Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Chemical Engineering Science 233 (2021) 116347

Contents lists available at ScienceDirect

Chemical Engineering Science

journal homepage: www.elsevier.com/locate/ces

A comprehensive spatial-temporal infection model

Harisankar Ramaswamy [a], Assad A. Oberai [a], Yannis C. Yortsos [b,*]

[a] Aerospace and Mechanical Engineering, Viterbi School of Engineering, University of Southern California, United States
[b] Mork Family Department of Chemical Engineering and Materials Science, Viterbi School of Engineering, University of Southern California, United States

HIGHLIGHTS

- Modeling infections as chemical processes captures their spatio-temporal spreading.
- Inclusion of areal densities in the model that lead to new conclusions regarding the effect of initial infections and heterogeneous environments.
- Inclusion of diffusion leads to propagating fronts of infection that display select wave-like behavior.

ARTICLE INFO

Article history:
Received 1 September 2020
Received in revised form 26 November 2020
Accepted 2 December 2020
Available online 6 January 2021

Keywords:
Infection models
Chemical processes
Wave propagation
COVID-19

ABSTRACT

Motivated by analogies between the spread of infections and of chemical processes, we develop a model that accounts for infection and transport where infected populations correspond to chemical species. Areal densities emerge as the key variables, thus capturing the effect of spatial density. 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. We present results for a batch reactor, the chemical process equivalent of the SIR model, where we examine how the dependence of R_0 on process extent, the initial density of infected individuals, and fluctuations in population densities effect the progression of the disease. We then consider spatially distributed systems. Diffusion generates traveling waves that propagate at a constant speed, proportional to the square root of the diffusivity and R_0. Preliminary analysis shows a similar behavior for the effect of stochastic advection.

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 epidemiological models have been proposed, developed and tested, e.g. see (Kermack and McKendrick, 1927; Croccolo and Roman, 2020; Harko et al., 2014) among many others. 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 convert into another.

Widely used is the celebrated SIR model (Kermack and McKendrick, 1927; Anderson and May, 1979), 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 limitations, as they assume sufficiently large diffusion or mobility.

Another category of models is based on computing individual trajectories of typical members of the various populations and on subsequently postulating probabilities of collisions and infection e.g. (Burghardt and Lerman, 2020) and many others. 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. Importantly, 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 correspond-

* Corresponding author.
E-mail addresses: hramaswa@usc.edu (H. Ramaswamy), aoberai@usc.edu (A.A. Oberai), yortsos@usc.edu (Y.C. Yortsos).

https://doi.org/10.1016/j.ces.2020.116347
0009-2509/

ing transport and reaction models. This continuum 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 partial differential equations description. It is the objective of this paper to create such a methodology to model the spread of an epidemic in general, with application to COVID-19, in particular. Our work parallels related previous continuum models, such as (Noble, 1974) or particularly the very recent work of Viguerie et al. (2020), but differs in a number of aspects. For example, the recent work of Viguerie et al. (2020) uses a compartmental modeling approach using continuum mechanics, that allows one to obtain results including diffusion and reaction similar to the work here. An important difference is that the view of contagion in the form of chemical reaction processes allows non-trivial additional insights, including the ab initio representation of the kinetics of the infection, as well as the effect of mobility through advection.

As in any continuum model, the fundamental underlying assumption is that one can homogenize population distributions by defining continuum variables, expressed as continuous functions of position and time. This allows postulating continuum conservation laws in terms of their rate of change, transport and reaction. We recognize limitations in the analogy between transport and interaction of human populations on the one hand and molecular transport and reaction on the other. Indeed, given the relatively sparse areal densities of human populations, compared to molecular densities, the law of large numbers to obtain well-defined spatial averages (homogenization) (Ross, 2009; Bensoussan et al., 2011) may in fact not be in effect. In such cases, the equivalent distributions might be more akin to rarefied gas dynamics (Sharipov and Strapasson, 2012). Further, human movement is influenced by behavioral drivers, rather than random walks, and it is likely that a diffusion or dispersion process for human population mobility would require a description different than Brownian motion. These alternatives will be pointed out where appropriate but will not be further pursued. Indeed, we believe that the analogies made here are useful, novel and instructive, and allow important new insights on fundamental spatial aspects of the problem.

The commonly used SIR model arises from the present formulation in the limit of "perfect mixing", namely the absence of spatial gradients, over specific control areas (e.g. a building, a manufacturing plant, a school, a city, a state, a country, etc.) and assume perfect mixing, induced by fast diffusion or advection. The chemical engineering analogue is the "batch reactor" model (Fogler, 2010). We will examine its validity, as infection rates, depending crucially on spatial (areal) density, lead to spatial non-uniformities, while finite transport leads to infection waves. Both impact the stationary assumptions in the SIR models. First, we find that the key variable R_0 does not actually remain constant during the process, contrary to the commonly made 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 then consider the solution of spatially-dependent problems by including diffusion in both 1-D (rectilinear) and in 2-D (including radial symmetry) geometries. We discuss the asymptotic emergence of traveling infection waves, the speed and shape of which are found to be independent of geometry, but to also depend on the relative importance of transport, with an intensity that is lower, hence corresponding to a smaller value of R_0, when diffusion is weaker. Advection is discussed in the context of its most important effect, namely in a heterogeneous, stochastic environment. As in related problems (Sreenivasan, 2019), it leads to an effective macrodispersion, which typically dominates over diffusion and leads to enhanced mixing of populations, hence an increase in the intensity of the contagion.

The paper is organized as follows: We first proceed with constructing the analogous 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, which range from a "batch reactor" model (the SIR equivalent), to the propagation of infection in both 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 to be largely 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 already implicitly assumed in many SIR-type models.

2. Formulation

Consider the conservation of mass of the three areal quantities of interest, namely, susceptible, infected, and recovered (including perished), and associate to them three equivalent "chemical species". In writing the corresponding mass balances, we associate species mobility through both advection and diffusion mechanisms, and infection and recovery rates through equivalent chemical reactions. Mass conservation equations, then, read as follows

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

where ρ_i is the population density (number/area) for species i, T is dimensional time, spatial coordinates are in dimensional notation, the spatial coordinates are only in two dimensions, $\boldsymbol{q}_i$ is the advective velocity vector of species i, $\boldsymbol{D}_i$ is the diffusive (or dispersive) flux, and r_i is the net reaction rate of species i, that convert populations into one another due to infection and recovery. The overall species balance dictates $r_S + r_I + r_R = 0$.

In (1) the area is defined as the surface over which populations reside, work or interact (e.g. work floor area, geographical urban area, etc.). The advective velocity $\boldsymbol{q}_i$ denotes an advective transport mobility term. While we allow in (1) the three velocities to be different, in the remainder we will only take $\boldsymbol{q}_i = \boldsymbol{q}$ for all populations. This common velocity is assumed independent of ρ_i or its gradients, although it can be a function of space and/or time. Next, define

$$\rho_S + \rho_I + \rho_R = \rho, \tag{2}$$

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 (Bird et al., 1961), one may take

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

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, Eq. (5) yields $\boldsymbol{D} = -D\nabla\rho$ and (4) becomes

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

This suggests 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 notion that the overall density diffuses is contrary to the common continuity equation for physical phenomena in fluids, e.g. incompressible fluids (Lamb, 1993). This contradiction is removed if one takes the alternative approach and defines instead

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

Then, 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 (Coifman and Lafon, 2006; Hoffmann et al., 2019). While either equations (5 and 6) or (7 and 8) can be taken to describe diffusive transport, in the examples to be discussed below we will only assume the more familiar continuity equation versions of equations (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, since the equations now become decoupled. It is important to note that while the overall density in (8) does not diffuse, the individual species (populations) continue to do so via (7).

As remarked above 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) (Mandelbrot, 1982). Modeling such motions can be handled by an integro-differential equation description through fractional derivatives (Chaves, 1998; del Castillo-Negrete et al., 2003). While worth considering, such an approach will not be pursued here. We will also not explore additional, possibly interesting, cases, e.g. of non-linear diffusion, in which the diffusion coefficients are functions of the individual densities, e.g. see (Viguerie et al., 2020). In crowded environments of high density, diffusion will certainly be affected by the total density, ρ, although it will likely be independent of the *individual* density fractions. In problems of constant overall density, therefore, those types of non-linear effects will be absent. Conversely, one must consider cases where the advection velocity is a stochastic variable. The resulting fluctuations will be expected to lead to enhanced macrodispersion, e.g. as in turbulence or in stochastic transport in porous media (Sreenivasan, 2019; Gelhar and Axness, 1983), dominating over diffusion, thus contributing to enhanced spreading of the contagion.

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 10 m, over an $8-$ hr period, one obtains $D \approx 10^{-2}$ m^2/s. By comparison, molecular diffusion coefficients in gases are about three orders of magnitude smaller. Dispersion at much larger scales can be obtained 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. We show that values of the order of $D \approx 10^3$ m^2/s or higher are needed in order to match the observed propagation velocities. Clearly, such large values do not only reflect diffusion, but rather the aggregate mix of transport activities (including macrodispersion through stochastic advection).

2.2. Reaction rates

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

$$S + I \rightarrow 2I \tag{9}$$

$$I \rightarrow R \tag{10}$$

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 (Érdi and Tóth, 1989) in the reactions (9) and (10) provides expressions for the reaction rates in terms of the respective concentrations or densities, namely,

$$r_S = -K\rho_S\rho_I, \qquad r_I = K\rho_S\rho_I - \Lambda\rho_I, \qquad r_R = \Lambda\rho_I, \tag{11}$$

where we introduced the reaction rate constants K and Λ. Eq. (11) are the same as those for the SIR model, except that here the rates are correctly expressed here in terms of areal densities, rather than in terms of the total populations, as assumed in typical models. This has significant implications on the kinetic parameters, as discussed below. The two kinetic constants have different dimensions, Λ expressed in inverse time, and K in inverse (time $\times$ number/(length)^2).

We hasten to note that in reality a more fine-grained model would be applicable, the reaction rates also depending on demographics and/or health conditions of the individual species. Such fine-graining is possible with the present framework, by further partitioning the populations into subgroups (Liu and Jiang, 2019). The corresponding description would then be cast in terms of an equivalent "multicomponent mixture" (Wesselingh et al., 2000), where Eqs. (1) are expressed in terms of an extended species vector, with reactions (9) and (10) extended appropriately, possibly involving a product of a reaction matrix with the species vector. A related multicomponent description, although based on a different continuum model using applied mechanics methodologies, was presented in Viguerie et al. (2020). 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} (Noble, 1974; Gog et al., 2014; Viguerie et al., 2020). This kinetic parameter expresses the rate at which infected individuals recover (or perish), on average, and it 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, 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. Most controllable among these factors are the frequency and degree 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 being inversely proportional to density (Hirschfelder et al., 1964). At least in some domain of the values of ρ, K should be increasing with spatial density, a feature ignored in previous SIR models. Directly applying Maxwell–Boltzmann-type kinetic theories is unrealistic in the present context however, and must be further 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 accurately captured. Indeed, it is by now widely accepted that infection rates are negligible for densities below a limiting value (corresponding to a separation of 2 m or 6*ft*, as also supported by fluid mechanical models of droplet propagation, and recommended in health policy guidelines (Wells, 1934; Bourouiba et al., 2014; Xie et al., 2007)).

We incorporate all these aspects by postulating the following dependence

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

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

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

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

When the rate of infection is relatively low (and $r \ll 1$), the above correction is negligible. For strong infection rates, on the other hand, it can be quite significant, as also shown 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 can still be applicable: For example, assume that the infected species ρ_I is further subdivided into one category containing asymptomatic individuals (denoted by A, with corresponding density ρ_A) and another containing quarantined individuals (denoted by Q, with corresponding density ρ_Q). The associated infection reaction rate will now read $K\rho_S\rho_A$. 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 $Ka\rho_S\rho_I$. This is of the same dependence as before, hence the previous holds, subject to modifying the kinetic constant with the 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 does not 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 such geographic areas (states, countries) never on average reach the minimum density required for infection (e.g. $0.1\ \mathrm{m}^{-2}$), area-wide averages with substantially heterogeneous density distributions are not very informative. Rather, distinguishing high-density areas (e.g. urban places, stadiums, schools, retirement homes, etc.), or events attracting high densities, from low-density ones (e.g. farms, rural) is essential. Such reporting of more fine-grained area statistics is much more informative. Connecting the transmission of infection between areas of different density is feasible with the present formalism, which includes spatial transport and/or diffusion, and where K is space-dependent. These features are further discussed below.

2.3. Dimensionless formulation

We will next proceed with rendering Eq. (1) dimensionless. First, we introduce the notation $s = \frac{\rho_s}{\rho}, i = \frac{\rho_i}{\rho}$, and $r = \frac{\rho_r}{\rho}$, namely we normalize species densities by ρ. Then, we use Eq. (7) for the diffusion terms and Eq. (11) for the reaction rates. Finally, we dimensionalize time by $1/\Lambda$, space by a characteristic external length scale L, K by K_0, and velocities by a characteristic velocity U, and use lower case symbols for all dimensionless variables. We obtain the following set of differential equations

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

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

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

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

Here, we defined the dimensionless Damkohler number $Da = \frac{U}{L\Lambda}$, the dimensionless diffusion number $C = \frac{D}{\Lambda L^2} = \phi^{-2}$, and the rescaled velocity $\boldsymbol{v}$. We recognize ϕ as the Thiele modulus of the chemical reaction engineering literature (Fogler, 2010). From this dimensionless formulation arises naturally the most important parameter R_0, commonly associated with the spreading of epidemics, namely

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

where

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

Eq. (18) shows that even in spatially homogeneous systems, $R_0(\rho, r)$ is not constant, as typically assumed, but also depends both on the areal density and on a measure of the intrinsic 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 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 $\rho_0/\rho \ll 1$ and a linear function for F. For consistency in the remainder, we will denote the R_0 dependence of various solutions to the problem, e.g. of r_∞, through its value at the onset of the process, namely, through $R_0(\rho, 0)$.

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 on the prediction of infection results, as will be shown below.

For completeness, we also remark that had we used the different version for diffusion (5), instead of (7), Eqs. (14)–(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

We are now in a position to consider a number of interesting applications. With the exception of the analytical results for the batch reactor problem, all other numerical results were obtained by solving Eqs. (14)–(16) using a finite-element method with the weak form of these equations implemented and solved in the Fenics software package (Alnæs et al., 2015). The spatial domain was discretized using linear finite elements, and the resulting equations were integrated in time using the trapezoidal rule. The mesh size and the time-step were chosen to be sufficiently small such that further refinement did not lead to significant changes in the resulting solution.

3.1. Perfect Mixing: The SIR model

Consider, first, the zero-dimensional ("batch reactor") problem, with no input or output in the control volume, at conditions of perfect mixing. Dependent variables, as well as density, are not functions of the spatial coordinates, the implication being that mobility effects, such as advection and diffusion are sufficiently large to homogenize the system. The partial differential Eqs. (14)–(16) then become ordinary differential equations in time. Using $\dot{(\cdot)}$ 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

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

We must point out that even though the overall density in such an SIR-like model is time-independent, spatial effects do enter through the effect of density (hence of the extent of contagion) on R_0 (Eqs. (21), (22)).

Eqs. (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 find:

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

thus,

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

Further substitution into (23) gives

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

which can be integrated to provide the final solution

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

Results are shown in the sections to follow. In (24) we have tacitly assumed that at the onset of the process, all individuals (minus an initial infected seed) are susceptible. This assumption does not account for the possibility of immune individuals, e.g. via vaccination. We will discuss this important point later below. We also 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

$$\dot{r}(t) = i_0 - r(1 - s_0 R_0(\rho, 0)) - \frac{s_0 R_0(\rho, 0)^2}{2} r^2 \tag{30}$$

accurate for small R_0. This equation can be used as a simpler model for quick insights. Riccati equations have been used as a simpler alternative to the SIR description (Marmarelis, 2020; Fokas et al., 2020a; Fokas et al., 2020b).

3.1.1. Infection curves, herd immunity and effective R_0

The solutions of Eqs. Eqn 21, 22 and 23 are shown in Fig. 1. We first remark that from Eq. (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 Fig. 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 with significantly larger values if a constant R_0 value is used. One could define an effective constant R_0, using the average

$$R_0 \equiv \frac{\int_0^{r_\infty} R_0(\rho, r')dr'}{r_\infty} \tag{31}$$

Its relation to $R_0(\rho, 0)$ is inferred from Fig. 2, which for example shows that to a constant value of $R_0 = 2.5$ throughout the process corresponds the twice as large initial value of $R_0(\rho, 0) = 4.5$.

Parameter R_0 is typically interpreted as the number of new infections caused on average by an infected individual. Such an interpretation is useful as long as one can specify the interval of time, over which such infections will occur. In general, it will be variable, depending on the frequency and the ultimate exposure of susceptible populations to the infected individual. If we assume for simplicity an exponential rise of infections with constant $s = 1$ and constant R_0 in (22), the interpretation of R_0 as the number of new infections caused on average by an infected individual requires an exposure time of $t_{exp} = \frac{lnR_0}{R_0 - 1}$, e.g. $t_{exp} = 0.549$ for $R_0 = 3$, or $t_{exp} = 0.46$ for $R_0 = 4$, both roughly corresponding to one week, for typical values. By analogy, in *super-spreader* events, for an infected individual to infect a number of others, during the relatively short period of time of such an event, the corresponding

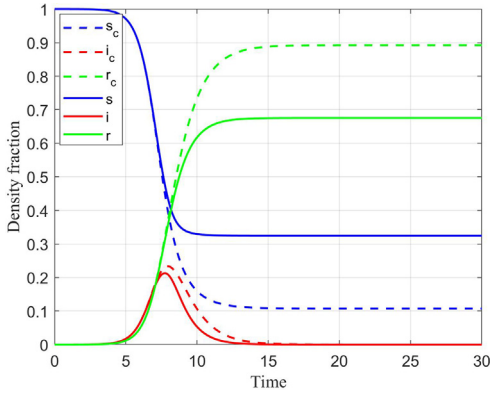

Fig. 1. Variation of susceptible (s), infected (i) and recovered (r) fractions as a function of non-dimensional time for the case when R_0 varies following Eq. (19) (solid lines) with $R_0(\rho, 0) = 2.5$, and for the case of a constant value of $R_0 = R_0(\rho, 0) = 2.5$ (dashed lines), which is the SIR assumption. Note the substantial difference between the results obtained between a constant and a variable R_0. The initial infected fraction is $i(0) = 10^{-5}$.

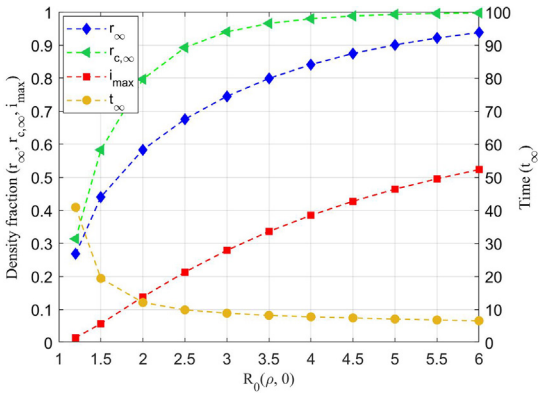

Fig. 2. Variation of herd immunity h (denoted here as r_∞), maximum infection fraction (i_{max}) and the duration of a pandemic (t_∞) as a function of $R_0(\rho, 0)$. Plotted also is the herd immunity ($r_{c,\infty}$) assuming a constant R_0. The initial infection fraction is $i(0) = 0.001$.

value of R_0 must be substantially larger (of the order of tens) than 1.

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

$$t_\infty = \int_0^{r_e} \frac{du}{1 - u - s_0 \exp\left(-\int_0^u R_0(\rho, r')dr'\right)}. \tag{32}$$

where $r_e = 0.99 r_\infty$, and r_∞ is the solution of (33). Results are shown in Fig. 2. We note that 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 lasts longer 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 than intuitively expected.

Plotted in Fig. 2 is also the maximum in the infected fraction, $i_{\max}$, which is an increasing function of $R_0(\rho, 0)$, as expected. Eq. (22) shows that the maximum is reached when $R_0(\rho, r^*)s^* = 1$, where superscript * corresponds to the values at that point.

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_∞, based on a given value of $R_0(\rho, 0)$. It is the solution of the equation

$$1 - h - s_0 \exp\left(-\int_0^h R_0(\rho, r')dr'\right) = 0. \tag{33}$$

Results are shown in Fig. 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. It is important to realize that herd immunity has the commonly accepted interpretation, if and only if the corresponding value of R_0 that resulted in the arresting of the contagion is maintained at long times. Otherwise, and if behavioral changes cause R_0 to increase, because of relaxation in behavior, then contagion will commence again (e.g. second or third waves, etc). We can explain this further by considering the stability of such an asymptotic state.

Consider a small resurgence of infections after an asymptotic state (and its associated herd immunity is reached. As long as the asymptotic condition $R_0(\rho, h)(1 - h) < 1$ is satisfied (which happens to always be the case), such fluctuations will decay exponentially fast, as shown in Eq. (22): the final state is asymptotically stable. This will not be the case, however, when the fluctuation is instead in $R_0(\rho, h)$, e.g. when a new value, e.g. R_0', such that $R_0'(1 - h) > 1$, sets in. For example, this could be the result of relaxations on spatial distancing, and/or of abandoning due caution, after the asymptotic state is reached. Under such conditions, 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. Fig. 3 demonstrates 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 in the specific example 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 the relaxation of restrictions at $t = 10$ and a return to a higher value of $R_0(\rho, 0) = 3$, 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 derives from the fact that the infection reaction (9) is auto-catalytic, 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 desired herd immunity levels to be sustained, at least until immunization reduces considerably the susceptible fraction.

We close by noting that condition $R_0(\rho, r_\infty)(1 - r_\infty) < 1$, also written as $R_0(1 - h) < 1$, namely $h > 1 - \frac{1}{R_0}$ delineates the state of "herd immunity" typically understood by the public. In the absence of a vaccine or of means (e.g. spatial distancing) that keep R_0 small, such herd immunity requires that a non-trivial fraction of the population was infected. For example, for a behavior corresponding to $R_0 = 4$, the associated herd immunity is 0.75. More generally, assume that a fraction of the initial population has been immunized, e.g. through vaccination. A way to model this problem would be to assign this fraction initially to the "recovered" population fraction, namely to take instead of (25) the following initial conditions

$$i(0) = i_0, \qquad s(0) \equiv s_0 = 1 - i_0 - r_0, \qquad r(0) = r_0. \tag{34}$$

Then, the previous solution applies, by using in all integrals over r, in Eqs. (27)–(29), the lower limit of r_0 instead of 0. For infection to not spread, and thus be contained from its onset the following must apply: $R_0(\rho, r_0)s_0 < 1$, which in view of (34), reads $R_0(\rho, r_0)(1 - r_0) < 1$ assuming a small initial fraction of infections. This is the same condition as above for the onset of herd immunity, namely $h > 1 - \frac{1}{R_0}$. For example, if pre-pandemic behavior in terms

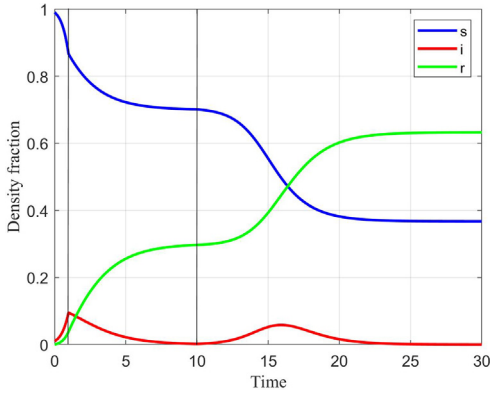

Fig. 3. Variation of susceptible (s), infected (i) and recovered (r) fractions as a function of non-dimensional time with $R_0(\rho, 0) = 3, t \in (0, 1), R_0(\rho, 0) = 0.8, t \in (1, 10), R_0(\rho, 0) = 3, t \in (10, 30)$. The initial infected fraction is $i(0) = 0.01$.

of spatial distance and facial covering has an associated value of $R_0 = 5$, then the corresponding requirement on immunization is $> 80\%$.

3.1.2. Universal scaling

The auto-catalytic 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 effect of the initial condition i_0. Fig. 4 is a plot of the infection curve for different values of the initial condition and for $R_0(\rho,0) = 2.5$. For the typical values considered, 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 are almost identical. The shift is approximately 2 non-dimensional time units for each decrease in the initial condition by a factor of 10.

We can explain these findings by considering the early-time solution of Eq. (22). At small times we obtain

$$i(t) \approx i_0 \exp\left((R_0(\rho,0)-1)t\right) = \exp\left((R_0(\rho,0)-1)(t-t_0)\right), \qquad (35)$$

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 provides a non-trivial interval of time to contain the infection, e.g. by educating the public to modify behavior, thus to lower R_0 and ultimately mitigate 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 regardless, ultimately dependent only on R_0.

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 Eq. (15) over spatial dimensions and assuming perfect mixing we obtain to leading order

$$\dot{i}(t) = (R_0(\rho,0)-1)i + j_i \qquad (36)$$

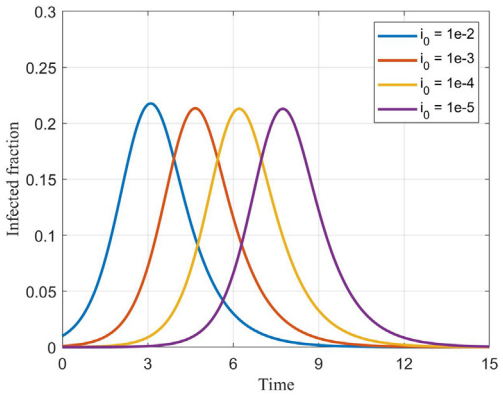

Fig. 4. Variation of infected (i) fraction as a function of non-dimensional time for different values of the initial infected fraction i_0 and $R_0(\rho,0) = 2.5$.

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 Eq. (36) subject to the initial condition $i(0) = 0$, we obtain, to leading-order (small τ)

$$i = j_i t + \cdots \qquad (37)$$

essentially providing an equivalent initial condition, $i(0) = i_0 = j_i\tau$. It follows that as long as $(R_0(\rho,0)-1)\tau \ll j_i$, the impact of "imported infections" is simply to nucleate the process, and the problem reverts to the previous. Namely, the net impact of a ban is to provide additional time for necessary public health measures to be implemented and to lead to as small a value of $R_0(\rho,0)$ as possible, thereby limiting or even preventing an inevitable spread of 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 this assumption in the presence of fluctuations or other non-uniformities, by considering the solution of a typical 1-D problem in the absence of advection. We have

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

$$\frac{\partial i}{\partial t} - C\frac{\partial \ln\rho}{\partial x}\frac{\partial i}{\partial x} = C\frac{\partial^2 i}{\partial x^2} + R_0(\rho,r)si - i \qquad (39)$$

$$\frac{\partial \rho}{\partial t} = 0 \qquad (40)$$

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 $x = \{0,1\}$. Consider either the case where the initial fluctuations are variable, but the density is uniform,

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

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

$$i(x,0) = i_0, \qquad s(x,0) = 1 - i(x,0), \qquad \rho(x,t) = \rho_m(1+\epsilon g(x)), \qquad (42)$$

where $\epsilon \ll 1$. Expecting that diffusion ($C > 0$) will help to smooth spatial non-uniformities, we will focus on the solution in its absence ($C = 0$). Thus, any arising effects 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 averages by superscript bar and fluctuations by superscript prime to obtain

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

$$\frac{\partial \bar{i}}{\partial t} = +R_0(\rho_m)\left(\bar{s}\bar{i} + \overline{s'i'}\right) - \bar{i} \qquad (44)$$

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

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

$$\frac{d\eta}{dt} = +R_0(\rho_m)(\bar{s}\eta + \bar{i}\sigma) - \eta \qquad (46)$$

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

We first note that fluctuations contribute to the rate expression, Eqs. (43) and (44), through the average of a product of fluctuations $\overline{s'i'}$. This contribution is proportional to the square of the amplitude of the fluctuations $\overline{(i_0\epsilon g(x))^2}$. A first conclusion 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 fluctuations are bounded, we must find the eigenvalues ω of the matrix of the linear system (43) and (44),

$$\omega^2 + \left(R_0(\rho_m)\bar{i} - R_0(\rho_m)\bar{s} + 1\right) + R_0(\rho_m)\bar{i} = 0. \tag{47}$$

This equation has two negative eigenvalues as long as

$$R_0(\rho_m)\bar{i} - R_0(\rho_m)\bar{s} + 1 > 0. \tag{48}$$

Eqs. (21)–(23) show that for some time before the infection fraction i reaches its maximum, and always after that time, condition (48) is indeed valid. One concludes that fluctuations will remain bounded, if not altogether decay, hence they have little effect on the average behavior. This result is demonstrated in Fig. 5, where we plot ensemble average responses. The curves obtained are practically the same either in the presence or in the absence of fluctuations.

3.2.2. Spatial heterogeneity

The results are quite different, however, when spatial heterogeneity is stronger. 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 $i_2 = 10^{-3}$)

$$i(x,0) = i_{0,1}H(\xi - x) + i_{0,2}H(x - \xi) \tag{49}$$

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. as suggested in Fig. 4 and Eq. (35)). 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. Fig. 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 fraction of the 24-h 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 the conservation equation $\rho_s = \rho_u \frac{\xi}{1-\xi}$. This setting is intended to model a regular commute between two regions ("home" and "work", or "home" and "school") where we expect significantly different R_0 values. For further simplification, we will take $\xi \ll 1$, and 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 is to represent it 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. (Note the assumption of no infection at home made.) The results of this hypothetical "commute", for an 8-h work period ($\lambda = 1/3$), are shown in Fig. 7. Superimposed are also the results that would have emerged if the "work" parameter $R_{0,w}(\rho, r)$ was applied in both regions (or, equivalently, when $\lambda = 1$). As expected for this example, commuting leads to a lower effective value $R_{0,eff}(\rho, 0) = 1.67$, although it does not suppress infection: As long as $R_{0,eff}(\rho, 0) > 1$ infection will occur, although by reducing exposure (decreasing λ) the resulting effective rates are lower. We can show using the asymptotic method of "two-timing" (Holmes, 2012) that for all values of λ the corresponding effective value is the arithmetic mean, $R_{0,eff}(\rho, 0) = \lambda R_{0,w}$. This suggests that there

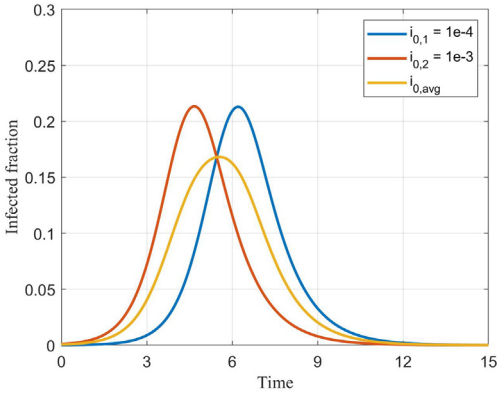

Fig. 6. The infected (i) fraction curves as functions of non-dimensional time for the two different regions with different initial infections, and the composite curve ($R_0(\rho, 0) = 2.5$).

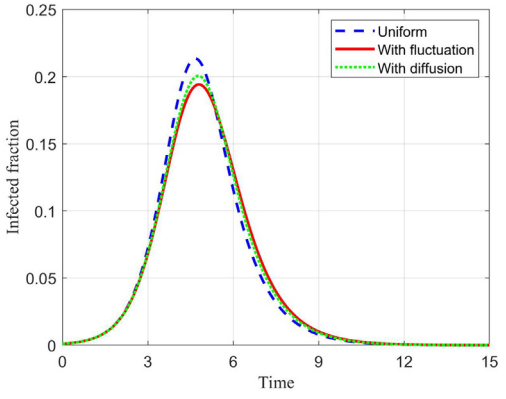

Fig. 5. Variation of the ensemble average of the infected (i) fraction as a function of non-dimensional time for $R_0(\rho, 0) = 2.5$. The initial condition is Eq. (41) with $i_0 = 0.05, g(x) = \sin(x)$, and $\epsilon = 0.04$.

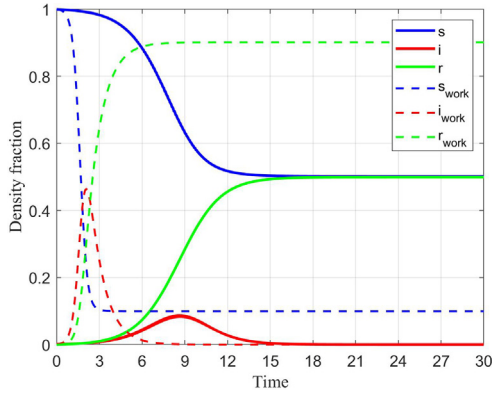

Fig. 7. The infection curves for the case of a commute between two regions of high and low values of R_0 ($R_{0,w}(\rho, 0) = 5$) and 0, and exposure frequency of $\lambda = 1/3$ (solid lines). Plotted also are the infection curves corresponding to a uniform $R_{0,w}(\rho, 0) = 5$ (equivalently to $\lambda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

is a critical exposure value $\lambda_{crit} = \frac{1}{R_{0,w}}$, below which contagion might be suppressed. These findings must be interpreted with due caution, as they are subject to the assumptions made, and specifically the lack of possible infection while at home. This can be relaxed, as needed, with a corresponding effect on the effective value, which remains the arithmetic mean.

3.3. The effect of transport

The preceding sections dealt with applications in zero-dimensional space (batch reactors or their ensembles), namely under the assumption of fast diffusion or other transport. Transport (via advection, diffusion or dispersion) plays two roles: The first is to reduce the effect of small spatial non-uniformities, as discussed above. The second is in the direction of further spreading the infection spatially. This section will consider such mobility effects, starting first with diffusion.

3.4. Effect of diffusion

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

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

$$\frac{\partial i}{\partial t} - C\nabla(\ln\rho)\cdot\nabla i = C\nabla^2 i + R_0(\rho, r)si - i \tag{51}$$

$$\frac{\partial r}{\partial t} - C\nabla(\ln\rho)\cdot\nabla r = C\nabla^2 r + i \tag{52}$$

We will first consider diffusion in one spatial dimension. Before proceeding, we note that the SIR solution is obtained in the limit of large C, where from Eqs. (50)–(52) all spatial profiles become flat, and all variables become only functions of time, in a dependence which is identical to the SIR problem.

3.4.1. 1-D geometries

Assume finite values of C, 1-D rectilinear geometries, and spatially constant density. The initial infection conditions are taken to be non-zero in a specified interval, e.g. $-1 < x < 1$, 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, and particularly, whether or not traveling waves develop.

Introducing a moving coordinate $\xi = x - Vt$, where V is the wave velocity, and assuming that a steady-state (denoted by tilde) is reached in these coordinates one obtains,

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

$$-V\frac{\partial \tilde{i}}{\partial \xi} = C\frac{\partial^2 \tilde{i}}{\partial \xi^2} + R_0\tilde{s}\tilde{i} - \tilde{i} \tag{54}$$

$$-V\frac{\partial \tilde{r}}{\partial \xi} = C\frac{\partial^2 \tilde{r}}{\partial \xi^2} + \tilde{i} \tag{55}$$

subject to no-flux conditions at the ends

$$\frac{\partial \tilde{s}}{\partial \xi} = \frac{\partial \tilde{i}}{\partial \xi} = \frac{\partial \tilde{r}}{\partial \xi} = 0, \text{ at } \xi = \pm\infty. \tag{56}$$

The invariance of (53)–(55) 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 the opposite velocity $-V$. Let ξ_u and ξ_d be two locations sufficiently upstream and downstream, respectively, of the wave front. We then expect

$$\tilde{r}(\xi_u) = r_{V,\infty}, \tilde{i}(\xi_u) = 0, \text{ and } \tilde{r}(\xi_d) = 0, \tilde{i}(\xi_d) = 0 \tag{57}$$

where we have anticipated that $r_{V,\infty}$ might not be identical to the r_∞ of the batch reactor result, Eq. (33).

The forward wave velocity can be determined by integrating (55) between ξ_u and ξ_d,

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

and, more generally,

$$V = \frac{1}{r_{V,\infty}}\int_{-\infty}^{\infty} \tilde{i}d\xi. \tag{59}$$

As long as the integral in (59) is non-zero, which is always the case in a contagion, a wave solution exists. This wave spreads at a constant velocity, driven by diffusion, with a magnitude that expresses the intensity of the contagion.

3.4.2. Invariance

We can explicitly remove the C-dependence by introducing rescaled space coordinates and velocities, $\xi = \sqrt{C}\zeta$ and $V = W\sqrt{C}$. Denoted by superscript hat, the profiles satisfy

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

$$-W\frac{\partial \hat{i}}{\partial \zeta} = \frac{\partial^2 \hat{i}}{\partial \zeta^2} + R_0\hat{s}\hat{i} - \hat{i} \tag{61}$$

$$-W\frac{\partial \hat{r}}{\partial \zeta} = \frac{\partial^2 \hat{r}}{\partial \zeta^2} + \hat{i} \tag{62}$$

subject to no-flux conditions at the two ends. The velocities now become

$$W = \frac{1}{r_{V,\infty}}\int_{-\infty}^{\infty} \hat{i}d\zeta, \tag{63}$$

$$\mathscr{V} = \frac{\sqrt{D\Lambda}}{r_{V,\infty}}\int_{-\infty}^{\infty} \hat{i}d\zeta. \tag{64}$$

Although the initially infected region $x \in (a, b)$ transforms into $\zeta \in \left(aC^{-1/2}, bC^{-1/2}\right)$, thus containing a C-dependence, the asymptotic traveling wave is independent of initial conditions, hence of C. The resulting infection curves in the presence of diffusion as well as the wave velocity dependence on R_0 are further explored below.

Numerical results of the solution of (50)–(52) are shown in Fig. 8, for a problem in which the initial infection region is near the left boundary, for $R_0(\rho, 0) = 2.5$. As anticipated, the infection profiles evolve as a function of time, and reach an asymptotic traveling wave after about $t = 3$. The wavelike nature of the solution is evident if one plots the infected fraction in space–time coordinates (Fig. 9). A ridge with a constant slope is clearly seen.

Consider next, the question of how different are the infection profiles when diffusion is included. Fig. 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 included and $C = 1.0$. Plotted also are the results of the batch reactor (SIR) problem, for the same value of $R_0(\rho, 0) = 2.5$. Diffusion does affect the shape of the curves obtained, leading to a slightly smaller (for this value of $R_0(\rho, 0)$) infection intensity, essentially corresponding to a slightly smaller effective $R_0(\rho, 0)$. Fig. 11 is a plot of the asymptotic

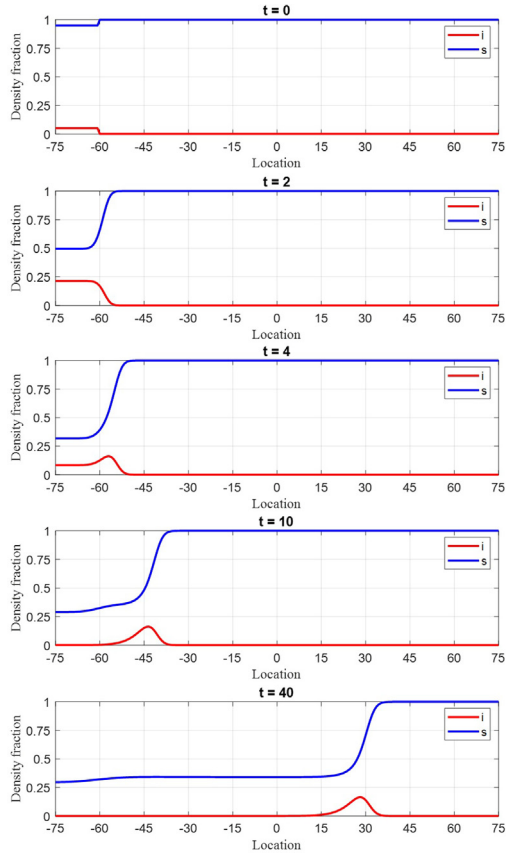

Fig. 8. Infected and susceptible density fraction profiles at different values of time ($R_0(\rho,0)=2.5$).

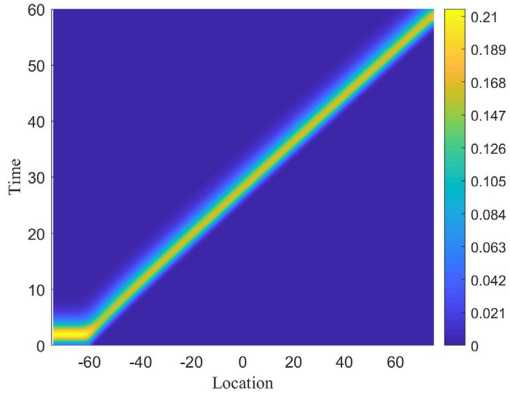

Fig. 9. Infected density fraction as a function of space and time coordinates ($R_0(\rho,0)=2.5$).

value r_∞ and of the maximum in infections i_{max} as a function of $R_0(\rho,0)$ for the respective two cases. Diffusion acts to moderate somewhat the contagion intensity. Overall, the two solutions, corresponding to the contagion wave (Eqs. (50) and (51)), now denoted by $i_D(t)$ and to the batch (SIR) problem (Eqs. (21) and (22)), now denoted by $i_B(t)$, are approximately equal, for this value of $R_0(\rho,0)$, although they are not identical

$$i(x,t) \to i(x-Vt) \equiv \hat{i}(\zeta) \equiv i_D\left(const - \frac{x}{V} + t\right)$$
$$\approx i_B\left(const - \frac{x}{V} + t\right) \tag{65}$$

Also note that the constant in (65) can absorbed in x/V.

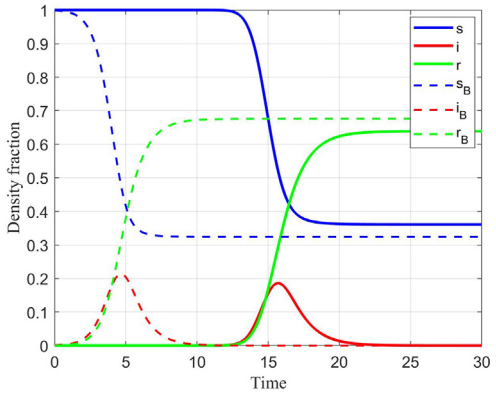

Fig. 10. Variation of susceptible (s), infected (i) and recovered (r) fractions (steady state) as a function of non-dimensional time with diffusion included (solid lines) and for the batch reactor (SIR) problem (dashed lines) for the same value of $R_0(\rho,0)=2.5$. The initial infection fraction for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

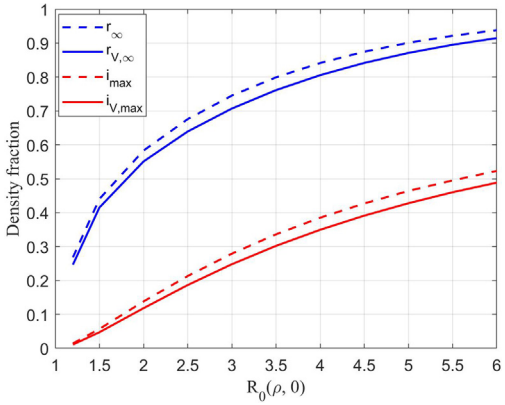

Fig. 11. Variation of herd immunity (r_∞) and maximum infection fraction (i_{max}) as a function of $R_0(\rho,0)$ for the two cases of a batch system (dashed lines) and when diffusion is included (solid lines). The initial infection fraction for the batch system is $i_B(0)=0.001$.

The variation of the dimensionless velocity, Eq. (63) with $R_0(\rho,0)$ is shown in Fig. 12. As expected it increases with $R_0(\rho,0)$. Interestingly, it is varying roughly in a linear fashion at relatively large values of the latter. Even for a finite value of the diffusion parameter C, the time profiles at a fixed spatial location are close to those for the SIR model, assuming sufficiently large

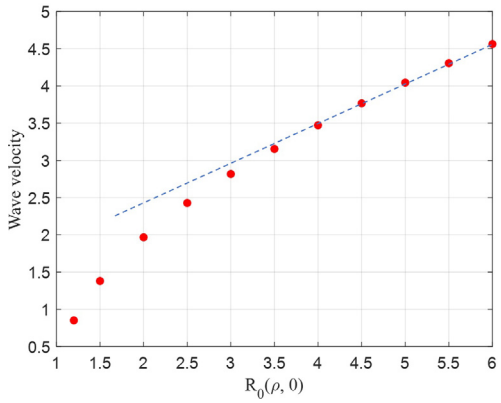

Fig. 12. Variation of non-dimensional wave speed in a one-dimensional problem as a function of $R_0(\rho,0)$.

$R_0(\rho,0)$. We can show this by introducing the new spatial variable $\chi = \zeta/W$ in Eqs. (60)–(62). For example, (62) will read as

$$-\frac{\partial \hat{r}}{\partial \chi} = \frac{1}{W^2}\frac{\partial^2 \hat{r}}{\partial \chi^2} + \hat{i} \tag{66}$$

Now, at relatively large values of $R_0(\rho,0)$, we can ignore the second derivative in the RHS of (66), which then becomes a set of equations identical to the SIR problem, subject to the equivalence $\chi \to -t$. This equivalence does not carry over when $R_0(\rho,0)$ is smaller and closer to 1, in which case the intensity of infection is smaller. Because this regime is of lesser interest, however, it will not be further explored here.

We conclude that, even in the absence of advection, diffusion is sufficient to spread the contagion at a constant velocity, with the rate of spreading increasing with the square root of the diffusion coefficient and with the value of $R_0(\rho,0)$. Diffusion affects the overall infection intensity, although only mildly, and at larger $R_0(\rho,0)$. 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. It is also worth pointing out that because of the presence of the reaction terms, the effect of diffusion is to lead to fronts of constant velocity, as opposed to a velocity that varies with the square root of time, as is the case in the absence of reactions.

For completeness, we also simulated the "collision" of two waves, one moving from the left and the other from the right. Fig. 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. For a number of reasons, these waves are different from those observed in other wave problems (e.g. solitary waves, where the wave velocities increase with the wave amplitude (Drazin and Johnson, 1989)): In the present context, the wave amplitude, e.g. i_{max} cannot exceed the value of 1.

To conclude this section we next consider effects of heterogeneity in the 1-D rectilinear geometry. Figs. 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 figures this occurs at $x = -25$. As it enters the low-density region the wave decelerates, and the magnitude of the infected fraction decreases. The slowing of the wave is indicated by the increase of the slope in the $x - t$ domain. Conversely, at $x = 25$, when the wave re-enters a region of higher $R_0(\rho,0)$, it accelerates to a larger velocity, the intensity of the infected fraction correspondingly increasing.

3.4.3. Radial geometries

The results of the 1-D rectilinear geometry apply almost identically to radial geometries. Consider first an extension to radially symmetric 2-D geometries. The application to more general 2-D geometries is considered in the subsequent section. Now, Eqs. (50) and (51) become

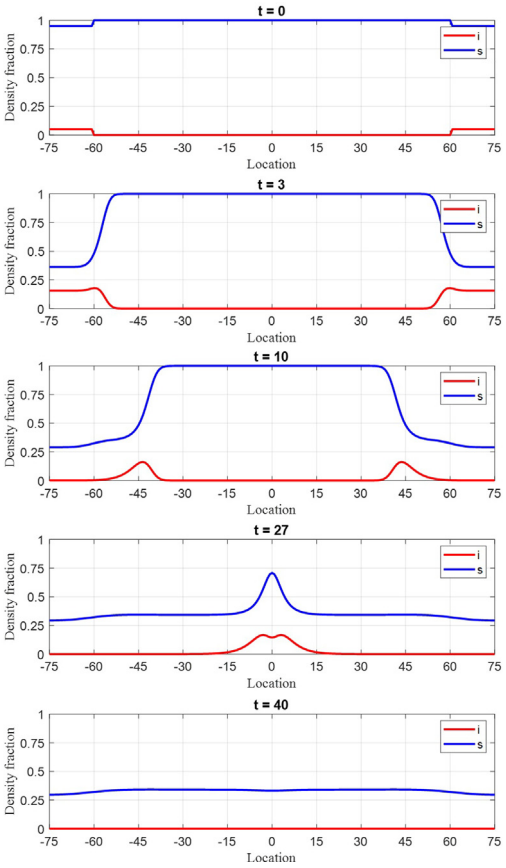

Fig. 13. Infected and susceptible density fraction profiles at different values of time for two waves propagating towards the center of the domain ($R_0(\rho,0) = 2.5$).

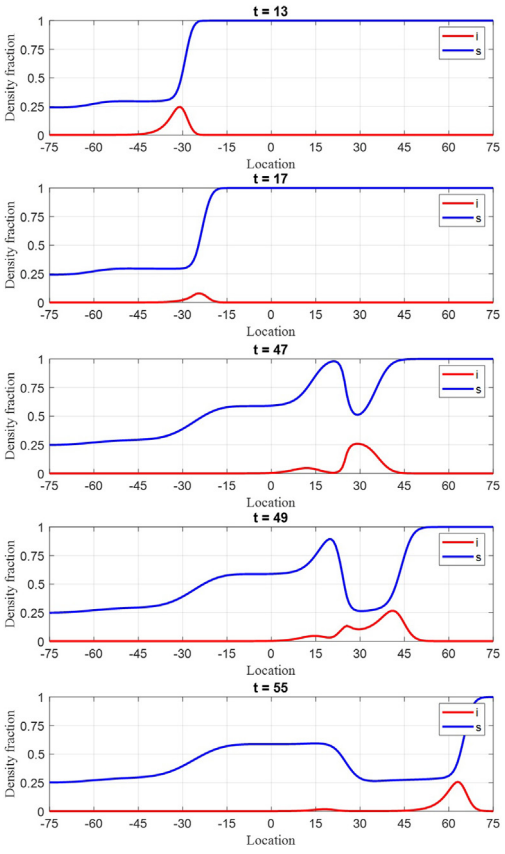

Fig. 14. Infected and susceptible density fraction profiles at different values of time in a 1-D heterogeneous system. For $x \in (-80,-25)$ and $x \in (25,80), R_0(\rho,0) = 3$, while for $x \in (-25,25), R_0(\rho,0) = 1.5$.

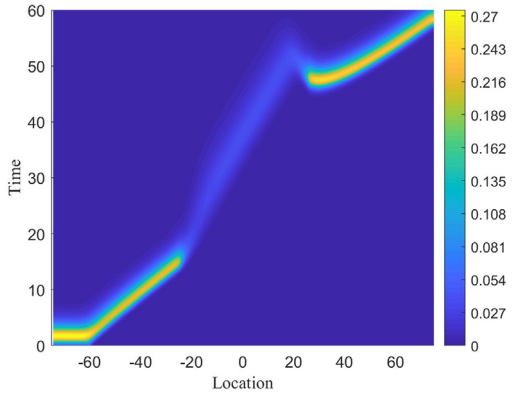

Fig. 15. Infected density fraction as a function of space and time coordinates in a 1-D heterogeneous system: For $x \in (-80,-25)$ and $x \in (25,80)$, $R_0(\rho,0)=3$, while for $x \in (-25,25)$, $R_0(\rho,0)=1.5$.

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

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

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

where we used μ to denote the radial coordinate. 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

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

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

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

By further taking the limit of large t', Eqs. (70) and (71) transform to the same equations as (53) and (54). It follows that the same results hold for radial geometries as for the rectilinear problem. The simulations for the 1-D radial geometry, shown in Fig. 16, confirm these findings. As before, the asymptotic solution in this case of reaction–diffusion system does not enter in the 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.

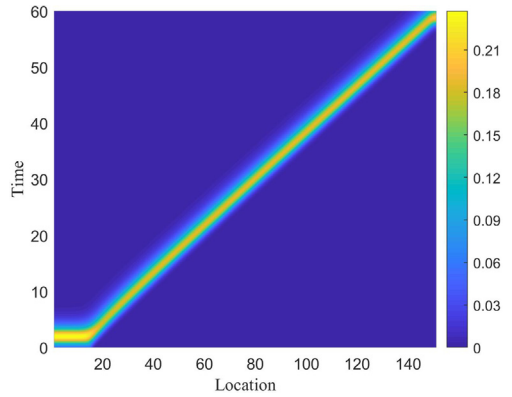

Fig. 16. Infected density fraction as a function of radial (space) and time coordinates ($R_0(\rho,0)=2.5$).

3.4.4. Effect of heterogeneity in two dimensions

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

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 Fig. 17). As expected, an infection wave emerges in the middle layer, traveling with a velocity v_1 corresponding to the high R_0 value in that layer. This is indeed equal to the corresponding asymptotic 1-D wave velocity of 3.15 (see Fig. 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 also 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 are manifested as linear fronts, 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, Eq. (73) gives

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

therefore, we find

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

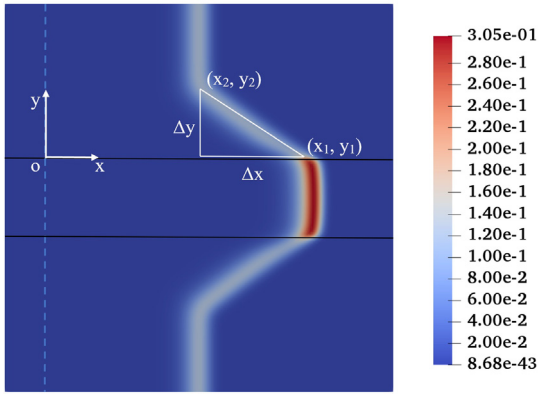

Fig. 17. Infected density fraction (i) at $t = 40$ in a layered medium with an infection wave starting at $x = 0$. In the outer layers $R_0(\rho,0)=2$ while in the inner layer $R_0(\rho,0)=3.5$. The spatial dimension of the domain is 150×150.

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 Fig. 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 above, (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

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

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-quadrant systems. A different illustration of these effects is shown in Fig. 18, where infection initiates in the upper right corner of a rectangular domain with four different quadrants (NE, SE, NW and SW), each taken with a different density (hence different values of $R_0(\rho, 0)$). The idea is to simulate how infection can spread in a "country" consisting of four hypothetical quadrants, with different densities, hence different values of $R_0(\rho, 0)$. A similar simulation of the spreading of infection in the Lombard region of Italy using a different continuum model was provided in Viguerie et al. (2020). In our work, we have assumed that the two diagonally opposite quadrants (e.g. NE and SW) have relatively larger values of $R_0(\rho, 0) = 3.5$, hence an asymptotic wave velocity of 3.15, the other two (SE, NW) being of relatively smaller values of $R_0(\rho, 0) = 2$, and with 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 Fig. 18). When the SE and NW boundaries are encountered (lower right panel in Fig. 18), the waves slow down and start spreading in a radial manner as they enter the two quadrants at the lower 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 Fig. 18). Upon touching the boundaries with the NW and SE regions, diffusion causes infection waves to start emanating from the higher to the lower density regions, resulting into a flat front, similar to that for the layered system, which connects the two waves in the two different regions. For the same reasons as in the layered system, this front has a slope equal to the ratio of the two respective velocities, namely as in (76). The simulations confirm these results. Similar results can be obtained when the four quadrants are heterogeneous, where the values of $R_0(\rho, 0)$ are distributed in a stochastic manner in the regions.

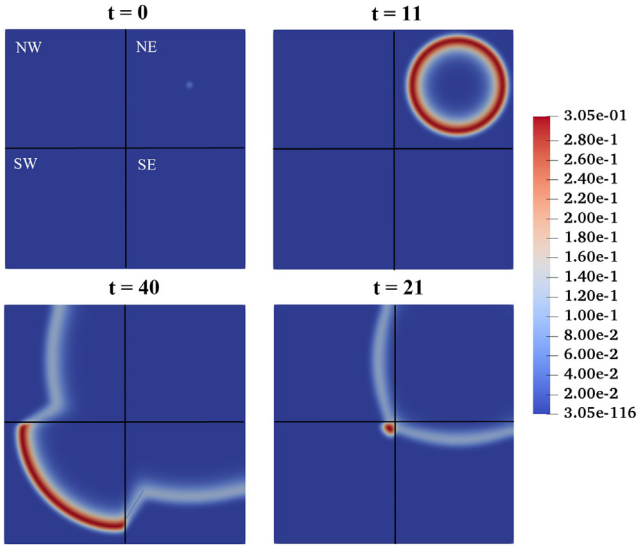

Fig. 18. Contour plot of the infected density fraction a region with four quadrants with different values of $R_0(\rho, 0)$ at different values of time. $R_0(\rho, 0) = 3.5$ for NE and SW, and $R_0(\rho, 0) = 2$ for NW and SE, respectively. The spatial dimension of the region is 150×150.

We end this section by noting that one can further explore the use of a wave (or eikonal) equation for the description of the spreading of contagion. This would be the objective of a future study.

3.5. Effect of advection

Consider, next, the effect of advection. For simplicity, we proceed by keeping the same advective velocity for all three populations. The corresponding dimensionless equations read as follows

$$\frac{\partial s}{\partial t} + Da\boldsymbol{v} \cdot \nabla s = C\nabla^2 s - R_0(\rho, r) s i \tag{80}$$

$$\frac{\partial i}{\partial t} + Da\boldsymbol{v} \cdot \nabla i = C\nabla^2 i + R_0(\rho, r) s i - i \tag{81}$$

$$\frac{\partial r}{\partial t} + Da\boldsymbol{v} \cdot \nabla r = C\nabla^2 r + i \tag{82}$$

along with (17). The simplest result is obtained when one assumes that the velocity is constant in space and time. Consider the case of 1-D geometries, and take without loss $\boldsymbol{v} = \boldsymbol{e}_x$, where $\boldsymbol{e}_x$ is the unit vector in the x-direction. By following the same reasoning as in the case of diffusion, we arrive at the same results corresponding to the traveling waves, Eq. (63), except that in the forward moving wave the velocity is enhanced

$$W = v + \frac{1}{r_{V,\infty}} \int_{-\infty}^{\infty} \hat{i} d\zeta, \tag{83}$$

where we defined $v = Da/\sqrt{C}$. Conversely, the backwards moving wave is retarded leading to a propagation velocity of

$$W = -v + \frac{1}{r_{V,\infty}} \int_{-\infty}^{\infty} \hat{i} d\zeta \tag{84}$$

As expected, the contribution of advection in this constant advection case is to speed up (or slow down) the spreading velocity in the direction of advection with a contribution that is equal to $Da/\sqrt{C} = U/\sqrt{D\Lambda}$.

More interesting is the problem when the velocity vector is a stochastic variable, distributed heterogeneously in space. Such a distribution will render all species to stochastic variables as well, and the relevant variables are ensemble averages. We have encountered some of this feature in a previous section dealing with fluctuations. The effect of heterogeneity will enter in two ways: as a correction to the non-linear reaction terms, via the ensemble average of the product of fluctuations, as discussed previously; and as the ensemble average of the product of the fluctuations of the velocity with the gradient of the fluctuation, namely $< \boldsymbol{v}' \cdot \nabla s' >$. The latter effect arises in many related contexts, for example in the advection of a scalar in turbulent flows (Sreenivasan, 2019), or in transport in heterogeneous porous media (Gelhar and Axness, 1983) and extensions to reacting flows.

It will manifest itself in the form of an effective macrodispersivity, which is dominant over molecular diffusion, namely

$$< \boldsymbol{v}' \cdot \nabla s' >= D_{macro} \nabla^2 < s > \tag{85}$$

and likewise for the other species. For example, in the case of flow in heterogeneous porous media, the macrodispersivity is proportional to the mean velocity and to a characteristic correlation length λ,

$$D_{macro} = \alpha |\boldsymbol{v}| \lambda. \tag{86}$$

We have simulated the solution of a 2-D problem, with $Da = 1, C = 0.1$, and where the velocity field is distributed stochastically in a divergence-free field, with zero mean, a dimensionless correlation length of 0.05, and $R_0(\rho, 0) = 2.5$. Fig. 19 shows in one of the panels the velocity field. The remaining panels show the state of the infected fraction at different time instances. We note that the wave propagates approximately radially; however the front bends and folds at several places determined by the local velocity field.

In Fig. 20 we have plotted the ensemble average of the infection front obtained from ten such stochastic velocity fields. The results suggest a wave that moves radially at a constant speed, reminiscent of the spreading in the radial geometry driven by diffusion alone, as would be expected for the case of macrodispersion. The thickness of the front is substantially larger, however, indicating the enhanced dispersion, as anticipated. More quantitative analyses of this important effect will be the subject of a future communication.

3.6. Some final remarks

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 $R_0 \approx 1.5 - 2.5$ (Noble, 1974; Nikbakht et al., 2009). Further, for both diseases the infectious period is around two weeks, and therefore $\Lambda \approx 1/14\ \text{days}^{-1}$. To predict the wave speed, we need an estimate of the respective diffusion (or dispersion) coefficients. It is to be expected that in the medieval times, the mobility of humans, therefore of infection, was much smaller. In Noble (1974), Rashevsky (1968) this was determined by how quickly information was disseminated (in that case by word of mouth), resulting into an estimate of about 27 miles^2/day. This can be interpreted as either a diffusivity or a macrodispersivity. Using the obtained estimate in our estimate for the speed of the contagion (see Fig. 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 (Langer, 1964).

For the 2009 influenza pandemic, we can make a similar back-of-the-envelope calculation. Because of the much higher increased mobility, and the lack of significant public health awareness, we can estimate that the equivalent random walk radius was an order of magnitude larger compared to the medieval times. We will assume a diffusion coefficient of $2,500\ \text{miles}^2\text{day}^{-1}$, or equivalently, a diffusivity of about $10^5\ \text{m}^2/\text{sec}$. Clearly, the mean radius of 50 miles so assumed incorporates a mix of various travel or other commute activities in the modern world. 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 $30\ \text{days}^{-1}$, and therefore a diffusivity (or macrodispersivity) of around $3,000\ \text{miles}^2$/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

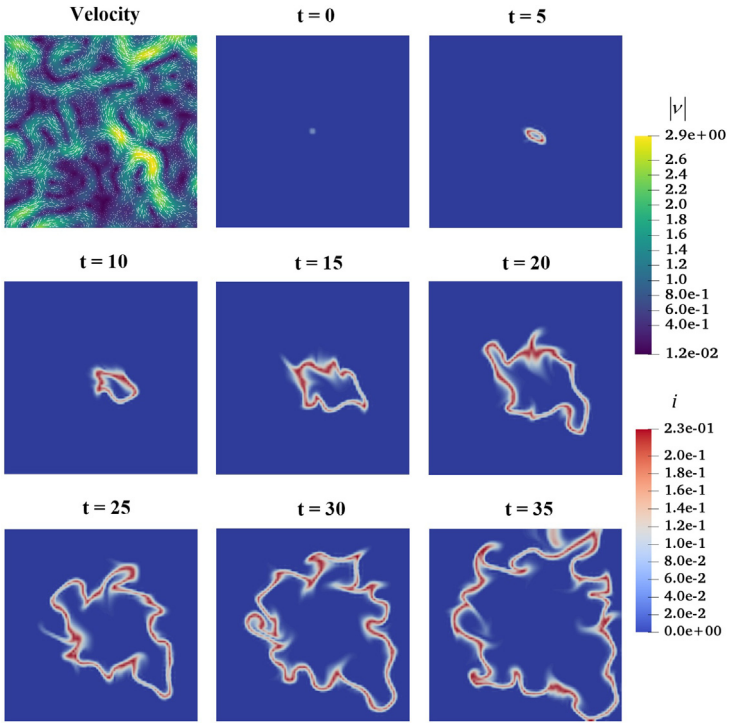

Fig. 19. Simulation in a domain with random velocity field and constant diffusion and R_0. Top left: velocity field for the simulation. From top left to bottom right: images of the infected fraction at different time instances.

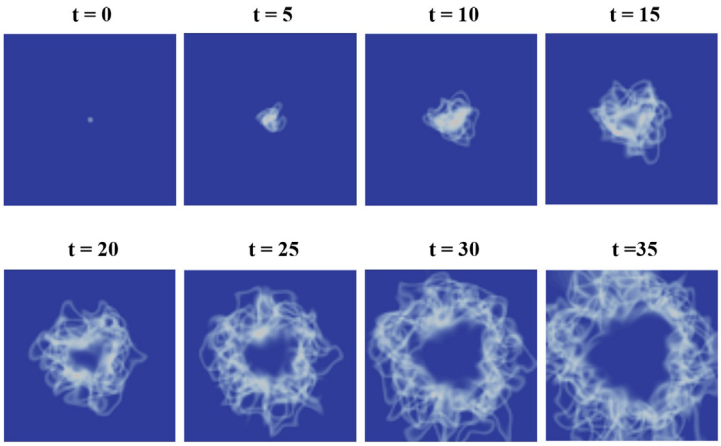

Fig. 20. Simulation in ten different domains with random velocity field and constant diffusion and R_0. From top left to bottom right: ensemble-averaged images of the infected fraction at different time instances.

ball-park of the observed wave speed of around 23 miles/day for the 2009 influenza pandemic (Gog et al., 2014). The higher diffusion/dispersion coefficient for the 2009 influenza pandemic leads to a significantly faster wave speed. All these results should be also interpreted based on the understanding that the enhanced diffusion coefficient reflects in all likelihood macrodispersion driven by the fluctuating advective velocity field. Further work in this area is needed.

4. Conclusions

In this paper 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 case of perfect mixing, where profiles are spatially uniform, 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 (namely, density effects) on the kinetic coefficient 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 mobility, by first considering diffusion in 1-D geometries. We showed the emergence of infection waves in rectilinear and also in 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 curves at a fixed time is similar to that for an SIR system, which they approach at relatively large values of R_0. 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 equivalently by wave propagation, the speed of which only depends on the value of R_0 in that region. Finally, we explored aspects of advection in the case of a stochastic velocity distribution. Such advection leads to an effective macrodispersion, which dominates over diffusion, and leads to enhanced "mixing", hence effectively enhances the spreading of the infection. More work is required to fully explore fully these effects, however. Spatial effects via density, diffusion or advection considerations are important and must be considered in the description of contagion and its epidemics.

The fundamental purpose of this work was to provide a robust framework that allows the modeling of the spreading of human-to-human infections by accounting for mobility (diffusion and advection) and chemical reaction-like attributes. A number of new insights have been uncovered, from how to represent the kinetic parameters and how to include density dependence, to the role of diffusion and advection in the spatial propagation of infection. This has captured essential, although not all relevant variables. For example, limiting the various sub-populations to three ignores important demographic attributes as well as the possibility of correlations in infections, e.g. through family or work relationships. Reflecting human behavior, the important parameter R_0 can vary in space and time. Such variations were considered in many of our applications. Solving the inverse problem of matching real data has not been attempted in this work, other than in a qualitative way. This task is to be considered next, by possibly extending the approach to additional populations. Regardless, we believe that the approach presented here provides a fundamentally sound framework, based on extending fundamental physico-chemical principles to model collective human behavior, and should be useful to the further understanding of the spreading of infections.

CRediT authorship contribution statement

Harisankar Ramaswamy: Methodology, Software, Validation, Data curation, Visualization. **Assad A. Oberai:** Methodology, Validation, Writing - review & editing, Supervision. **Yannis C. Yortsos:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Alnæs, M., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M.E., Wells, G.N., 2015. The fenics project version 1.5. Arch. Num. Softw. 3 (100).

Anderson, R.M., May, R.M., 1979. Population biology of infectious diseases: Part i. Nature 280 (5721), 361–367.

Bensoussan, A., Lions, J.-L., Papanicolaou, G., 2011. Asymptotic Analysis for Periodic Structures, vol. 374, American Mathematical Soc.

Bird, R.B., Stewart, W.E., Lightfoot, E.N., Meredith, R.E., 1961. Transport phenomena. J. Electrochem. Soc. 108 (3), 78C.

Bourouiba, L., Dehandschoewercker, E., Bush, J.W., 2014. Violent expiratory events: on coughing and sneezing. J. Fluid Mech. 745, 537–563.

Burghardt, K., Lerman, K., 2020. Unequal impact and spatial aggregation distort covid-19 growth rates, arXiv preprint arXiv:2004.12994.

Chaves, A., 1998. A fractional diffusion equation to describe lévy flights. Phys. Lett. A 239 (1–2), 13–16.

Coifman, R.R., Lafon, S., 2006. Diffusion maps. Appl. Comput. Harmonic Anal. 21 (1), 5–30.

Croccolo, F., Roman, H.E., 2020. Spreading of infections on random graphs: A percolation-type model for covid-19. Chaos, Solitons & Fractals, 110077.

del Castillo-Negrete, D., Carreras, B., Lynch, V., 2003. Front dynamics in reaction-diffusion systems with levy flights: a fractional diffusion approach. Phys. Rev. Lett. 91 (1), 018302.

Drazin, P.G., Johnson, R.S., 1989. Solitons: An Introduction, vol. 2. Cambridge University Press.

Érdi, P., Tóth, J., 1989. Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models. Manchester University Press.

Fogler, H.S., 2010. Essentials of Chemical Reaction Engineering: Essenti Chemica Reactio Engi. Pearson Education.

Fokas, A.S., Dikaios, N., Kastis, G.A., 2020a. Covid-19: Predictive mathematical models for the number of deaths in south korea, italy, spain, france, uk, germany, and usa, medRxiv.

Fokas, A.S., Cuevas-Maraver, J., Kevrekidis, P.G., 2020b. Two alternative scenarios for easing covid-19 lockdown measures: one reasonable and one catastrophic, medRxiv.

Gelhar, L.W., Axness, C.L., 1983. Three-dimensional stochastic analysis of macrodispersion in aquifers. Water Resour. Res. 19 (1), 161–180.

Gog, J.R., Ballesteros, S., Viboud, C., Simonsen, L., Bjornstad, O.N., Shaman, J., Chao, D. L., Khan, F., Grenfell, B.T., 2014. Spatial transmission of 2009 pandemic influenza in the us. PLoS Comput. Biol. 10 (6), e1003635.

Harko, T., Lobo, F.S., Mak, M., 2014. Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death and birth rates. Appl. Math. Comput. 236, 184–194.

Hirschfelder, J.O., Curtiss, C.F., Bird, R.B., Mayer, M.G., 1964. Molecular Theory of Gases and Liquids, vol. 165. Wiley, New York.

Hoffmann, F., Hosseini, B., Oberai, A.A., Stuart, A.M., 2019. Spectral analysis of weighted laplacians arising in data clustering, arXiv preprint arXiv:1909.06389.

Holmes, M.H., 2012. Introduction to Perturbation Methods, vol. 20, Springer Science & Business Media.

Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. Proc. Roy. Soc. London. Series A, Contain. Pap. Mathe. Phys. Character 115 (772), 700–721.

Lamb, H., 1993. Hydrodynamics. Cambridge University Press.

Langer, W.L., 1964. The black death. Sci. Am. 210 (2), 114–121.

Liu, Q., Jiang, D., 2019. Dynamical behavior of a stochastic multigroup sir epidemic model. Physica A 526, 120975.

Mandelbrot, B.B., 1982. The fractal geometry of. Nature, 394–397.

Marmarelis, V., 2020. Predictive modeling of covid-19 data in the us: Adaptive phase-space approach. IEEE Open J. Eng. Med. Biol.

Nikbakht, R., Baneshi, M.R., Bahrampour, A., Hosseinnataj, A., 2009. Comparison of methods to estimate basic reproduction number (r0) of influenza, using canada 2009 and 2017–18 a (h1n1) data, J. Res. Med. Sci.: Off. J. Isfahan Univ. Med. Sci. 24.

Noble, J.V., 1974. Geographic and temporal development of plagues. Nature 250 (5469), 726–729.

Rashevsky, N., 1968. Looking at history through mathematics, Cambridge. MIT Press, Mass.

Ross, S., 2009. A First Course in Probability. Pearson.

Sharipov, F., Strapasson, J.L., 2012. Ab initio simulation of transport phenomena in rarefied gases. Phys. Rev. E 86 (3), 031130.

Sreenivasan, K.R., 2019. Turbulent mixing: A perspective. Proc. Nat. Acad. Sci. 116 (37), 18175–18183.

Viguerie, A., Lorenzo, G., Auricchio, F., Baroli, D., Hughes, T.J., Patton, A., Reali, A., Yankeelov, T.E., Veneziani, A., 2020. Simulating the spread of covid-19 via spatially-resolved susceptible-exposed-infected-recovered-deceased (seird) model with heterogeneous diffusion, arXiv preprint arXiv:2005.05320.

Viguerie, A., Veneziani, A., Lorenzo, G., Baroli, D., Aretz-Nellesen, N., Patton, A., Yankeelov, T.E., Reali, A., Hughes, T.J., Auricchio, F., 2020. Diffusion–reaction compartmental models formulated in a continuum mechanics framework: application to covid-19, mathematical analysis, and numerical study. Comput. Mech. 66 (5), 1131–1152.

Wells, W.F., 1934. On air-borne infection: Study ii. droplets and droplet nuclei. Am. J. Epidemiol. 20 (3), 611–618.

Wesselingh, J., Krishna, R., et al., 2000. Mass Transfer in Multicomponent Mixtures. Delft University Press Delft.

Xie, X., Li, Y., Chwang, A., Ho, P., Seto, W., 2007. How far droplets can move in indoor environments–revisiting the wells evaporation–falling curve. Indoor air 17 (3), 211–225.