

---

## MATHEMATICAL MODELLING OF INFECTIOUS DISEASES

*Coronavirus (COVID-19, SARS-CoV-2)*

---

Ali Al-Hayki\*

✉ ali.hayki@gmail.com

➲ alihayki

㏌ ali-al-hayki

23/03/2020

---

\*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies.

## Abstract

In this paper we are going to explain the approaches for the mathematical modelling of the spread of infectious diseases such as Coronavirus (COVID-19, SARS-CoV-2). This paper will explain the classic compartmental models in the epidemics literature such as SIR, SIS, SEIR etc. It also explains key concepts such as the *reproduction number*, the *final size equation*, and the key theorems that are implied by these concepts.

## Acknowledgement

I am writing this acknowledgement on exactly the morning of **1st of April 2020**. The below is based up to this time. Any update or new development will be under an **update** (highlighted in bold) explaining the development of the spread with a corresponding date. Consider this section as a ‘diary’ preserving the events what might go down in history as modern humanity’s most critical events.

Technology allowed working from home possible, Last time this was done was in University (Possibly the instinct of writing this paper is due to that time in history)

I have started writing this paper on the 18th of March 2020, the date that the UK government enforced ‘lock-down’ - where all the UK work force was compelled to work from home if they could.

**01.04.2020:** The contactless payment limit for in-store spending is raised from £30 to £45.

**06.04.2020:** Prime Minister Boris Johnson is taken into intensive care at St Thomas’ Hospital.

**17.04.2020:** The lock down has been extended for another 3 weeks.

**23.04.2020:** The first human trials of a coronavirus vaccine in Europe begin in Oxford.

**01.05.2020:** Hancock announces that fertility clinics will be allowed to open again from 11 May.

**15.06.2020:** Non-essential shops are opening.

**Keywords** First keyword · Second keyword · More

## Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Epidemic Theory Key Terms</b>	<b>4</b>
2.1	Basic Reproduction Number ( $\mathcal{R}_0$ ) . . . . .	4
2.2	Effective Reproductive Number ( $\mathcal{R}$ ) . . . . .	5
2.3	Herd Immunity . . . . .	5
2.4	Epidemics . . . . .	6
2.5	Epidemic curves . . . . .	6

---

2.6	Index Case and Generation Time . . . . .	6
2.7	Exception Reporting . . . . .	7
2.8	Significant Clusters . . . . .	7
<b>3</b>	<b>Compartmental Modelling</b>	<b>7</b>
3.1	Introduction . . . . .	7
3.2	Level of Compartments . . . . .	7
<b>4</b>	<b>Compartmental Framework</b>	<b>8</b>
4.1	SIR Framework . . . . .	8
4.2	SIR Framework Overview . . . . .	8
4.3	Mathematical Modelling of the SIR Framework . . . . .	11
4.4	SIS Framework Solution . . . . .	12
4.5	SIR Framework Solution . . . . .	15

## 1 Introduction

Coronaviruses are a group of related viruses that cause diseases in mammals and birds. In humans, coronaviruses cause respiratory tract infections that can be mild, such as some cases of the common cold (among other possible causes, predominantly rhinoviruses), and others that can be lethal, such as SARS, MERS, and COVID-19. Symptoms in other species vary: in chickens, they cause an upper respiratory tract disease, while in cows and pigs they cause diarrhea. As of the date of writing this paper, there are yet to be vaccines or antiviral drugs to prevent or treat human coronavirus infections.

Coronavirus has been causing a lot of suffering, public health emergencies, anxiety and economic challenges. It will be of great interest to discuss how mathematical modelling is used to understand and contain the spread of such infectious diseases.

Infectious diseases spread via an infected host to a non-infected host - it very similar to a reproduction process. The agent could be a virus, bacteria, protozoa etc. The role of mathematical modelling is to help estimate the number of infectives over time. Amature modelers tend to use exponential curves to describe the spread of the desease. But this is not how actual mathematical modelling is done in practice. The increase in the number of infectives in the intial stages of the spread looks exponantial, hence the rush to fit an exponential curve. Curve fitting might be good for some applications such as scaremongering<sup>2</sup>. However the essential dynamics of the essential dynamics of the essential disease process needs to be modelled in order to gain insights into the drivers of the spread and inform the containment strategies. The questions that one can answer with mathematics could be:

Once we have reasonable mathematical model of the spread, then we can evaluate the relative effectiveness of various containment strategies. Mathematical models don't require a lot of data, and hence can be quickly put to use. The list of uses of mathematics is boundless, for example at later stages when one has a vaccine, then models can also help determine the optimal vaccination strategy to achieve herd immunity. For example, you are not supposed to capture all the details in a mathematical model - just the essentials. Infectious diseases involve microscopic processes, the little creatures such as virus and bacteria, and their actions and macro variables such as impact on communities. Capturing everything is not going to give you a tractable model, so the principle of parsimony is the key.

Before going into the mathematical modelling the next sections outlines some key Epidemic theory terms.

## 2 Epidemic Theory Key Terms

### 2.1 Basic Reproduction Number ( $\mathcal{R}_0$ )

The basic reproduction number ( $\mathcal{R}_0$ ) is used to measure the transmission potential of a disease. It is the average number of secondary infections produced by a typical case of an infection in a population where everyone is susceptible.<sup>1</sup> For example, if the  $R_0$  for measles in a population is 15, then we would expect each new case of measles to produce 15 new secondary cases (assuming everyone around the case was susceptible).  $R_0$  excludes new cases produced by the secondary cases.

<sup>2</sup>The spreading of frightening or ominous reports or rumours.

The basic reproductive number is affected by several factors:

The rate of contacts in the host population The probability of infection being transmitted during contact The duration of infectiousness. In general, for an epidemic to occur in a susceptible population  $\mathcal{R}_0$  must be  $>1$ , so the number of cases is increasing.<sup>1</sup>

In many circumstances not all contacts will be susceptible to infection. This is measured by the effective reproductive rate ( $\mathcal{R}$ )

## 2.2 Effective Reproductive Number ( $\mathcal{R}$ )

A population will rarely be totally susceptible to an infection in the real world. Some contacts will be immune, for example due to prior infection which has conferred life-long immunity, or as a result of previous immunisation. Therefore, not all contacts will become infected and the average number of secondary cases per infectious case will be lower than the basic reproduction number. The effective reproductive number ( $\mathcal{R}$ ) is the average number of secondary cases per infectious case in a population made up of both susceptible and non-susceptible hosts. If  $\mathcal{R} > 1$ , the number of cases will increase, such as at the start of an epidemic. Where  $R=1$ , the disease is endemic, and where  $R<1$  there will be a decline in the number of cases.

The effective reproduction number can be estimated by the product of the basic reproductive number and the fraction of the host population that is susceptible ( $x$ ). So:

$$R = R_0 X_s$$

For example, if  $R_0$  for influenza is 12 in a population where half of the population is immune, the effective reproductive number for influenza is  $12 \times 0.5 = 6$ . Under these circumstances, a single case of influenza would produce an average of 6 new secondary cases.<sup>1</sup>

To successfully eliminate a disease from a population, R needs to be less than 1.

## 2.3 Herd Immunity

Herd immunity occurs when a significant proportion of the population (or the herd) have been vaccinated (or are immune by some other mechanism), resulting in protection for susceptible (e.g. unvaccinated) individuals. The larger the number of people who are immune in a population, the lower the likelihood that a susceptible person will come into contact with the infection. It is more difficult for diseases to spread between individuals if large numbers are already immune as the chain of infection is broken.

The herd immunity threshold is the proportion of a population that need to be immune in order for an infectious disease to become stable in that community. If this is reached, for example through immunisation, then each case leads to a single new case ( $R=1$ ) and the infection will become stable within the population.

If the threshold for herd immunity is surpassed, then  $R<1$  and the number of cases of infection decreases. This is an important measure used in infectious disease control and immunisation and eradication programmes.

## 2.4 Epidemics

An epidemic is defined as an increase in the frequency of occurrence of a disease in a population above its baseline, or expected level, in a given time period.<sup>2</sup> The term is used broadly and the number of cases and time period are often unspecified. It is generally more widespread than an outbreak, which usually implies two or more epidemiologically linked cases, although the two terms have been used interchangeably. Additionally, the term has also been used to describe increasing levels of non-communicable disease, such as an ‘epidemic of cardiovascular disease.’

The definition above is general, but the term has been defined quantitatively for certain infections and a threshold is selected above which the term ‘epidemic’ is applied. For example, in England levels of influenza are routinely monitored drawing on data from GP consultations and lab diagnoses. The Royal College of General Practitioners (RCGP) has defined the baseline threshold for ‘normal seasonal activity’ in England as 30 to 200 GP consultations for influenza-like illness per week per 100,000 population. The epidemic threshold would be reached if the number of consultations surpassed 200 per week per 100,000.<sup>3</sup>

Other thresholds are used in epidemic theory. The Critical Community Size (CCS) is the total population size needed to sustain an outbreak once it has appeared, and the Outbreak Threshold is the number of infected individuals that are needed to ensure that an outbreak is unlikely to go extinct without intervention.<sup>4</sup>

## 2.5 Epidemic curves

An epidemic curve is a graph that illustrates the distribution of the onset of new cases of an infectious disease in relation to the onset of illness. The time interval for the onset of illness used will be determined by the incubation period (see “Definitions including: incubation, communicability and latent period; susceptibility, immunity, and herd immunity” in Section 2G for a definition of this and related terms).

Epidemic curves are a useful tool in outbreak investigations, helping to:

Determine the type of epidemic (continuous source, point source, propagated) Determine the difference between the maximum and minimum incubation period Estimate the likely time of exposure, and thus help focus investigation on a particular time period Determine the incubation period in cases where the time of exposure is known Identify outliers (below) It would be worth looking here for examples of different types of epidemic curve<sup>5</sup>.

## 2.6 Index Case and Generation Time

The original case of an outbreak is labelled as the primary case. Secondary cases contract the infection from primary cases, and tertiary cases contracted theirs from secondary cases, and so on. The index case is the term given to the first recognised case, or cases, in an outbreak. Note that the index case may not turn out to be a primary case, and the primary case of an outbreak may only be identified on further investigation, if at all. The generation time describes the duration from the onset of infectiousness in the primary case to the onset of infectiousness in a secondary case (infected by the primary case).

## 2.7 Exception Reporting

Infectious disease surveillance ensures that the frequency of certain diseases or symptoms are monitored. If there is an abrupt increase in the frequency of a particular disease, outside of predefined limits, it will be flagged as an “exception” and thus functions as an early indicator that further investigation is required.

## 2.8 Significant Clusters

A cluster, or significant cluster, is an aggregation of cases related in time or place that is suspected to be greater than the number expected (although the “expected” number may not be known). The term can relate to both communicable and non-communicable disease. Clusters can be identified using spot maps (where each case is represented on a map by a coloured dot), although such maps may show apparent “clusters” in areas that are densely populated (and thus would have a higher number of expected cases). Alternatively, maps which colour areas in different shades depending on the rate of disease in each area can be used, although if the defined areas are too large it will mask real clusters.

1. Will the disease become an epidemic? (i.e. large increase in the number of cases within a short period of time)
2. Will the disease become an endemic? (i.e. disappearance or constant prevalence)

## 3 Compartmental Modelling

### 3.1 Introduction

The classical framework for modelling infectious diseases is the so called compartmental modelling approach. Compartmental models are a technique used to simplify the mathematical modelling of infectious disease.

Compartmental models may be used to predict properties of how a disease spreads, for example the prevalence (total number of infected) or the duration of an epidemic. Also, the model allows for understanding how different situations may affect the outcome of the epidemic, e.g., what the most efficient technique is for issuing a limited number of vaccines in a given population.

The models are usually investigated through ordinary differential equations (which are deterministic), but can also be viewed in a stochastic framework, which is more realistic but also more complicated to analyse.

### 3.2 Level of Compartments

The population is divided into compartments, with the assumption that every individual in the same compartment has the same characteristics, below are the compartments:

**Susceptibles** Individuals capable of contracting the disease and becoming themselves infectives;

**Infactives** Individuals capable of transmitting the disease to susceptibles;

**Removed** Individuals which, having contracted the disease, have died or, if recovered, are permanently immune, or have been isolated, thus being unable to further transmit the disease.

## 4 Compartmental Framework

### 4.1 SIR Framework

1. **Susceptibles:** In the beginning we have the population of susceptibles (i.e. the individuals that can get infected), which is the first compartment illustrated in figure 1.
2. **Outbreak:** Suddenly there is a virus outbreak<sup>3</sup> and one of the individuals gets infected, which is the second compartment illustrated in figure 2.
3. **Spread of Infection:** The infected individual then starts to infect other individuals, effectively increasing the number of infective (i.e. expanding the second compartment), this expansion is illustrated in figure 3.
4. **Stability:** The spread of infection stabilises or plate's at some point in time, this is illustrated in figure 4.
5. **Removed (Recovery or Death):** Some of the infective recovers or die. This is illustrated in figure 5.
6. **Expansion of Removed:** The recovered compartment grows until covers all the infected compartment, this is both people who have recovered with immunity to the virus, or have died.

This is called the SIR Framework, one of the most elementary compartmental models. The goal of this model is to track the number of hosts in each of the three compartments at any given time  $t$ , and we denote these numbers by  $S_t$ ,  $I_t$ , and  $R_t$  accordingly. The general flow is from susceptibles to infective to removed, this is illustrated in figure 10. This framework is for diseases where upon recovery, one develops long term immunity to the disease. Generally this is the case for virus based infectious diseases. For diseases caused by bacteria one generally does not develop immunity, so the SIS framework is used for the modelling of these diseases, so the flow is from susceptibles to infective and then back to Susceptibles see figure 9. Many variations are possible within the compartmental framework. For example, for some diseases, one only gets short term immunity, so there would be an additional flow from R to S, leading to the SIRS framework, this is illustrated in figure 8. It could be that there is no recovery , and you will then have the SI framework, with just two compartments. This same framework can also be used to model the spread of fake news, so ‘infected’ would then be the people who have the fake news and are spreading it, and the ‘recovered’ are the enlightened. So many variations are possible, but one has to base the choice on how the dynamics of how the infection works.

### 4.2 SIR Framework Overview

Generally, the way infectious diseases spread is through the following:

1. **Contact with Infective:** you get the agent, virus or bacteria, via contact with an infective.
2. **Infectious:** the virus will try to overcome the immune system, and at some point will become Infectious.
3. **Symptoms:** You may or may not have the symptoms yet, but you can spread the disease to others, though the infectivity might be lower than when you start showing symptoms.
4. **End of Infectiousness:** As you recover, you will become non-infective.
5. **Removed/Recovered:** You recover.

<sup>3</sup>the reason for the outbreak is beyond the scope of this paper

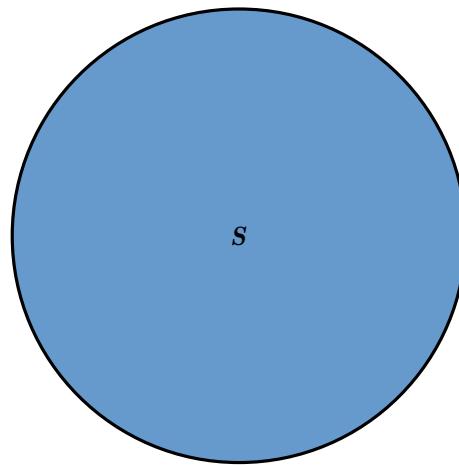


Figure 1: Population of Susceptibles

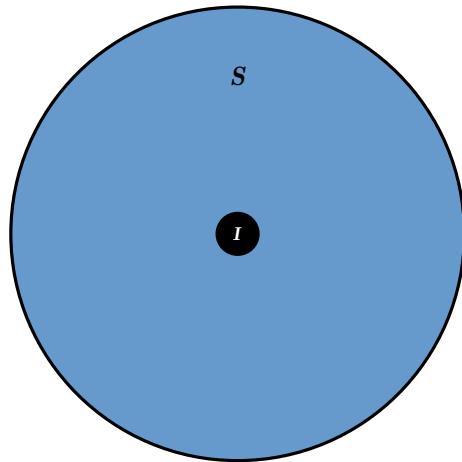


Figure 2: A Single Infection

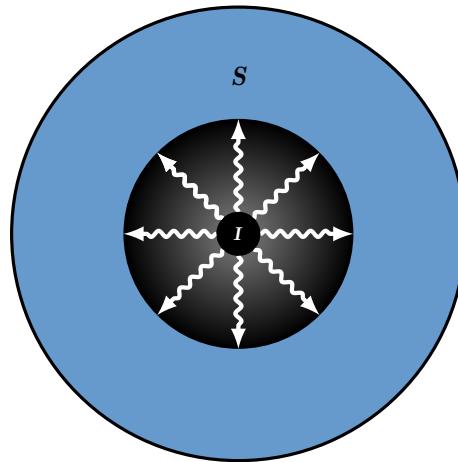


Figure 3: Infection Spreading

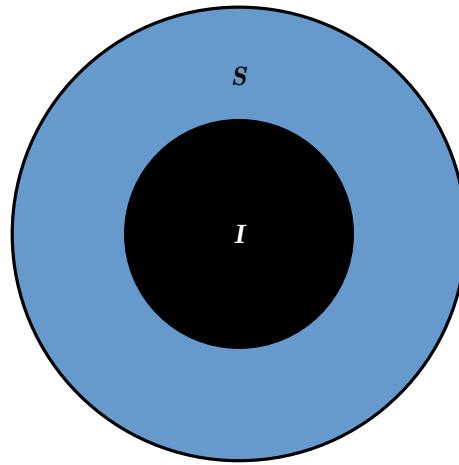


Figure 4: Infection Settles

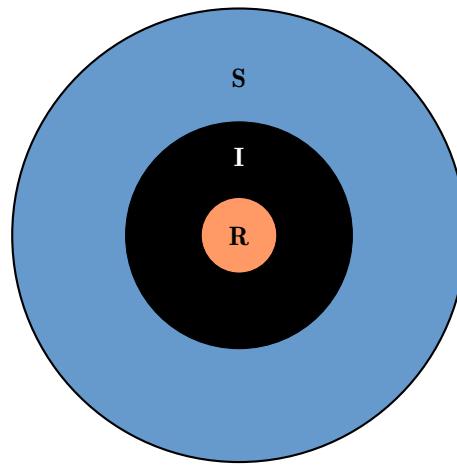


Figure 5: Infected start to be ‘Removed’ (i.e. either died or recovered)

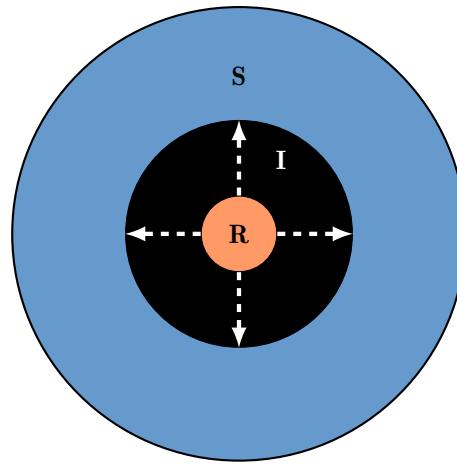


Figure 6: ‘Removed’ start to expand within infected

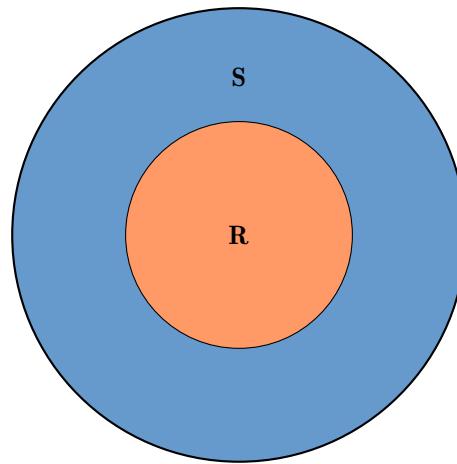


Figure 7: All infected are Removed (i.e. either died or recovered)

You can add another compartment - ‘Exposed’. Then the SIR will become SEIR, and SIS will become SEIS etc.

### 4.3 Mathematical Modelling of the SIR Framework

Here we are going to be using deterministic calculus, which means that the follows will take the form of differential equations. The model is known as Kermack–McKendrick model. Let’s start with the flow from  $S$  to  $I$ . Individuals flow from  $S$  to  $I$  when infectives mix with susceptibles, this is similar to a chemical reaction and the law of mass action applies. Now if there are  $n$  susceptibles  $\{S_1, S_2, S_3 \dots S_n\}$ , and  $m$  infectives  $\{I_1, I_2, I_3 \dots I_m\}$ , then the number of interactions will be of the scale of  $S_t I_t$ . These will be vary with time as the individuals flow between the compartments. So the rate of change from  $S$  to  $I$ , which is the change in the number of susceptibles with respect to time, will be proportional to:

$$\frac{dS_t}{dt} = -\beta S_t I_t \quad (1)$$

Here the  $\beta$  represents both the contact rate between the individuals and the probability of transmission of the virus. Similarly, let’s represent the rate of flow from  $I$  to  $R$ , by  $\gamma$ . The rate of change of  $I$  is captured as the inflow from  $S$  minus the outflow to the recovery or dead compartment:

$$\frac{dI_t}{dt} = \beta S_t I_t - \gamma I_t \quad (2)$$

Finally the rate of change  $R$  is just equal to the inflow from the infectives compartment:

$$\frac{dR_t}{dt} = \gamma I_t \quad (3)$$

The sum of these three compartments gives us the total population. The total population:

$$N_t = S_t + I_t + R_t \quad (4)$$

The total population can change over time due to vital population dynamics, births and deaths, but since we are interested in modelling epidemics, which only lasts for a few months, we can ignore the natural births and deaths, and assume  $N$  is constant over time. But it is easy to add the births and deaths if one likes.

These differential equations are most commonly viewed in terms of population density instead of the size of the compartments. For example, you will have division by  $N$  on the right hand side of the first differential, but it does not much difference to the analysis as the total population is constant, so it is just a scaling issue. Other possible variations make the real difference by enriching the framework. For example, one can make the infectivity a function of time since infection, so now we can get a more realistic model of the actual infection process.

$$\int \beta(\tau) \frac{dS(t-\tau)}{dt} d\tau \quad (5)$$

So this equation is essentially taking the infected population, along with the time since they got infected, which is captured by  $(t - \tau)$ . The derivative  $\frac{dS_t}{dt}$  captures the flows to the infected compartment, and then multiplying by the infectivity  $\beta$ , which is now a function of time since infection. This system is depicted graphically in figure 11

In the simpler version we are considering here, we assumed all currently infected people have the same infectivity  $\beta$ , irrespective of time since infection.

For another variation, it is expected that the contact rate, the infectivity rate and the recovery rate could be different for different strata, such as different age groups. So multiple strata can be introduced into the framework. However solving the system must be known, which may not be as simple, which may lead to resorting to numerical methods.

#### 4.4 SIS Framework Solution

We can solve the SIR system, but let's consider the equivalent SIS first, which is easier because we only have two compartments, the susceptibles and the infected in equation ???. The susceptibles compartment equation is modified as the recovered move back into the susceptibles:

$$\frac{dS_t}{dt} = -\beta S_t I_t - \gamma I_t \quad (6)$$

Now  $N$ , which is the total population, is fixed:

$$N_t = S_t + I_t \quad (7)$$

So we need to solve for either  $S$  or  $I$ .

$$\begin{aligned} \frac{dI_t}{dt} &= \beta S_t I_t - \gamma I_t \\ &= \beta(N - I_t) I_t - \gamma I_t && \text{substitute } (N_t - I_t) \text{ into } S_t \\ &= \beta N I_t - \beta I_t^2 - \gamma I_t && \text{open brackets} \\ &= (\beta N - \gamma) I_t - \beta I_t^2 && \text{combine items} \\ \frac{dI_t}{dt} &= (\beta N - \gamma) I_t - \beta I_t^2 \end{aligned} \quad (8)$$

We just have to solve for  $I$  now, and we  $S$  plus  $I$  equals  $N$ , which is a constant, so we can determine  $S$  easily. This flow is predicted in figure 12

Now, the differential equation 8 is the famous Bernoulli equation. It is not a linear equation but one can easily linearise it by considering the following transformation:

$$I_t = \frac{1}{y} \quad (9)$$

$$\frac{dI_t}{dt} = \frac{1}{y^2} \frac{dy}{dt} \quad (10)$$

Now let's substitute these into equation 8:

$$\frac{1}{y^2} \frac{dy}{dt} = (\beta N - \gamma) \frac{1}{y} - \beta \frac{1}{y^2} \quad (11)$$

$$-\frac{dy}{dt} = (\beta N - \gamma)y - \beta \quad \text{multiply by } y^2 \quad (12)$$

$$\frac{dy}{dt} + (\beta N - \gamma)y = \beta \quad \text{rearrange to show in the familiar linear differential equation form} \quad (13)$$

This equation can be solved via the integrating factor method. The integrating factor for this particular case is defined as the exponential of the integral of the coefficient of  $y$ :

$$f = e^{\int_0^t (\beta N - \gamma) ds} = e^{(\beta N - \gamma)t} \quad (14)$$

$$\frac{dy}{dt} + (\beta N - \gamma)y = \beta \quad (15)$$

$$e^{(\beta N - \gamma)t} \frac{dy}{dt} + e^{(\beta N - \gamma)t} (\beta N - \gamma)y = e^{(\beta N - \gamma)t} \beta \quad \text{Multiplying by integrating factor} \quad (16)$$

$$\frac{d}{dt}(e^{(\beta N - \gamma)t} y) = e^{(\beta N - \gamma)t} \beta \quad \text{rewrite as product of 2 terms} \quad (17)$$

$$e^{(\beta N - \gamma)t} y \Big|_0^t = \beta \int_0^t e^{(\beta N - \gamma)u} du \quad \text{Integrate from 0 to } t \quad (18)$$

The result of the integral on the right hand side will differ depending on whether the term in the exponent is zero. If it is not equal to zero, then we know the integral of the exponential will produce this expression:

$$\beta \int_0^t e^{(\beta N - \gamma)u} du = \frac{\beta}{\beta N - \gamma} e^{(\beta N - \gamma)u} y \Big|_0^t \quad \text{if } \beta N - \gamma \neq 0 \quad (19)$$

$$e^{(\beta N - \gamma)t} y_t - y_0 = \frac{\beta}{\beta N - \gamma} (e^{(\beta N - \gamma)t} - 1) \quad \text{Evaluate expression at the upper and lower limits} \quad (20)$$

Notice exponential of zero gives 1. On the other hand, if the term in the exponent equals zero  $\beta N - \gamma = 0$ , then the exponential term will become:

$$\frac{dy_t}{dt} = \beta \quad (21)$$

$$y_t - y_0 = \beta t \quad \text{Taking integral from 0 to } t \quad (22)$$

Now we introduced this  $y$  to linearise the equation. Our object of interest is  $I$ , so lets reverse the transformation by recalling equation 9.

$$e^{(\beta N - \gamma)t} y_t - y_0 = \frac{\beta}{\beta N - \gamma} (e^{(\beta N - \gamma)t} - 1) \quad (23)$$

$$e^{(\beta N - \gamma)t} \frac{1}{I_t} - \frac{1}{I_0} = \frac{\beta}{\beta N - \gamma} (e^{(\beta N - \gamma)t} - 1) \quad \text{Reverse transformation and substitute } \frac{1}{y} \text{ for } y \quad (24)$$

$$I_t = \frac{e^{(\beta N - \gamma)t}}{\frac{1}{I_0} + \frac{\beta}{\beta N - \gamma} (e^{(\beta N - \gamma)t} - 1)} \quad \text{Rearrange to isolate } I_t \quad (25)$$

We can do the same for the solution where  $\beta N - \gamma = 0$ :

$$y_t - y_0 = \beta t \quad (26)$$

$$\frac{1}{I_t} - \frac{1}{I_0} = \beta t \quad \text{Reverse transformation and substitute } \frac{1}{y} \text{ for } y \quad (27)$$

$$I_t = \frac{1}{\frac{1}{I_0} + \beta t} \quad \text{Rearrange to isolate } I_t \quad (28)$$

So given the parameters,  $\beta$  and  $\gamma$ , and the population size  $N$ , we have an equation that tells us the number of infectives over time:

$$I_t = \frac{1}{\frac{1}{I_0} + \beta t} \quad \beta N - \gamma = 0 \quad (29)$$

$$I_t = \frac{e^{(\beta N - \gamma)t}}{\frac{1}{I_0} + \frac{\beta}{\beta N - \gamma} (e^{(\beta N - \gamma)t} - 1)} \quad \beta N - \gamma \neq 0 \quad (30)$$

(31)

By knowing  $I_t$  we know  $S_t$ . We can plot  $S_t$  and  $I_t$  over time and we will be able to tell the number of cases, whether the disease is going to evolve into an epidemic. One of the most useful comparisons with more general systems is to plot  $I_t$  versus  $S_t$ , where it will be a straight downward sloping line, but this diagram is more important when one moves to more complex models.

Now let's see what happens when  $t$  goes to infinity. The behaviour of equation 29 is trivial as it goes to zero. As for behaviour of equation 30, we shift the exponential term down so that it is easy to take the limit:

$$I_t = \frac{1}{\frac{1}{I_0} e^{-(\beta N - \gamma)t} + \frac{\beta}{\beta N - \gamma} (1 - e^{(\beta N - \gamma)t})} \quad (32)$$

We now have two branches and we also introduce a third and forth branch where the condition is less than zero and larger than zero:

- The third condition lets the limit go to zero similar to the first condition
- The fourth condition lets the exponential term go to zero where we can simplify the equation

$$I_t = \frac{1}{\frac{1}{I_0} + \beta t} \quad \beta N - \gamma = 0 \quad (33)$$

$$\lim_{t \rightarrow \infty} I_t = \frac{1}{\frac{1}{I_0} e^{-(\beta N - \gamma)t} + \frac{\beta}{\beta N - \gamma} (1 - e^{(\beta N - \gamma)t})} \quad \beta N - \gamma \neq 0 \quad (34)$$

$$\lim_{t \rightarrow \infty} I_t = 0 \quad \beta N - \gamma < 0 \quad (35)$$

$$\lim_{t \rightarrow \infty} I_t = \frac{1}{0 + \frac{\beta}{\beta N - \gamma} (1 - 0)} = \frac{\beta N - \gamma}{\beta} = N - \frac{\gamma}{\beta} \quad \beta N - \gamma > 0 \quad (36)$$

Now we can rearrange the inequality to get the memorable result:

$$\beta N - \gamma > 0 \quad N \frac{\beta}{\gamma} > 1 \quad (37)$$

So we can interoperate  $\mathcal{R}_0 = N \frac{\beta}{\gamma}$  as the threshold for whether the disease will become an epidemic. This ratio is called the *reproduction number* or *reproduction ratio* or the *R number* and it is a very important concept in the study of epidemics. This reproduction number has a very easy interpretation. We know  $N\beta$  really means the average number of people infected by one infective per unit of time. Remember the  $\gamma$  represents the rate of flow from the infective compartment, so you can see  $\frac{1}{\gamma}$  really represents the average time an infective spends in the infectious compartment.

So the reproduction number is the number of secondary infections produced by one primary infective. If one infectious produces more than one infective  $\mathcal{R}_0 > 1$ , then we have an epidemic, but if it is less than 1 infective  $\mathcal{R}_0 < 1$ , then we don't have an epidemic.

#### 4.5 SIR Framework Solution

Section 4.4 discusses and derived the solution for the SIS framework. This section discusses and derives the slightly complicated SIR framework. Now recall the SIR population equation 4 and its 3 equation (i.e. susceptibles, infected and recovered - equations 1, 2 and 3 respectively). We can now reduce this to two equations which is 1 more equation than the previous case.

Lets start with susceptible equation 1:

$$\frac{dS_t}{dt} = -\beta S_t I_t \quad (38)$$

$$\frac{dS_t}{S_t} = -\beta I_t dt \quad \text{Separate } S \text{ on the lhs} \quad (39)$$

$$d\ln(S_t) = -\beta I_t dt \quad d\ln S_t = d\ln S_t \quad (40)$$

$$\ln S_t - \ln S_0 = -\beta \int_0^t I_s ds \quad \text{Integrate from 0 to } t \quad (41)$$

Now we need to calculate the integral of  $I$ , which can be determined by considering the sum of equation 1 and equation 2:

$$\frac{dS_t}{dt} + \frac{dI_t}{dt} = -\gamma I_t \quad \text{Sum of } S \text{ and } I \quad (42)$$

$$dS_t + dI_t = -\gamma I_t dt \quad \text{In differential form} \quad (43)$$

$$S_t - S_0 + I_t - I_0 = -\gamma \int_0^t I_s ds \quad \text{Integrate from 0 to } t \quad (44)$$

We know the sum of  $S_0$  and  $I_0$  is the total population  $S_0 + I_0 = N$  so we can substitute N for the sum:

$$S_t + I_t - N = -\gamma \int_0^t I_s ds \quad (45)$$

We are interested in the integral of  $I$ :

$$S_t + I_t - N = -\gamma \int_0^t I_s ds \quad (46)$$

$$-\gamma \int_0^t I_s ds = \frac{S_t + I_t - N}{\gamma} \quad \text{Isolate to the rhs} \quad (47)$$

$$\ln S_t - \ln S_0 = \beta \frac{S_t + I_t - N}{\gamma} \quad \text{Substitute lhs into equation 41} \quad (48)$$

$$\frac{\gamma}{\beta} \ln S_t - S_t - I_t = \frac{\gamma}{\beta} \ln S_0 - N \quad \text{Rearrange to get constants on rhs and variables on lhs} \quad (49)$$

Now, equation 49 can be used to generate the scatter plot of  $S$  and  $I$  (i.e. plot the solution in the  $S - I$  plane,  $I$  on the horizontal plane and  $S$  on the vertical axis). Following from previous analysis, there are 2 things we would like to know.

1. How the infected population changes over time? and;
2. What is the fate of the different compartments?

Studying the behaviours of the infectives population is easy!

$$\frac{\gamma}{\beta} \ln S_t - S_t - I_t = \frac{\gamma}{\beta} \ln S_0 - N \quad \text{Recall equation 49} \quad (50)$$

$$I_t = \frac{\gamma}{\beta} \ln S_t - S_t - \frac{\gamma}{\beta} \ln S_0 + N \quad \text{Isolate } I_t \text{ to the lhs} \quad (51)$$

Looking at the differential equation 2:  $\frac{dI_t}{dt} = \beta S_t I_t - \gamma I_t$  and factoring  $I_t(\beta S_t - \gamma)$ . If the term in the brackets is greater than zero, then the number of infectives will increase, and if it is less than zero, then the number of infectives will decrease. Notice that  $\beta$  and  $\gamma$  are constant, but  $S_t$  changes over time as the individuals get infected. So we can deduce that the  $I$  will attain a maximum when the derivative is equal to 0. So if we substitute this for  $S_t = \frac{\gamma}{\beta}$ , we get the maximum number of infectives:

$$I_{max} = \frac{\gamma}{\beta} \ln \left( \frac{\gamma}{\beta} \right) - \frac{\gamma}{\beta} - \frac{\gamma}{\beta} \ln S_0 + N \quad (52)$$

$$= \frac{\gamma}{\beta} \left( \ln \frac{\gamma}{\beta} - 1 - \ln S_0 \right) + N \quad \text{Re-arrange} \quad (53)$$

So the key to the growth in infectives is again when  $\frac{\gamma}{\beta} S_t > 1$ , which is more or less the reproduction number we saw earlier  $\mathcal{R}_0 = N \frac{\beta}{\gamma}$ .

Now lets' deduce the fate of the disease by letting  $t$  go to infinity:

$$\lim_{t \rightarrow \infty} \frac{\gamma}{\beta} \ln S_t - S_t - I_t = \frac{\gamma}{\beta} \ln S_0 - N \quad \text{Recall equation 49 and take it to } \infty \quad (54)$$

$$\frac{\gamma}{\beta} \ln S_\infty - S_\infty - I_\infty = \frac{\gamma}{\beta} \ln S_0 - N \quad \text{Replace } t \text{ in the subscript with } \infty \quad (55)$$

Recall equation 42 which is the sum of the  $S$  and  $I$  differential equations  $\frac{dS_t}{dt} + \frac{dI_t}{dt}$ , we see the sum is a decreasing function, and since the sum of  $S$  and  $I$  is non-negative, because the population can't become negative. It has a

limiting value meaning its derivative will go to zero and we thus deduce that  $I_\infty$  goes to zero eventually. It might increase initially depending on the reproduction number, but it ultimately goes to zero. So we can set  $I_\infty = 0$  and shift the terms:

$$\lim_{t \rightarrow \infty} \frac{\gamma}{\beta} \ln S_\infty - \frac{\gamma}{\beta} S_0 - I_\infty = S_\infty - N \quad (56)$$

$$\ln \frac{S_0}{S_\infty} = N \frac{\gamma}{\beta} \left(1 - \frac{S_\infty}{N}\right) \quad \text{Re-arrange to get the final size equation} \quad (57)$$

Equation 57 is the *final size equation*. It gives the relationship between the initial population and the final population of the susceptibles,  $\frac{S_0}{S_\infty}$ , in terms of the two terms on the rhs:

$\mathcal{R}_0 = N \frac{\beta}{\gamma}$  **Reproduction Number**, represents the expected number of cases directly generated by one case in a population where all individuals are susceptible to infection.

$\left(1 - \frac{S_\infty}{N}\right)$  **Attack Ratio**, which represents the proportion of the population that will get infected at some point.

Now the rhs of the final size equation is a finite number, so we deduce that the lhs must also be finite, means that  $S_\infty \neq 0$  which means no matter what happens, there will be some susceptibles who will never get infected. This result again applies quite generally across models.

So we saw that the reproduction number is the key!

## References

- [1] Vincenzo Capasso (Oct. 2008). *Mathematical Structures of Epidemic Systems (Lecture Notes in Biomathematics)* Springer Berlin Heidelberg.
- [2] Emilia Vynnycky, Richard G White (May 2008). *An Introduction to Infectious Disease Modelling* Oxford University Press, USA; 1 edition.
- [3] Michael Y. Li (Jun 2019). *An Introduction to Mathematical Modeling of Infectious Diseases (Mathematics of Planet Earth)* Springer; Softcover reprint of the original 1st ed. 2018 edition (4 Jun. 2019)

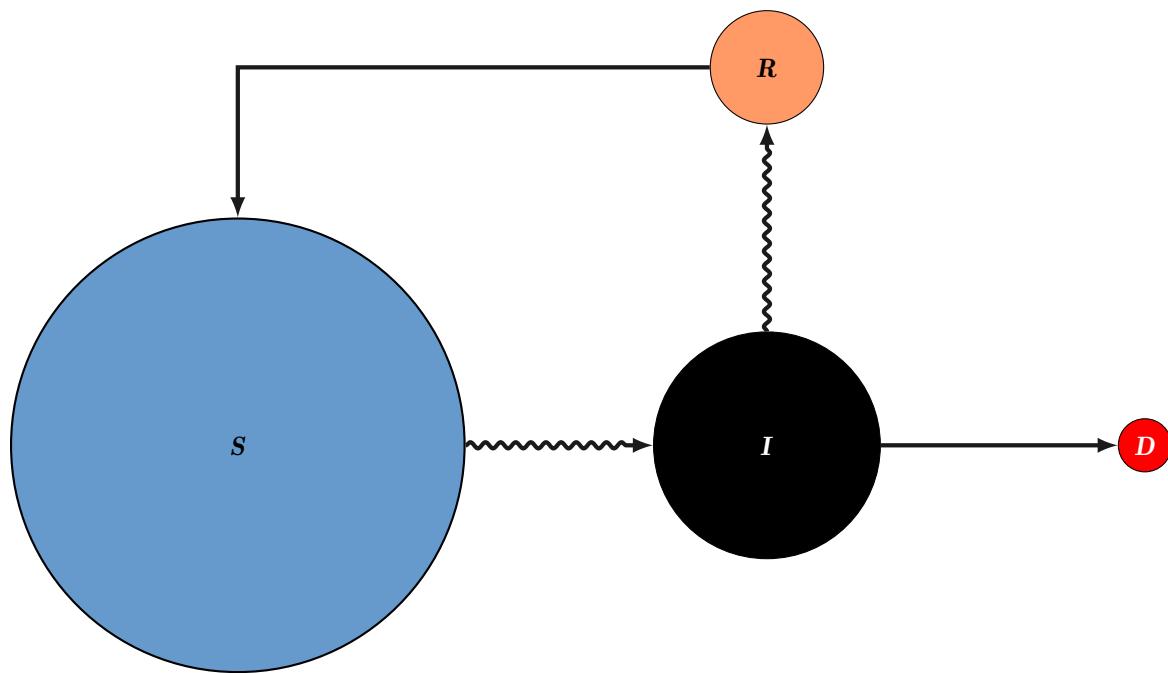


Figure 8: Infected start to be ‘Removed’ (i.e. either died or recovered)

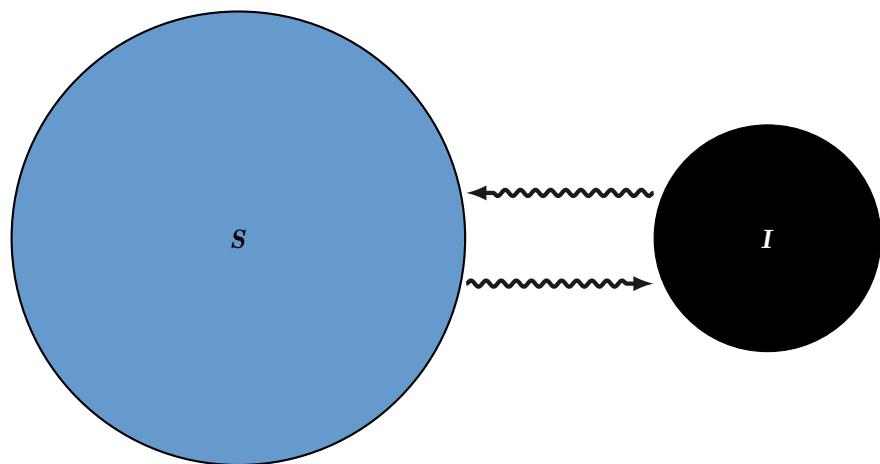


Figure 9: Infected start to be ‘Removed’ (i.e. either died or recovered)

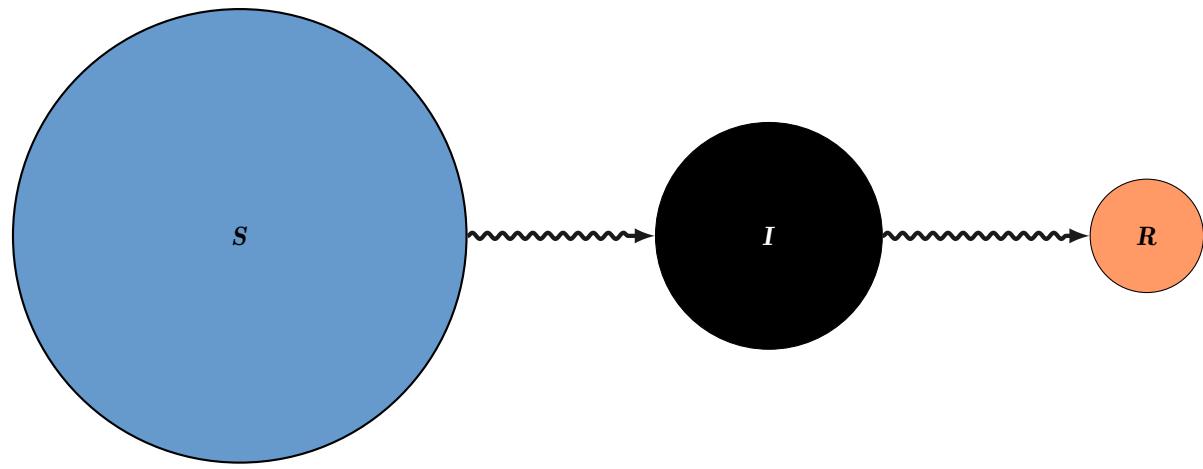


Figure 10: SIR Flow: Susceptibles → Infected → Removed

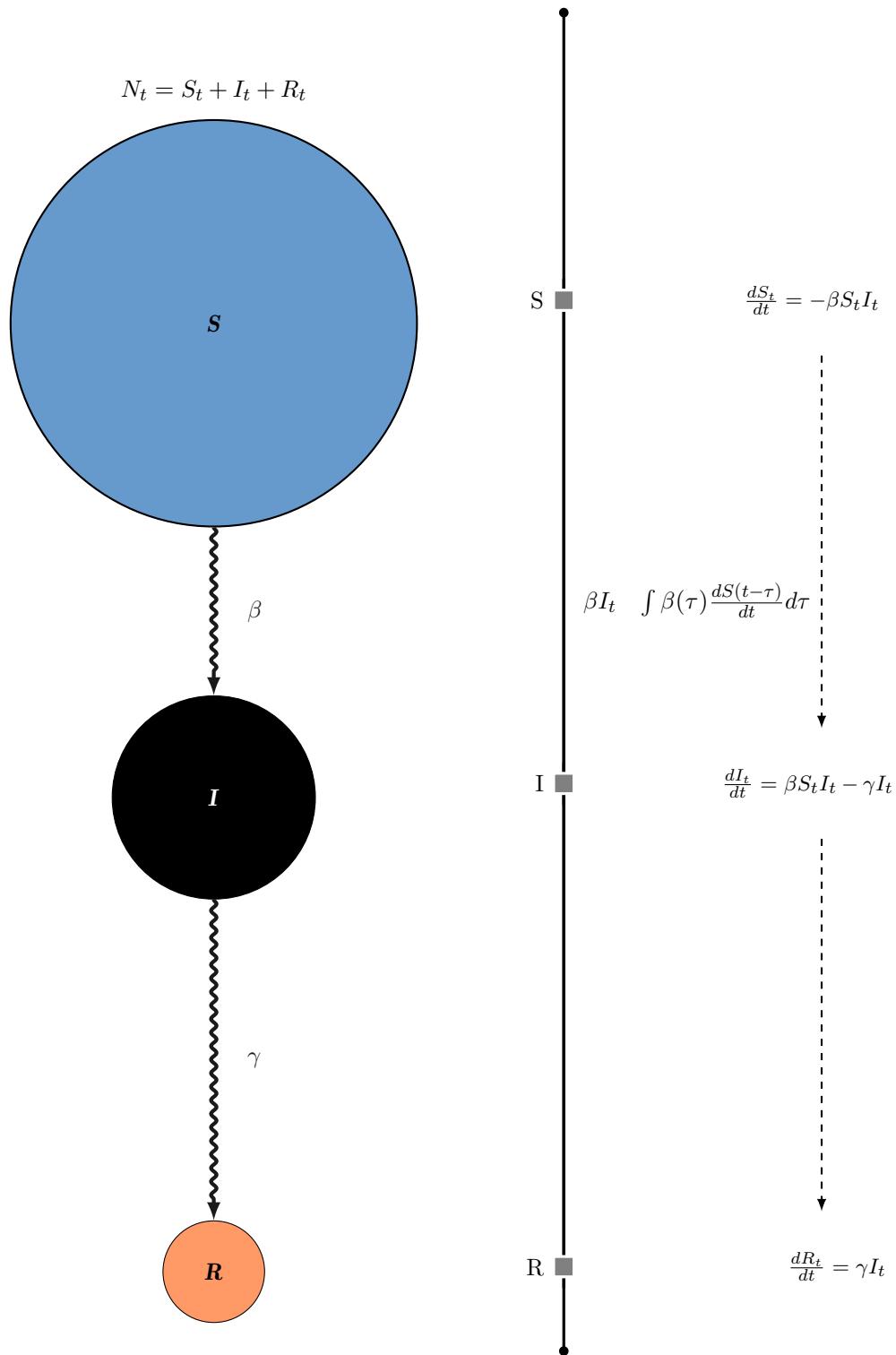


Figure 11: SIR Mathematical Flow: Susceptibles → Infected → Removed

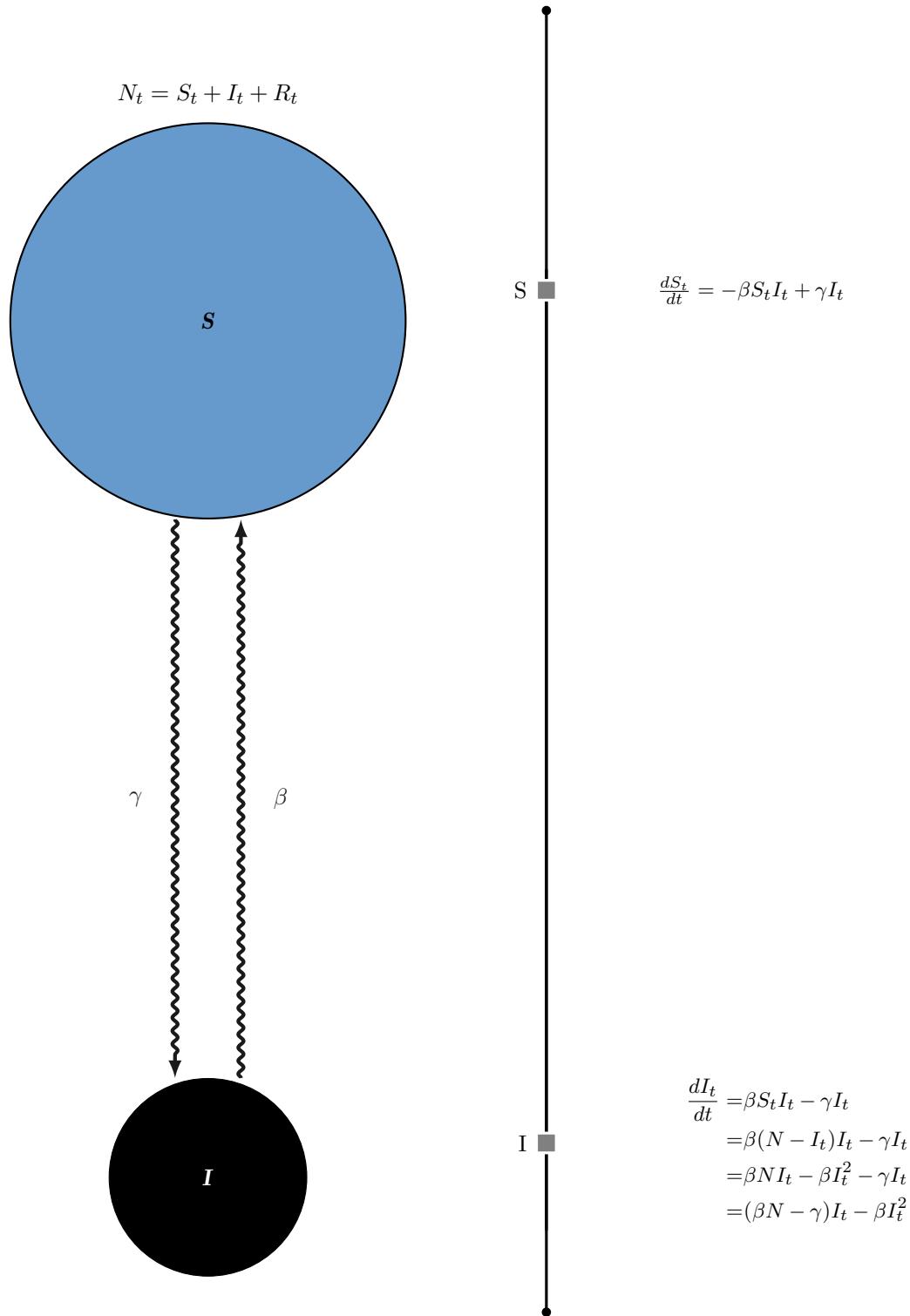


Figure 12: SIS Mathematical Flow: Susceptibles → Infected → Removed

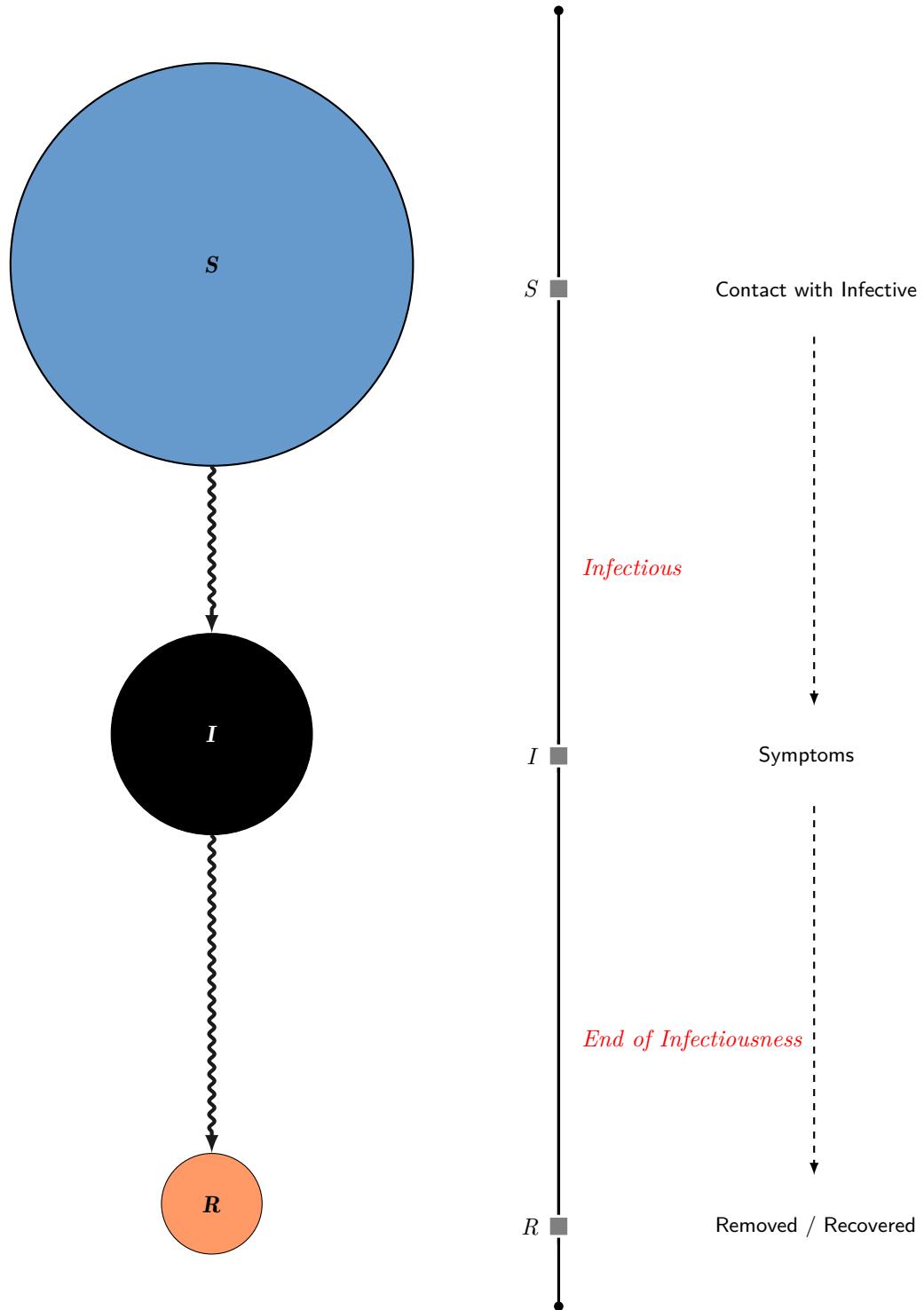


Figure 13: SIR Descriptive Flow: Susceptibles → Infected → Removed