

## Article

# A Mathematical Model of Epidemics—A Tutorial for Students

Yutaka Okabe \*  and Akira Shudo 

Department of Physics, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan; shudo@tmu.ac.jp

\* Correspondence: okabe@phys.se.tmu.ac.jp

Received: 10 June 2020; Accepted: 14 July 2020; Published: 17 July 2020



**Abstract:** This is a tutorial for the mathematical model of the spread of epidemic diseases. Beginning with the basic mathematics, we introduce the susceptible-infected-recovered (SIR) model. Subsequently, we present the numerical and exact analytical solutions of the SIR model. The analytical solution is emphasized. Additionally, we treat the generalization of the SIR model including births and natural deaths.

**Keywords:** SIR model; numerical solution; exact solution; Bernoulli differential equation; Abel differential equation

## 1. Introduction

The coronavirus disease 2019 (COVID-19) is an infectious disease that emerged in Wuhan, China in December 2019 and has rapidly spread in China and worldwide. On 30 January 2020, the World Health Organization (WHO) declared the outbreak as a public health emergency of international concern. Furthermore, the WHO announced the COVID-19 outbreak as a “pandemic” on 11 March 2020 [1,2]. The prevention of infection is the responsibility of every human being.

A few mathematical models have been developed to describe the dynamics of the evolution of COVID-19 [3–7]. Among them, the susceptible-infected-recovered (SIR) model is treated as a basic mathematical model for the spread of epidemic diseases. This model was proposed in 1927 by Kermack and McKendrick [8,9], and has been used as a basic model for epidemics [10–14].

The SIR model is not difficult from the viewpoint of mathematics, and can be understood even by high-school students. We present a numerical solution and an exact analytical solution [15] of the SIR model herein. The analytical solution is emphasized because it is not well known. We provide a more straightforward derivation of the analytical solution compared with that in the original paper; it will serve as a good opportunity for students to learn differentiation and integration. Furthermore, we treat the generalization of the SIR model with vital dynamics including birth and natural death processes. We demonstrate that the complex SIR model can be reduced to an Abel-type differential equation.

## 2. Exponential Growth

The spread of an epidemic disease depends on both the amount of contact between individuals and the possibility of an infected person transmitting the disease to another person. If each infected person is in contact with two other people before they recover, the disease will soon begin to spread rapidly. Assuming that recovery requires one day, the number of sick people will double each day. This situation is illustrated schematically in Figure 1. A numerical series for which the ratio of each two consecutive terms is a constant  $r$  (in this case,  $r = 2$ ) is called a geometric series. The  $n$ -th term is expressed as follows:

$$A(n) = A(0) r^n, \quad (1)$$

where the initial value is set to be  $A(0)$ . If  $r$  is greater than 1, then the series increases and approaches positive infinity rapidly (we assume  $A(0) > 0$ ); if  $r$  is smaller than 1, the series decreases and approaches zero. At the border of  $r = 1$ , the series remains constant. As an example, when one borrows money, the interest is typically a compound interest, i.e., a geometric series. When the annual interest is 30%, the return in five years will increase rapidly as follows:

$$\frac{A(5)}{A(0)} = 1.3^5 \simeq 3.7. \quad (2)$$

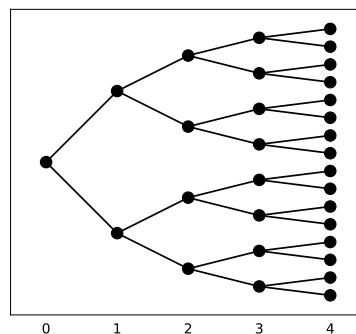


Figure 1. Schematic illustration of spread of epidemic disease.

Extending  $n$  to a real number, we introduce the exponential function

$$x = r^t. \quad (3)$$

The differentiation of the exponential function can be calculated as follows:

$$\frac{dx}{dt} = \lim_{h \rightarrow 0} \frac{r^{t+h} - r^t}{h} = r^t \lim_{h \rightarrow 0} \frac{r^h - 1}{h}. \quad (4)$$

The special constant  $r$  that satisfies the limiting equation

$$\lim_{h \rightarrow 0} \frac{r^h - 1}{h} = 1 \quad (5)$$

is called Napier's constant  $e$ . With this definition of  $e$ , the differentiation of the exponential function becomes

$$\frac{d}{dt} e^t = e^t. \quad (6)$$

Hereafter, we treat the form of the exponential function as

$$x(t) = Ce^{at}. \quad (7)$$

The exponential function grows for  $a > 0$ , and decays for  $a < 0$ . When  $a = 0$ , the function becomes constant. We have assumed that  $C > 0$ .

The inverse function of  $x$ ,  $t = t(x)$ , is a logarithmic function written as

$$t = \log_r x, \quad (8)$$

where  $r$  is the base. The logarithmic function with the base  $e$  is called a natural logarithmic function and is denoted as follows:

$$\log_e x = \ln x. \quad (9)$$

The exponential function always explodes for  $a > 0$ . To verify whether it grows rapidly or slowly, we may consider the doubling time. From the condition

$$\frac{x(t_2)}{x(t_1)} = \frac{e^{at_2}}{e^{at_1}} = e^{a(t_2-t_1)} = 2, \quad (10)$$

we obtain that the doubling time is a time difference that satisfies

$$a(t_2 - t_1) = \ln 2. \quad (11)$$

It is instructive to use graph plots to observe the growth behavior. Both the linear and semi-log plots are shown in Figure 2. A semi-log plot has one axis ( $t$ ) on a linear scale, and the other on a logarithmic scale. When the function grows exponentially, a linear increase is observed in a semi-log plot.

Next, we explain the differential equation. Let us consider the case in which the increase in a variable  $x(t)$  is proportional to  $x(t)$ , i.e.,

$$\Delta x(t) = x(t + \Delta t) - x(t) = ax(t) \Delta t. \quad (12)$$

The difference in the left-hand side becomes larger when the time difference  $\Delta t$  becomes longer; hence,  $\Delta t$  is multiplied to the right-hand side. If we divide both sides by  $\Delta t$  and take the limit of  $\Delta t \rightarrow 0$ , the following differentiation definition is obtained:

$$\lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} = \frac{dx(t)}{dt} = ax(t). \quad (13)$$

It becomes a  $dx(t)/dt$  equation, which is known as a differential equation. Because Equation (7) satisfies the differential equation, it is regarded as a general solution to the differential equation shown in Equation (13). Here,  $C$  is an arbitrary constant and is known as an integral constant. Next, we discuss the mathematical model of the spread of epidemic diseases.

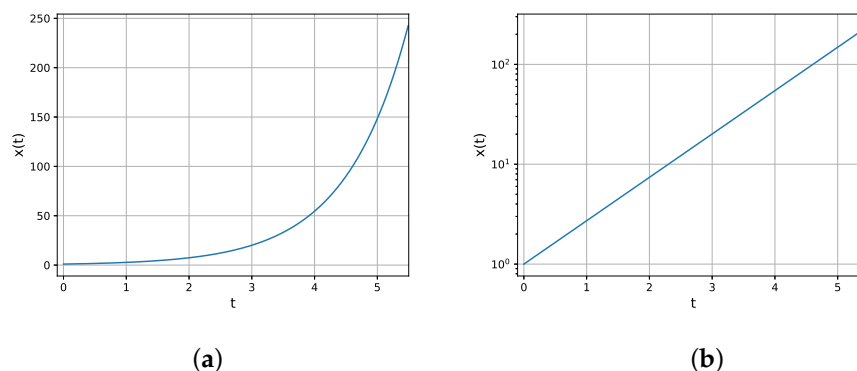


Figure 2. Exponential growth  $e^t$ . (a) Linear plot. (b) Semi-log plot.

### 3. Mathematical Model of Epidemics

If exponential growth continues, the number of the infected individuals increases. We may consider that the infected individuals recover and develop immunity. Subsequently, the number of recovered individuals increases, but we hope that the spread of the disease will subside. We may consider a closed society of  $N$  individuals and classify individuals into three types: susceptible (S), infected (I), and recovered (R). The SIR model [8,9] is a mathematical model for the spread of epidemic diseases that describes the changes in the numbers of the three types of individuals.

We denote the number of susceptible (S) individuals by  $x(t)$ , the number of infected (I) individuals by  $y(t)$ , and the number of recovered or removed (R) individuals by  $z(t)$ . The time  $t$  is measured in days,  $t$  (day). We may consider that each person is in contact with  $m$  persons per day on average. Because the proportion of infected individuals is  $y(t)/N$ , the number of contacts between the infected and susceptible persons becomes  $m(y(t)/N)x(t)$  per day. If we set the probability of infection for each contact as  $p$ , the number of the newly infected individuals within  $\Delta t$  days becomes

$$m \frac{y(t)}{N} x(t) p \Delta t \quad (14)$$

in total. If we set  $\beta = mp$ , the number of non-infected individuals (S) from the  $t$ -th day through the  $(t + \Delta t)$ -th day changes as

$$x(t + \Delta t) - x(t) = -\frac{\beta}{N} x(t) y(t) \Delta t. \quad (15)$$

For the limit of  $\Delta t \rightarrow 0$ , the derivative can be obtained as follows:

$$\frac{dx(t)}{dt} = -\frac{\beta}{N} x(t) y(t), \quad (16)$$

which is the same procedure as that derived for Equation (13).

Meanwhile, the infected individuals recover at a removal rate of  $\gamma$  per day. Subsequently, the increase in the number of the recovered persons becomes

$$\frac{dz(t)}{dt} = \gamma y(t). \quad (17)$$

The inverse of  $\gamma$  is the expected duration of infection. Here, the number of the recovered persons includes the number of the deceased persons because it refers to the number of individuals who cannot possibly infect others.

Because the total number of individuals is set to be constant such that

$$x(t) + y(t) + z(t) = N, \quad (18)$$

we express  $y(t)$  as follows:

$$y(t) = N - x(t) - z(t). \quad (19)$$

The change in the number of the infected individuals can be written as

$$\frac{dy(t)}{dt} = -\frac{dx(t)}{dt} - \frac{dz(t)}{dt} = \frac{\beta}{N} x(t) y(t) - \gamma y(t). \quad (20)$$

Gathering the equations, we have

$$\frac{dx(t)}{dt} = -\frac{\beta}{N} x(t) y(t), \quad (21)$$

$$\frac{dy(t)}{dt} = \frac{\beta}{N} x(t) y(t) - \gamma y(t) = \left( \frac{\beta}{N} x(t) - \gamma \right) y(t), \quad (22)$$

$$\frac{dz(t)}{dt} = \gamma y(t). \quad (23)$$

These are simultaneous equations, where  $\beta$  is the rate of infection and  $\gamma$  the rate of recovery (the rate of quarantine).

If we introduce the transformations of variables, such as  $\tilde{x} = x/N$ ,  $\tilde{y} = y/N$ ,  $\tilde{z} = z/N$ , and  $\tilde{t} = \beta t$ , then Equations (21)–(23) can be rewritten as

$$\frac{d\tilde{x}(t)}{d\tilde{t}} = -\tilde{x}(t)\tilde{y}(t), \quad (24)$$

$$\frac{d\tilde{y}(t)}{d\tilde{t}} = \left(\tilde{x}(t) - \frac{1}{R_0}\right)\tilde{y}(t), \quad (25)$$

$$\frac{d\tilde{z}(t)}{d\tilde{t}} = \frac{1}{R_0}\tilde{y}(t). \quad (26)$$

The equations are in non-dimensional forms. The use of  $\tilde{x}$ ,  $\tilde{y}$ , and  $\tilde{z}$  implies that we only need to consider the fraction of the number of individuals. A single parameter representing the ratio of  $\beta$  and  $\gamma$  appears as follows:

$$R_0 = \frac{\beta}{\gamma}. \quad (27)$$

The number  $R_0$  is known as the basic reproduction number [16–18], and the number of infected individuals increases when  $R_0 > 1$  (precisely, in the limit of  $\tilde{x}(0) \rightarrow 1$ ), whereas it decreases when  $R_0 < 1$ .

#### 4. Numerical Solution of SIR Model

The SIR equations (Equations (21)–(23)) were derived; subsequently, the differential equations can be solved numerically. The simplest method to solve the differential equation is as follows: It is noteworthy that

$$x(t + \Delta t) - x(t) \simeq \frac{dx(t)}{dt} \Delta t. \quad (28)$$

Therefore, we may calculate

$$x(t + \Delta t) \simeq x(t) + \frac{dx(t)}{dt} \Delta t \quad (29)$$

with a time step size of  $\Delta t$ . This first-order numerical procedure is known as the Euler method. However, it has a convergence problem; therefore, the errors will be larger. Although many methods have been proposed to reduce errors, we employ the second-order Runge–Kutta method.

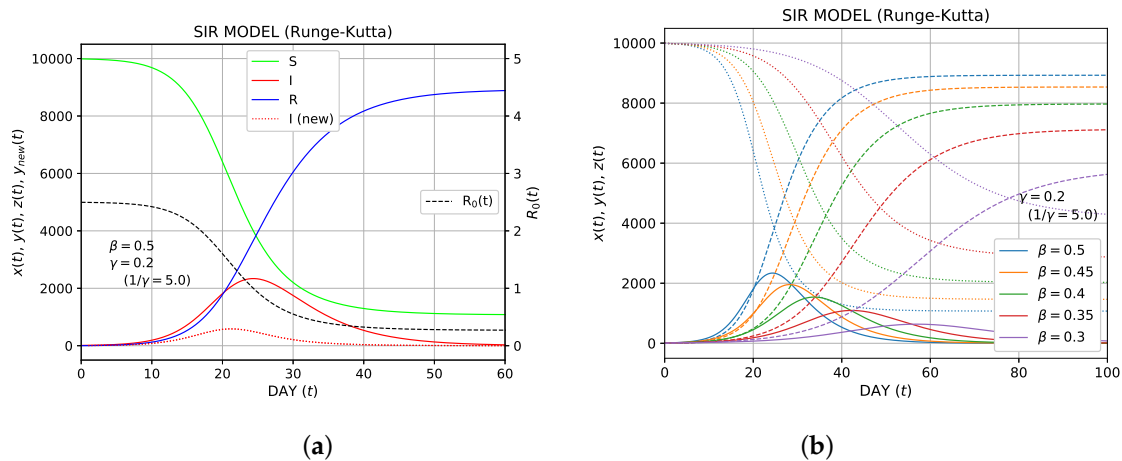
The python program is given in Appendix A. As initial conditions, we use

$$x(0) = N_1, \quad (30)$$

$$y(0) = N_2, \quad (31)$$

$$z(0) = 0 \quad (32)$$

with  $N = 10,000$ ,  $N_1 = 9900$ ,  $N_2 = 10$ , and  $N_3 = 0$ . The calculated results using parameters  $\beta = 0.5$  and  $\gamma = 0.2$  are shown in Figure 3a. As shown, the number of infected individuals (I) increases with time, reaches a peak, and gradually decreases. Meanwhile, the number of the recovered individuals (R) increases gradually.



**Figure 3.** The numerical solution of the SIR model using the Runge–Kutta method;  $N = 10,000$ ,  $N_1 = 9900$ ,  $N_2 = 10$ ,  $N_3 = 0$ . (a) Plot of the temporal variation of the susceptible (S), the infected (I), and the recovered (R) for  $\beta = 0.5$  and  $\gamma = 0.2$ . In addition, the number of the newly infected individuals is shown by the red dotted curve. The time-dependent  $R_0(t)$  is also shown in the black dashed curve. (b) The comparison of S, I, R for  $\beta=0.5, 0.45, 0.4, 0.35$ , and  $0.3$ ;  $\gamma$  is fixed as  $0.2$ .

Thus far, we have discussed the number of the infected individuals,  $y(t)$ . The number of newly infected individuals is reported daily. How can we express the number of the newly infected individuals in the framework of the SIR model? The newly infected individuals are added to  $y(t)$ , but the recovered individuals  $z(t)$  are subtracted from  $y(t)$  as well. Therefore, the number of the newly infected individuals,  $y_{\text{new}}(t)$ , is expressed as

$$y_{\text{new}}(t) = (y(t) + z(t)) - (y(t - \Delta t) + z(t - \Delta t)) \quad (33)$$

with  $\Delta t = 1$ . Because  $x(t) + y(t) + z(t) = N$ , we can express

$$y_{\text{new}}(t) = -(x(t) - x(t - \Delta t)). \quad (34)$$

Based on Equation (21), we obtain

$$y_{\text{new}}(t) \simeq -\frac{dx(t)}{dt} \Delta t = \frac{\beta}{N} x(t) y(t) \Delta t. \quad (35)$$

Because  $y_{\text{new}}(t)$  is proportional to  $y(t)$ , the behavior of the temporal variation of  $y_{\text{new}}(t)$  is similar to that of  $y(t)$ . In Figure 3a,  $y_{\text{new}}(t)$  is plotted as a dotted red curve.

In Figure 3b, the temporal variations of the numbers of individuals for the three types (S, I, R) are compared for  $\beta = 0.5, 0.45, 0.4, 0.35$ , and  $0.3$ . Here,  $\gamma$  is fixed as  $0.2$ . The delay of the epidemic and the reduction in the peak height (number of the infected individuals) were shown as  $\beta$  decreases. This trend is typically associated with the effect of social distancing.

The basic reproduction number  $R_0$  is frequently used as the “R naught” or “R zero”. Sometimes, it is treated as a time-varying variable and used as a measure for easing lockdowns. In the SIR model,  $R_0$  is a constant and defined as shown in Equation (27). If we formally define a time-varying number as

$$R_0(t) = \left(\frac{\beta}{\gamma}\right) \left(\frac{x(t)}{N}\right), \quad (36)$$

in the framework of the SIR model, it judges whether  $y(t)$  increases ( $R_0(t) > 1$ ) or decreases ( $R_0(t) < 1$ ) in Equation (22). In Figure 3a, this time-varying  $R_0(t)$  is also plotted as a dotted black curve. The initial value  $R_0$  was set to be  $2.5$ ; as shown,  $R_0(t)$  gradually decreases, becomes  $1$  at the point where  $y(t)$

reaches a peak, and further decreases below 1. The terminology “effective reproduction number” is sometimes used, but it has several meanings.

Finally, we remark on the cumulative number of the infected individuals, which is reported daily. In this model, the cumulative number is expressed as  $y(t) + z(t)$  because it includes the number of the recovered individuals. In the initial stage of the outbreak of an epidemic, this number increases exponentially, and it is appropriate to use the semi-log plot as shown in Figure 2b. The initial increase in the cumulative number of the infected individuals,  $y(t) + z(t)$ , is plotted in a semi-log plot for several  $\beta$  in Figure 4, where  $\gamma$  is fixed. As shown, the initial increase is linear, and the slope decreases as  $\beta$  decreases.

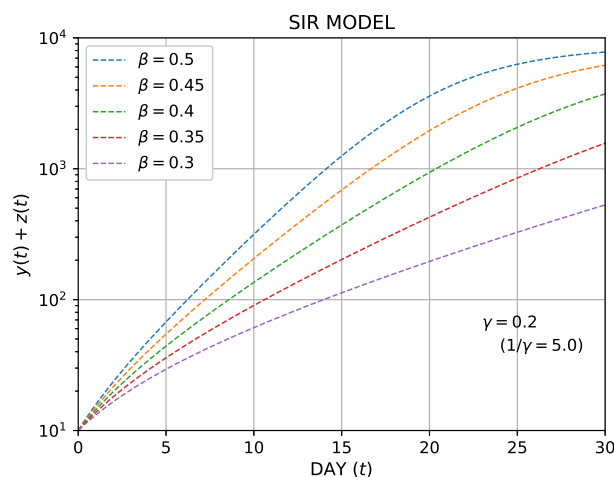


Figure 4. The semi-log plot of the cumulative number of the infected,  $(y(t) + z(t))$ .

## 5. Analytical Solution of SIR Model

We have demonstrated the numerical solution of the SIR model. Although the SIR model was proposed in 1927 [8,9], the analytical solution of the differential equations was only presented in 2014 [15]. In the analytical solution, only exponential and logarithmic functions are involved. Therefore, high-school students will be able to understand the analytical solution of the SIR model.

Next, we present the main stream of the derivation of the analytical solution. Some detailed calculations are given as questions, and the solutions can be found in Appendix B. Equations (21)–(23) are simultaneous differential equations; we can rewrite them as the differential equation of a single function. The procedure is the same as that used to solve a linear simultaneous equation.

We employ a more direct method of deriving the analytical solution compared with that in the original paper [15]. First, we rewrite Equation (21) as

$$y = -\frac{1}{\hat{\beta}} \left( \frac{x'}{x} \right), \quad (37)$$

where  $\hat{\beta} = \beta/N$ . If we differentiate both sides, we obtain

$$y' = -\frac{1}{\hat{\beta}} \left( -\frac{x'^2}{x^2} + \frac{x''}{x} \right). \quad (38)$$

Inserting Equation (37) into Equation (21) yields

$$y' = -(\hat{\beta}x - \gamma) \frac{1}{\hat{\beta}} \left( \frac{x'}{x} \right). \quad (39)$$

Comparing Equation (38) and Equation (39), we have

$$x \frac{d^2x}{dt^2} - \left( \frac{dx}{dt} \right)^2 + (\gamma - \hat{\beta}x)x \frac{dx}{dt} = 0. \quad (40)$$

Therefore, we have obtained a single differential equation of  $x(t)$ .

Next we introduce the following function:

$$\phi = \frac{dt}{dx}. \quad (41)$$

Therefore, we have the following equation for  $\phi$ :

$$\frac{d\phi}{dx} + \frac{\phi}{x} = (\gamma - \hat{\beta}x)\phi^2. \quad (42)$$

An ordinary differential equation of this form is known as the Bernoulli differential equation. The general solution is obtained as follows:

$$\phi = \frac{1}{x(C_1 - \gamma \ln x + \hat{\beta}x)}, \quad (43)$$

where  $C_1$  is an integral constant.

**Problem 1.** Show that the general solution of the Bernoulli differential Equation (42) is given by Equation (43).

From the relation of the inverse function of Equation (41), we have

$$\frac{1}{\phi} = \frac{dx}{dt}. \quad (44)$$

Using Equation (21), we obtain

$$y = -\frac{1}{\hat{\beta}}(C_1 - \gamma \ln x + \hat{\beta}x). \quad (45)$$

Moreover, from Equations (21) and (23), we have

$$\frac{dz}{dt} = -\frac{\gamma}{\hat{\beta}} \left( \frac{x'}{x} \right). \quad (46)$$

Subsequently, the relation between  $x(t)$  and  $z(t)$  becomes

$$z(t) = -\frac{\gamma}{\hat{\beta}} \ln(x(t)/C_2), \quad (47)$$

where  $C_2$  is an integral constant. If  $x(0) = N_1$  and  $z(0) = 0$  for  $t = 0$ , then  $C_2$  becomes  $N_1$ . Therefore, we have

$$z = -\frac{\gamma}{\hat{\beta}} \ln(x/N_1). \quad (48)$$

From the relation  $x(0) + y(0) + z(0) = N$ , we obtain the integral constant  $C_1$  as

$$C_1 = -\hat{\beta}N + \gamma \ln N_1. \quad (49)$$



If we substitute  $C_1$  into Equation (43), we have

$$\frac{dt}{dx} = \frac{1}{x(-\hat{\beta}N - \gamma \ln(x/N_1) + \hat{\beta}x)}. \quad (50)$$

By integrating this equation,  $t$  is obtained as a function of  $x$  as follows:

$$t = \int_{x(0)}^{x(t)} \frac{d\zeta}{\zeta(-\hat{\beta}N - \gamma \ln(\zeta/N_1) + \hat{\beta}\zeta)}. \quad (51)$$

**Problem 2.** Confirm that Equation (51) satisfies differential Equations (21)–(23).

For the convenience of calculation, we change a variable such that  $\zeta/N_1 \rightarrow \zeta$ . Subsequently, we rewrite it as

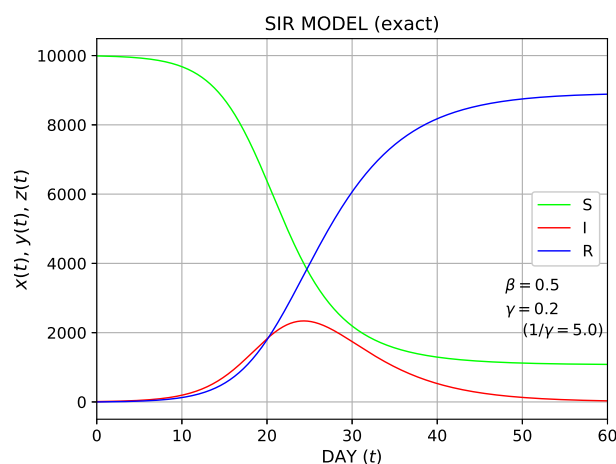
$$t = \int_1^{x(t)/N_1} \frac{d\zeta}{\zeta(-\beta - \gamma \ln \zeta + \beta(N_1/N)\zeta)}. \quad (52)$$

It is noteworthy that  $\beta$  appears here instead of  $\hat{\beta}$ .

Because we have obtained the form of the integral of one variable, we then calculate  $t(x)$ . We can perform a numerical integration by summing up an integrand with a sufficiently small step size, as follows:

$$\int_a^b f(\zeta) d\zeta \simeq \sum_{i=0}^{n-1} f(a + ih) h, \quad h = (b - a)/n. \quad (53)$$

It is noteworthy that the direction of the integral is negative because  $x(t)/N_1 < 1$  in Equation (52), and the integrand has a negative value. If  $t(x)$  is obtained, then  $y$  and  $z$  can be calculated using Equations (45) and (48), respectively. The calculated results of  $x(t)$ ,  $y(t)$ , and  $z(t)$  are shown in Figure 5, where the same conditions are used as in Figure 3a. As expected, the analytical solution and the numerical solution obtained using the Runge–Kutta method are the same.



**Figure 5.** The analytical solution of the SIR model. The conditions are the same as in Figure 3a.

Next, we analyze the property of the differential equation. We can investigate the behavior of  $t \rightarrow \infty$ . The variable  $\zeta$  in the limit of  $t \rightarrow \infty$ ,  $x(\infty)/N_1$ , is obtained when the denominator of the integrand of Equation (52) becomes zero. If we write the solution to the equation

$$\beta + \gamma \ln \zeta - \beta(N_1/N)\zeta = 0 \quad (54)$$

as  $\zeta^*$ , we have  $x(\infty) = N_1\zeta^*$ ,  $y(\infty) = 0$ , and  $z(\infty) = N - N_1\zeta^*$ . To solve an equation such as Equation (54) numerically, an iterative method such as Newton's method is used. However, for the case of a single variable, it is sufficient to plot the left-hand side of Equation (54) and find an intersection

with a zero. Figure 3b shows a comparison of the behaviors of the temporal evolutions of  $x(t)$ ,  $y(t)$ , and  $z(t)$  for  $\beta = 0.5, 0.45, 0.4, 0.35$ , and  $0.3$ , with  $\gamma = 0.2$  ( $N = 10,000, N_1 = 9900$ ). Figure 6 shows the solutions of Equation (54) under the same conditions. The solutions reproduce the limit values of  $x(t)$  for  $t \rightarrow \infty$  in Figure 3b. Using  $R_0$  defined in Equation (27), we can rewrite Equation (54) as

$$R_0 = -\frac{\ln \xi}{1 - \xi} \quad (55)$$

in the limit of  $N_1/N \rightarrow 1$ . This relation is known as the solution of the final size equation [19].

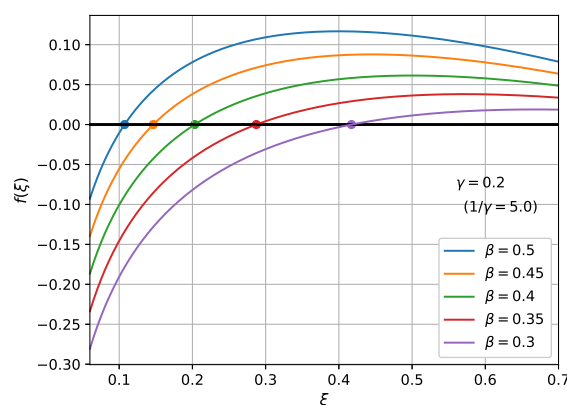


Figure 6. The solution of Equation (54).

The final number of the infected individuals,  $(z(\infty) = N - x(\infty))$ , is a variable of interest. The  $\beta$  dependence of this number is tabulated in Table 1. As shown, the final number of the infected individuals decreases with the infection rate  $\beta$ . Furthermore, the table presents the peak values of the infected individuals.

**Table 1.** The final number of the infected individuals:  $N = 10,000, N_1 = 9900, N_2 = 10, N_3 = 0, \gamma$  is fixed as 0.2. The peak values of the infected individuals are also shown.

	The Final Number of the Infected	The Infected ( $y(t)$ ) Peak Day	Peak Number	The Newly Infected ( $y_{\text{new}}(t)$ ) Peak Day	Peak Number
$\beta = 0.50$ ( $R_0 = 2.5$ )	8916	24.4	2339	21.2	587
$\beta = 0.45$ ( $R_0 = 2.25$ )	8524	28.3	1956	24.9	468
$\beta = 0.40$ ( $R_0 = 2.0$ )	7958	33.8	1539	30.2	351
$\beta = 0.35$ ( $R_0 = 1.75$ )	7117	42.2	1094	38.4	238
$\beta = 0.30$ ( $R_0 = 1.5$ )	5818	57.0	637	52.8	133

## 6. Analytical Approach of SIR Model with Vital Dynamics

The standard SIR model has been extended to include some other factors. One example is a model that considers birth and natural death processes [20,21]. The nonlinear equations are as follows:

$$\frac{dx(t)}{dt} = -\hat{\beta}x(t)y(t) + \mu(N - x(t)), \quad (56)$$

$$\frac{dy(t)}{dt} = \hat{\beta}x(t)y(t) - (\gamma + \mu)y(t), \quad (57)$$

$$\frac{dz(t)}{dt} = \gamma y(t) - \mu z(t), \quad (58)$$

where  $\mu$  represents the natural death rate, which is assumed to be equal to the birth rate. Harko et al. [15] studied this model and demonstrated that these simultaneous equations can be reduced to an Abel-type equation.

Next, we show a more straightforward derivation than that presented in the original paper [15]. Similar to the procedure used for the standard SIR model, we first eliminate a variable  $y(t)$  from Equations (56) and (57);

$$x \frac{d^2 x}{dt^2} - \left( \frac{dx}{dt} \right)^2 + (-\hat{\beta}x^2 + (\gamma + \mu)x + \mu N) \frac{dx}{dt} + Ax^3 + Bx^2 + Cx = 0, \quad (59)$$

where  $A = \mu\hat{\beta}$ ,  $B = -\mu N\hat{\beta} - \mu(\gamma + \mu)$ ,  $C = \mu N(\gamma + \mu)$ .

Next we introduce a function

$$\phi = \frac{dt}{dx}. \quad (60)$$

Subsequently, we obtain an equation for  $\phi$  as follows:

$$\frac{d\phi}{dx} + \frac{\phi}{x} = (-\hat{\beta}x + (\gamma + \mu) + \frac{\mu N}{x})\phi^2 - (Ax^2 + Bx + C)\phi^3. \quad (61)$$

By introducing a function

$$v = \phi x, \quad (62)$$

we obtain the equation for  $v$  as follows:

$$\frac{dv}{dx} = (-\hat{\beta} + (\gamma + \mu)\frac{1}{x} + \frac{\mu N}{x^2})v^2 - (A + \frac{B}{x} + \frac{C}{x^2})v^3. \quad (63)$$

An ordinary differential equation of this form is known as the Abel-type first-order differential equation of the first kind. The mathematical properties of the Abel-type equations have been investigated intensively. One may use an iterative solution, as demonstrated by Harko et al. [15]. If we set  $\mu = 0$ , this model becomes the standard SIR model; Equation (63) is reduced to Equation (A4) in the case of  $\mu = 0$ . Furthermore, some recent developments regarding the solution of the Abel equation of the first kind have been reported [22,23].

Next, we briefly analyze the properties of the SIR model with vital dynamics. We may solve Equations (56) and (57) simultaneously or solve an equation comprising one variable, i.e., Equation (63). For the analytical approach, it is beneficial to introduce the following function:

$$u = \frac{1}{v}. \quad (64)$$

Subsequently, we can rewrite Equation (63) as

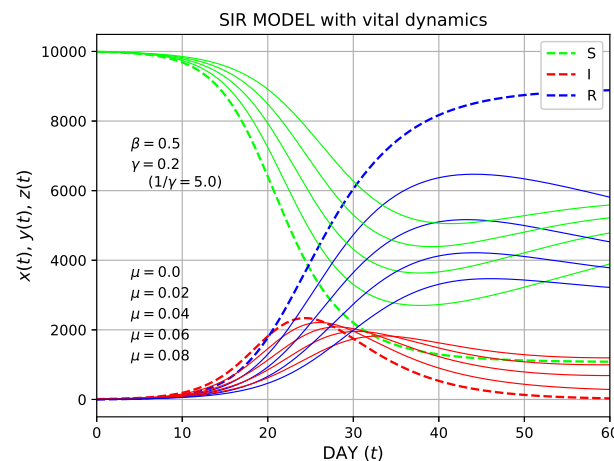
$$\frac{du}{dx} = (\hat{\beta} - (\gamma + \mu)\frac{1}{x} - \frac{\mu N}{x^2}) + (A + \frac{B}{x} + \frac{C}{x^2})\frac{1}{u}, \quad (65)$$

or

$$u \frac{du}{dx} = (\hat{\beta} - (\gamma + \mu)\frac{1}{x} - \frac{\mu N}{x^2})u + (A + \frac{B}{x} + \frac{C}{x^2}). \quad (66)$$

This type of equation is known as the Abel equation of the second kind.

We plot  $x(t)$ ,  $y(t)$ , and  $z(t)$  of the SIR model with vital dynamics by solving Equations (56) and (57) simultaneously or solving Equation (65), in Figure 7. Both approaches yielded the same results, as expected. The parameters  $\mu$  varied as 0.0, 0.02, 0.04, 0.06, and 0.08, whereas  $\beta$  and  $\gamma$  were fixed as 0.5 and 0.2, respectively. In the plot, the data of  $\mu = 0.0$  are shown by thick dashed curves. As shown, a minimum occurred in  $x(t)$ , and  $y(t)$  remained non-zero for  $t \rightarrow \infty$ . This represents a disease in the endemic steady state. New births can provide more susceptible individuals, thereby resulting in a sustained epidemic.



**Figure 7.** The solution of the SIR model with vital dynamics. The parameter  $\mu$  is varied as 0.0, 0.02, 0.04, 0.06, and 0.08, and the data of  $\mu = 0.0$  are shown by thick dashed curves.

## 7. Discussion and Conclusions

We explained the mathematical model of the spread of epidemic diseases. It is noteworthy that in the SIR model, a closed society was assumed, and all the individuals have the same contact possibility. Hence, the obtained results are valid only within such a model.

Next, we consider the meanings of the infected (I) and the recovered (R). In this model, recovered (R) refers to individuals who do not infect others. Hence, the individuals who have been quarantined are regarded as recovered individuals. Therefore,  $\gamma$  is called the rate of recovery or the rate of quarantine. To reduce  $R_0$ , we may reduce  $\beta$  because  $R_0$  is expressed by  $\beta/\gamma$ . This can be realized by social distancing (physical distancing). Because the inverse of the denominator,  $\gamma$ , is the average time of infection, shortening the time of confirming the infection by testing provides the same effect as social distancing. In Equations (21)–(23),  $\beta$  is multiplied by  $x(t)y(t)$ , whereas  $\gamma$  is only multiplied by  $y(t)$ .

Next, we discuss the reproduction number. The Robert Koch Institute (RKI) of Germany publicizes the reproduction number  $R$  daily. The point estimate of  $R$  for a specific day is estimated as the quotient of the number of incident cases on that day divided by the number of incident cases four days earlier. The incident cases are not directly extracted from the notification system, but they are estimated using statistical analysis to predict the true onset of illness. To reduce random effects, a moving 4-day or 7-day average is used. The reproduction number  $R$  of the RKI is clearly defined and reflects the measure of disease spread; i.e.,  $R$  is either greater or smaller than 1. However, it is noteworthy that the absolute value of such a time-varying reproduction number depends on the definition.

As for the mathematical models of epidemics, some extended models have been developed, such as the susceptible-infectious-recovered-deceased (SIRD) model [24–26]. This model differentiates between recovered (i.e., individuals who have survived the disease and are now immune) and deceased individuals. Another example is the susceptible-exposed-infectious-recovered (SEIR) model [27–30], where an “exposed” category is added. It considers the situation involving a significant incubation period, during which individuals have been infected but are not yet infectious.

We learned not only the numerical solution of the SIR model, but also the exact analytical solution of the model. Furthermore, we treated the generalized SIR model with vital dynamics. Students can learn differentiation and integration using this tutorial. Furthermore, it will be beneficial if they indicate interest in social problems.

**Author Contributions:** These authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors thank Hiroyuki Mori and Yoshinori Hayakawa for valuable discussions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Python Code of the Numerical Solution of the SIR Model Using the Runge–Kutta Method

```
# coding: utf-8
import math
import matplotlib.pyplot as plt

N = 10000
x = N-10
y = 10
z = 0

taxis=[ ]
xaxis=[ ]
yaxis=[ ]
zaxis=[ ]

beta=0.5/N
gamma=0.2
dt=0.001

t = 0
cnt=0
while t<60:
    if cnt%100==0:
        taxis.append(t)
        xaxis.append(x)
        yaxis.append(y)
        zaxis.append(z)
    # step 1
    kx1 = - beta*x*y
    ky1 = beta*x*y - gamma*y
    # step 2
    t2 = t+dt
    x2 = x + kx1*dt
    y2 = y + ky1*dt
    kx2 = - beta*x2*y2
    ky2 = beta*x2*y2 - gamma*y2
    # update
    x = x + (kx1+kx2)*dt/2
    y = y + (ky1+ky2)*dt/2
    z = N - x - y
    t = t + dt
    cnt = cnt + 1

plt.title("SIR MODEL")
plt.plot(taxis,xaxis, color=(0,1,0), linewidth=1.0, label='S')
plt.plot(taxis,yaxis, color=(1,0,0), linewidth=1.0, label='I')
plt.plot(taxis,zaxis, color=(0,0,1), linewidth=1.0, label='R')
```

```
plt.xlim(0,60)
plt.legend()
plt.xlabel('DAY')
plt.grid(True)
plt.show()
```

## Appendix B. Solutions of Problems

**Solution A1** (Solution of Problem 1). We start with the equation

$$\frac{d\phi}{dx} + \frac{\phi}{x} = (\gamma - \hat{\beta}x)\phi^2. \quad (\text{A1})$$

If we introduce  $v = \phi x$ , we have

$$\frac{d}{dx}\left(\frac{v}{x}\right) = \frac{1}{x} \frac{dv}{dx} - \frac{v}{x^2} = \frac{1}{x} \frac{dv}{dx} - \frac{\phi}{x}. \quad (\text{A2})$$

Then, Equation (A1) can be rewritten as

$$\frac{1}{x} \frac{dv}{dx} = (\gamma - \hat{\beta}x) \frac{v^2}{x^2}. \quad (\text{A3})$$

If we divide both sides by  $v^2/x$ , we obtain

$$-\frac{1}{v^2} \frac{dv}{dx} = \frac{d}{dx}\left(\frac{1}{v}\right) = -\frac{\gamma}{x} + \hat{\beta}. \quad (\text{A4})$$

The integration with respect to  $x$  yields

$$\frac{1}{v} = -\gamma \ln x + \hat{\beta}x + C_1, \quad (\text{A5})$$

where  $C_1$  is an integral constant. Then, we obtain  $\phi$  as

$$\phi = \frac{v}{x} = \frac{1}{x(C_1 - \gamma \ln x + \hat{\beta}x)}. \quad (\text{A6})$$

**Solution A2** (Solution of Problem 2). If we differentiate the equation of the solution, we can confirm that the given solution satisfies the differential equation. Differentiating the both sides of Equation (51), we obtain

$$1 = \frac{dx}{dt} \frac{1}{x(-\hat{\beta}N - \gamma \ln(x/N_1) + \hat{\beta}x)}. \quad (\text{A7})$$

Using Equation (48), we solve  $dx/dt$  as

$$\frac{dx}{dt} = x(-\hat{\beta}N + \gamma(\hat{\beta}/\gamma)z + \hat{\beta}x) = \hat{\beta}x(-N + z + x) = -\hat{\beta}xy. \quad (\text{A8})$$

We can also show for  $dz/dt$  that

$$\frac{dz}{dt} = -\frac{\gamma}{\hat{\beta}} \frac{x'}{x} = -\frac{\gamma}{\hat{\beta}}(-\hat{\beta}y) = \gamma y. \quad (\text{A9})$$

Thus, we confirmed that the solution satisfies the differential equation.

## References

- Harapan, H.; Itoh, N.; Yufika, A.; Winardi, W.; Keam, S.; Te, H.; Megawati, D.; Hayati, Z.; Wagner, A.L.; Mudatsir, M. Coronavirus disease 2019 (COVID-19): A literature review. *J. Infect. Public Health* **2020**, *13*, 667–673. [\[CrossRef\]](#)
- Burki, T.K. Coronavirus in China. *Lancet Respir. Med.* **2020**, *8*, 223. [\[CrossRef\]](#)
- Tang, B.; Bragazzi, N.L.; Li, Q.; Tang, S.; Xiao, Y.; Wu, J. An updated estimation of the risk of transmission of the novel coronavirus (2019-nCoV). *Infect. Dis. Model.* **2020**, *5*, 248–255. [\[CrossRef\]](#)
- Atkeson, A. *What Will Be the Economic Impact of COVID-19 in the US? Rough Estimates of Disease Scenarios*; NBER Working Paper No. 26867; The National Bureau of Economic Research: Cambridge, MA, USA, 2020.
- Stock, J.H. *Data Gaps and the Policy Response to the Novel Coronavirus*; NBER Working Paper No. 26902; The National Bureau of Economic Research: Cambridge, MA, USA, 2020.
- Ivorra, B.; Ferrandez, M.R.; Vela-Perez, M.; Ramos, A.M. Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China. *Commun. Nonlinear Sci. Numer. Simul.* **2020**, *88*, 105303. [\[CrossRef\]](#)
- Prem, K.; Liu, Y.; Russell, T.W.; Kucharski, A.J.; Eggo, R.M.; Davies, N.; Jit, M.; Klepac, P. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *Lancet Public Health* **2020**, *5*, e261–e270. [\[CrossRef\]](#)
- Kermack, W.O.; McKendrick, A.G. A Contribution to the Mathematical Theory of Epidemics, I. *Proc. Roy. Soc. Lond. A* **1927**, *115*, 700–721.
- Kermack, W.O.; McKendrick, A.G. A Contribution to the Mathematical Theory of Epidemics, II. *Proc. Roy. Soc. Lond. A* **1932**, *138*, 55–83.
- Diekmann, O. Thresholds and travelling waves for the geographical spread of infection. *J. Math. Biol.* **1978**, *6*, 109–130. [\[CrossRef\]](#)
- Metz, J.A.J. The epidemic in a closed population with all susceptibles equally vulnerable; some results for large susceptible populations and small initial infections. *Acta Biotheor.* **1978**, *27*, 75–123. [\[CrossRef\]](#)
- Thieme, H.R. A model for the spread of an epidemic. *J. Math. Biol.* **1977**, *4*, 337–351. [\[CrossRef\]](#)
- Inaba, H. Kermack and McKendrick revisited: The variable susceptibility model for infectious diseases. *Jpn. J. Indust. Appl. Math.* **2001**, *18*, 273–292. [\[CrossRef\]](#)
- Diekmann, O.; Heesterbeek, J.A.P.; Metz, J.A.J. The Legacy of Kermack and McKendrick. In *Epidemic Models, their Structure and Relation to Data*; Mollison, D., Ed.; Cambridge University: Cambridge, UK, 1994.
- Harko, T.; Lobo, F.S.N.; Mak, M.K. Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Appl. Math. Comput.* **2014**, *236*, 184–194. [\[CrossRef\]](#)
- Diekmann, O.; Heesterbeek, J.A.P.; Metz, J.A.J. On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **1990**, *28*, 365–382. [\[CrossRef\]](#) [\[PubMed\]](#)
- Diekmann, O.; Heesterbeek, J.A.P. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*; John Wiley and Sons: Chichester, UK, 2000.
- Dietz, K. The estimation of the basic reproduction number for infectious diseases. *Stat. Methods Med. Res.* **1993**, *2*, 23–41. [\[CrossRef\]](#) [\[PubMed\]](#)
- Metz, J.A.J.; Diekmann, O. (Eds.) The Dynamics of Physiologically Structured Populations. In *Lecture Notes in Biomathematics* 68; Springer: Heidelberg, Germany, 1986.
- Brauer, F.; Castillo-Chávez, C. *Mathematical Models in Population Biology and Epidemiology*; Springer: New York, NY, USA, 2001.
- Murray, J.D. *Mathematical Biology: I. An Introduction*; Springer: New York, NY, USA, 2002.
- Panayotounakos, D.E.; Zampoutis, T.I. Construction of exact parametric or closed form solutions of some unsolvable classes of nonlinear ODEs (Abel's nonlinear ODEs of the first kind and relative degenerate equations). *Int. J. Math. Math. Sci.* **2011**, *2011*, 387429. [\[CrossRef\]](#)
- Salinas-Hernández, E.; Muñoz-Vega, R.; Sosa, J.C.; López-Carrera, B. Analysis to the Solutions of Abel's Differential Equation of the First Kind under the Transformation  $y = u(x)z(x) + v(x)$ . *Appl. Math. Sci.* **2013**, *7*, 2075–2092. [\[CrossRef\]](#)
- Hethcote, H.W. The Mathematics of Infectious Diseases. *SIAM Rev.* **2000**, *42*, 599–653. [\[CrossRef\]](#)

25. Fernández-Villaverde, J.; Jones, C.I. *Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities*; NBER Working Paper No. 27128; The National Bureau of Economic Research: Cambridge, MA, USA, 2020.
26. Batista, M. Estimation of the final size of the COVID-19 epidemic. *medRxiv* **2020**.10.1101/2020.02.16.20023606. [[CrossRef](#)]
27. Liu, W.-M.; Hethcote, H.W.; Levin, S.A. Dynamical behavior of epidemiological models with non-linear incidence rate. *J. Math. Biol.* **1987**, *25*, 359–380. [[CrossRef](#)]
28. Hethcote, H.W.; van den Driessche, P. Some epidemiological models with nonlinear incidence. *J. Math. Biol.* **1991**, *29*, 271–287. [[CrossRef](#)]
29. Li, M.Y.; Graef, J.R.; Wang, L.; Karsai, J. Global dynamics of a SEIR model with varying total population size. *Math. Biosci.* **1999**, *160*, 191–213. [[CrossRef](#)]
30. Peng, L.; Yang, W.; Zhang, D.; Zhuge, C.; Hong, L. Epidemic analysis of COVID-19 in China by dynamical modeling. *arXiv* **2020**, arXiv:2002.06563.



© 2020 by the authors. 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 (<http://creativecommons.org/licenses/by/4.0/>).