0.0.1   SIRS Model

We now introduce another model, in which we take into account that during the years the **immune system may lose the ability to recognize a known pathogen**. This immunity could have been acquired via either a vaccine, or having recovered from that disease itself. Moreover, there could be the possibility that viruses mutate, as it occurs with the seasonal influenza, and so antibodies are not able to recognize it any more. Hence, let us build a model in which after an individual is recovered, can become again susceptible after a certain period of time.

The SIRS Model allows to interpolate between SIR ($w = 0$) and SIS ($w \to \infty$), where w is the **waning immunity rate**, namely the rate at which we lose our ability to defend ourselves from a certain pathogen. We can end up again into either an absorbing, with no more disease, or endemic state, where it keeps on circulating. The transitions for this model are:

$$\begin{aligned} S + I &\xrightarrow{\beta} I + I \\ I &\xrightarrow{\mu} R \\ R &\xrightarrow{w} S \end{aligned} \tag{1}$$

In particular, the differential equations that describe the model are:

$$\begin{aligned} \frac{\mathrm{d}s}{\mathrm{d}t} &= \alpha + wr - \beta si - \alpha s \\ \frac{\mathrm{d}i}{\mathrm{d}t} &= \beta si - \mu i - \alpha i \\ \frac{\mathrm{d}r}{\mathrm{d}t} &= \mu i - wr - \alpha r \end{aligned} \tag{2}$$

In this case, the **endemic state** can be found by setting the derivatives equal to zero.

One may note that the transition $R \to S$ does not affect the I, so it holds that for the **infectious period**:

$$\tau = \frac{1}{\alpha + \mu} \tag{3}$$

while the **R$_0$** factor is:

$$R_0 = \frac{\beta}{\alpha + \mu} \tag{4}$$

In addition, the equilibrium values s^*, i^* and r^* can be easily obtained using the same arguments as of the SIR model with demography.

0.0.2   SEIR Model

In reality people do not become instantaneously infectious, but there is a **latent period** which is the time between infection and becoming infectious. Indeed, the pathogen replication takes time, i.e. viral load is too low to be able to transmit the infection. This argument leads us to introduce the **S**usceptible, **E**xposed, **I**nfected, **R**ecovered model, where the class **E** takes into account that a person has already contracted the disease, hence is not susceptible anymore, but is not able to spread it yet.

Moreover, this period can be extremely heterogeneous depending on the disease: it can take from few hours to years, such as the case for HIV or, even longer, TBC. In the latter, latent periods might appear to be even longer than an individual's

lifespan, with the result that he may have contracted the disease, but the death occurs for other causes before the onset of any symptom.

It is important to remind that the **latent period** is **not the same** of the **incubation period** (see Fig. 1). An individual can be infectiuous before symptoms. For instance, there might be a **pre-syntomatic infection period** as it occurs in the case of COVID-19! This explains, once again, why medical status is different from the infection status.
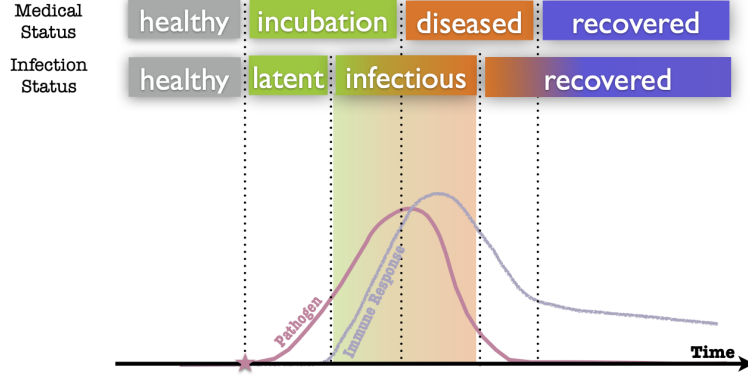

Figure 1: Difference between infection status and medical status.

The transition for the **SEIR** model are:

$$S + I \xrightarrow{\beta} I + E$$
$$E \xrightarrow{\sigma} I \tag{5}$$
$$I \xrightarrow{\mu} R$$

with the equations:

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \alpha - \beta si - \alpha s$$
$$\frac{\mathrm{d}i}{\mathrm{d}t} = \beta si - (\alpha + \sigma)e$$
$$\frac{\mathrm{d}i}{\mathrm{d}t} = \sigma e - (\alpha + \mu)i \tag{6}$$
$$\frac{\mathrm{d}r}{\mathrm{d}t} = \mu i - \alpha r$$

Hence, the spreading is delayed due to the time spent in E class.

The **endemic state** is:

$$s^* = \frac{(\alpha + \mu)(\alpha + \sigma)}{\beta \sigma} = \frac{1}{R_0}$$
$$e^* = \frac{\alpha(\alpha + \mu)}{\beta \sigma}(R_0 - 1) \tag{7}$$
$$i^* = \frac{\alpha}{\beta}(R_0 - 1)$$

For very short latent time ($\sigma \to \infty$) we recover the endemic state of the SIR.

The **R_0** factor is:

$$R_0 = \frac{\beta \sigma}{(\alpha + \mu)(\alpha + \sigma)} \tag{8}$$

Since latent time is way shorter than demography one, usually $\frac{\sigma}{\sigma + \alpha} \simeq 1$, hence $R_0 = \frac{\beta}{\alpha + \mu}$ as in the SIR with demography.

One may object that, given that the infectious period and R_0 are similar between SEIR and SIR, adding the Exposed class may seem an unnecessary complication. However, if we look at the time evolution, at the **early stages** there is a huge difference between SEIR and SIR model:

$$i_{SEIR}(t) \approx e^{\left(\sqrt{4(R_0-1)\sigma\mu+(\sigma+\mu)^2}-(\sigma+\mu)\right)t/2} \approx i_0 e^{\left(\sqrt{R_0}-1\right)\mu t}$$
$$i_{SIR}(t) \approx i_0 e^{(R_0-1)\mu t} \tag{9}$$

Even if the behavior at the steady state is similar, the temporal evolution of the prevalence of SEIR model is actually slower than the one using SIR. This has surely to be taken into account in policy making, given its important implications.

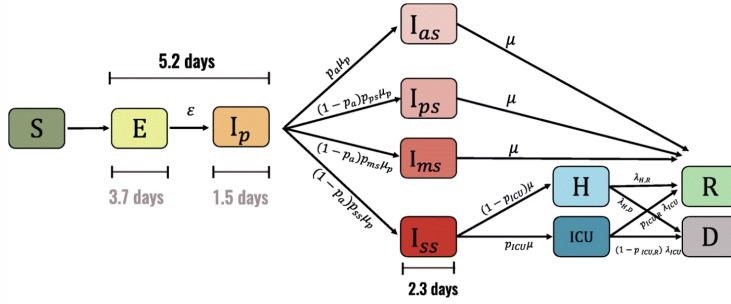The SEIR can be the starting point for modeling realistic diseases: i.e. Covid-19 (see Fig. 2).

Figure 2: Model for Covid-19.

0.1 Summary of compartmental models in well-mixed populations

Let us summarize all the compartmental models in well-mixed populations we have tackled so far:

- we solved the **SI model** analytically, and observed that the growth is the one of a sigmoid:

$$i(t) = \frac{i_0 e^{\beta t}}{1 - i_0 + i_0 e^{\beta t}}$$

In the early stages we observe an exponential growth, governed by β, that always saturates at 1;

- in the **SIS model** things starts to change. We have and **endemic** (meta-stable) **state**:

$$i = \frac{\beta - \mu}{\beta}$$

and we reach a sort of **dynamical equilibrium**. We can define an **epidemic threshold**:

$$\beta > \beta_c = \frac{\mu}{\langle k \rangle}$$

- for the **SIR model** equations cannot be solved analytically. However, we observe **no endemic state** and the **epidemic threshold** is once again:

$$\beta > \beta_c = \frac{\mu}{\langle k \rangle}$$

- then, in the **SIRS model** we introduced **waning immunity**. This model interpolates between SIR and SIS model. We do observe **endemic state** and the **infectious period** is:
$$\tau = \frac{1}{\alpha + \mu}$$

  and moreover:
$$R_0 = \frac{\beta}{\alpha + \mu}$$

- finally, we discussed **SEIR model** in which we included a **latent period**. We have that:
$$R_0 = \frac{\beta\sigma}{(\alpha + \mu)(\alpha + \sigma)}$$

  and the Exposed class has the effect to slow down the spreading.

1

Network Science - Basics

1.1 Main definitions

When we talk about Network Science, as the name would suggest, we study **Networks** that, in math, are also known as graph. A **Graph** $G(V, E)$ is simply an object that is composed by a set of **nodes** (vertices) V and a set of **links** (edges) E:

- **nodes** represent the *entities* $V = [\ldots, i, j, k, \ldots]$ involved in some relationship. These might be entries, people belonging to a social network and so forth. The **number of nodes** is $N = |V|$;

- **links** represent the relationships between entities $E = [\ldots, (i, j), (i, k), \ldots]$. The **number of links** is $L = |E|$.

Links can be of different kinds and so networks: the basic distinction is between **undirected** and **directed** links. The former ones can be thought as directed edges, but with arrows pointing in both directions, i.e. to both node of the pair. While the second ones do have a direction according to which sense the relationship represented by the link holds.

Another important distinction is between **unweighted** and **weighted** links. The latter ones can be exploited to take into account the possibility that some nodes can be more connected than the others, therefore **weights** follow. In a certain sense, it describes the "strength" of the link between two nodes.

Another important quantity is the **network density** (connectance), that is the fraction of links present normalized to all the possible pairs:

$$d = \frac{L}{N(N-1)} \tag{1.1}$$

Real networks usually have a very low density, so are **sparse systems** ($L \ll N^2$).

A graph, mathematically, can be represented by the mean of a matrix. It is the so called **adjacency matrix** A of the network, where:

- $a_{ij} = 1$, if a link between nodes i and j exists;

- $a_{ij} = 0$ otherwise.

Many mathematical tools can be used to determine the properties of the system alongside with this matrix, as an example we may want to compute its spectrum in order to obtain the largest eigenvalue. Moreover, one should note that the matrix is symmetrical for undirected and unweighted graphs, i.e. $a_{ij} = a_{ji}$. However, as we already told, real networks are usually sparse, therefore the adjacency matrix will be

filled for large part by zeros. Hence in order to store graphs in a computer efficiently, it is better to use other tools such as adjacency lists, etc.

Two nodes that share a link are defined "connected", "adjacent", "neighbors". In particular, the **neighborhood** of node i is the set of nodes connected to i. The number of neighbors k_i of each node i is what is called the **degree** of the node i. This is the basic measure that we are going to encounter so many times. Once we have defined the degree, the next step is to define what is the **average degree** over the entire network:

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i, \qquad \text{or} \qquad \langle k \rangle = \frac{2L}{N} = d(N-1) \qquad (1.2)$$

The next definition is the one of **path**, which is a sequence of links which permits to go from node i to node j following edges. Another relevant quantity is the so called **shortest path** between i and j, it is important since it gives us the idea of how big the network is. In particular, the **distance** l_{ij} represents the length of the shortest path between i and j. There could be multiple shortest paths between i and j. The shortest path of maximum length in the network is defined as **diameter**:

$$l_{max} = \max_{ij} l_{ij}$$

Another measure we may want to introduce is the **average** (shortest) **path length**:

$$\langle I \rangle = \frac{\sum_{ij} l_{ij}}{N(N-1)}$$

The network is said to be **connected** if every possible couple of nodes is reachable trough a path. Otherwise, each connected part is defined as a **connected component**.

Now, let us see some examples of networks, such as "The Oracle of Bacon", or the so called "Erdos Number". The first one is a site that, given the name of an actor, returns the distance between this actor and Kevin Bacon, in unit of costarring movies. This quantity is indeed computed by taking into account the network of actors, linked by common movies in which they starred. The **Erdos Number** instead is the "academical version" for the "Oracle of Bacon": we compute the distance, in terms of collaborations in publications, between a given researcher and the mathematician Paul Erdos through the publications network. The most surprising fact is that, for both examples, the distance is very low! Therefore a question arises: why such short distances in such large networks? In particular, real networks are smaller (i.e. shorter) than one would expect. This is pointed out by the idea of the "Six degrees of separation". It refers to an experiment that was run in the '60s by Stanley Milgram: he gave a postcard to a person on the West Coast, with the instructions that it had to be delivered to a place situated in the East Coast. The main goal was to count how many people would receive that postcard, given the rule that it was allowed to give the postcard to acquaintances of the actual possessor. It was discovered that this postcard actually was delivered to 6 people before reaching the destination. This is what is called the **small world phenomena**. When we study the average path length, for some networks we may find that $\langle I \rangle \sim ln(N)$ or, in some cases even $\langle I \rangle \sim ln(ln(N))$. This is extremely important in the spreading of diseases, since we are able to cover the whole system in few steps.