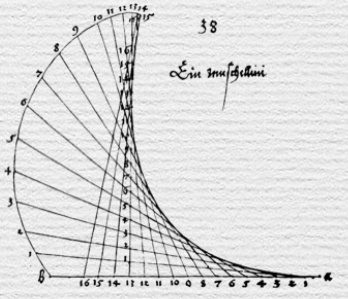
Modelling Infectious Disease

Gabriel Wittum
AMCS, CEMSE, KAUST

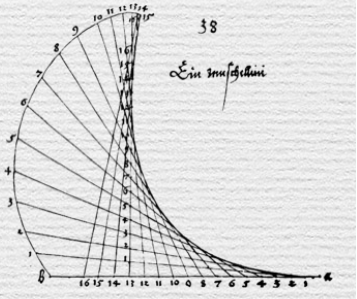
for internal use only



Modelling Infectious Disease

The simplest way to model the spreading of infectious diseases is to set up a spatially concentrated problem, i.e. drop the spatial dependencies and just describe the dynamical behaviour by a system of ordinary differential equations (ODEs) in time.

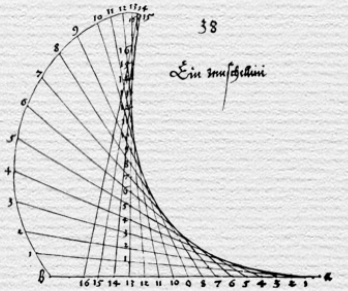
Gabriel Wittum
AMCS, CEMSE, KAUST



Modelling Infectious Diseases

The basic behaviour of an infection like the SARS-CoV-2, which is currently spreading, is as follows. The whole population of the region where the disease is emerging is divided into five groups.

Gabriel Wittum
AMCS, CEMSE, KAUST

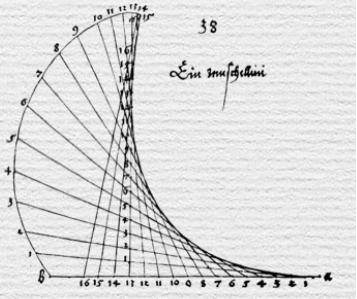


Modelling Infectious Diseases

1. The healthy people. Their number is denoted by $g(t)$.

2. People, who are infected, but do not show signs of sickness. They are contagious and dangerous, since they move around like all others. Their number is denoted by $a(t)$.

Gabriel Wittum
AMCS, CEMSE, KAUST



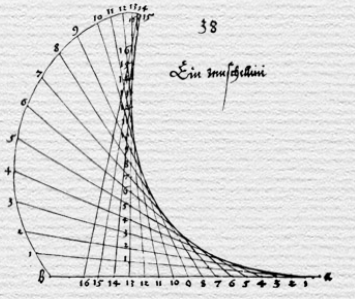
Modelling Infectious Diseases

3. People, who are sick. They show symptoms and will mostly stay at home, in bed or even in hospital. So they will not infect of the healthy ones. Their number is denoted by $k(t)$.

4. People, who have recovered. They have been sick and successfully overcame the infection. They are now immune against the disease. Their number is described by $w(t)$

5. The dead people, $v(t)$.

Gabriel Wittum
AMCS, CEMSE, KAUST



Modelling Infectious Diseases

Initially, all people are healthy, except one guy, who is already infected, i.e., the disease exists already. We call the total population number

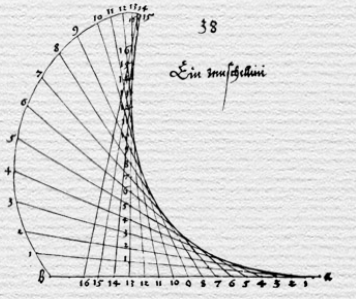
$$N = g_0 + 1$$

with $g_0 = g(t=0)$. One requirement for our model is

$$N = g(t) + a(t) + k(t) + w(t) + v(t) = \text{const}$$

for all $t \in [0, T]$, denoting the time interval under consideration.

Gabriel Wittum
AMCS, CEMSE, KAUST



Modelling Infectious Diseases

The initial values at the very beginning of the spreading of the infection is:

$$g(0) = N\text{-}1$$

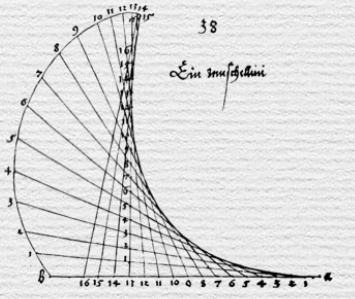
$$a(0) = 1$$

$$k(0) = 0$$

$$w(0) = 0$$

$$v(0) = 0$$

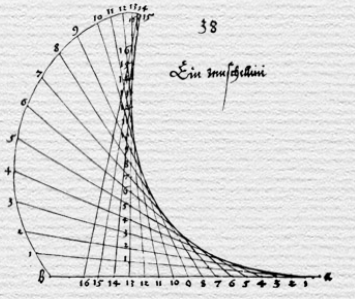
Gabriel Wittum
AMCS, CEMSE, KAUST



Modelling Infectious Diseases

Due to the infection, people change from group to other groups. Healthy people will be infected by virtue of the already infected ones with a given rate of infection, describing the speed with which the infection spreads. This rate of infection can be converted into the number of healthy people one infected person infects during staying in the group of infected people.

Gabriel Wittum
AMCS, CEMSE, KAUST



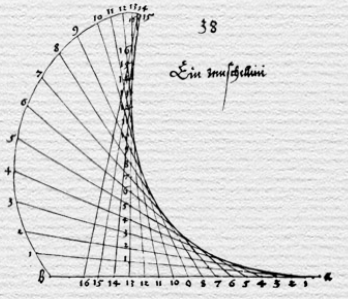
Modelling Infectious Diseases

Consequently, we model the change of the number of healthy people by

$$\frac{dg(t)}{dt} = -\alpha g(t) \cdot a(t)$$

where $\alpha > 0$ is the rate of infection. This leads to an increase of $a(t)$ by the same rate, but with the opposite sign.

Gabriel Wittum
AMCS, CEMSE, KAUST



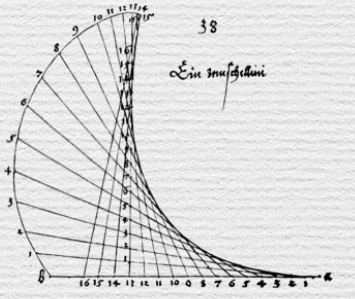
Modelling Infectious Diseases

This leads to an increase of $a(t)$ by the same rate, but with the opposite sign.

$$\frac{da(t)}{dt} = \alpha g(t) \cdot a(t) + \ldots$$

After incubation time, however, a certain number of infected people will show symptoms of the illness, i.e. they transit into the number of sick people.

Gabriel Wittum
AMCS, CEMSE, KAUST



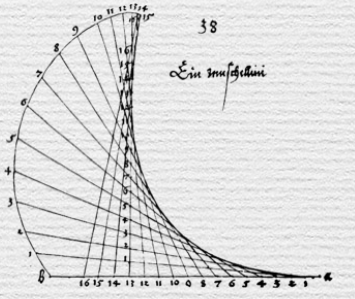
Modelling Infectious Diseases

This is described by

$$\frac{da(t)}{dt} = \alpha g(t) \cdot a(t) - \frac{\kappa}{q} a(t) - \ldots$$

Here, κ means the rate of transition between infection and sickness and q is the incubation time. This is the sole source for the number of sick people $k(t)$.

Gabriel Wittum
AMCS, CEMSE, KAUST

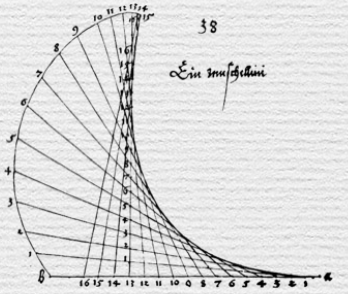


Modelling Infectious Diseases

Not all infected people will become sick, some of them carry the virus, but their immune system is able to keep it at bay without falling into sickness. The complete equation for the change of the number of infected people $a(t)$ is then

$$\frac{da(t)}{dt} = \alpha g(t)a(t) - \frac{\kappa}{q}a(t) - \frac{(1-\kappa)}{q}a(t) = \alpha g(t)a(t) - \frac{1}{q}a(t)$$

Here, $\kappa \in [0,1]$ is the rate of transition from infection to sickness, q the incubation time and

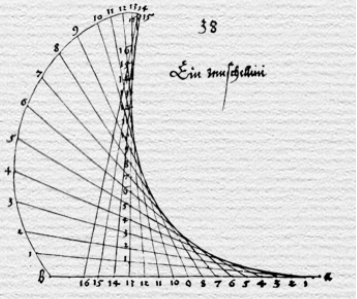


Gabriel Wittum
AMCS, CEMSE, KAUST

Modelling Infectious Diseases

How long persons stay in the infected group depends on the incubation time q. This is the latency of the infection. The larger q is, the more dangerous is the spread, since it happens without outer signs. This allows the infection to spread like an avalanche. Likewise, the illness has a certain duration p. Both durations show up in the equations as a factor $1/q$ or $1/p$ resp. The mean time a person is infected, but not ill, is q.

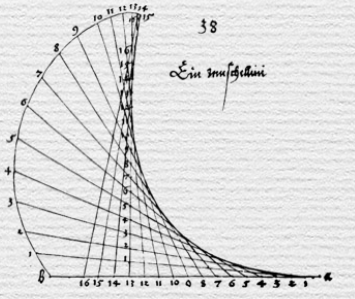
Gabriel Wittum
AMCS, CEMSE, KAUST



Modelling Infectious Diseases

The equation for the number of sick people, $k(t)$, can be derived now. The source are the infected people transiting into sickness. This group has two sinks. People are recovering and now being immune to the illness, another part of the sick people may die. Death occurs with a rate of θ. Those, who are not dying recover.

Gabriel Wittum
AMCS, CEMSE, KAUST



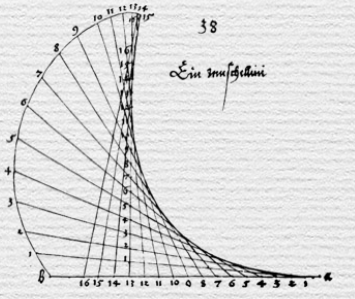
Modelling Infectious Diseases

The equation for $k(t)$ then reads

$$\frac{dk(t)}{dt} = \frac{\kappa}{q}a(t) - \frac{1-\theta}{p}k(t) - \frac{\theta}{p}k(t) = \frac{\kappa}{q}a(t) - \frac{1}{p}k(t)$$

The mean time a person stays sick is ρ.

Gabriel Wittum
AMCS, CEMSE, KAUST



Modelling Infectious Diseases

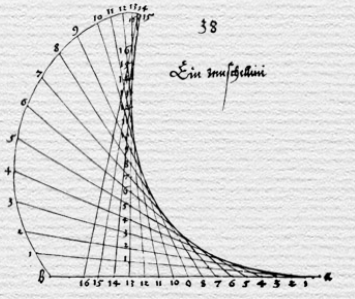
One source for $w(t)$ is the group of infected people $a(t)$ who do not get sick, but transit into recovered state directly

$$\frac{dw(t)}{dt} = \frac{(1-\kappa)}{q} a(t) + \ldots$$

including the incubation time q. The other source of $w(t)$ is sick people recovering.

$$\frac{dw(t)}{dt} = \frac{1-\kappa}{q} a(t) + \frac{1-\theta}{p} k(t)$$

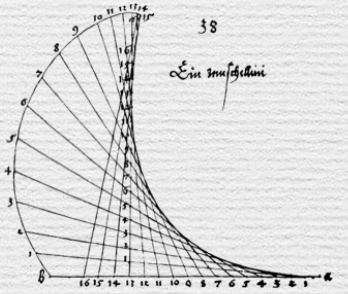
Gabriel Wittum
AMCS, CEMSE, KAUST



Modelling Infectious Disesases

The number of dead people is given by

$$\frac{dv(t)}{dt} = \frac{\theta}{p} k(t)$$



Gabriel Wittum
AMCS, CEMSE, KAUST

Modelling Infectious Disesases

In total, we have the following eqations

$$\frac{dg(t)}{dt} = -\alpha g(t) \cdot a(t)$$

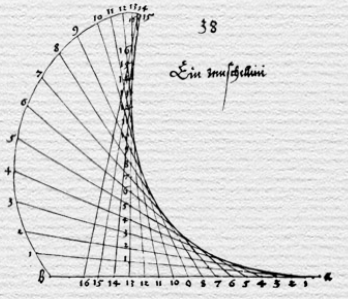
$$\frac{da(t)}{dt} = \alpha g(t) a(t) - \frac{1}{q} a(t)$$

$$\frac{dk(t)}{dt} = \frac{\kappa}{q} a(t) - \frac{1}{p} k(t)$$

$$\frac{dw(t)}{dt} = \frac{1-\kappa}{q} a(t) + \frac{1-\theta}{p} k(t)$$

$$\frac{dv(t)}{dt} = \frac{\theta}{p} k(t)$$

Gabriel Wittum
AMCS, CEMSE, KAUST



Modelling Infectious Disesases

With the following parameters:

infection rate $\alpha = \sigma/N > 0$,

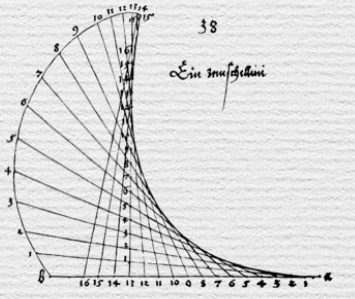
sickness rate $\kappa \in [0,1]$,

death rate $\theta > 0$,

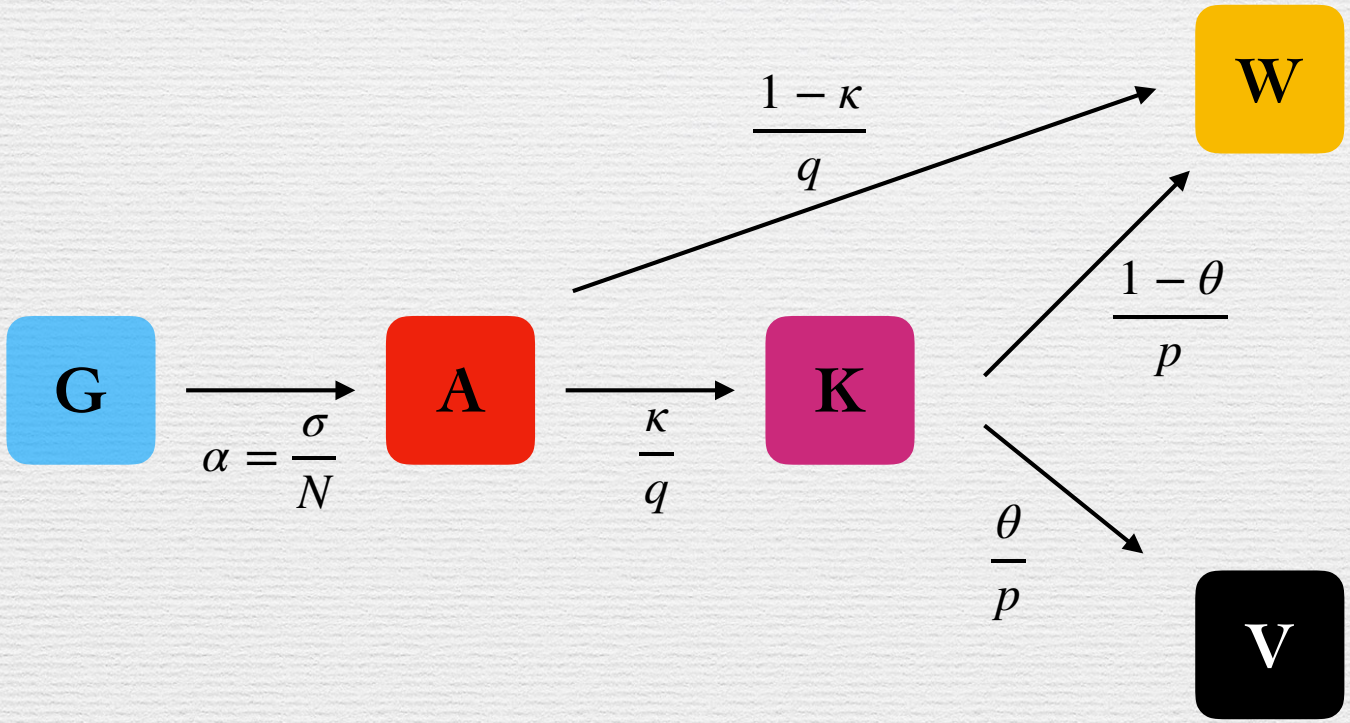
incubation time q,

duration of sickness ρ.

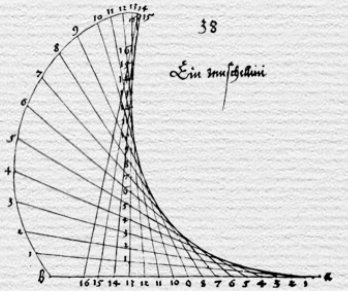
Gabriel Wittum
AMCS, CEMSE, KAUST



The Model Structure



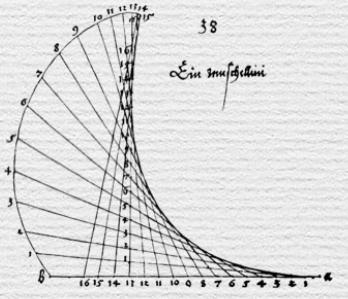
Gabriel Wittum
AMCS, CEMSE, KAUST



Modelling Infectious Diseases

Due to the infection, people change from one group to other groups. Healthy people will be infected by virtue of the already infected ones with a given rate of infection σ, which is usually measured by number of people infected per person. This counts over the whole time the infecting person belongs to the corr. group. That means, if every infected person infects one other person, this rate is 1. In this case, the disease will be endemic. The number of infected people will remain constant.

Gabriel Wittum
AMCS, CEMSE, KAUST



Numerics of ODE

In general, we have

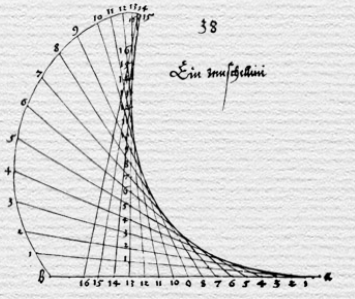
$$u'(t) = f(t,u) \text{ for } t > 0$$

and the initial value

$$u(0) = u_0$$

with, u: $\mathbb{R}^m \rightarrow \mathbb{R}^m$. *This problem is called initial value problem (IVP)*

Gabriel Wittum
AMCS, CEMSE, KAUST



The Explicit Euler Method

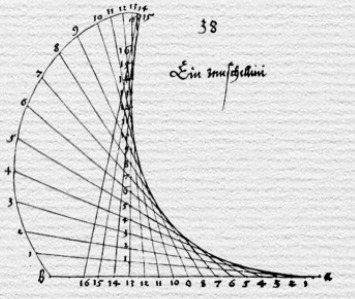
To solve this IVP we use the simplest method available, the explicit Euler method. Replace the continuous interval [0,T] by a grid

$$0 = t_0 < t_1 < t_2 < \ldots < t_1 = T$$

and the differential quotient

$$\frac{du(t)}{dt} = \lim_{\tau \to 0} \frac{u(t+\tau) - u(t)}{\tau} = \frac{u(t+\tau) - u(t)}{\tau} + O(\tau)$$

Gabriel Wittum
AMCS, CEMSE, KAUST



The Explicit Euler Method

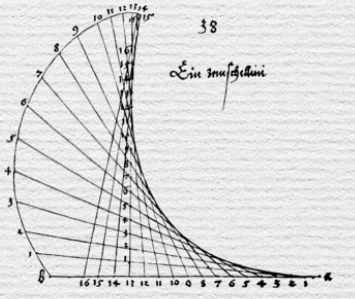
Dropping the error term $O(\tau)$ and using the IVP, we obtain

$$\frac{u(t+\tau)-u(t)}{\tau} = f(t,u)$$

resolving that for $u(t+\tau)$ yields

$$u(t+\tau) = u(t) + \tau f(t,u)$$

Gabriel Wittum
AMCS, CEMSE, KAUST



The Explicit Euler Method

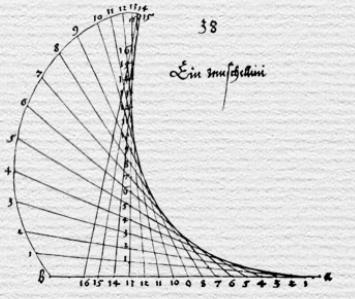
In the sense of a time-stepping iteration we write

$$u(t_{i+1}) = u(t_i) + \tau f(t_i, u(t_i)) \,.$$

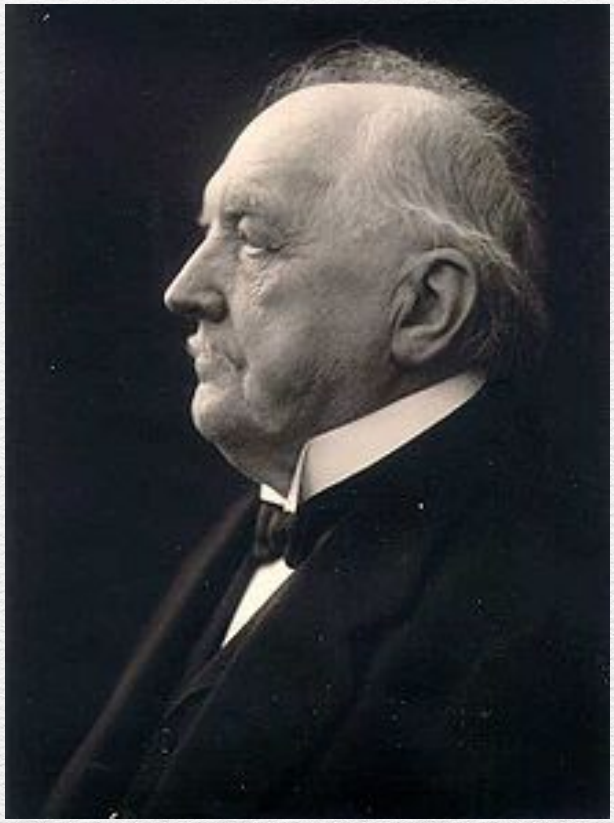
The whole iteration starts with the initial value $u(t_0) = u(0) = u_0$.

The numerical error is of the order of the omitted term $O(\tau)$.

Gabriel Wittum
AMCS, CEMSE, KAUST



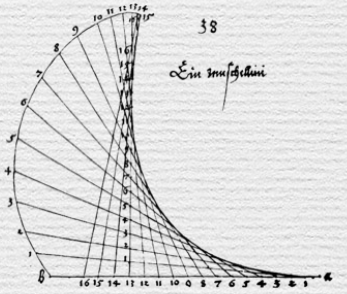
4th order Runge-Kutta Method



Carl Runge (1856-1927) and Wilhelm Kutta (1867-1944)

Gabriel Wittum
AMCS, CEMSE, KAUST

for internal use only



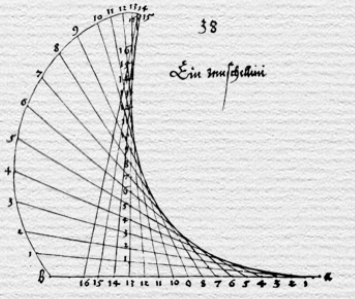
4th order Runge-Kutta Method

had the idea to improve this method by replacing $f(t_i,u(t_i))$ in this formula by a sum of four coefficients:

$$u(t_{i+1}) = u(t_i) + \tau \frac{k_1 + 2k_2 + 2k_3 + k_4}{6}$$

The coefficients k_i are chosen by comparing with coefficients of a Taylor series up to order four.

Gabriel Wittum
AMCS, CEMSE, KAUST



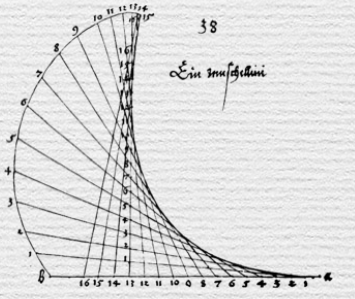
4th order Runge-Kutta Method

This leads to

$$k_1 = f(t_i, u(t_i)), k_2 = f(t_i + \frac{\tau}{2}, u(t_i) + k_1\frac{\tau}{2}),$$
$$k_3 = f(t_i + \frac{\tau}{2}, u(t_i) + k_2\frac{\tau}{2}), k_4 = f(t_i + \tau, u(t_i) + k_3\tau)\,.$$

The method is explicit and of fourth order, i.e. the error is $O(\tau^4)$.

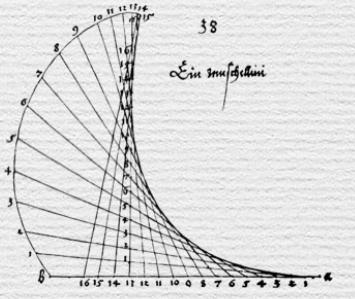
Gabriel Wittum
AMCS, CEMSE, KAUST



How to Get the Parameters?

A lot of data about the number of infected people, the number of deaths, and the number of recovered people can be found in the internet, in particular on github.com. Some parameters are mentioned quite often in the press and the media, in particular the *basic reproduction number* R_0. Sometimes also the *manifestation rate* shows up. What is this and what does it have to do with our parameters?

Gabriel Wittum
AMCS, CEMSE, KAUST

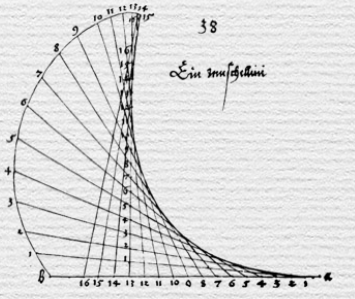


How to Get the Parameters?

We already looked at a problem of parameter identification with a partial differential equation. This was a constraint quadratic optimisation problem.

Systems of ODE like here are usually handled by solving an optimisation problem, trying to find the parameters, such that the solution of the ODEs is as close as possible to a measured one.

Gabriel Wittum
AMCS, CEMSE, KAUST



Parameter Identification

Parameter identification consists of solving the least-squares minimization problem

$$f(p) := \frac{1}{m}\sum_{j=1}^{m} \|D_j^{-1}(x(t_j,p) - z_j)\|_2^2 \to \min_p$$

with diagonal weighting

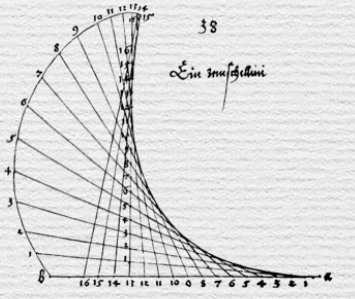
$$D_j := \mathrm{diag}\{(\delta z_j)_1,\ldots,(\delta z_j)_n\} \in \mathcal{M}_n(\mathbb{R});\ j = 1,\ldots,m.$$

That means, we want to minimize the relative deviation of model and data at the measurement time points t_j. Again in short-hand notation, the minimisation problem can be written as

$$f(p) = F(p)^T \cdot F(p) \to \min_p,$$

where $F(p) := (F_1(p),\ldots,F_m(p))$ is a vector of length $N = m{\cdot}n$ with entries defined by

Gabriel Wittum
AMCS, CEMSE, KAUST



Parameter Identification

$$F(p) = \begin{pmatrix} D_1^{-1} \cdot (x(t_1, p) - z_1) \\ \vdots \\ D_m^{-1} \cdot (x(t_m, p) - z_m) \end{pmatrix}.$$

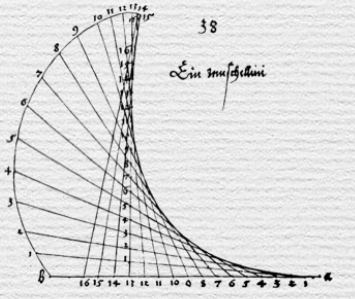
and $F : \mathbb{R}^q \to \mathbb{R}^N$ is a non-linear mapping and structured as a stacked vector. The above problem, which is highly nonlinear in p, can be solved by a Gauss-Newton iteration, as introduced in Chapter 6 "Least squares". There, each iteration step k requires the solution of a linear least-squares problem,

$$\|J(p^k) \cdot \Delta p^k + F(p^k)\|_2 \to \min_{p^k},$$

$$p^{k+1} = p^k + \Delta p^k$$

where $J(p^k) = F'(p^k) \in \mathbb{R}^{N \times q}$ denotes the Jacobian matrix of F.

Gabriel Wittum
AMCS, CEMSE, KAUST



Parameter Identification

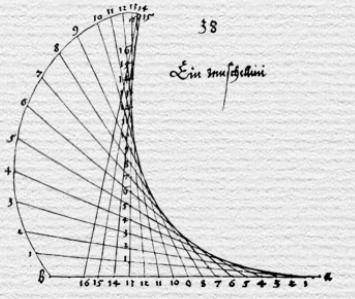
The Jacobian is is also called *sensitivity matrix*, since its elements

$$J(p^k) = (s_{ij})_{i,j=1}^{N,q} = \frac{\partial x_i(t)}{\partial p_j}$$

are the *sensitivities* computed in all time-points. These entries represent the sensitivity of the solution x with respect to the parameters p at the time points of measurements. In our model, values of the parameters and population classes can vary over orders of magnitude. To achieve comparability, the sensitivity values have to be normalised by the absolute values of species and parameters to obtain scaled sensitivities

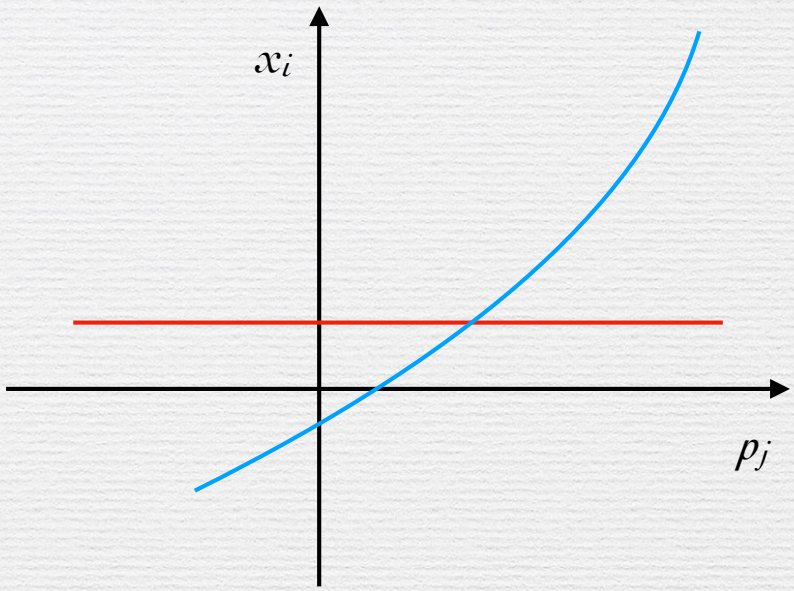
$$\overline{s_{ij}(t)} = \frac{\partial x_i(t)}{\partial p_j} \frac{|p_{\text{scal}}|}{|x_{\text{scal}}|},$$

Gabriel Wittum
AMCS, CEMSE, KAST



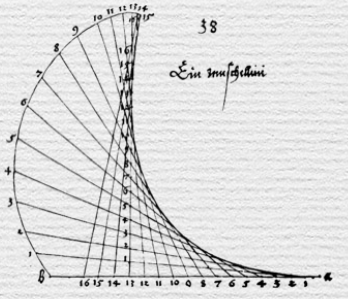
Sensitivities

$$J(p^k) = (s_{ij})_{i,j=1}^{N,q} = \frac{\partial x_i(t)}{\partial p_j}$$



An analysis of the scaled matrix $J(p)$ gives some hints whether the current combination of model and data will allow an identification of a given parameter. Parameters with **very small sensitivity** have nearly no influence on the solution at the measurement time points and therefore cannot be estimated. In this case the entries of the corresponding column in $J(p)$ are almost zero. Furthermore, some of the parameters might be linearly dependent, which leads to nearly identical columns in $J(p)$. In both cases the matrix $J(p)$ will be (nearly) singular. In order to reveal such properties, the linear least squares problem is solved by QR factorization with column pivoting.

Gabriel Wittum
AMCS, CEMSE, KAUST



Gauss-Newton Method

- Quasi-Newton method with line search

$$\mathbf{p}_k^{GN} = -\left(\mathbf{J}(\mathbf{x}_k)^T\mathbf{J}(\mathbf{x}_k)\right)^{-1}\mathbf{J}(\mathbf{x}_k)^T\mathbf{r}(\mathbf{x}_k)$$

  is a descent direction if: $\mathbf{J}(\mathbf{x}_k)^T\mathbf{J}(\mathbf{x}_k) \succ 0$

- Then do exact or inexact line search (remember Wolfe conditions) on

$$f(\mathbf{x}_k + \alpha\mathbf{p}_k^{GN})$$

Gabriel Wittum
AMCS, CEMSE, KAUST