Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2025, 13(8), 2578-2592
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

The Spreading Profile for an Emerging Infectious Disease and Its Resemblance to the KdV Solution

Zach Fendler, Valipuram S. Manoranjan*

Department of Mathematics, Washington State University, Pullman, USA
Email: *mano@wsu.edu

How to cite this paper: Fendler, Z. and Manoranjan, V.S. (2025) The Spreading Profile for an Emerging Infectious Disease and Its Resemblance to the KdV Solution. *Journal of Applied Mathematics and Physics*, **13**, 2578-2592.
https://doi.org/10.4236/jamp.2025.138146

Received: June 6, 2025
Accepted: August 15, 2025
Published: August 18, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Open Access

Abstract

Modern times have shown that understanding infectious diseases is imperative to our survival. In this study, we gain an understanding of the SIR model for novel or emerging infectious diseases by deriving time-dependent solutions for the spreading profiles using a second-order approximation. We noticed that we got a solution that resembles the well-known soliton solution of the Korteweg-de Vries (KdV) equation. The KdV equation is a deterministic nonlinear partial differential equation that possesses a solitary wave solution known as a soliton. Using phase portrait analysis, we can show that the graphical time profile of the number of infected in the SIR model is qualitatively the same as the KdV wave profile.

Keywords

SIR Model, Second-Order Approximation, Korteweg-de Vries Equation, Nonlinear PDE, Solitary Wave

1. Introduction

Disease dynamics is the study of how an infectious disease behaves in a population over time. One such model separates a population into three compartments: susceptible to infection, infected, and removed. In the classical model, the removed population is assumed to consist of those who have one of the following statuses:

1) Survived the infection and is immune.
2) Did not survive the infection (death).
3) Quarantined and in recovery.
4) Vaccinated.

The change in the number of susceptible $S(t)$, infected $I(t)$, and removed

$R(t)$ with respect to time in this model can be described by the following system of equations.

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta SI \tag{1}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta SI - \gamma I \tag{2}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I. \tag{3}$$

It is assumed that $N = S(t) + I(t) + R(t)$ where N is a constant representing the entire population. The constant β is the likelihood of contracting the infection from adequate contact and the constant γ represents the recovery rate in number of days [1]. This classical form of the SIR model originated from the seminal papers of Ross and Ross and Hudson in 1916-1917. Yet, it was Kermack and McKendrick in 1927-1932 that provided the fundamental contributions towards its application [2]. The SIR model describes the transmission of infectious diseases between susceptible and infected individuals and provides the basic framework for almost all subsequent epidemic models [3].

2. Disease Model at the Early Stages

We are interested in deriving solutions with respect to time. To accomplish this, we use the approach of small perturbations in time for a *good* approximation of the S, I, R solutions. In reality, this means we are looking at the early stages of the onset of the disease. Now, dividing (1) by (3) and using separation of variables, we get

$$\frac{\frac{\mathrm{d}S}{\mathrm{d}t}}{\frac{\mathrm{d}R}{\mathrm{d}t}} = \frac{-\beta SI}{\gamma I}$$

$$\frac{\mathrm{d}S}{\mathrm{d}R} = \frac{-\beta}{\gamma} S$$

$$\frac{1}{S} \cdot \mathrm{d}S = \frac{-\beta}{\gamma} S \mathrm{d}R \cdot \frac{1}{S}.$$

Integrating,

$$\int \frac{1}{S} \mathrm{d}S = \int \frac{-\beta}{\gamma} \mathrm{d}R$$

$$\ln S + \ln K = -\frac{\beta}{\gamma} R$$

$$S = K \mathrm{e}^{-\frac{\beta}{\gamma} R},$$

where K is an arbitrary constant from integration. Our assumption in this paper is that we are dealing with a novel or emerging infectious disease. One implication of this assumption is that at time $t = 0$, $R(0) = 0$. Then at time $t = 0$ the number of susceptibles is equal to K. So,

$$S(t) = S_0 \mathrm{e}^{-\frac{\beta}{\gamma}R}. \tag{4}$$

Now, using the second order series expansion of $\mathrm{e}^{-\frac{\beta}{\gamma}R}$ we get

$$S = S_0\left[1-\frac{\beta}{\gamma}R+\frac{\beta^2}{\gamma^2(2!)}R^2\right]. \tag{5}$$

Here, we are assuming the efficacy of our second-order model only when the number of removed is small. That is, as R gets larger in the function $S(t)$, our second-order approximation becomes less reliable.

2.1. Deriving Time-Dependent Solutions

We will use the system of differential equations to derive R as a function of time. Then we will use R to find S and I as functions of time. From Equations (1), (2) and (3) we have:

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{1}{\gamma}\frac{\mathrm{d}^2R}{\mathrm{d}t^2} \tag{6}$$

$$\beta SI = -\frac{\mathrm{d}S}{\mathrm{d}t} \tag{7}$$

$$-\gamma I = -\frac{\mathrm{d}R}{\mathrm{d}t} \tag{8}$$

$$S = S_0\left[1-\frac{\beta}{\gamma}R+\frac{\beta^2}{2\gamma^2}R^2\right]. \tag{9}$$

Differentiating (9) with respect to t,

$$\frac{\mathrm{d}S}{\mathrm{d}t} = S_0\left[-\frac{\beta}{\gamma}\frac{\mathrm{d}R}{\mathrm{d}t}+\frac{\beta^2}{\gamma^2}R\frac{\mathrm{d}R}{\mathrm{d}t}\right]. \tag{10}$$

Then, we substitute from Equations (6)-(10) to get

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta SI - \gamma I$$

$$\frac{1}{\gamma}\frac{\mathrm{d}^2R}{\mathrm{d}t^2} = -\frac{\mathrm{d}S}{\mathrm{d}t}-\frac{\mathrm{d}R}{\mathrm{d}t}.$$

The above simplifies to

$$\frac{\mathrm{d}^2R}{\mathrm{d}t^2} = (S_0\beta-\gamma)\frac{\mathrm{d}R}{\mathrm{d}t}-\frac{S_0\beta^2}{\gamma}R\frac{\mathrm{d}R}{\mathrm{d}t},$$

and integrating once to get

$$\frac{\mathrm{d}R}{\mathrm{d}t} = (S_0\beta-\gamma)R-\frac{S_0\beta^2}{2\gamma}R^2+Q$$

where Q is the constant of integration. Let $D = S_0\beta-\gamma$ and $F = \frac{S_0\beta^2}{2\gamma}$. Using separation of variables,

$$\frac{\mathrm{d}R}{-\left[FR^2-DR-Q\right]} = \mathrm{d}t$$

$$\frac{\mathrm{d}R}{\left[-R^2+\frac{D}{F}R+\frac{Q}{F}\right]}=F\mathrm{d}t.$$

If we write $-R^2+\frac{D}{F}R+\frac{Q}{F}=\left(\lambda_1-R\right)\left(\lambda_2+R\right)$, then

$$\frac{\mathrm{d}R}{\left(\lambda_1-R\right)\left(\lambda_2+R\right)}=F\mathrm{d}t.$$

Using partial fractions and integrating, we have

$$\int\frac{1}{\lambda_1+\lambda_2}\left[\frac{1}{\lambda_1-R}+\frac{1}{\lambda_2+R}\right]\mathrm{d}R=\int F\mathrm{d}t$$

$$\frac{1}{\lambda_1+\lambda_2}\left[-\ln\left(\lambda_1-R\right)+\ln\left(\lambda_2+R\right)\right]=Ft-G$$

Simplifying,

$$\frac{\lambda_2+R}{\lambda_1-R}=\mathrm{e}^{(\lambda_1+\lambda_2)Ft-\tilde{G}}$$

and solving for R to obtain

$$R=\frac{\lambda_1\mathrm{e}^{(\lambda_1+\lambda_2)Ft-\tilde{G}}-\lambda_2}{1+\mathrm{e}^{(\lambda_1+\lambda_2)Ft-\tilde{G}}}$$

where G is the constant of integration and $\tilde{G}=\left(\lambda_1+\lambda_2\right)G$. We can derive S from (10) using R to get

$$S=S_0\left[1-\frac{\beta}{\gamma}R+\frac{\beta^2}{\gamma^2\left(2!\right)}R^2\right]$$

$$S\left(t\right)=S_0-\mathbf{R}_0\left(\frac{\lambda_1\mathrm{e}^{(\lambda_1+\lambda_2)Ft-\tilde{G}}-\lambda_2}{1+\mathrm{e}^{(\lambda_1+\lambda_2)Ft-\tilde{G}}}\right)+\frac{F}{\gamma}\left(\frac{\lambda_1\mathrm{e}^{(\lambda_1+\lambda_2)Ft-\tilde{G}}-\lambda_2}{1+\mathrm{e}^{(\lambda_1+\lambda_2)Ft-\tilde{G}}}\right)^2$$

where $\mathbf{R}_0=S_0\frac{\beta}{\gamma}$. Note that $\mathbf{R}_0$ is *not* the recovered function at time $t=0$. $\mathbf{R}_0$ is referred to as the *basic reproduction number* and it represents the number of individuals that an infected person infects during the infectious period, when $\mathbf{R}_0>1$ the disease will spread [4]. Let $\mu=\left(\lambda_1+\lambda_2\right)F$. Then, we use (3) to find I as a function of time

$$\gamma I=\frac{\mathrm{d}R}{\mathrm{d}t}$$

$$I=\frac{1}{\gamma}\frac{\mathrm{d}R}{\mathrm{d}t}=\frac{1}{\gamma}\frac{\mu\mathrm{e}^{\mu t-\hat{G}}\left(\lambda_1+\lambda_2\right)}{\left(1+\mathrm{e}^{\mu t-\hat{G}}\right)^2}$$

We obtain

$$I\left(t\right)=\frac{F\left(\lambda_1+\lambda_2\right)^2}{4\gamma}\mathrm{sech}^2\left(\frac{\left(\lambda_1+\lambda_2\right)F}{2}t-\hat{G}\right)$$

where

$$\hat{G} = \frac{\tilde{G}}{2} = \frac{\lambda_1 + \lambda_2}{2} G$$

for some constant G. We arrive at our SIR model with second-order approximated solutions as

$$S(t) = S_0 - \mathbf{R}_0 \left(\frac{\lambda_1 e^{(\lambda_1+\lambda_2)Ft-\tilde{G}} - \lambda_2}{1 + e^{(\lambda_1+\lambda_2)Ft-\tilde{G}}} \right) \tag{11}$$

$$+ \frac{F}{\gamma} \left(\frac{\lambda_1 e^{(\lambda_1+\lambda_2)Ft-\tilde{G}} - \lambda_2}{1 + e^{(\lambda_1+\lambda_2)Ft-\tilde{G}}} \right)^2 \tag{12}$$

$$I(t) = \frac{F(\lambda_1 + \lambda_2)^2}{4\gamma} \operatorname{sech}^2 \left(\frac{(\lambda_1 + \lambda_2)Ft}{2} - \hat{G} \right) \tag{13}$$

$$R(t) = \frac{\lambda_1 e^{(\lambda_1+\lambda_2)Ft-\tilde{G}} - \lambda_2}{1 + e^{(\lambda_1+\lambda_2)Ft-\tilde{G}}} \tag{14}$$

and constants

$$\mathbf{R}_0 = S_0 \frac{\beta}{\gamma} \tag{15}$$

$$D = S_0 \beta - \gamma \tag{16}$$

$$F = \frac{S_0 \beta^2}{2\gamma} \tag{17}$$

$$\hat{G} = \frac{\tilde{G}}{2} = \frac{\lambda_1 + \lambda_2}{2} G \tag{18}$$

with Q a constant of integration and λ_1, λ_2 are roots to $-R^2 + \frac{D}{F}R + \frac{Q}{F} = 0$.

2.2. Long-Term Behavior of the Time-Dependent Solutions

With the second-order approximated solutions for the disease model derived, we can investigate how the functions behave over time. Understanding what happens in the long-term may give us insight into the impact the disease will have on the population. First, we must rearrange $R(t)$ for exponential decay over time, like so

$$R(t) = \frac{\lambda_1 e^{(\lambda_1+\lambda_2)Ft-\tilde{G}} - \lambda_2}{1 + e^{(\lambda_1+\lambda_2)Ft-\tilde{G}}} = \frac{\lambda_1 - \lambda_2 e^{-\left[(\lambda_1+\lambda_2)Ft-\tilde{G}\right]}}{e^{-\left[(\lambda_1+\lambda_2)Ft-\tilde{G}\right]} + 1},$$

then as $t \to \infty$, we have

$$R_\infty = \frac{\lambda_1 - \lambda_2 e^{\tilde{G}}}{e^{\tilde{G}} + 1}. \tag{19}$$

Starting with Equation (11)

$$S(t) = S_0 - \mathbf{R}_0 \left(\frac{\lambda_1 e^{(\lambda_1+\lambda_2)Ft-\tilde{G}} - \lambda_2}{1 + e^{(\lambda_1+\lambda_2)Ft-\tilde{G}}} \right) + \frac{F}{\gamma} \left(\frac{\lambda_1 e^{(\lambda_1+\lambda_2)Ft-\tilde{G}} - \lambda_2}{1 + e^{(\lambda_1+\lambda_2)Ft-\tilde{G}}} \right)^2$$

which can be re-written as

$$S(t) = S_0 - \mathbf{R}_0\left(\frac{\lambda_1 - \lambda_2 \mathrm{e}^{-\left[(\lambda_1+\lambda_2)Ft-\tilde{G}\right]}}{\mathrm{e}^{-\left[(\lambda_1+\lambda_2)Ft-\tilde{G}\right]}+1}\right) + \frac{F}{\gamma}\left(\frac{\lambda_1 - \lambda_2 \mathrm{e}^{-\left[(\lambda_1+\lambda_2)Ft-\tilde{G}\right]}}{\mathrm{e}^{-\left[(\lambda_1+\lambda_2)Ft-\tilde{G}\right]}+1}\right)^2$$

and taking time $t \to \infty$ to get

$$S_\infty = S_0 - \mathbf{R}_0\left(\frac{\lambda_1 - \lambda_2 \mathrm{e}^{\tilde{G}}}{\mathrm{e}^{\tilde{G}}+1}\right) + \frac{F}{\gamma}\left(\frac{\lambda_1 - \lambda_2 \mathrm{e}^{\tilde{G}}}{\mathrm{e}^{\tilde{G}}+1}\right)^2. \tag{20}$$

So, R_∞ and S_∞ estimate the number of people in the removed and susceptible categories respectively at the end of the disease epidemic. Because the infected function is in hyperbolic secant squared form, we know that $I_\infty = 0$. This implies that the disease dies out in the long run. Furthermore, $N = S_\infty + R_\infty$.

2.3. Determining Constants

Understanding how the starting number of infected and recovered can influence the behavior of a disease is one of the primary goals of this paper. To gain some insight, we investigate the values of constants Q and G.

Consider $\frac{\mathrm{d}R}{\mathrm{d}t}$ as

$$\begin{aligned}\frac{\mathrm{d}R}{\mathrm{d}t} &= \left(S_0\beta - \gamma\right)R - \frac{S_0\beta^2}{2\gamma}R^2 + Q \\ &= -FR^2 + DR + Q.\end{aligned}$$

At time $t = 0$, $R(0) = 0$, $I(0) = I_0$. Using Equation (3) with the form of $\frac{\mathrm{d}R}{\mathrm{d}t}$ above,

$$-FR^2 + DR + Q = \gamma I$$

$$-F\left(R(0)\right)^2 + DR(0) + Q = \gamma I_0$$

$$Q = \gamma I_0.$$

So Q simply relies upon the starting number of infected and the rate of recovery. Investigating G, we use Equation (13) at time $t = 0$,

$$I(0) = \frac{F\left(\lambda_1+\lambda_2\right)^2}{4\gamma}\operatorname{sech}^2\left(\hat{G}\right)$$

$$\frac{4I_0\gamma}{F\left(\lambda_1+\lambda_2\right)^2} = \operatorname{sech}^2\left(\hat{G}\right)$$

which becomes

$$\hat{G} = \operatorname{sech}^{-1}\left(\sqrt{\frac{4I_0\gamma}{F\left(\lambda_1+\lambda_2\right)^2}}\right)$$

$$= \operatorname{sech}^{-1}\left(\sqrt{\frac{4I_0\gamma^2}{S_0\beta^2\left(\lambda_1+\lambda_2\right)^2}}\right)$$

$$\hat{G} = \operatorname{sech}^{-1}\left(\sqrt{\frac{4I_0\gamma}{\mathbf{R}_0\beta\left(\lambda_1+\lambda_2\right)^2}}\right).$$

Notice that the horizontal shift of the *infected solution* depends on the initial conditions of S_0 and I_0.

2.4. Finding Peak Data

Our goal is that, when given a starting value for infected individuals at time $t=0$, we can find at what time $t=T$ we get the maximum infected $I_{\max}$. Having a predictive model would allow health professionals to be proactive toward an emerging disease in the early stages.

Recall the following equation and its result.

$$\begin{aligned}\left(\lambda_1-R\right)\left(\lambda_2+R\right) &= \lambda_1\lambda_2+\left(\lambda_1-\lambda_2\right)R-R^2\\ &= \frac{Q}{F}+\frac{D}{F}R-R^2.\end{aligned}$$

Matching coefficients from above, we see that

$$\lambda_1\lambda_2 = \frac{Q}{F} \tag{21}$$

$$\lambda_1-\lambda_2 = \frac{D}{F}. \tag{22}$$

Now recall the *infected solution* Equation (13). The peak of infection $I_{\max}$ occurs when $\operatorname{sech}^2(0)$, then

$$I_{\max} = \frac{F\left(\lambda_1+\lambda_2\right)^2}{4\gamma}\operatorname{sech}^2(0)$$

$$I_{\max} = \frac{F\left(\lambda_1+\lambda_2\right)^2}{4\gamma}.$$

Now notice

$$\begin{aligned}\left(\lambda_1+\lambda_2\right)^2 &= \lambda_1^2+\lambda_2^2+2\lambda_1\lambda_2\\ &= \lambda_1^2+\lambda_2^2+2\lambda_1\lambda_2+\left(-2\lambda_1\lambda_2+2\lambda_1\lambda_2\right)\\ &= \lambda_1^2+\lambda_2^2-2\lambda_1\lambda_2+4\lambda_1\lambda_2\end{aligned}$$

$$\left(\lambda_1+\lambda_2\right)^2 = \left(\lambda_1-\lambda_2\right)^2+4\lambda_1\lambda_2.$$

Then, substituting in Equations (21) and (22),

$$\left(\lambda_1+\lambda_2\right)^2 = \left(\lambda_1-\lambda_2\right)^2+4\lambda_1\lambda_2$$

$$\left(\lambda_1+\lambda_2\right)^2 = \frac{D^2}{F^2}+\frac{4Q}{F}.$$

We now have

$$I_{\max} = \frac{F}{4\gamma}\left[\frac{D^2}{F^2}+\frac{4Q}{F}\right] \tag{23}$$

Substituting in for $Q=\gamma I_0$ note that $D=S_0\beta-\gamma=S_0\frac{\beta}{\gamma}-1=\mathbf{R}_0-1$, we find the maximum infected as

$$I_{\max}=\frac{F}{4\gamma}\left[\frac{D^2}{F^2}+\frac{4\gamma I_0}{F}\right]=\frac{D^2}{4\gamma F}+I_0$$

$$I_{\max}=\frac{\gamma^2\left(\mathbf{R}_0-1\right)^2}{2S_0\beta^2}+I_0$$

$$=I_0+\left(\frac{S_0}{S_0}\right)\frac{\gamma^2\left(\mathbf{R}_0-1\right)^2}{2S_0\beta^2}$$

$$=I_0+\frac{S_0\gamma^2\left(\mathbf{R}_0-1\right)^2}{2S_0^2\beta^2}$$

$$=I_0+\left(\frac{S_0}{2}\right)\frac{\left(\mathbf{R}_0-1\right)^2}{\mathbf{R}_0^2}$$

becoming

$$I_{\max}=I_0+\left(\frac{S_0}{2}\right)\left(\frac{\mathbf{R}_0-1}{\mathbf{R}_0}\right)^2. \tag{24}$$

Now we can determine the maximum infected using Equation (24) with initial values and the basic reproduction number $\mathbf{R}_0$. Furthermore, we see that I_0 has an additive effect on $I_{\max}$. For finding time $t=T$, we solve for when $\operatorname{sech}^2(0)$, that is, when $\frac{\left(\lambda_1+\lambda_2\right)Ft}{2}-\hat{G}=0$. Then

$$0=\frac{\left(\lambda_1+\lambda_2\right)FT}{2}-\hat{G}$$

$$\frac{\left(\lambda_1+\lambda_2\right)FT}{2}=\hat{G}$$

$$T=\frac{2}{\left(\lambda_1+\lambda_2\right)F}\hat{G}$$

$$T=\left(\frac{2}{\left(\lambda_1+\lambda_2\right)F}\right)\operatorname{sech}^{-1}\left(\sqrt{\frac{4I_0\gamma}{F\left(\lambda_1+\lambda_2\right)^2}}\right).$$

Therefore, using our solution, we can estimate the maximum number of infected and when that occurs. Using Equations (11)-(14), we get the graph in Figure 1. This illustrates the second-order approximated solutions for the SIR model with the COVID ancestral strain basic reproduction number $\mathbf{R}_0=2.71$ [5]. In Figure 2, we compare *infected solutions* with different COVID variants' basic reproduction numbers (Table 1).

2.5. A Link to the KdV Equation

Looking a little deeper into the *infected solution*, we notice that if we can write the following equation

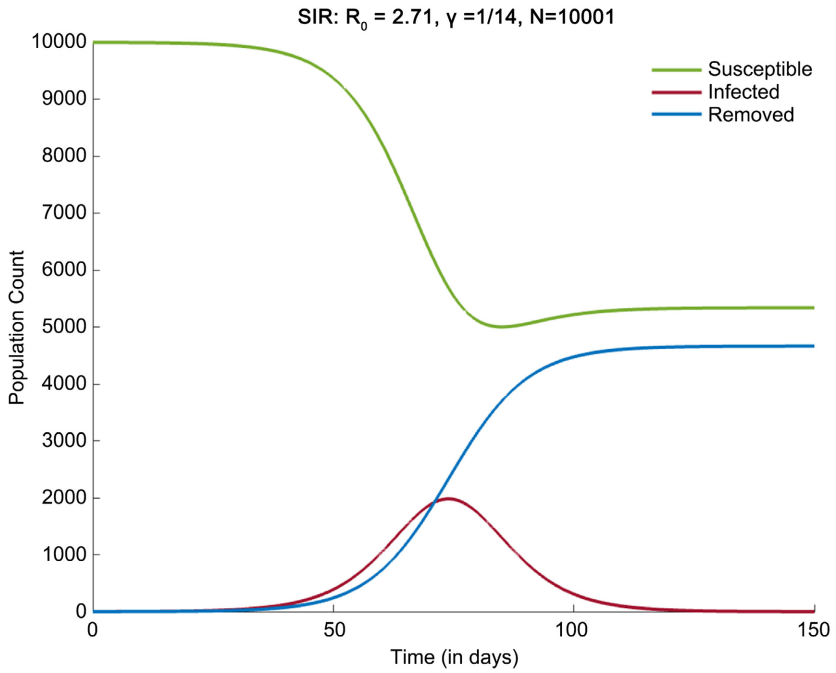

Figure 1. SIR solutions based on second-order approximation.

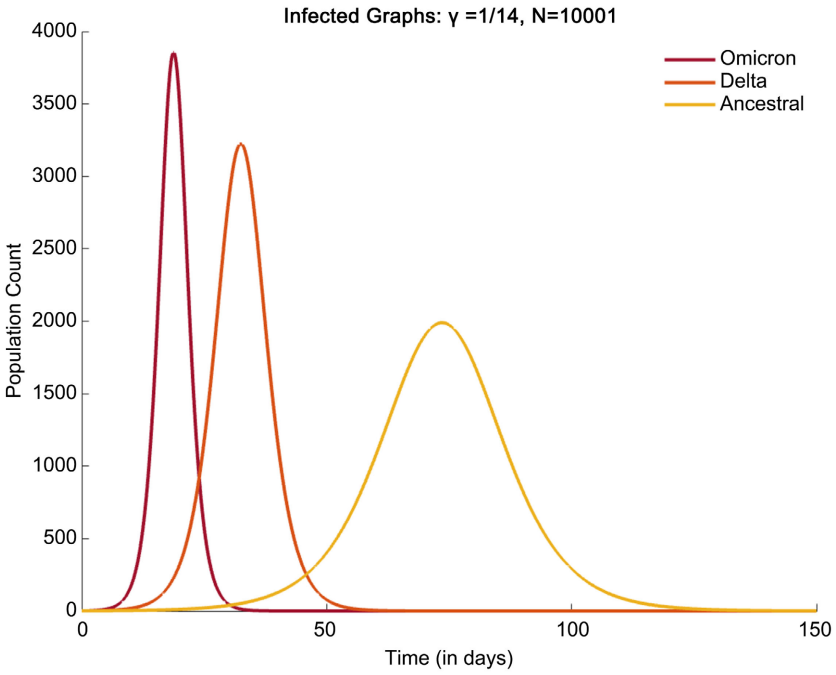

Figure 2. *Infected solutions* for different COVID variants.

Table 1. COVID case study with basic reproduction numbers and peak data [5]-[7].

Variation	$\mathbf{R}_0$	Time of Peak	Peak Value
Ancestral	2.71	73.55	1990
Delta	5.08	32.48	3225
Omicron	8.20	18.75	3854

$$aI(t)+bI^2(t)+cI''(t)=0, \tag{25}$$

then we get a form that reminds us of the Korteweg-de Vries (KdV) equation. The KdV equation is a nonlinear partial differential equation that models long wave motion in shallow water. Because of a delicate balance between dispersion and nonlinearity this equation is known to posses a special solution called soliton.

$$I(t)=\frac{F(\lambda_1+\lambda_2)^2}{4\gamma}\operatorname{sech}^2\left(\frac{(\lambda_1+\lambda_2)Ft}{2}-\hat{G}\right)$$

$$I^2(t)=F^2\frac{(\lambda_1+\lambda_2)^4}{16\gamma^2}\operatorname{sech}^4\left(\frac{(\lambda_1+\lambda_2)Ft}{2}-\hat{G}\right)$$

$$I''(t)=\frac{-F^3(\lambda_1+\lambda_2)^4}{8\gamma}\left[-2\operatorname{sech}^2\left(\frac{(\lambda_1+\lambda_2)Ft}{2}-\hat{G}\right)\right.$$
$$\left.+3\operatorname{sech}^4\left(\frac{(\lambda_1+\lambda_2)Ft}{2}-\hat{G}\right)\right].$$

Let $\alpha=\frac{F(\lambda_1+\lambda_2)}{2}$ to simplify some notation, then

$$0=a\frac{F(\lambda_1+\lambda_2)^2}{4\gamma}\operatorname{sech}^2\left(\alpha t-\hat{G}\right)+b\frac{F^2(\lambda_1+\lambda_2)^4}{16\gamma^2}\operatorname{sech}^4\left(\alpha t-\hat{G}\right)$$
$$+c\frac{-F^3(\lambda_1+\lambda_2)^4}{8\gamma}\left[-2\operatorname{sech}^2\left(\alpha t-\hat{G}\right)+3\operatorname{sech}^4\left(\alpha t-\hat{G}\right)\right]$$

factoring to obtain

$$0=\frac{F(\lambda_1+\lambda_2)^2}{4\gamma}\operatorname{sech}^2\left[a+cF^2(\lambda_1+\lambda_2)^2\right]$$
$$+\frac{F^2(\lambda_1+\lambda_2)^4}{8\gamma}\operatorname{sech}^4\left[\frac{b}{2\gamma}-3cF\right].$$

What's inside the brackets must equal zero resulting in the following system:

$$\frac{b}{2\gamma}-3cF=0$$

$$a+cF^2(\lambda_1+\lambda_2)^2=0.$$

Solving this system, we get

$$b=(6\gamma F)c \tag{26}$$

$$a=\left(F^2(\lambda_1+\lambda_2)^2\right)c. \tag{27}$$

One form of the KdV equation is

$$-vu+\frac{1}{2}u^2+u''=0, \tag{28}$$

where v is a constant (see Appendix for derivation). Comparing with (28)

$$aI(t)+bI^2(t)+cI''(t)=0,$$

we set $c=1$. Now,

$$\left(-F^2\left(\lambda_1+\lambda_2\right)^2\right)I(t)+\left(6\gamma F\right)I^2(t)+I''(t)=0.$$

Further, recall $F=\frac{S_0\beta^2}{2\gamma}$ so b becomes

$$6\gamma F=6\gamma\frac{S_0\beta^2}{2\gamma}=3S_0\beta^2.$$

Again, comparing with Equation (32),

$$3S_0\beta^2=\frac{1}{2}$$

$$S_0\beta^2=\frac{1}{6}.$$

For a, we have

$$\begin{aligned}F^2\left(\lambda_1+\lambda_2\right)^2&=\left(\frac{S_0\beta^2}{2\gamma}\right)^2\left(\lambda_1+\lambda_2\right)^2\\&=\frac{\left(S_0\beta^2\right)^2}{4\gamma^2}\left(\lambda_1+\lambda_2\right)^2\\&=\frac{\frac{1}{36}}{4\gamma^2}\left(\lambda_1+\lambda_2\right)^2\\&=\left(\frac{\lambda_1+\lambda_2}{12\gamma}\right)^2.\end{aligned}$$

Therefore, $v=\left(\frac{\lambda_1+\lambda_2}{12\gamma}\right)^2$ when compared to Equation (25). Finally,

$$-\left(\frac{\lambda_1+\lambda_2}{12\gamma}\right)^2I(t)+3S_0\beta^2I^2(t)+I''(t)=0. \tag{29}$$

This satisfies the condition of $\frac{1}{2}$ in front of the u^2 verifying the resemblance to the KdV form in Equation (28). It is well known that the solution for the KdV is also in hyperbolic secant squared form. However, there is an important distinction between Equations (28) and (29). In (28), the independent variable is the traveling wave coordinate z which is given by x − vt and in (29), the independent variable is just time t. Therefore, in a transformed space z, the wave profile for the infected could be thought of as traveling with a speed v.

3. Conclusion

With this second-order approximation for SIR solutions, we can get an estimate of the behavior of a novel or emerging disease. In particular, the *infected solution* that we obtain provides data on the disease. Data that will be valuable to health professionals preparing for the disease outbreak. Further, we can do *real-time analysis* to keep an updated estimate of the disease behavior. Such an application

would look like:

At the beginning of a novel or emergent disease breakout, we have $S(t), I(t), R(t)$ for time t. After some small but arbitrary amount of days, we have $t = t_{\text{old}}$ and $t_{\text{new}} = t_{\text{old}} - t_{\text{elapsed}}$. The updated S, I, R system becomes

$$S_{\text{new}} = S_{\text{old}} - S_{\text{elapsed}}$$

$$I_{\text{new}} = I_{\text{old}} - I_{\text{elapsed}}$$

$$R_{\text{new}} = R_{\text{old}} - R_{\text{elapsed}}$$

A process that can be repeated as data is reported and updated during the disease breakout.

It should be noted that our approximate solution will become unreliable when the principal error associated with the second-order approximation becomes larger. For example, if one wants the principal error to be less than $10^{(-n)}$, then provided the number of removed R is less than 1.82 $(S_0/R_0)10^{(-n/3)}$, the approximate solution could be considered reliable. It will be up to the healthcare professional or user to choose the exponent n, depending on the accuracy they are looking for. Otherwise, further work could be done making use of third-order approximation or higher. It is somewhat interesting that the second-order solution for the infected resembles the soliton solution of the KdV equation.

Authors' Contributions

Both V.S. Manoranjan and Zachary Fendler contributed equally throughout the development of this paper. V.S. Manoranjan designed the problem and provided ideas, derivations, and edits to the manuscript. Zachary Fendler worked on derivations, and implementation using MATLAB, and wrote the manuscript. Both authors did the literature review.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Hethcote, H.W. (2000) The Mathematics of Infectious Diseases. *SIAM Review*, **42**, 599-653. https://doi.org/10.1137/s0036144500371907

[2] Kermack, W.O. and McKendrick, A.G. (1927) A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A*: *Mathematical, Physical and Engineering Sciences*, **115**, 700-721.

[3] Magal, P. and Ruan, S. (2014) Susceptible-Infectious-Recovered Models Revisited: From the Individual Level to the Population Level. *Mathematical Biosciences*, **250**, 26-40. https://doi.org/10.1016/j.mbs.2014.02.001

[4] van den Driessche, P. (2017) Reproduction Numbers of Infectious Disease Models. *Infectious Disease Modelling*, **2**, 288-303. https://doi.org/10.1016/j.idm.2017.06.002

[5] Rahman, B., Sadraddin, E. and Porreca, A. (2020) The Basic Reproduction Number of SARS-CoV-2 in Wuhan Is about to Die out, How about the Rest of the World? *Reviews in Medical Virology*, **30**, e2111. https://doi.org/10.1002/rmv.2111

[6] Liu, Y. and Rocklöv, J. (2021) The Reproductive Number of the Delta Variant of SARS-CoV-2 Is Far Higher Compared to the Ancestral SARS-CoV-2 Virus. *Journal of Travel Medicine*, **28**, taab124. https://doi.org/10.1093/jtm/taab124

[7] Liu, Y. and Rocklöv, J. (2022) The Effective Reproductive Number of the Omicron Variant of SARS-CoV-2 Is Several Times Relative to Delta. *Journal of Travel Medicine*, **29**, taac037. https://doi.org/10.1093/jtm/taac037

[8] Lewis, B.J., Onder, E.N. and Prudil, A.A. (2022) Nonlinear Differential Equations. In: *Advanced Mathematics for Engineering Students*, Elsevier, 329-347. https://doi.org/10.1016/b978-0-12-823681-9.00020-4.

Appendix

This appendix provides a derivation of the particular form of the KdV equation and its solution used in the paper.

The KdV Equation

The Korteweg-de Vries (KdV) equation is a well-known equation that models long waves in shallow water [8]. This equation can be written as

$$u_t + uu_x + u_{xxx} = k$$

for some constant k. Then

$$u_t + \left(\frac{1}{2}u^2\right)_x + u_{xxx} = k.$$

Due to the nature of the KdV equation, we know that at $-\infty$ and ∞, $k = 0$. We are interested in the profile of this equation as the wave is moving with a constant speed v so let $z = x - vt$. Then

$$-v\frac{\mathrm{d}u}{\mathrm{d}z} + \left(\frac{1}{2}u^2\right)\frac{\mathrm{d}}{\mathrm{d}z} + \frac{\mathrm{d}^3u}{\mathrm{d}z^2} = 0$$

$$-v\frac{\mathrm{d}u}{\mathrm{d}z} + u\frac{\mathrm{d}u}{\mathrm{d}z} + \frac{\mathrm{d}^3u}{\mathrm{d}z^2} = 0.$$

Simplifying notation and integrating once, we get

$$-vu + \frac{1}{2}u^2 + u'' = 0 \tag{30}$$

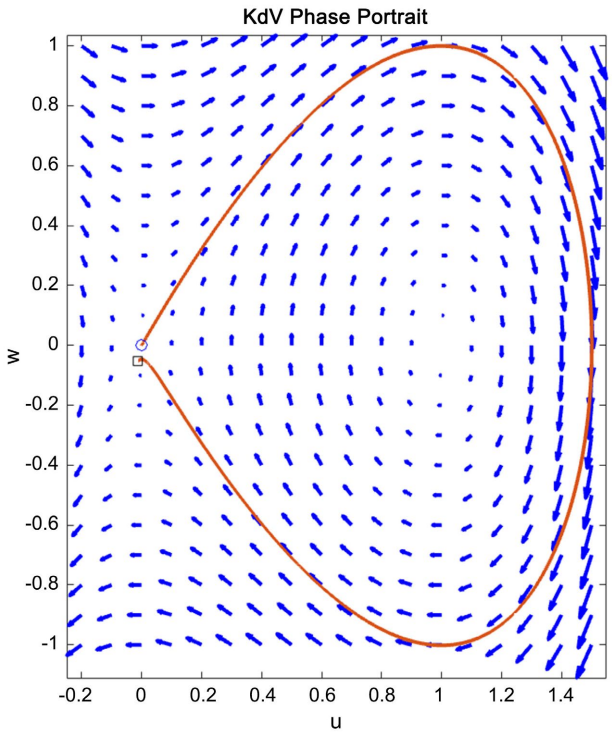

Figure A1. Phase portrait of *infected solution*.

Now let $w = u'$ and,

$$w = u' \tag{31}$$

$$w' = vu - \frac{1}{2}u^2. \tag{32}$$

This allows us to solve numerically in a phase portrait. We can see a homoclinic orbit in Figure A1.