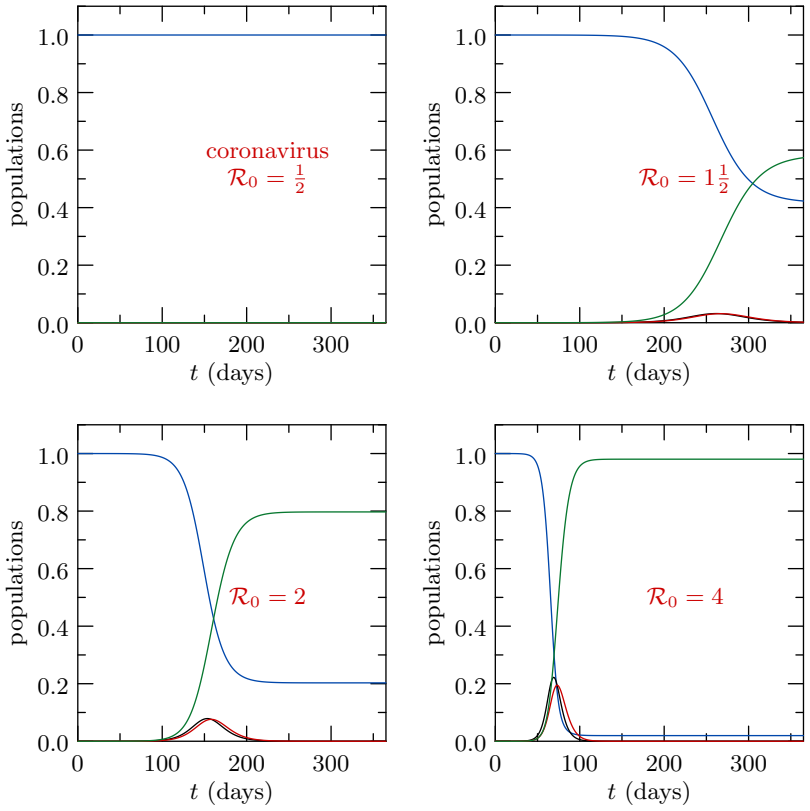
Mathematical Methods *for* Molecular Science

John E. Straub



Infectious Disease Supplement Complete Answers

Version 2.0 released June 2020.

This work is licensed under a Creative Commons "Attribution-NonCommercial-ShareAlike 4.0 International" license.

Solutions to exercises

E.1 The parameter b defining number of contacts per day between individuals can be reduced by practices such as *social distancing*. The parameter k defining the rate of recovery can be increased by effective medical care. The parameter q is related to the incubation period is largely a constant defined by the illness.

E.2 Increasing the number of contacts per day leads to an earlier onset of the disease and a larger peak number $i_{\max}$ of infected individuals. Reducing the number of contacts per day can prevent the onset of an epidemic by making the rate of recovery exceed the rate of infection or

$$\text{rate of infection} = b\,\mathrm{s}(t)\mathrm{i}(t) < k\,\mathrm{i}(t) = \text{rate of recovery}$$

This can also be written

$$b\frac{1}{k}\,\mathrm{s}(t) = \mathcal{R}_0\,\mathrm{s}(t) < 1$$

after which the disease will recede.

E.3 (a) Using the chain rule we find

$$\frac{d\mathrm{i}}{d\mathrm{s}} = \frac{d\mathrm{i}}{dt}\left(\frac{d\mathrm{s}}{dt}\right)^{-1} = (b\,\mathrm{s}(t)\mathrm{i}(t) - k\,\mathrm{i}(t))\left(\frac{-1}{b\,\mathrm{s}(t)\mathrm{i}(t)}\right) = -1 + \frac{k}{b\mathrm{s}(t)}$$

Separating the dependence on $\mathrm{i}(t)$ to the left and $\mathrm{s}(t)$ to the right we find

$$d\mathrm{i} = \left(-1 + \frac{k}{b\mathrm{s}(t)}\right) d\mathrm{s}$$

Integrating the left side of the equation leads to

$$\int d\mathrm{i} = \mathrm{i}(t) + C'$$

while integrating the right side of the equation leads to

$$\int \left(-1 + \frac{k}{b\mathrm{s}(t)}\right) d\mathrm{s} = -\mathrm{s}(t) + \frac{k}{b}\ln \mathrm{s}(t) + C''$$

Combining the two results we find

$$\mathrm{i}(t) + \mathrm{s}(t) - \frac{k}{b}\ln \mathrm{s}(t) = C$$

where $C = C'' - C'$ is a constant independent of time. To determine the constant C, we evaluate the expression at $t = 0$ and find

$$\mathrm{i}(0) + \mathrm{s}(0) - \frac{k}{b}\ln \mathrm{s}(0) = 1 = C$$

where we use the exact relation $\mathrm{i}(0) + \mathrm{s}(0) = 1$ and the approximation that $\mathrm{s}(0) = 1$. Evaluating the expression at $t = \infty$ we find

$$\mathrm{i}(\infty) + \mathrm{s}(\infty) - \frac{k}{b}\ln \mathrm{s}(\infty) = \mathrm{s}(\infty) - \frac{k}{b}\ln \mathrm{s}(\infty) = 1$$

where we use the fact that $\mathrm{i}(\infty) = 0$. Rearranging terms we find

$$\frac{b}{k} = \frac{\ln \mathrm{s}_\infty}{\mathrm{s}_\infty - 1} = \mathcal{R}_0$$

where $\mathrm{s}(\infty) = \mathrm{s}_\infty$.

(b) In the SIR model we found

$$\frac{d}{dt}\mathrm{i}(t) = b\,\mathrm{s}(t)\mathrm{i}(t) - k\,\mathrm{i}(t) = \left(\frac{b}{k}\,\mathrm{s}(t) - 1\right) k\mathrm{i}(t) = (\mathcal{R}_0\,\mathrm{s}(t) - 1)\,k\mathrm{i}(i)$$

When $\mathcal{R}_0\,\mathrm{s}(t) > 1$, the fraction of infected individuals increases. When $\mathcal{R}_0\,\mathrm{s}(t) < 1$, the fraction of infected individuals decreases. The critical value of $\mathrm{s}(t)$ below which herd immunity is achieved is

$$s_c = \frac{1}{\mathcal{R}_0}$$

For example, when $\mathcal{R}_0 = 3$ we find $s_c = \frac{1}{3}$ so that two-thirds of the population must be immune to achieve the state of *herd immunity*.

(c) We can write the uncertainty in the estimate of the reproduction number $\mathcal{R}_0$ as

$$\epsilon_R^2 = \left|\frac{d\mathcal{R}_0}{d\mathrm{s}_\infty}\right|^2 \epsilon_s^2$$

where ϵ_R^2 is the mean square error in $\mathcal{R}_0$ and ϵ_s^2 is the mean square error in s_∞. Noting the near linear dependence of $\mathcal{R}_0$ on $\log_{10}(\mathrm{s}_\infty)$ we can write

$$\mathcal{R}_0(\mathrm{s}_\infty) = \mathcal{R}_0(10^{-8}) - m\left(\log_{10}(\mathrm{s}_\infty) + 8\right)$$

where we have arbitrarily picked a constant reference value $\mathcal{R}_0(10^{-8}) = 18.4$ and introduced a constant m proportional to the slope. This leads to

$$\epsilon_R^2 = \left|\frac{d\mathcal{R}_0}{d\mathrm{s}_\infty}\right| \epsilon_s^2 = \left(\frac{m}{\ln(10)\mathrm{s}_\infty}\right)^2 \epsilon_s^2$$

Suppose we take the error in s_∞ to be proportional to s_∞. Then $\epsilon_s = \alpha \mathrm{s}_\infty$ and α is the relative error in s_∞. We find

$$\epsilon_R^2 = \left(\frac{m}{\ln(10)\mathrm{s}_\infty}\right)^2 \epsilon_s^2 = \left(\frac{m}{\ln(10)\mathrm{s}_\infty}\right)^2 (\alpha \mathrm{s}_\infty)^2 = \left(\frac{m\alpha}{\ln(10)}\right)^2 = \text{constant}$$

Remarkably, the uncertainty in $\mathcal{R}_0$ is largely independent of the size of $\mathcal{R}_0$.

E.4 The onset of the disease is enhanced by an increase in the initial number of infected individuals. However, onset is a relatively insensitive function of N_I. As the data shows, the time at which the peak number of infected individuals is observed is reduced by roughly a factor of 2 following an increase in N_I by three orders of magnitude.

E.5 At the peak of the infection the fraction of infected individuals reaches a maximum value $\mathrm{i}(t^*) = i_{max}$ at t^* where

$$\frac{d}{dt}\mathrm{i}(t^*) = b\,\mathrm{s}(t^*)\mathrm{i}(t^*) - k\,\mathrm{i}(t^*) = 0$$

so that

$$\frac{b}{k} = \mathcal{R}_0 = \frac{1}{\mathrm{s}(t^*)}$$

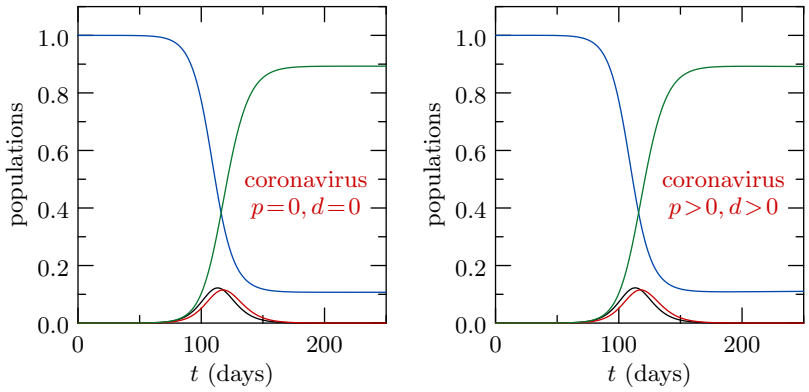
where $\mathrm{s}(t^*)$ is the fraction of susceptible individuals remaining uninfected at t^*. With this relation, we can estimate the reproduction number during the epidemic rather than after the fact. Using the data for coronavirus from Figure 1 we find that $t^* \simeq 45$ days where $\mathrm{s}(t^*) = 0.4$. Using the relation

$$\mathcal{R}_0 = \frac{1}{\mathrm{s}(t^*)} = \frac{1}{0.4} = 2.5$$

which equals the value of $\mathcal{R}_0 = 2.5$ used in our model.

E.6 The modified SEIR model that includes a birth rate p and death rate d takes the form

$$\begin{aligned}
\frac{d}{dt}\mathrm{s}(t) &= p - b\,\mathrm{s}(t)\mathrm{i}(t) - d\,\mathrm{s}(t) \\
\frac{d}{dt}\mathrm{e}(t) &= b\,\mathrm{s}(t)\mathrm{i}(t) - q\,\mathrm{e}(t) - d\,\mathrm{e}(t) \\
\frac{d}{dt}\mathrm{i}(t) &= q\,\mathrm{e}(t) - k\,\mathrm{i}(t) - d\,\mathrm{i}(t) \\
\frac{d}{dt}\mathrm{r}(t) &= k\,\mathrm{i}(t) - d\,\mathrm{r}(t)
\end{aligned}$$



where the new terms appear in red.

Let's explore the impact of the addition of birth and death rates on the predictions of the SEIR model using the parameters for coronavirus. We use 2019 parameters for birth rate $p = 3.18 \times 10^{-5}\text{day}^{-1}$ and natural death rate $d = 2.35 \times 10^{-5}\text{day}^{-1}$ in the United States. The results are shown below.
Inclusion of natural birth and death rates leads to no observable changes in the predictions of the model. The birth and death rates impact the overall population on the order of 1% per year. Therefore, changes due to the inclusion of birth and death are relatively insignificant on the time scale explored in the model.

E.7 (a) The product $N_T\,\mathrm{r}(t)$ is the cumulative number of individuals who have passed the stage of infection and $\mathcal{M}_0$ is the fraction of those individuals who die rather than recover. Therefore $\dagger(t) = \mathcal{M}_0\,N_T\,\mathrm{r}(t)$.

(b) The death rate as a function of time is given by

$$\frac{d\dagger(t)}{dt} = \mathcal{M}_0\,N_T\,\frac{d\mathrm{r}(t)}{dt}$$

Inserting the definition of $\dfrac{d\mathrm{r}(t)}{dt}$ from the SEIR model we find

$$\Delta(t) = \frac{d\dagger(t)}{dt} = \mathcal{M}_0\,N_T\,\frac{d\mathrm{r}(t)}{dt} = \mathcal{M}_0\,N_T\,k\,\mathrm{i}(t)$$

(c) Integrating over the death rate $\Delta(t)$ gives

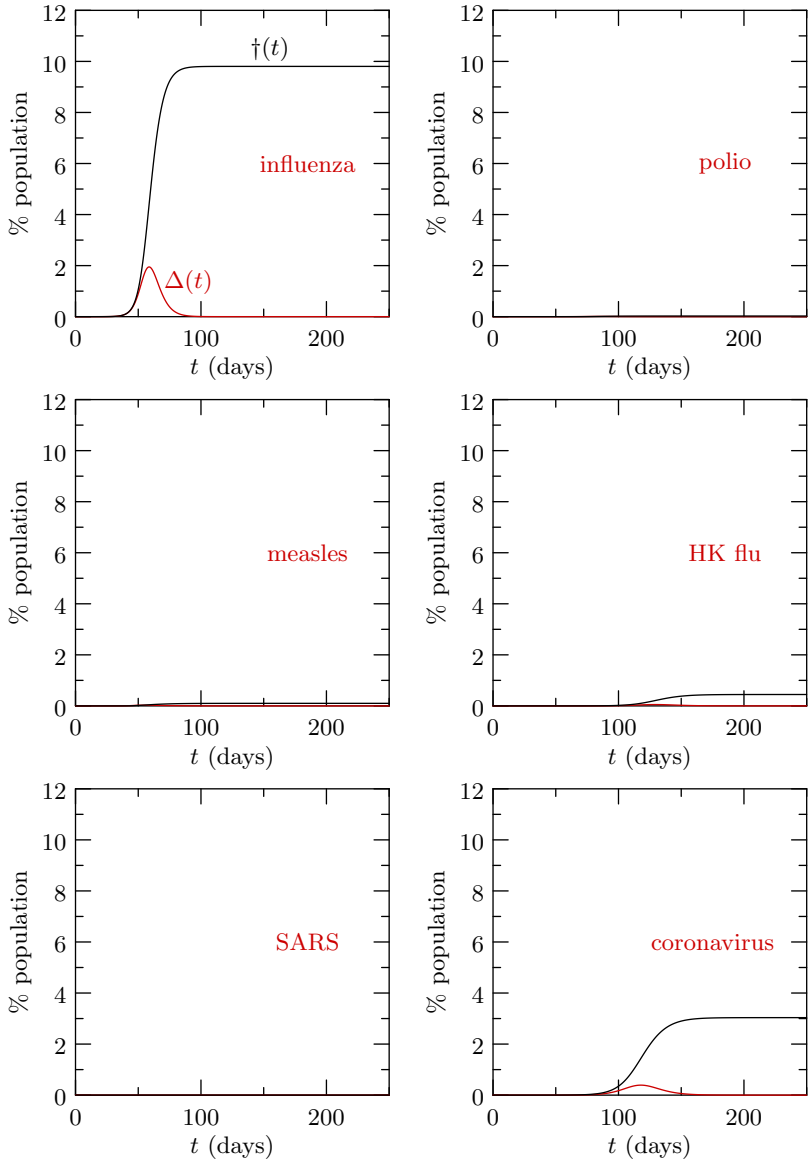
$$\dagger(t) = \int_0^t \Delta(\tau)\,d\tau = \int_0^t \mathcal{M}_0\,N_T\,\frac{dr}{d\tau}\,d\tau = \mathcal{M}_0\,N_T\,\mathrm{r}(t)$$

This result agrees with the original expression in (a).

The predictions of the simple extended SEIR model for the death rate $\Delta(t)$ and cumulative deaths $\dagger(t)$ were computed using the percent mortality values $\mathcal{M}_0$ of 10% for influenza (1918), 0.025% for polio (1952), 0.1% for measles (1960), 0.5% for Hong Kong flu (1968), 11% for SARS (2003), and 3.4% for coronavirus (2019). The dramatic variations in the predicted deaths per day and cumulative deaths as a function of time are shown below.

We can make a number of observations. (1) While the percent mortality of influenza (1918) and SARS (2003) are comparable, far fewer deaths were observed in the SARS epidemic due to early isolation of infected individuals. (2) The mortality resulting from measles and polio are low in spite of the wide spread of the disease. (3) While the profile of the HK flu and coronavirus were similar, the predicted mortality rate is far higher for coronavirus. Using parameters derived from February 2020 before the widespread implementation of social distancing measures in the United States, with an assumed population of 4,875,390 representing metropolitan Boston, the predicted total number of deaths from coronavirus is over 60,000.

The cumulative number of deaths from the six illnesses range from 50 million for the influenza outbreak of 1918, to 1 million for the Hong Kong flu of 1968, to several thousand due to measles in 1960, to on the order of thousands for polio in 1952, to on the order of 1 thousand for the SARS outbreak of 2003. The predictions of the simple model are consistent with those historic numbers.



E.8 Our data are well described by an exponential function

$$\dagger(t) = \dagger(0)e^{t/\tau}$$

Taking the base-10 logarithm of the function leads to

$$\log_{10}(\dagger(t)) = \log_{10}(\dagger(0)) + \left(\frac{1}{\ln(10)\tau}\right) t$$

where the slope m is

$$\tau = \frac{1}{m \ln(10)}$$

and $\ln(10) = 2.303$.

E.9 Power-law kinetics converges more slowly than exponential kinetics leading to a slower progression of the epidemic and an increased likelihood of long distance transmission of the infection.

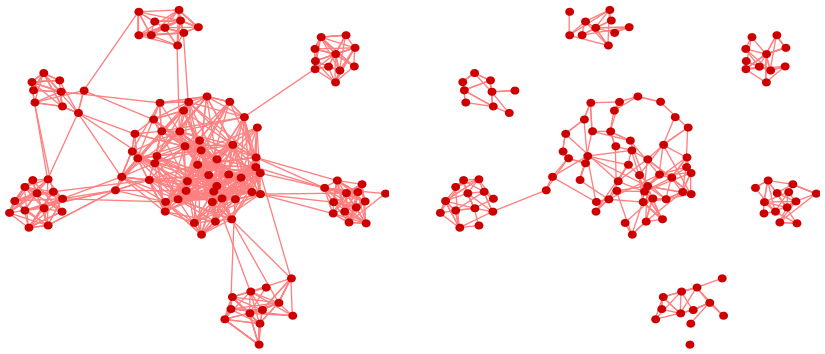


Figure 1: A social network of 114 individuals (red dots) forming 7 clusters. The distance between dots represents social distance.(Left) All clusters are connected and every individual is linked to every other individual. (Right) Clusters are disconnected and there is no longer connectivity between all individuals.

E.10 In the earlier realization of the small world network, every cluster was connected to every other cluster. The plots below show the same network of 114 individuals forming 7 clusters with lines representing $1,000$ encounters.
On the left the network remains above the percolation threshold. Every individual is connected to every other individual. On the right the network is below the percolation threshold. An infection starting in one cluster is isolated from infecting other clusters.

E.11 In the case of chemical kinetics, the reactant species are assumed to be dilute and uniformly distributed in solution. While it is not equally probably for one reactant to encounter another reactant in solution, the probability of reactant encounters is uniform throughout the solution. In such cases, there is one average time scale for reaction and homogeneous exponential kinetics is observed. In contrast, in the small world network model individuals are non-uniformly distributed in space. In such cases, there is a distribution of reaction times and heterogeneous power-law kinetics is observed.

E.12 The *mad cow disease* is a protein-only disease. It evolves by having a normal form of prion protein, $\mathrm{PrP^C}$, contact a pathogenic scrapie form of prion protein, $\mathrm{PrP^{Sc}}$, and undergo a conformational transition to the scrapie form. The rate of contact is taken to be b. Subsequently, the pathogenic scrapie form, $\mathrm{PrP^{Sc}}$, aggregates to form protein fibril $\mathrm{PrP^F}$. The rate of aggregation is taken to be k. As such, the kinetics of prion protein interactions are isomorphic with the kinetics of the SIR model

$$\begin{aligned}
\frac{d}{dt}[\mathrm{PrP^C}](t) &= -b\,[\mathrm{PrP^C}](t)[\mathrm{PrP^{Sc}}](t) \\
\frac{d}{dt}[\mathrm{PrP^{Sc}}](t) &= b\,[\mathrm{PrP^C}](t)[\mathrm{PrP^{Sc}}](t) - k\,[\mathrm{PrP^{Sc}}](t) \\
\frac{d}{dt}[\mathrm{PrP^F}](t) &= k\,[\mathrm{PrP^{Sc}}](t)
\end{aligned}$$

where $k = [\mathrm{PrP^F}]k_F$ is a rate constant proportional to the concentration of fibril $[\mathrm{PrP^F}]$ that is assumed to be high. The elementary rate constant for adding a scrapie form of the protein to a preexisting fibril is k_F.

E.13 The models we have considered assume rate constants that are independent of time. Implementing measures to enhance social distancing, medical intervention in treating patient symptoms to facilitate recovery, and isolation of infected individuals from the susceptible population impact the evolution of an epidemic and the parameters of our models. The formulation of kinetic models with time-dependent rate constants would add detail that could be useful in both prediction and post-epidemic analysis of the progression of disease.