

A Fundamental Inconsistency in the SIR Model Structure and Proposed Remedies

Michael Nikolaou¹

¹ Chemical & Biomolecular Engineering Department, University of Houston, Houston, TX 77204-4004
E-mail: nikolaou@uh.edu

Abstract – The susceptible-infectious-removed (SIR) compartmental model structure and its variants are a fundamental modeling tool in epidemiology. As typically used, however, this tool may introduce an inconsistency by assuming that the rate of depletion of a compartment is proportional to the content of that compartment. As mentioned in the seminal SIR work of Kermack and McKendrick, this is an assumption of mathematical convenience rather than realism. As such, it leads to underprediction of the infectious compartment peaks by a factor of about two, a problem of particular importance when dealing with availability of resources during an epidemic. To remedy this problem, we develop the dSIR model structure, comprising a single delay differential equation and associated delay algebraic equations. We show that SIR and dSIR fully agree in assessing stability and long-term values of a population through an epidemic, but differ considerably in the exponential rates of ascent and descent as well as peak values during the epidemic. The novel Padé-SIR structure is also introduced as an approximation of dSIR by ordinary differential equations. We rigorously analyze the properties of these models and present a number of illustrative simulations, particularly in view of the recent coronavirus epidemic. Suggestions for further study are made.

Keywords – Epidemics, SIR model, Coronavirus, Flattening the Curve, Social Distancing, Population Dynamics, Control

1. Introduction

In their landmark 1927 publication *Contribution to the Mathematical Theory of Epidemics*,^{1,2} Kermack and McKendrick developed a general, if elaborate model structure to capture the dynamics of a fixed-size population comprising compartments of individuals susceptible (S) to a spreading infection, infectious (I), and removed (R) from the preceding two compartments by recovery or death. Propagation of an individual from S to I to R underlies the basic context of the exercise. In a modeling tour-de-force, the authors eventually present, in equations (11) through (15) of their paper (*ibid.*), the general structure of the elaborate mathematical model they derive. They proceed to examine the implications of their model for special cases, and finally present a very special case resulting in a set of three relatively simple ordinary differential equations (ODEs, equations (29) *ibid.*), which were destined to form the basis for a genre of mathematical models in epidemiology, the celebrated SIR model and its many variants.³ The three ODEs are

$$s' = -\beta si \quad (1)$$

$$i' = \beta si - \gamma i = \beta \left(s - \frac{1}{R_0} \right) i \quad (2)$$

$$r' = \gamma i \quad (3)$$

where prime denotes time derivative.^{2,4,5} Consistent with the importance of the SIR ODEs, the *basic reproductive ratio*, $R_0 \stackrel{\text{def}}{=} \beta/\gamma$, is widely considered “one of the most critical epidemiological parameters”² and has even become a household name in the recent coronavirus epidemic.⁶

In the sentence right before they present their SIR model in equation (29), (*ibid.*, cf. the above eqns. (1)-(3)) Kermack and McKendrick explain that this is a

special case in which ϕ and ψ are constants κ and l respectively

with (κ, l) referring to (β, γ) of eqns. (1)-(3), respectively.

The assumption about constant ϕ is plausible, as it refers to the rate of spread of the epidemic (cf. eqn. (1)). While that

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

taken to curb an epidemic, such a change could easily be reflected in the SIR model by a time-varying ϕ (cf. β in eqns. (1) and (2)).

The assumption about constant ψ , however, as fruitful as it may have proved, is chosen for mathematical convenience rather than for intent to describe the system as realistically as possible. Indeed, a clear definition of ψ is provided by the authors as follows (p. 703, *ibid.*):

If ψ_θ denotes the rate of removal, ..., then the number who are removed from each θ group at the end of the interval t is $\psi_\theta v_{t,\theta}$,

where (p. 702, *ibid.*)

$v_{t,\theta}$ shall denote the number of individuals in unit area at the time t who have been infected for θ intervals.

However, the rate of removal depends more on the duration over which individuals have remained infected and less on the size of that group. In the conceptually simple case where all individuals recover or die at a single number of days, D , after their infection, the rate of removal is 1 or 0, depending on whether θ is greater than D or not (Figure 1).

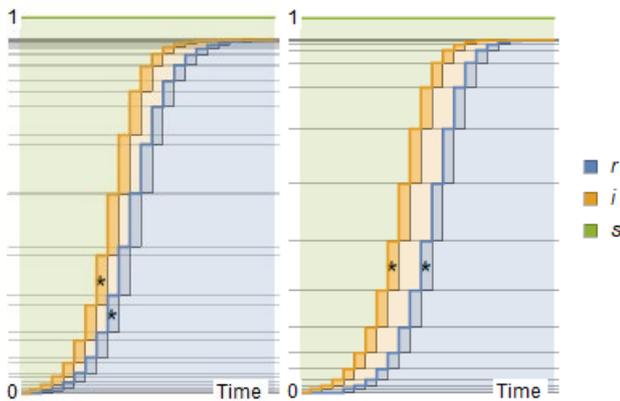


Figure 1. Stacked profiles of $\{s, i, r\}$ over discrete time in a constant-size population over time. Left: At each time step, r increases in proportion to the previous i . Right: At each time step, r increases by the decrease in s at a fixed number of steps before. Stars indicate correspondence in the i to r transition.

Of course, in reality D will likely follow a distribution (Figure 2) rather than being a single number. However, in that case as well, the removal rate will depend on comparison between θ and the distribution of D , rather than on the size of the group remaining infectious for time θ .

Starting with the assumption that individuals leave the infectious group at time D after infection, we develop in this paper a corresponding mathematical model structure, named delay SIR (dSIR), in the form of a single delay differential equation (DDE) for s , and two associated delay algebraic equations, for i and r in terms of s .

In the rest of the paper we first introduce the dSIR model structure in section 2, and provide an intuitive exposition of its basic properties in section 3, where we also introduce the Padé SIR model structure as an ODE approximation of dSIR. Rigorous analysis follows in section 4. In that section we explain that certain SIR and dSIR properties are exactly similar (e.g. herd immunity, total number of infected), while others are quite different (e.g. exponential rates, peak of the infectious group). Extension to models with additional compartments beyond S, I, and R is discussed in section 5, with presentation of the dSPIR model. Finally, the significance of this work and future extensions are discussed.

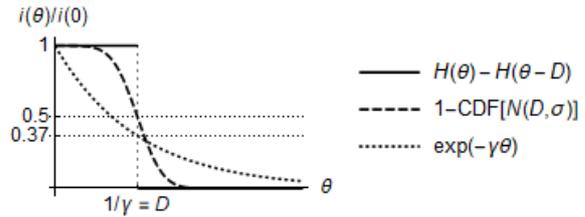


Figure 2. Reduction of the infectious fraction i from $i(0)$ to 0 (a) in a single front in terms of the Heaviside step function H with $D = 1/\gamma$, (b) in a diffused front in terms of the cumulative density function (CDF) of the normal distribution $N(D, \sigma)$, and (c) at rate $-\gamma i$, yielding exponential decline, eqn. (2).

2. The dSIR model structure

To explain the derivation of the dSIR model structure, we will rely on the detailed version of Figure 2 shown in Figure 3. The schematic shows the evolution of s_k, i_k, r_k over discrete time steps $k = 0, 1, 2, \dots$ of length δt each.

The thick-green bordered rectangle in Figure 3 suggests by visual inspection that

$$s_k - s_{k+1} = \delta t f(i_k, s_k) = \delta t \beta i_k s_k \quad (4)$$

where $f(r_k, s_k)$ is the rate at which the infection spreads, for example in proportion to the product $i_k s_k$, as shown above.

Assuming that each new part of the infectious fraction, i_k , gained at a given time step, k , moves to the removed fraction, r , in n time steps, the thick-red bordered rectangle in Figure 3 suggests that

$$i_k = s_{k-n} - s_k \quad (5)$$

Eqns. (4) and (5) yield

$$\begin{aligned} \frac{s_{k+1} - s_k}{\delta t} &= -f(s_{k-n} - s_k, s_k) \\ &= -\beta s_k (s_{k-n} - s_k) \Rightarrow \end{aligned} \quad (6)$$

$$\boxed{s' = -f(s_{-D} - s, s) = -\beta s (s_{-D} - s)} \quad (7)$$

where $s' \stackrel{\text{def}}{=} ds/dt$ and $s_{-D} \stackrel{\text{def}}{=} s(t - D)$.

Eqn. (7) is a single nonlinear DDE that involves only s and is decoupled from equations for i and r . As such, it captures the entire dynamics of the dSIR system.

The remaining two fractions of the population, i and r , can simply be inferred by algebraic equations, as eqn. (5) implies

$$\boxed{i = s_{-D} - s} \quad (8)$$

and $i + s + r = 1$ implies

$$\boxed{r = 1 - s_{-D}} \quad (9)$$

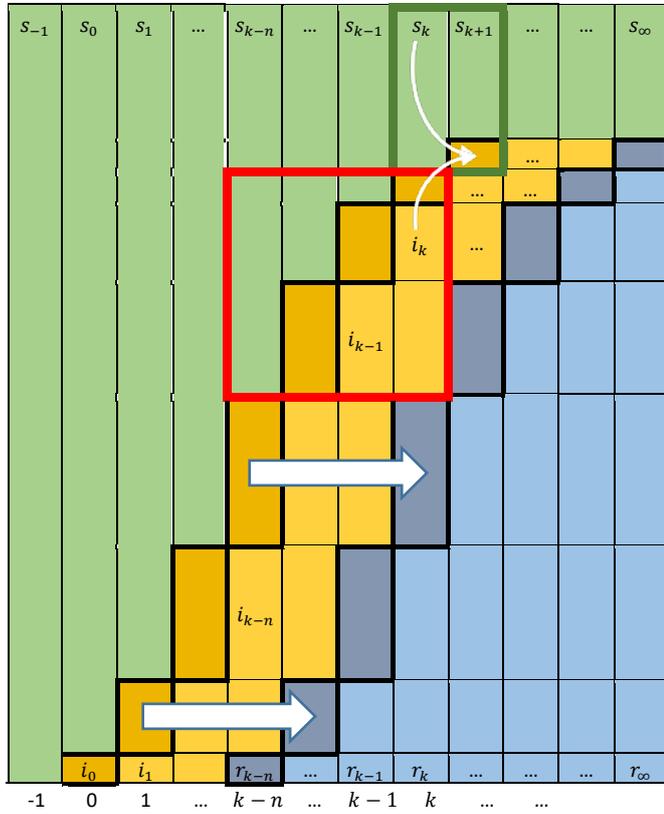


Figure 3. Schematic of evolving susceptible (green), infectious (orange), and removed (blue) fractions of a fixed-size population after an initial infection, i_0 . Each new part of the infectious fraction (thick-black bordered orange rectangles) moves to the removed fraction (thick-black bordered blue rectangles) in n time steps. The population eventually reaches a steady state at $s_\infty, r_\infty = 1 - s_\infty$, and $i_\infty = 0$.

To put the schematic in Figure 3 in context, observe that summation of eqns. (1)-(3) and discretization yields

$$r_{k+1} - r_k = \delta t \gamma i_k \quad (10)$$

This suggests that, according to the SIR model, each thick-black bordered green rectangle in Figure 3 would have to be a constant fraction $\delta t \gamma$ of the immediately previous orange column. However, with each new part of the infectious fraction moving to the removed fraction after n steps (thick-black bordered orange rectangles moving to thick-black bordered blue rectangles) this cannot be true as an emerging fact, by the simple observation of the shapes of r and $r + i$, which follow a convex to concave pattern after an inflection point. Therefore, the assumption of constant ψ , equivalent to constant γ in eqn. (10), is not compatible with the assumption

of time to transition from infectious to removed being independent of the infectious fraction size. This will be further illustrated with simulations in the next section.

Incidentally, eqns. (1)-(3) of the SIR model can also be decoupled to a single if non-intuitive nonlinear ODE in terms of s , solved in a form implicit in time, t .⁷

3. From SIR to dSIR and beyond: Visualization

Before any theoretical properties of the dSIR model structure are analyzed, a simple visual comparison between SIR and dSIR is presented. Unless specifically stated otherwise, all numerical simulations have been conducted with^{8,9}

$$D = \frac{1}{\gamma} = 8.4 \text{ days}, \quad \beta = \gamma R_0 \text{ days}^{-1} \quad (11)$$

Figure 4 is the standard $\{s, i, r\}$ plot, for SIR and dSIR. Note the faster dynamics of dSIR compared to SIR, the approximately twice as high peak of i for dSIR compared to SIR, as well as the asymmetric profile of i for SIR, compared to the symmetric dSIR profile of i . (Details in section 4.)

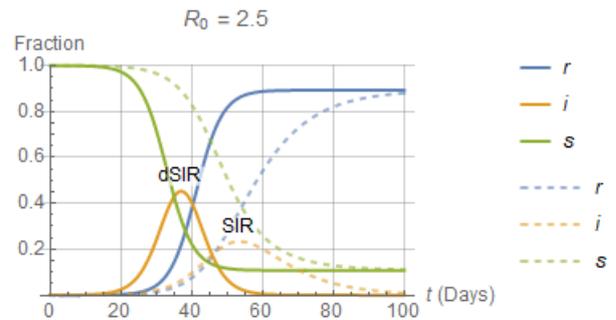


Figure 4. Profiles of s, i, r of a population through an epidemic according to the dSIR and SIR models, eqns. (7)-(9) and (1)-(3) respectively. The correspondence $\gamma = 1/D$ was used.

To further illustrate the dSIR/SIR relationship, Figure 5 and Figure 6 show the continuous-time counterparts of Figure 3 for the SIR and dSIR models, respectively. In this stacked representation of $\{s, i, r\}$ over time, it is clear that the SIR model corresponds to a time-varying infectious period for each individual. In fact, the value of γ calculated according to eqn. (3) as

$$\gamma = r' / i \quad (12)$$

with r' and i produced by the dSIR model, is time varying, as shown in Figure 7. This discrepancy suggests that the standard interpretation of $1/\gamma$ as

the average infectious period ... estimated relatively precisely from epidemiological data²

is increasingly inaccurate as R_0 increases above 1. What is estimated from epidemiological data is the delay D , rather than γ , and if $1/D$ is used as an estimate of γ , as is typically done, the SIR model response will be too slow, with a peak value for i lower than its dSIR counterpart by about half. Depending on

the R_0 value considered, this may significantly affect the SIR model value for “Flattening the Curve”, an aim that has become a household name in the latest coronavirus epidemic.¹⁰

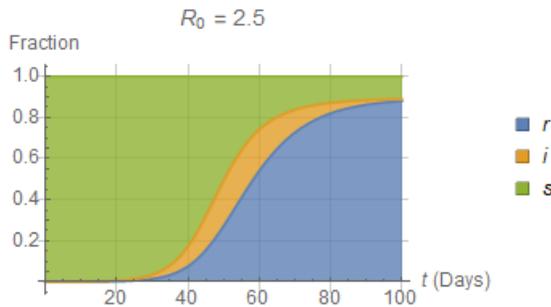


Figure 5. Stacked fractions $\{s, i, r\}$ of a population through an epidemic according to the SIR model, eqns. (1)-(3). Note that the duration from entrance to the infectious group until removal from it (horizontal slices in orange area) is variable.

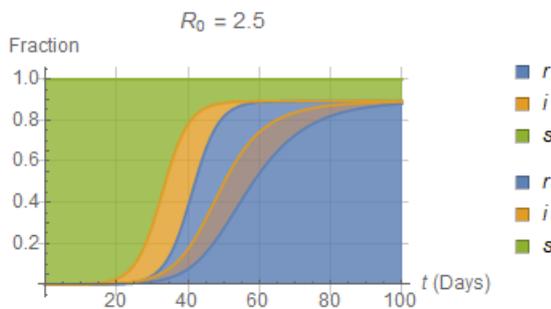


Figure 6. Stacked fractions $\{s, i, r\}$ of a population through an epidemic according to the dSIR model, eqns. (7)-(9). Note that D remains constant. Figure 5 is also superimposed, for comparison. The correspondence $\gamma = 1/D$ was used.

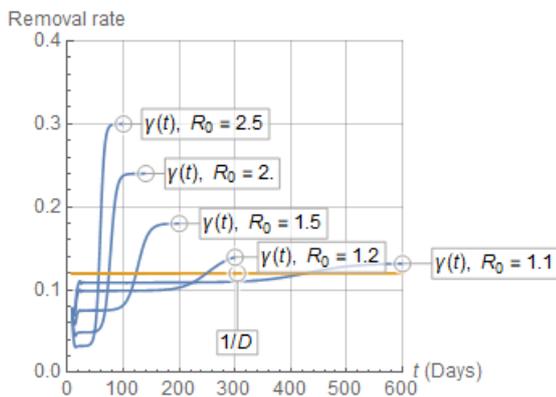


Figure 7. Comparison between $\gamma(t) \stackrel{\text{def}}{=} r'(t)/i(t)$ and the corresponding $1/D = 0.12 \text{ day}^{-1}$ for various R_0 in numerical integration of the dSIR eqns. (7)-(9). Note that all lines start at $t = D$, as $r'(t < D) = 0$. Note also small spurious deviations from constant shortly after $t = D$, due to numerical approximation of $i(t \leq 0)$, by the continuous function $(1 - \epsilon) \exp(\lambda t)$, $t \leq 0$ with $\lambda \gg 1$.

3.1 The Padé SIR model structure

As pointed out in the preceding sections, interpreting the value of γ in the SIR model as the average infectious period is not accurate and produces misleading results. It turns out (Appendix A) that the following simple remedy can be used to retain the ODE structure of the standard SIR model, while better approximating the DDE dynamics of the more realistic dSIR model structure: The SIR equations for $\{i', s'\}$, eqns. (2) and (3), can be replaced by the equally simple ODEs

$$\begin{cases} i' = 2 \left(\beta s - \frac{1}{D} \right) i \\ r' = \left(-\beta s + \frac{2}{D} \right) i \end{cases} \quad (13)$$

or by the more accurate set of ODEs

$$\begin{cases} i' = z_2, \quad z_2' = \frac{12}{D^2} \left(D\beta s i - i - \frac{D}{2} z_2 \right) \\ r' = -z_2 + \beta s i \end{cases} \quad (14)$$

with $z_2(0) = 0$.

Figure 8 confirms that the i trajectories for the dSIR and the modified SIR model structures – eqns. (7)-(9) and (1)-(3) with eqn. (2) replaced by (13) or (14), respectively – are close to one another, both in terms of the time to peak and the value of the peak. As already mentioned, these properties are important when such models are used to anticipate hospitalization needs for the infected during an epidemic.

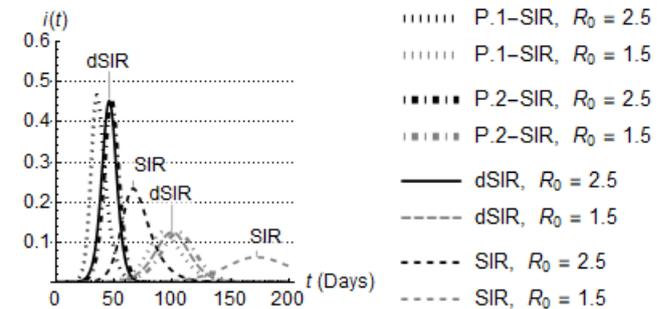


Figure 8. Comparison of the i profiles for the first-order Padé SIR, second-order Padé SIR, dSIR, and SIR models. Note the improve approximation of dSIR by the second-order Padé SIR, compared to first-order Padé SIR, as anticipated by eqns. (13) or (14), respectively.

Note (Appendix A) that the essence of eqns. (13) and (14) is in approximating the pulse profile $H(\theta) - H(\theta - D)$ in Figure 2 by a transfer function approximation based on first- or second-order Padé-approximants (in the Laplace domain)

$$\frac{i(\theta)}{i(0)} \approx \mathcal{L}^{-1} \left[\frac{D}{1+qD/2} \right] = 2e^{-\frac{2\theta}{D}} \quad (15)$$

$$\frac{i(\theta)}{i(0)} \approx \mathcal{L}^{-1} \left[\frac{D}{1+qD/2+q^2D^2/12} \right] = 4\sqrt{3}e^{-\frac{3\theta}{D}} \sin\left(\frac{\sqrt{3}\theta}{D}\right) \quad (16)$$

(rather than by the decaying exponential of Figure 2) as shown in Figure 9. Padé-approximation has long been a popular

approach for approximating transcendental transfer functions by polynomial rational fractions in process control.¹¹

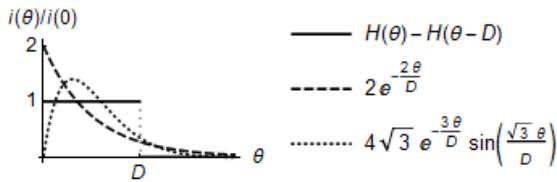


Figure 9. Reduction of the infectious fraction i from $i(0)$ to 0 (a) in a single front in terms of the Heaviside step function H with $D = 1/\gamma$, (b) following a first-order Padé approximation, and (c) following a second-order Padé approximation. (cf. Figure 2).

Because of the critical role of Padé approximation in deriving eqn. (14) for the ODEs of the modified SIR model to approximate the dSIR model, we will use the term Padé SIR to denote the modified SIR model structure.

Finally, it should be noted that eqn. (13) may seem counter-intuitive, as it appears to suggest that the generation and depletion rates of i are $2\beta si$ and $-2i/D$, respectively. However, eqn. (13) rather suggests that while the susceptible depletion rate remains $-\beta si$, the infectious depletion rate appears as $(\beta s - 2/D)i$ (which is s -dependent) rather than $-i/D$ (which is not) due simply to replacement of the delay term e^{-qD} in the Laplace domain by its Padé approximant. Similarly, the rate of increase for r in eqn. (13) is s -dependent, rather than not.

4. Dynamics of the dSIR and Padé SIR models

Standard theory of DDEs (e.g. equation (9.42) in Ch. 4 and equation (1.2) in Ch. 5 of Kuang,¹² or Gopalsamy¹³) can be applied to establish rigorous properties for the dSIR model, such as global stability, convergence to a final steady state, and others. A complete analysis is beyond the scope of this paper. However, some important theoretical properties of the dSIR model structure of practical interest are discussed next, particularly in comparison to their SIR counterparts.

The analysis of the Padé SIR model follows standard ODE analysis and is presented more briefly, except when it has important implications for either theoretical or practical issues.

4.1 Stability at equilibrium and epidemic outbreak

Eqn. (7) can be used to show (Appendix B) that an equilibrium point \bar{s} is stable and an epidemic outbreak does *not* occur iff

$$\bar{s} < \frac{1}{\beta D} \stackrel{\text{def}}{=} \frac{1}{R_0} \quad (17)$$

Note that the stability upper bound $1/(\beta D)$ for the dSIR model in the above eqn. (17) coincides with the well known bound $\gamma/\beta = 1/R_0$ dictated by the SIR model under the widely used correspondence

$$D = 1/\gamma \quad (18)$$

The same stability bound can be derived for the Padé SIR model structure using standard ODE analysis based on linearization around an equilibrium point.

4.2 Final values of $\{s, i, r\}$

Accepting for now without proof the global stability of eqn. (7) and existence of $\{s_\infty, i_\infty = 0, r_\infty\}$ for initial values $i_0 \approx \epsilon \approx 0$, $s_0 = 1 - \epsilon \approx 1$, one can show (Appendix C) that

$$\frac{1}{\beta D} \ln s_\infty - s_\infty + 1 = 0 \Leftrightarrow s_\infty = \frac{-W[-R_0 \exp(-R_0)]}{R_0} \quad (19)$$

$$r_\infty = 1 - s_\infty \quad (20)$$

where $R_0 \stackrel{\text{def}}{=} \beta D$ and W is the Lambert function^{14,15} of order 0.

The standard plot for r_∞ , equal to the total fraction of infected by the end of an epidemic,² is shown in Figure 10 for completeness. Note that the plot is valid for both SIR and dSIR models, with $R_0 = \beta/\gamma$ and $R_0 = 1/D$, respectively.

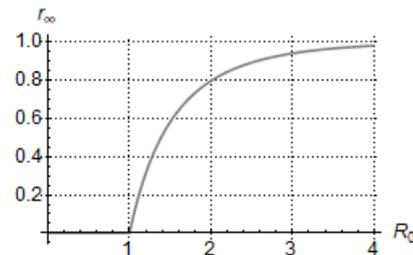


Figure 10. Total fraction of a population infected by the end of an epidemic, r_∞ , as a function of the basic reproductive ratio $R_0 = \beta/\gamma = 1/D$ according to eqn. (20), similar for both the SIR and dSIR model structures. For $R_0 < 1$ the epidemic is contained.

Interestingly, while use of the Lambert function to solve problems such as the above was pointed out as early as 1996,¹⁴ it may have escaped the attention of most literature in this field.² The Lambert function in its various forms will turn up in a number of results below.

It should also be noted that eqns. (19) and (20) are the same for the SIR and Padé SIR model (Appendix C) under the correspondence between D and γ in eqn. (18).

4.3 Exponential rates of ascent

For the initial part of a spreading epidemic, starting from a perturbation of the steady state $(s, i, r) = (\bar{s}, 0, 1 - \bar{s})$ as

$$\begin{aligned} i(t < 0) &= 0, & i(0) &= \epsilon, \\ s(t < 0) &= \bar{s}, & s(0) &= \bar{s} - \epsilon \\ r(t < 0) &= 1 - \bar{s}, & r(0) &= 1 - \bar{s} \end{aligned}$$

with

$$\bar{s}\beta D \stackrel{\text{def}}{=} \bar{s}R_0 > 1$$

it can be shown (Appendix D) that the infectious fraction initially grows approximately at an exponential rate as

$$\Delta i \approx \Delta i_0(a + e^{p_0 t} b) \quad (21)$$

where $\Delta s \stackrel{\text{def}}{=} s - \bar{s}$ and the constants a, b are in terms of $R_0 \stackrel{\text{def}}{=} \beta D$ (Appendix D) with

$$p_0 = \frac{\bar{s} R_0}{D} \left(1 + \frac{W[-(\bar{s} R_0) e^{-\bar{s} R_0}]}{\bar{s} R_0} \right) \quad (22)$$

as shown in Figure 11.

Similarly, the initial exponential rate of the Padé SIR model is given by

$$\Delta i = \Delta i_0 \exp\left(\frac{2\gamma(R_0 \bar{s} - 1)t}{p_{0, \text{PadéSIR}}}\right) \quad (23)$$

where $R_0 = \beta/\gamma$, and for the standard SIR model, the exponential rate is

$$\Delta i = \Delta i_0 \exp\left(\frac{\gamma(R_0 \bar{s} - 1)t}{p_{0, \text{SIR}}}\right) \quad (24)$$

where $R_0 = \beta/\gamma$, shown in Figure 11 as well.

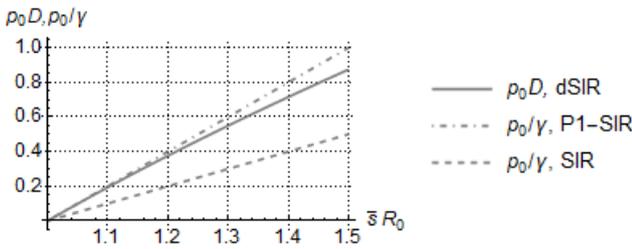


Figure 11. Dimensionless exponential rates $\{p_0 D, p_0/\gamma\}$ for the dSIR, first-order Padé SIR, and SIR models, by eqns. (22), (23), and (24), respectively. The first two coincide fairly well as $\bar{s}R$ approaches 1, whereas the SIR rate remains about half of the other two.

This has immediate implications for the early rate of rise of the infectious fraction to its peak, i^* , as illustrated in Figure 12.

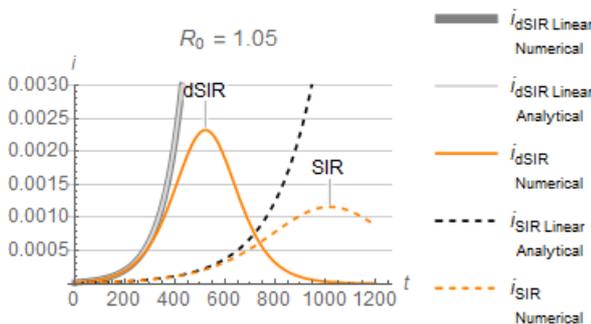


Figure 12. Exponential rate of increase at the early stages of an epidemic for the dSIR and SIR model structures.

Note that while the initial SIR rate is half of the Padé SIR rate and about half of the dSIR rate, all three models eventually reach the exact same steady-state values, as captured by eqns. (19) and (20).

4.4 Peak of infectious fraction

While it is not obvious to the author whether the peak i^* can be easily obtained for the dSIR model, a good approximation can be obtained (Appendix E) through the Padé SIR model, as

$$s^* = \frac{1}{\beta D}, \quad i^* = 2 \left(\frac{-1}{\beta D} \ln(\beta D) - \frac{1}{\beta D} + 1 \right) \quad (25)$$

Note that the above i^* , exact for the first-order Padé SIR model and approximate for the second-order Padé SIR and dSIR models, is double the i^* of the standard SIR model, as confirmed in Figure 8. Once more, there are obvious practical implications from this discrepancy.

Note also that in case an upper bound is placed on i^* , to avoid overwhelming hospitalization facilities during an epidemic, eqn. (25) has an explicit solution for the corresponding maximum $R_0 \stackrel{\text{def}}{=} \beta D$ as

$$\beta D \stackrel{\text{def}}{=} R_0 = \frac{2W_{-1}\left(\frac{-2+i^*}{2e}\right)}{-2+i^*} \quad (26)$$

where $W_{-1}\left(\frac{-2+i^*}{2e}\right)$ is the Lambert function of order -1 .

For comparison, the standard SIR model yields

$$\frac{\beta}{\gamma} \stackrel{\text{def}}{=} R_0 = \frac{W_{-1}\left(\frac{-1+i^*}{e}\right)}{-1+i^*} \quad (27)$$

The values of R_0 indicated by eqns. (26) and (27), with corresponding definitions, are shown in Figure 13. It is evident that the Padé SIR model places twice as tight a restriction on R_0 as the standard SIR model, if i is not to exceed the i^* value.

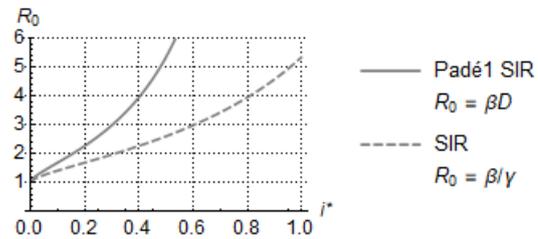


Figure 13. Maximum value of R_0 indicated by the Padé SIR and SIR models for i not to exceed i^* .

5. dSPIR, Padé SPIR, and variants

The dSIR model structure developed in Figure 3 can be easily extended to include additional compartments. In fact, practically all population models of infections developed to date using the concept of exchange between compartments³ can be immediately translated (a) from ODEs to DDEs through replacement of compartment drain rates proportional to the drained quantity by drainage of amounts that have resided for a certain time, D , in that compartment, or (b) from ODEs to Padé approximations that maintain the ODE structure but are more realistic. We illustrate these ideas on

the SPIR model structure, in view of its importance for the recent coronavirus epidemic.¹⁶⁻¹⁸

A distinct feature of the coronavirus causing COVID-19 is that it enables infection transmission at the pre-symptomatic stage.¹⁹ Therefore, the SPIR model structure (Figure 14) comprises the usual population fractions $\{s, i, r\}$ along with the pre-symptomatic infectious fraction, p , with symptomatic infectious being i . SPIR differs from the standard SEIR structure^{2,4,5,20} by the way the four compartments interact.

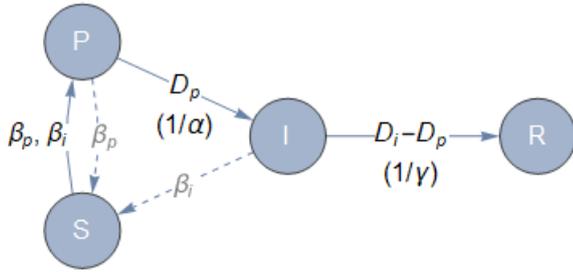


Figure 14. The SPIR and dSPIR model structure, with members of the pre-symptomatic infectious compartment, P, moving, after time D_p , to the symptomatic infectious compartment, I, and from there, after time $D_i - D_p$, to the removed compartment, R. Both P and I infect the susceptible group, S, at rates, β_p and β_i , respectively.

While it is formidable to practically monitor infections on pre-symptomatic infectious individuals, monitoring symptomatic infectious is more reasonable, as symptoms are clear and can be confirmed by testing. Therefore, tighter restrictions can be placed on the I group, in addition to the P group typically following general restrictions placed on the general population to curb the spread of the epidemic, as captured by the spread factors β_i and β_p , respectively, in Figure 14.

The dynamics of the dSPIR structure is shown in Figure 15, which follows the pattern of Figure 3, with the addition of the pre-symptomatic infectious fraction p .

Following the same technique as in section 2, we can immediately write the following equations for the dSPIR model structure by visual inspection of Figure 15:

$$s_{k+1} = s_k - \delta t s_k f(i_k, p_k, s_k) \quad (28)$$

$$s_k = s_{k-n_1} - p_k \quad (29)$$

$$s_k = s_{k-n_2} - p_k - i_k \quad (30)$$

Combining the above equations and letting $\delta t \rightarrow \infty$ yields the nonlinear DDE

$$s' = -f\left(\underbrace{s_{-D_i} - s_{-D_p}}_i, \underbrace{s_{-D_p} - s}_p, s\right) \quad (31)$$

and associated delay algebraic equations

$$p = s_{-D_p} - s \quad (32)$$

$$i = s_{-D_i} - s_{-D_p} \quad (33)$$

$$r = 1 - s_{-D_i} \quad (34)$$

where D_p and $D_i - D_p$ are the durations of an individual's staying in the P and I compartment, respectively, and typically

$$f(i, p, s) = \beta_i s i + \beta_p s p \quad (35)$$

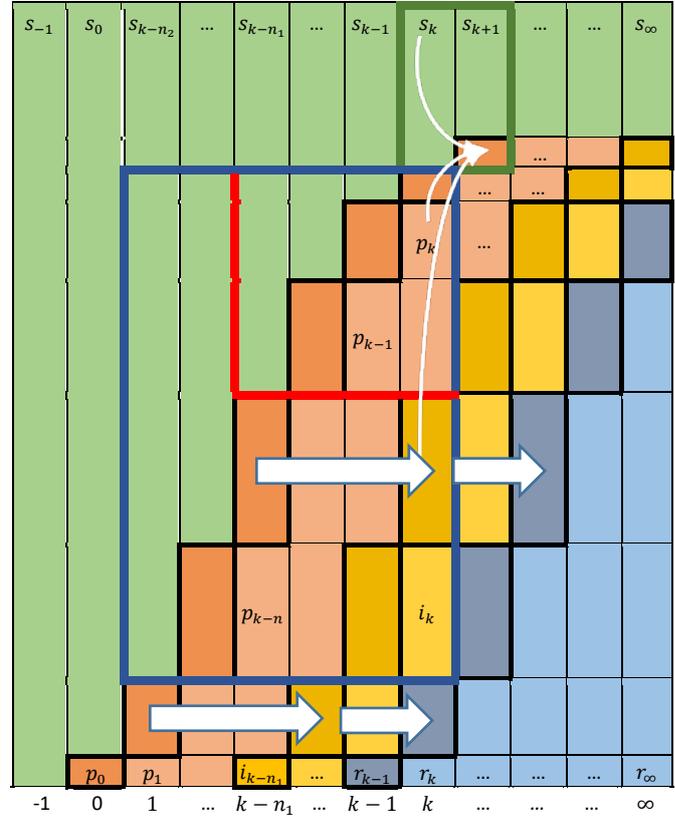


Figure 15. Schematic of evolving susceptible (green), pre-symptomatic infectious (red), symptomatic infectious (orange), and removed (blue) fractions of a fixed-size population after an initial infection, p_0 . Each new part of a fraction (thick-black bordered rectangles) moves to the next fraction (thick-black bordered rectangles to the right) in the same number of time steps. The population eventually reaches a steady state at $s_\infty, r_\infty = 1 - s_\infty, i_\infty = p_\infty = 0$.

Note that the form of the SPIR model following the standard SIR pattern, eqns. (1)-(3), is the coupled ODEs²¹

$$s' = -\beta_i s i - \beta_p s p \quad (36)$$

$$p' = \beta_i s i + \beta_p s p - \alpha p \quad (37)$$

$$i' = \alpha p - \gamma i \quad (38)$$

$$r' = \gamma i \quad (39)$$

In the form of these ODEs is the following more realistic Padé SPIR model, which approximates the dSPIR DDEs, eqns. (31)-(35), significantly better than the SIR ODEs (Appendix F):

$$s' = -f(i, p, s) \quad (40)$$

$$p' = 2 \left(f(i, p, s) - \frac{1}{D_p} p \right) \quad (41)$$

$$i' = 2 \left(\left(\frac{1}{D_p} - \frac{1}{D_i} \right) p - \frac{1}{D_i} i \right) \quad (42)$$

$$r = 1 - s - p - i \quad (43)$$

where

$$D_p = \frac{1}{\alpha}, \quad D_i = \frac{1}{\alpha} + \frac{1}{\gamma} \quad (44)$$

The illustration and analysis presented in sections 3 and 4 can be easily repeated for the dSPIR and Padé SPIR structures. To maintain the scope of this publication, only a few properties will be explored below. The rest will be explored in more detail in forthcoming publications.

5.1 From SPIR to dSPIR: Visualization

We present here only a few simulations comparing the $\{s, p, i, r\}$ profiles resulting from numerical solution of the SPIR, dSPIR, and Padé SPIR models. The values

$$D_p = \frac{1}{\alpha} = 5.1 \text{ days}, \quad D_i - D_p = \frac{1}{\gamma} = 3.3 \text{ days} \quad (45)$$

are used in all simulations.^{8,9}

Figure 16 illustrates the differences in the infectious peaks, i^* and p^* , similar to these in Figure 8.

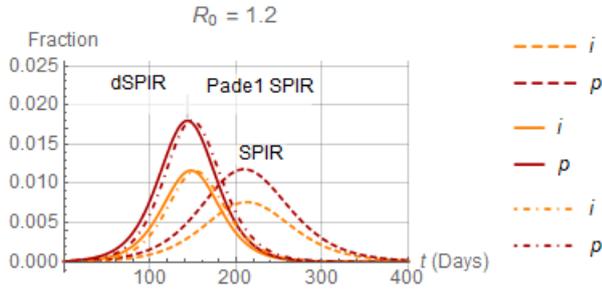


Figure 16. Comparison of the infectious peaks, i^* and p^* , among the SPIR (dashed), dSPIR (continuous) and Padé SPIR (Dot-dashed) model structures. Note the proximity between dSPIR and Padé SPIR.

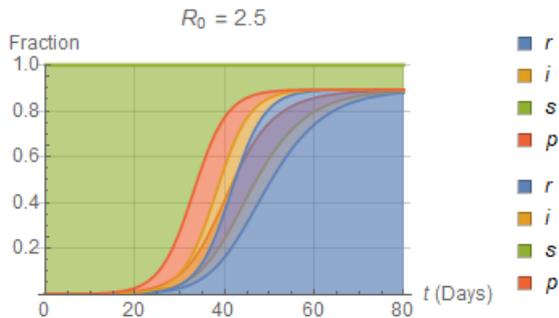


Figure 17. Comparison of stacked $\{s, p, i, r\}$ distribution for the SPIR and dSPIR model structures.

5.2 Stability at equilibrium and epidemic outbreak

Following the same approach as in section 4.1, eqn. (31) can be used to show (Appendix G) that an epidemic outbreak does not occur around an equilibrium point \bar{s} iff

$$\bar{s} < \frac{1}{\beta_i(D_i - D_p) + \beta_p D_p} \stackrel{\text{def}}{=} \frac{1}{R_0} \quad (46)$$

The proof in Appendix G is by approximation. It is conjectured that the bound in the above equation is exact. This conjecture will be examined in subsequent studies.

5.3 Final values of $\{s, p, i, r\}$

It can be shown (Appendix H) that the dSPIR counterpart of eqn. (19) is

$$\frac{1}{\beta_i(D_i - D_p) + \beta_p D_p} \ln s_\infty - s_\infty + 1 = 0 \Leftrightarrow s_\infty = \frac{-W[-R_0 \exp(-R_0)]}{R_0} \quad (47)$$

where $D_i > D_p$ and

$$R_0 \stackrel{\text{def}}{=} \beta_i(D_i - D_p) + \beta_p D_p \quad (48)$$

The counterpart of the SIR and dSIR plot (Figure 10) for r_∞ as a function of R_0 , eqn. (20), obviously remains the same.

Note that the choice of β_p, β_i has a significant effect on both the stability of the equilibrium point \bar{s} , eqn. (46), and on the value of $r_\infty = 1 - s_\infty$, eqn. (47), with obvious implications for distancing measures.²¹ For

$$\phi \stackrel{\text{def}}{=} \frac{\beta_i}{\beta_p}, \quad \lambda \stackrel{\text{def}}{=} \frac{D_p}{D_i} \quad (49)$$

with $0 \leq \phi, \lambda \leq 1$, eqn. (48) yields

$$R_0(\lambda, \phi) = \beta_p D_i (\phi + \lambda - \lambda \phi), \quad R_0(1, 1) = \beta_p D_i \quad (50)$$

as shown in Figure 18.

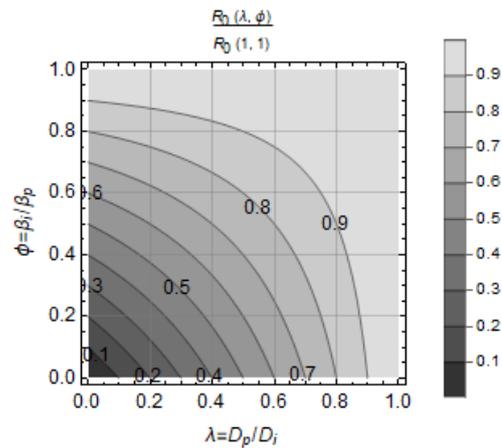


Figure 18. Contour plot of $R_0(\lambda, \phi)/R_0(1, 1)$, eqn. (50), for the range of additional restrictions, ϕ , placed on symptomatic infectious, and for the possible range of the pre-symptomatic infectious period, D_p , as a fraction, λ , of the total infectious period, D_i .

6. Discussion

A subtle inconsistency in the standard SIR model structure was pointed out. This inconsistency arises from misinterpretation of an assumption explicitly articulated in the original publication of Kermack and McKendrick:¹ that the depletion of the infectious compartment is proportional to the content of that compartment. The depletion rate constant, γ , is usually interpreted as equal to the inverse of the residence time in the infectious compartment, namely the duration of the infection, D , for each individual (eqn. (18)). To the extent that this duration is about constant, the analysis presented here suggests that the preceding interpretation is incorrect, leading to certain erroneous conclusions.

A corresponding model structure, termed dSIR, was developed, to account for each individual leaving the infectious compartment after a certain duration. The dSIR model structure comprises a single DDE for the susceptible fraction, s , and associated algebraic equations capturing the dependence of the remaining population fractions on s . While both SIR and dSIR produce the same results for assessment of stability and final values, the SIR model produces a maximum of the infectious fraction, i^* , about half of its dSIR counterpart. This has profound consequences if the SIR model is used to predict i^* during an epidemic. It is also noted that even if the SIR model parameters β and γ are estimated based on experimental data fit – albeit under the wrong interpretation – model predictions are still going to be inaccurate. This is because the standard SIR model, comprising the three ODEs in eqns. (1)-(3), is structurally different from the DDE form of the dSIR model, eqns. (7)-(9), or even from the ODE form of the Padé-based approximation of the dSIR model DDEs, eqns. (1) and (13) or (1) and (14).

The dSIR structure can be easily extended to other compartment-based population models.³ Such an extension to the dSPIR model structure was presented and briefly illustrated and analyzed. This model is important for infections transmitted by both pre-symptomatic and symptomatic infected individuals. Similarities and differences between SPIR and dSPIR models are of the same nature as between SIR and dSIR models.

Numerous additional issues related to this work can be considered, including the following: Rigorous analysis of DDE models; DDE modeling and analysis for a distribution rather than uniform delay (cf. Figure 2); resolution of conjectures presented in the text; and implications for different forms of infection kinetics. Such issues will be addressed in forthcoming publications.

7. Acknowledgements

All computations were done in *Mathematica*, available at the University of Houston. Sharing of teaching material about the SIR model on Github by Prof. Jeff Kantor of Notre Dame is also gratefully acknowledged. Research reported in this

publication was partially supported by the Institute of Allergy and Infectious Diseases of the National Institutes of Health under award number R01AI140287, financed with Federal money. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

8. References

- 1 Kermack, W. O. & McKendrick, A. G. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London Series A* **115**, 700 (1927).
- 2 Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals*. (Princeton University Press, 2008).
- 3 Anderson, R. M. & May, R. M. Population biology of infectious diseases: Part I. *Nature* **280**, 361-367, doi:10.1038/280361a0 (1979).
- 4 Murray, J. D. *Mathematical Biology: I. An Introduction*. (Springer, 2002).
- 5 Anderson, R. M., Anderson, B. & May, R. M. *Infectious diseases of humans: dynamics and control*. (Oxford University Press, 1992).
- 6 Wang, C. *et al.* Evolving Epidemiology and Impact of Non-pharmaceutical Interventions on the Outbreak of Coronavirus Disease 2019 in Wuhan, China. *medRxiv*, 2020.2003.2003.20030593, doi:10.1101/2020.03.03.20030593 (2020).
- 7 Harko, T., Lobo, F. S. N. & Mak, M. K. Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Applied Mathematics and Computation* **236**, 184-194, doi:10.1016/j.amc.2014.03.030 (2014).
- 8 Lauer, S. A. *et al.* The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*, doi:10.7326/M20-0504 (2020).
- 9 Boldog, P. *et al.* Risk Assessment of Novel Coronavirus COVID-19 Outbreaks Outside China. *Journal of Clinical Medicine* **9** (2020).
- 10 Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, doi:10.1016/S1473-3099(20)30120-1.
- 11 Stephanopoulos, G. *Chemical Process Control: An Introduction to Theory and Practice*. (Prentice Hall, 1984).
- 12 Kuang, Y. *Delay Differential Equations with Applications in Population Dynamics*. (Academic Press, 1993).
- 13 Gopalsamy, K. *Stability and Oscillations in Delay Differential Equations of Population Dynamics*. (Springer Netherlands, 1992).
- 14 Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J. & Knuth, D. E. On the LambertW function. *Advances in Computational Mathematics* **5**, 329-359, doi:10.1007/BF02124750 (1996).

- 15 Kesisoglou, I., Singh, G. & Nikolaou, M. The Lambert function belongs in the Function Hall of Fame. *Chemical Engineering Education In preparation* (2020).
- 16 Ferguson, N. M. *et al.* Report 9 - Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. (Imperial College, <https://doi.org/10.25561/77482>, 2020).
- 17 Institute for Health Metrics and Evaluation (IHME). *COVID-19 projections assuming full social distancing through May 2020*, <COVID-19 projections assuming full social distancing through May 2020> (2020).
- 18 Kucharski, A. J. *et al.* Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases*, doi:10.1016/S1473-3099(20)30144-4.
- 19 Zhang, J. *et al.* Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *The Lancet Infectious Diseases*, doi:10.1016/S1473-3099(20)30230-9.
- 20 Blackwood, J. C. & Childs, L. M. An introduction to compartmental modeling for the budding infectious disease modeler. *Letters in Biomathematics* **5**, 195-221, doi:10.1080/23737867.2018.1509026 (2018).
- 21 Nikolaou, M. Using Feedback on Symptomatic Infections to Contain the Coronavirus Epidemic: Insight from a SPIR Model. *medRxiv*, 2020.2004.2014.20065698, doi:10.1101/2020.04.14.20065698 (2020).

Appendix A. Proof of eqns. (13)-(16)

Define the variable

$$u = s - \bar{s} \quad (51)$$

which trivially satisfies $u(t) = 0$ for $t < 0$, for simple application of the Translation Theorem of Laplace transforms. Then, take Laplace transforms, \mathcal{L} , of eqn. (8) with $I(q) \stackrel{\text{def}}{=} \mathcal{L}[i(t)]$, $U(q) \stackrel{\text{def}}{=} \mathcal{L}[u(t)]$, to get

$$I = (e^{-qD} - 1)U \quad (52)$$

Using the second-order Padé approximation

$$e^{-qD} \approx \frac{1 - \frac{D}{2}q + \frac{D^2}{12}q^2}{1 + \frac{D}{2}q + \frac{D^2}{12}q^2} \quad (53)$$

in eqn. (52) yields

$$I = \frac{-Dq}{1 + \frac{D}{2}q + \frac{D^2}{12}q^2} U \Rightarrow \left(1 + \frac{D}{2}q + \frac{D^2}{12}q^2\right) I = -DqU \Rightarrow$$

$$i + \frac{D}{2}i' + \frac{D^2}{12}i'' = -Du' = -Ds' = D\beta si \quad (54)$$

Introducing the new variables $z_1 \stackrel{\text{def}}{=} i$, $z_2 \stackrel{\text{def}}{=} i'$ immediately yields eqn. (14).

Eqn. (13) can be proved similarly.

To prove eqn. (16) observe that if $s(0)$ changes from its previous value of 1 by a step, $-\epsilon$, then $U(q) = -\epsilon/q$. Because $r(t \leq 0) = 0$ and $s + i + r = 1$, it follows that i jumps from $i(t < 0) = 0$ to $i(0) = \epsilon$. Therefore, the above eqn. (54) yields

$$\left(1 + \frac{D}{2}q + \frac{D^2}{12}q^2\right) I = D\epsilon \quad (55)$$

which immediately implies eqn. (16).

Eqn. (15) can be proved similarly.

Appendix B. Proof of eqn. (17)

Eqns. (7)-(9) imply that $(\bar{s}, 0, 1 - \bar{s})$ is an equilibrium point, where \bar{s} can be arbitrary between 0 and 1. Then, local stability analysis of eqn. (7) by approximate linearization of $g(s_{-D}, s) = f(s, i)$ around the steady state $s_{-D} = s = \bar{s}$ yields

$$\Delta s' \approx -\left(\frac{\partial g}{\partial s}\right)_{\bar{s}} \Delta s - \left(\frac{\partial g}{\partial s_{-D}}\right)_{\bar{s}} \Delta s_{-D}$$

$$= \bar{s}\beta \Delta s - \bar{s}\beta \Delta s_{-D} \quad (56)$$

where $\Delta s \stackrel{\text{def}}{=} s - \bar{s}$, with $\Delta s(t < 0) = 0$, $\Delta s(0) = \Delta s_0$.

Taking Laplace transforms with $\Delta S(q) \stackrel{\text{def}}{=} \mathcal{L}[\Delta s(t)]$ yields $q\Delta S - \Delta s_0 = \bar{s}\beta \Delta S - \bar{s}\beta e^{-qD} \Delta S \Rightarrow$

$$\Delta S = \frac{1}{q - \bar{s}\beta + e^{-qD}\bar{s}\beta} \Delta s_0 \quad (57)$$

The roots p_m of the transcendental characteristic equation

$$q - \bar{s}\beta + e^{-qD}\bar{s}\beta = 0 \quad (58)$$

(poles of ΔS) can be obtained in terms of the Lambert function,^{14,15} W , as follows: The last equation implies

$$e^{qD - \bar{s}\beta D} (qD - \bar{s}\beta D) = -\bar{s}\beta D e^{-\bar{s}\beta D} \Rightarrow$$

$$p_m D - \bar{s}\beta D = W_m[-(\bar{s}\beta D) e^{-\bar{s}\beta D}] \Rightarrow$$

$$p_m = \bar{s}\beta \left(1 + \frac{W_m[-(\bar{s}\beta D)e^{-\bar{s}\beta D}]}{\bar{s}\beta D} \right) \quad (59)$$

For stability, all p_m must be in the left-half of the complex plane.

We show first that for any positive value of \bar{s} , no complex root $p_m = \lambda + j\omega$ can cross the imaginary axis to move from stability to instability. Because, if $p_m = j\omega$ were a root of eqn. (58) for $\bar{s} > 0$, it would be

$$\begin{aligned} j\omega - \bar{s}\beta + e^{-j\omega D} \bar{s}\beta &= 0 \\ \Rightarrow \begin{cases} -\bar{s}\beta + \bar{s}\beta \cos(\omega D) = 0 \Rightarrow \omega D = 2m\pi \\ \omega - \bar{s}\beta \sin(\omega D) = 0 \Rightarrow \omega = 0 \end{cases} \end{aligned} \quad (60)$$

Therefore, only real roots should be considered in eqn. (59) for stability analysis. Furthermore, since

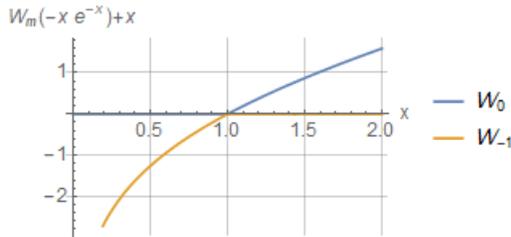
$$-e^{-1} \leq -e^{-\bar{s}\beta D} \bar{s}\beta D < 0 \quad (61)$$

the relevant values of m are 0, -1 in eqn. (59) for $p_m \in \mathbb{R}$, and $p_m < 0$ implies

$$W_m(-xe^{-x}) + x < 0, \quad m = 0, -1 \quad (62)$$

$$x \stackrel{\text{def}}{=} \bar{s}\beta D$$

which is satisfied for $x < 1$ (see figure below) leading immediately to eqn. (17).



Appendix C. Proof of eqn. (19)

Dividing eqn. (7) by s yields

$$(\ln s)' = -\beta(s_{-D} - s) \quad (63)$$

Using eqn. (51) yields $s_{-D} - s = u_{-D} - u$. Then taking Laplace transforms, \mathcal{L} , of both sides of eqn. (63) yields

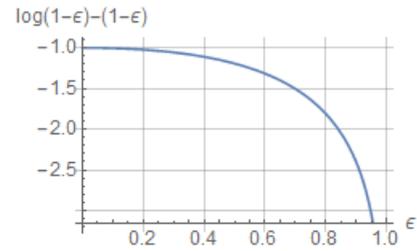
$$\begin{aligned} \mathcal{L}[(\ln s)'] &= \beta(1 - e^{-qD})\mathcal{L}[u] \Rightarrow \\ q\mathcal{L}[\ln s] - \ln s_0 &= \beta \frac{(1 - e^{-qD})}{q} q\mathcal{L}[u] \Rightarrow \\ \lim_{q \rightarrow 0} q\mathcal{L}[\ln s] - \underbrace{\ln s_0}_{\approx 0} &= \beta \lim_{q \rightarrow 0} \frac{(1 - e^{-qD})}{q} \lim_{q \rightarrow 0} q\mathcal{L}[u] \Rightarrow \\ \ln s_\infty - \ln s_0 &= \beta D u_\infty \end{aligned}$$

by the Final Value Theorem. Therefore

$$\frac{1}{\beta D} \ln s_\infty - s_\infty = \ln s_0 - s_0 \approx -1 \quad (64)$$

for “small” values of ϵ , as shown in the following figure.

(Recall that $s_0 > \frac{1}{\beta D} \Leftrightarrow \epsilon < 1 - \frac{1}{\beta D} = \frac{1}{R_0}$ for the epidemic to spread.)



Continuing on the last equation we get

$$\begin{aligned} (-\beta D s_\infty) \exp(-\beta D s_\infty) &= (-\beta D) \exp(-\beta D) \Rightarrow \\ -\beta D s_\infty &= W[-(\beta D) \exp(-\beta D)] \end{aligned}$$

which is eqn. (19).

The Padé SIR model also reaches the same result:

Dividing eqn. (1) by eqn. (13) and rearranging yields

$$\int_{s_0}^s \frac{\beta s - \gamma}{\beta s} ds = -2\beta \int_{i_0}^i di \Rightarrow s - s_0 - \frac{\gamma}{\beta} \ln\left(\frac{s}{s_0}\right) = 2(i_0 - i)$$

Taking the limit as $t \rightarrow \infty$ with $1/\gamma = D$, $s_0 \approx 1$, and $i_0 \approx 0$ yields eqn. (19).

Proof that the standard SIR model also reaches the same result follows the same pattern and is omitted for brevity.

Appendix D. Proof of eqns. (21), (22), and (24)

Using the Residue Theorem for Laplace Transforms, eqn. (57) implies

$$\begin{aligned} \Delta s &= \mathcal{L}^{-1}[\Delta S] = \\ &= \Delta s_0 \sum_m e^{p_m t} \lim_{q \rightarrow p_m} \left[\frac{q - p_m}{q - \bar{s}\beta + e^{-qD} \bar{s}\beta} \right] \end{aligned} \quad (65)$$

where p_m are the poles of ΔS , as shown in eqn. (59).

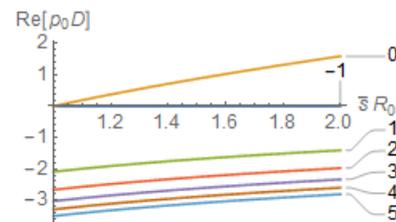
With $\bar{s}\beta D > 1$ (spreading epidemic) the argument $-(\bar{s}\beta D)e^{-\bar{s}\beta D}$ of W_m in eqn. (59) satisfies the inequality in eqn. (61). Consequently, based on the properties of the Lambert function, the summation in eqn. (65) contains two terms with real poles, for $m = 0, -1$, with values

$$p_{-1} = \bar{s}\beta \left(1 + \frac{W_{-1}[-(\bar{s}\beta D)e^{-\bar{s}\beta D}]}{\bar{s}\beta D} \right) = 0 \quad (66)$$

(because $W_{-1}(-xe^{-x}) = -x$ for $x \geq 1$) and

$$p_0 = \bar{s}\beta \left(1 + \frac{W_0[-(\bar{s}\beta D)e^{-\bar{s}\beta D}]}{\bar{s}\beta D} \right) \quad (67)$$

The remaining poles are complex and with negative real parts (see Appendix B) as shown in the following figure of $p_m D, m = -1, 0, 1, 2, \dots$



Therefore, the terms $e^{p_m t}$ rapidly decay for $m \geq 1$, and the summation in eqn. (65) quickly becomes approximately equal to

$$\begin{aligned} \Delta s &\approx \\ &\approx \Delta s_0 \left(\lim_{q \rightarrow 0} \left[\frac{q}{q - \bar{s}\beta + e^{-qD}\bar{s}\beta} \right] + e^{p_0 t} \lim_{q \rightarrow p_0} \left[\frac{q - p_0}{q - \bar{s}\beta + e^{-qD}\bar{s}\beta} \right] \right) \Rightarrow \\ \Delta s &\approx \Delta s_0 \left(\frac{1}{1 - \bar{s}\beta D} + e^{p_0 t} \frac{1}{1 - e^{-p_0 D}\bar{s}\beta D} \right) \end{aligned} \quad (68)$$

The time dependence of i can be obtained in a similar fashion, as eqn. (8), combined with eqn. (57) implies

$$\begin{aligned} i &= \Delta s_{-D} - \Delta s \Rightarrow \\ I &= (e^{-qD} - 1)\Delta s = \frac{e^{-qD} - 1}{q - \bar{s}\beta + e^{-qD}\bar{s}\beta} \Delta s_0 \end{aligned} \quad (69)$$

which eventually yields

$$i \approx i_0 \left(\frac{D}{\bar{s}\beta D - 1} + e^{p_0(t-D)} \frac{D}{e^{-p_0 D}\bar{s}\beta D - 1} \right) \quad (70)$$

The rates for Padé SIR and standard SIR are obtained by standard ODE linearization analysis and omitted for brevity.

Appendix E. Proof of eqn.(25)

At the peak value i^* of i , eqn. (13) of the first-order Padé SIR model implies

$$s^* = \frac{1}{\beta D} \quad (71)$$

Dividing eqn. (1) by eqn. (13) and integrating yields

$$\frac{ds}{di} = \frac{-\beta s}{2(\beta s - \frac{1}{D})} \Rightarrow s - s_0 - \frac{1}{\beta D} \ln \frac{s}{s_0} = -\frac{i - i_0}{2} \quad (72)$$

For $i_0 = 1 - s_0 \approx 0$, combining eqn. (71) with the above yields

$$\begin{aligned} i^* &= 2 \left(\frac{-1}{\beta D} \ln(\beta D) - \frac{1}{\beta D} + 1 \right) \Leftrightarrow \\ \beta D &\stackrel{\text{def}}{=} R_0 = \frac{2W_{-1}\left(\frac{-2+i^*}{2e}\right)}{-2+i^*} \end{aligned} \quad (73)$$

where $W_{-1}\left(\frac{-2+i^*}{2e}\right)$ is the Lambert function of order -1 , as $-e^{-1} < \frac{-2+i^*}{2e} < 0$ for $0 \leq i^* \leq 1$, and $W_0\left(\frac{-2+i^*}{2e}\right)$ does not yield feasible values above 1.

The peak value i^* of i for the standard SIR model is obtained in an entirely similar way.

Appendix F. Proof of eqns. (40)-(44)

Following the same approach as in Appendix A for eqn. (32) yields

$$P = (e^{-qDp} - 1)U \quad (74)$$

Using the first-order Padé approximation

$$e^{-qD} \approx \frac{1 - \frac{Dp}{2}q}{1 + \frac{D}{2}q} \quad (75)$$

in eqn. (52) yields

$$P = \frac{-Dp q}{1 + \frac{D}{2}q} U \Rightarrow \left(1 + \frac{Dp}{2}q\right)P = -Dp q U \Rightarrow$$

$$p + \frac{Dp}{2}p' = -Dp u' = -Dp s' \Rightarrow$$

$$p' = 2 \left(f(i, p, s) - \frac{1}{Dp} p \right) \quad (76)$$

Similarly, adding eqns. (32) and (33), taking Laplace transforms, and using first-order Padé approximation yields

$$p + i = s_{-D_i} - s \Rightarrow$$

$$p' + i' = 2 \left(f(i, p, s) - \frac{1}{D_i} (p + i) \right) \Rightarrow$$

$$i' = 2 \left(\left(\frac{1}{D_p} - \frac{1}{D_i} \right) p - \frac{1}{D_i} i \right) \quad (77)$$

Second-order Padé approximations can be developed in a similar fashion, as shown in Appendix A.

Appendix G. Proof of eqn. (46)

Eqns. (31)-(34) imply that $(\bar{s}, 0, 0, 1 - \bar{s})$ is an equilibrium point, where \bar{s} can be arbitrary between 0 and 1. Then, local stability analysis of eqn. (31) by approximate linearization of $g(s_{-D_i}, s_{-D_p}, s) = f(s, p, i)$ around the steady state $s_{-D_i} = s_{-D_p} = s = \bar{s}$ yields

$$\begin{aligned} \Delta s' &\approx - \left(\frac{\partial g}{\partial s} \right)_{\bar{s}} \Delta s - \left(\frac{\partial g}{\partial s_{-D_p}} \right)_{\bar{s}} \Delta s_{-D_p} - \\ &\quad \left(\frac{\partial g}{\partial s_{-D_i}} \right)_{\bar{s}} \Delta s_{-D_i} \quad (78) \\ &= \bar{s}\beta_p \Delta s + \bar{s}(\beta_i - \beta_p) \Delta s_{-D_p} - \bar{s}\beta_i \Delta s_{-D_i} \end{aligned}$$

where $\Delta s \stackrel{\text{def}}{=} s - \bar{s}$, with $\Delta s(t < 0) = 0$, $\Delta s(0) = \Delta s_0$.

Taking Laplace transforms with $\Delta S(q) \stackrel{\text{def}}{=} \mathcal{L}[\Delta s(t)]$ yields

$$\Delta S = \frac{1}{q - \bar{s}\beta_p - \bar{s}(\beta_i - \beta_p)e^{-qDp} + \bar{s}\beta_i e^{-qD_i}} \Delta s_0 \quad (79)$$

The roots p_m of the transcendental characteristic equation

$$q - \bar{s}\beta_p - \bar{s}(\beta_i - \beta_p)e^{-qDp} + \bar{s}\beta_i e^{-qD_i} = 0 \quad (80)$$

(poles of ΔS) cannot be obtained in terms of standard functions known to the author. However, the linear approximation $e^{-x} \approx 1 - x$ in eqn. (80) immediately yields the result.

Appendix H. Proof of eqn.(47)

The proof follows the same pattern as in Appendix C, and is presented here briefly.

Dividing eqn. (31) by s , using eqn. (51), and taking Laplace transforms yields

$$\begin{aligned} q\mathcal{L}[\ln s] - \underbrace{\ln s_0}_{\approx 0} &= \\ &= \left[\beta_i \frac{e^{-qDp} - e^{-qD_i}}{q} + \beta_p \frac{1 - e^{-qDp}}{q} \right] q\mathcal{L}[u] \Rightarrow \end{aligned}$$

$$\begin{aligned} \lim_{q \rightarrow 0} q\mathcal{L}[\ln s] &= \\ &= \lim_{q \rightarrow 0} \left[\beta_i \frac{e^{-qDp} - e^{-qD_i}}{q} + \beta_p \frac{1 - e^{-qDp}}{q} \right] \lim_{q \rightarrow 0} q\mathcal{L}[u] \Rightarrow \end{aligned}$$

$$\ln s_\infty = (\beta_i(D_i - D_p) + \beta_p D_p)(s_\infty - 1)$$

by the Final Value Theorem.