



Analytical approximations for the phase plane trajectories of the SIR model with vital dynamics

Todd L Parsons, David J D Earn

► To cite this version:

Todd L Parsons, David J D Earn. Analytical approximations for the phase plane trajectories of the SIR model with vital dynamics. 2023. hal-04178969

HAL Id: hal-04178969

<https://cnrs.hal.science/hal-04178969>

Preprint submitted on 8 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analytical approximations for the phase plane trajectories of the SIR model with vital dynamics

Todd L. Parsons¹ and David J. D. Earn²

¹Laboratoire de Probabilités, Statistique et Modélisation (LPSM), Sorbonne Université, CNRS UMR 8001, Paris, France, 75005

²Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada, L8S 4K1

August 8, 2023 @ 17:49

Abstract

We derive accurate, closed-form analytical approximations for the phase-plane trajectories of the standard susceptible-infectious-removed (SIR) epidemic model, including host births and deaths. Our approximations for the SIR ordinary differential equations also allow us to provide convenient, accurate analytical approximations for the associated Poincaré map, and the minimum and maximum susceptible and infected host densities in each epidemic wave. Our analysis involves matching asymptotic expansions across branch cuts of the Lambert W function.

1 Introduction

Infectious disease transmission dynamics have been modelled with a susceptible-infectious-removed (SIR) compartmental framework since the early 20th century [18, 3]. What has become the standard SIR ordinary differential equation (ODE) model [Equation (1) below] was first published by Kermack and McKendrick in 1927 [12] (hereafter KM).

An exact solution for the time course of the state variables in the standard SIR ODEs has never been found, though KM did find an approximate analytical solution that is reasonably accurate for weakly transmissible diseases, and has often been used to obtain (usually qualitative) insights (*e.g.*, [4, 24, 20]).

If we ignore the recruitment of new susceptible individuals—whether they result from births, immigration, or decay of immunity—then the SIR ODEs can be solved exactly in the susceptible-infectious phase plane. No such exact analytical solution in the phase plane is available for the more realistic situation in which susceptible recruitment is not negligible.

In this paper, we use multiple scale and singular perturbation methods [22, 13] to obtain accurate analytical approximations to the phase-plane trajectories of the standard SIR ODEs with vital dynamics (*i.e.*, with host births and deaths). A natural small parameter is the infectious period relative to host lifetime (denoted ε), but we find that substantially simpler expressions are obtained by expanding in the (even smaller) parameter $\epsilon = \varepsilon/\mathcal{R}_0$, where \mathcal{R}_0 is the basic reproduction number.

Our interest in approximating the phase-plane solutions of the SIR ODEs was motivated by a problem in stochastic epidemic theory, namely estimating the probability of persistence of a pathogen after an initial epidemic in a naïve population. Our solution to that problem [23] depends on an analytical estimate of the fraction susceptible near the end of a major outbreak [Equation (117) below], which is one of many expressions that we derive here.

Key words and phrases. epidemics, SIR model, matched asymptotics, Poincaré map.
1991 *Mathematics subject classification.* 34E05, 34E13, 37N25, 92D30

We are not the first to attempt to approximate solutions of the SIR ODEs in the phase plane. In particular, van Herwaarden [25]—who was also motivated by (a different approach to) the stochastic extinction problem—presented some closely related approximations that are valid in specific regions of the phase plane. Here, we provide approximations that are uniformly valid throughout the phase plane, and we derive approximate expressions for a number of epidemiologically important quantities (*e.g.*, the peak and minimum prevalence, and the minimum and maximum susceptible frequency, following initial disease invasion and following subsequent epidemics).

Our matching yields simple expressions that can, unlike previous results, be applied to an arbitrary number of epidemic cycles and allows us to find an accurate analytical approximation to the Poincaré map for the SIR model.

The related expressions that van Herwaarden [25] has presented previously—in addition to applying only to restricted subregions of the phase plane—depend on the numerical evaluation of integrals for which explicit analytical forms cannot be found. In contrast, our results are fully analytical closed-form expressions that are valid everywhere in phase space. We succeed by exploiting Lambert’s W function [7] to invert implicit relations, and by asymptotic matching across branch cuts of the W function. The asymptotic techniques that we use are non-trivial and not part of the standard technical toolbox employed in mathematical epidemiology. Consequently, we present our analyses without assuming any familiarity with matched asymptotics, and hope that in so doing we have made it easier for readers to apply the methods to other problems.

For convenience in using our approximations, all of our major results are summarized in Tables 2 and 3.

1.1 The SIR Model with Vital Dynamics

Writing the proportions of the host population that are susceptible, infective, and removed as X , Y , and Z , respectively, the standard SIR ordinary differential equations (ODEs) are [2]

$$\frac{dX}{dt} = \mu(1 - X) - \beta XY, \quad (1a)$$

$$\frac{dY}{dt} = (\beta X - \gamma - \mu)Y, \quad (1b)$$

$$\frac{dZ}{dt} = \gamma Y - \mu Z, \quad (1c)$$

where μ is the *per capita* rate of birth and death, β is the transmission rate, and γ is the recovery (or removal) rate. Our focus in this paper is on solutions of these deterministic equations (1), but elsewhere [23] we show how they can be used together with a branching process approximation to obtain accurate analytical results for the fully stochastic model (including the probability of pathogen extinction due to stochastic effects). In keeping with our stochastic analysis [23], we reserve S , I , and R for the *number* of individuals in each state. For the sake of clarity, we adopt the convention of using lower case letters to indicate independent variables and capitals to indicate dependent variables.

We henceforth work with dimensionless parameters. To that end, we first note that a natural timescale is the expected duration of an individual’s infectious period,

$$T_{\text{inf}} = \frac{1}{\gamma + \mu}. \quad (2)$$

Expressing the **expected infectious period** in units of the expected host lifetime ($1/\mu$), we define

$$\varepsilon = \frac{\mu}{\gamma + \mu}. \quad (3)$$

The **basic reproduction number** (\mathcal{R}_0) is the product of the transmission rate β and the mean infectious period,

$$\mathcal{R}_0 = \frac{\beta}{\gamma + \mu}, \quad (4)$$

which gives the expected total number of new infections caused by a single infective individual introduced into a naïve population. We define a new time variable, $\tau = t/(\gamma + \mu)$, so that one time unit corresponds to the expected duration of an individual's infectious period. The SIR model then becomes

$$\frac{dX}{d\tau} = \varepsilon(1 - X) - \mathcal{R}_0 XY, \quad (5a)$$

$$\frac{dY}{d\tau} = (\mathcal{R}_0 X - 1)Y. \quad (5b)$$

Since $X(\tau) + Y(\tau) + Z(\tau) = 1$ for all $\tau \geq 0$, the two equations above completely describe the dynamical system (1). The ratio of ε and \mathcal{R}_0 ,

$$\epsilon = \frac{\varepsilon}{\mathcal{R}_0}, \quad (6)$$

turns out to be a better choice of small parameter to use in our analysis, because it leads to simpler asymptotic expressions.¹

Equations (5) have two nullclines. The y nullcline, $\frac{dY}{d\tau} = 0$, is the line

$$x = \frac{1}{\mathcal{R}_0}, \quad (7a)$$

whereas the x nullcline, $\frac{dX}{d\tau} = 0$, is the curve

$$y = \epsilon \left(\frac{1}{x} - 1 \right). \quad (7b)$$

When $\mathcal{R}_0 > 1$, these two nullclines intersect at (x_*, y_*) , where

$$x_* = \frac{1}{\mathcal{R}_0}, \quad (8a)$$

$$y_* = \varepsilon \left(1 - \frac{1}{\mathcal{R}_0} \right) = \varepsilon(1 - x_*). \quad (8b)$$

It is well known that any trajectory departing from an initial point (x_i, y_i) in the positive cone $\{(x, y) : x > 0, y > 0\}$ eventually converges on a globally asymptotically stable **endemic equilibrium** (EE) at (x_*, y_*) , whereas the **disease-free equilibrium** (DFE) at $(1, 0)$ is a saddle attracting the set $\{(x, y) : y = 0\}$. Approach to the EE occurs via damped oscillations provided [10, eq. 13]

$$\varepsilon < \frac{4(\mathcal{R}_0 - 1)}{\mathcal{R}_0^2} = 4x_*(1 - x_*). \quad (9)$$

This condition is satisfied for most diseases of interest; we restrict our focus to this typical behaviour.

We will be primarily concerned with the phase-plane trajectories of (5), so rather than using those equations directly, we will make use of the phase-plane equations,

$$\frac{dY}{dx} = \frac{(\mathcal{R}_0 x - 1)Y}{\varepsilon(1 - x) - \mathcal{R}_0 x Y} = \frac{(x - x_*)Y}{\epsilon(1 - x) - xY} \quad (10)$$

and

$$\frac{dX}{dy} = \frac{\varepsilon(1 - X) - \mathcal{R}_0 X y}{(\mathcal{R}_0 X - 1)y} = \frac{\epsilon(1 - X) - X y}{(X - x_*)y}. \quad (11)$$

Thus, $Y(x)$ and $X(y)$ indicate phase-plane solutions where x (resp. y) is the independent variable, whereas we use $X(\tau)$ and $Y(\tau)$ to indicate the solution to the time-parametrized equations (5). We shall need both

¹The biological meaning of ϵ is the time—in units of the mean infectious period—that can be expected to elapse before an infected individual that enters a fully susceptible population infects one other individual.

equations (10) and (11). Equation (10) [(11)] is singular where the susceptible [infected] hosts attain their minimum and maximum density, \underline{x} and \bar{x} [\underline{y} and \bar{y}], for a given cycle (these are points where nullclines, (7a) and (7b), are crossed). These singular points (see Figures 1 and 2) divide different branches of the multi-functions $Y(x)$ and $X(y)$, and we will need each equation to extend the solution beyond the singularities of the other. No exact analytical form is available for any of these turning points, but with our matched asymptotic expansions, we obtain asymptotic approximations to all of them (see Table 2 for a summary). We shall refer to the part of the trajectory above the x nullcline (7b) as the **epidemic phase** or simply as **epidemic** (e.g., we will speak of the **initial epidemic**, **second epidemic**, and so on) and the part below the x nullcline as the **trough**.

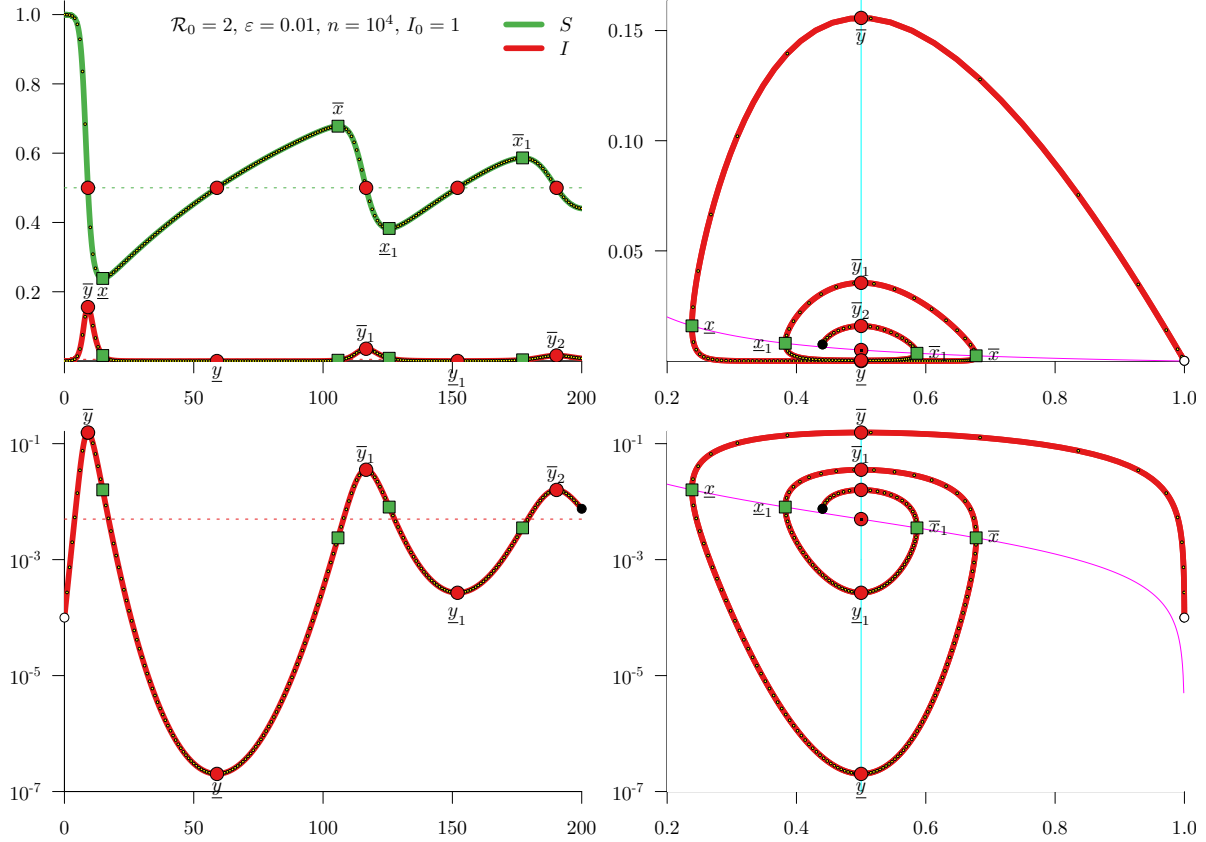


Figure 1: Sample solution of the SIR model Equation (1) for $\mathcal{R}_0 = 2$. *Top left*: the frequencies of susceptible and infected individuals. Symbols indicate the critical points of the curve of the corresponding colour. The j^{th} local minimum (maximum) of the susceptible [infective] frequency is labelled \underline{x}_j (\bar{x}_j) [\underline{y}_j (\bar{y}_j)]; we derive approximations to these optima and denote them $\underline{x}_{0,j}$, etc. (see Table 2). Dashed lines indicate the endemic equilibrium (8) of the model (1). *Bottom left*: the frequency of infected individuals (on a log scale) as a function of time. *Right*: trajectories in the susceptible-infected phase plane with the nullclines (*top right*: linear scale, *bottom right*: log scale).

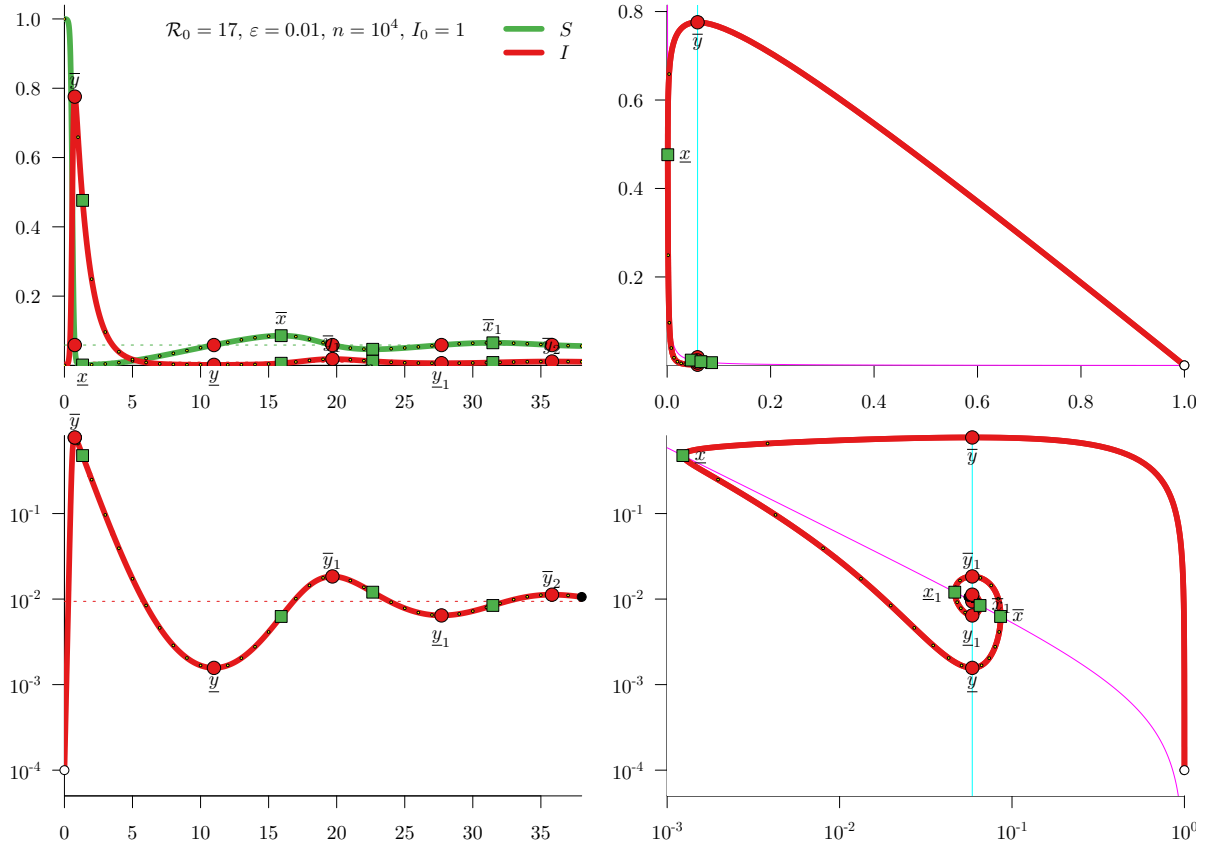


Figure 2: Sample solution of the SIR model Equation (1) for $\mathcal{R}_0 = 17$. See caption to Figure 1, but note that unlike Figure 1 both axes are logarithmic in the bottom right panel above.

1.2 The Method of Matched Asymptotic Expansions

Given a small parameter ϵ , an **asymptotic sequence** is a collection of functions²

$$\{\varphi_j(\epsilon), j = 1, 2, \dots\}, \quad \text{such that} \quad \varphi_{j+1}(\epsilon) \ll \varphi_j(\epsilon) \quad (12)$$

(a typical choice is $\varphi_j(\epsilon) = \epsilon^{j-1}$, but as we shall see below, other choices are often necessary). Given an equation—which may be an ordinary or partial differential equation—for an unknown function Y that depends on ϵ , we can look for a formal **asymptotic series** solution:

$$Y(x; \epsilon) = \sum_{j=0}^{\infty} Y_j(x) \varphi_j(\epsilon). \quad (13)$$

Note that we do not demand that the series converge; the infinite upper limit on the sum is to indicate that one could in principle compute an arbitrary number of terms.

Substituting the formal series into the exact equation (10) and equating terms of common order $\varphi_j(\epsilon)$, we obtain a hierarchy of equations that can be solved sequentially for the $Y_j(x)$. The first few terms of such a series solution often yield a good approximation to the exact solution. However, this approach may fail in regions of rapid change (*e.g.*, near boundaries or singularities) called **layers**³, especially in differential equations where one or more derivatives have coefficients depending on ϵ , which can make imposing all boundary (or initial or terminal) conditions impossible. The **method of matched asymptotic expansions** begins by identifying layers where different approximations better capture the correct quantitative and qualitative behaviour of the solution. These layers coincide with regions where some term in the equation is implicitly of size comparable to ϵ (*i.e.*, of order $\eta(\epsilon)$ for some function η), even when it doesn't explicitly contain ϵ , so that neglecting terms including ϵ fails to capture the correct dynamics. **Boundary layers** refer to subsets of the domain adjacent to the boundary (where one of the dependent variables is small compared to some function of ϵ) whereas **interior layers** occur away from the boundaries (*e.g.*, shear flow⁴ in a vector field could result in a region where one or more derivatives is small compared to a function of ϵ). **Corner layers** arise where boundary and/or interior layers intersect. (For the SIR model that we study here, there are boundary and corner layers, but no interior layers.)

Outside of such layers, the so-called **outer solution** provides a good approximation to the exact solution, but inside the layers it fails to capture the correct qualitative behaviour. Within a layer, rescaling of the independent and/or dependent variables by appropriately chosen factors $\eta(\epsilon)$ is used to amplify the local behaviour, and the resulting equation is then solved via another asymptotic series to obtain an **inner solution** that performs well in the layer, but typically is a poor approximation outside the layer. Inner and outer solutions are bridged by determining values for the constants of integration in the inner and outer solutions so that the various solutions intersect, and can thus be combined into a continuous approximation to the solution. When a combined solution is continuous, but not differentiable, a **corner layer** (and **corner layer solution**) near the point of intersection can be included in the matching to form a smooth approximation to the true trajectory that performs equally well across the entire domain (a **uniform asymptotic solution**). We discuss the matching procedure in greater detail in §3 below.

Table 1 lists estimates of natural history of infection parameters for a variety of common diseases. For the problem at hand, observing that infectious periods are typically of order days or weeks, whereas typical human lifetimes are longer than 50 years, we see that ϵ (6) will generally be quite small. Consequently, we anticipate that matched asymptotic solutions $Y(x; \epsilon)$ and $X(y; \epsilon)$ will provide good approximations to the exact solutions of (10) and (11). In subsequent sections, we derive and compare these asymptotic solutions to numerically computed trajectories for an illustrative parameter set (chosen with ϵ larger than for any acute infectious disease of humans so that discrepancies between the exact and asymptotic solutions are visible).

²We use the Hardy-Vinogradov notation: $f(\epsilon) \ll g(\epsilon)$ if and only if $\lim_{\epsilon \rightarrow 0} \frac{f(\epsilon)}{g(\epsilon)} = 0$. We also use Landau's "O" notation where convenient: $f(\epsilon) = o(g(\epsilon))$ if $f(\epsilon) \ll g(\epsilon)$, and $f(\epsilon) = \mathcal{O}(g(\epsilon))$ if there exists a positive constant C such that $|f(\epsilon)| \leq C|g(\epsilon)|$ as $\epsilon \rightarrow 0$.

³The terminology reflects the origins of the method in fluid mechanics [1].

⁴Shearing flow refers to adjacent regions where a vector field has substantially different magnitudes.

Disease	\mathcal{R}_0	x_*	T_{lat} [days]	T_{inf} [days]	$\varepsilon \times 10^3$	$\epsilon \times 10^3$	Source
measles	17	0.059	8	5	0.71	0.042	[2]
pertussis	17	0.059	8	14	1.2	0.071	[2]
mumps	12	0.08	15	6	1.1	0.092	[2]
chickenpox	11	0.091	10	5	0.82	0.075	[2]
COVID-19 (Delta)	6.8	0.15	5.8	14	1.1	0.16	[16]
rubella	6.5	0.15	10	7	0.93	0.14	[2]
scarlet fever	5.5	0.18	1.5	18	1	0.19	[2]
smallpox	4.5	0.22	15	7	1.2	0.27	[14]
COVID-19 (ancestral)	3	0.33	3.7	14	0.97	0.32	[16]
HIV	2.2	0.47	87	270	19	8.9	[11]
influenza (1918)	1.8	0.56	2	2.5	0.25	0.14	[19, 2]
Ebola	1.6	0.62	9.3	7	0.89	0.56	[26]
pneumonic plague	1.3	0.77	4.3	2.5	0.37	0.29	[9]

Table 1: Estimates of parameters associated with the natural history of infection for a variety of diseases. The observed parameters are the basic reproduction number (\mathcal{R}_0), the mean latent period (T_{lat}), and the mean infectious period (T_{inf}). The values of the other parameters were derived using Equations (2), (3), (6) and (8a). Note that the mean intrinsic generation interval in the SIR model (1) is $T_{\text{gen}} = 1/\gamma \simeq T_{\text{inf}}$, whereas in the SEIR model (which includes an exposed state in which individuals are not yet infectious), $T_{\text{gen}} \simeq T_{\text{lat}} + T_{\text{inf}}$ [15, 6]; consequently, SIR and SEIR dynamics correspond most closely if we set $1/\gamma$ in the SIR model to be the sum of the observed mean latent and infectious periods. We set $\mu = 0.02/\text{year}$ to mimic human birth and death rates, and compute $\varepsilon = \mu/(\gamma + \mu)$ [Equation (3)]. Where original sources present a range, we have listed the midpoint. Many of the estimates come from Anderson and May [2] (\mathcal{R}_0 from their Table 4.1 [2, p. 70]; the mean latent and infectious periods from their Table 3.1 [2, p. 31]).

2 Outer and Inner Solutions

As a first step in our analysis, in this section we derive solutions in the various subdomains where different approximations are natural. Then, in §3, we match local approximations to obtain a single global approximation that is uniformly valid throughout phase space.

2.1 Outer Solution

An outer solution is an asymptotic solution—obtained in the original variables—that captures the behaviour of the exact solution in the majority of phase space. Replacing Y in Equation (10) with

$$Y^{\text{out}}(x; \epsilon) = \sum_{j=0}^{\infty} Y_j^{\text{out}}(x) \epsilon^j, \quad (14)$$

and equating terms of similar order in ϵ , we obtain a hierarchy of equations that can be solved inductively for $Y_0^{\text{out}}(x), Y_1^{\text{out}}(x)$, and so on. The lowest order equation is

$$\frac{dY_0^{\text{out}}}{dx} = \frac{x_*}{x} - 1, \quad (15)$$

which has generic solution

$$Y_0^{\text{out}}(x) = C_0 - x + x_* \ln x, \quad (16)$$

where C_0^{out} is an arbitrary constant⁵. We can determine C_0^{out} by specifying initial conditions (x_i, y_i) , in which case Equation (16) gives the phase-plane trajectory of the SIR ODEs without vital dynamics, *i.e.*, for $\epsilon = 0$; see §3.1.1 below. In §3.1.2, we derive “effective initial conditions” that allow us to approximate subsequent epidemic waves.

2.2 x -axis Boundary Layer

The outer solution $Y^{\text{out}}(x; \epsilon)$ [Equation (14)] is a *function* of x , and cannot have the same qualitative behaviour as the exact trajectories, which have multiple branches as the fraction susceptible decreases to a minimum (\underline{x}) and then recovers (Figures 1 and 2). We thus seek a boundary layer solution along the x -axis that will capture the dynamics when the infected hosts are rare (*i.e.*, $y \ll 1$). Knowing that the equilibrium infective frequency y_* is $\mathcal{O}(\epsilon)$ [Equation (8b)], we might plausibly posit a solution

$$Y^{\text{b}}(x; \epsilon) = \epsilon \Upsilon(x; \epsilon), \quad (17)$$

where $\Upsilon = \mathcal{O}(1)$. Substituting $\epsilon \Upsilon$ for Y in Equation (10) yields

$$\epsilon \left((1-x) - x \Upsilon(x; \epsilon) \right) \frac{d\Upsilon}{dx} = (x - x_*) \Upsilon(x; \epsilon). \quad (18)$$

Searching for a series solution,

$$\Upsilon(x; \epsilon) = \sum_{j=0}^{\infty} \Upsilon_j(x) \epsilon^j, \quad (19)$$

and collecting terms of common order ϵ^j yields, for $j = 0$,

$$0 = (x - x_*) \Upsilon_0(x), \quad (20a)$$

which implies that $\Upsilon_0(x) \equiv 0$. Now, suppose that $\Upsilon_0(x) \equiv \dots \equiv \Upsilon_{j-1}(x) \equiv 0$; then, for $j \geq 1$, collecting terms of order ϵ^j in Equation (18) yields

$$0 = (x - x_*) \Upsilon_j(x). \quad (20b)$$

By induction, it follows that $\Upsilon_j(x) \equiv 0$ for all j , and hence that $Y^{\text{b}}(x; \epsilon) \equiv 0$, which is a contradiction. We must conclude that our assumption in Equation (19) that $\Upsilon(x; \epsilon)$ can be expanded in a series of powers of ϵ is incorrect. Instead, it must be that $\Upsilon(x; \epsilon)$ is **transcendentally small**, *i.e.*, vanishes more rapidly than any power ϵ^j as $\epsilon \rightarrow 0$ [13, p. 4].

As a means to guess the asymptotic dependence of $\Upsilon(x; \epsilon)$ on ϵ , we formally solve Equation (18) from the point of entry to the boundary layer (x_{in}) to an arbitrary point within it (x), which yields

$$\Upsilon(x; \epsilon) = \Upsilon(x_{\text{in}}; \epsilon) \exp \left(-\frac{1}{\epsilon} \int_{x_{\text{in}}}^x \frac{x_* - u}{(1-u) - u \Upsilon(u; \epsilon)} du \right). \quad (21)$$

We first note that $x < x_*$ in the x -axis boundary layer, so $1 - \xi > x_* - \xi > 0$ in the integrand above. Moreover, since $\xi \Upsilon(\xi; \epsilon)$ is transcendentally small we can expect that $\xi \Upsilon(\xi; \epsilon) \ll 1 - \xi$. Thus, the integrand can be expected to be positive throughout the layer, and the integral can be expected to be $\mathcal{O}(1)$ with respect to ϵ . Finally, the coefficient $\Upsilon(x_{\text{in}}; \epsilon)$ is evaluated at the edge of the boundary layer and can therefore be expected to be $\mathcal{O}(1)$. Putting these heuristic insights together, we hypothesize that $\Upsilon(x; \epsilon)$ is exponentially small in x with rate proportional to $\frac{1}{\epsilon}$. Consequently, we are led to what is known as a **WKB ansatz** [5, Chapter 10], that is, we postulate a solution of the form

$$Y^{\text{b}}(x; \epsilon) = e^{-\frac{\phi(x; \epsilon)}{\epsilon}} \quad (22)$$

⁵Throughout, we adopt the convention that C , with any combination of sub- and/or superscripts, indicates a constant of integration.

for some non-negative function $\phi(x; \epsilon)$ that can be expanded in an asymptotic series in ϵ . Substituting this ansatz (22) into Equation (10) gives us

$$\frac{1}{\epsilon} \frac{d\phi}{dx} = \frac{x_* - x}{\epsilon(1-x) - xe^{-\frac{\phi(x; \epsilon)}{\epsilon}}}, \quad (23)$$

for which we posit a series solution,

$$\phi(x; \epsilon) = \sum_{j=0}^{\infty} \phi_j(x) \epsilon^j, \quad (24)$$

so that

$$\begin{aligned} e^{-\frac{\phi(x; \epsilon)}{\epsilon}} &= e^{-\frac{\phi_0(x)}{\epsilon}} e^{-\phi_1(x)} e^{-\epsilon\phi_2(x) - \epsilon^2\phi_3(x) + \dots} \\ &= e^{-\frac{\phi_0(x)}{\epsilon}} e^{-\phi_1(x)} \left(1 - (\epsilon\phi_2(x) + \epsilon^2\phi_3(x) + \dots) + \frac{1}{2!} (\epsilon\phi_2(x) + \epsilon^2\phi_3(x) + \dots)^2 + \dots \right) \\ &= e^{-\frac{\phi_0(x)}{\epsilon}} e^{-\phi_1(x)} \left(1 - \epsilon\phi_2(x) + \epsilon^2 \left(\frac{\phi_2(x)^2}{2} - \phi_3(x) \right) + \dots \right), \end{aligned} \quad (25)$$

which is transcendentally small on any set where $\phi_0(x)$ is strictly positive (below we determine conditions under which $\phi_0(x) > 0$ on at least part of the interval $(0, 1)$). If $\phi_0(x) > 0$ then the term $e^{-\frac{\phi(x; \epsilon)}{\epsilon}}$ in Equation (23) is transcendentally small and we can omit it to obtain

$$\frac{d\phi_0}{dx} = \frac{x_* - x}{1 - x}, \quad (26)$$

and hence

$$\phi_0(x) = C_0^\phi - x - (1 - x_*) \ln(1 - x). \quad (27)$$

In contrast, we obtain $\frac{d\phi_1}{dx} = 0$, and hence

$$\phi_1(x) = C_1^\phi \quad (28)$$

(and similarly $\phi_j(x) = C_j^\phi$ for all $j > 1$). From the derivative (26), we see that $\phi_0(x)$ is increasing for $x < x_*$, decreasing for $x > x_*$, and has a vertical asymptote for $x = 1$. In particular, provided

$$C_0^\phi < x_* + (1 - x_*) \ln(1 - x_*), \quad (29)$$

$\phi_0(x)$ will be strictly positive on some subinterval of $(0, 1)$, as required. Inserting Equations (27) and (28) in Equation (25), we have an approximate solution in the x -axis boundary layer,

$$\bar{Y}(x; \epsilon) = (1 - x)^{-\frac{1}{\epsilon}(1 - x_*)} e^{\frac{1}{\epsilon}(C_0^\phi - x) + C_1^\phi}. \quad (30)$$

2.3 Corner layers

Close to the minimum susceptible frequency (\underline{x}), our outer (16) and inner (30) solutions have tangents of positive and negative slope respectively (see Figure 4), and thus meet in a non-differentiable corner, which we address by seeking a corner layer solution [13, p. 67]. The trajectories have vertical slope at \underline{x} , *i.e.*, $\frac{dY}{dx}$ is singular, so we will instead work with $\frac{dX}{dy}$. The increase from \underline{x} is caused by importation of new susceptible hosts at a rate proportional to ϵ [see Equation (11)], so we would expect the “turn-around” (change in sign of derivative) only when the frequency of infectives is sufficiently low that the first term dominates the second term in the numerator of Equation (11), *i.e.*, when y is $\mathcal{O}(\epsilon)$. We therefore amplify the behaviour near the x -axis, making a change of variable, $y = \epsilon v$, which converts Equation (11) to

$$\frac{dX}{dv} = \epsilon \frac{1 - X(v) - X(v)v}{(X(v) - x_*)v}. \quad (31)$$

Positing a series solution

$$X(v; \epsilon) = \sum_{j=0}^{\infty} X_j^{\text{cor}}(v) \epsilon^j, \quad (32)$$

yields, to lowest order, $\frac{dX_0^{\text{cor}}}{dv} = 0$, so X_0^{cor} is constant,

$$X_0^{\text{cor}}(v) = C_0^{\text{cor}}. \quad (33)$$

The next order term is

$$\frac{dX_1^{\text{cor}}}{dv} = -\left(\frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right) \frac{1}{v} + \left(\frac{C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right), \quad (34)$$

with solution

$$X_1^{\text{cor}}(v) = -\left(\frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right) \ln v + \left(\frac{C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right) v + C_1^{\text{cor}}. \quad (35)$$

We now insert Equations (33) and (35) in Equation (32), and convert back to the original variable y . Noting that $\ln v = \ln y + \ln \epsilon^{-1}$, we obtain

$$X(y; \epsilon) = C_0^{\text{cor}} + \left(\frac{C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right) y + \epsilon \ln \epsilon^{-1} \left(\frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right) - \epsilon \left[\left(\frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right) \ln y - C_1^{\text{cor}} \right] + \mathcal{O}(\epsilon^2). \quad (36)$$

This solution contains a term of order $\epsilon \ln \epsilon^{-1}$, which is intermediate between $\mathcal{O}(1)$ and $\mathcal{O}(\epsilon)$, and which we did not include in our proposed asymptotic series in v [Equation (32)]. The emergence of such a term when switching back to original variables is known as “transcendental switchback” [22, p. 71], and experience has shown that when this phenomenon occurs it is best to go back and include the intermediate order explicitly in the asymptotic sequence at the outset. Thus, we replace our initial ansatz (32) with

$$X(v; \epsilon) = X_0^{\text{cor}}(v) + \epsilon \ln \epsilon^{-1} X_{\ln}^{\text{cor}}(v) + \epsilon X_1^{\text{cor}}(v) + \mathcal{O}(\epsilon^2), \quad (37)$$

which includes a term of order $\epsilon \ln \epsilon^{-1}$ in addition to powers of ϵ . Inserting Equation (37) in Equation (31) we obtain a new hierarchy of ODEs, with one new equation,

$$\frac{dX_{\ln}^{\text{cor}}}{dy} = 0, \quad (38)$$

which yields a new constant,

$$X_{\ln}^{\text{cor}}(y) = C_{\ln}^{\text{cor}}. \quad (39)$$

Our revised expansion is

$$X(y; \epsilon) = C_0^{\text{cor}} + \left(\frac{C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right) y + \epsilon \ln \epsilon^{-1} \left[C_{\ln}^{\text{cor}} + \left(\frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right) \right] - \epsilon \left[\left(\frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}}\right) \ln y - C_1^{\text{cor}} \right] + \mathcal{O}(\epsilon^2), \quad (40)$$

which differs from Equation (36) only in that it includes C_{\ln}^{cor} in the $\mathcal{O}(\epsilon \ln \epsilon^{-1})$ term. Below, in §3.1.2, we will see that the additional degree of freedom provided by the constant C_{\ln}^{cor} is essential to successfully match other asymptotic solutions.

Remark 1 While this section concerns dynamics near \underline{x} , a virtually identical analysis provides a corner solution near \bar{x} , where the fraction susceptible is at its maximum prior to a second epidemic wave. In particular, we again have a solution of the form (40), albeit for different values of the constants C_0^{cor} and C_1^{cor} . We will refer to the corner solutions near \underline{x} and \bar{x} as the “left” and “right” corner solutions, X^{lc} and X^{rc} , respectively.

2.4 y -axis Boundary Layer

For large values of \mathcal{R}_0 , the trajectories of the SIR model approach the y -axis very closely, even when y is far from 0 (compare Figure 1 with $\mathcal{R}_0 = 2$ to Figure 2 with $\mathcal{R}_0 = 17$). Consequently, if \mathcal{R}_0 is large then there will be substantial periods during which x will be $\mathcal{O}(\epsilon)$ or smaller, and we can no longer assume that the effects of terms proportional to ϵ in Equation (10) can be safely neglected. Instead, we consider a boundary layer solution along the y -axis, making a change of variables $x = \epsilon\xi$ in Equation (10) to get

$$\frac{dY}{d\xi} = \frac{(\epsilon\xi - x_*)Y(\xi)}{1 - \epsilon\xi - \xi Y(\xi)}. \quad (41)$$

Positing a series solution,

$$Y(\xi; \epsilon) = \sum_{j=0}^{\infty} Y_j^{xb}(\xi) \epsilon^j, \quad (42)$$

we get, to lowest order,

$$\frac{dY_0^{xb}}{d\xi} = \frac{x_* Y_0^{xb}(\xi)}{\xi Y_0^{xb}(\xi) - 1}. \quad (43)$$

It is not immediately obvious how to solve this equation. However, if we invert the ODE, we have

$$\frac{d\Xi_0}{dy} = \frac{\Xi_0(y)}{x_*} - \frac{1}{x_* y}, \quad (44)$$

and we can now find a solution, $\xi = \Xi_0(y)$, using the method of integrating factors, which yields

$$X_0^{yb}(y) = \epsilon \Xi_0(y) = \frac{\epsilon}{x_*} e^{\frac{y}{x_*}} \left(E_1\left(\frac{y}{x_*}\right) + \epsilon \right), \quad (45)$$

where $E_1(z) = \int_z^\infty \frac{e^{-u}}{u} du$ is the exponential integral function (see, *e.g.*, [21, §6.2(i)]).

2.5 Scaled exponential and Lambert W functions

To match our various solutions, some functions of x , some functions of y , we need to write all expressions as functions of the same variable. For our matching, the common variable will be y , which is facilitated by two functions that we introduce in this section.

The expression ze^z occurs frequently in our analysis, with u often being a non-trivial expression itself. Consequently, in order to reduce equation clutter and emphasize patterns in expressions more clearly, we define the **scaled exponential function**,

$$\mathcal{E}(z) = ze^z. \quad (46)$$

The second function is the **Lambert W function** [7], a transcendental multi-function defined by the implicit relation

$$\mathcal{E}(W(z)) = z. \quad (47)$$

Thus, W is a Lambert W function if \mathcal{E} is its left inverse. There are countably many such functions of a complex argument, leading to countably many branches $W_i(z)$. There are two branches that are real-valued for real arguments. We will need both real-valued branches, which are denoted W_{-1} and W_0 . They have overlapping domains, but non-overlapping ranges,

$$W_{-1} : \left[-\frac{1}{e}, 0\right) \rightarrow (-\infty, -1], \quad (48a)$$

$$W_0 : \left[-\frac{1}{e}, \infty\right) \rightarrow [-1, \infty). \quad (48b)$$

Thus, these two branches meet at -1 ,

$$W_{-1}(-1) = W_0(-1) = -\frac{1}{e} = \mathcal{E}(-1). \quad (49)$$

For any i , \mathcal{E} is also a *partial* right inverse of W_i , i.e., $W_i(\mathcal{E}(z)) = z$ on *part* of the domain of \mathcal{E} . The set on which $W_i(\mathcal{E}(z)) = z$ depends on i ; for $i \in \{-1, 0\}$,

$$\begin{aligned} W_{-1}(\mathcal{E}(z)) &= z & \text{if } z \leq -1, \\ W_0(\mathcal{E}(z)) &= z & \text{if } z \geq -1. \end{aligned} \quad (50)$$

However, $W_i(\mathcal{E}(z))$ is well-defined outside the region on which \mathcal{E} is a right inverse, and it is on the domain where $W_i(\mathcal{E}(z)) \neq z$ that we frequently need to evaluate it. Indeed, the explicit final size formula for the SIR model (and many other models) is [17, Eq. (A.2)]

$$Z(\mathcal{R}_0) = 1 + \frac{1}{\mathcal{R}_0} W_0(\mathcal{E}(-\mathcal{R}_0)). \quad (51)$$

Graphs of $\mathcal{E}(z)$ and $W_i(\mathcal{E}(z))$ are shown in Figure 3.

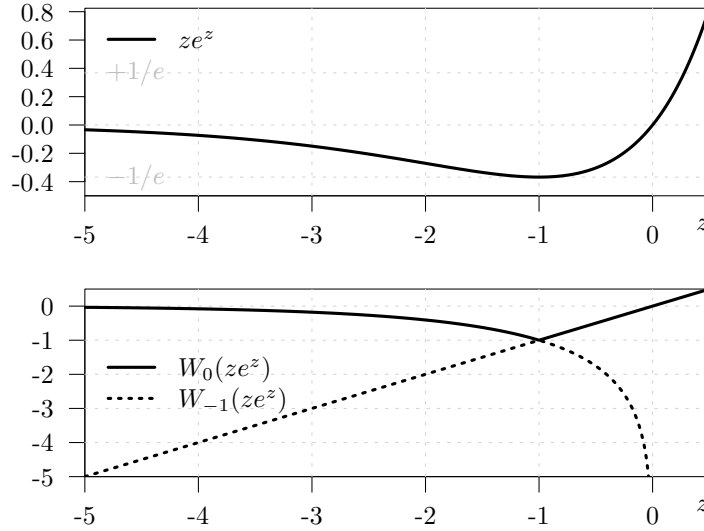


Figure 3: $\mathcal{E}(z) = ze^z$ (46) and Lambert W functions (47) evaluated at $\mathcal{E}(z)$.

We briefly recall some series and asymptotic expansions of the Lambert W function that we will need below. See [7] for details and proofs.

(i) Implicitly differentiating the identity $\mathcal{E}(W_i(z)) = W_i(z)e^{W_i(z)} = z$ and solving for $\frac{dW_i}{dz}$ one finds that

$$\frac{dW_i}{dz} = \frac{W_i(z)}{z(1 + W_i(z))}. \quad (52)$$

(ii) Applying the Lagrange inversion theorem (see, e.g., [8, p. 180]) to the power series for $\mathcal{E}(x)$, one finds that near 0,

$$W_0(z) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} z^n = z + O(z^2), \quad (53)$$

with radius of convergence $\frac{1}{e}$.

(iii) For large z , there is a (convergent) asymptotic series representation,

$$W_0(z) = L_1 - L_2 + \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \frac{(-1)^m}{m!} \frac{L_2^n}{L_1^{m+n}}, \quad z \rightarrow \infty, \quad (54)$$

where $L_1 = \ln z$ and $L_2 = \ln(\ln z)$. The same expansion applies for $W_{-1}(z)$ as $z \rightarrow -0$ if one replaces L_1 and L_2 with $\tilde{L}_1 = \ln(-z)$ and $\tilde{L}_2 = \ln(-\ln(-z))$.

(iv) We will also find it useful to expand

$$f_i(z) = W_i(-Ae^{-A+z}) = W_i(\mathcal{E}(-A)e^z) \quad (55)$$

in a series about $z = 0$ for various values of A . Using Equation (52), we find

$$f_i(z) = W_i(\mathcal{E}(-A)) + \frac{W_i(\mathcal{E}(-A))}{1 + W_i(\mathcal{E}(-A))} z + \mathcal{O}(z^2). \quad (56)$$

In particular, Equation (50) tells us that

$$f_i(z) = -A - \frac{A}{1-A} z + \mathcal{O}(z^2), \quad \text{if } i = 0 \text{ and } A \leq 1 \text{ or } i = -1 \text{ and } A \geq 1. \quad (57)$$

2.6 Inverting the outer and inner solutions

To facilitate matching, we now exploit the scaled exponential \mathcal{E} [Equation (46)] and Lambert's W [Equation (47)] to invert some of our asymptotic solutions.

Our outer solution (16) can be rearranged by setting $Y_0^{\text{out}}(x) = y$ and using Equation (46),

$$-\frac{1}{x_*} e^{\frac{y - C_0}{x_*}} = \mathcal{E}\left(-\frac{x}{x_*}\right). \quad (58)$$

Next, using Equation (50), we have

$$W_i\left(-\frac{1}{x_*} e^{\frac{y - C_0}{x_*}}\right) = W_i\left(\mathcal{E}\left(-\frac{x}{x_*}\right)\right) = -\frac{x}{x_*}, \quad (59)$$

where the $i = -1$ branch is to be used in the half-plane $x \geq x_*$ (since $W_{-1} \leq -1$, Equation (48a)), whereas the $i = 0$ branch is to be used in the half-plane $x \leq x_*$ (since $W_0 \geq -1$, Equation (48b)). From Equation (59), we obtain

$$X_0^i(y) = x = -x_* W_i\left(-\frac{1}{x_*} e^{\frac{y - C_0}{x_*}}\right). \quad (60)$$

Thus, the inversion yields a multi-function $X_0^i(y)$ with two branches, both of which we need. These branches correspond to the growth ($i = -1$) and decline ($i = 0$) phases of the epidemic, which meet at $y = Y_0^{\text{out}}(x_*)$ (the maximum of $Y_0^{\text{out}}(x)$, and thus also the upper bound of the domain of its inverse, $X_0^i(y)$).

Similarly, we can invert our inner solution in the x -axis boundary layer [Equation (30)] to get

$$X_0^i(y) = 1 + (1 - x_*) W_i\left(-\frac{1}{1 - x_*} e^{-\frac{1 - C_0^{\phi}}{1 - x_*}} \left(\frac{e^{C_1^{\phi}}}{y}\right)^{\frac{\epsilon}{1 - x_*}}\right), \quad (61)$$

where now the $i = -1$ and $i = 0$ branches give the solution for $x \leq x_*$ and $x \geq x_*$, respectively (the opposite of the situation for $X_0^i(y)$ in Equation (60), where $i = -1$ and $i = 0$ correspond to the right and left half-planes, respectively).

Remark 2 Once the constants have been determined by matching, we will be able to exploit the resulting symmetry (and the function \mathcal{E}) to simplify the inverted expressions considerably.

3 Matched asymptotic solutions

We now turn to the task of identifying the unknown constants in the solutions above and combining these local approximations into a uniform approximation of the trajectory. Suppose we have two or more solutions defined at different asymptotic scales (*e.g.*, our outer (§2.1) and corner (§2.3) solutions). The local solutions are obtained by imposing a scale on dependent or independent variables (*e.g.*, we supposed x is $\mathcal{O}(1)$ to get the outer solution, and assumed y was $\mathcal{O}(\epsilon)$ to get the corner solution, whereas we obtained the y -axis boundary layer solution (§2.4) by assuming X is $\mathcal{O}(\epsilon)$). In practice, however, each local solution remains valid over some larger domain that can be characterized with another asymptotic scale. Matching (see, *e.g.*, [13, §2.1]) is achieved by considering an intermediate scale $\eta = \eta(\epsilon)$ on which all solutions remain valid. Solutions to be matched are evaluated at $x = \eta x_\eta$ (or $y = \eta y_\eta$) for some x_η (or y_η) independent of ϵ (for readability, we suppress the explicit dependence of η itself on ϵ). The constants of integration (*e.g.*, C_0^{cor} , C_1^ϕ , *etc.*) are then chosen so that the two solutions agree as well as possible (*i.e.*, so that they coincide on as many orders as possible when both are expanded as an asymptotic series). The **matched solution** is obtained by summing the component solutions—with the choice of constants of integration that maximizes their mutual agreement—and subtracting their common **overlap** (the sum of all terms occurring in both asymptotic series).

To illustrate the process without getting bogged down in details, suppose $F(x)$ and $G(x)$ are outer and inner (*e.g.*, boundary layer) asymptotic series approximations for a given (exact) function $E(x)$. Moreover, suppose we have

$$F(x) = f_0(x) + f_1(x)\varphi_1(\epsilon) + f_2(x)\varphi_2(\epsilon) + f_3(x)\varphi_3(\epsilon) + \mathcal{O}(\varphi_4(\epsilon)), \quad (62)$$

$$G(x) = g_2(x)\varphi_2(\epsilon) + g_3(x)\varphi_3(\epsilon) + \mathcal{O}(\varphi_4(\epsilon)), \quad (63)$$

where the functions f_j and g_j contain arbitrary constants, and $\{\varphi_j\}$ is an asymptotic sequence [§1.2] (which typically refines the asymptotic sequences initially defined for the outer and inner solutions). If, say, the constants in $f_2(x)$ and $g_2(x)$ can be chosen so that these functions coincide exactly for all x , then $f_2(x)\varphi_2(\epsilon) = g_2(x)\varphi_2(\epsilon)$ is the overlap, and our matched approximation to $E(x)$ would be

$$F(x) + G(x) - f_2(x)\varphi_2(\epsilon). \quad (64)$$

If it were possible to choose the values of constants so that F and G agree in more than one order, so the overlap contains multiple orders (the more the better) then the resulting matched solution would be smoother (just as matching a function and its derivative at a single point leads to a smoother approximation at that point).

The matched solution has the virtue of being a valid approximation in both the inner and outer domains, so that one does not need to decide *a priori* which local solution (*e.g.*, outer, boundary, or corner) best approximates a given part of the trajectory.

In what follows, we will focus on matching the trajectory that escapes the DFE $[(x_i, y_i) = (1, 0)]$ along its unstable manifold, because that is the trajectory corresponding to disease invasion. We will give a detailed treatment of the first matching, and describe the second matching much more briefly. For easy reference, we summarize our results in Tables 2 and 3, while the matching constants are listed in Table 4.

3.1 Matching outer, corner and x -axis boundary layer solutions

3.1.1 Kermack and McKendrick's phase plane solution

Given an initial condition (x_i, y_i) , the constant of integration in Equation (16) is readily found to be

$$C_0^{\text{out}} = y_i + x_i - x_\star \ln x_i, \quad (65)$$

$$\text{so that } Y_0^{\text{out}}(x, x_i, y_i) = y_i + x_i - x + x_\star \ln\left(\frac{x}{x_i}\right), \quad (66)$$

which is the phase plane solution first discovered by KM. Equivalently, using Equation (60), we can express the solution as a function of y ,

$$X_0^{\text{out}}(y, x_i, y_i) = -x_* W_i\left(-\left(x_i/x_*\right) e^{\frac{y-y_i-x_i}{x_*}}\right) = -x_* W_i\left(\mathcal{E}(-x_i/x_*) e^{\frac{y-y_i}{x_*}}\right). \quad (67)$$

Provided $x_i \geq x_*$ and $y_i \geq 0$, the solution (66) is non-negative and concave, with two positive roots $[x_f \in (0, x_*)$ and another in $(x_*, 1)]$, and a unique maximum \bar{y}_0 at x_* ,

$$\bar{y}_0(x_i, y_i) = Y_0^{\text{out}}(x_*, x_i, y_i) = y_i + x_i - x_* (1 + \ln(x_i/x_*)). \quad (68)$$

Note that $\bar{y}_0(x_i, y_i)$ and $x_f(x_i, y_i)$ are the true peak prevalence and final size for the SIR model without vital dynamics ($\epsilon = 0$) started from (x_i, y_i) ; \bar{y}_0 only approximates the peak prevalence for the model with demography, and there is no “final” size if there is a continuous source of new susceptibles. Nevertheless, we informally refer to x_f as the “final size” for convenience (note that there *is* a minimum fraction susceptible, \underline{x} , near $x_f(x_i, y_i)$; see Equation (92) below).

Using the $i = 0$ branch to give the solution in the half-plane $x \leq x_*$, Equation (67) gives us an explicit expression for $x_f(x_i, y_i)$ [17],

$$x_f(x_i, y_i) = X_0^{\text{out}}(0) = -x_* W_0\left(\mathcal{E}(-x_i/x_*) e^{-\frac{y_i}{x_*}}\right). \quad (69)$$

The series expansion for Lambert’s W function (53) then yields

$$x_f(x_i, y_i) = x_i e^{-(x_i+y_i)/x_*} + \mathcal{O}\left((x_{i,2}/x_*) e^{-2(x_i+y_i)/x_*}\right) = x_i e^{-\mathcal{R}_0(x_i+y_i)} + \mathcal{O}\left(\mathcal{R}_0(x_i) e^{-\mathcal{R}_0(x_i+y_i)}\right)^2, \quad (70)$$

so the final size is exponentially small in \mathcal{R}_0 , with a correction of exponentially smaller order.

When referring specifically to the focal solution emanating from the DFE $[(x_i, y_i) = (1, 0)]$ we suppress the dependence on x_i and y_i and write $Y_0^{\text{out}}(x)$, $X_0^{\text{out}}(x)$, \bar{y}_0 , and x_f . While approximating this focal solution is our principal goal, in §3.1.3 below we will need to use solutions with $y_i = 0$ and $0 < x_i < 1$; for these solutions with infection-free endpoints, we write $Y_0^{\text{out}}(x, x_i)$ or $X_0^{\text{out}}(y, x_i)$, and refer to $\bar{y}_0(x_i)$ and $x_f(x_i)$. $[x_f(x_i, 0), x_i]$ is the interval on which $Y_0^{\text{out}}(x, x_i)$ is non-negative, and thus its domain for practical purposes; its range, $[0, \bar{y}_0(x_i)]$, is the domain for its inverse $X_0^{\text{out}}(y, x_i)$.

3.1.2 Matching in the left half-plane ($x \leq x_*$)

Our next step is to match the outer solution $X_0^{\text{out}}(y)$ [Equation (67) with $(x_i, y_i) = (1, 0)]$ with the corner solution $X^{\text{cor}}(y; \epsilon)$ [Equation (40)]. Note that the DFE lies in the right half-plane (since $x_* < 1$), but in the present subsection it is the $i = 0$ branch of the outer solution that we need because we are investigating only the part of the trajectory that lies in the left half-plane.

In order to consider the behaviour near $(x_f, 0)$ on scales intermediate between the outer solution (67) and the corner solution (40), we take $y = \eta y_\eta$ with

$$\epsilon \ll \eta \ll 1. \quad (71)$$

We haven’t yet identified an appropriate intermediate scale, but examples that satisfy Equation (71) include $\eta = \epsilon^{1/2}$ and $\eta = \epsilon \ln \epsilon^{-1}$.

Inserting $(x_i, y_i) = (1, 0)$ in the outer solution (67), and expanding it using Equation (56), we obtain

$$X_0^{\text{out}}(\eta y_\eta) = -x_* W_i\left(\mathcal{E}(-1/x_*) e^{\frac{\eta y_\eta}{x_*}}\right) = x_f + \eta \frac{x_f}{x_* - x_f} y_\eta + \mathcal{O}(\eta^2). \quad (72)$$

On the other hand, inserting $y = \eta y_\eta$ in the corner solution (40) and expanding

$$\ln(\eta y_\eta) = -\ln \eta^{-1} + \ln y_\eta, \quad (73)$$

we can write

$$X^{\text{cor}}(\eta y_\eta; \epsilon) = C_0^{\text{cor}} + \eta \frac{C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}} y_\eta + \epsilon \ln \epsilon^{-1} \left(C_{\text{ln}}^{\text{cor}} + \frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}} \right) + \epsilon \ln \eta^{-1} \left(\frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}} \right) - \epsilon \left(\frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}} \ln(y_\eta) - C_1^{\text{cor}} \right) + \mathcal{O}(\epsilon^2). \quad (74)$$

If we now refine our initial assumption (71) to

$$\epsilon \ln \epsilon^{-1} \ll \eta \ll (\epsilon \ln \epsilon^{-1})^{1/2} \ll 1 \quad (75)$$

then in the expansions (72,74) each term has a distinct asymptotic order. Our scaling assumptions ensure that

$$\epsilon \ll \epsilon \ln \eta^{-1} \ll \epsilon \ln \epsilon^{-1} \ll \eta \ll (\epsilon \ln \epsilon^{-1})^{1/2} \ll 1, \quad (76)$$

so if we take

$$C_0^{\text{cor}} = x_f \quad \text{and} \quad C_{\text{ln}}^{\text{cor}} = - \left(\frac{1 - C_0^{\text{cor}}}{x_* - C_0^{\text{cor}}} \right) = - \frac{1 - x_f}{x_* - x_f} \quad (77)$$

then the two solutions (72,74) coincide to⁶ $\mathcal{O}(\epsilon \ln \epsilon^{-1})$. For the moment, C_1^{cor} remains undetermined, but we will use it to match with the inner solution.

To match the corner layer solution with the inner (x -axis boundary layer) solution we now let η denote a different asymptotic order,

$$e^{-\frac{C}{\epsilon}} \ll \eta \ll \epsilon \ll 1 \quad \text{for all } C > 0. \quad (78)$$

Since we are interested in $x \leq x_*$, as noted after Equation (61) we must use the $i = -1$ branch of the inner solution $X_0^{x_b, i}(y)$. (Once we have determined the constants C_0^ϕ and C_1^ϕ in Equation (61) by matching with $i = -1$, the $i = 0$ branch will give the solution for $x \geq x_*$ without further work). We will use Equation (56) to derive an asymptotic expansion, which motivates us—after some algebraic exploration—to set

$$C_0^\phi = c_0^\phi - (1 - x_*) \ln(1 - c_0^\phi). \quad (79)$$

This choice for C_0^ϕ in Equation (61) leads to

$$X_0^{x_b, i}(\eta y_\eta) = 1 + (1 - x_*) W_i \left(- \frac{1 - c_0^\phi}{1 - x_*} e^{-\frac{1 - c_0^\phi}{1 - x_*}} \left(\frac{e^{C_1^\phi}}{\eta y_\eta} \right)^{\frac{\epsilon}{1 - x_*}} \right) = 1 + (1 - x_*) W_i \left(\mathcal{E} \left(- \frac{1 - c_0^\phi}{1 - x_*} \right) e^z \right), \quad (80a)$$

$$\text{where} \quad z = \frac{\epsilon}{1 - x_*} \ln \left(\frac{e^{C_1^\phi}}{\eta y_\eta} \right) = \frac{\epsilon}{1 - x_*} (C_1^\phi - \ln \eta - \ln y_\eta). \quad (80b)$$

Because z contains a factor ϵ , it is guaranteed to be small, so we can apply Equations (55) and (57) to Equation (80) to obtain

$$X_0^{-1}(\eta y_\eta) = c_0^\phi + \epsilon \left(\frac{1 - c_0^\phi}{x_* - c_0^\phi} \right) (C_1^\phi - \ln y_\eta) + \epsilon \ln \eta^{-1} \left(\frac{1 - c_0^\phi}{x_* - c_0^\phi} \right) + \mathcal{O}(\epsilon^2). \quad (81)$$

Our assumption that $\eta \ll \epsilon \ll 1$ [Equation (78)] implies that

$$\eta \ll \epsilon \ll \epsilon \ln \epsilon^{-1} \ll \epsilon \ln \eta^{-1} \ll 1. \quad (82)$$

⁶The assumptions $\epsilon \ll \eta$ and $\eta \ll (\epsilon \ln \epsilon^{-1})^{1/2}$ [Equation (76)] together imply that $\epsilon^2 \ll \eta^2 \ll \epsilon \ln \epsilon^{-1}$, which is necessary to ensure that the $\mathcal{O}(\eta^2)$ terms in Equation (72) are negligible in comparison to $\epsilon \ln \epsilon^{-1}$.

Consequently, comparing Equations (74) and (81) to order $\mathcal{O}(\epsilon)$, we see that the overlap is maximized by taking

$$c_0^\phi = C_0^{\text{cor}} = x_f \quad (\text{see Equation (77)}), \quad \text{and} \quad C_1^{\text{cor}} = \frac{1 - x_f}{x_* - x_f} C_1^\phi, \quad (83)$$

whereas C_1^ϕ is yet to be determined.

With the values of the constants determined above, the outer (72) and corner (74) solutions have a common overlap of

$$x_f + \eta \frac{x_f}{x_* - x_f} y_\eta, \quad (84)$$

whereas for the corner and inner (81) solutions, the overlap is

$$x_f - \epsilon \frac{1 - x_f}{x_* - x_f} (C_1^\phi - \ln y_\eta) + \epsilon \ln \eta^{-1} \frac{1 - x_f}{x_* - x_f}. \quad (85)$$

Summing the outer (72), inner (81) and corner (74) solutions and subtracting these two overlaps yields a matched solution to the left of x_* ,

$$\overleftarrow{X}(y; \epsilon) = X_0^{\text{out}}(y; \epsilon) + X_0^{\text{b}}(y; \epsilon) - x_f. \quad (86)$$

(Subtracting the overlaps (84,85) removes the corner solution (74) from the matched solution; the corner was nonetheless necessary to determine the matching constant c_0^ϕ .) For the matched solution to be continuous, \overleftarrow{X} must agree at (x_*, \bar{y}_0) with the outer solution (67) in the right half plane, *i.e.*, we need

$$\overleftarrow{X}(\bar{y}_0; \epsilon) = x_* = X_0^{\text{out}}(\bar{y}_0). \quad (87)$$

This requirement is satisfied provided

$$C_1^\phi = \ln \bar{y}_0. \quad (88)$$

Substituting the values of the matching constants C_0^ϕ and C_1^ϕ into Equation (61), we find that the boundary layer inner solution expressed as a function of y is

$$X_0^{\text{b}}(y; \epsilon) = 1 + (1 - x_*) W_i \left(- \frac{1 - x_f}{1 - x_*} \right) \left(\frac{\bar{y}_0}{y} \right)^{\frac{\epsilon}{1 - x_*}}. \quad (89)$$

Substituting the values of C_0^ϕ and C_1^ϕ into Equation (30) gives us an alternative description of the boundary layer dynamics as a function of x ,

$$Y^{\text{b}}(x; \epsilon) = \bar{y}_0 \left(\frac{1 - x_f}{1 - x} \right)^{\frac{1 - x_*}{\epsilon}} e^{-\frac{x - x_f}{\epsilon}}. \quad (90)$$

Inserting the values of C_0^{cor} , $C_{\text{ln}}^{\text{cor}}$, and C_1^{cor} , into Equation (40), we see that the corner solution near x_f is

$$X^{\text{lc}}(y; \epsilon) = x_f + \frac{x_f}{x_* - x_f} y + \epsilon \frac{x_f}{x_* - x_f} \ln \left(\frac{\bar{y}_0}{y} \right), \quad (91)$$

where, as noted in ?? 1, we use “lc” to emphasize that this is a “left corner” solution lying in the left half-plane. We will consider a right corner solution below.

Minimum susceptible proportion The left corner solution (91) approximately characterizes the trajectory near \underline{x} , the point where the fraction susceptible is minimized. Solving $\frac{dX}{dv} = 0$ [Equation (31)], we find that its minimum occurs at $v = \left(\frac{1}{x_f} - 1\right)$, whence the minimum fraction susceptible is approximately

$$\begin{aligned} \underline{x}_0 = x_0(x_f) &= \overset{\text{lc}}{X}\left(\epsilon\left(\frac{1}{x_f} - 1\right); \epsilon\right) = x_f + \epsilon\left(\frac{1 - x_f}{x_* - x_f}\right)\left[1 + \ln\left(\frac{\epsilon\left(\frac{1}{x_f} - 1\right)}{\bar{y}_0}\right)\right] \\ &= x_f - \epsilon \ln \epsilon^{-1}\left(\frac{1 - x_f}{x_* - x_f}\right) + \epsilon\left(\frac{1 - x_f}{x_* - x_f}\right)\left[1 + \ln\left(\frac{\frac{1}{x_f} - 1}{\bar{y}_0}\right)\right]. \end{aligned} \quad (92)$$

Prevalence trough Substituting $x = x_*$ into Equation (90) gives us an approximation to \underline{y} , the minimum fraction infected after the initial epidemic,

$$\underline{y}_0 = \bar{y}_0 \left(\frac{1 - x_f}{1 - x_*}\right)^{\frac{1 - x_*}{\epsilon}} e^{-\frac{x_* - x_f}{\epsilon}}. \quad (93)$$

Point of entry into the boundary layer With the known values of the matching constants c_0^ϕ and C_0^ϕ , we can write the leading term (27) in the asymptotic series (25) by which we obtained the inner solution as

$$\phi_0(x) = (x - x_f) - (1 - x_*) \ln\left(\frac{1 - x_f}{1 - x}\right). \quad (94)$$

The inner solution (22,30) is proportional to $e^{-\frac{\phi_0(x)}{\epsilon}}$ [see Equation (25)]. Consequently, as we observed in §2.2, the inner solution is transcendentally small in ϵ^{-1} on the set of x where $\phi_0(x) > 0$, whereas we see from Equation (94) that $\phi_0(x_f) = 0$. Thus, x_f is effectively the point of entry into the boundary layer: $e^{-\phi_0(x_f)/\epsilon} = 1 = \mathcal{O}(1)$, whereas for $x > x_f$ (near x_f), $e^{-\phi_0(x)/\epsilon}$ is transcendentally small.

Point of exit from the boundary layer In addition to x_f , $\phi_0(x)$ has a second root that we denote $x_{i,2}$ (for reasons that will become clear when we complete the matching). The point $x_{i,2}$ is where the trajectory exits the boundary layer: $\phi_0(x) > 0$ for $x \in (x_f, x_{i,2})$, and $\phi_0(x) < 0$ for $x > x_{i,2}$. Just as we used a corner solution at x_f to characterize the transition from the outer solution to the inner solution entering the boundary layer,⁷ a right corner solution at $x_{i,2}$ allows us to match the inner solution to a new outer solution corresponding to the second epidemic wave. To find an expression for $x_{i,2}$ we substitute $x = x_{i,2}$ into Equation (94) and obtain

$$(1 - x_*) \ln\left(\frac{1 - x_f}{1 - x_{i,2}}\right) - (x_{i,2} - x_f) = 0. \quad (95)$$

As in §(2.6), we solve this for $x_{i,2}$ using the Lambert W -function and find

$$x_{i,2} = 1 + (1 - x_*) W_0\left(\mathcal{E}^\epsilon\left(-\frac{1 - x_f}{1 - x_*}\right)\right). \quad (96)$$

3.1.3 Matching in the right half-plane ($x \geq x_*$)

Our matched inner layer solution (89) can be continued to the right of x_* by switching from the $i = -1$ to the $i = 0$ branch of Lambert's W . As we observed above, this boundary layer solution is transcendentally

⁷The pedantic reader (or author) might observe that the corner solution was obtained by assuming that $y = \mathcal{O}(\epsilon)$, whereas the outer and inner solutions correspond to $y = \mathcal{O}(1)$ and transcendentally small y , respectively. This apparent incongruity is reconciled by considering the solutions in a very small neighbourhood of x_f : for $C > 0$, $\exp\left(-\frac{\phi_0(x_f + C\epsilon \ln \epsilon^{-1})}{\epsilon}\right) = \mathcal{O}(\epsilon)$, and it is in this $\mathcal{O}(\epsilon \ln \epsilon^{-1})$ neighbourhood of x_f that the solutions match, which is reflected in the scaling (75) required when matching the solutions in §3.1.2.

small for $x < x_{i,2}$, *i.e.*, for all y such that $\bar{X}_0^{\text{b}}(y; \epsilon) < x_{i,2}$. As the trajectory leaves the boundary layer, the fraction infected goes from transcendentally small to $\mathcal{O}(1)$, until eventually the rate of infection exceeds the rate of replenishment of susceptible hosts by host demography, causing a second **turn-around**, where now the fraction susceptible starts to decrease. Our inner solution fails to capture this turn-around, which we now address, as in §3.1.2, with a (right) corner solution near $x_{i,2}$.

We begin by considering our solutions on a scale η that is intermediate between $\mathcal{O}(\epsilon)$ and transcendentally small [Equation (78)]. Expanding the inner solution $\bar{X}_0^{\text{b}}(\eta y_\eta)$ as in Equation (81), we find that

$$\bar{X}_0^{\text{b}}(\eta y_\eta) = x_{i,2} + \epsilon \frac{1 - x_{i,2}}{x_\star - x_{i,2}} (\ln \bar{y}_0 - \ln y_\eta) + \epsilon \ln \eta^{-1} \frac{1 - x_{i,2}}{x_\star - x_{i,2}} + \mathcal{O}(\epsilon^2). \quad (97)$$

Comparing this expansion to the corner series (74), we see that a maximal matching is obtained by taking

$$\bar{C}_0^{\text{cor}} = x_{i,2}, \quad \bar{C}_{\text{ln}}^{\text{cor}} = -\frac{1 - x_{i,2}}{x_\star - x_{i,2}}, \quad \text{and} \quad \bar{C}_1^{\text{cor}} = \frac{1 - x_{i,2}}{x_\star - x_{i,2}} \ln \bar{y}_0. \quad (98)$$

Substituting these values in Equation (74) gives us the right corner solution,

$$\bar{X}^{\text{rc}}(y; \epsilon) = x_{i,2} + \frac{x_{i,2}}{x_\star - x_{i,2}} y + \epsilon \frac{1 - x_{i,2}}{x_\star - x_{i,2}} \ln \left(\frac{\bar{y}_0}{y} \right) + \mathcal{O}(\epsilon^2). \quad (99)$$

Beyond the turn-around at the corner, we are again in the domain of validity of the outer solution $\bar{X}_0^{\text{out}}(y)$ [Equation (60)], where we now use the $i = -1$ branch as we are matching in the right half-plane. To match corner and outer solutions, we choose

$$\epsilon \ll \eta \ll 1 \quad (100)$$

and set $y = \eta y_\eta$. As we did for C_0^ϕ in the inner solution in §3.1.2 [Equation (79)], we make a change of constants,

$$\bar{C}_0^{\text{out}} = c_0^{\text{out}} - x_\star \ln c_0^{\text{out}}, \quad (101)$$

in Equation (60) to get

$$\bar{X}_0^{\text{out}-1}(\eta y_\eta) = -x_\star W_{-1} \left(-\left(c_0^{\text{out}}/x_\star \right) e^{-\frac{c_0^{\text{out}} - \eta y_\eta}{x_\star}} \right) = -x_\star W_{-1} \left(\mathcal{E} \left(-\left(c_0^{\text{out}}/x_\star \right) \right) e^{\frac{\eta y_\eta}{x_\star}} \right). \quad (102)$$

Expanding this expression using Equations (55) and (57) then gives us

$$\bar{X}_0^{\text{out}-1}(\eta y_\eta) = c_0^{\text{out}} + \eta \frac{c_0^{\text{out}}}{x_\star - c_0^{\text{out}}} y_\eta + \mathcal{O}(\eta^2). \quad (103)$$

Substituting $y = \eta y_\eta$ in Equation (99) and expanding exactly as in Equation (74), we find that Equations (99) and (103) agree to order $\mathcal{O}(\eta)$ provided

$$c_0^{\text{out}} = x_{i,2}. \quad (104)$$

Thus, the matched solution is

$$\bar{X}_0^{\text{out}-1}(y, x_{i,2}) = -x_\star W_{-1} \left(\mathcal{E} \left(-x_{i,2}/x_\star \right) e^{\frac{y}{x_\star}} \right). \quad (105)$$

As in our derivation of the left solution $\bar{X}(y; \epsilon)$ [Equation (86)], we now sum the outer (105), inner (89), and corner (99) solutions, and subtract their overlaps to obtain a uniform asymptotic solution to the right of x_\star ,

$$\vec{X}(y; \epsilon) = \bar{X}_0^{\text{out}-1}(y, x_{i,2}) + \bar{X}_0^{\text{b}}(y; \epsilon) - x_{i,2}. \quad (106)$$

Consistent approximations We now have consistent approximations to the trajectory that starts from the DFE, $(x_i, y_i) = (1, 0)$. From the initial time until the peak prevalence is reached, the trajectory is in the right half-plane and we use the KM solution (66) for the model without vital dynamics ($\epsilon = 0$). We then continue into the left half-plane using $\overleftarrow{X}(y; \epsilon)$ [Equation (86)] until the first prevalence trough is reached at $x = x_*$, where we switch to $\overrightarrow{X}(y; \epsilon)$ [Equation (106)] to approximate the rising segment of the second epidemic. The switches from one approximation to another are differentiable and always occur when $x = x_*$, and the combined approximation is uniformly valid (*i.e.*, valid to the same order throughout the phase plane).

Effective initial conditions Comparing $\overleftarrow{X}_0^{-1}(y, x_{i,2})$ [(105)] to $\overrightarrow{X}_0^i(y, x_i, y_i)$ [(67)], we see that $(x_{i,2}, 0)$ is an *effective initial condition* for the second epidemic: if $(x_{i,2}, 0)$ were used as the initial state in the KM ($\epsilon = 0$) solution, the resulting trajectory would meet the second rise of the actual solution as it curves up from the left in the phase plane at (approximately) \overline{x}_0 (see below). Thus, while $(x_{i,2}, 0)$ is not a point on the actual trajectory, it represents an “effective” initial condition that would give rise to the true dynamics after the end (*i.e.*, trough) of the first epidemic. This observation motivates our choice of notation $x_{i,2}$.

Maximum fraction susceptible Just as the minimum value of the left corner solution near x_f (91) gives an estimate of \underline{x} [Equation (92)], the maximum value for the right corner solution near $x_{i,2}$ (99) gives us an estimate of \overline{x} , the maximum fraction susceptible before a second epidemic wave,

$$\overline{x}_0 = \overline{x}_0(x_{i,2}) = x_{i,2} - \epsilon \frac{1 - x_{i,2}}{x_{i,2} - x_*} \left(1 + \ln \left(\epsilon \left(\frac{1}{x_{i,2}} - 1 \right) / \overline{y}_0 \right) \right), \quad (107)$$

which occurs at $y = \epsilon \left(\frac{1}{x_{i,2}} - 1 \right)$.

Peak prevalence for the second wave Writing Equation (105) as a function of x via KM’s formula (66),

$$\overrightarrow{Y}_0^{\text{out}}(x) = x_{i,2} - x + x_* \ln(x/x_{i,2}), \quad (108)$$

we can also obtain an approximation of the second epidemic’s prevalence peak,

$$\overline{y}_{0,2} \approx x_{i,2} - x_* (1 + \ln(x_{i,2}/x_*)). \quad (109)$$

3.1.4 Matching beyond the first epidemic wave

Our uniform matched asymptotic solutions, \overleftarrow{X} and \overrightarrow{X} (86,106), were derived starting from the DFE at $(x_i, y_i) = (1, 0)$. However, a straightforward observation allows us to use the formulae for \overleftarrow{X} and \overrightarrow{X} for the entire trajectory (*i.e.*, all epidemic waves). Other than x_* and ϵ (or the more fundamental parameters \mathcal{R}_0 and ε), the only parameters on which our approximations depend are the initial condition (x_i) , the approximate maximum size of the epidemic (\overline{y}_0) [(68)], the final size of the epidemic without vital dynamics (x_f) [(69)], and the effective initial condition for the next epidemic $(x_{i,2})$ [(96)].

Epidemic iteration Setting $x_{i,1} = x_i$, we iteratively obtain $x_{i,j+1}$ from $x_{i,j}$ by computing

$$x_{f,j} = x_f(x_{i,j}, 0) = -x_* W_0(\mathcal{E}(-x_{i,j}/x_*)), \quad (110a)$$

$$\overline{y}_{0,j} = \overline{y}_0(x_{i,j}, 0) = x_{i,j} - x_* (1 + \ln(x_{i,j}/x_*)), \quad (110b)$$

$$x_{i,j+1} = 1 + (1 - x_*) W_0 \left(\mathcal{E} \left(-\frac{1 - x_{f,j}}{1 - x_*} \right) \right). \quad (110c)$$

The intermediate quantities in this recurrence relation, $x_{f,j}$ and $\overline{y}_{0,j}$, are the final fraction susceptible and maximal fraction infected, respectively, for the SIR model *without* vital dynamics ($\varepsilon = 0$) with initial condition $(x_{i,j}, 0)$.

Substituting these expressions (110) into our uniform matched asymptotic solutions (86,106) provides approximations to the full j^{th} epidemic wave for all $j \geq 1$, \overleftarrow{X} (\overrightarrow{X}) valid on $[\underline{x}_{0,j}, x_*]$ (resp., $[x_*, \overline{x}_{0,j}]$), where

$$\underline{x}_{0,j} = \underline{x}_0(x_{f,j}, \overline{y}_{0,j}) = x_{f,j} + \epsilon \frac{1 - x_{f,j}}{x_* - x_{f,j}} \left(1 + \ln \left(\epsilon \left(\frac{1}{x_{f,j}} - 1 \right) / \overline{y}_{0,j} \right) \right), \quad (111a)$$

$$\overline{x}_{0,j} = \overline{x}_0(x_{i,j+1}, \overline{y}_{0,j}) = x_{i,j+1} - \epsilon \frac{1 - x_{i,j+1}}{x_{i,j+1} - x_*} \left(1 + \ln \left(\epsilon \left(\frac{1}{x_{i,j+1}} - 1 \right) / \overline{y}_{0,j} \right) \right). \quad (111b)$$

Poincaré map If we think of the y nullcline ($x = x_*$) as a surface of section, we can use Equation (110) to explicitly write down the associated Poincaré map. Using Equation (93), we define

$$\underline{y}_{0,j} = \underline{y}_0(x_{f,j}, \overline{y}_{0,j}) = \overline{y}_{0,j} \left(\frac{1 - x_{f,j}}{1 - x_*} \right)^{\frac{1 - x_*}{\epsilon}} e^{-\frac{x_* - x_{f,j}}{\epsilon}}. \quad (112)$$

We can then iteratively define the time-forward Poincaré map on the y nullcline via

$$\overline{y}_{0,1} \rightarrow \underline{y}_{0,1} \rightarrow \overline{y}_{0,2} \rightarrow \underline{y}_{0,2} \rightarrow \overline{y}_{0,3} \rightarrow \underline{y}_{0,3} \rightarrow \dots \quad (113)$$

3.2 Matching with the y -axis boundary layer solution

We conclude our analysis with a matching between the outer solution on the unstable manifold, $Y_0^{\text{out}}(x)$ (66), and the boundary layer solution along the y -axis, $X_0^{\text{yb}}(y)$ (45). For biologically relevant parameters, this new matching improves significantly upon our formulae for \overleftarrow{X} (86) and \overrightarrow{X} (106).

Having a boundary layer along the y -axis is sensible only for trajectories that approach the y -axis. Consequently, when studying this layer, we are assuming implicitly that $x_f = \mathcal{O}(\epsilon)$ (since $x_f < \underline{x}$ and no trajectory gets closer than \underline{x} to the y -axis). In particular, since

$$x_f = -x_* W_0 \left(\mathcal{E} \left(-\frac{1}{x_*} \right) \right) = -\frac{1}{\mathcal{R}_0} W_0(\mathcal{E}(-\mathcal{R}_0)) = e^{-\mathcal{R}_0} + \mathcal{O}(\mathcal{R}_0 e^{-2\mathcal{R}_0}), \quad (114)$$

we are implicitly assuming that $e^{-\mathcal{R}_0} = \mathcal{O}(\epsilon)$ or, equivalently, $\mathcal{R}_0 = \mathcal{O}(\ln \epsilon^{-1})$. For the diseases listed in Table 1, $\mathcal{R}_0 / \ln \epsilon^{-1}$ ranges from ~ 0.17 (for pneumonic plague and influenza) to ~ 1.7 (for measles and pertussis), suggesting it is not unreasonable to assume $\mathcal{R}_0 / \ln \epsilon^{-1}$. For the example in the top panels of Figures 4 and 5, $\mathcal{R}_0 / \ln \epsilon^{-1} \simeq 0.38$.

As with our original matching (§3.1), we use the outer solution expressed as a function of y , $X_0^{\text{out}}(y, x_i)$ (67), and since we are matching in the left half-plane ($x \leq x_*$), we set $i = 0$. Matching $X_0^{\text{out}}(y, x_i)$ with the y -axis boundary layer solution $X_0^{\text{yb}}(y)$ (45) is uncharacteristically simple: the two are of different asymptotic orders ($\mathcal{O}(1)$ and $\mathcal{O}(\epsilon)$), and the solutions have no overlap. The matched solution is thus their sum, with the constant C_0^{yb} as yet undetermined.

We set

$$C_0^{\text{yb}} = -E_1 \left(\frac{\overline{y}_0(x_i)}{x_*} \right) \quad (115)$$

so that $X_0^{\text{yb}}(y)$ vanishes when evaluated at (x_*, \overline{y}_0) and consequently the sum $X_0^{\text{out}}(y, x_i) + X_0^{\text{yb}}(y)$ agrees at (x_*, \overline{y}_0) with $X_0^{\text{out}}(y)$ (67), the corresponding focal approximation in the right half-plane ($x \geq x_*$). Thus, $X_0^{\text{yb}}(y)$ becomes

$$X_0^{\text{yb}}(y) = \frac{\epsilon}{x_*} e^{-\frac{y}{x_*}} \left(E_1 \left(\frac{y}{x_*} \right) - E_1 \left(\frac{\overline{y}_0(x_i)}{x_*} \right) \right). \quad (116)$$

516 With the choice (115) for $\overset{yb}{C}_0$, the sum

$$517 \quad X_{\text{in}}(y; \epsilon) := \overset{\text{out}}{X}_0^0(y, x_i) + \overset{yb}{X}_0^b(y) = -x_* W_0 \left(\mathcal{E} \left(-\frac{x_i}{x_*} \right) e^{\frac{y}{x_*}} \right) + \epsilon \frac{e^{\frac{y}{x_*}}}{x_*} \left(E_1 \left(\frac{y}{x_*} \right) - E_1 \left(\frac{\bar{y}_0(x_i)}{x_*} \right) \right) \quad (117)$$

518 is a very good approximation to the trajectory, except in the x -axis boundary layer. Elsewhere [23], we use
 519 Equation (117) to approximate the fraction susceptible at the point of entry into the set $\{y \leq y_*\}$ (hence
 520 “in”).

521 We next match with the inner solution expressed as a function of y , $\overset{xb}{X}_0^{-1}(y)$ (61) (now the $i = -1$ branch
 522 gives the solution with $x \leq x_*$), for which we obtained the asymptotic expansion for $y = \eta y_\eta$ previously
 523 [Equation (81)].

524 To expand the matched outer and y -axis boundary layer solutions (117), we introduce the **complemen-**
 525 **tary exponential integral** ([21, 6.2.4]),

$$\text{Ein}(z) = \int_0^z \frac{1 - e^{-u}}{u} du, \quad (118)$$

526 an entire function that satisfies $\text{Ein}(z) = z + \mathcal{O}(z^2)$ and

$$E_1(z) = \text{Ein}(z) - \ln z - \gamma, \quad (119)$$

527 where $\gamma \simeq 0.57721$ is the Euler-Mascheroni constant [21, 5.9.18] (not the recovery rate in the SIR model (1)).
 528 Using Equation (119), Equation (117) becomes

$$X_{\text{in}}(y; \epsilon) = x_f + \eta \left(\frac{x_f}{x_* - x_f} \right) y_\eta + \epsilon \ln \eta^{-1} \frac{1}{x_*} + \epsilon \frac{1}{x_*} \left(\ln \bar{y}_0 - \ln y_\eta + \text{Ein} \left(\frac{\bar{y}_0}{x_*} \right) \right) + \mathcal{O}(\epsilon \eta). \quad (120)$$

529 Comparing this expression with the asymptotic series expansion for the x -axis boundary layer solution (81),
 530 we see that the coefficient of $\ln y_\eta$ appears to be different in the two expansions ($\frac{\epsilon}{x_*}$ in (120) versus $\epsilon \frac{1-x_f}{x_*-x_f}$
 531 in (81)). This apparent difference is a consequence of the assumption implicit throughout this section that
 532 $x_f = \mathcal{O}(\epsilon)$, which implies that

$$\epsilon \frac{1-x_f}{x_*-x_f} = \frac{\epsilon}{x_*} + \mathcal{O}(\epsilon^2), \quad (121)$$

533 so the two coefficients are in fact *asymptotically* equal. With this in mind, we see that, as in the original
 534 matching,

$$c_0^\phi = x_f, \quad (122)$$

535 from which we obtain C_0^ϕ via Equation (79), whereas now

$$C_1^\phi = \ln \bar{y}_0 - \text{Ein} \left(\frac{\bar{y}_0}{x_*} \right). \quad (123)$$

536 Substituting c_0^ϕ and c_1^ϕ into Equation (61) yields

$$\overset{yb}{X}_i^0(y) = 1 + (1 - x_*) W_i \left(\mathcal{E} \left(-\frac{1-x_f}{1-x_*} \right) \left(\frac{\bar{y}_0}{y} \right)^{\frac{\epsilon}{1-x_*}} e^{-\frac{\epsilon}{1-x_*} \text{Ein} \left(\frac{\bar{y}_0}{x_*} \right)} \right). \quad (124)$$

537 This expression differs from the matched boundary layer solution $\overset{yb}{X}_0^{-1}(y)$ (89) by an additional factor

$$e^{-\frac{\epsilon}{1-x_*} \text{Ein} \left(\frac{\bar{y}_0}{x_*} \right)} = 1 - \frac{\epsilon}{1-x_*} \text{Ein} \left(\frac{\bar{y}_0}{x_*} \right) + \mathcal{O}(\epsilon^2), \quad (125)$$

which multiplies the argument of the W -function, giving an $\mathcal{O}(\epsilon)$ refinement to $X_0^{-1}(y)$.

Summing the outer and two inner solutions and subtracting the common overlap yields a solution uniformly valid to the left of $x = x_*$,

$$\overleftarrow{X}(y; \epsilon) = X_0^{\text{out}}(y) + X_0^{\text{yb}}(y) + X_0^{\text{yb}-1}(y) - x_{\text{f}} + \frac{\epsilon}{x_*} \left(\ln \left(\frac{y}{\bar{y}_0} \right) + \text{Ein} \left(\frac{\bar{y}_0}{x_*} \right) \right). \quad (126)$$

This solution can be extended the right half-plane ($x \geq x_*$) using the $i = 0$ branch of Equation (61), with c_0^ϕ , C_0^ϕ , and C_1^ϕ as determined above. The matching to the second epidemic wave then proceeds identically to §3.1.3 (except that the argument of the inner solution now has the additional factor $e^{-\frac{\epsilon}{1-x_*} \text{Ein}(\frac{\bar{y}_0}{x_*})}$), resulting in

$$\overrightarrow{X}(y; \epsilon) = X_0^{\text{out}-1}(y) + X_0^{\text{yb}}(y) - x_{i,2}. \quad (127)$$

We summarize these results in Table 3 and compare them to the numerically evaluated trajectory in Figure 5.

Subsequent epidemic waves Just as before (§3.1.4), these solutions can be extended to subsequent epidemic waves, replacing \bar{y}_0 , x_{f} and $x_{i,2}$ by $\bar{y}_{0,j}$, $x_{\text{f},j}$ and $x_{i,j+1}$ defined using the iterative scheme in Equations (110a) to (110c).

Quantity		Expression	Equation
Equilibrium susceptible density	x_*	$\frac{1}{\mathcal{R}_0}$	(8a)
Peak prevalence (KM)	\bar{y}_0	$1 - x_* (1 - \ln(x_*))$	(68)
Final size (KM)	x_f	$-x_* W_0(\mathcal{E}(-1/x_*))$	(69)
Minimum prevalence	\underline{y}_0	$\bar{y}_0 \left(\frac{1-x_f}{1-x_*} \right)^{\frac{1-x_*}{\epsilon}} e^{-\frac{x_*-x_f}{\epsilon}} e^{-\text{Ein}\left(\frac{\bar{y}_0}{x_*}\right)}$	(93)
Minimum susceptible density	\underline{x}_0	$x_f + \epsilon \frac{1-x_f}{x_*-x_f} \left(1 + \ln\left(\epsilon\left(\frac{1}{x_f} - 1\right)/\bar{y}_0\right) \right)$	(92)
Effective initial condition	$x_{i,2}$	$1 + (1 - x_*) W_0\left(\mathcal{E}\left(-\frac{1-x_f}{1-x_*}\right)\right)$	(96)
Maximum susceptible density	\bar{x}_0	$x_{i,2} - \epsilon \frac{1-x_{i,2}}{x_{i,2}-x_*} \left(1 + \ln\left(\epsilon\left(\frac{1}{x_{i,2}} - 1\right)/\bar{y}_0\right) \right)$	(107)
Peak prevalence, second wave	$\bar{y}_{0,2}$	$x_{i,2} - x_* (1 + \ln(x_{i,2}/x_*))$	(109)

Table 2: Approximations of quantities of epidemiological interest for disease invasions, *i.e.*, on the trajectory that emanates from the disease-free equilibrium (DFE), $(x_i, y_i) = (1, 0)$. Each entry may depend upon entries above it in the table (but never on entries below). These quantities are used in our approximations to the full trajectories in Table 3. We use “(KM)” to indicate quantities that are exact for the Kermack-McKendrick SIR model *without* vital dynamics ($\epsilon = 0$). With vital dynamics ($\epsilon > 0$), the peak prevalence \bar{y}_0 is an approximation, and there is no “final” size, but the quantity x_f appears in the approximation to the minimum fraction susceptible (\underline{x}_0). Replacing x_i by $x_{i,j}$ as defined in Equation (110) gives asymptotic approximations for the j^{th} epidemic wave, given that the first wave starts on the trajectory emanating from the DFE. We discuss the effective initial condition (x_i^1) in §3.1.3. The expressions for the minimum and maximum susceptible densities (\underline{x}_0 , \bar{x}_0) are identical except that \underline{x}_0 is evaluated at x_f and \bar{x}_0 is evaluated at $x_{i,2}$. In these expressions, note that $(x_f, \epsilon(\frac{1}{x_f} - 1))$ and $(x_{i,2}, \epsilon(\frac{1}{x_{i,2}} - 1))$ are points on the x nullcline $dX/d\tau = 0$. We write the formulae for \underline{x}_0 and \bar{x}_0 as compactly as possible here; see Equation (92) for the same expression written out with separate terms for each asymptotic order.

Solution	Notation	Expression	Branch	Domain	Equation
Outer	$X_0^{\text{out},i}(y, x_i)$	$-x_* W_i \left(\mathcal{E}(-x_i/x_*) e^{\frac{y}{x_*}} \right)$	$i = \begin{cases} 0 & x \leq x_* \\ -1 & x \geq x_* \end{cases}$	$[0, \bar{y}_0]$	(60), (67)
y -axis bdry	$X_0^{yb}(y)$	$\frac{\epsilon}{x_*} e^{-\frac{y}{x_*}} \left(E_1 \left(\frac{y}{x_*} \right) - E_1 \left(\frac{\bar{y}_0}{x_*} \right) \right)$	–	$[\underline{y}_0, \bar{y}_0]$	(116)
Left corner	$X^{\text{lc}}(y; \epsilon)$	$x_f + \frac{x_f}{x_* - x_f} y + \epsilon \frac{1 - x_f}{x_* - x_f} \ln(\bar{y}_0/y)$	–	$[\underline{y}_0, \bar{y}_0]$	(36), (91)
x -axis bdry	$X_0^{xb}(y)$	$1 + (1 - x_*) W_i \left(\mathcal{E} \left(-\frac{1 - x_f}{1 - x_*} \right) \left(\frac{\bar{y}_0}{y} \right)^{\frac{\epsilon}{1 - x_*}} e^{-\frac{\epsilon}{1 - x_*} \text{Ein} \left(\frac{\bar{y}_0}{x_*} \right)} \right)$	$i = \begin{cases} -1 & x \leq x_* \\ 0 & x \geq x_* \end{cases}$	$[\underline{y}_0, \bar{y}_0]$	(124)
Right corner	$X^{\text{rc}}(y; \epsilon)$	$x_{i,2} - \frac{x_{i,2}}{x_{i,2} - x_*} y - \epsilon \frac{1 - x_{i,2}}{x_{i,2} - x_*} \ln(\bar{y}_0/y)$	–	$[\underline{y}_0, \bar{y}_{0,2}]$	(36), (99)
Matched, left	$\bar{X}(y; \epsilon)$	$X_0^{\text{out},0}(y, 1) + X_0^{xb}(y) - x_f$ $+ X_0^{yb}(y) + \frac{\epsilon}{x_*} \left(\ln \left(\frac{y}{\bar{y}_0} \right) + \text{Ein} \left(\frac{\bar{y}_0}{x_*} \right) \right)$	–	$[\underline{y}_0, \bar{y}_0]$	(126)
Matched, right	$\vec{X}(y; \epsilon)$	$X_0^{\text{out},-1}(y, x_{i,2}) + X_0^{xb}(y; \epsilon) - x_{i,2}$	–	$[\underline{y}_0, \bar{y}_{0,2}]$	(106)

Table 3: Matched solutions for disease invasions $[(x_i, y_i) = (1, 0)]$. The quantities x_* , x_f , \bar{y}_0 , $x_{i,2}$, and $\bar{y}_{0,2}$ are expressed in terms of the parameters \mathcal{R}_0 and ϵ in Table 2. Left and right corner solutions are valid in a neighbourhood of the points $(x_f, 0)$ and $(x_{i,2}, 0)$, respectively (and are identical up to swapping x_f and $x_{i,2}$). Left and right matched solutions are uniformly valid to the left and right of the y nullcline ($x = x_*$). Setting $x_i = 1$ gives the first epidemic wave emanating from the DFE. Asymptotic approximations for the j^{th} epidemic wave (for $j \geq 1$) are obtained by replacing x_i with $x_{i,j}$, x_f by x_f^j , $x_{i,2}$ by $x_{i,j+1}$, \bar{y}_0 by $\bar{y}_{0,j}$, and $\bar{y}_{0,2}$ and $\bar{y}_{0,j+1}$ in the expressions and domains [Equation (110)]. For each of the outer, inner, and left and right corner solutions, the equation reference first is the “raw” expression with undetermined matching constants (the matched values are listed in Table 4) and then the “matched” expression with the matched values of the constants inserted. Grey text is used to emphasize a factor in the x -axis boundary layer approximation and a term in the left matched solution that appear when we include the y -axis boundary layer approximation in the matching; Figure 4 shows results without the grey quantities, whereas Figure 5 shows the improvement obtained by including them.

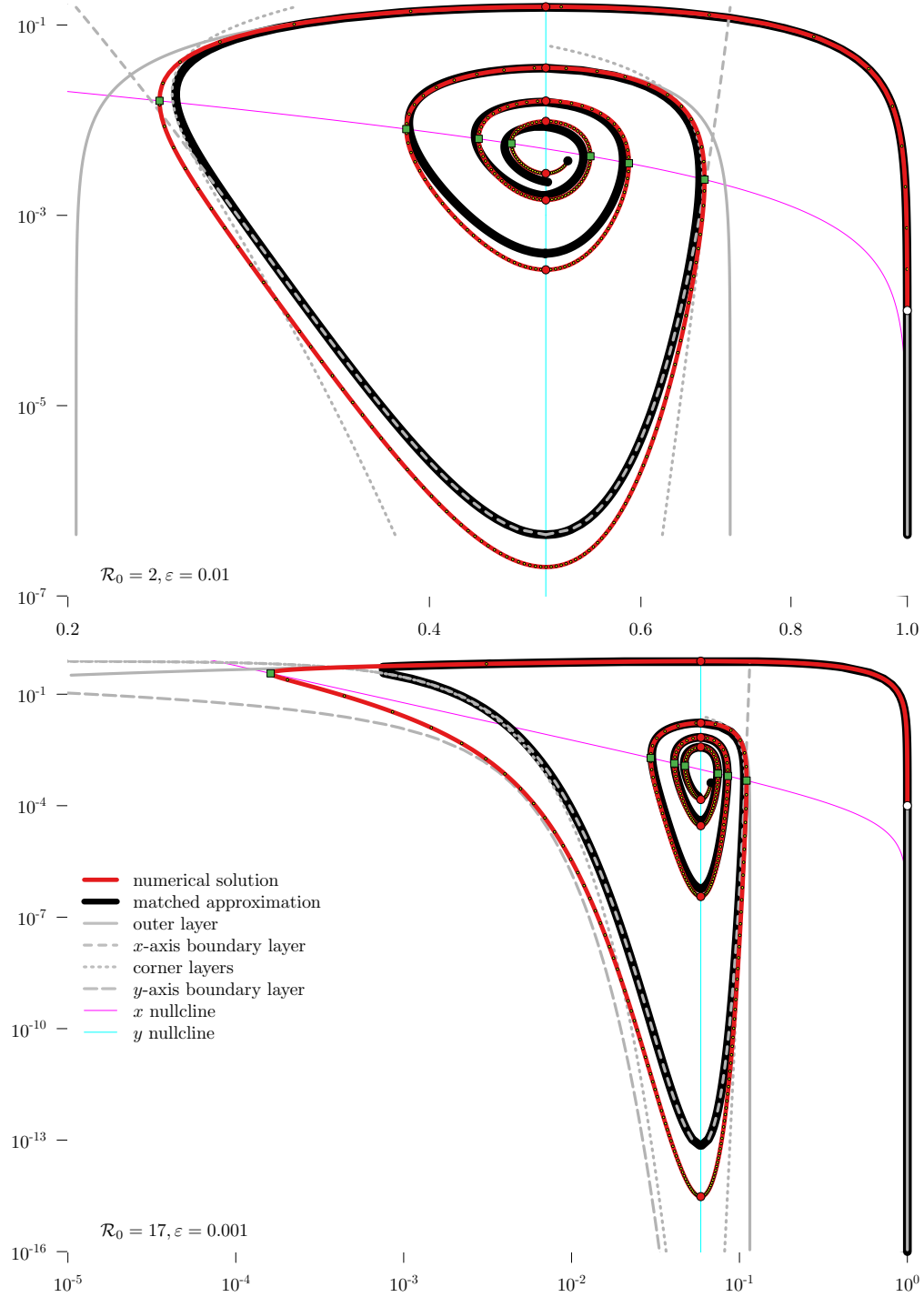


Figure 4: Solutions of the SIR ODEs Equation (1) and approximations (Table 3). *Top panel:* $\mathcal{R}_0 = 2$, $\varepsilon = 0.01$. *Bottom panel:* $\mathcal{R}_0 = 17$, $\varepsilon = 0.001$; similar to measles and whooping cough (Table 1). Various outer and inner approximations are shown in grey, and the matched approximation is in black. The numerically computed solutions are red, as in Figures 1 and 2.

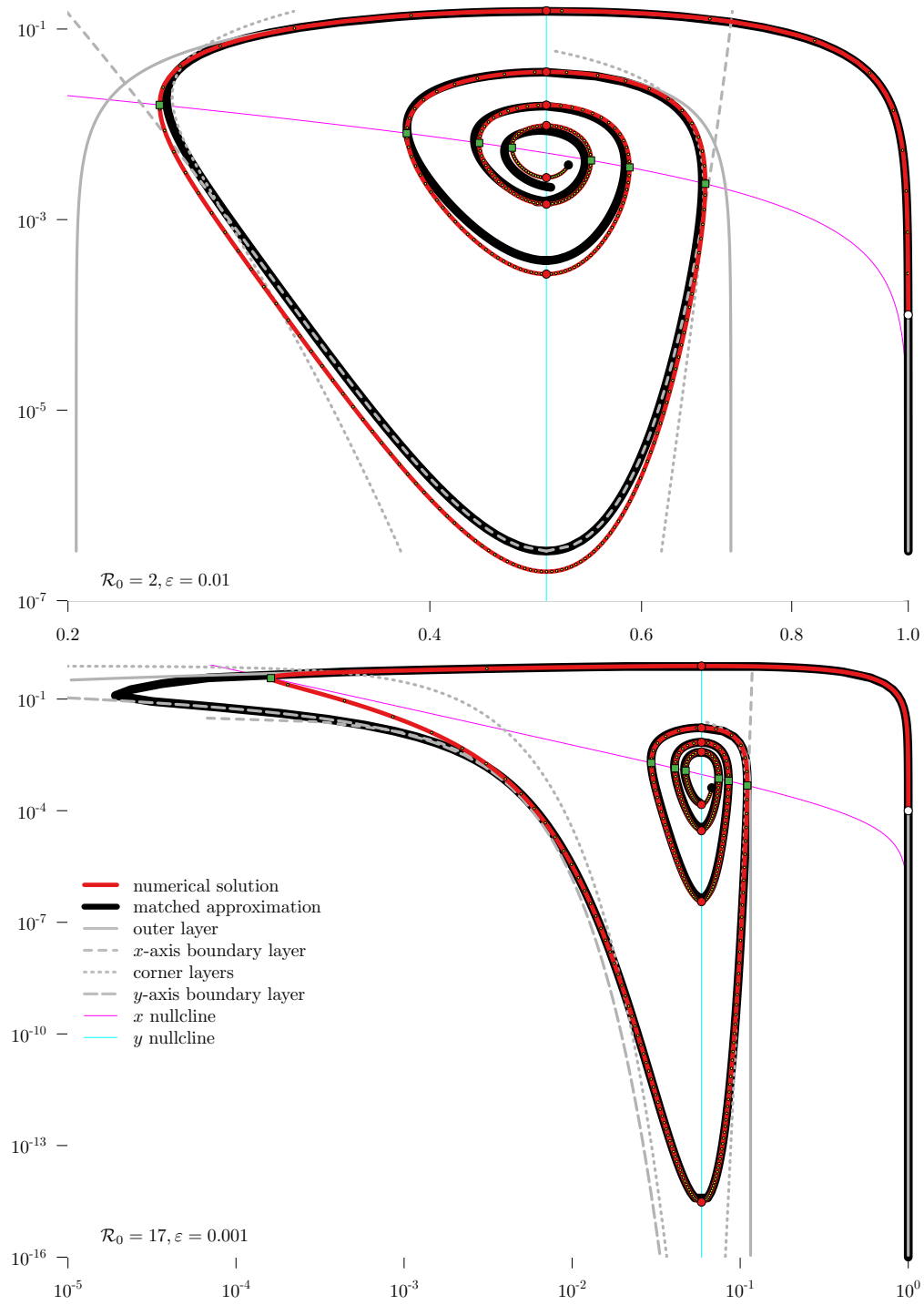


Figure 5: Similar to Figure 4, but including the y -axis boundary layer approximation (Y^b) in the matching.

4 Discussion

Nonlinear differential equations can rarely be solved exactly. Creative analyses leading to approximate analytical solutions were once the only way to study nonlinear systems (see [22, pp. 201–204] has a very concise history), but interest in such approximations has diminished as computers have become more powerful and software for efficient and accurate numerical solution of differential equations has become so easily accessible. However, closed-form analytical expressions can often lead to valuable insights, and can facilitate further analyses that would be impossible or exceedingly challenging to conduct numerically.

We have derived new, fully analytical approximations for the phase plane dynamics of the SIR model with vital dynamics. In Table 2, we list our expressions for key epidemiological quantities, including peaks and troughs of the susceptible and infectious proportions of the host population. We present a closed-form analytical approximation to the Poincaré map for the SIR model in §3.1.4. A highly accurate approximation to the susceptible proportion as the trajectory enters the x -axis boundary layer is given in Equation (117), and is a critical ingredient in a stochastic disease persistence analysis that we present elsewhere [23].

Our approach has involved matching asymptotic expansions across branch cuts of a special function (the Lambert W function, §2.5). To our knowledge, this is the first example of asymptotic matching across branch cuts.

We have focused on the most relevant case of the invasion of a novel pathogen into an epidemiologically naïve population *i.e.*, the trajectory escaping the DFE along its unstable manifold. Nonetheless, the solutions presented in §2 and the matching methodology remain valid for arbitrary initial conditions (x_i, y_i) . Obtaining approximations for any other initial condition would be a matter of determining new matching constants and modifying the expressions accordingly. We have considered only the standard SIR model, but the techniques we have presented can be adapted to other compartmental ODE models, which will explore in further work.

Acknowledgements

This project was partially supported by the CNRS IEA grant “Structured Populations, Epidemics & Control Strategies (SPECS)”. DJDE was supported by an NSERC Discovery Grant.

References

- [1] J. D. ANDERSON, *Ludwig Prandtl’s boundary layer*, Physics Today, 58 (2005), pp. 42–48.
- [2] R. M. ANDERSON AND R. M. MAY, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, Oxford, 1991.
- [3] N. BACAËR, *A Short History of Mathematical Population Dynamics*, Chapter 16: “McKendrick and Kermack on epidemic modelling (1926–1927)”, Springer London, London, 2011.
- [4] N. BACAËR, *The model of Kermack and McKendrick for the plague epidemic in Bombay and the type reproduction number with seasonality*, J. Math. Biol., 64 (2012), pp. 403–422.
- [5] C. M. BENDER AND S. A. ORSZAG, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, New York, 1978.
- [6] D. CHAMPREDON, J. DUSHOFF, AND D. J. D. EARN, *Equivalence of the Erlang SEIR epidemic model and the renewal equation*, SIAM J. Appl. Math., 78 (2018), pp. 3258–3278.
- [7] R. M. CORLESS, G. H. GONNET, D. E. G. HARE, D. J. JEFFREY, AND D. E. KNUTH, *On the Lambert W function*, Adv. Comput. Math., 5 (1996), pp. 329–359.
- [8] P. FLAJOLET AND R. SEDGEWICK, *Analytic Combinatorics*, Cambridge University Press, 2009.

- [9] R. GANI AND S. LEACH, *Epidemiologic determinants for modeling pneumonic plague outbreaks*, Emerg. Infect. Dis., 10 (2004), pp. 608–614.
- [10] D. HE AND D. J. D. EARN, *Epidemiological effects of seasonal oscillations in birth rates*, Theor. Popul. Biol., 72 (2007), pp. 274–291.
- [11] T. D. HOLLINGSWORTH, R. M. ANDERSON, AND C. FRASER, *HIV-1 transmission, by stage of infection*, J. Infect. Dis., 198 (2008), pp. 687–693.
- [12] W. O. KERMACK AND A. G. MCKENDRICK, *A contribution to the mathematical theory of epidemics*, Proc. R. Soc. A, 115 (1927), pp. 700–721.
- [13] J. KEVORKIAN AND J. D. COLE, *Multiple Scale and Singular Perturbation Methods*, Springer-Verlag New York Inc., New York, 1996.
- [14] O. KRYLOVA, *Predicting epidemiological transitions in infectious disease dynamics: Smallpox in historic London (1664-1930)*, PhD, McMaster University, Canada, 2011.
- [15] O. KRYLOVA AND D. J. D. EARN, *Effects of the infectious period distribution on predicted transitions in childhood disease dynamics*, J. R. Soc. Interface, 10 (2013), p. 20130098.
- [16] Z. LEVINE AND D. J. D. EARN, *Face masking and COVID-19: Potential effects of variation on transmission dynamics*, J. R. Soc. Interface, 19 (2022), p. 20210781.
- [17] J. MA AND D. J. D. EARN, *Generality of the final size formula for an epidemic of a newly invading infectious disease*, Bull. Math. Biol., 68 (2006), pp. 679–702.
- [18] A. G. MCKENDRICK, *Applications of mathematics to medical problems*, Proc. Edinburgh Math. Soc., 13 (1926), pp. 98–130.
- [19] C. E. MILLS, J. M. ROBINS, AND M. LIPSITCH, *Transmissibility of 1918 pandemic influenza*, Nature, 432 (2004), pp. 904–906.
- [20] R. OLINKY, A. HUPPERT, AND L. STONE, *Seasonal dynamics and thresholds governing recurrent epidemics*, J. Math. Biol., 56 (2008), pp. 827–839.
- [21] F. W. J. OLVER, D. W. LOZIER, R. F. BOISVERT, AND C. W. CLARK, eds., *NIST Handbook of Mathematical Functions*, National Institute of Standards and Technology (NIST) and Cambridge University Press, New York, 2010.
- [22] R. E. O’MALLEY, JR., *Singular Perturbation Methods for Ordinary Differential Equations*, vol. 89, Springer, 1991.
- [23] T. L. PARSONS, B. M. BOLKER, J. DUSHOFF, AND D. J. D. EARN, *The probability of epidemic burnout in the stochastic SIR model with demography*, in preparation, xx (2023), p. xx.
- [24] L. STONE, R. OLINKY, AND A. HUPPERT, *Seasonal dynamics of recurrent epidemics*, Nature, 446 (2007), pp. 533–536.
- [25] O. A. VAN HERWAARDEN, *Stochastic epidemics: the probability of extinction of an infectious disease at the end of a major outbreak*, J. Math. Biol., 35 (1997), pp. 793–813.
- [26] Z. Y. WONG, C. M. BUI, A. A. CHUGHTAI, AND C. R. MACINTYRE, *A systematic review of early modelling studies of Ebola virus disease in West Africa*, Epidemiol. Infect., 145 (2017), pp. 1069–1094.

A Summary Table of Matching Constants

Constant	Expression	Equations
C_0^{out}	$y_i + x_i - x_* \ln x_i$	(16), (60), (65)
C_0^{cor} (left)	x_f	(33), (36), (77)
C_{\ln}^{cor} (left)	$-\frac{1-x_f}{x_*-x_f}$	(36), (77)
C_1^{cor} (left)	$\frac{1-x_f}{x_*-x_f} \ln \bar{y}_0$	(36), (83), (88),
c_0^ϕ	x_f	(83)
C_0^ϕ	$x_f - (1-x_*) \ln(1-x_f)$	(27), (30), (61), (79), (83)
C_1^ϕ	$\ln \bar{y}_0$	(28), (30), (61), (88)
C_0^{cor} (right)	$x_{i,2}$	(33), (36), (98)
C_{\ln}^{cor} (right)	$-\frac{1-x_{i,2}}{x_*-x_{i,2}}$	(36), (98)
C_1^{cor} (right)	$\frac{1-x_{i,2}}{x_*-x_{i,2}} \ln \bar{y}_0$	(36), (98)
C_0^{out} (2 nd)	$x_{i,2}$	(16), (60), (104)
c_0^{out} (2 nd)	$x_{i,2} - x_* \ln x_{i,2}$	(102), (101), (104)
C_0^{yb}	$-E_1\left(\frac{\bar{y}_0}{x_*}\right)$	(115), (45)
C_1^ϕ (y-axis matching)	$\ln \bar{y}_0 - \text{Ein}\left(\frac{\bar{y}_0}{x_*}\right)$	(28), (30), (61), (123)

Table 4: Matching constants for disease invasions $[(x_i, y_i) = (1, 0)]$. Left and right indicate constants appearing in the left and right corner solutions, while 2nd indicates constants for the second epidemic wave. See Table 2 for x_* , x_f , \bar{y}_0 , and $x_{i,2}$ expressed in terms of \mathcal{R}_0 and ϵ . The final expressions with these values for the matching constants are listed in Table 3.