MATH 8xxx – Lecture 01

Epidemiology & A brief history of Mathematical Epidemiology

Julien Arino

September 2023

Among the first events recorded

- Epidemic events are among the first events recorded in History

- Indeed their effect was devastating at a time when medicine virtually did not exist and thus easily observed

First known epidemics (from Wikipedia)

Epidemics with major human cost

“Forgotten” killers

- Tuberculosis (TB). In 2020, estimated 10 M cases of active TB, leading to 1.5 M deaths

- Malaria: 229 M cases and 409 000 deaths in 2019

Neglected tropical diseases (NTD)

Often endemic diseases, sometimes major causes of death, but out of sight of rich countries. De Wikipedia, sachant que la liste précise varie selon les auteurs :

The Black Death

The British Plague of 1547

The Plague of Marseille of 1720

Definition

Wiki **Epidemiology** is the study and analysis of the distribution (who, when, and where), patterns and determinants of health and disease conditions in defined populations

BMJ **Epidemiology** is the study of how often diseases occur in different groups of people and why. Epidemiological information is used to plan and evaluate strategies to prevent illness and as a guide to the management of patients in whom disease has already developed

Etymology: *the study of what is upon the people*, derived from the Greek *epi* (*upon*, *among*), *demos* (*people*, *district*) and *logos* (*study*, *word*, *discourse*)

Who, when and where

Recall part of the definition on Wikipedia

> ***Epidemiology*** *is the study and analysis of the distribution (who, when, and where)*

A terminologically heavy domain

A few pointers:

- Moghadas and Laskowski. Review of terms used in modelling influenza infection. NCCID 2014

- Milwid et al. Toward standardizing a lexicon of infectious disease modeling terms. Frontiers in Public Health 2016

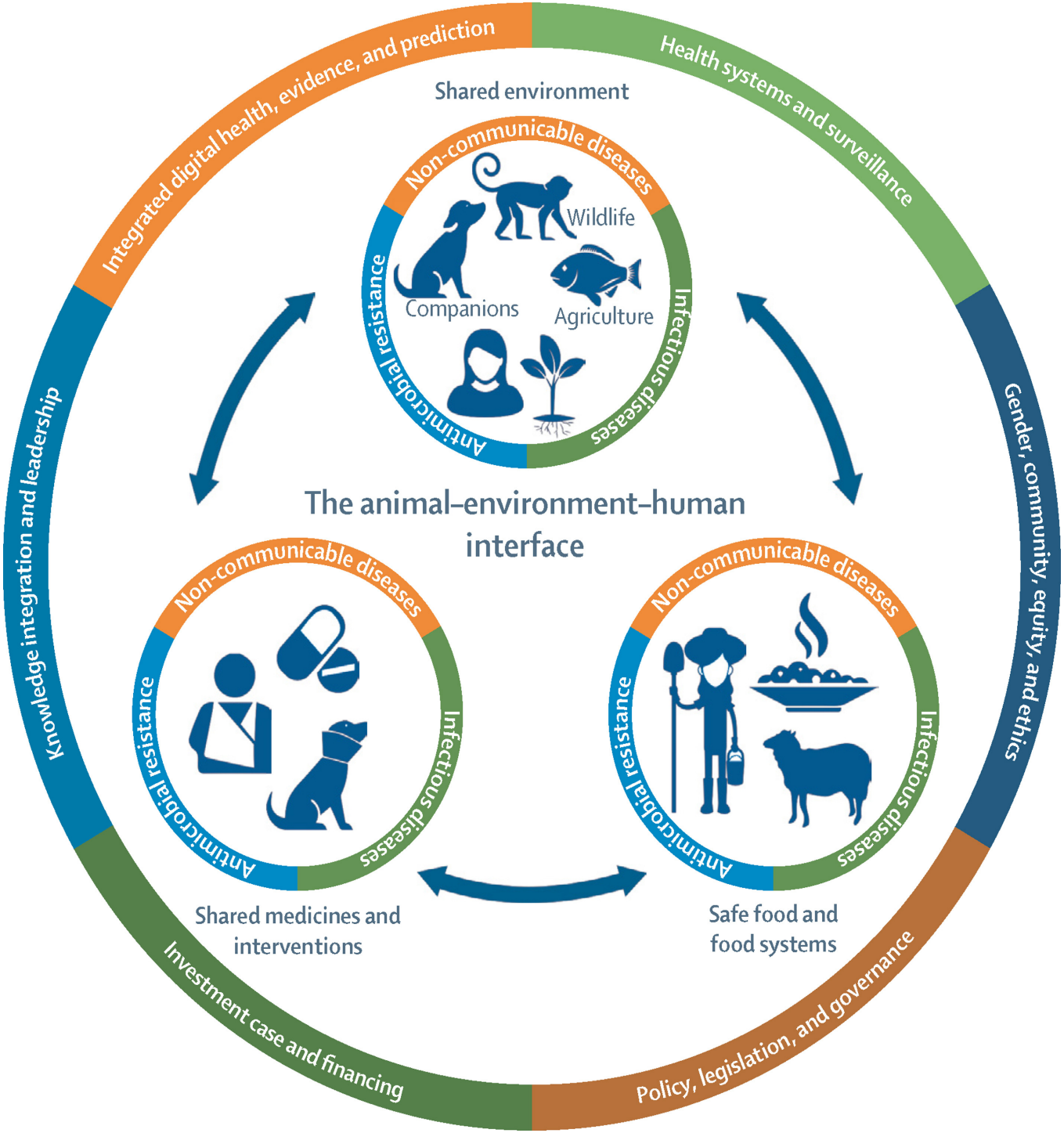
Who

- **Epidemiology** typically used when dealing with humans, but sometimes also generically when an easy description is sought; e.g., plant disease epidemiology

- **Epizootic**: denoting or relating to a disease that is temporarily prevalent and widespread in an animal population

- **Panzootic** is like a pandemic for animals

- **One Health**: considers health of humans, animals and their environment (including plants)


Integrated digital health, evidence, and prediction
Health systems and surveillance
Gender, community, equity, and ethics
Policy, legislation, and governance
Investment case and financing
Knowledge integration and leadership
Shared environment
Non-communicable diseases
Infectious diseases
Antimicrobial resistance
Wildlife
Companions
Agriculture
The animal–environment–human interface
Shared medicines and interventions
Safe food and food systems

Incidence & Prevalence (when?)

Incidence: number of new cases in a population generated within a certain time period

Prevalence: number of cases of a disease at a single time point in a population

$\Longrightarrow$ $I(t)$ in an epidemiological model is **prevalence**, not **incidence**

Exposition versus Exposed

- Some bright bulb (not sure who) in days of yore: let's call **exposed** someone who has contracted the disease but is not yet showing symptoms ($\Longrightarrow$ SEIR model)

- "Real" epidemiologist: let's trace people who were exposed to the virus, i.e., people having come into contact with the virus (whether they have contracted the disease or not)

- Interestingly, I have embarked on a quixotic quest to make people use L instead of E, only to be told by real epidemiologists that they don't care :)

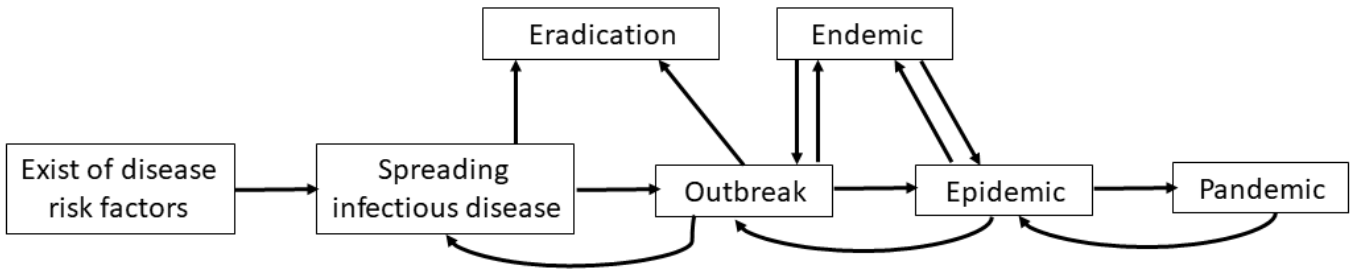
The different stages of propagation

Outbreak
sudden increase in occurrences of a disease in a particular time and place.

Endemic
constant maintained increase in occurrences of a disease in a geographic area.

Epidemic
rapid spread of disease to a large number of people in a given population within a short period of time.

Pandemic
spread across a large region, for instance multiple continents, or worldwide.

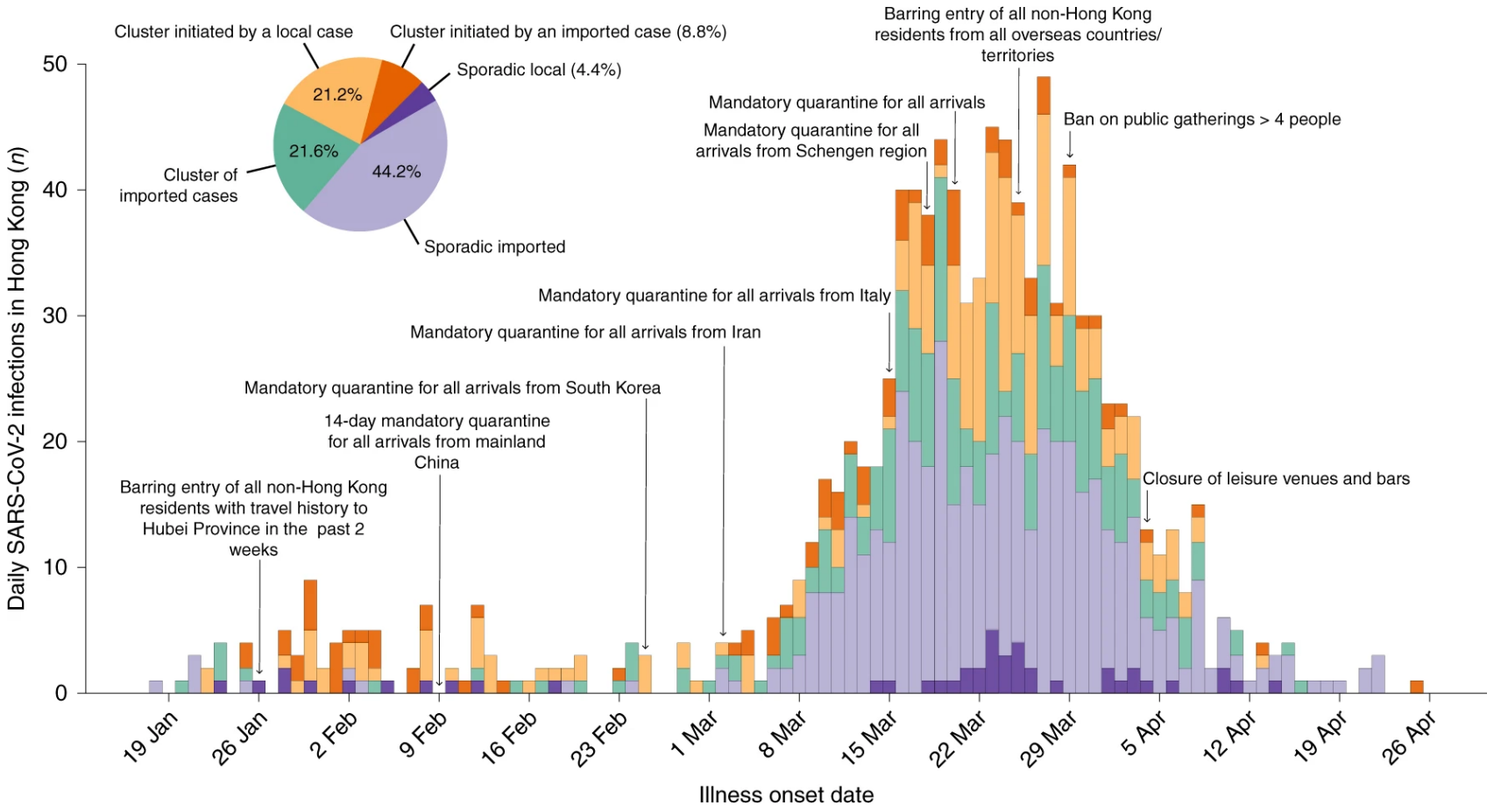


Epidemic curves

- Used to record the occurrence of new cases as a function of time

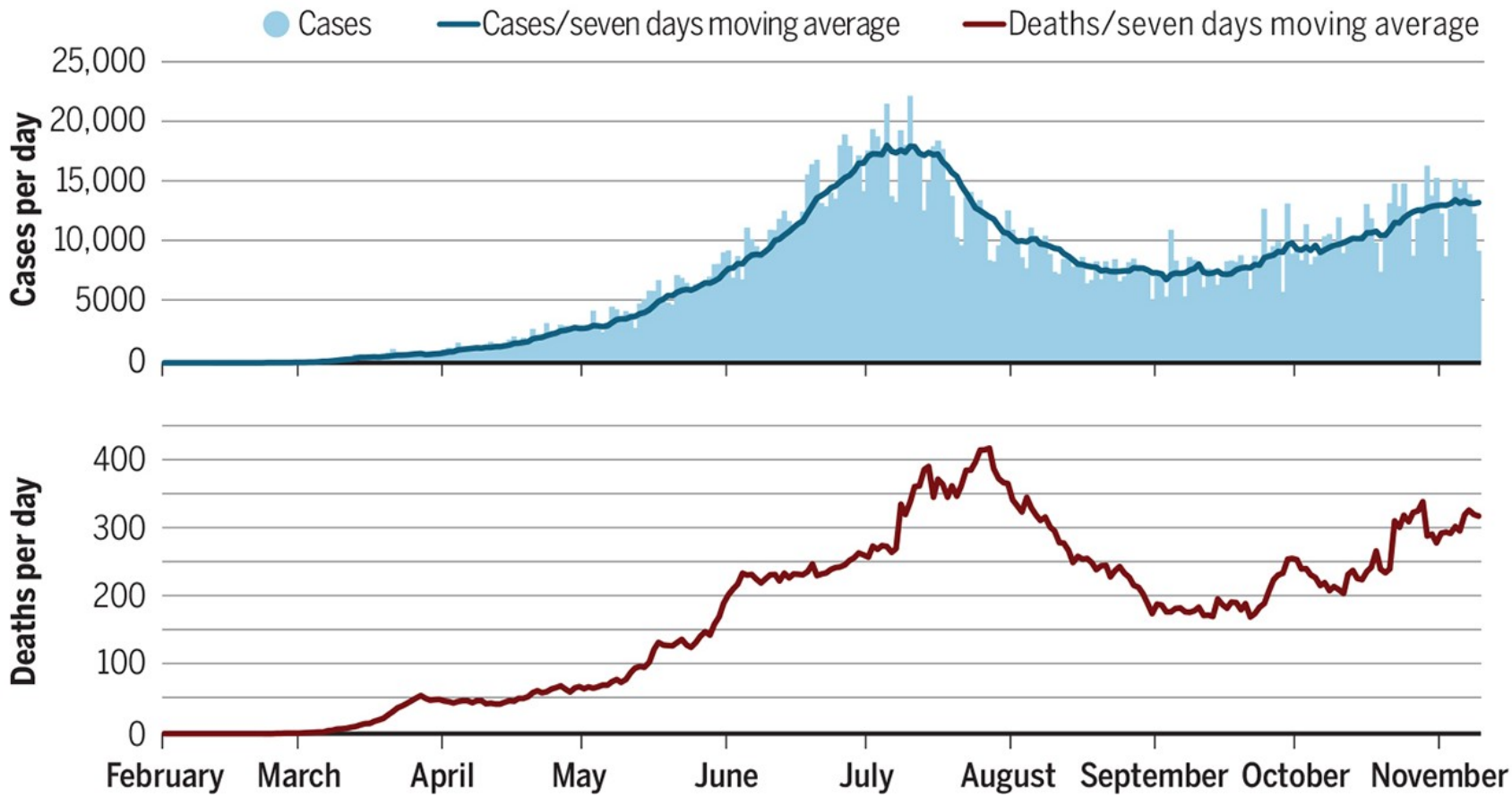
- When not too many cases, usually "individualised" (bar plots)

- When number of cases is large, continuous curve


Cluster initiated by a local case
Cluster initiated by an imported case (8.8%)
Sporadic local (4.4%)
21.2%
21.6%
44.2%
Cluster of imported cases
Sporadic imported
Barring entry of all non-Hong Kong residents from all overseas countries/ territories
Mandatory quarantine for all arrivals
Mandatory quarantine for all arrivals from Schengen region
Ban on public gatherings > 4 people
Mandatory quarantine for all arrivals from Italy
Mandatory quarantine for all arrivals from Iran
Mandatory quarantine for all arrivals from South Korea
14-day mandatory quarantine for all arrivals from mainland China
Barring entry of all non-Hong Kong residents with travel history to Hubei Province in the past 2 weeks
Closure of leisure venues and bars
Daily SARS-CoV-2 infections in Hong Kong (n)
0
10
20
30
40
50
19 Jan
26 Jan
2 Feb
9 Feb
16 Feb
23 Feb
1 Mar
8 Mar
15 Mar
22 Mar
29 Mar
5 Apr
12 Apr
19 Apr
26 Apr
Illness onset date

COVID-19 cases and deaths in Africa

The trend of daily reported cases of COVID-19 for the African continent, February to November 2020, shows the first peak of cases occurred July to August (mostly attributed to the Southern African Region) followed by a second peak, which started in October (mostly attributed to the Northern Region).



Some terminology for "where"

- **Epidemic**: diseases that are *visited upon* a population

- **Pandemic**: (will revisit this later in the course) epidemic that has spread across a large region, e.g., multiple continents or worldwide

- **Endemic**: diseases that *reside within* a population

- We don't say "panendemic"

Where? 1854 cholera outbreak

Cholera outbreak near Broad Street, London (UK)
Studied by John Snow

> *I found that nearly all the deaths had taken place within a short distance of the [Broad Street] pump*

WHO pandemic (influenza) phases

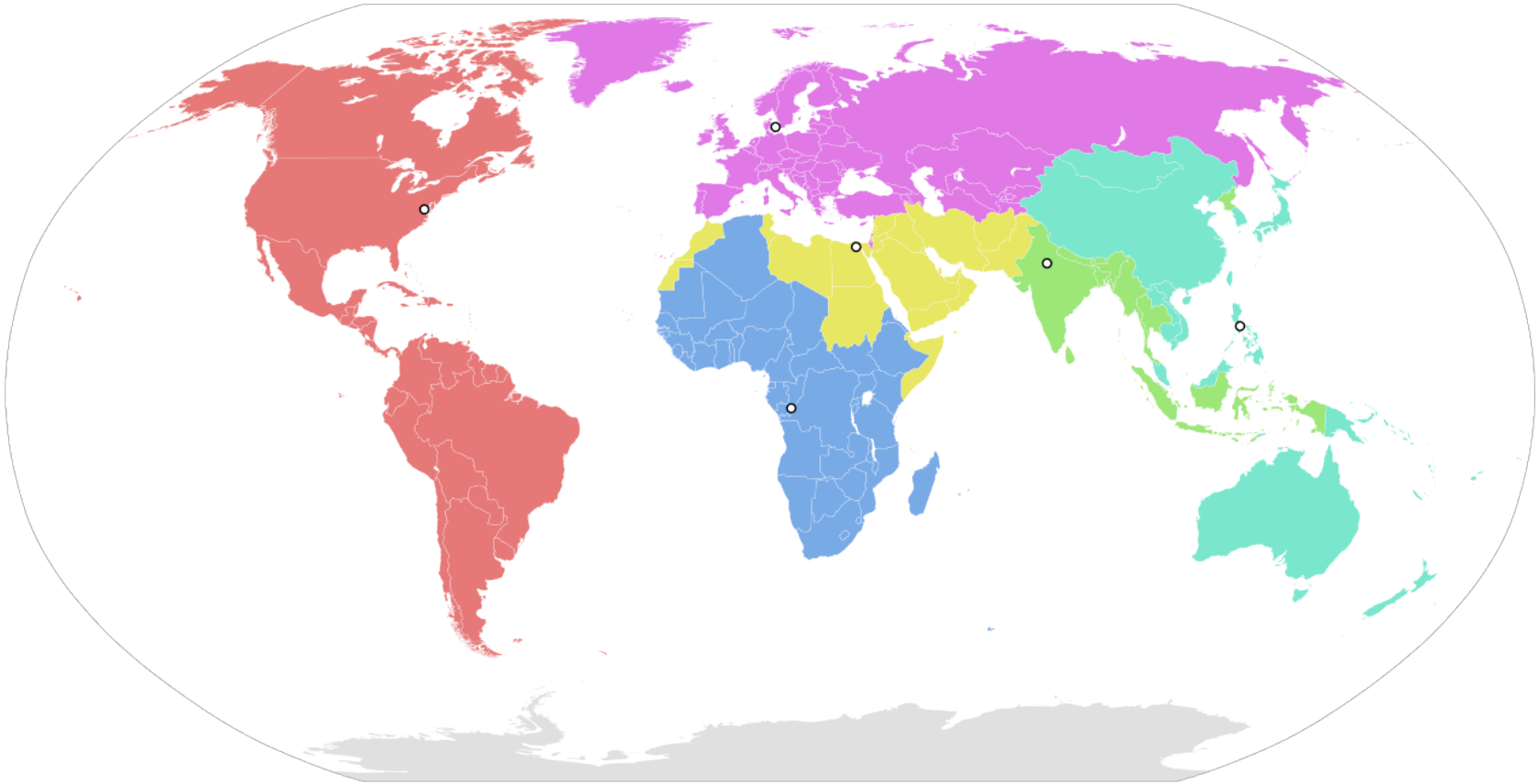
Period	Phase	Description
Interpandemic	1	No animal influenza virus circulating among animals has been reported to cause infection in humans
	2	Animal influenza virus circulating in domesticated or wild animals known to have caused infection in humans and therefore considered a specific potential pandemic threat

WHO pandemic (influenza) phases

Period	Phase	Description
Pandemic alert	3	Animal or human-animal influenza reassortant virus has caused sporadic cases or small clusters of disease in people, but has not resulted in H2H transmission sufficient to sustain community-level outbreaks
	4	Human-to-human transmission of an animal or human-animal influenza reassortant virus able to sustain community-level outbreaks has been verified

WHO pandemic (influenza) phases

Period	Phase	Description
Pandemic alert	5	Same identified virus has caused sustained community-level outbreaks in at least 2 countries in 1 WHO region
Pandemic	6	In addition to criteria in Phase 5, same virus has caused sustained community-level outbreaks in at least 1 other country in another WHO region



Fighting against infections

Epidemiological information is used to plan and evaluate ***strategies to prevent illness*** *and as a guide to the* ***management of patients*** *in whom disease has already developed*

- Preventing illness
  - Prophylactic measures
  - Vaccination

- Managing illness
  - Prevention of further spread (e.g., in hospital)
  - Treatment

Immunisation

- Smallpox first disease for which it was known

- Mentioned in a 1549 Chinese book

- China: powdered smallpox scabs blown up noses of the healthy; variolation-induced mortality not negligible (0.5-2%) but lower than normal (20%)

- 1798: Edward Jenner introduces safer inoculation with cowpox (vaccination)

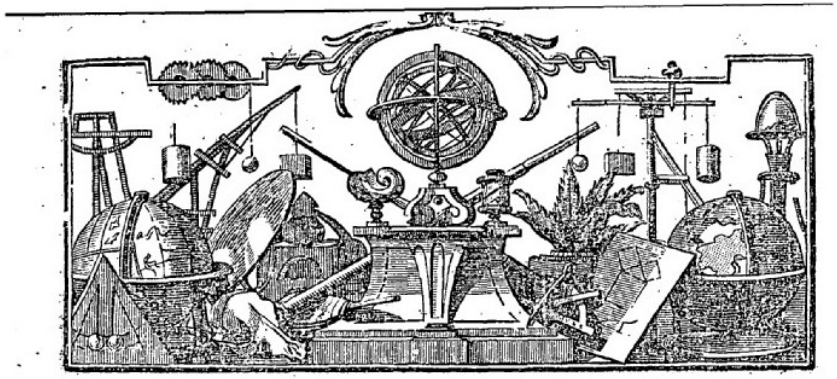
- 1880s: Pasteur extends vaccination to chicken cholera and anthrax in animals and human rabies

At the time, *herd immunity* was not understood so this was for personal protection

The domain is quite old ..

.. but has only become a thing in recent years!

Daniel Bernoulli (1760)



MÉMOIRES
DE
MATHÉMATIQUE
ET
DE PHYSIQUE,
TIRÉS DES REGISTRES de l'Académie Royale des Sciences;
De l'Année M. DCCLX.

ESSAI D'UNE NOUVELLE ANALYSE
De la mortalité causée par la petite Vérole, & des avantages de l'Inoculation pour la prévenir.

Par M. DANIEL BERNOULLI.

INTRODUCTION APOLOGÉTIQUE.*

CEUX qui ont senti tout l'avantage de l'Inoculation, ont imaginé différentes façons de représenter cet avantage, qui, quoique revenant au même, ne laissent pas de faire une

* Cette Introduction n'a été faite que long-temps après le Mémoire, étant du 16 Avril 1765.

Mém. 1760. A

- BNF scan or pdf
- Probably the first epidemic model
- About petite vérole (smallpox) inoculation

Ross (early 1900)

- On 20 August 1897, observed malaria parasites in the gut of a mosquito fed several days earlier on a malaria positive human
- Nobel Prize for Medicine 1902
- Started considering malaria eradication using mathematical models; for some history, read this 2012 paper

Kermack and McKendrick (1927+)

- We spend a lot more time on this later

- Groundbreaking set of papers starting in 1927

- We will see one particular case, the most well known, but this is just the tip of the iceberg of their work

Macdonald, Dietz and malaria

- Read for instance this paper, which presents a history of the development of the so-called Ross-Macdonald model

- Klaus Dietz also worked a lot on malaria

Some activity later, but not much until 1990s

- In recent years, explosion

- Since the beginning of COVID-19: just nuts..

Some landmarks in mathematical epidemiology (IMBO)

- Macdonald. The epidemiology and control of malaria. 1957

- Baroyan, Rvachev et al. Deterministic epidemic models for a territory with a transport network. Kibernetika, 1967

- Hethcote & Yorke. Gonorrhea Transmission Dynamics and Control. LNBM 56, 1984

- Anderson & May. Infectious diseases of humans: dynamics and control. 1991

- Capasso. Mathematical Structures of Epidemic Systems. LNBM 97, 1993

- Hethcote. The mathematics of infectious diseases. SIAM Review, 2000

- van den Driessche & Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. MBS, 2002

A more recent trend

- Some rare numerical work $\leq$ 1980s, mostly simulation of math models

- Baroyan, Rvachev et al. Computer modelling of influenza epidemics for the whole country (USSR). *Advances in Applied Probability* (1971)
- Rvachev & Longini. A mathematical model for the global spread of influenza. *Mathematical Biosciences* (1986)
- Flahault, Letrait et al. Modelling the 1985 influenza epidemic in France. *Statistics in Medicine* (1988)

- More and more frequent now, to the point that some modelling studies are purely simulation-based

Agent-based models (ABM)

- Early in the life of these models, they were called IBM (individual-based models)

- Over the years, a "philosophical" distinction has emerged: - IBM are mathematical models that consider individuals as the units; e.g., DTMC, CTMC, branching processes, etc. - ABM are computational models whose study is, for the most part, only possible numerically

Network models

- Network models endow vertices with simple systems and couple them through graphs

- Can be ABM, but some networks can also be studied analytically

Has happened all along, undergoing a transformation

- Epidemiology has long relied on data

- Many developments in statistics originate there

- Data has traditionally been better for chronic diseases than for infectious ones

- Near-real-time surveillance of infectious diseases ongoing since the 1980s (e.g., Réseau Sentinelles)

- SARS-CoV-1 saw the beginning of a move towards real-time emerging infectious disease data

- With SARS-CoV-2, the system has really progressed a lot, both in terms of "citizen science" and governmental initiatives

Compartmental models

- Have become synonymous with epidemiological models

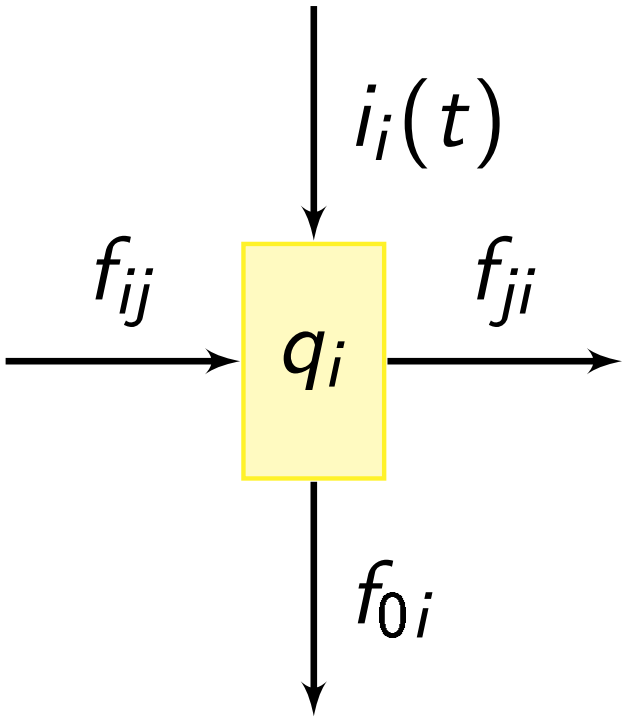
- Many epidemiological models are compartmental models, but the development of compartmental models in the 1970-1980s was not at all specific to epidemiology

- See in particular the works of John Jacquez, Carl Simon, GG Walter

- Unjustly fell into disuse: there are some very nice results in the area

Compartment (Jacquez 1979)

*A **compartment** is an amount of some material which acts kinetically like a distinct, homogeneous, well-mixed amount of material. A **compartmental system** consists of one or more compartments which interact by exchanging the material. There may be inputs into one or more compartments from outside the system and there may be excretions from the compartments of the system.*



- q_i size of the compartment, i.e., quantity of kinetically homogeneous material present in i; $q_i \geq 0$
- f_{ij} and f_{ji} transfer coefficients/functions
- f_{0i} excretion coefficient/function
- $i_i(t)$ entries from outside the system

Above is a **flow diagram**, which summarises the different flows acting on the compartment

Paper series worth reading

Model here is a particular case in

- Kermack & McKendrick. A contribution to the mathematical theory of epidemics (1927)

That paper was followed by a series of "Contributions to the mathematical theory of epidemics."

- II. The problem of endemicity (1932)
- III. Further studies of the problem of endemicity (1933)
- IV. Analysis of experimental epidemics of the virus disease mouse ectromelia (1937)
- V. Analysis of experimental epidemics of mouse-typhoid; a bacterial disease conferring incomplete immunity (1939)

What is the *size* of an epidemic?

- If we are interested in the possibility that an epidemic occurs
  - Does an epidemic peak always take place?
  - If it does take place, what is its size?

- If an epidemic traverses a population, is everyone affected/infected?

The Kermack-McKendrick SIR model without demography

- The period of time under consideration is sufficiently short that demography can be neglected (we also say the model has *no vital dynamics*)

- Individuals are either *susceptible* to the disease or *infected* by (and *infectious* with) the disease

- After recovering or dying from the disease, individuals are *removed* from the infectious compartment (R)

- Incidence is of **mass action** type and takes the form βSI

Important – Incidence functions

Incidence is the rate at which new cases arise, the incidence function then describes how contacts lead to new infections

If there are S susceptible individuals and I infectious individuals in the population, we use a function of the form

$$f(S, I)$$

The function can also explicitly depend on the total population N, i.e., $f(S, I, N)$

We return to incidence functions in Lecture XY

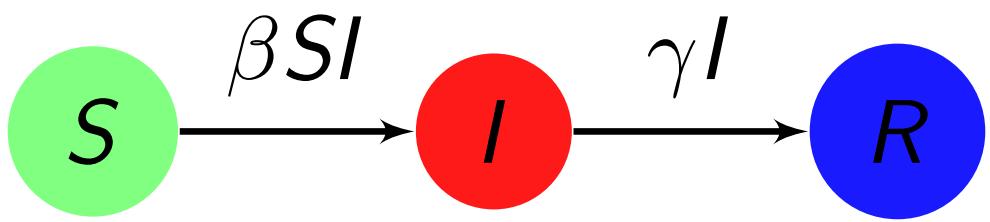
For now, just know the most common incidence functions are

- **mass action incidence** $f(S, I, N) = \beta SI$
- **standard** (or **proportional**) **incidence** $f(S, I, N) = \beta SI/N$

The Kermack-McKendrick model

This model is typically called the Kermack-McKendrick (KMK) SIR model

$$S' = -\beta SI \tag{1a}$$
$$I' = \beta SI - \gamma I \tag{1b}$$
$$R' = \gamma I \tag{1c}$$



Reduction of the model

3 compartments, but when considered in detail, we notice that *removed* do not have a direct influence on the dynamics of S or I, in the sense that R does not appear in (1a) or (1b)

Furthermore, the total population (including deceased who are also in R) $N = S + I + R$ satisfies

$$N' = (S + I + R)' = 0$$

Thus, N is constant and the dynamics of R can be deduced from $R = N - (S + I)$

So we now consider

$$S' = -\beta SI \tag{2a}$$
$$I' = \beta SI - \gamma I \tag{2b}$$

Equilibria

Let us consider the equilibria of

$$S' = -\beta SI \tag{2a}$$
$$I' = (\beta S - \gamma)I \tag{2b}$$

From (2b)

- either $\bar{S} = \gamma/\beta$
- or $\bar{I} = 0$

Substitute into (2a)

- in the first case, $(\bar{S}, \bar{I}) = (\gamma/\beta, 0)$
- in the second case, any $\bar{S} \geq 0$ is an EP

The second case is an *issue*: the usual linearisation does not work when there is a *continuum* of equilibria as the EP are not *isolated*

Another approach – Study dI/dS

$$S' = -\beta SI \tag{2a}$$
$$I' = \beta SI - \gamma I \tag{2b}$$

What is the dynamics of dI/dS?

$$\frac{dI}{dS} = \frac{dI}{dt}\frac{dt}{dS} = \frac{I'}{S'} = \frac{\beta SI - \gamma I}{-\beta SI} = \frac{\gamma}{\beta S} - 1 \tag{4}$$

provided $S \neq 0$

Note – Recall that S and I are $S(t)$ and $I(t)$.. (4) thus describes the relation between S and I over solutions to the original ODE (2)

Integrate (4) and obtain trajectories in state space

$$I(S) = \frac{\gamma}{\beta} \ln S - S + C$$

with $C \in \mathbb{R}$

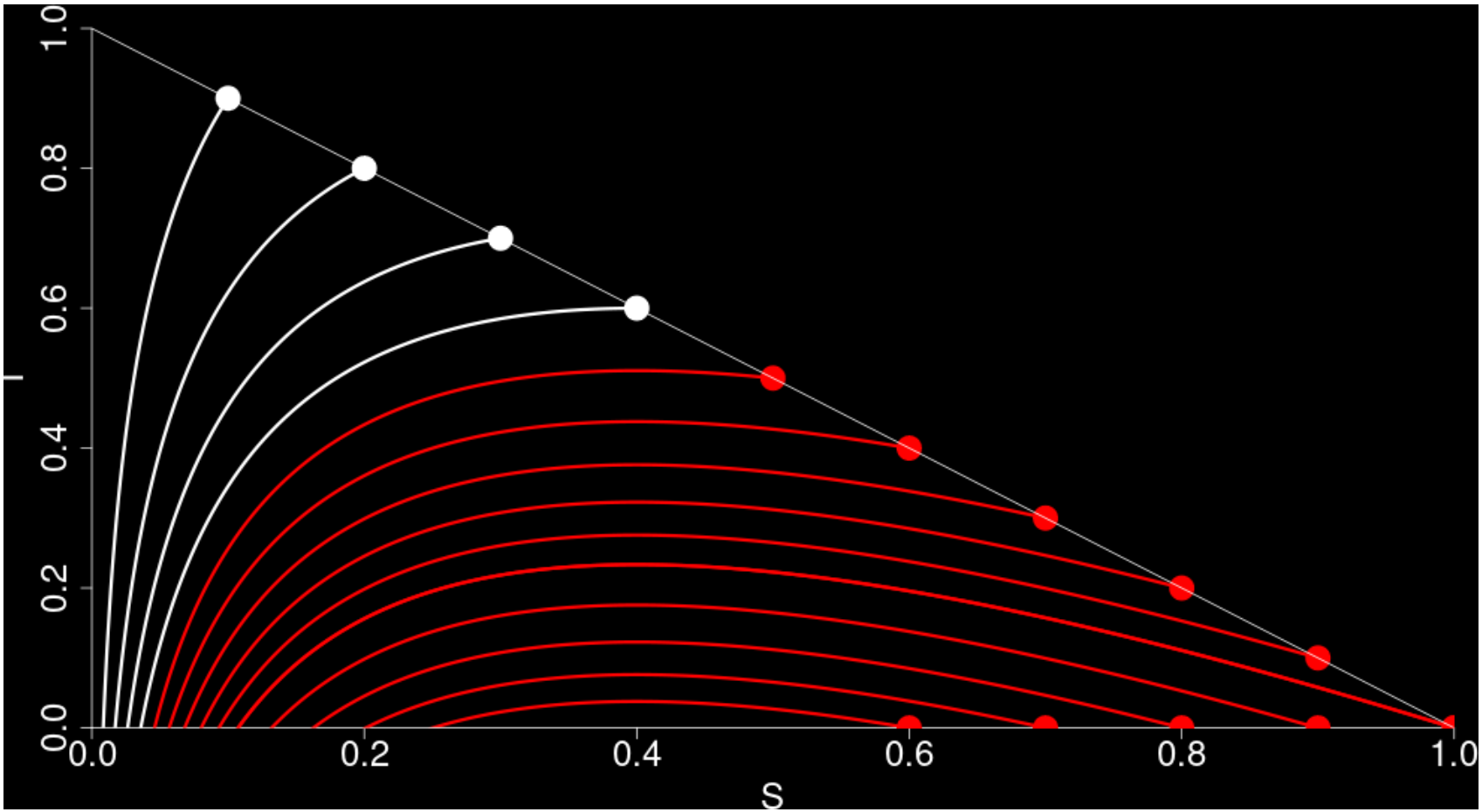
IC $I(S_0) = I_0 \Rightarrow C = S_0 + I_0 - \frac{\gamma}{\beta} \ln S_0$ and the solution to (1) is, as a function of S

$$I(S) = S_0 + I_0 - S + \frac{\gamma}{\beta} \ln \frac{S}{S_0}$$

$$R(S) = N - S - I(S) = R_0 - \frac{\gamma}{\beta} \ln \frac{S}{S_0}$$

(since $N_0 = S_0 + I_0 + R_0$)

Trajectories of (2) in (S, I)-space, normalised, with IC $(S_0, 1 - S_0)$ and $\beta/\gamma = 2.5$



Let us study

$$I(S) = S_0 + I_0 - S + \frac{\gamma}{\beta} \ln \frac{S}{S_0}$$

We have

$$\frac{d}{dS} I(S) = \frac{\gamma}{\beta S} - 1$$

So, in the previous curves, the max of $I(S)$ happens when $S = \gamma/\beta$ ($S = 0.4$ in the example)

At that point,

$$I(S) = I_0 + \left(1 - \frac{1}{\mathcal{R}_0} - \frac{\ln(\mathcal{R}_0)}{\mathcal{R}_0}\right) S_0$$

Theorem 1 (Epidemic or no epidemic?)

Let $(S(t), I(t))$ be a solution to (2) and $\mathcal{R}_0$ defined by

$$\mathcal{R}_0 = \frac{\beta}{\gamma} S_0 \tag{5}$$

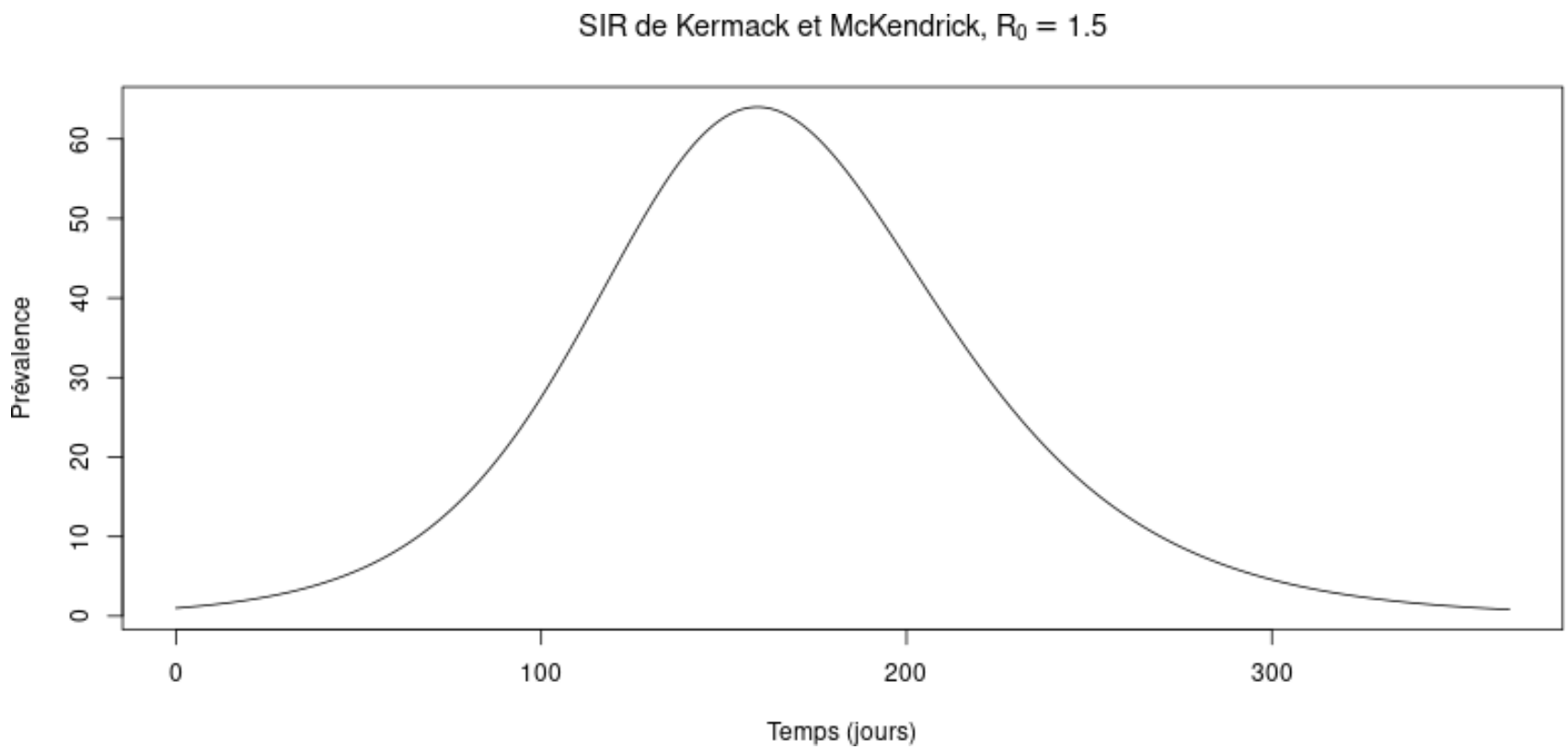
- *If $\mathcal{R}_0 \leq 1$, then $I(t) \searrow 0$ when $t \to \infty$*
- *If $\mathcal{R}_0 > 1$, then $I(t)$ first reaches a maximum*

$$I_0 + \left(1 - \frac{1}{\mathcal{R}_0} - \frac{\ln(\mathcal{R}_0)}{\mathcal{R}_0}\right) S_0 \tag{6}$$

then goes to 0 as $t \to \infty$

```r
rhs_SIR_KMK <- function(t, x, p) {
  with(as.list(c(x, p)), {
    dS = - beta * S * I
    dI = beta * S * I - gamma * I
    dR = gamma * I
    return(list(c(dS, dI, dR)))
  })
}
# Condition initiale pour S (pour calculer R_0)
S0 = 1000
gamma = 1/14
# Set beta so that R_0 = 1.5
beta = 1.5 * gamma / S0
params = list(gamma = gamma, beta = beta)
IC = c(S = S0, I = 1, R = 0)
times = seq(0, 365, 1)
sol <- ode(IC, times, rhs_SIR_KMK, params)
```

```
plot(sol[, "time"], sol[, "I"], type = "l",
main = TeX("Kermack-McKendrick␣SIR,␣$R_0=1.5$"),
xlab = "Time␣(days)", ylab = "Prevalence")
```



The basic reproduction number $\mathcal{R}_0$

- Indicator often used in epidemiology. Verbally

  average number of secondary cases of infection produced when a single infectious individual is introduced in a wholly susceptible population

- If $\mathcal{R}_0 < 1$, then each infectious individual infects on average less than 1 person and the epidemic is quite likely to go extinct

- If $R_0 > 1$, then each infectious individual infects on average more than 1 person and an epidemic is quite likely to occur

A few sample values of $\mathcal{R}_0$

$\mathcal{R}_0$ can be estimated from data

Infection	Location	Period	$\mathcal{R}_0$
Measles	Cirencester, England	1947-50	13-14
	England and Wales	1950-68	16-18
	Kansas, USA	1918-21	5-6
	Ontario, Canada	1912-3	11-12
	Willesden, England	1912-3	11-12
	Ghana	1960-8	14-15
	East Nigeria	1960-8	16-17

Final size of an epidemic

For a nonnegative valued integrable function $w(t)$, denote

$$w_\infty = \lim_{t\to\infty} w(t), \qquad \hat{w} = \int_0^\infty w(t)\, dt$$

Denote $w_0 = w(0)$. In the subsystem

$$S' = -\beta SI \tag{2a}$$
$$I' = \beta SI - \gamma I \tag{2b}$$

compute the sum of (2a) and (2b), making sure to show time dependence

$$\frac{d}{dt}(S(t) + I(t)) = -\gamma I(t)$$

Integrate from 0 to ∞:

$$\int_0^\infty \frac{d}{dt}(S(t) + I(t))\ dt = -\int_0^\infty \gamma I(t) dt$$

The left hand side gives

$$\int_0^\infty \frac{d}{dt}(S(t) + I(t))\ dt = S_\infty + I_\infty - S_0 - I_0 = S_\infty - S_0 - I_0$$

since $I_\infty = 0$

The right hand side takes the form

$$-\int_0^\infty \gamma I(t) dt = -\gamma \int_0^\infty I(t) dt = -\gamma \hat{I}$$

We thus have

$$S_\infty - S_0 - I_0 = -\gamma \hat{I} \tag{7}$$

Now consider (2a):

$$S' = -\beta SI$$

Divide both sides by S:

$$\frac{S'(t)}{S(t)} = -\beta I(t)$$

Integrate from 0 to ∞:

$$\ln S_\infty - \ln S_0 = -\beta \hat{I} \tag{8}$$

Express (7) and (8) in terms of $-\hat{I}$ and equate

$$\frac{\ln S_\infty - \ln S_0}{\beta} = \frac{S_\infty - S_0 - I_0}{\gamma}$$

Thus we have

$$(\ln S_0 - \ln S_\infty)S_0 = (S_0 - S_\infty)\mathcal{R}_0 + I_0\mathcal{R}_0 \tag{9}$$

Theorem 2 (Final size relation)

Let $(S(t), I(t))$ *be a solution to* (2) *and* $\mathcal{R}_0$ *defined by* (5)

The number $S(t)$ *of susceptible individuals is a nonincreasing function and its limit* S_∞ *is the only solution in* $(0, S_0)$ *of the transcendental equation*

$$(\ln S_0 - \ln S_\infty)S_0 = (S_0 - S_\infty)\mathcal{R}_0 + I_0\mathcal{R}_0 \tag{9}$$

The (transcendantal) final size equation

Rewrite the final size equation

$$(\ln S_0 - \ln S_\infty)S_0 = (S_0 - S_\infty)\mathcal{R}_0 + I_0\mathcal{R}_0 \tag{9}$$

as

$$T(S_\infty) = (\ln S_0 - \ln S_\infty)S_0 - (S_0 - S_\infty)\mathcal{R}_0 - I_0\mathcal{R}_0 \tag{10}$$

Thus, we seek the zeros of the function $T(S_\infty)$

We seek S_∞ in $(0, S_0]$ s.t. $T(S_\infty) = 0$, with

$$T(S_\infty) = (\ln S_0 - \ln S_\infty)S_0 - (S_0 - S_\infty)\mathcal{R}_0 - I_0\mathcal{R}_0 \qquad (10)$$

Note to begin that

$$\lim_{S_\infty \to 0} T(S_\infty) = \lim_{S_\infty \to 0} -S_0 \ln(S_\infty) = \infty$$

Differentiating T with respect to S_∞, we get

$$T'(S_\infty) = \mathcal{R}_0 - S_0/S_\infty$$

When $S_\infty \to 0$, $\mathcal{R}_0 - S_0/S_\infty < 0$, so T decreases to $S_\infty = S_0/\mathcal{R}_0$

So if $\mathcal{R}_0 \leq 1$, the function T is decreasing on $(0, S_0)$, while it has a minimum if $\mathcal{R}_0 > 1$

Case $\mathcal{R}_0 \leq 1$

$$T(S_\infty) = (\ln S_0 - \ln S_\infty)S_0 - (S_0 - S_\infty)\mathcal{R}_0 - I_0\mathcal{R}_0 \qquad (10)$$

- We have seen that T decreases on $(0, S_0]$
- Also, $T(S_0) = -I_0\mathcal{R}_0 < 0$ ($I_0 = 0$ is trivial and not considered)
- T is continuous

$\Longrightarrow$ there exists a unique $S_\infty \in (0, S_0]$ s.t. $T(S_\infty) = 0$

Case $\mathcal{R}_0 > 1$

$$T(S_\infty) = (\ln S_0 - \ln S_\infty)S_0 - (S_0 - S_\infty)\mathcal{R}_0 - I_0\mathcal{R}_0 \qquad (10)$$

- We have seen that T decreases on $(0, S_0/\mathcal{R}_0]$
- For $S_\infty \in [S_0/\mathcal{R}_0]$, $T' > 0$
- As before, $T(S_\infty) = -I_0\mathcal{R}_0$
- T is continuous

$\Longrightarrow$ there exists a unique $S_\infty \in (0, S_0]$ s.t. $T(S_\infty) = 0$. More precisely, in this case, $S_\infty \in (0, S_0/\mathcal{R}_0)$

We solve numerically. We need a function

```
final_size_eq = function(S_inf, S0 = 999, I0 = 1, R_0 = 2.5) {
  OUT = S0*(log(S0)-log(S_inf)) - (S0+I0-S_inf)*R_0
  return(OUT)
}
```

and solve easily using `uniroot`, here with the values by default that we have set for the function

```
uniroot(f = final_size_eq, interval = c(0.05, 999))
$root
[1] 106.8819
$f.root
[1] -2.649285e-07
$iter
[1] 10
$init.it
[1] NA
$estim.prec
[1] 6.103516e-05
```

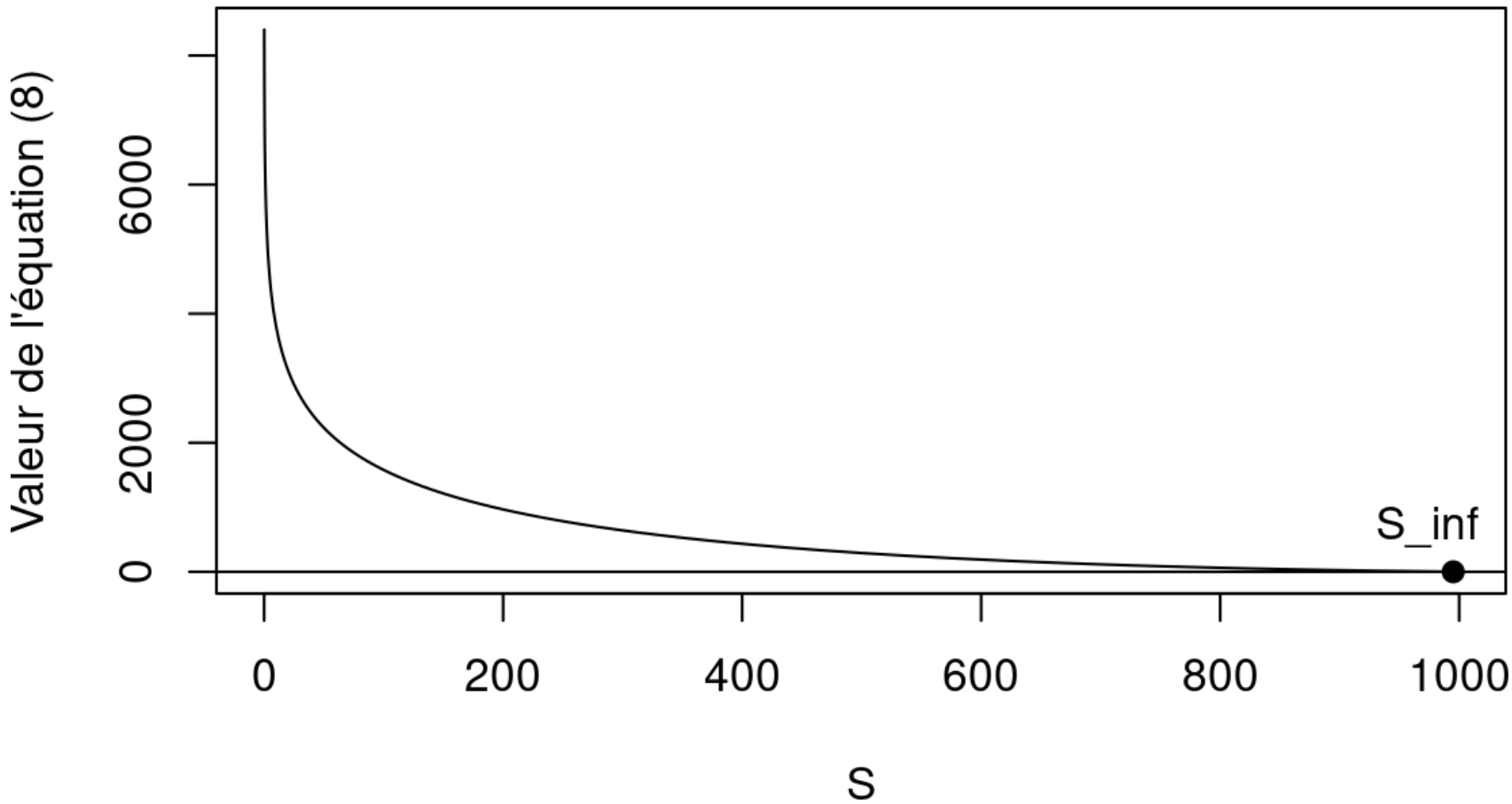
To use something else than the default values, e.g.,

```
N0 = 1000
I0 = 1
S0 = N0-I0
R_0 = 2.4
uniroot(
  f = function(x)
    final_size_eq(S_inf = x,
                  S0 = S0, I0 = I0,
                  R_0 = R_0),
  interval = c(0.05, S0))
```

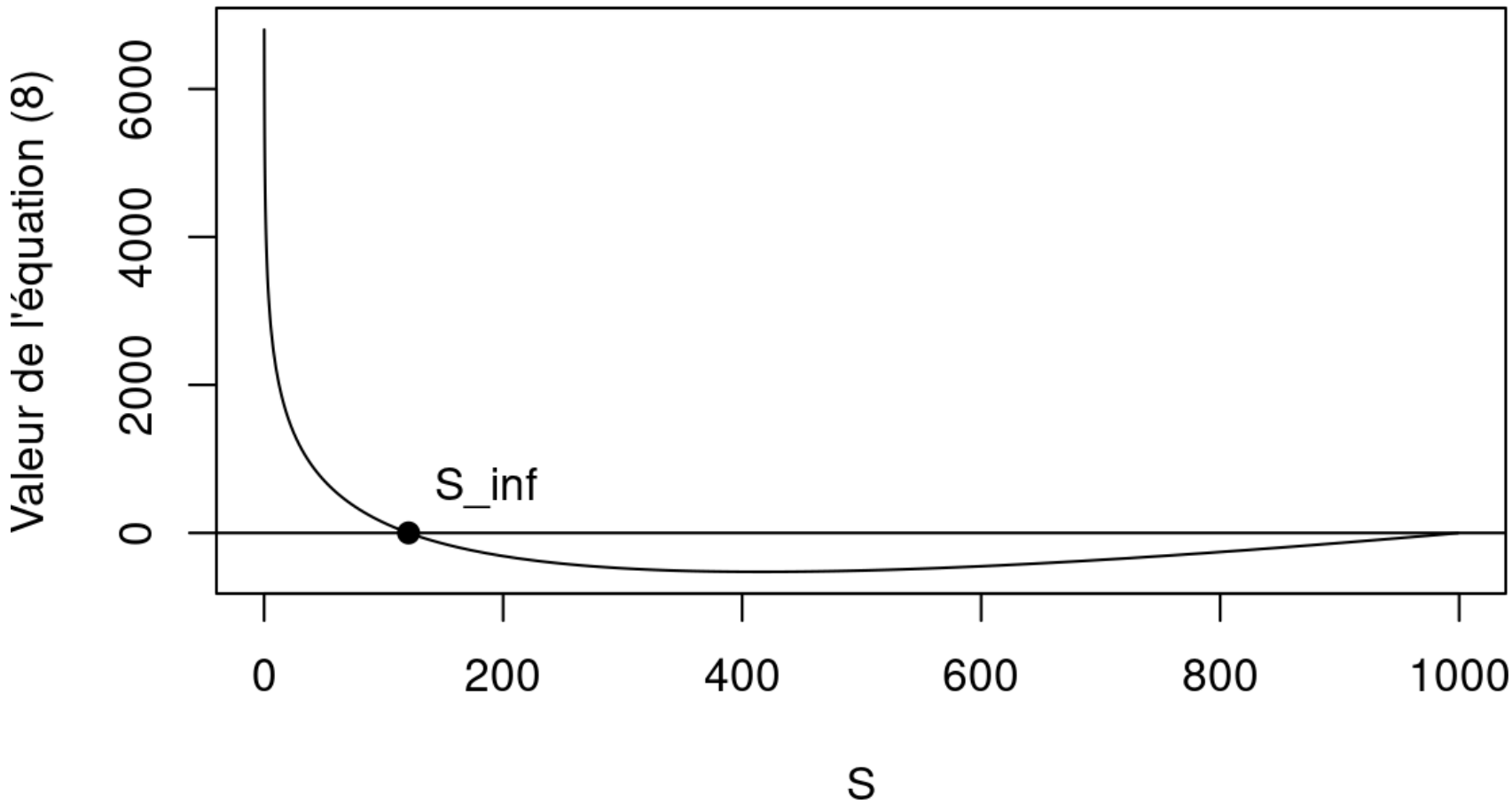
A figure with all the information

```
S = seq(0.1, S0, by = 0.1)
fs = final_size(S, S0 = S0, I0 = I0, R_0 = R_0)
S_inf = uniroot(f = function(x) final_size_eq(S_inf = x,
                                              S0 = S0, I0 = I0,
                                              R_0 = R_0),
                interval = c(0.05, S0))
plot(S, fs, type = "l", ylab = "Value of equation (10)")
abline(h = 0)
points(x = S_inf$root, y = 0, pch = 19)
text(x = S_inf$root, y = 0, labels = "S_inf", adj = c(-0.25,-1))
```

$\mathcal{R}_0 = 0.8$



$\mathcal{R}_0 = 2.4$

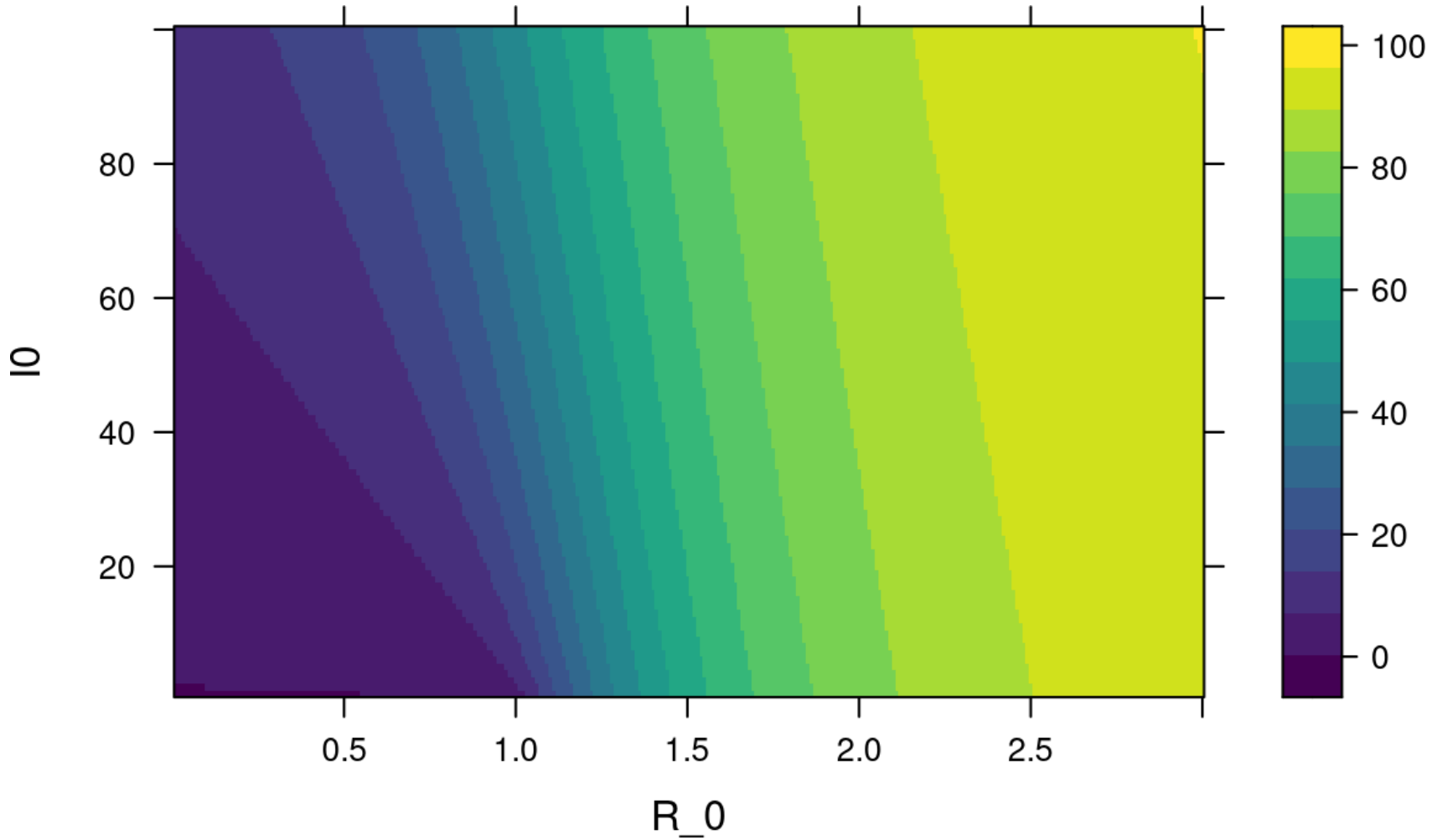


A little nicer

```
values = expand.grid(
  R_0 = seq(0.01, 3, by = 0.01),
  I0 = 1:100
)
values$S0 = N0-values$I0
L = split(values, 1:nrow(values))

values$S_inf = sapply(X = L, FUN = final_size)

values$taille_finale = values$S0-values$S_inf+values$I0
values$taux_attaque = (values$taille_finale / N0)*100

levelplot(taux_attaque ~ R_0*I0, data = values,
          xlab="R_0", ylab = "I0",
          col.regions = viridis(100))
```

(requires `lattice` and `viridis` librairies)

Attack rate (in %)



The simplest vaccination model

To implement vaccination in KMK, assume that vaccination reduces the number of susceptibles

Let total population be N with S_0 initially susceptible

Vaccinate a fraction $p \in [0, 1]$ of susceptible individuals

Original IC (for simplicity, $R(0) = 0$)

$$IC : (S(0), I(0), R(0)) = (S_0, I_0, 0) \tag{11}$$

Post-vaccination IC

$$IC : (S(0), I(0), R(0)) = ((1 - p)S_0, I_0, pS_0) \tag{12}$$

Vaccination reproduction number

Without vaccination

$$\mathcal{R}_0 = \frac{\beta}{\gamma} S_0 \tag{5}$$

With vaccination, denoting $\mathcal{R}_0^{\mathrm{v}}$ the reproduction number,

$$\mathcal{R}_0^{\mathrm{v}} = \frac{\beta}{\gamma}(1-p)S_0 \tag{13}$$

Since $p \in [0,1]$, $\mathcal{R}_0^{\mathrm{v}} \leq \mathcal{R}_0$

Herd immunity

Therefore

- $\mathcal{R}_0^{\mathrm{v}} < \mathcal{R}_0$ if $p > 0$
- To control the disease, $\mathcal{R}_0^{\mathrm{v}}$ must take a value less than 1

To make $\mathcal{R}_0^{\mathrm{v}}$ less than 1

$$\mathcal{R}_0^{\mathrm{v}} < 1 \iff p > 1 - \frac{1}{\mathcal{R}_0} \tag{14}$$

By vaccinating a fraction $p > 1 - 1/\mathcal{R}_0$ of the susceptible population, we thus are in a situation where an epidemic peak is precluded (or, at the very least, the final size is reduced)

This is **herd immunity**

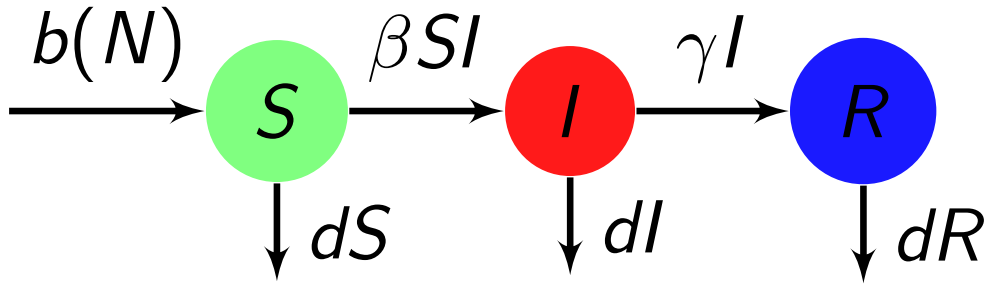
Two potential variations on the Kermack-McKendrick model

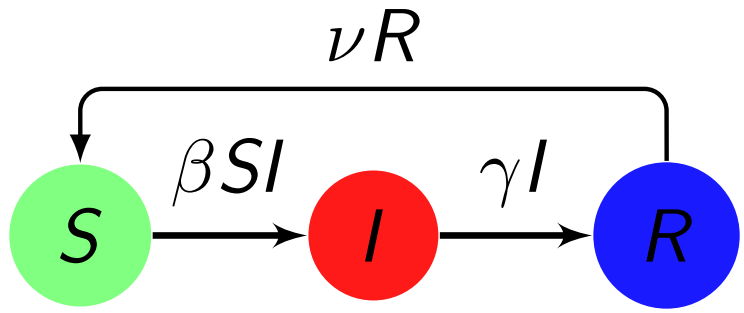
- Add *vital dynamics*, i.e., consider demographic processes

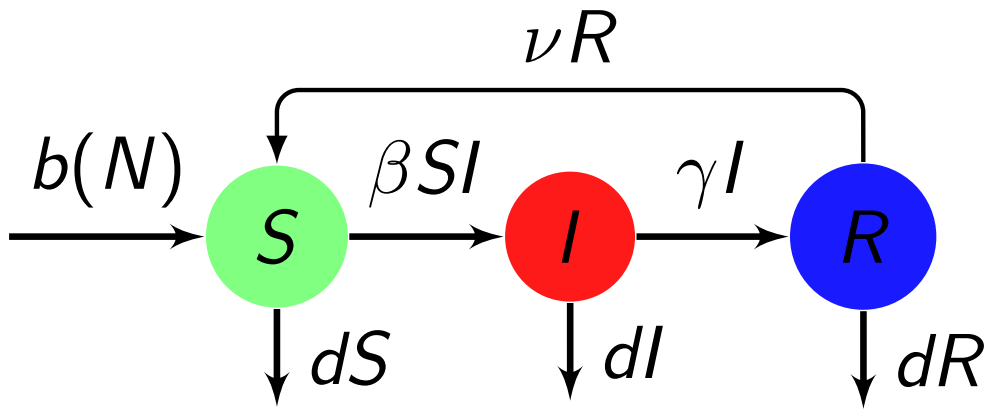
- Individuals do not die from the disease; after recovering, individuals are *immune* from infection for some time

- We can of course combine both!

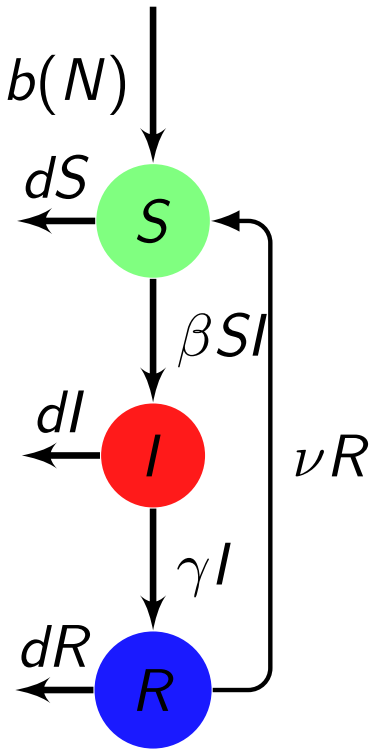
Potential variations







The model



$$S' = b(N) + \nu R - dS - \beta SI \tag{15a}$$
$$I' = \beta SI - (d + \gamma)I \tag{15b}$$
$$R' = \gamma I - (d + \nu)R \tag{15c}$$

Consider the initial value problem consisting in (15) to which we couple initial conditions $S(0) = S_0 \geq 0$, $I(0) = I_0 \geq 0$ and $R(0) = R_0 \geq 0$

Typically, we assume $N_0 = S_0 + I_0 + R_0 > 0$ to avoid a trivial case

Birth and death are *relative*

Remark that the notions of *birth* and *death* are relative to the population under consideration

E.g., consider a model for human immunodeficiency virus (HIV) in an at-risk population of intravenous drug users. Then

- birth is the moment the at-risk behaviour starts
- death is the moment the at-risk behaviour stops, whether from "real death" or because the individual stops using drugs

Choosing a form for demography

Before we proceed with the analysis proper, we must discuss the nature of the assumptions on demography

To do this, we consider the behaviour of the total population

$$N(t) = S(t) + I(t) + R(t)$$

Behaviour of the total population

Summing the equations in (15)

$$N' = b(N) - dN \tag{16}$$

There are three common ways to define $b(N)$ in (16)

1. $b(N) = b$
2. $b(N) = bN$
3. $b(N) = bN - cN^2$

Case 3 leads to logistic dynamics of the total population and is not discussed here

Case of a birth rate constant *per capita*

If $b(N) = bN$, then birth in (16) satisfies $N'/N = b$; we say that birth is **constant *per capita***

In this case, (16) takes the form

$$N' = bN - dN = (b - d)N$$

with initial condition $N(0) = N_0$

The solution to this scalar autonomous ODE is easy

$$N(t) = N_0 e^{(b-d)t}, \quad t \geq 0$$

Thus there are 3 possibilities:

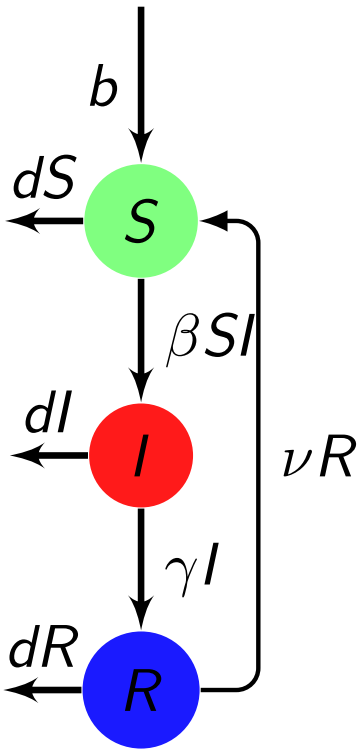
- if $b > d$, $N(t) \to \infty$, the total population explodes
- if $b = d$, $N(t) \equiv N_0$, the total population remains constant
- if $b < d$, $N(t) \to 0$, the total population collapses

From now on, assume $b(N) = b$

- We want a reasonable case, we could therefore suppose that $b(N) = d$, which would lead to a constant total population

- However, this is a little reductive, so we choose instead $b(N) = b$, which, we will see, works as well even though it can initially be thought of as not being very realistic

The model (for good this time)



$$S' = b + \nu R - dS - \beta SI \tag{17a}$$
$$I' = \beta SI - (d + \gamma)I \tag{17b}$$
$$R' = \gamma I - (d + \nu)R \tag{17c}$$

Consider the initial value problem consisting in (17) to which we couple initial conditions $S(0) = S_0 \geq 0$, $I(0) = I_0 \geq 0$ and $R(0) = R_0 \geq 0$

Typically, we assume $N_0 = S_0 + I_0 + R_0 > 0$ to avoid a trivial case

Is the system well-posed?

For an ODE epidemiological model

- Do solutions to (17) exist and are they unique?
- Is the positive cone invariant under the flow of (17)?
- Are solutions to (17) bounded? Some models have unbounded solutions but they are rare and will need to be considered specifically

Solutions exist and are unique

- The vector field is always C^1, implying that solutions exist and are unique

If we had instead considered an incidence of the form $f(S, I, N) = \beta SI/N$ and, say, demography with $b(N) = bN$, then some discussion might have been needed if $b < d$

To simplify or not to simplify

Invariance of $\mathbb{R}^3_+$ under the flow (1)

Let us start by assuming that $I = 0$. Then (??) reduces to the scalar equation

$$S' = bN - dS = 0$$

since $S = N - I = N$ here

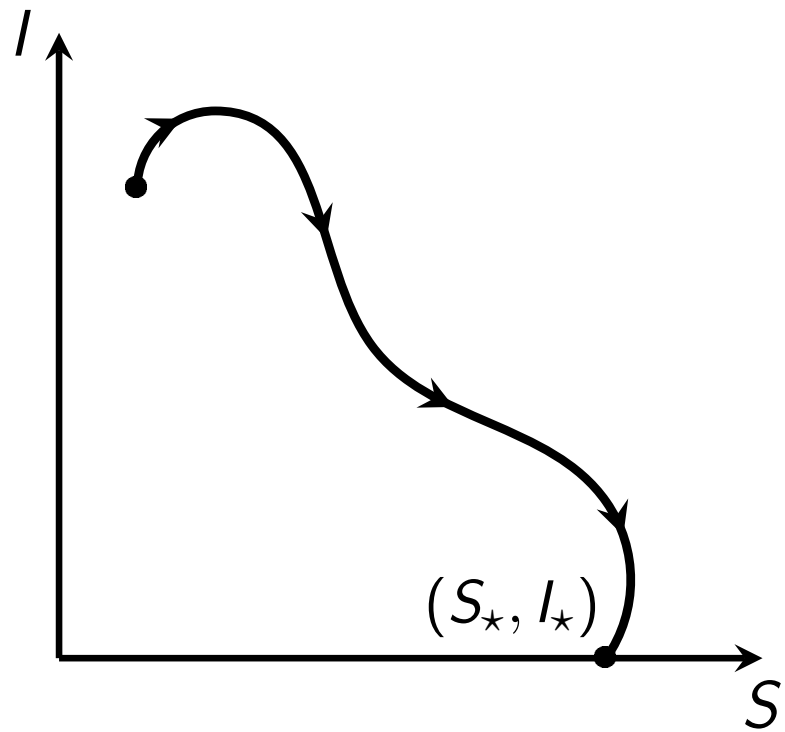
It follows that the axis $\{I = 0\}$ is invariant and any solution initiated in this set is constant

This implies that a solution with $S(0) > 0$ and $I(0) > 0$ cannot reach the axis $\{I = 0\}$

Invariance of $\mathbb{R}^3_+$ under the flow (2)

This implies that a solution with $S(0) > 0$ and $I(0) > 0$ cannot reach the axis $\{I = 0\}$

Suppose that $S(0) = S_0 > 0$ and $I(0) = I_0 > 0$ and that there exists $t_\star > 0$ such that $S(t_\star) = S_\star$ and $I(t_\star) = 0$



Invariance of $\mathbb{R}^2_+$ under the flow (3)

However, in $S = S_\star$ and $I = 0$, another solution is present, the one such that $S(0) = S_\star$ and $I(0) = 0$

Indeed, with these IC, we have $S(t) = S_\star$ and $I(t) = 0$ for all $t \geq 0$

This contradicts uniqueness of solutions

$\Rightarrow I(t) > 0$ if $I(0) > 0$

Invariance of $\mathbb{R}^2_+$ under the flow (4)

We saw that $I(t) > 0$ if $I(0) > 0$

Suppose now that $S = 0$. Equation (**??**) is then

$$S' = dN + \gamma I > 0$$

So if $S(0) > 0$, then $S(t) > 0$ for all t. If, on the other hand, $S(0) = 0$, then $S(t) > 0$ for $t > 0$ small; from what we just saw, this is then also true for all $t > 0$

We say the vector field points *inward*

$\Longrightarrow$ S cannot become zero

To summarise, for invariance

- If $(S(0), I(0)) \in \mathbb{R}_+ \times (\mathbb{R}_+ \setminus \{0\})$, then $\forall t > 0$,

$$(S(t), I(t)) \in (\mathbb{R}_+ \setminus \{0\})^2$$

- If $(S(0), I(0)) \in \mathbb{R}_+ \times \{0\}$, then $\forall t \geq 0$,

$$(S(t), I(t)) = (S(0), 0)$$

The model is therefore satisfactory in that it does not allow solutions to become negative

Remark – Know your audience

This reasoning has its place in an MSc of PhD manuscript: you need to demonstrate that you know what to do and how to do it

In a research paper, this is not really necessary and actually often superfluous; *it is easy to show that solutions exist uniquely and that the positive orthant is invariant under the flow of the system* is typically sufficient

Boundedness

It follows from what we just saw that the positive quadrant $\mathbb{R}^2_+$ is (positively) invariant under the flow of (??)

From invariance and boundedness (constance) of the total population N, we deduce that solutions of (??) are bounded

Seeking equilibria (1)

We find equilibria of the system by seeking values of S and I such that $S' = I' = 0$

To do that, we set $S' = I' = 0$ and solve for S and I

At an EP, (??) becomes

$$0 = d(N - S) - \beta \frac{SI}{N} + \gamma I \tag{18a}$$

$$0 = \left(\beta \frac{S}{N} - (d + \gamma) \right) I \tag{18b}$$

Seeking equilibria (2)

We consider

$$0 = d(N - S) - \beta \frac{SI}{N} + \gamma I \tag{18a}$$

$$0 = \left(\beta \frac{S}{N} - (d + \gamma) \right) I \tag{18b}$$

From (18b), we deduce that either

$$\beta \frac{S}{N} - (d + \gamma) = 0 \iff S = \frac{d + \gamma}{\beta} N$$

or $I = 0$

Substituting $I = 0$ in (18a) gives $d(N - S) = 0$, i.e, $S = N$. This is the disease-free equilibrium (DFE)

$$E_0 : (S, I) = (N, 0) \tag{19}$$

Seeking equilibria (3)

From

$$S = \frac{d+\gamma}{\beta} N$$

we deduce the endemic equilibrium (EEP) $E_\star$: substituting this value of S into (18a),

$$\begin{aligned} 0 &= d\left(N - \frac{d+\gamma}{\beta}N\right) - \beta\frac{\frac{d+\gamma}{\beta}NI}{N} + \gamma I \\ &= dN\left(1 - \frac{d+\gamma}{\beta}\right) - (d+\gamma)I + \gamma I \\ &= dN\left(1 - \frac{d+\gamma}{\beta}\right) - dI \end{aligned}$$

whence

$$I = \left(1 - \frac{d+\gamma}{\beta}\right) N$$

Seeking equilibria (4)

As a consequence, the EEP is

$$E_\star : (S, I) = \left(\frac{d+\gamma}{\beta} N, 1 - \frac{d+\gamma}{\beta} N \right)$$

Define, as previously, the basic reproduction number by (**??**) and noting that $(d+\gamma)/\beta = 1/\mathcal{R}_0$, it follows that

$$E_\star : (S, I) = \left(\frac{1}{\mathcal{R}_0} N, \left(1 - \frac{1}{\mathcal{R}_0} \right) N \right) \tag{20}$$

Classic method for computing $\mathcal{R}_0$

$\mathcal{R}_0$ is the surface in parameter space where the DFE loses its LAS

To find $\mathcal{R}_0$, we therefore study the LAS of the DFE

In an arbitrary (S, I), the Jacobian matrix of (??) takes the form

$$J_{(S,I)} = \begin{pmatrix} -d - \beta \frac{I}{N} & \gamma - \beta \frac{S}{N} \\ \beta \frac{I}{N} & \beta \frac{S}{N} - (d + \gamma) \end{pmatrix} \tag{21}$$

The LAS of the DFE depends on the sign of the real parts of the eigenvalues of (21)at that equilibrium point, so we evaluate

$$J_{E_0} = \begin{pmatrix} -d & \gamma - \beta \\ 0 & \beta - (d + \gamma) \end{pmatrix}$$

Triangular matrix $\implies$ eigenvalues are $-d < 0$ and $\beta - (d + \gamma)$

$\implies$ LAS of the DFE determined by sign of $\beta - (d + \gamma)$, giving (??)

A faster way to compute $\mathcal{R}_0$

Diekmann & Heesterbeek, caracterised in the ODE case by PvdD & Watmough (2002)

We only consider compartments with *infected* individuals and write

$$x' = \mathcal{F} - \mathcal{V}$$

- $\mathcal{F}$ entry into infected compartments because of new infections
- $\mathcal{V}$ contains all other flows (beware the $-$ sign)

Compute the (Fréchet) derivatives $F = D\mathcal{F}$ and $V = D\mathcal{V}$ with respect to the infected variables x (the Jacobian matrices) and evaluate at the DFE

Then

$$\mathcal{R}_0 = \rho(FV^{-1}) \tag{22}$$

where ρ is the spectral radius

A result of PvdD & Watmough (2002)

Theorem 3

Suppose a DFE exists and denote

$$\mathcal{R}_0 = \rho(FV^{-1}) \tag{22}$$

where matrices F and V are obtained as indicated. Suppose that conditions (A1) to (A5) are satisfied. Then

- *if $\mathcal{R}_0 < 1$, then the DFE is LAS*
- *if $\mathcal{R}_0 > 1$, then the DFE is unstable*

(Conditions (A1)-(A5) are made explicit in Lecture XY; we discuss why it is important to check these conditions in Lecture XY)

Using the next generation method here

This computation replaces the one involving the Jacobian matrix. Once the DFE has been found, consider only the infected variables in (??), i.e., I, and its corresponding equation (??):

$$I' = \beta \frac{SI}{N} - (d + \gamma)I$$

We write this equation in the form

$$x' = \mathcal{F} - \mathcal{V}$$

with $x = I$, $\mathcal{F}$ the new infections, i.e.,

$$\mathcal{F} = \beta \frac{SI}{N}$$

and $\mathcal{V}$ all other flows, being careful about the $-$ sign:

$$\mathcal{V} = (d + \gamma)I$$

Compute the Jacobian matrices $\mathcal{F}$ and $\mathcal{V}$ (scalars here since $\mathcal{F}$ and $\mathcal{V}$ are scalar)

$$D\mathcal{F} = \frac{\partial}{\partial I}\mathcal{F} = \beta\frac{S}{N}$$

and

$$D\mathcal{V} = \frac{\partial}{\partial I}\mathcal{V} = d + \gamma$$

We obtain F and V by evaluating these derivatives at the DFE,

$$F = D\mathcal{F}_{E_0} = \beta\frac{S}{N} = \beta, \qquad V = D\mathcal{V}_{E_0} = d + \gamma$$

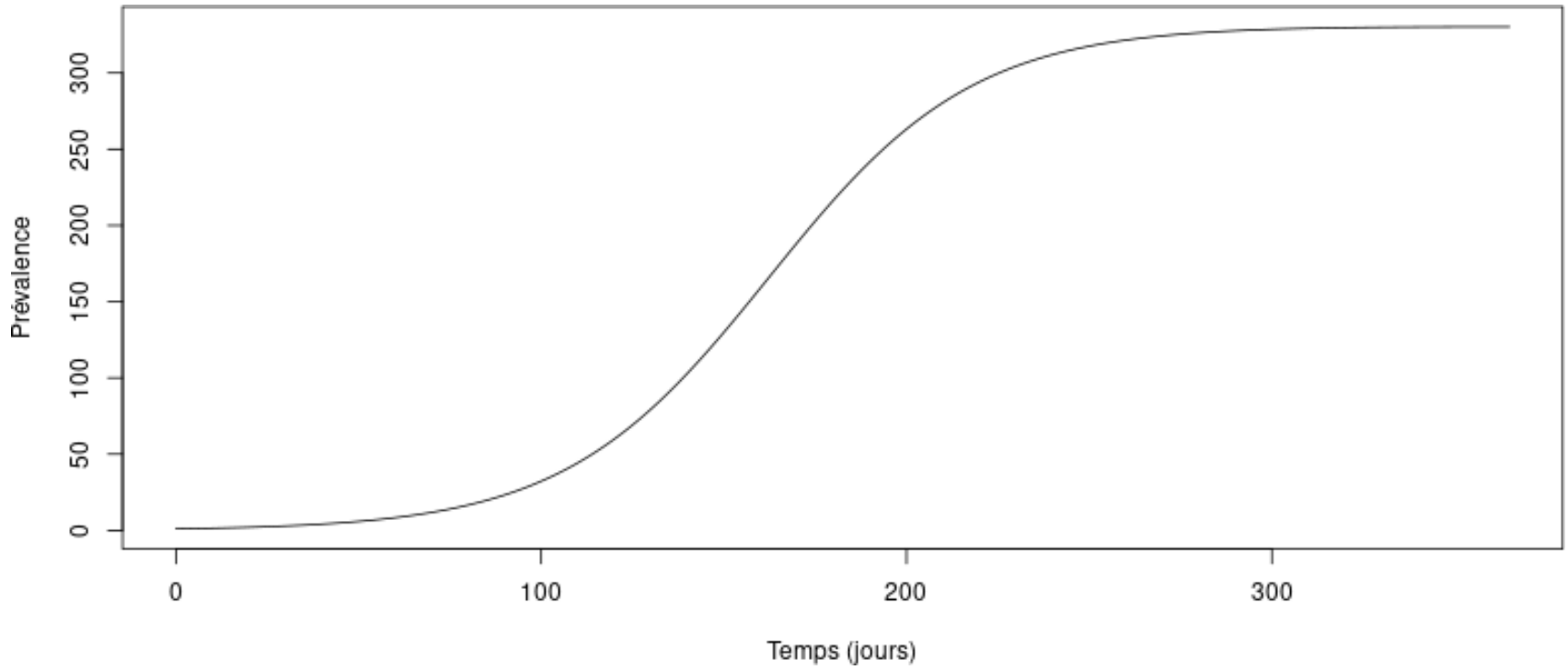
Finally, invert V, i.e., here, $V^{-1} = 1/(d + \gamma)$ The conditions (A1)-(A5) of Theorem 3 are verified, so we also get LAS of the DFE

Thus, just like (??),

$$\mathcal{R}_0 = \rho(FV^{-1}) = \rho\left(\frac{\beta}{d+\gamma}\right) = \frac{\beta}{d+\gamma}$$

The benefit here is not necessarily obvious, but we will see later that this method can greatly simplify computations

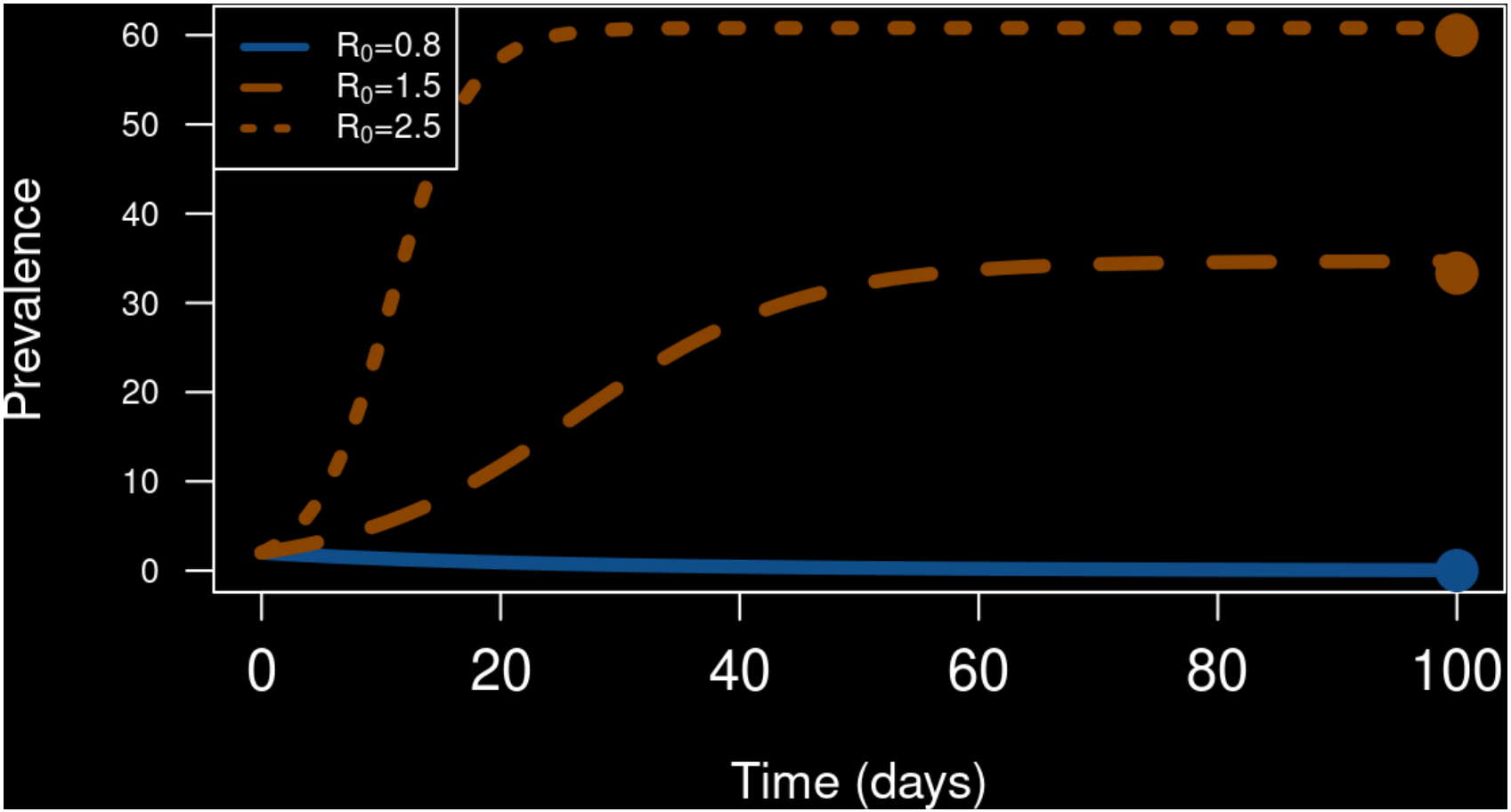
```r
library(deSolve)
rhs_SIS <- function(t, x, p) {
  with(as.list(c(x, p)), {
    N = S + I
    dS = d*(N-S) + gamma * I - d * S - beta * S * I / N
    dI = beta * S * I / N - (d + gamma) * I
    return(list(c(dS, dI)))
  })
}
# "Known" parametres
d = 1/(80 * 365.25)
gamma = 1/14
# Set beta s.t. R_0 = 1.5
R_0 = 1.5
beta = R_0 * (d + gamma)
params = list(d = d, gamma = gamma, beta = beta)
IC = c(S = 1000, I = 1)
times = seq(0, 365, 1)
# Call the numerical integrator
sol <- ode(y = IC, times = times, func = rhs_SIS,
           parms = params, method = "ode45")
```


Prévalence
0
50
100
150
200
250
300
0
100
200
300
Temps (jours)

I just did ...

What I advise not to do: illustrate a mathematical result without adding anything to the result itself

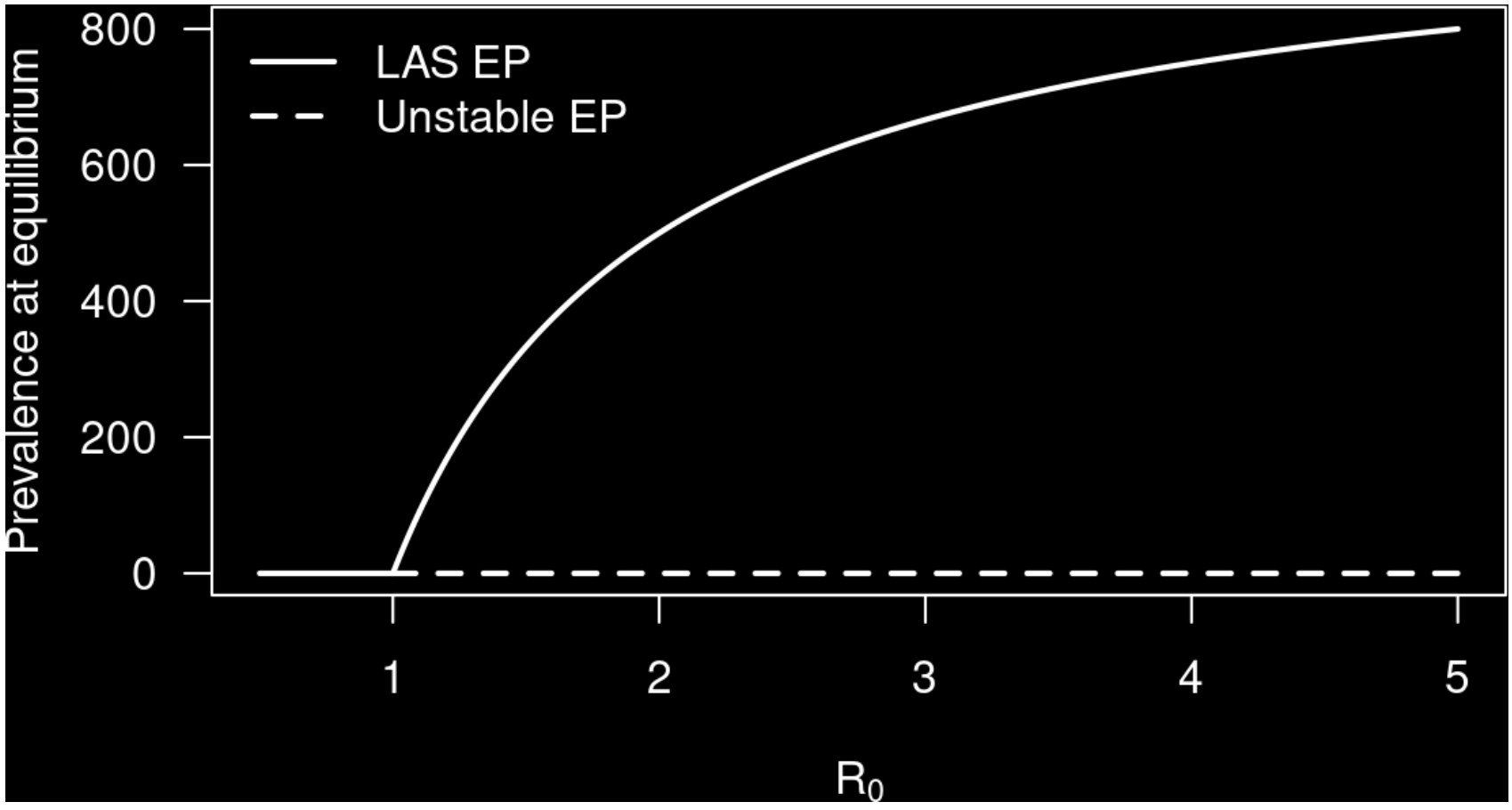
Let us make things a bit better. See the code


R_0=0.8
R_0=1.5
R_0=2.5
Prevalence
0
10
20
30
40
50
60
0
20
40
60
80
100
Time (days)

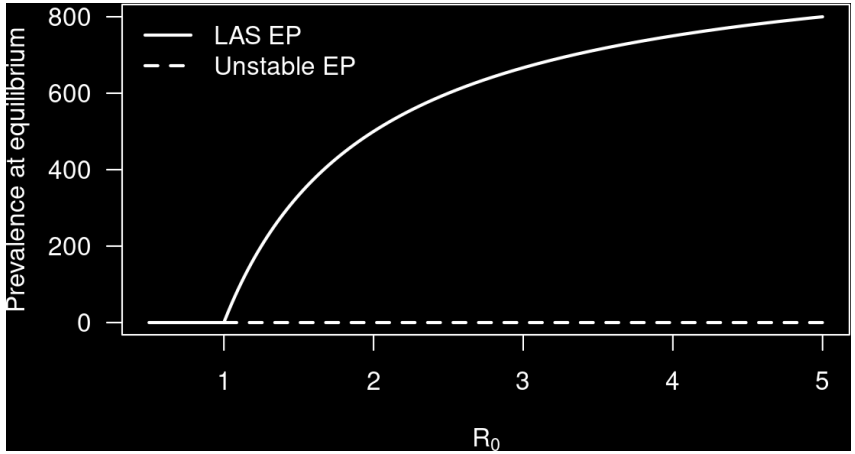
We could continue, but with a model this simple, there is little more to do: the 3 parameters of the system are combined within $\mathcal{R}_0$ and the latter summarises the dynamics well

We are going to show something important: the bifurcation diagram

We saw that when $\mathcal{R}_0 < 1$, $I \to 0$, whereas when $\mathcal{R}_0 > 1$, $I \to (1 - 1/\mathcal{R}_0)N$. Let us represent this (code)


800
600
400
200
0
Prevalence at equilibrium
LAS EP
Unstable EP
1
2
3
4
5
R_0

I "cheated" a bit...



In the second analysis, all we showed is LAS/instability of the DFE

In the present case, Analysis V1.0 tells us everything (Theorem **??**)

But in a model without explicit solutions, we may want to think about global stability properties

What remains to be done (2nd analysis)

- LAS of EEP when it is biologically relevant

- GAS of DFE when $\mathcal{R}_0 \leq 1$

- GAS of EEP when $\mathcal{R}_0 > 1$

Biologically relevant EP

Remember we had found

$$E_\star : (S, I) = \left(\frac{1}{\mathcal{R}_0} N, \left(1 - \frac{1}{\mathcal{R}_0} \right) N \right) \tag{20}$$

Here, the value of $I_\star$ only has meaning as the EP of a compartmental model if $I_\star \geq 0$, i.e., if $\mathcal{R}_0 \geq 1$

In this case, we say the EEP is **biologically relevant**

Otherwise the EEP is **not biologically relevant** and by abuse of language, we often say it **does not exist**

LAS of EEP when biologically relevant

Additionally to the EEP

$$E_\star : (S, I) = \left(\frac{1}{\mathcal{R}_0} N, \left(1 - \frac{1}{\mathcal{R}_0} \right) N \right) \tag{20}$$

we had also computed the Jacobian matrix (21) at an arbitrary point

$$J_{(S,I)} = \begin{pmatrix} -d - \beta \frac{I}{N} & \gamma - \beta \frac{S}{N} \\ \beta \frac{I}{N} & \beta \frac{S}{N} - (d + \gamma) \end{pmatrix} \tag{21}$$

Evaluate J at $E_\star$:

$$J_{E_\star} = \begin{pmatrix} -d - \beta\frac{\left(1-\frac{1}{\mathcal{R}_0}\right)N}{N} & \gamma - \beta\frac{\frac{1}{\mathcal{R}_0}N}{N} \\ \beta\frac{\left(1-\frac{1}{\mathcal{R}_0}\right)N}{N} & \beta\frac{\frac{1}{\mathcal{R}_0}N}{N} - (d+\gamma) \end{pmatrix}$$

$$= \begin{pmatrix} -d - \beta\left(1-\frac{1}{\mathcal{R}_0}\right) & \gamma - \frac{\beta}{\mathcal{R}_0} \\ \beta\left(1-\frac{1}{\mathcal{R}_0}\right) & \frac{\beta}{\mathcal{R}_0} - (d+\gamma) \end{pmatrix}$$

Find eigenvalues $-d < 0$ and $2\beta/\mathcal{R}_0 - (\beta + d + \gamma)$, i.e., since $\mathcal{R}_0 = \beta/(d+\gamma)$, $d + \gamma - \beta$

The second eigenvalue is therefore < 0 when $\mathcal{R}_0 > 1$, so $E_\star$ is LAS when it is biologically relevant

GAS of the DFE when $\mathcal{R}_0 < 1$

We use a very simple trick: take $L(S, I) = I$ as a Lyapunov function. From invariance of $\mathbb{R}^2_+$ under the flow of (**??**), $L(S, I) > 0$ unless $I = 0$. We have

$$I' = \left(\beta \frac{S}{N} - (d + \gamma)\right) I$$

Note that $S/N \leq 1$ and as a consequence

$$I' \leq (\beta - (d + \gamma))\, I$$

If $I(0) > 0$ and $\mathcal{R}_0 < 1$, we therefore have, for all $t > 0$,

$$I' < 0$$

and $L(S, I) = I$ is a Lyapunov function for system (**??**) $\implies$ DFE is GAS when $\mathcal{R}_0 < 1$

Case $\mathcal{R}_0 = 1$

When $\mathcal{R}_0 = 1$, reasoning as previously, we only have

$$I' \leq 0$$

Note, however, that if $I(0) > 0$, then $I(t) > 0$ for all t and as a consequence, we actually have $S/N < 1$, whence

$$I' < (\beta - (d + \gamma))I = 0 \text{ when } \mathcal{R}_0 = 1$$

so $L(S, I) = I$ is still a Lyapunov function

An SIR model with vaccination

Take SIR model and assume the following

- Vaccination takes susceptible individuals and moves them directly into the recovered compartment, without them ever becoming infected/infectious

- Birth = death

- A fraction p is vaccinated at birth

- $f(S, I, N) = \beta SI$

$$S' = (1-p)dN - dS - \beta SI \tag{23a}$$
$$I' = \beta SI - (d+\gamma)I \tag{23b}$$
$$R' = pdN + \gamma I - dR \tag{23c}$$

Computation of $\mathcal{R}_0$

- DFE, SIR:

$$E_0 := (S, I, R) = (N, 0, 0)$$

- DFE, SIR with vaccination

$$E_0^v := (S, I, R) = ((1-p)N, 0, pN)$$

Thus, - In SIR case

$$\mathcal{R}_0 = \frac{\beta N}{d+\gamma}$$

- In SIR with vaccination case, denote $\mathcal{R}_0^{\mathrm{v}}$ and

$$\mathcal{R}_0^{\mathrm{v}} = (1-p)\mathcal{R}_0$$

Herd immunity

Therefore - $\mathcal{R}_0^{\mathrm{v}} < \mathcal{R}_0$ if $p > 0$ - To control the disease, $\mathcal{R}_0^{\mathrm{v}}$ must take a value less than 1, i.e.,

$$\mathcal{R}_0^{\mathrm{v}} < 1 \iff p > 1 - \frac{1}{\mathcal{R}_0} \tag{24}$$

By vaccinating a fraction $p > 1 - 1/\mathcal{R}_0$ of newborns, we thus are in a situation where the disease is eventually eradicated

This is **herd immunity**

To normalise or not to normalise?

- In the SIS of Lecture 05 and here, since the total population is constant, it is possible to normalise to $N = 1$

- This can greatly simplify some computations

- However, I am not a big fan: it is important to always have the "sizes" of objects in mind

- If you do normalise, at least for a paper destined to mathematical biology, always do a "return to biology", i.e., interpret your results in a biological light, which often implies to return to original values

Where we are

- An *endemic* SIS model in which the threshold $\mathcal{R}_0 = 1$ is such that, when $\mathcal{R}_0 < 1$, the disease goes extinct, whereas when $\mathcal{R}_0 > 1$, the disease becomes established in the population

- An *epidemic* SIR model (the KMK SIR) in which the presence or absence of an epidemic wave is characterised by the value of $\mathcal{R}_0$

- The SIS and the KMK SIR have explicit solutions (in some sense). **This is an exception!**