, developed and tested [1, 2, 3]. These are based on three essential components: (i) Identifying different populations (typically, susceptible, infected, and recovered (or perished), and their extensions to demographic or health history subcategories); (ii) Describing ways by which the various populations come in proximity with one another; and (iii) Postulating rates by which members from one population covert into another.

Widely used is the celebrated SIR model [1, 4], which captures essential aspects of contagion using three populations: susceptible (denoted by S), infected (denoted by I) and recovered (including perished) (denoted by R). The original model, and many of its recent variants, account for items (i) and (iii) above, but not for (ii). More importantly, the SIR representation is in terms of the total populations, rather than their areal densities (people/area), which are expected to be the most important contagion variables. Such models also ignore spatial transfer, or the areal spreading of infections, that ultimately causes populations of different densities to interact with one another.

Another category is based on computing individual trajectories of typical members of the various populations and on subsequently postulating probabilities of collisions and infection [5]. These agent-based models implicitly account for spatial density effects through various mobility, propagation and collision rules. In particular, for specific individuals, whose trajectories and potential contacts are known, one can predict or retrace past infection paths. This forms the basis of contact tracing methods.

A different and, we believe, more fundamental way to model the problem, is to follow a chemical reaction engineering approach, in which one can cast contagion spreading in terms of corresponding transport and reaction models. This approach relies on the following analogies: populations map into chemical species; densities (specifically, areal densities) into molecular concentrations; infection rates into chemical reaction rates (where mass action kinetics apply); and spatial transport into advective and diffusive (or dispersive) fluxes. Then, one can express the relevant population (species) balances using a partial differential equation description. It is the objective of this paper to create such a methodology for the description of the spread of an epidemic in general, with application to COVID-19, in particular. The work presented parallels related previous continuum models [6] but differs in a number of aspects to be explicitly discussed below.

A key, fundamental underlying assumption is that we can homogenize population distributions by defining continuum variables in terms of their areal averages. This allows one to postulate continuum conservation laws in terms of their rate of change, transport and reaction. We recognize the limitations in this analogy between transport and interaction of human populations on the one hand and molecular transport and reaction on the other. Given the relatively sparse areal densities of human populations, compared to molecular densities, the law of large numbers may in fact not be in effect, in order to obtain well-defined averages (homogenization) [7], in which case the corresponding distributions might be akin to rarefied gas dynamics [8]. In addition, human movement is often influenced by behavioral drivers, rather than random walks, and it is possible that, e.g. a diffusion or dispersion process would require a description different than Brownian motion. These alternatives will be pointed out where appropriate. Despite such limitations, however, we believe that the analogy pursued here is useful and instructive, as it allows one to obtain important new insights on fundamental spatial aspects of the problem.

The commonly used SIR model arises from our formulation by integrating the relevant partial differential equations over specific control areas (e.g. a building, a manufacturing plant, a school, a city, a state, a country, etc.), namely by considering a "batch reactor" model, where one assumes perfect mixing [9]. We examine its validity. Infection rates, depending crucially on spatial (areal) density, lead to spatial non-uniformities, while transport leads to infection waves. Both impact the stationary assumption of the SIR models. More importantly, we find that the key variable R_0 is not constant during the process, contrary to the common assumption. Instead, it is found to depend on the areal spatial density and to decrease as a function of the extent of the contagion. We consider the solution of spatially-dependent problems by including diffusion in both 1-D (rectilinear or radial) and in 2-D geometries. We discuss the asymptotic emergence of traveling infection waves, the speed and shape of which are found to be independent of geometry, and also correspond to a slightly smaller effective value of R_0, than for the case of the batch SIR model.

The paper is organized as follows: We first proceed with constructing an equivalent chemical reaction and transport process. After recasting all conservation equations in dimensionless form, we derive the associated key parameter R_0. Then, we consider the solution of a number of specific problems, from a "batch reactor" model (the SIR equivalent), to the propagation of infection in time and space (including both 1-D and 2-D domains). Infection waves are found to depend only on R_0, or its effective equivalent, and for batch or 1-D processes, to be independent of initial conditions. This lack of influence of initial conditions on the shape of the infection curve (in space and/or in time) is notable, as it signals universal scaling properties, a property implicitly assumed in many SIR-type models.

2 Formulation

Consider the conservation of mass of the three quantities of interest, susceptible, infected, and recovered (including perished), and associate to them three equivalent "chemical species". The corresponding mass balances read

$$\frac{\partial N_i}{\partial T} + \nabla \cdot (\boldsymbol{q} N_i) = -\nabla \cdot (\boldsymbol{D}_i) + r_i, \qquad i = I, S, R. \tag{1}$$

where N_i is population density (number/area), $\boldsymbol{q}_i$ (if any) is an advective velocity vector, $\boldsymbol{D}_i$ is the diffusive (or dispersive) flux, and r_i is the net reaction rate of species i, and converts populations into one another due to infection and recovery, respectively. The advective velocity $\boldsymbol{q}_i$ denotes a transport or mobility term. While we allow in (1) the three velocities to be different, in the remainder we will take $\boldsymbol{q}_i = \boldsymbol{q}$ for all populations. This common velocity is assumed independent of N_i or its gradients, although it can be a function of space and/or time. But we also note in advance, that in the remainder we will not explicitly examine any effects of advection, included here only for the sake of completeness.

Next, define

$$N_S + N_I + N_R = \rho, \tag{2}$$

and

$$\boldsymbol{D}_S + \boldsymbol{D}_I + \boldsymbol{D}_R = \boldsymbol{D} \tag{3}$$

where ρ is the total population density (total number/area) and $\boldsymbol{D}$ is the total diffusive (or dispersive flux), to obtain

$$\frac{\partial \rho}{\partial T} + \nabla \cdot (\boldsymbol{q}\rho) = -\nabla \cdot (\boldsymbol{D}) \tag{4}$$

as the equation governing the evolution of total density.

2.1 Diffusion

How to represent the diffusive (or dispersive) fluxes requires some discussion. For a typical Fick's law type of diffusion [10], one may take

$$\boldsymbol{D}_i = -D\nabla N_i, \tag{5}$$

where we defined a constant diffusion (or dispersion) coefficient D, and assumed that all species are indistinguishable as far as their physical properties is concerned. Then, Equation (4) becomes

$$\frac{\partial \rho}{\partial T} + \nabla \cdot (\boldsymbol{q}\rho) = \nabla \cdot (D\nabla\rho) \tag{6}$$

suggesting that the total population density ρ, in addition to being advected, also diffuses in the direction of a negative spatial gradient. While possible and perhaps even likely in human dynamics, the suggestion that the overall density may diffuse, is contrary to the common continuity equation for fluids, e.g. incompressible fluids [11]. If we consider an alternative approach and take instead

$$\boldsymbol{D}_i = -D\rho\nabla(N_i/\rho), \tag{7}$$

the more familiar form emerges

$$\frac{\partial \rho}{\partial T} + \nabla \cdot (\boldsymbol{q}\rho) = 0 \tag{8}$$

Diffusive fluxes of the type (7) do in fact arise in random walks on non-uniform lattices, in the limit when the lattice becomes continuous [12, 13]. While either (5-6) or (7-8) can be taken to describe diffusive transport, in the examples below we will only assume the more familiar continuity equation versions (7) and (8). This has the additional advantage that in the absence of advection, the total density is time-independent and a stationary function of space, which is useful for the solution of many problems of interest, as the equations now become decoupled.

It must be also noted that diffusion or dispersion can occur by different than a Brownian motion type random walk, e.g. by Levy flights (where the distribution of diffusion steps has a heavy tail, thus allowing for occasional, although rare, large steps) [14]. Modeling such motions can be handled by an integro-differential equation description through fractional derivatives [15, 16]. While worth considering, such an approach will not be pursued here.

Diffusion coefficients can be estimated by considering the ratio of the square of the average radius of a random walk over an associated time interval. For instance, for an office type environment where random walks may have a mean radius of $10m$, over an $8 - hr$ period, one obtains $D \approx 10^{-2} m^2/s$, which is about three orders of magnitude larger than that for molecular diffusion in gases. A different approach to evaluate dispersion at much larger scales is by inference from the spatial spread of the epidemic. As we describe later, diffusion leads to infection waves, whose propagation speed depends on the assumed diffusion coefficient. Values of the order of $D \approx 10^3 m^2/s$ or higher are derived in order to match the observed propagation velocities. Clearly, these very large values reflect an aggregate mix of transport activities (including advection), here expressed only through a diffusion mechanism, which, therefore must be viewed only as an approximation.

2.2 Reaction Rates

Consider, next, contagion rates. Following our chemical reaction analogy, we postulate the following two chemical reactions

$$
\begin{aligned}
S + I &\rightarrow 2I & (9) \\
I &\rightarrow R & (10)
\end{aligned}
$$

that convert the three species to one-another. The stoichiometric coefficient 2 in the RHS of (9) indicates that one member of infected species I is produced as a result of its interaction with one member of susceptible species S. In turn, species I, consumed in reaction (10), leads to species R, produced in one-to-one stoichiometry. Both reactions are irreversible.

Applying mass action kinetics [17] in the reactions (9) and (10) provides expressions for the reaction rates in terms of the respective concentrations or densities

$$
r_S = -K N_S N_I, \qquad r_I = K N_S N_I - \Lambda N_I, \qquad r_R = \Lambda N_I, \tag{11}
$$

where we introduced the reaction rate constants K and Λ. Equations (11) are the same as those for the SIR model, except that here the rates are correctly expressed in terms of areal densities, rather than in terms of the total populations, which is inaccurately assumed in the typical models.

We hasten to note that for a more fine-grained model, the reaction rates would also depend on additional demographics and/or health conditions of the individual species. Such fine-graining is possible by further partitioning the populations into subgroups [18]. The corresponding description would then be in terms of an equivalent "multicomponent mixture" [19], where equations (1) are recast in terms of a larger species vector, with reactions (9)-(10) extended appropriately, possibly involving a product of a reaction matrix with the species vector. For simplicity, this generalization will not be further considered here.

Relevant to linear ("first-order") reactions, Λ has dimensions of inverse time, a typical value being $1/14 \ days^{-1}$ [6, 20, 21]. The kinetic parameter expresses the rate at which infected individuals recover (or die), on average, and is intrinsic to the infected fraction. This is not the case for non-linear (e.g. "second-order") reactions, like those involved in the rate of generation of new infections (reaction (9)). As reflected in the kinetic parameter K, which has dimensions of inverse (time $\times$ (number/area)), the infection kinetics will depend on the duration, method and type of human-to-human contact, the protective gear (PPE) of susceptible and infected individuals, various biological and physiological variables, the ambient environmental conditions (room air conditioning), spatial distancing and other parameters. The most controllable among these factors are the frequency and degree of of encounters (collisions) between individuals, as well as the intensity of interaction.

Possible approaches to estimating K include kinetic theory models (e.g. similar to Maxwell-Boltzmann models, where the kinetic parameter is inversely proportional to the molecular mean free path (average length between collisions), itself inversely proportional to density [22]). At least in some domain, K should be increasing with spatial density, something ignored in previous SIR models. Directly applying Maxwell-Boltzmann-type kinetic theories, however, is not possible in the present context and must be refined: Human encounters are not elastic collisions, and typically last over finite time intervals. More importantly, effects of spatial distancing, a recognized key to the kinetics of contagion, must be captured in any model. Indeed, it is by now widely accepted that infection rates are negligible for densities below a limiting value (corresponding to a separation of $2m$ or $6ft$, as also supported by fluid mechanical models of droplet propagation, and recommended in health policy guidelines [23, 24, 25]).

We incorporate such aspects by postulating the following dependence

$$
K(\rho) = \left\{ \begin{array}{ll} 0, & \rho < \rho_0 \\ K_0 F(\frac{\rho - \rho_0}{\rho_1 - \rho_0}), & \rho \geq \rho_0 \end{array} \right. \tag{12}
$$

where $F(x)$ is an increasing function of x, $F(0) = 0$, $F(1) \approx 1$. The value $\rho_0 \approx 0.1 m^{-2}$ represents the density below which the reaction rate is negligible, while the upper limit $\rho_1 \approx 1 m^{-2}$ corresponds to the maximum "packing" density. In (12) we have separated spatial distancing effects, included through ρ, from factors, such as biological, environmental, facial covering, etc., which enter through K_0 only (with K_0 decreasing substantially as facial covering is applied).

Equation (12) captures both spatial distancing and biological and environmental effects. Upon more careful inspection, however, we realize that the density dependence must be further modified to not include the recovered fraction component, since that fraction does not affect contagion (assuming immunity in all recovered individuals). Therefore, we must modify the way density enters in (12) by replacing it with $\rho(1 - r)$ (where we defined $r = \frac{R}{\rho}$, see also below). The new expression (replacing $K(\rho)$ with $K(\rho(1 - r))$) reads

$$K = K(\rho, r) = \begin{cases} 0, & \rho(1-r) < \rho_0 \\ K_0 F(\frac{(1-r)\rho - \rho_0}{\rho_1 - \rho_0}), & \rho(1-r) \geq \rho_0 \end{cases} \tag{13}$$

When the rate of infection is relatively low (and $r \ll 1$), the correction is negligible. For strong infection rates, on the other hand, it can be quite significant, as will be demonstrated later in the paper.

Some additional remarks are pertinent: The above assume that an infected individual can infect a susceptible one at the same constant rate. This is true either for asymptomatic, infected individuals, or for those who do not exhibit symptoms until a few days following infection. It is not true, however, when infected individuals are isolated. Nonetheless, the previous formalism still applies: Assume that the infected species N_I is further subdivided into one category containing asymptomatic individuals (denoted by A, with corresponding density N_A) and another containing quarantined individuals (denoted by Q, with corresponding density N_Q). The associated infection reaction rate is now $K N_S N_A$. Based further on the reasonable assumption that the percentage of those infected, but are asymptomatic or with mild symptoms, is a fixed fraction (e.g. a) of the total fraction of the infected, the infection rate becomes $K a N_S N_I$. This is of the same dependence as before, hence the previous holds, except that the kinetic constant is now modified by parameter a. The same reasoning (with an additional parameter included) holds when infected individuals are contagious, but not identified as such, until sometime after infection. On the other hand, our approach cannot as easily account for correlations between infected and susceptible individuals, for example when susceptible individuals have increased contact, hence higher probability of infection, with specific infected individuals related to them, e.g. by family, work or other proximity relations.

A final remark relates to the practice of reporting area-wide averages (e.g. for states or countries). Given that almost all areas will never on average reach the minimum density required for infection (e.g. $0.1 m^{-2}$), area-wide averages over substantially heterogeneous density distributions are not very informative. Rather, distinguishing high-density areas (e.g. urban places, stadiums, schools, retirement homes, etc.) from low-density ones (e.g. farms, rural) is essential, with the reporting of much more fine-grained area statistics being much more informative. Connecting the transmission of infection between areas of different density is possible using the present formalism, which includes spatial transport and/or diffusion, and where K is space-dependent, as further discussed below.

2.3 Dimensionless formulation

We next proceed with a dimensionless notation. Denote dimensionless variables by lower case symbols and normalize species densities by ρ, time by $1/\Lambda$, space by a characteristic external length scale l, K by K_0, and velocities by a characteristic velocity q. Using equation (7) for diffusion we obtain

$$\frac{\partial s}{\partial t} + Da \nabla \cdot (\boldsymbol{v}s) - C\nabla(\ln \rho) \cdot \nabla s = \nabla \cdot (C\nabla s) - R_0(\rho, r)si \tag{14}$$

$$\frac{\partial i}{\partial t} + Da \nabla \cdot (\boldsymbol{v}i) - C\nabla(\ln \rho) \cdot \nabla i = \nabla \cdot (C\nabla s) + R_0(\rho, r)si - i \tag{15}$$

$$\frac{\partial r}{\partial t} + Da \nabla \cdot (\boldsymbol{v}r) - C\nabla(\ln \rho) \cdot \nabla r = \nabla \cdot (C\nabla r) + i \tag{16}$$

$$\frac{\partial \rho}{\partial t} + Da \nabla \cdot (\boldsymbol{v}\rho) = 0 \tag{17}$$

Here, we defined the dimensionless Damkohler number $Da = \frac{ql}{\Lambda}$, the dimensionless diffusion number $C = \frac{D}{\Lambda l^2} = \phi^{-2}$, and the rescaled velocity $\boldsymbol{v}$. Note that in the chemical reaction engineering literature, ϕ is known as the Thiele modulus [9]. From the dimensionless formulation arises a most important parameter, and one most commonly associated with the SIR model, namely

$$R_0(\rho, r) = \frac{K_0 \rho}{\Lambda} \kappa(\rho, r) \tag{18}$$

5

where

$$\kappa(\rho, r) = \begin{cases} 0, & (1-r)\rho < \rho_0 \\ F(\frac{(1-r)\rho-\rho_0}{\rho_1-\rho_0}), & (1-r)\rho \geq \rho_0 \end{cases} \tag{19}$$

Equation (18) shows that $R_0(\rho, r)$ is not constant, as typically assumed, but also depends both on the areal density and on the extent of the process r. For example, the ratio of the final value $R_0(\rho, r_\infty)$ to its initial $R_0(\rho, 0)$ can be approximated (assuming a linear function for F) by

$$\frac{R_0(\rho, r_\infty)}{R_0(\rho, 0)} = \frac{(1-r_\infty)\rho - \rho_0}{\rho - \rho_0} \approx 1 - r_\infty \tag{20}$$

assuming $\rho_0/\rho \ll 1$. For consistency, in the remainder, we will express the R_0 dependence of various solutions to the problem, e.g. r_∞, through its value at the onset of the process, $R_0(\rho, 0)$. The finding that $R_0(\rho, r)$ is not constant is consistent with spatial distancing health policy guidelines. To our knowledge this property has not been noted before, even though it has important implications to the prediction of infection results, as shown below.

For completeness, we also remark that had we used the different version for diffusion (5), instead, then equations (16), (15) and (16) would still remain in effect, subject to multiplying by a factor of 2 the third term on the LHS of each equation.

3 Applications

Consider, now, the application of the previous formulation to special cases, starting first with the zero-dimensional ("batch reactor") perfect-mixing problem.

3.1 Perfect Mixing: The SIR Model

Assume a "batch reactor" (closed model, with no input or output) and conditions of perfect mixing, in which all spatial densities only depend on time. By integrating the differential equations (14)-(16) and using (˙) to denote time derivatives, we obtain the following system of ordinary differential equations

$$\dot{s}(t) = -R_0(\rho, r)si \tag{21}$$
$$\dot{i}(t) = R_0(\rho, r)si - i \tag{22}$$
$$\dot{r}(t) = i \tag{23}$$

subject to the closure

$$s + i + r = 1 \tag{24}$$

and the initial conditions

$$i(0) = i_0, \qquad s(0) \equiv s_0 = 1 - i_0, \qquad r(0) = 0. \tag{25}$$

Even though the overall density is time-independent, in such an SIR-like model, spatial effects enter through the effect of the extent of contagion on R_0 (equations (21), (22)).

Equations (21)-(23) produce non-trivial results when an initial, even infinitesimally small, seed of infected individuals (i_0) is present. An analytical solution is possible. Substitute (23) into (21) and integrate to give:

$$s = s_0 \exp(- \int_0^r R_0(\rho, r')dr') \tag{26}$$

thus,

$$i = 1 - r - s_0 \exp(- \int_0^r R_0(\rho, r')dr') \tag{27}$$

Further substitution into (23) gives

$$\dot{r}(t) = 1 - r - s_0 \exp(- \int_0^r R_0(\rho, r')dr') \tag{28}$$

which can be integrated to the final solution

$$t = \int_0^r \frac{du}{1 - u - s_0 \exp(- \int_0^u R_0(\rho, r')dr')}. \tag{29}$$

We remark, in passing, that by expanding the exponential in (28) in a Taylor series, assuming constant R_0 and keeping the first three terms in the expansion, leads to a Riccati equation

$$\dot{r}(t) = 1 - r(1 - s_0 R_0(\rho, 0)) - \frac{s_0 R_0(\rho, 0)^2}{2} r^2 \tag{30}$$

accurate for small R_0. Such a model has been used as another alternative to the SIR description [26, 27, 28].

3.1.1 Infection Curves, Herd Immunity and Effective R_0

Results from the solution of the above equations are shown in Figure 1. We first remark that from Equation (22), $R_0(\rho, 0) = 1$ is the boundary demarcating two regions, where an initial infection either decays ($R_0(\rho, 0) < 1$) or grows ($R_0(\rho, 0) > 1$). We will focus on the latter case ($R_0(\rho, 0) > 1$). Plotted in Figure 1 is the time variation of the three different populations for $R_0(\rho, 0) = 2.5$, as well as of the curves obtained under the SIR assumption of constant R_0 (taken throughout the process to be equal to $R_0(\rho, 0) = 2.5$). The infection curves are of a similar shape, but significantly larger when the constant R_0 value is used.
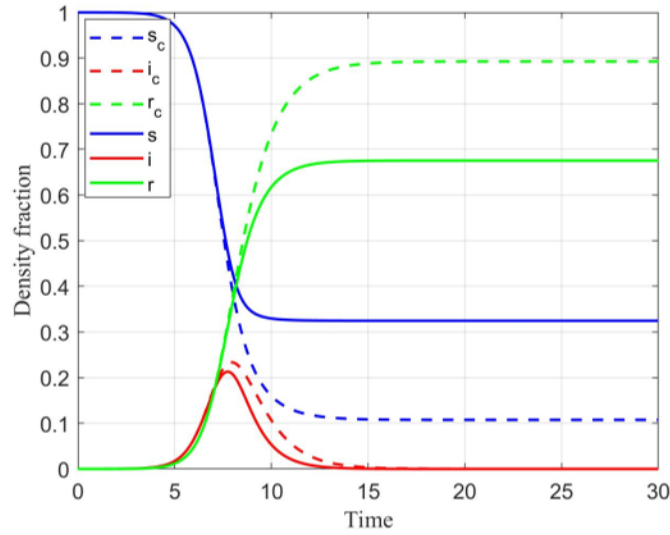
Figure 1: Variation of susceptible (s), infected (i) and recovered (r) fractions as a function of non-dimensional time for the case when R_0 varies following equation (19) (solid lines) with $R_0(\rho, 0) = 2.5$, and for the case of a constant value of $R_0 = R_0(\rho, 0) = 2.5$ (dashed lines), which is the SIR assumption. Note the substantial difference between the results obtained. The initial infected fraction is $i(0) = 10^{-5}$.

Of importance is the concept of *herd immunity*, denoted here by $h = r_\infty$, and defined as the asymptotic value of the recovered individuals, r_∞. The latter is the solution of the equation

$$1 - r_\infty - s_0 \exp(-\int_0^{r_\infty} R_0(\rho, r')dr') = 0. \tag{31}$$

Results are shown in Figure 2. We note that h is an increasing function of $R_0(\rho, 0)$ (with $h \to i_0$ as $R_0(\rho, 0) \to 0$). Also plotted in the same figure is the herd immunity calculated under the SIR assumption of a constant $R_0 = R_0(\rho, 0)$. The corresponding values are significantly higher, even for relatively mild rates of infection. An effective constant R_0 can be defined using the following average

$$R_0 \equiv \frac{\int_0^{r_\infty} R_0(\rho, r')dr'}{r_\infty} \tag{32}$$

to relate an effective constant R_0 to $R_0(\rho, 0)$. This relationship can be also inferred from Figure 2, which for example shows that to a constant value of $R_0 = 2.5$ corresponds a twice as large value of $R_0(\rho, 0) = 4.5$. Plotted in Figure 2 is also the maximum in the infected fraction, $i_{\max}$, which is an increasing function of $R_0(\rho, 0)$, as expected.

A related quantity is the length of the epidemic, here approximated by

$$t_\infty = \int_0^{r_e} \frac{du}{1 - u - s_0 \exp(-\int_0^u R_0(\rho, r')dr')}. \tag{33}$$

7

where $r_e = 0.99 r_\infty$, and r_∞ is the solution of (31). Results are shown in Figure 2. The dependence of t_∞ on $R_0(\rho, 0)$ is monotonically decreasing, with smaller values of $R_0(\rho, 0)$ (and lower infection levels) resulting into a longer duration, as long as 30 in dimensionless time (corresponding roughly to 15 months) or longer. A more typical duration, but with higher infection rates, is about 12 dimensionless time units (6 months). The fact that the epidemic is higher at smaller values of R_0 is counter-intuitive, and calls for the need to educate the public that epidemics on the border of being under control will last longer.
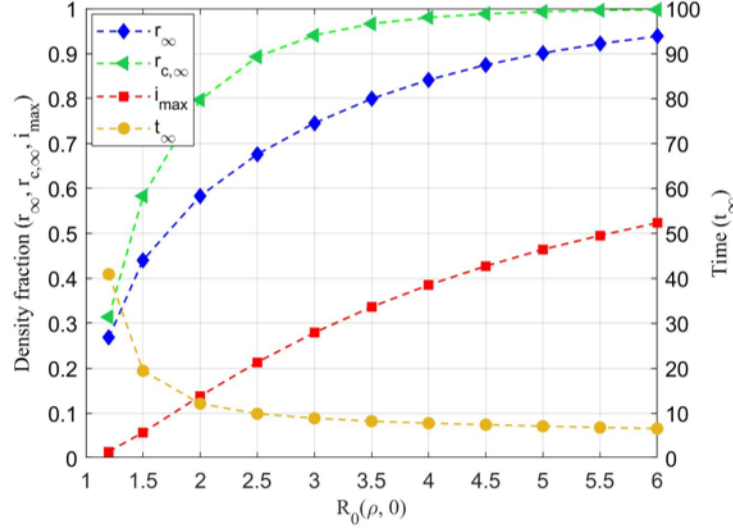
Figure 2: Variation of herd immunity (r_∞), maximum infection fraction (i_{max}) and the duration of a pandemic (t_∞) as a function of $R_0(\rho, 0)$. Plotted also is the herd immunity ($r_{c,\infty}$) calculated assuming a constant R_0. The initial infection fraction is $i(0) = 0.001$.

Of relevance is the stability of the asymptotic state. Consider a small resurgence of infections after an asymptotic state is reached. As long as condition $R_0(\rho, r_\infty)(1 - r_\infty) < 1$ is satisfied (which is always the case), such fluctuations will decay exponentially fast, as shown in equation (22) demonstrating that the final state is asymptotically stable. Such will not be the case, however, when the fluctuation is instead in $R_0(\rho, r_\infty)$, e.g. when a new value, e.g. R_0', such that $R_0'(1 - r_\infty) > 1$, sets in. For example, this could be the result of relaxations to spatial distancing, and of abandoning due caution, after the asymptotic state is reached. Then, there will be an eruption of new cases, resulting into a new spreading of infections (a "second wave"), which will in turn reach a new asymptotic state, and a correspondingly new, higher herd immunity. Figure 3 demonstrates the development of such a second wave. The figure shows three different regimes: An initial one with $R_0(\rho, 0) = 3$ for which infection grows; a second regime following the imposition of restrictions at $t = 1$ (which leads to $R_0 = 0.8$), and which results in the "flattening of the curve" with a corresponding herd immunity of 0.3; and a regime in which relaxation of restrictions at $t = 10$ and a return to a higher $R_0(\rho, 0) = 3$, leads to a second wave. The new epidemic lasts at least as long as the first one, and contributes to substantial new infections (almost double the initial). This lack of structural stability of the asymptotic state is derives from the fact that the infection reaction (9) is autocatalytic, and leads, at least initially, to exponential rises. It illustrates the importance of closely adhering to a consistent, and as low as possible, value of R_0, for a desired herd immunity to be sustained.

3.1.2 Universal Scaling

The autocatalytic nature of reaction (9) raises the additional question as to whether or not the infection curves depend on parameters other than R_0. Consider, first, the impact of the initial condition i_0. Figure 4 is a plot of the infection curve for different values of the initial condition for $R_0(\rho, 0) = 2.5$. For the typical values considered, we observe that a decrease in the initial condition fraction leads to a shift in the infection curve $i(t)$ to the right, but otherwise produces shapes that remain invariant. The shift is approximately 2 non-dimensional time units for each decrease in the initial condition by a factor of 10.

To explain these findings we consider the early-time solution of equation (22). At small times we obtain

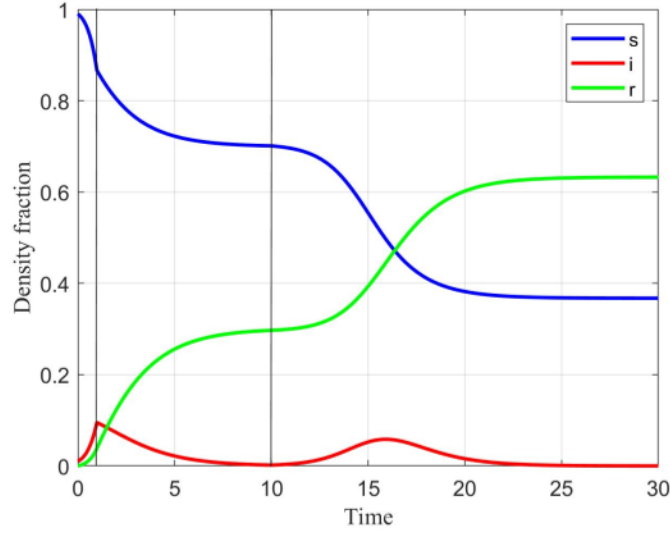$$i(t) \approx i_o \exp((R_0(\rho, 0) - 1)t) = \exp((R_0(\rho, 0) - 1)(t - t_0)), \tag{34}$$

8

Figure 3: Variation of susceptible (s), infected (i) and recovered (r) fractions as a function of non-dimensional time with $R_0(\rho, 0) = 3, t \in (0, 1), R_0(\rho, 0) = 0.8, t \in (1, 10), R_0(\rho, 0) = 3, t \in (10, 30)$. The initial infected fraction is $i(0) = 0.01$.

where $t_0 = -\frac{\log(i_0) \ln(10)}{R_0(\rho, 0) - 1} \approx -2.3 \frac{\log(i_0)}{R_0(\rho, 0) - 1}$. This confirms the existence of a time shift t_0 of the same magnitude as in the figure, consistent with the simulations. For the same reasons, the maximum infection fraction i_{max} is independent of the initial condition i_0, assuming that the latter remains relatively small. The invariance in the shape of the infection curve and the fact that a decrease in the number of initial infections only acts to delay the onset of infection, are significant from a health policy perspective: containing the initial number of infections helps to provide a non-trivial cushion of time to contain the infection and educate the public to modify behavior by lowering R_0, thus ultimately mitigating the intensity of the incoming infection. Absent such preparation or behavior modification, will negate any beneficial effect of the lower number of initial infections: A corresponding contagion wave will emerge, ultimately dependent only on R_0.
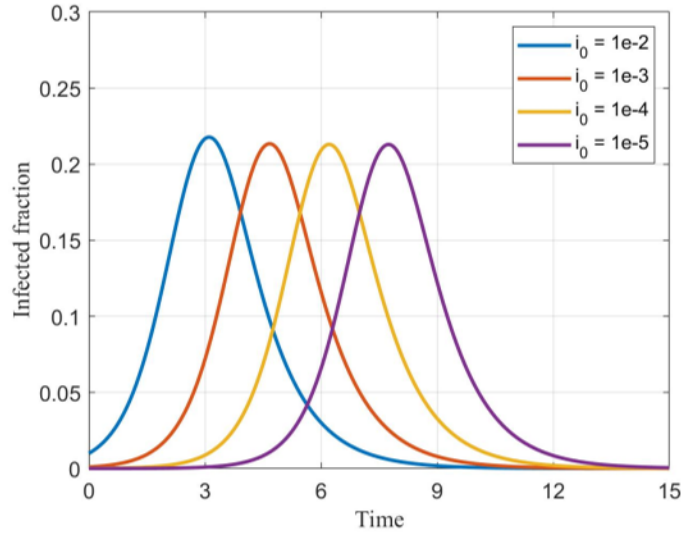

Figure 4: Variation of infected (i) fraction as a function of non-dimensional time for different values of the initial infected fraction i_0 and $R_0(\rho, 0) = 2.5$.

3.1.3 Imported Infection

The same conclusions apply for "imported infection". With this term, we refer to the case where a control area (e.g. a country), originally without any infected individuals, receives a constant influx of infected and susceptible individuals over a finite (and small, reflecting a prompt public authority response) time interval τ, following which such influx is stopped (e.g. via a "flight ban"). By integrating equation (15) over spatial dimensions and assuming perfect mixing we find to leading order

$$\dot{i}(t) = (R_0(\rho, 0) - 1)t + j_i \tag{35}$$

valid for $0 < t < \tau$. Here, we denoted the influx rate by j_i and also assumed that the corresponding changes in density are negligible, as long as τ is small. Solving equation (35) subject to the initial condition $i(0) = 0$, we obtain, to leading-order (small τ)

$$i = j_i t + \cdots \tag{36}$$

It follows that as long as $R_0(\rho, 0)\tau \ll j_i$, the impact of "imported infections" is simply to nucleate the process, essentially providing the equivalent initial condition, $i(0) = i_0 = j_i\tau$, and the problem reverts to the previous. We conclude that the net impact of a ban is to provide additional time (e.g. a month) for the necessary public health measures to be implemented and to possibly lead to as small a value of $R_0(\rho, 0)$ as possible, thereby limiting the spread of the inevitable infection. In the absence of additional such prophylactic measures, a ban can only serve to delay the onset of contagion.

3.2 Fluctuations

The batch (SIR-like) model rests on the assumption that all profiles are spatially uniform. We explore the validity of this assumption in the presence of fluctuations or other non-uniformities, by considering the solution of a typical 1-D problem. We take

$$\frac{\partial s}{\partial t} - C\frac{\partial \ln \rho}{\partial x}\frac{\partial s}{\partial x} = C\frac{\partial^2 s}{\partial x^2} - R_0(\rho, r)si \tag{37}$$

$$\frac{\partial i}{\partial t} - C\frac{\partial \ln \rho}{\partial x}\frac{\partial i}{\partial x} = C\frac{\partial^2 i}{\partial x^2} + R_0(\rho, r)si - i \tag{38}$$

$$\frac{\partial \rho}{\partial t} = 0 \tag{39}$$

subject to no-flux conditions at the two ends, $\frac{\partial s}{\partial x} = \frac{\partial i}{\partial x} = \frac{\partial \rho}{\partial x} = 0$, at $x = \{0, 1\}$. We consider the case where the initial fluctuations are variable, but the density is uniform,

$$i(x, 0) = i_0(1 + \epsilon g(x)), \qquad s(x, 0) = 1 - i(x, 0), \qquad \rho(x, t) = \rho_m, \tag{40}$$

or where the non-uniformity is in the density profile only

$$i(x, 0) = i_0, \qquad s(x, 0) = 1 - i(x, 0), \qquad \rho(x, t) = \rho_m(1 + \epsilon g(x)), \tag{41}$$

where $\epsilon \ll 1$. Expecting that diffusion ($C > 0$) will help to smooth spatial non-uniformities, we will focus in this section on the solution in the absence of diffusion ($C = 0$). Thus, any effects to arise will correspond to the spatial averaging of an ensemble of batch problems.

3.2.1 Stability

When $C = 0$ the relevant equations revert to (21)-(23). Consider a small perturbation expansion and denote means by superscript bar and fluctuations by superscript prime to obtain

$$\frac{\partial \bar{s}}{\partial t} = -R_0(\rho_m)(\bar{s}\bar{i} + \overline{s'i'}) \tag{42}$$

$$\frac{\partial \bar{i}}{\partial t} = +R_0(\rho_m)(\bar{s}\bar{i} + \overline{s'i'}) - \bar{i} \tag{43}$$

where we ignored the effect of the fluctuations on R_0. By further taking the representation $s' = \epsilon i_0 \sigma(t)g(x)$ and $s' = \epsilon i_0 \eta(t)g(x)$, the fluctuations to leading order satisfy

$$\frac{d\sigma}{dt} = -R_0(\rho_m)(\bar{s}\eta + \bar{i}\sigma) \tag{44}$$

$$\frac{d\eta}{dt} = +R_0(\rho_m)(\bar{s}\eta + \bar{i}\sigma) - \eta \tag{45}$$

with corresponding initial conditions $\sigma(0) = -1$ and $\eta(0) = 1$.

We first note that fluctuations contribute to the rate expression, Equations (42)-(43), through the mean of the product $\overline{s'i'}$. This is negative and proportional to the square of the amplitude of the fluctuations $\overline{(i_0 \epsilon g(x))^2}$. A first conclusion, therefore, is that fluctuations lower the effective rate of the infection reaction (although only to order ϵ^2 assuming that the fluctuations remain bounded). Adding diffusion will further reduce any such impact.

To determine whether or not the fluctuations are bounded, we must find the eigenvalues ω of the matrix of the linear system (44)-(45), namely the solutions of

$$\omega^2 + (R_0(\rho_m)\bar{i} - R_0(\rho_m)\bar{s} + 1) + R_0(\rho_m)\bar{i} = 0. \tag{46}$$

The equation has two negative eigenvalues as long as

$$R_0(\rho_m)\bar{i} - R_0(\rho_m)\bar{s} + 1 > 0. \tag{47}$$

Equations (21)-(23) show that for some time before the infection fraction i reaches its maximum, and always after that time, condition (47) is indeed valid. We conclude that the fluctuations will remain bounded, if not altogether decay, hence they have little effect on the average behavior. This finding is demonstrated in Figure 5, where we plot ensemble average responses. The curves obtained are practically the same either in the presence or the absence of fluctuations.
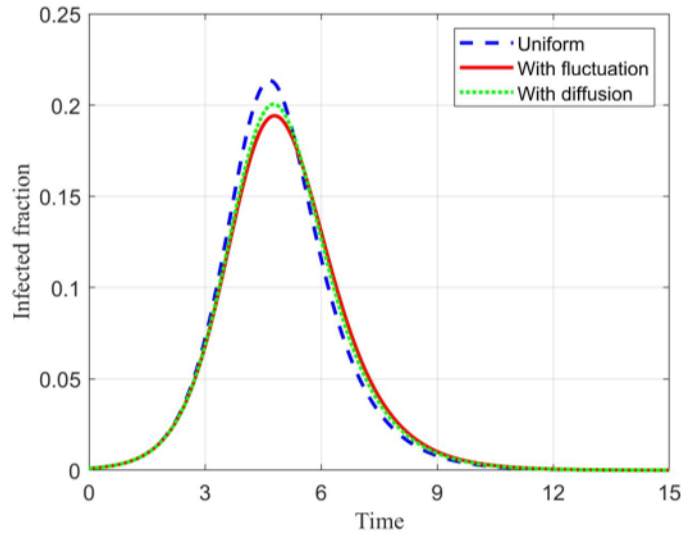

Figure 5: Variation of the ensemble average of the infected (i) fraction as a function of non-dimensional time for $R_0(\rho, 0) = 2.5$. The initial condition is equation (40) with $i_0 = 0.05$, $g(x) = \sin(x)$, and $\epsilon = 0.04$.

3.2.2 Spatial Heterogeneity

Results are quite different when the spatial heterogeneity is larger, however. Here, the interpretation of the composite average must be done with a careful understanding of the underlying heterogeneities.

Consider, first, two regions with the same density, but with different initial conditions that differ by an order of magnitude (e.g. $i_1 = 10^{-4}$ and $i_2 = 10^{-3}$)

$$i(x, 0) = i_1 H(\xi - x) + i_2 H(x - \xi) \tag{48}$$

where $0 < \xi < 1$ and H is the Heaviside step function. From the previous analysis we expect that the region with higher initial infections will respond faster, the other response trailing by a time lag (e.g. see Figure 4 and Equation (34)). Accordingly, the composite behavior will be controlled initially by the region with the larger number of initial infections, but at later times by the second region. Figure 6 shows that this is indeed the case.

More interesting, perhaps, is the case where the area of interest consists again of two regions, e.g. "urban" in the interval $0 < x < \xi$ and "suburban" in the interval $\xi < x < 1$ We are interested in the infection curve, under the conditions of a "commute" between the two regions: During the day, and for a certain time interval of duration λ (expressed as a
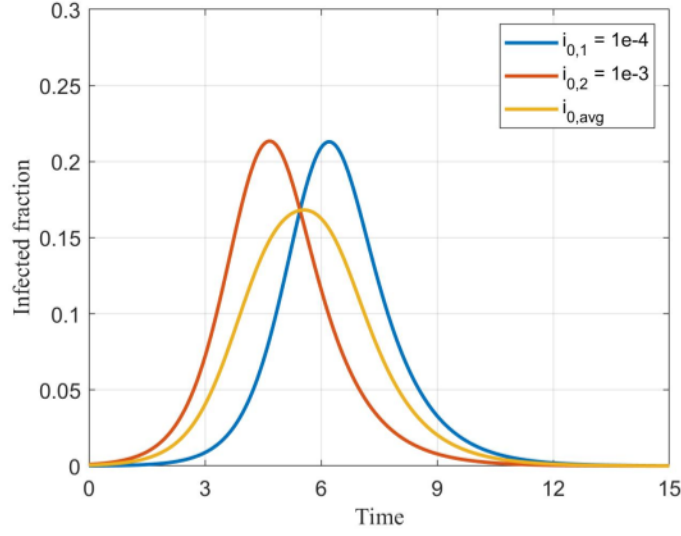
Figure 6: Variation of the infected (i) fraction as a function of non-dimensional time for the two different regions with different initial infections and the composite curve ($R_0(\rho, 0) = 2.5$).

fraction of the 24-hr period), the density in the urban region, ρ_u, is high, and in the suburban region it is negligible; while during the remaining part of the day and at night, the density in the urban region is zero, and in the suburban is ρ_s. These two densities are related by a conservation equation $\rho_s = \rho_u \frac{\xi}{1-\xi}$. This setting is intended to model the commute between two regions ("home" and "work", or "home" and "school") where we expect significantly different R_0 values. By further taking $\xi \ll 1$, we will simply assume $R_0 = 0$ for conditions at "home".

This problem can be solved based on the detailed transport and reaction equations derived earlier, that include advection and diffusion. A simpler alternative, on the other hand, is to represent the problem as two "batch reactors", whose R_0 values oscillate between the two values, $R_{0,w}(\rho, r)$ and 0, when the population is at "work" (or "school") or at "home", respectively. The results of this hypothetical "commute", for an 8-hr work period ($\lambda = 1/3$), are shown in Figure 7. Superimposed are also the results corresponding to the value of the "work" parameter $R_{0,w}(\rho, r)$ applied everywhere (equivalently, by taking $\lambda = 1$). As expected, while commuting leads to a much lower effective value $R_{0,eff}(\rho, 0) = 1.67$, it does not suppress infection, for this example. As long as $R_{0,eff}(\rho, 0) > 1$ infection will occur, although by reducing exposure (decreasing λ) the effect on decreasing the resulting effective rates is non-trivial. Indeed, we have found (not shown here for the sake of brevity) that for all value of λ examined, the corresponding effective is roughly equal to the arithmetic mean, $R_{0,eff}(\rho, 0) \approx \lambda R_{0,w}$. This finding suggests that there is a critical exposure value $\lambda_{crit} = \frac{1}{R_{0,w}}$, below which contagion is suppressed. These findings are new, and must also be interpreted with due caution, and subject to the assumptions made.

3.3 The Effect of Transport

The preceding sections dealt with applications in zero-dimensional space (batch reactors or their ensembles). In problems that involve spatial dimensions, transport (via advection, diffusion or dispersion) plays two different roles: One is stabilizing, reducing the effect of small spatial non-uniformities, as discussed above. The other is in the opposite direction, however, and helps spread the infection spatially. This section will only consider mobility effects due to diffusion.

For a constant diffusion coefficient, in the absence of advection ($Da\boldsymbol{v} = \boldsymbol{0}$), the relevant equations become

$$\frac{\partial s}{\partial t} - C\nabla(\ln \rho) \cdot \nabla s = C\nabla^2 s - R_0(\rho, r)si \tag{49}$$

$$\frac{\partial i}{\partial t} - C\nabla(\ln \rho) \cdot \nabla i = C\nabla^2 i + R_0(\rho, r)si - i \tag{50}$$

$$\frac{\partial r}{\partial t} - C\nabla(\ln \rho) \cdot \nabla r = C\nabla^2 r + i \tag{51}$$
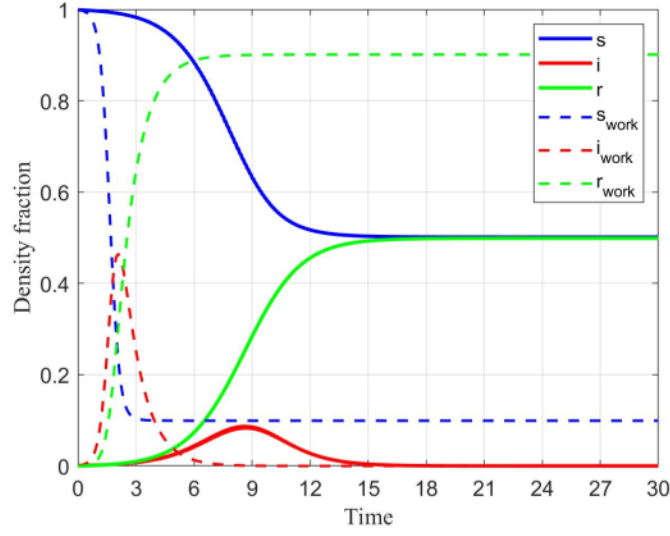
We will first consider problems in one spatial dimension.

Figure 7: The infection curves for the case of a commute between two regions of high and low values of R_0 ($R_{0,w}(\rho,0) = 5$) and exposure frequency of $\lambda = 1/3$ (solid lines). Plotted also are the infection curves corresponding to a uniform $R_{0,w}(\rho,0) = 5$ (equivalently to $\lambda = 1$) (dashed lines). The initial infection fraction is $i(0) = 0.001$.

3.3.1 1-D geometries

Assume 1-D rectilinear geometries, with spatially constant density. The initial infection conditions are non-zero in a specified interval, $-1 < x < 1$, the rest of the domain being free of infections. We are interested in exploring how diffusion leads to the spreading of the contagion from the infected region. Specifically, we will explore if traveling waves develop at large times.

To this effect, we introduce moving coordinate $\xi = x - Vt$, where V is the wave velocity. Assuming that a steady-state is reached in these coordinates (denoted by tilde) we obtain,

$$-V\frac{\partial \tilde{s}}{\partial \xi} = C\frac{\partial^2 \tilde{s}}{\partial \xi^2} - R_0\tilde{s}\tilde{i} \tag{52}$$

$$-V\frac{\partial \tilde{i}}{\partial \xi} = C\frac{\partial^2 \tilde{i}}{\partial \xi^2} + R_0\tilde{s}\tilde{i} - \tilde{i} \tag{53}$$

$$-V\frac{\partial \tilde{r}}{\partial \xi} = C\frac{\partial^2 \tilde{r}}{\partial \xi^2} + \tilde{i} \tag{54}$$

subject to no-flux conditions at the ends

$$\frac{\partial \tilde{s}}{\partial \xi} = \frac{\partial \tilde{i}}{\partial \xi} = \frac{\partial \tilde{r}}{\partial \xi} = 0, \text{ at } \xi = \pm\infty. \tag{55}$$

We first observe that the invariance of (52)-(54) to the transformation $\xi \to -\xi$, $V \to -V$, suggests that there will be two asymptotic waves, one moving to the right, with velocity V, and one to the left, with velocity $-V$. Let ξ_u and ξ_d be two locations sufficiently upstream and and downstream, respectively, of the wave fronts. We then expect

$$\tilde{r}(\xi_u) = r_{V,\infty}, \tilde{i}(\xi_u) = 0, \text{ and } \tilde{r}(\xi_d) = 0, \tilde{i}(\xi_d) = 0 \tag{56}$$

where we have anticipated that $r_{V,\infty}$ might not be identical to the r_∞ of the batch reactor result (equation (31)).

The wave velocity can be determined by integrating (54) between ξ_u and ξ_d,

$$V = \frac{1}{r_{V,\infty}} \int_{\xi_u}^{\xi_d} \tilde{i} d\xi. \tag{57}$$

and, equivalently,

$$V = \frac{1}{r_{V,\infty}} \int_{-\infty}^{\infty} \tilde{i} d\xi. \tag{58}$$

13

We find, therefore, that a wave solution exists, as long as the integral in (58) is non-zero, which is true if there is a contagion. This contagion wave spreads, driven by diffusion, with a velocity that expresses the intensity of the contagion.

3.3.2 Invariance

In the above, we can explicitly remove the C-dependence by introducing rescaled space coordinates and velocities, $\xi = \sqrt{C}\zeta$ and $V = W\sqrt{C}$. The equations remain the same with C set as $C = 1$. Namely, the profiles, now, denoted by superscript hat, satisfy

$$-W\frac{\partial \hat{s}}{\partial \zeta} = \frac{\partial^2 \hat{s}}{\partial \zeta^2} - R_0 \hat{s}\hat{i} \tag{59}$$

$$-W\frac{\partial \hat{i}}{\partial \zeta} = \frac{\partial^2 \hat{i}}{\partial \zeta^2} + R_0 \hat{s}\hat{i} - \hat{i} \tag{60}$$

$$-W\frac{\partial \hat{r}}{\partial \zeta} = \frac{\partial^2 \hat{r}}{\partial \zeta^2} + \hat{i} \tag{61}$$

subject to no-flux conditions at the two ends. The velocity becomes

$$W = \frac{1}{r_{V,\infty}} \int_{\zeta_u}^{\zeta_d} \hat{i}\,d\zeta, \tag{62}$$

in dimensional form,

$$\mathcal{V} = \frac{\sqrt{D\Lambda}}{r_{V,\infty}} \int_{\zeta_u}^{\zeta_d} \hat{i}\,d\zeta. \tag{63}$$

and by extension

$$W = \frac{1}{r_{V,\infty}} \int_{-\infty}^{\infty} \hat{i}\,d\zeta, \tag{64}$$

$$\mathcal{V} = \frac{\sqrt{D\Lambda}}{r_{V,\infty}} \int_{-\infty}^{\infty} \hat{i}\,d\zeta. \tag{65}$$

We note that although the initially infected region $x \in (a,b)$ transforms into $\zeta \in (aC^{-1/2}, bC^{-1/2})$, thus containing a C-dependence, the asymptotic traveling wave is independent of initial conditions, and hence of C. This result is significant: it shows that the limit $C \to 0$ is a singular one, which means that the batch reactor (SIR) problem is also a singular limit. The question then becomes, how different are the resulting curves in the presence of diffusion, and whether or not one can define an equivalent, effective, value of R_0. These are explored below.

Numerical results of the solution of (49)-(51) are shown in Figure 8, for a problem in which the initial infection region is near the left boundary, and $R_0(\rho, 0) = 2.5$. We observe that the infection profile evolves as a function of time, and reaches the asymptotic traveling wave after about $t = 3$. The wavelike nature of the solution is clear if one plots the infected fraction in space-time coordinates (Figure 9), where a ridge with a constant slope is clearly seen.

Consider now, the answer to the question in how different are the infection profiles with diffusion included. Figure 10 shows the relevant wave profiles at a fixed value of x (taken at $x = -30$), calculated using the full system of equations, with diffusion considered. Plotted also are the results of the batch reactor (SIR) problem, for the same value of $R_0(\rho, 0)$. While the results are quite similar, diffusion does affect the shape of the curves obtained, leading to a slightly smaller infection intensity, essentially corresponding to a slightly smaller effective $R_0(\rho, 0)$. Figure 11 is a plot of the asymptotic value r_∞ and of the maximum in infections i_{max} as a function of $R_0(\rho, 0)$ for the respective two cases. We find that diffusion acts to moderate somewhat the contagion intensity.

We conclude that the two solutions, corresponding to the contagion wave (equations (49)-(50)), now denoted by $i_D(t)$. and to the batch (SIR) problem (equations (21)-(22), now denoted by $i_B(t)$), are approximately equal, although not identical

$$i(x,t) \to i(x - Vt) \equiv \hat{i}(\zeta) \equiv i_D(const - \frac{x}{V} + t) \approx i_B(const - \frac{x}{V} + t) \tag{66}$$

Note that the constant in (66) can absorbed in x/V.

The variation of the dimensionless velocity, equation (64) with $R_0(\rho, 0)$ is shown in Figure 12. As expected it increases with $R_0(\rho, 0)$, varying roughly in a linear fashion at relatively large values of the latter. We conclude that diffusion is
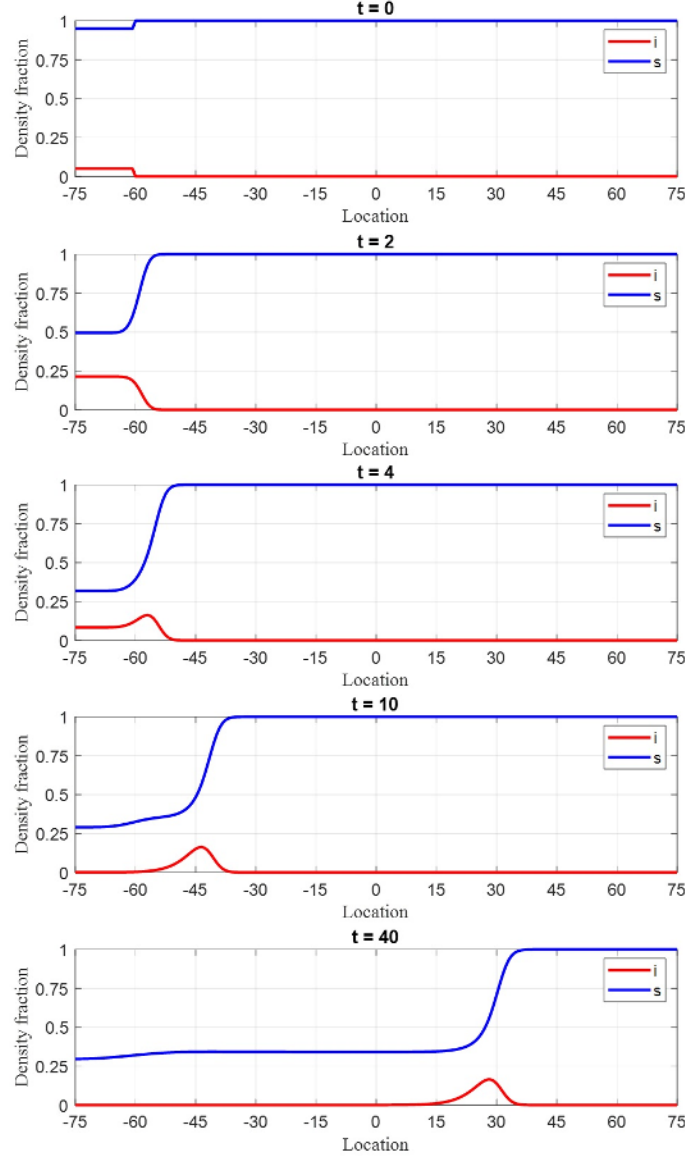
Figure 8: Infected and susceptible density fraction profiles at different values of time ($R_0(\rho, 0) = 2.5$).

fundamental to the spreading of the contagion, the rate of spreading increasing with the square root of the diffusion coefficient and with the value of $R_0(\rho, 0)$. It also affects, although only mildly, the overall infection profiles, by lowering their intensity. Restricting mobility, here represented by D, confirms a most important policy effect for containing infection. The finding that diffusion also affects somewhat the effective value of $R_0(\rho, 0)$, is a result intuitively expected, but not previously identified. And the fact that the limit of small diffusion is a singular one calls for some caution on the interpretation of any batch problem results.

The wavelike spread of infections predicted by the model has been observed in several pandemics. In this context it is interesting to compare the spread of the 1347-1350 black death (pneumonic plague) in Europe and that of the 2009 pandemic influenza in the US. While there is significant debate regarding the effective R_0 value for these pandemics, it is generally accepted that they are in the range of $R_0 \approx 1.5 - 2.5$ [6, 29]. Further, for both diseases the infectious period is around two weeks, and therefore $\Lambda \approx 1/14 \ days^{-1}$. To predict the wave speed, we need an estimate of the respective diffusion coefficients. It is to be expected that in the medieval times, the diffusion of humans, therefore of infection, was much smaller. In [6, 30] this was determined by how quickly information was disseminated (in that case by word of mouth), resulting into a diffusivity estimate of about $27 miles^2/day$. Using the obtained estimate in
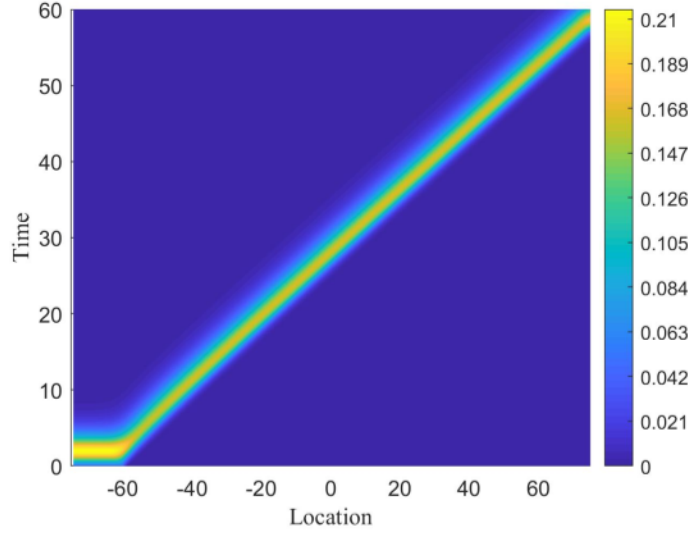
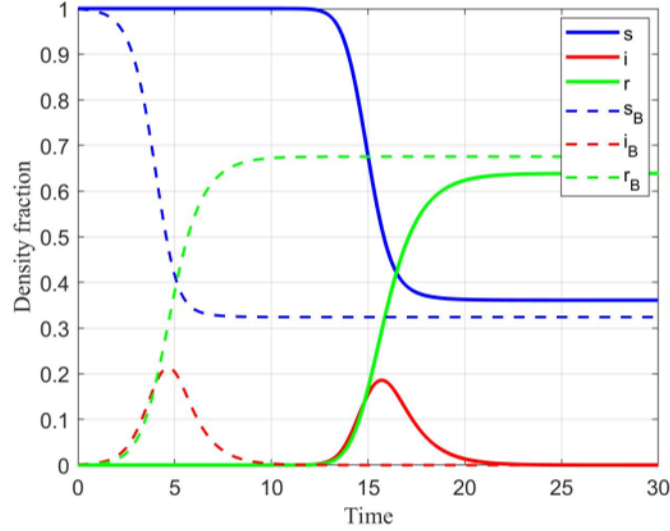Figure 9: Infected density fraction as a function of space and time coordinates ($R_0(\rho, 0) = 2.5$).

Figure 10: Variation of susceptible (s), infected (i) and recovered (r) fractions (steady state) as a function of non-dimensional time with diffusion included (solid lines) and for the batch reactor (SIR) problem (dashed lines) for the same value of $R_0(\rho, 0) = 2.5$. The initial infection fraction for the batch system is $i_B(0) = 0.001$

our estimate for the speed of the contagion (see Figure 12) we arrive at a speed of $2.7 \; miles/day$ for the black death pandemic, which is in the ball-park of the actual observed speed of around $1 \; mile/day$ [31].

For the 2009 influenza pandemic, we can make a similar back-of-the-envelope calculation. Since children play an important role in the spread of influenza, we derive this estimate based on their activities. We assume that a typical child will see 30 other children in school and playgrounds during a typical day. Further, they will travel an average of 10 miles to get to these places. In a coarse analogy to an ideal gas this gives a mean-free path of 10 miles and a frequency of collision of $30 \; days^{-1}$, and therefore a diffusivity of around $3,000 miles^2/day$. Using these parameters in our estimate of the speed of the contagion wave, we arrive at a speed of $29 \; miles/day$ for the 2009 influenza pandemic. This is also in the ball-park of the observed wave speed of around $23 \; miles/day$ for the 2009 influenza pandemic [20]. The higher diffusion coefficient for the 2009 influenza pandemic leads to a significantly faster wave speed.
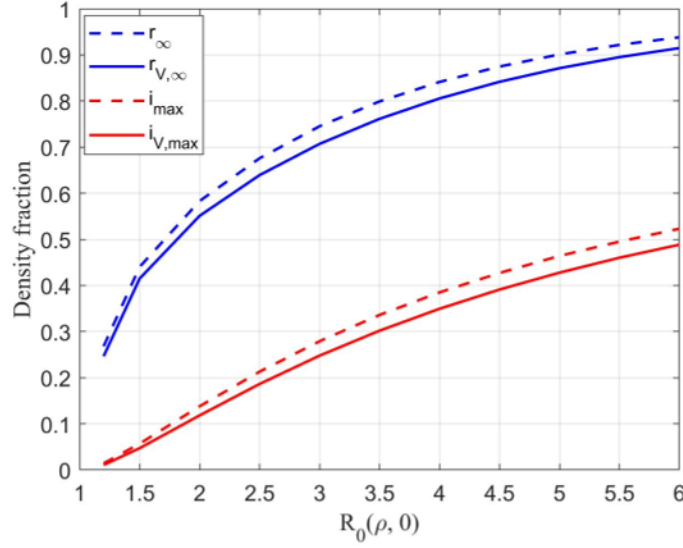
Figure 11: Variation of herd immunity (r_∞) and maximum infection fraction (i_{max}) as a function of $R_0(\rho, 0)$ for the two cases of a batch system (dashed lines) and when diffusion is included (solid lines). The initial infection fraction for the batch system is $i_B(0) = 0.001$.
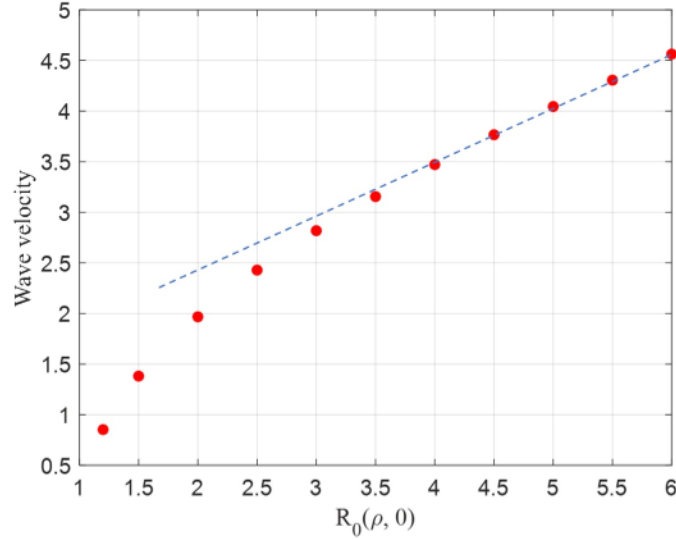

Figure 12: Variation of non-dimensional wave speed in a one-dimensional problem as a function of $R_0(\rho, 0)$.

For completeness, we simulated the "collision" of two waves, one moving from the left and the other from the right. Figure 13 shows how the two waves amplify as they interact, then ultimately decay as infection has spread fully and the population reaches conditions of herd immunity at the corresponding value of R_0. This wave interaction is different, for a number of reasons, from what is observed in other wave problems (e.g. solitary waves, and where for example, the wave velocities increase with the wave amplitude [32]). In the present context, the wave amplitude, e.g. i_{max} cannot exceed the value of 1.

We conclude this section by considering effects of heterogeneity in this 1-D rectilinear geometry. Figures 14 and 15 show results when an infection wave enters a region with a lower value of $R_0(\rho, 0)$, e.g. one of lower spatial density, from a region of a higher value of $R_0(\rho, 0)$, e.g. one of higher spatial density. For example, such could be the case of contagion spreading from an urban to a rural area. In the figure this occurs at $x = -25$. As it enters the low-density region the wave decelerates, with the magnitude of infected fraction decreasing. The slowing of the wave is indicated
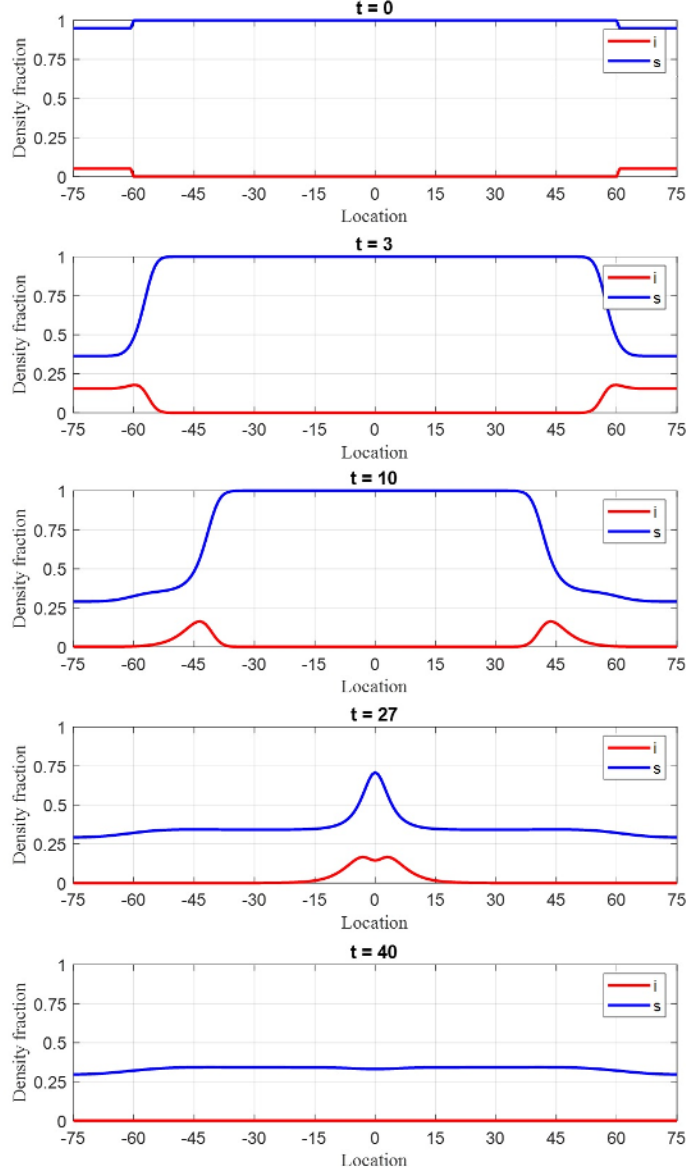
Figure 13: Infected and susceptible density fraction profiles at different values of time for two waves propagating towards the center of the domain ($R_0(\rho, 0) = 2.5$).

by the increase of the slope in the $x - t$ domain. Conversely, at $x = 25$, when the wave now enters a region of higher $R_0(\rho, 0)$, it accelerates to a larger velocity, the intensity of the infected fraction correspondingly increasing.

3.3.3 Radial Geometries

The results of the 1-D rectilinear geometry also apply almost identically to radial geometries. Now, equations (49)-(50) become

$$\frac{\partial s}{\partial t} = \frac{C}{\mu}\frac{\partial}{\partial \mu}(\mu\frac{\partial s}{\partial \mu}) - R_0(\rho, r)si \tag{67}$$

$$\frac{\partial i}{\partial t} = \frac{C}{\mu}\frac{\partial}{\partial \mu}(\mu\frac{\partial i}{\partial \mu}) + R_0(\rho, r)si - i \tag{68}$$

$$\frac{\partial r}{\partial t} = \frac{C}{\mu}\frac{\partial}{\partial \mu}(\mu\frac{\partial r}{\partial \mu}) + i \tag{69}$$
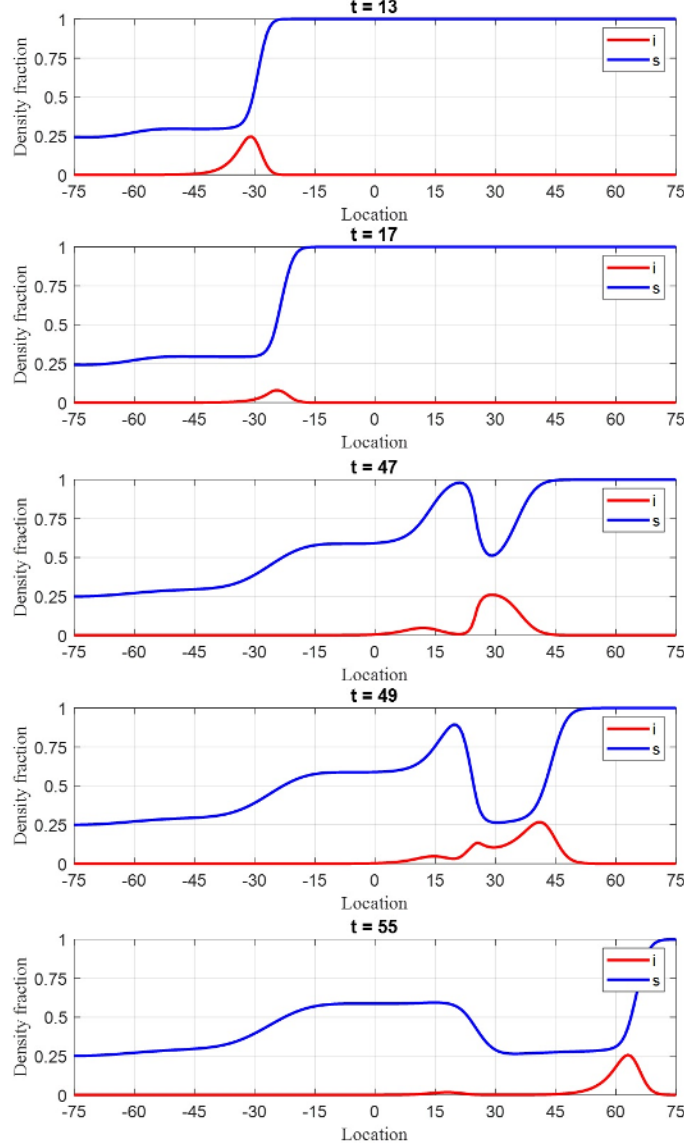
18

Figure 14: Infected and susceptible density fraction profiles at different values of time in a 1-D heterogeneous system. For $x \in (-80, -25)$ and $x \in (25, 80)$, $R_0(\rho, 0) = 3$, while for $x \in (-25, 25)$, $R_0(\rho, 0) = 1.5$.

where we denoted the radial coordinate with μ. We look for the solution of the problem when a region around the origin ($\mu = 0$) is initially infected. As in the rectilinear geometry case, we expect that the solution will evolve in terms of a traveling wave, therefore, we consider a transformation to the moving coordinates $\xi = \mu - Vt$, and $t' = t$. In these coordinates, the equations become

$$\frac{\partial s}{\partial t'} - V\frac{\partial s}{\partial \xi} = C\frac{\partial^2 s}{\partial \xi^2} + \frac{C}{\xi + Vt'}\frac{\partial s}{\partial \xi} - R_0(\rho, r)si \tag{70}$$

$$\frac{\partial i}{\partial t'} - V\frac{\partial i}{\partial \xi} = C\frac{\partial^2 i}{\partial \xi^2} + \frac{C}{\xi + Vt'}\frac{\partial i}{\partial \xi} + R_0(\rho, r)si - i \tag{71}$$

$$\frac{\partial r}{\partial t'} - V\frac{\partial r}{\partial \xi} = C\frac{\partial^2 r}{\partial \xi^2} + \frac{C}{\xi + Vt'}\frac{\partial r}{\partial \xi} + i \tag{72}$$

We note that in the limit of large t', equations (70)-(71) transform to the same equations as (52)-(53), with the first term on the left and the second term on the right of the equations vanishing. Therefore, the same results hold for
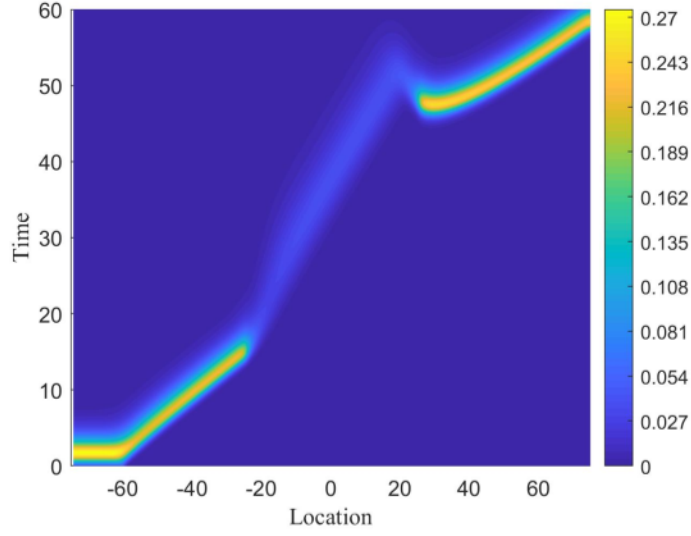
19

Figure 15: Infected density fraction as a function of space and time coordinates in a 1-D heterogeneous system: For $x \in (-80, -25)$ and $x \in (25, 80)$, $R_0(\rho, 0) = 3$, while for $x \in (-25, 25)$, $R_0(\rho, 0) = 1.5$.

radial geometries as those for the rectilinear problem. The simulations for the 1-D radial geometry, shown in Figure 16, confirm the theoretical findings. Interestingly, diffusion in the above reaction systems does not enter in the more
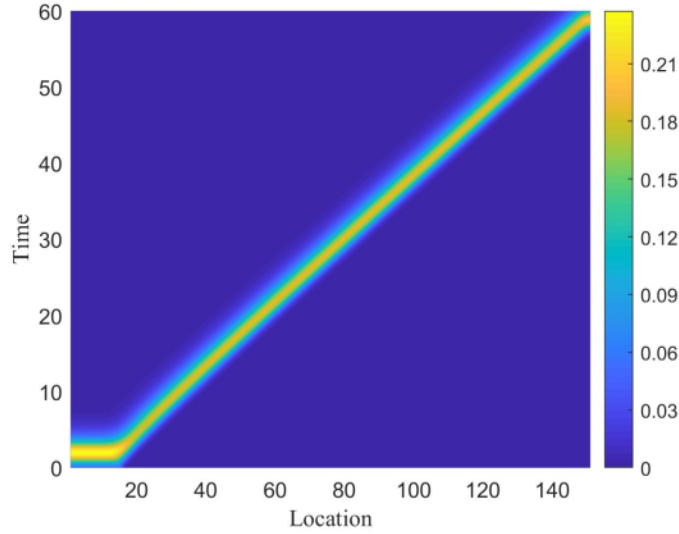

Figure 16: Infected density fraction as a function of radial (space) and time coordinates ($R_0(\rho, 0) = 2.5$).

familiar form of a similarity variable $\eta = \frac{\mu}{\sqrt{t}}$, but rather in terms of a wave that propagates with a constant linear speed.

3.3.4 Effect of Heterogeneity in Two Dimensions

Consider, now infection propagation in a general, heterogeneous 2-D system. We are interested in understanding how propagation occurs and whether or not one can use the asymptotic wave solutions obtained for the 1-D geometries. In particular, we are interested in knowing if one can use a wave equation to describe the evolution of the contagion fronts. To explore this question we consider three different geometries, two corresponding to a layered (stratified) system, and one corresponding to a set of four quadrants.

Layered System Consider propagation in a stratified layered system, the middle layer having a larger density value (thus, a larger $R_0(\rho, 0)$), compared to its two adjacent layers. Infection is initiated uniformly on the left boundary (see Figure 17). As expected, an infection wave emerges in the middle layer, traveling with a velocity v_1 corresponding
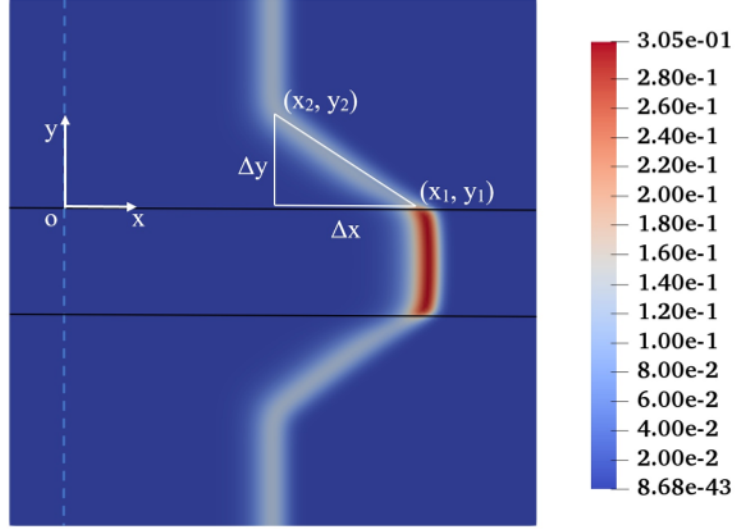

Figure 17: Infected density fraction (i) at $t = 40$ in a layered medium with an infection wave starting at $x = 0$. In the outer layers $R_0(\rho, 0) = 2$ while in the inner layer $R_0(\rho, 0) = 3.5$. The spatial dimension of the domain is 150×150.

to the high R_0 value in that layer, and equal to the corresponding asymptotic 1-D wave velocity of 3.15 (see Figure 12) for $R_0(\rho, 0) = 3.5$. A slower moving contagion front develops in the adjacent layers, with corresponding velocity v_2 equal to 1.96 which is the asymptotic value of the 1-D wave velocity corresponding to $R_0(\rho, 0) = 2$. Because of diffusion, however, transverse waves also emerge, emanating at the interface of the layers, propagating in the transverse (y) direction. These waves manifest as a straight line, slanted with respect to the transverse direction, but moving forward with speed v_1.

If we set the origin of the coordinates at the intersection of the interface between the two layers and the initial infection boundary), the straight line connecting the two fronts in the two layers is described by the following equation

$$F \equiv y(v_1 - v_2) + xa - v_1 at = 0 \tag{73}$$

where the leading edge travels with the middle layer velocity v_1, the trailing edge with the adjacent layer velocity v_2, and the tip of the connecting layer at $y = y_2(t)$ with velocity a, to be determined. Because the waves emanating from the interface travel in the adjacent layers at fixed x with transverse velocity v_2, equation (73) gives

$$v_2(v_1 - v_2) = v_1 a \tag{74}$$

therefore we find

$$a = \frac{v_2}{v_1}(v_1 - v_2) \tag{75}$$

The slope of the straight line must then be

$$slope = \frac{a}{(v_1 - v_2)} = \frac{v_2}{v_1} \tag{76}$$

The simulations of (17) show that this is indeed the case.

More generally, if we define a front by $F(\boldsymbol{x}, t) = 0$ where $\boldsymbol{x}$ is the space vector, its evolution satisfies

$$F_t(\boldsymbol{x}, t) + \boldsymbol{v} \cdot \nabla F = 0 \tag{77}$$
$$F_t(\boldsymbol{x}, t) + v_n|\nabla F| = 0 \tag{78}$$

where subscript t denotes time derivative, and v_n is its component in the direction of ∇F. For rectilinear or radial geometries, as well as in each of the two layers discussed, (78) reduces to a linear wave propagating at a constant

velocity. However, for the straight line connecting the two waves, resulting from a sequence of waves emanating from the layer interface, the relevant equation is (73), which leads to a normal velocity equal to

$$v_n = \frac{v_2}{\sqrt{(1 + \frac{v_2^2}{v_1^2})}} \tag{79}$$

The interdependence between the three velocities (leading front, v_1, normal to the surface, v, and trailing front, v_2) is to be noted.

Four heterogeneous quadrants   A different illustration of these effects is shown in Figure 18, where infection initiates in the upper right corner of a rectangular domain with four different quadrants, each taken with a different density (hence different values of $R_0(\rho, 0)$). We have selected the two diagonally opposite quadrants (e.g. NE and SW) to have relatively larger values of $R_0(\rho, 0) = 3.5$, hence an asymptotic wave velocity of 3.15, while the other two (SE, NW) to be of relatively smaller values of $R_0(\rho, 0) = 2$, hence an asymptotic wave velocity of about 2. The infection waves follow the expected pattern. Infection grows first radially in the NE quadrant with the asymptotic speed of 3.15 (upper right panel in Figure 18). When the SE and NW boundaries are encountered (lower right panel in Figure 18), the waves slow down and start spreading in a radial manner as they enter the two quadrants at the slower speed of 2. Subsequently, the infection wave enters the high density SW region, in which it starts spreading radially, now moving with the higher velocity of 3.15 (lower left panel in Figure 18). Upon touching the boundaries with the NW and SE regions, however, diffusion causes infection waves to start emanating from the higher to the lower density regions, resulting into a straight line segment, similar to that for the layered system, connecting the two different waves in the two different regions. For the same reasoning as in the layered system, this straight line segment has a slope equals to the ratio of the two respective velocities, namely (76). The simulations confirm these results.
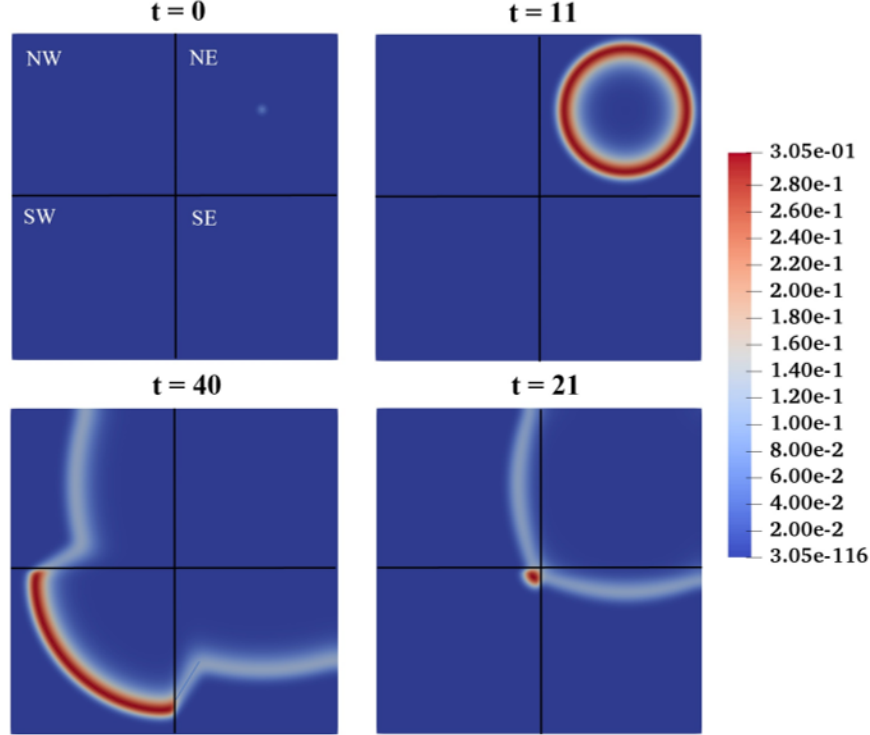

Figure 18: Contour plot of the infected density fraction a region with four quadrants with different values of $R_0(\rho, 0)$ at different values of time. $R_0(\rho, 0) = 3.5$ for NE and SW, and $R_0(\rho, 0) = 2$ for NW and SE, respectively. The spatial dimension of the region is 150×150.

We end this section by noting that one can further explore the use of a wave equation, or perhaps the eikonal equation, for the description of the spreading of contagion. This would be the goal of a separate study.

4 Conclusions

We used an analogy with chemical reaction engineering processes to model the growth and spreading of human-transmitted infections, such as COVID-19 and influenza. The basis of the model is the assumption of three distinct populations, as in the celebrated SIR model, which are mapped into equivalent chemical species. An important first result from this formulation is that the relevant quantities are population densities (specifically, areal spatial densities) and not total populations, as is in the SIR model. These satisfy partial differential equations, with spatial mobilities included through diffusion and advection. Spatial density effects are then incorporated into the kinetic constant of the infection growth, hence into the key parameter R_0, which becomes an explicitly function of spatial densities, and found to decrease with the extent of the contagion. Incorporating a spatial density dependence in contagion kinetics is necessary, and consistent with health guidelines and droplet dynamics.

In the absence of diffusion, and assuming spatially uniform profiles, the results revert to a modified version of the SIR model, in a geometry recognized as a "batch reactor". We find that if were to identify an effective SIR-like R_0, it would be smaller than when spatial effects are ignored, which is the standard SIR case. The infection curves are found to be largely independent of initial conditions, the effect of which is simply to delay the onset of the infection. Analogous results hold when a ban on "imported infection" is applied, which also only affects to delay the onset of the epidemic, assuming that R_0 is not modified. Small fluctuations in the initial conditions have a minimal effect on the ensemble behavior. However, this is not the case when the fluctuations are significant in space or in time, in which case they result into non-trivial ensemble averages. Using the density-dependent R_0 we can readily model the effect of the equivalent of a commute between "home" and "work", or "suburban" and "urban", to obtain an effective R_0, which is found to be equal to the arithmetic mean weighted by the exposure time.

We subsequently considered the effect of diffusion in 1-D geometries. We showed the emergence of infection waves, whether in rectilinear or radial geometries, which asymptotically travel with constant velocity, the dimensional value of which scales with the square root of diffusivity, and increases with R_0. The behavior of the infection waves at a fixed time is very similar to that for an SIR system, although the shape and the effective value of R_0 are slightly different than in the absence of diffusion, leading to lower intensities, than in the absence of diffusion. This result is important, as it shows that the SIR limit (absence of diffusion) is a singular limit. We then examined how distributed densities in different geometries affect the propagation of infection waves. We find that for all practical purposes, the contagion propagation can be modeled by equivalent waves, the speed of which only depends on the value of R_0 in that region. The overall conclusion is that spatial effects via density or diffusion considerations are important and must be considered in the description of contagion and its epidemics.

References

[1] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[2] Fabrizio Croccolo and H Eduardo Roman. Spreading of infections on random graphs: A percolation-type model for covid-19. *Chaos, Solitons & Fractals*, page 110077, 2020.

[3] Tiberiu Harko, Francisco SN Lobo, and MK Mak. Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death and birth rates. *Applied Mathematics and Computation*, 236:184–194, 2014.

[4] Roy M Anderson and Robert M May. Population biology of infectious diseases: Part i. *Nature*, 280(5721):361–367, 1979.

[5] Keith Burghardt and Kristina Lerman. Unequal impact and spatial aggregation distort covid-19 growth rates. *arXiv preprint arXiv:2004.12994*, 2020.

[6] Julia V Noble. Geographic and temporal development of plagues. *Nature*, 250(5469):726–729, 1974.

[7] Sheldon Ross. *A First Course in Probability 8th Edition*. Pearson, 2009.

[8] Felix Sharipov and José L Strapasson. Ab initio simulation of transport phenomena in rarefied gases. *Physical Review E*, 86(3):031130, 2012.

[9] H Scott Fogler. *Essentials of Chemical Reaction Engineering: Essenti Chemica Reactio Engi*. Pearson Education, 2010.

[10] R Byron Bird. We stewart, and en lightfoot. *Transport phenomena*, 11:5, 1960.

[11] Hydrodynamics Lamb and Sixth Ed Hydrodynamics. Cambridge university press, 1932. *et seq*, page 254, 1966.

[12] Ronald R Coifman and Stéphane Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006.

[13] Franca Hoffmann, Bamdad Hosseini, Assad A Oberai, and Andrew M Stuart. Spectral analysis of weighted laplacians arising in data clustering. *arXiv preprint arXiv:1909.06389*, 2019.

[14] Benoit B Mandelbrot. The fractal geometry of. *Nature*, pages 394–397, 1982.

[15] AS Chaves. A fractional diffusion equation to describe lévy flights. *Physics Letters A*, 239(1-2):13–16, 1998.

[16] Diego del Castillo-Negrete, BA Carreras, and VE Lynch. Front dynamics in reaction-diffusion systems with levy flights: a fractional diffusion approach. *Physical Review Letters*, 91(1):018302, 2003.

[17] Péter Érdi and János Tóth. *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*. Manchester University Press, 1989.

[18] Qun Liu and Daqing Jiang. Dynamical behavior of a stochastic multigroup sir epidemic model. *Physica A: Statistical Mechanics and its Applications*, 526:120975, 2019.

[19] JA Wesselingh, Rajamani Krishna, et al. *Mass transfer in multicomponent mixtures*. Delft University Press Delft, 2000.

[20] Julia R Gog, Sébastien Ballesteros, Cécile Viboud, Lone Simonsen, Ottar N Bjornstad, Jeffrey Shaman, Dennis L Chao, Farid Khan, and Bryan T Grenfell. Spatial transmission of 2009 pandemic influenza in the us. *PLoS Comput Biol*, 10(6):e1003635, 2014.

[21] Alex Viguerie, Guillermo Lorenzo, Ferdinando Auricchio, Davide Baroli, Thomas JR Hughes, Alessia Patton, Alessandro Reali, Thomas E Yankeelov, and Alessandro Veneziani. Simulating the spread of covid-19 via spatially-resolved susceptible-exposed-infected-recovered-deceased (seird) model with heterogeneous diffusion. *arXiv preprint arXiv:2005.05320*, 2020.

[22] Joseph O Hirschfelder, Charles F Curtiss, Robert Byron Bird, and Maria Goeppert Mayer. *Molecular theory of gases and liquids*, volume 165. Wiley New York, 1964.

[23] William F Wells. On air-borne infection: Study ii. droplets and droplet nuclei. *American journal of Epidemiology*, 20(3):611–618, 1934.

[24] Lydia Bourouiba, Eline Dehandschoewercker, and John WM Bush. Violent expiratory events: on coughing and sneezing. *Journal of Fluid Mechanics*, 745:537–563, 2014.

[25] X Xie, Y Li, ATY Chwang, PL Ho, and WH Seto. How far droplets can move in indoor environments–revisiting the wells evaporation–falling curve. *Indoor air*, 17(3):211–225, 2007.

[26] Vasilis Marmarelis. Predictive modeling of covid-19 data in the us: Adaptive phase-space approach. *IEEE Open Journal of Engineering in Medicine and Biology*, 2020.

[27] Athanasios S Fokas, Nikolaos Dikaios, and George A Kastis. Covid-19: Predictive mathematical models for the number of deaths in south korea, italy, spain, france, uk, germany, and usa. *medRxiv*, 2020.

[28] Athanasios S Fokas, Jesús Cuevas-Maraver, and Panayotis G Kevrekidis. Two alternative scenarios for easing covid-19 lockdown measures: one reasonable and one catastrophic. *medRxiv*, 2020.

[29] Roya Nikbakht, Mohammad Reza Baneshi, Abbas Bahrampour, and Abolfazl Hosseinnataj. Comparison of methods to estimate basic reproduction number (r0) of influenza, using canada 2009 and 2017-18 a (h1n1) data. *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences*, 24, 2019.

[30] Nicolas Rashevsky. *Looking at history through mathematics*. Cambridge, Mass: MIT Press, 1968.

[31] William L Langer. The black death. *Scientific American*, 210(2):114–121, 1964.

[32] Philip G Drazin and Robin S Johnson. *Solitons: an introduction*, volume 2. Cambridge university press, 1989.