

A SIR Epidemic Model Allowing Recovery

Anthony G. Pakes

Department of Mathematics & Statistics, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia; tony.pakes@uwa.edu.au

Abstract: The deterministic SIR model for disease spread in a closed population is extended to allow infected individuals to recover to the susceptible state. This extension preserves the second constant of motion, i.e., a functional relationship of susceptible and removed numbers, $S(t)$ and $R(t)$, respectively. This feature allows a substantially complete elucidation of qualitative properties. The model exhibits three modes of behaviour classified in terms of the sign of $-S'(0)$, the initial value of the epidemic curve. Model behaviour is similar to that of the SIS model if $S'(0) > 0$ and to the SIR model if $S'(0) < 0$. The separating case is completely soluble and $S(t)$ is constant-valued. Long-term outcomes are determined for all cases, together with determination of the rate of convergence. Determining the shape of the epidemic curve motivates an investigation of curvature properties of all three state functions and quite complete results are obtained that are new, even for the SIR model. Finally, the second threshold theorem for the SIR model is extended in refined and generalised forms.

Keywords: SIR epidemic model; recovery from infection; curvature properties; threshold phenomena

MSC: 26A06; 26A51; 92D30



Citation: Pakes, A.G. A SIR Epidemic Model Allowing Recovery. *Axioms* **2024**, *13*, 115. <https://doi.org/10.3390/axioms13020115>

Academic Editors: Babak Shiri and Zahra Alijani

Received: 20 December 2023

Revised: 29 January 2024

Accepted: 2 February 2024

Published: 8 February 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The very well known general or SIR model of the spread of a disease in a large population lends itself to a fairly complete qualitative understanding owing to the fact that its fundamental system of three differential equations (DEs) possess two constants of motion. See [1] or [2] for accounts of this topic. The first of these is common to very many models: the size of the population is constant in time. The second constant of motion occurs because the time variable can be eliminated between the DEs for the number of susceptible and removed individuals, $S(t)$ and $R(t)$, respectively, to give a functional relation between them that holds along each system trajectory.

Recall that the SIR model describes progression of susceptible individuals to possible infection and then ultimate removal with no chance of infecting any in the susceptible subpopulation. The mathematically simpler SIS model ([1,2]) allows for recovery of infected individuals back to full susceptibility. These models are frequently named, respectively, the general and the simple epidemic. In this paper, we consider an extension that interpolates between these models. This extension allows for infected individuals, $I(t)$ in number, to either be removed at rate γ or to recover with immediate susceptibility at rate ρ . We assume that $\gamma + \rho > 0$. The immediacy of recovery to a susceptible state can, perhaps, be justified on the assumption that any period of immunity is so brief as to be negligible in comparison with typical periods of infectiousness and susceptibility and that final states are essentially achieved by times that are short relative to demographic changes.

So, assuming infection occurs from homogeneous mixing of infected and susceptible individuals with rate $\beta > 0$, the governing DEs are

$$S'(t) = -\beta S(t)I(t) + \rho I(t), \quad (1)$$

$$I'(t) = \beta S(t)I(t) - \gamma I(t) - \rho I(t), \quad (2)$$

and

$$R'(t) = \gamma I(t). \tag{3}$$

Initial values are denoted by $S_0 > 0$, $I_0 > 0$ and $R_0 = 0$. Thus, if all rate parameters are positive, transitions between compartments occur as for the SIR model with an additional flux of individuals at rate ρ from the infected compartment back to the susceptible compartment.

This model could be denoted by SIRVS, pronounced ‘serves’, and this is, to the author’s knowledge, due to Mulkern and Nosrati [3] (whose notation for the parameters differs). They pursue the analytic consequences of the system equations to the extent of noting the second constant of motion (with a minor error) and illustrate its behaviour with some numerical calculation and graphical displays. The aim of this paper is to derive as far as is possible the analytical consequences of the above system.

Before describing the structure of the paper, we recall the so-called first threshold theorem for the SIR model. It is known for this model that $\lim_{t \rightarrow 0} I(t) = 0$ and $I(t)$ either decreases or initially increases and then decreases. These modes are taken to represent, respectively, a minor or major epidemic. Their occurrence is determined by the sign of the initial rate of infection, $I'(0)$, specifically $I'(0) > 0$ for a major epidemic. These initial conditions are characterised by the first threshold theorem asserting the obvious conclusion from (2) that $I'(0) > 0$ iff $S_0 > \gamma/\beta$. The lower bound is called the relative removal rate and it is the threshold that separates a minor/major outbreak. The threshold criterion is usually expressed in terms of the basic reproduction ratio $\mathcal{R}_0 = \beta S_0/\gamma$, i.e., a major outbreak occurs iff $\mathcal{R}_0 > 1$.

The basic reproduction ratio is similarly defined for other epidemic models in terms of the sign of $I'(0)$, and for the SIRVS model it is clear from (2) that

$$\mathcal{R}_0 = \frac{\beta S_0}{\gamma + \rho}. \tag{4}$$

Restricted to the SIS or SIR models, \mathcal{R}_0 can be interpreted as the average number of secondary infections caused by introducing a single infective into a large susceptible population. This extends to the SIRVS model, as follows. The number of infections, \mathcal{I}_t , by time, t , caused by a single infective arriving at $t = 0$ equals $\mathcal{N}(\beta S_0 t)$, where $\mathcal{N}(\cdot)$ is a unit-rate Poisson process. The duration, T , of infectivity is modelled as a competing risks situation, i.e., $T = \min(\epsilon_\gamma, \epsilon_\rho)$, where ϵ_γ and ϵ_ρ are independent random variables having exponential distributions with parameters γ and ρ and they are independent of $\mathcal{N}(\cdot)$. Hence, T has an exponential distribution with parameter $\gamma + \rho$. It thus follows that

$$E(\mathcal{I}_T) = E[\mathcal{N}(\beta S_0 T)] = \beta S_0 E(T) = \frac{\beta S_0}{\gamma + \rho}.$$

Many of the formulae derived for the SIRVS model have $\gamma > 0$ appearing as a denominator. Hence, the results for the SIS model cannot simply be read off from the general results established in Sections 3–6. In addition, if $\gamma > 0$, then the fundamental property of the SIR model that $I(t) \rightarrow 0$ subsists for the SIRVS model (Lemma 1), whereas the SIS model admits an endemic level of infection under some circumstances. Consequently, we begin in Section 2 by recalling (mostly) known facts about the SIS model. In addition to behaviour determined by the sign of $\mathcal{R}_0 - 1$, where $\mathcal{R}_0 = \beta S_0/\rho$, there are curvature behaviours of $I(t)$ that qualitatively differ according to the sign of the difference $\delta := \beta N - \rho$.

In Sections 3–6, we assume that $\gamma, \rho > 0$, the full SIRVS model, and derive (in Section 3) the second constant of motion, a functional relation between $S(t)$ and $R(t)$, and its consequence for the limiting values S_∞ and R_∞ of susceptible and removed numbers. This reveals three cases (denoted 1, 2 and 3, respectively) according as $\beta S_0 > \rho$, $= \rho$ or $< \rho$, where $S_0 = S(0)$. It is clear from (1) that these cases correspond, respectively, to $S'(0) < 0$, $S'(0) = 0$ and $S'(0) > 0$. All three of these can occur for the SIS model and only the first is

possible for the SIR model. Case 2 is a degenerate case having a complete elementary solution. We also establish the existence of the long-term limits.

We show in Section 4 that the three state functions converge exponentially quickly to their limiting values, with a common time constant, τ , that depends on the three parameters and the final number S_∞ of susceptibles.

In Section 5, we examine the curvature properties of the state functions. To appreciate why, recall that the epidemic curve is defined as the production rate of new infectives, $-S'(t)$. Thus, the maximum c^* of the epidemic curve will occur in $(0, \infty)$ iff $S(t)$ has an infection point therein (Theorem 5 and Lemma 5). This invites a more thorough investigation of second-order properties. Observe too that, e.g., $S''(t)$ measures the instantaneous acceleration of susceptible numbers. In particular, the speeding up, or slowing down, of the state functions corresponds to regions of their convexity, or concavity. Inflection points delineate, e.g., transition times from acceleration to deceleration. Much of this is new, even for the SIR model. Case 3 exhibits the simplest behaviour—there are no inflection points and, quite unlike the SIR model, $S(t)$ is increasing. Case 1 shares behaviour with the SIR model: $S(t)$ decreases and both it and $R(t)$ have at most one inflection and $I(t)$ has one or two inflections.

The second threshold theorem for the SIR model is an assertion of the final size, R_∞ (i.e., total number of removed infected individuals), in the case that S_0 is just a little larger than the threshold, γ/β . More precisely, if $S_0 = \gamma/\beta + \epsilon$, where $0 < \epsilon \ll 1$, and if $I_0 \ll \epsilon^2$, then $R_\infty \approx 2\epsilon$. Equivalently, $S_\infty \approx \gamma/\beta - \epsilon$, i.e., final susceptible numbers end as far below the threshold as they began above it. See [1,2]. This fundamental result, going back to the pioneering work of Kermack and McKendrick, is usually derived by using a Maclaurin expansion of the exponential function occurring in the functional equation relating S_∞ and R_∞ . See, e.g., p. 84 in [1] (where the constraint on I_0 is stated a little obscurely—see the equation just after (6.9)) or p. 28 in [2] (an imprecise assertion in that I_0 is specified as being ‘small relative to ϵ ’) and the proof on pp. 30, 31 where the assumption is $I_0 \ll \epsilon^2$.

In Theorem 8, we follow this mode of proof for the SIRVS model with the precise specification that $I_0 = c\epsilon^2$, where $0 \leq c < \infty$. The consequences of the general outcome show that the general form of the second threshold theorem subsists for the SIRVS model, even if $c > 0$. If $c = 0$, then the above outcome for the SIR model is unchanged, i.e., the parameter ρ is absent from the first-order of approximation. Lemma 6 asserts a consequence of the proof of Theorem 9 in which we allow a larger initial number of infectives: $I_0 \sim C\epsilon^{1+\omega}$, where $\omega \in [0, 1)$. Finally, by using a branch-point expansion of the Lambert function we extend, in Theorem 9, the approximation (34) by estimating the quadratic and cubic components of the $O(\epsilon^2)$ term.

A concluding section ends the paper. The following derived constants frequently occur:

$$\alpha = \beta/\gamma \quad \text{and} \quad a = \rho/\gamma. \tag{5}$$

The first of these is the reciprocal of the more commonly occurring relative removal rate.

2. The SIS Model

If $\gamma = 0$, then $\mathcal{R}_0 = \beta S_0/\rho$ and the DE (2) reduces to a logistic DE because in that case $S(t) + I(t) \equiv N$. Defining $\delta = \beta N - \rho$, its solution is

$$I(t) = \begin{cases} \frac{I_0 e^{\delta t}}{1 + (I_0 \beta / \delta)(e^{\delta t} - 1)} > 0 & \text{if } \delta \neq 0, \\ \frac{I_0}{1 + \beta t} & \text{if } \delta = 0. \end{cases}$$

As is well known and clear from these expressions,

$$\lim_{t \rightarrow \infty} I(t) = \begin{cases} I_\infty = N - \rho/\beta & \text{if } \delta > 0, \\ 0 & \text{if } \delta \leq 0. \end{cases}$$

However, the finer aspects of $I(t)$ (and $S(t)$) also depend on cases analogous to 1–3 specified above.

We observe first that if $\delta \leq 0$ (implying $\mathcal{R}_0 < 1$), then it is easy to check that $I(t)$ is convex-decreasing to zero, and hence $S(t)$ is concave-increasing to N .

The case $\delta > 0$ exhibits richer behaviour, as follows (but omitting algebraic detail). Clearly, $\mathcal{R}_0 = \beta S_0 / \rho > 1$ implies that $\delta > 0$ and that this latter can hold if $\mathcal{R}_0 \leq 1$. The above limit result for infective numbers takes the monotone form: As $t \uparrow \infty$,

$$I(t) \begin{cases} \downarrow I_\infty & \text{if } I_0 > I_\infty, \text{ i.e., } \mathcal{R}_0 < 1; \\ \equiv I_\infty & \text{if } I_0 = I_\infty, \text{ i.e., } \mathcal{R}_0 = 1; \\ \uparrow I_\infty & \text{if } I_0 < I_\infty, \text{ i.e., } \mathcal{R}_0 > 1. \end{cases}$$

Differentiation will show (for all possible values of δ) that $I''(t)$ has the same sign as the product

$$(I_\infty - I_0) \times (I_\infty - I_0 - I_0 e^{\delta t}). \tag{6}$$

Hence, if $\delta > 0$ and $\mathcal{R}_0 < 1$, i.e., $I_0 > I_\infty$, then $I(t)$ is convex-decreasing to I_∞ .

If $\mathcal{R}_0 > 1$, then the second factor in (6) decreases from $I_\infty - 2I_0$ to $-\infty$. Hence, if $2I_0 < I_\infty$, equivalently,

$$\mathcal{R}_0 > 1 + \beta I_0 / \rho, \tag{7}$$

then this second factor has a unique zero $t_i \in (0, \infty)$. We thus conclude: *If $1 < \mathcal{R}_0 \leq 1 + \beta I_0 / \rho$ then $I(t)$ is concave-increasing, and if (7) holds then there is a time of inflection,*

$$t_i = \delta^{-1} \log \frac{\mathcal{R}_0 - 1}{\beta I_0 / \rho} > 0,$$

such that

$$I(t) \text{ is } \begin{cases} \text{convex - increasing in } (0, t_i), \\ \text{concave - increasing in } (t_i, \infty). \end{cases}$$

These results give a precise expression to the idea that the infection initially accelerates if \mathcal{R}_0 is sufficiently larger than unity before necessarily slowing as it approaches its endemic level, I_∞ . Recalling the interpretation of \mathcal{R}_0 as the initial relative per-capita rate of increase of infectives, the dual quantity $\widehat{\mathcal{R}}_0 = \beta I_0 / \rho$ is the initial per-capita relative rate of increase of recoveries. Hence, we have a threshold result for the SIS model asserting that $I(t)$ is initially convex-increasing iff $\mathcal{R}_0 - \widehat{\mathcal{R}}_0 > 1$.

3. The Second Constant of Motion

We assume in the sequel that $\gamma > 0$ and $\rho \geq 0$. Observe that, exactly as for the SIR model, $R'(t) > 0$ for all t and hence $R(t) \uparrow R_\infty$, say, and $R_\infty \leq N$. Next, dividing (3) into (1) yields

$$\frac{dS}{dR} = -\frac{\beta S - \rho}{\gamma}, \tag{8}$$

whose solution for $t > t' > 0$ is

$$\log \left| \frac{\beta S(t) - \rho}{\beta S(t') - \rho} \right| = -\frac{\beta}{\gamma} (R(t) - R(t')).$$

This is valid for all parameter combinations because (3) implies that $R(t)$ is strictly increasing and hence the argument of the logarithm function must be less than unity. Assuming that $R(0+) = 0$, then letting $t' \rightarrow 0$ and exponentiating the result we obtain for $t \geq 0$ that

$$|\beta S(t) - \rho| = |\beta S_0 - \rho| e^{-(\beta/\gamma)(R(t))}. \tag{9}$$

It follows that the sign of $\beta S(t) - \rho$ is the same as the sign of $\beta S_0 - \rho$, i.e., the sign of $-S'(0)$. This invariance distinguishes three cases.

Case 1: $\beta S_0 > \rho \geq 0$. Here, susceptibles are initially being infected at a faster rate than they are recovering. Recalling (5), the solution (9) becomes

$$\beta S(t) = \rho + (\beta S_0 - \rho)e^{-\alpha R(t)}. \tag{10}$$

It is clear that susceptible numbers preserve the classical SIR behaviour: $S(t) \downarrow S_\infty$, say, as $t \uparrow \infty$. Setting $\rho = 0$ yields the long-known relation for the SIR model.

Case 2: $\beta S_0 = \rho$. In this case, we have the degenerate outcome that $\beta S(t) \equiv \rho$ because, independently of the value of γ , the rate of infection is exactly balanced by the rate of recovery to the susceptible state. It then follows that (2) reduces to $I' = -\rho I$, whence

$$I(t) = I_0 e^{-\gamma t}$$

and (3) yields

$$R(t) = I_0(1 - e^{-\gamma t}).$$

Hence, infected numbers decrease convexly to zero and the final size of the epidemic is $R_\infty = I_0$.

Case 3: $\beta S_0 < \rho$, implying that $\rho > 0$. Here, (7) subsists with the result that $S(t) \uparrow S_\infty$ as $t \uparrow \infty$.

We observe in passing that (25) in [3] is a rearrangement of (10), with a minor error in that the first occurrence of γ_2 on their right-hand side should be γ_1 . Observe too that it is not obvious what should result from (10) if $\gamma \rightarrow 0$ because then $\alpha \rightarrow \infty$ and, of course, $R(t) \equiv 0$ in the actual limit value $\gamma = 0$ described in Section 2.

We wish to let $t \rightarrow \infty$ in (10). As previously remarked, $R_\infty = \lim_{t \rightarrow \infty} R(t)$ exists and hence so does $I_\infty = \lim_{t \rightarrow \infty} I(t)$. The following fundamental result generalises the well-known, long-term outcome for the SIR model.

Lemma 1. *If $\gamma > 0$, then the limit $I_\infty = 0$.*

Proof. If $I_\infty > 0$, it follows from (3) that there exists $t' > 0$ such that $R'(t) > \frac{1}{2}I_\infty$ if $t > t'$. This implies that $R(t) \uparrow \infty$, a clear contradiction. \square

This result shows that even the slightest degree of removal prevents an endemic level of infection. In particular, $S_\infty + R_\infty = N$, so taking the limit in (10) yields the functional equation

$$\beta S_\infty - \rho = (\beta S_0 - \rho)e^{-\alpha(N - S_\infty)}. \tag{11}$$

Defining $x = S_\infty - \rho/\beta$ and $x_0 = S_0 - \rho/\beta$, and recalling (5), identity (11) can be recast as

$$-\alpha x e^{-\alpha x} = -\alpha x_0 e^{a - \alpha N}.$$

This functional equation has the form attributed (incorrectly) to J. Lambert, and we write its solution as

$$\beta S_\infty - \rho = -\gamma W\left(-(\alpha S_0 - a)e^{a - \alpha N}\right), \tag{12}$$

where $W(\cdot)$ denotes the principal branch of the Lambert W -function [4]. This is the appropriate choice because, if not, then the left-hand side would be an unbounded function of the argument of $W(\cdot)$. See [5] for the version with $\rho = 0$ and the references there.

Observe that we can let $\gamma \rightarrow 0$ in (11) because $\alpha \rightarrow \infty$, which yields $S_\infty \rightarrow \rho/\beta$, agreeing with the analytical outcome for the SIS model. The explicit solution (12) is also consistent in this regard. To see this, observe for the SIS case that $\rho < \beta N$, implying that the argument of $W(\cdot)$ tends to zero as $\gamma \rightarrow 0$. In fact, we have the approximation

$$S_\infty = \rho/\beta + (\beta S_0 - \rho)e^{-(\rho - \beta N)/\gamma}(1 + o(1)),$$

as $\gamma \rightarrow 0$.

Another consequence of (12) is that

$$S_\infty \begin{cases} > \rho/\beta & \text{in Case 1,} \\ < \rho/\beta & \text{in Case 3.} \end{cases}$$

4. The Rate of Convergence

We begin with a lemma that introduces the convergence exponent, which we denote by τ .

Lemma 2. *The function $I(t)$ is log-concave in Case 1 and it is log-convex (hence convex) in Case 3. In addition, the following limit exists:*

$$\tau := - \lim_{t \rightarrow \infty} I'(t)/I(t) = -(\beta S_\infty - \gamma - \rho).$$

Proof. Simply observe from (2) that

$$\frac{d^2}{dt^2} \log I(t) = \beta S'(t)$$

and recall implications from Section 3. \square

Next, we establish that $\tau > 0$.

Lemma 3. *In all cases $\tau > 0$, equivalently,*

$$\frac{\beta S_\infty}{\gamma + \rho} < 1.$$

Proof. For Cases 2 and 3, it follows from (11) that $\beta S_\infty - \rho \leq 0$ and hence that $\tau = \gamma - (\beta S_\infty - \rho) > 0$. Assume now that Case 1 holds and recast (12) as

$$\beta S_\infty - \rho + \gamma W\left(-(\alpha S_0 - a)e^{a-\alpha N}\right) = 0.$$

The argument of the Lambert function is negative, and hence $W(\cdot) \geq -1$, implying that $0 \geq \beta S_\infty - \rho - \gamma$, i.e., $\tau \geq 0$. This inequality is strict provided that the positive-valued term $T = (\alpha S_0 - a)e^{a-\alpha N} < e^{-1}$, but $S_0 < N$ and hence $T < (\alpha N - a)e^{-(\alpha N - a)} \leq e^{-1}$. \square

In what follows, we need only, and will, consider Cases 1 and 3. It follows from (10) and its limiting form that

$$\begin{aligned} \beta(S_t - S_\infty) &= (\beta S_0 - \rho)\left(e^{-\alpha R(t)} - e^{-\alpha R_\infty}\right) \\ &= (\beta S_0 - \rho)e^{-\alpha R_\infty}\left(e^{-\alpha(R(t) - R_\infty)} - 1\right) \\ &= (\beta S_\infty - \rho)(-\alpha(R(t) - R_\infty))(1 + o(1)), \end{aligned}$$

i.e.,

$$S(t) - S_\infty \sim A(R_\infty - R(t)),$$

where

$$0 \neq A = \frac{\beta S_\infty - \rho}{\gamma} = 1 - \tau/\gamma < 1.$$

Hence, $S(t)$ and $R(t)$ approach their limits at the same rate. In addition,

$$I(t) = N - S(t) - R(t) \sim (1 - A)(R_\infty - R(t)). \tag{13}$$

Hence, the state variables approach their limits at the same rate and hence it suffices to determine how quickly $I(t) \rightarrow 0$.

Theorem 1. *If $\gamma > 0$, then*

$$0 < \lim_{t \rightarrow \infty} e^{\tau t} I(t) < \infty.$$

Proof. Let $\ell(t) = e^{\tau t} I(t)$ and observe from Lemma 2 that

$$\ell'(t) = \tau \ell(t) + I'(t)e^{\tau t} = \varepsilon(t)\ell(t),$$

where

$$\varepsilon(t) = \tau + I'(t)/I(t) \rightarrow 0 \tag{14}$$

as $t \rightarrow \infty$.

We infer from the monotone nature of $S(t)$ and (2) that $I(t)$ has at most one critical point in $(0, t)$, and if such exists then it must be at a global maximum. Hence, there exists $t' > \text{Argmax}(I(t))$ such that $I'(t) < 0$ if $t > t'$. The first equality in (14) is a differential equation whose solution is

$$I(t) = \ell(t') \exp \left[-\tau t + \int_{t'}^t \varepsilon(v) dv \right]. \tag{15}$$

It follows from Lemma 2 and (2) that

$$\varepsilon(t) = \beta(S(t) - S_\infty) \sim \text{const} \cdot I(t), \quad (t \rightarrow \infty).$$

Next, it follows from (3) that $\int_0^\infty I(t) dt = R_\infty / \gamma < \infty$ and hence that $\int_{t'}^\infty |\varepsilon(t)| dt < \infty$. The assertion now follows from (15). \square

Approaching the convergence rate through $I(t)$ seems to be the simplest route. To appreciate this, we could begin by defining $\delta(t) = R_\infty - R(t)$ and observe that (3) yields

$$\delta'(t) = -\gamma(N - S(t) - R(t)) = -\gamma(S_\infty - S(t) + R_\infty - R(t)) \sim -\tau \delta(t),$$

where we have appealed to (13) and Lemma 3. The problem thus reduces to estimating $\tau + \delta'(t)/\delta(t)$. This looks to be less straightforward than the above proof.

5. Shape Properties

We now investigate the convexity/concavity properties of the state functions, beginning with the simplest case.

Theorem 2. *Assume Case 3, $\beta S_0 < \rho$. Then the functions $S(t)$ and $R(t)$ are concave-increasing and $I(t)$ is convex decreasing.*

Proof. Knowing that $S(t) \uparrow S_\infty$ in Case 3, it follows from the definition of τ and Lemma 3 that $\beta S(t) - \gamma - \rho < 0$ and hence that $I'(t) < 0$ for all $t > 0$. We conclude from (3) that $R''(t) = \gamma I'(t) < 0$.

Next, twice differentiating (10) yields

$$-\beta S''(t) = \alpha(-R''(t) + R'(t))(\rho - \beta S(t)).$$

Each factor on the right-hand side is positive, and hence $S''(t) < 0$. The convexity assertion for $I(t)$ is now obvious. \square

We now turn to Case 1. Recall that this case can occur if $\mathcal{R}_0 \leq 1$ and if $\mathcal{R}_0 > 1$. The next result concerns the first of these possibilities. We define the constant

$$R = \frac{a}{1+a} = \frac{\rho}{\gamma + \rho} < 1$$

Recalling the stochastic model of infection from Section 1, we observe that $R = P(\epsilon_\rho < \epsilon_\gamma)$, the probability of recovery occurring before removal. In relation to parts (ii) and (iii) of the next result, note that

$$\mathcal{R}_0 - R = \frac{\beta S_0 - \rho}{\gamma + \rho} > 0.$$

The following theorem embraces all three state functions.

Theorem 3. Assume Case 1, $\beta S_0 > \rho$, and that $\mathcal{R}_0 \leq 1$. The following hold.

(i) $I(t)$ is decreasing and it is ultimately convex-decreasing.

(ii) If

$$I_0 \leq S_0 \frac{(1 - \mathcal{R}_0)^2}{\mathcal{R}_0(\mathcal{R}_0 - R)}, \tag{16}$$

then $I(t)$ is convex-decreasing in $(0, \infty)$.

(iii) If

$$I_0 > S_0 \frac{(1 - \mathcal{R}_0)^2}{\mathcal{R}_0(\mathcal{R}_0 - R)}, \tag{17}$$

then $I(t)$ is concave-convex with a unique inflection point, t_i , which solves

$$(\mathcal{R}_0 S(t) - S_0)^2 = \mathcal{R}_0 I(t) (\mathcal{R}_0 S(t) - a S_0). \tag{18}$$

(iv) $R(t)$ is concave-increasing and $S(t)$ is convex-decreasing.

Proof. Observe first that, since $S(t)$ is decreasing and $\mathcal{R}_0 \leq 1$,

$$\beta S(t) - \gamma - \rho < \beta S_0 - \gamma - \rho \leq 0,$$

i.e., $I'(t) < 0$ for all $t > 0$. The concavity of $R(t)$ follows because, from (3), $R''(t) = \gamma I'(t) < 0$. A double differentiation of (10) yields the identity

$$S''(t) = (\beta S_0 - \rho) e^{-\alpha R(t)} \left(-\alpha R''(t) + (\alpha R'(t))^2 \right) > 0,$$

and (iv) follows.

We now consider $I(t)$. Differentiating (1) and using (2) to eliminate first-order derivatives on the right-hand side yields the identity

$$I''(t) = I(t) \left[(\beta S(t) - \gamma - \rho)^2 - \beta I(t) (\beta S(t) - \rho) \right]. \tag{19}$$

The right-hand side is asymptotically equal to $[\beta S_\infty - \gamma - \rho]^2 I(t) > 0$, and Assertion (i) follows.

It follows from (19) that $I(t)$ is convex or concave within intervals determined by the sign of the right-hand side of (19). In addition, such intervals are separated by points of inflection that are solutions of (18), and this equation is obtained by dividing (19) by $\gamma + \rho$ and equating the result to zero. We now show that (18) has at most one solution. As a function of t , the right-hand side of (18) is strictly decreasing to zero. So, its maximum value occurs at $t = 0$.

The derivative of the left-hand side is

$$\mathcal{R}_0 S'(t)(\mathcal{R}_0 S(t) - S_0) = \frac{\mathcal{R}_0 S_0 S'(t)}{\gamma + \rho} (\beta S(t) - \gamma - \rho).$$

However, $S(t)$ is strictly decreasing for Case 1, hence $S'(t) < 0$. In addition, the second factor on the right-hand side decreases from

$$\beta S_0 - \gamma - \rho = (\gamma + \rho)(\mathcal{R}_0 - 1) \leq 0.$$

Hence, $S'(t)(\mathcal{R}_0 S(t) - S_0) > 0$, i.e., the left-hand side of (19) increases from $S_0^2(\mathcal{R}_0 - 1)^2$ up to $(\mathcal{R}_0 S_\infty - S_0)^2 > 0$. It follows that (19) either has no solution in $(0, \infty)$, i.e., $I''(t) > 0$, or it has exactly one solution in $(0, \infty)$. The latter case holds if (17) holds. All assertions now are evident. \square

It follows from Theorem 3 (iii) that, although infected numbers decrease, if I_0 is sufficiently large then this decrease is initially quite slow—a concave decrease—before gathering speed and then slowing again. Observe that for the SIR model, $\rho = 0$, Case 1 is the only possibility, and then, e.g., (17) takes the form

$$I_0 > S_0(\mathcal{R}_0^{-1} - 1)^2.$$

Hence, the concave–convex behaviour is achieved if \mathcal{R}_0 is sufficiently close to unity.

We now consider the case $\mathcal{R}_0 > 1$ (implying Case 1), i.e., $I(t)$ is increasing for small t . The next result extends a fundamental result for the SIR model.

Theorem 4. *If $\mathcal{R}_0 > 1$, then the function $I(t)$ has a unique maximum, I_{\max} , at the point $t_m > 0$, a solution of $\beta S(t) = \gamma + \rho$, i.e.,*

$$\mathcal{R}(t) := \frac{\beta S(t)}{\gamma + \rho} = 1. \tag{20}$$

In addition, the maximum number of infectives is

$$I_{\max} = N - \frac{S_0}{\mathcal{R}_0} \left(1 + \frac{\log((1+a)\mathcal{R}_0 - a)}{1+a} \right) \tag{21}$$

and

$$S(t_m) = S_0/\mathcal{R}_0.$$

Proof. It follows from (2) that $I'(t) = 0$ for any t solving (20). In addition, $\mathcal{R}(0) = \mathcal{R}_0 > 1$ by hypothesis, and since $S(t) \downarrow S_\infty$, it follows from Lemma 3 that $\mathcal{R}(t) \downarrow \beta S_\infty/(\gamma + \rho) < 1$. Hence, (20) has a unique positive solution, t_m . Evaluation (21) follows from $I_{\max} = N - S(t_m) - R(t_m)$, (20), (10) and the definition of \mathcal{R}_0 . The final assertion follows from (2). \square

Remark 1. *Note the simplification of (21) for the SIR model where $a = 0$.*

The recovered numbers Equation (3) implies that $R''(t) = \gamma I'(t)$, resulting in the following consequence.

Lemma 4. *If $\mathcal{R}_0 > 1$, then the removed numbers function,*

$$R(t) \text{ is } \begin{cases} \text{convex - increasing in } (0, t_m) \\ \text{concave - increasing in } (t_m, \infty). \end{cases}$$

The susceptible numbers function behaves in a more complicated manner.

Theorem 5. (a) If $\mathcal{R}_0 > 1$, then the function $S(t)$ is:

(a.i) log-concave and decreasing in $(0, t_m)$, and

(a.ii) convex-decreasing in (t_m, ∞) .

(b) In addition:

(b.i) If

$$I_0 \geq (1 - \mathcal{R}_0^{-1})S_0, \tag{22}$$

then $S(t)$ is convex-decreasing in $(0, \infty)$; and

(b.ii) If

$$I_0 < (1 - \mathcal{R}_0^{-1})S_0, \tag{23}$$

then there is an inflection point $t_{si} \in (0, t_m)$, such that

$$S(t) \text{ is } \begin{cases} \text{concave - decreasing in } (0, t_{si}), \\ \text{convex - decreasing in } (t_{si}, \infty). \end{cases}$$

Proof. Twice differentiating the logarithm of each side of (10) gives the identity

$$\frac{d^2}{dt^2} \log S(t) = -\phi_1(t)R''(t) - \phi_2(t)(R'(t))^2,$$

where the $\phi_i(t)$ are positive-valued functions in $(0, \infty)$. Assertion (a.i) follows from Lemma 4.

Twice differentiating (10) yields the identity

$$\beta S''(t) = -\alpha(\beta S_0 - \rho)e^{-\alpha R(t)}B(t), \text{ where } B(t) = R''(t) - \alpha(R'(t))^2. \tag{24}$$

Assertion (a.ii) follows again from Corollary 5.1.

Next, it follows from (2) and (3) that

$$B(t) = \gamma I(t)[\beta S(t) - \gamma - \rho - \beta I(t)].$$

The right-hand side is decreasing in $(0, t_m)$. Moreover,

$$B(t_m) = -\beta\gamma I^2(t) < 0,$$

implying by continuity and (24) that $S(t)$ is convex-decreasing in an interval (t'', ∞) where $0 \leq t'' < t_m$.

Now,

$$S_0 - \gamma - \rho - \beta I_0 = (\gamma + \rho)(\mathcal{R}_0 - 1 - \mathcal{R}_0 I_0 / S_0).$$

The right-hand side ≤ 0 if (22) holds, and hence $B(t) < 0$ in $(0, t_m)$ and Assertion (b.i) follows.

If (23) holds, then $B(0) < 0$ and, since $B(t_m) < 0$ and $B(t)$ is decreasing, it follows that there exists exactly one number $t_{si} \in (0, t_m)$ such that $S''(t_{si}) = 0$, and (b.ii) follows. \square

It follows from Theorem 5 that the epidemic curve has its maximum at $t = 0$ if (22) holds and the maximum is at $t_{si} < t_m$ if (23) holds. The values of $c^* := \max_{t \geq 0} (-S'(t))$ are stated in our next result.

Lemma 5. If (22) holds, then

$$c^* = -S'(0) = I_0(\beta S_0 - \rho),$$

and if (23) holds, then

$$c^* = I(t_{si})(\beta S(t_{si}) - \rho) = I(t_{si})(\gamma - \beta I(t_{si})).$$

These outcomes make intuitive sense in as much as they imply that if I_0/S_0 is large (which is unlikely) then susceptible numbers initially fall at maximum speed, whereas if I_0/S_0 is less than the critical number $1 - \mathcal{R}_0^{-1}$ then the epidemic begins more slowly and, in terms of the loss of susceptibles, it gathers pace until the time t_{si} , after which it slows down. Observe too that the condition $B(t_{si}) = 0$ is equivalent to

$$R(t_{si}) = N - S_0/\mathcal{R}_0.$$

It then follows from (10) that

$$\beta S(t_{si}) = \rho + (\beta S_0 - \rho)e^{-\alpha(N - S_0/\mathcal{R}_0)}$$

and hence the second evaluation in Lemma 5 is made explicit because

$$I(t_{si}) = N - S(t_{si}) - R(t_{si}) = S_0/\mathcal{R}_0 - S(t_{si}).$$

Next, still assuming $\mathcal{R}_0 > 1$, we look at the shape of the infective curve in the interval $(0, t_m)$. Recall that the sign of $I''(t)$ is the same as the sign of

$$\sigma(t) := (\beta S(t) - \gamma - \rho)^2 - \beta I(t)(\beta S(t) - \rho). \tag{25}$$

Theorem 6. Assume that $\mathcal{R}_0 > 1$ and $0 \leq t \leq t_m$.

(i) If $\sigma(0) \leq 0$, then $I(t)$ is concave-increasing in $(0, t_m)$.

(ii) If $\sigma(0) > 0$, then there is a single point of inflection, $t_{ii} \in (0, t_m)$, such that

$$I(t) \text{ is } \begin{cases} \text{convex - increasing in } (0, t_{ii}), \\ \text{concave - increasing in } (t_{ii}, t_m). \end{cases}$$

Proof. The proof relies on determining how $\sigma(t)$ behaves along the trajectory of $S(t)$ in $(0, t_m)$. We know from (8) that

$$\frac{dR}{dS} = -\frac{\gamma}{\beta S - \rho}$$

and that the right-hand side < 0 in $(0, \infty)$. Hence,

$$\frac{dI}{dS} = -1 + \frac{\gamma}{\beta S - \rho} = \frac{\beta S - \gamma - \rho}{\beta S - \rho}$$

and the right-hand side is positive for $t \in (0, t_m)$.

It follows from (25) that

$$\frac{d\sigma}{dS} = \beta[3(\beta S - \gamma - \rho) - \beta I]. \tag{26}$$

Differentiating again yields

$$\frac{d^2\sigma}{dS^2} = \beta^2 \left[3 + \frac{\beta S - \gamma - \rho}{S - \rho} \right], \tag{27}$$

which is positive for $t \in [0, t_m]$, i.e., $\sigma(t)$ is convex in $S(t)$ as t traverses $[0, t_m]$, but $\sigma(t_m) = -\beta^2 I_{\max} < 0$. Assertions (i) and (ii) follow. In particular, if $\sigma(0) > 0$, then t_{ii} is the unique number in $(0, t_m)$ at which $\sigma(t)$ vanishes. \square

Remark 2. We note that $(d\sigma/dS)|_{t=t_m} = -\beta^2 I_{\max} < 0$.

Remark 3. The condition (22) for $S(t)$ to be convex-decreasing is equivalent to $\beta I_0 \geq \beta S_0 - \gamma - \rho$. Hence,

$$(\beta S_0 - \rho)\beta I_0 \geq (\beta S_0 - \rho)(\beta S_0 - \gamma - \rho) > (\beta S_0 - \gamma - \rho)^2,$$

implying that $I(t)$ is concave-increasing in $(0, t_m)$.

Remark 4. Condition (23) implies that $\sigma(t_{si}) < 0$, i.e., $I(t)$ is concave-increasing in a neighbourhood of t_{si} and hence $t_{si} > t_{ij}$.

Finally, consider the shape of $I(t)$ in (t_m, ∞) . We have shown it is concave around t_m and ultimately convex-decreasing. Hence, there is at least one inflection point in (t_m, ∞) . Recalling that $\beta S(t) < \gamma + \rho$ if $t > t_m$, it follows from (26) that $d\sigma/dS < 0$, i.e., that $\sigma(t)$ is increasing in (t_m, ∞) . In addition, it follows from (25) that $\sigma(t_m) < 0$ and $\sigma(\infty-) = \tau^2 > 0$. Hence, $\sigma(t)$ vanishes at exactly one point $t'_{ii} > t_m$. This yields our final shape result.

Theorem 7. If $\mathcal{R}_0 > 1$, then there exists a unique critical point $t'_{ii} \in (t_m, \infty)$, such that

$$I(t) \text{ is } \begin{cases} \text{concave - decreasing in } (t_m, t'_{ii}) \\ \text{convex - decreasing in } (t'_{ii}, \infty). \end{cases}$$

6. Threshold Theorems

The following result extends the well-known ‘second’ threshold theorem (p. 84 in [1] or p. 29 in [2]), which dates back to the pioneering work of Kermack and McKendrick in the 1920s. Our version is expressed in a slightly more precise manner than in [2]. Let $r = (\gamma + \rho)/\beta$, the relative rate of removal or recovery and recall notation (4).

Theorem 8. Let $\epsilon \in (0, 1)$,

$$S_0 = r + \epsilon \tag{28}$$

(implying $\mathcal{R}_0 = 1 + \epsilon/r$) and let $I_0 \sim c\epsilon^2$ ($\epsilon \rightarrow 0$), where $c \in [0, \infty)$ is a constant. Then the final size R_∞ satisfies

$$R_\infty \sim A(\epsilon) := \frac{\epsilon}{1 + \alpha\epsilon} \left[1 + \sqrt{1 + (2c/\alpha)(1 + \alpha\epsilon)} \right]. \tag{29}$$

Proof. Observe first that (28) is equivalent to

$$\alpha S_0 - a = 1 + \alpha\epsilon. \tag{30}$$

Dividing (11) by γ renders it as

$$\alpha S_\infty - a = \alpha(S_0 + I_0 - R_\infty) - a = (1 + \alpha\epsilon)e^{-\alpha R_\infty}.$$

Writing $y = \alpha R_\infty$ and expanding the exponential term in powers of y yields, after cancellation, the identity

$$(1 + \alpha\epsilon)y^2 - 2\alpha\epsilon y - 2\alpha c\epsilon^2 + O(y^3) = 0. \tag{31}$$

This implies that $y = O(\epsilon)$, which, neglecting the $O(y^3)$ term and solving the resulting quadratic equation, yields the exact result $R_\infty = A(\epsilon) + O(\epsilon^3)$. \square

Taking account of the last step of the proof, it follows from (29) that

$$R_\infty = \left(1 + \sqrt{1 + 2c\gamma/\beta} \right) \epsilon + O(\epsilon^2). \tag{32}$$

Not surprisingly, if $c > 0$ then the number removed exceeds the initial excess, ϵ , of susceptibles. What is surprising is that this dominant contribution is independent of the recovery rate, ρ .

If $c = 0$, then (29) simplifies to the approximation

$$R_\infty \approx \frac{2\epsilon}{1 + \alpha\epsilon} = \frac{2\epsilon}{\alpha S_0 - a}. \tag{33}$$

Using (28) and defining $\bar{\gamma} = \gamma/\beta$ and $\bar{\rho} = \rho/\beta$, respectively, the relative removal and recovery rates, then this approximation takes the form

$$R_\infty \approx 2\bar{\gamma} \left(1 - \frac{\bar{\gamma}}{S_0 - \bar{\rho}} \right)$$

which reduces to a classical form when $\rho = 0$; e.g., (2.3.14) in [2].

It also follows from (33), or (38) with $c = 0$, that

$$R_\infty = 2\epsilon + O(\epsilon^2) \tag{34}$$

and, since $S_\infty = N - R_\infty = S_0 - R_\infty + O(\epsilon^2)$, we obtain the approximation

$$S_\infty = r - \epsilon + O(\epsilon^2) \tag{35}$$

which generalises the known result for the SIR model. In particular, we still observe the incremental change $S_0 - S_\infty = R_\infty \approx 2\epsilon$.

The following result admitting a relatively larger value of I_0 than in Theorem 8 is a corollary of its proof. It asserts a substantially stronger progression of the epidemic than seen in Theorem 8.

Lemma 6. *Let $\epsilon \in (0, 1)$ and (28) hold. If $I_0 \sim C\epsilon^{1+\omega}$, where $C > 0$ and $\omega \in [0, 1)$, then*

$$R_\infty \sim \sqrt{\frac{2C}{\alpha}} \cdot \epsilon^{\frac{1}{2}(1+\omega)}.$$

Refinements to (35) can in principle be obtained by using more terms in the expansion of e^{-y} , giving rise to polynomial equations of an order exceeding two. A more direct approach is using (12) and observing that outcome (35) corresponds to expanding the Lambert function around its branch point at $-e^{-1}$. We specify the following constants:

$$c_1 = \sqrt{1 + 2c\gamma/\beta}, \quad d_1 = \frac{2\alpha}{3c_1^2} \quad \text{and} \quad d_2 = \frac{\alpha^2}{2c_1^2}.$$

Theorem 9. *If the assumptions of Theorem 8 hold, then*

$$R_\infty = A_1\epsilon + A_2\epsilon^2 + A_3\epsilon^3 + O(\epsilon^4), \tag{36}$$

where

$$A_1 = 1 + c_1, \quad A_2 = \frac{1}{3} \left(c - \alpha - \frac{\alpha}{c_1} \right) \quad \text{and} \quad A_3 = \frac{\alpha^2}{36} \left(\frac{7}{c_1} + c_1^3 + 8 \right).$$

Proof. The required Lambert function expansion is

$$W\left(-e^{-1-z^2/2}\right) = 1 - \sigma(z), \tag{37}$$

where

$$\sigma(z) = z - \frac{z^2}{3} + \frac{z^3}{36} + \frac{z^4}{270} + \dots \tag{38}$$

See #4.13.6 in [4], and [5] for a survey of Lambert function expansions. Write (12) in the form

$$\alpha R_\infty = \alpha N - a + W\left(-(\alpha S_0 - a)e^{a-\alpha N}\right). \tag{39}$$

Recalling (30), we have

$$a - \alpha N = a - \alpha S_0 - \alpha I_0 = -1 - \alpha\epsilon - \alpha I_0,$$

and hence in (38) we choose (as expected)

$$z^2 = 2[\alpha\epsilon + \alpha I_0 - \log(1 + \alpha\epsilon)] = O(\epsilon^2).$$

Recalling the assumptions and notation of Theorem 8, expanding the log-term yields

$$z^2 = (\alpha c_1 \epsilon)^2 [1 - d_1 \epsilon + d_2 \epsilon^2 + O(\epsilon^3)].$$

It follows that

$$z = \alpha c_1 \epsilon \left[1 - \frac{1}{2} d_1 \epsilon - \frac{1}{2} (d_2 - d_1^2/4) \epsilon^2 + O(\epsilon^3) \right],$$

whence

$$z^3 = (\alpha c_1 \epsilon)^3 + O(\epsilon^4).$$

Collecting powers of ϵ and substituting into (39) using (37) and (38) leads to Assertion (36). □

Observe that the leading term corresponds to (38). In the classical case, $c = 0$, (36) simplifies to

$$R_\infty = 2\epsilon - \frac{2\alpha}{3}\epsilon^2 + \left(\frac{2\alpha}{3}\right)^2 \epsilon^3 + O(\epsilon^4),$$

the same form as (4.4) in [5] (with differing notation) derived for the SIR model.

7. Concluding Remarks

We have examined the SIR model modified to allow recovery from the infected to the susceptible state at per-capita rate ρ . If $\gamma = 0$, then this is the SIS model that exhibits an endemic level of infection if the derived parameter $\delta > 0$. On the other hand, if $\gamma > 0$, then $I(t) \rightarrow 0$ —the epidemic ultimately fades. The extended model behaves similarly to the subcritical SIS model ($\delta \leq 0$) if ρ is sufficiently large (i.e., $\rho > \beta S_0$, implying $\mathcal{R}_0 < 1$), in that $S(t)$ and $R(t)$ both increase to positive-valued limits (Theorem 2).

Behaviour similar to the SIR model is manifested if ρ is small, i.e., $0 \leq \rho < \beta S_0$. If also $\mathcal{R}_0 \leq 1$, then $I(t)$ decreases, possibly with an inflection. The other state functions are monotone with no inflections (Theorem 3). If $\mathcal{R}_0 > 1$, then curvature properties are more complicated, as seen in Theorems 4–7. In particular, $I(t)$ has a single positive mode at t_m with exactly one inflection in (t_m, ∞) , and it may, or may not, have an inflection in $(0, t_m)$.

Elucidating these shape properties depends on the existence of two constants of motion—the total population size is constant the the relation (10) holds. The model can be generalised at the expense of one or both of these invariants by allowing for birth and death. In this case, it is usual to assume that newborns are susceptible. The first invariant is preserved under the assumption of balanced growth—the birth and death rates are equal. If they are not, then the population size either grows or diminishes exponentially fast. In either case, a successful analysis probably will be limited to identifying equilibria and determining their stability using results from the qualitative theory of differential equations. The case of balanced growth for the SIS and SIR models is surveyed in [6], and some more general models are treated there. The review paper [7] lists papers extending classical models to allow varying total population sizes.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Bailey, N. *The Mathematical Theory of Infectious Diseases*, 2nd ed.; Charles Griffin: London, UK, 1975.
2. Daley, D.; Gani, J. *Epidemic Modelling: An Introduction*; C.U.P.: Cambridge, UK, 1999.
3. Mulkern, R.V.; Nosrati, R. A tutorial on common differential equations and solutions useful for modeling epidemics like COVID-19: Linear and non-linear compartment models. *J. Appl. Math. Phys.* **2022**, *10*, 3053–3071. [[CrossRef](#)]
4. Olver, F.; Lozier, D.; Boisvert, R.; Clark, C. *NIST Handbook of Mathematical Functions*; C.U.P.: Cambridge, UK, 2010.
5. Pakes, A.G. Lambert's W meets the Kermack-McKendrick epidemics. *IMA J. Appl. Math.* **2015**, *80*, 1368–1386. [[CrossRef](#)]
6. Hethcote, H.W. Qualitative analysis of communicable disease models. *Math. Biosci.* **1976**, *28*, 335–356. [[CrossRef](#)]
7. Hethcote, H.W. The mathematics of infectious diseases. *SIAM Rev.* **2000**, *42*, 599–653. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.