SOCIEDADE BRASILEIRA DE MATEMÁTICA

ENSAIOS MATEMÁTICOS
2023, Volume **38**, 343–358
https://www.doi.org/10.21711/217504322023/em3813

A two level contagion process and its deterministic McKendrick limit with relevance for the Covid epidemic

Stephan Luckhaus and Angela Stevens

Abstract. We introduce a stochastic epidemiological model, where two infection scenarios alternate. The first is infection within separate groups of finite size, the second is infection at meeting places of finite capacity, where individuals meet randomly. This can be thought of as an epidemic, where e.g. members of households regularly use public transport. For this model we derive the hydrodynamic limit: a McKendrick system with polynomial infections force.

Keywords. Jump processes on general state spaces, Interacting particle systems in time-dependent statistical mechanics, Epidemiology.

2020 Mathematics Subject Classification: 60J76, 82C22, 92D30.
Submitted on November 3, 2022, final version accepted on April 28, 2023.

1 Introduction

Dedicated to Errico Presutti on the occasion of his 80^{th} birthday; an always generous friend, mathematically and otherwise.

In this paper we introduce a stochastic process of 'household' type, with a two level infection scenario, which leads to an age-structured SIR model of the type introduced by A.G. McKendrick in 1926 [11]. In mathematical terms the large N limit of the many body process is an independent evolution of the 'household' driven by a deterministic equation. This is similar to [4]. In our case the deterministic equation is a kinetic equation instead of an ordinary differential equation.

The special feature of our model is that the mixing process are contacts in randomly chosen groups of finite size m_2. That leads to a Kermack-McKendrick type of equation [5] in a state space of dimension m_1, the size of the households, with a nonlinearity of degree m_2, the size of the meeting places. This state space is the space of the times a_i elapsed since the time of infection of the individual of type i in the household. The rate of infection for the individuals is no longer linear in the percentage of infections in the population, but has an independent term for the infections in the household and a nonlinear term, a polynomial of degree $m_2 - 1$ in the percentage of the infectious individuals for the random contact. Its coefficients can be calculated from a dose response curve, linking the rate of infection to the viral load at the contact.

The type of mixing we are looking at here is special, adapted e.g. to a metro filling up during the rush hour. Other mixing processes could be considered too, and also intermediate scalings for the size of the meeting places. This article presents the stochastic process and deals with the problem of determining its 'hydrodynamic limit'. Other aspects, e.g. the long time behavior of the limiting McKendrick system, are not discussed here. Analytic aspects of McKendrick systems in multidimensional state spaces were discussed in [9] and [8]. The model presented in [8] and the model presented here share two features with relevance for the Covid epidemics, namely:

• Separating the effect of infections in the household and the one of the random contacts, it becomes possible to explain why lockdowns have very limited effects. They certainly decrease random contacts, but increase the ones in the households.

We have data from a study of infection in public transport conducted in the Frankfurt region at the beginning of the epidemic wave caused by the UK-variant of the virus in Germany. These data clearly show that the vast majority of infections were reinfections. So in spite of the lockdown, the majority of the active population in that region had been infected (asymptomatically).

• The other effect, that a Kermack-McKendrick model with infection rate depending linearly on the relative percentage of infected in the population cannot explain, is what we call deterministic metastability. A small number of imported infections in the population does not start an epidemic.

For this compare the situation in march and april 2020 in the two swiss cantons of Zurich and Geneva. Zurich had a minimal number of Covid cases, whereas Geneva had 1% of the population according to data of BAG, [1]. Taking into account asymptomatic cases in Geneva, more than 10% of the active population was infected in spring 2020 [13]. So a population may be in the state of 'herd immunity' with respect to a level of import of infections of the amount of 10^{-4} of the population per week, but not if the level is 10^{-3} of the population per week, one might roughly estimate.

Our claim is that modeling the Covid epidemics, the models have to include the possibility to separate the effects of deterministic or reoccuring contacts and those of random contacts. Further, the rate of infection for an individual, sometimes called the force of infection, should not depend in a linear but in a nonlinear way on the percentage of infectious individuals in the population. For a more a detailed discussion see [8]

2 Discussion on data from a study of the infection risk in public transport in Germany

We include this discussion of the data of the study here, since it is not a publication in the classical sense, but a link to a web page of the ministry of transport of the state (Land) of Baden-Württemberg, [3]. And it is crucial for our claim of the ineffectiveness of the lockdown in Germany in winter 2022/21. The authors of the study were from CRO (Charité research organization). The relevant data for us are the following:

At the beginning of the study in calendar week 7, 2021 ca. 700 people were enrolled, unvaccinated and not having been previously diagnosed with Covid. Initially, 3 of them had a positive PCR-test, one among these 3 individuals also a positive specific IgG test (Euroimmun). 35 more individuals had only a positive IgG test.

Among the participants who had tested negative at the beginning of the study, ca. 660 completed the study (in calendar week 12 of 2021). During the study questionnaires about symptoms were regularly sent out. And in this way 3 individuals among the 660 were found who tested PCR

positive. At the end of the study everyone else among the 660 was tested again by PCR and IgG test. Only one more (presymptomatic) case was found by the PCR tests, but 23 (newly) positive cases were found be IgG test.

Already the fact that one discovers an asymptomatic carrier with a positive specific IgG antibody response and a positive PCR test is a clear hint that this was a reinfection of an individual, who had already been infected with the wild type previously and consequently had already specific IgG antibodies or at least B memory cells. A reinfection with the wild type can practically be excluded [2]. Most probably it was the UK type which became dominant in Germany around that time. Secondly, that by the end of the study so many newly positive IgG tests were found, but all in all only 4 people with positive PCR tests, proves that PCR visibility on an average had decreased to a very short time.

We can compare these data, an infection rate of 4% over 5 weeks, with the official data of the RKI (Robert Koch Institute) on the number of recorded Covid cases per week and population in the Frankfurt region, where the study was conducted. The so-called 7 day incidence was increasing during the study, and it came to an overall infection rate of $0,2\%-0,3\%$ of the population for the 5 weeks covered by the study. It was $0,4\%-0,5\%$ for the 5 weeks before. So firstly we can safely conclude that in order to determine the complete number of infections, the RKI numbers have to be multiplied with a factor of 15-20 (not every asymptomatic carrier develops a specific IgG response).

One can also conclude that the positive IgG tests corresponds to not more than to the number of infections 5-6 weeks before. And one can conclude that among the 23 seropositive (IgG positive) individuals in the study, 10 or more had been infected in the last week before the final test. So probably the average PCR visibility was one to two days, due to the effectiveness of the antibody response developed by the previous infection with the wild type.

For us this is important, because we obtain two independent confirmations from this study, that the vast majority of the active population in the Frankfurt region has been infected and immunized by calendar week 7 in 2021; but of course immunized against the wild type not against the UK type of the virus.

If we assume that the seroconversion rate among the infected is about 80%, then this means for a bit less than 90% of the infected, that they were infectious themselves for not more than 2 days and had no chance to develop symptoms. But for the rest their previous infection did not make any difference.

3 A household type contagion process, modeling Covid-19

The process we consider in this paper has two stages, the first one modeling in group infection, the second one models aerosol infection in public transport, supermarkets, University lectures, etc. Specifically we have M_1 groups of m_1 distinguishable individuals of different types. $N = M_1 m_1$ is the populations size. Moreover we have meeting places of size m_2, for simplicity a fixed multiple of m_1, numbered $l = 1, \dots, M_2$, where $M_2 m_2 = M_1 m_1$.

The process consists of independent evolutions in the time intervals $(t_{2\nu}, t_{2\nu+1})$ and $(t_{2\nu+1}, t_{2\nu+2})$. The first is an infection process within the groups, the second an infection process at the meeting places. At time $t_{2\nu+1}$, consecutively and with equal probability individuals are assigned to the meeting places. This is the only mixing in the process. The meeting place l assigned to individual n at time $t_{2\nu+1}$ will be denoted by $l(\nu, n)$.

Each individual can get infected only once, at time t_n. Here t_n is exponentially distributed with rates $s_i X$, and s_i, $i = 1, \dots, m_1$ are constants representing the different susceptibilities of the different types. X, the force of infection, depends nonlinearly on the amount of virus spread by the infected individuals at the meeting places and in the groups respectively. If α_i denotes this amount as a function of time for an individual of type i, and t_n the time of infection, we assume an i.i.d. distribution P_i for the function $\alpha_i(\cdot + t_{i+m_1 r})$, $r = 0, \dots, M_1 - 1$, in L^∞, with in addition a Fréchet Kolmogoroff condition:

$$E\left(\int_0^\infty |\alpha_i * \varphi_\varepsilon - \alpha_i|\right) < \omega(\varepsilon)$$

for a Dirac sequence $\varphi_\varepsilon = \varepsilon^{-1}\varphi(\frac{1}{\varepsilon}) \in C_0^\infty$, and the modulus of continuitiy ω with $\omega(0) = 0$, and ω being continuous. Moreover we assume $|\alpha_i|_\infty < C$, and $spt(\alpha_i) \subset [t_n + \delta, t_n + \bar{L}]$. So there is a delay between the time an individual gets infected and the time it gets infectious. The assumption on P_i means that the evolution of the virus production $\alpha\left((t - \bar{t})_+\right)$ in an individual infected at time $\bar{t}$, depends only on the type, and is independent of $\bar{t}$ and the environment.

With this notation the formula for X, the 'force of infection', is

$$X_r(t) = f\left(\sum_{i=1}^{m_1} \alpha_{m_1(r-1)+i}((t - t_{m_1(r-1)+i})_+)\right) \tag{3.1}$$

for $t \in (t_{2\nu}, t_{2\nu+1})$ in the groups, and

$$X_l(t) = \bar{f}\left(\sum_{l(\nu,n)=l} \alpha_n((t-t_n)_+)\right) \tag{3.2}$$

for $t \in (t_{2\nu+1}, t_{2\nu+2})$ at the meeting places.

Here f is the dose response curve for the in group infection, and $\bar{f}$ is the dose response curve for the infection at the meeting places.

The process is one with a delay bounded from below. For Covid this delay is approximately 2 days, the time between infection and the first positive PCR test. And this delay is more or less deterministic [7]. Covid also does not have reinfections (with the same virus variant and at least up to 7 months after recovery). Covid has though competing virus variants. So after a short period of crossimmunization, approx. 2 months, a new variant can spread among the recovered. The data from which one can infer this, are the Covid deaths in London compared to West Midlands in the fall/winter 2020/2021. As can be seen in Figure 1 up to calendar week 50/2020 the Covid death rates in London as opposed to West Midlands do not show any epidemic increase. But after calendar week 52 both curves look very similar. The delay between infection and demise is about 3 weeks. So that would mean that in London up to calendar week 49 approximately, the population was in a state of herd immunity. But immunization against the wild type of Covid-19 did not give any protection against the then new UK or *alpha* virus variant. Variants will be discussed in the next chapter.

Our process has two features which simplify the treatment a lot. The selection process is easily seen to produce just a permutation in $\{1, \dots, N\}$ and more precisely the Haar measure on the symmetric group $\sum_N$. And the bound on the delay from below means that the infection process defined by (3.1), respectively (3.2) in the time interval $(t, t+\delta/2)$ is independent from the infection process in the time interval $(t-\delta/2, t)$. So the correlation the process of selection produces can be estimated. First, if we have w.l.o.g. that

$$\sigma(n) = 1, \quad \sigma \in \Sigma_N,$$

the other $m_2 - 1$ individuals at the meeting place will be $\sigma^{-1}(2), \dots,$ $\sigma^{-1}(m_2)$, where σ is a random element of $\sum_{N-1}$. Each permutation $\sigma^{-1}(2), \dots, \sigma^{-1}(m_2)$ of $\{2, \dots, m_2\}$ into $\{1, \dots, N\} \setminus \{n\}$ will occur with equal probability. If one replaces that by all maps in $\{1, \dots, N\}^{m_2-1}$ that increases the cardinality by a factor $\frac{N}{N-1} \cdot \dots \cdot \frac{N}{N-m_2+1}$, so $1+O(\frac{1}{N})$. More

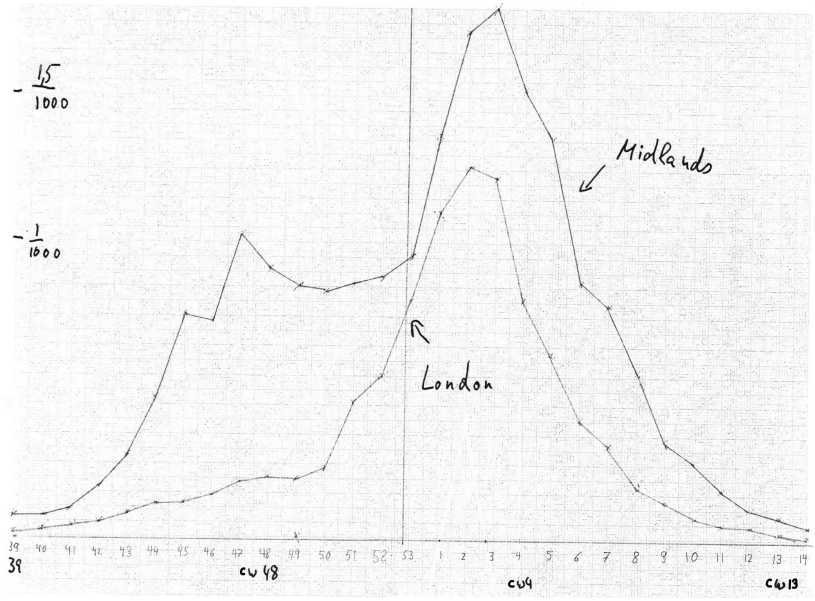

Figure 1: shows the Covid deaths per calendar week (cw) and 1000 inhabitants for the two English regions, London and West Midlands in the winter 2020/21. Up to the beginning of december 2020 the wild type was still dominant among infections. For the death toll a delay of ca. 3 weeks has to be taken into account. (Source: daily Covid deaths per region, NHS England [12].)

precisely if $\nu : \{1, \dots, m_2 - 1\} \to \{1, \dots, N\}$ denotes multiindices

$$\sum_{\nu \in \{1,\dots,N\}^{m_2-1}} \Big| P\left(\{l(\nu_1) = l(\nu_2) = \cdots = l(\nu_{m_2-1}) = l(n)\ ,\ \nu_i \neq \nu_j \neq n\}\right) - N^{1-m_2} \Big|$$

$$= \left[\frac{(N-m_2)!}{(N-1)!} - N^{1-m_2}\right] \frac{(N-1)!}{(N-m_2)!} + \left[N^{m_2-1} - \frac{(N-1)!}{(N-m_2)!}\right] N^{1-m_2}$$

$$= O\left(\frac{1}{N}\right)$$

This means that in the large N-limit the infection force at the meeting places visited by a susceptible individual under the actual mixing process can be replaced by one choosing the other $m_2 - 1$ individuals independently from all $N - 1$ individuals.

For the modified process the rate of infection for a susceptible individual $n = m_1(r-1)+i$ of type i is given by $s_i \bar{X}_{\vec{j}}(t)$, where

$$\bar{X}_{\vec{j}}(t) = \bar{f}\left(\sum_{k=1}^{m_2-1} \alpha_{j(k)}((t-t_{j(k)})_+)\right) \tag{3.3}$$

for $t \in (t_{2\nu+1}, t_{2\nu+2})$, and where $\vec{j}$ is any multiindex in N^{m_2-1}. $\bar{X}_{\vec{j}}(t)$ is the force of infection in the randomly chosen group for the individual n, up to the time when the individual n gets infected. It also defines an evolution in the group without the individual and therefore a transition probability in state space from time $t_{2\nu+1}$ to time t, where $t \in (t_{2\nu+1}, t_{2\nu+2})$.

From now on we will discuss the modified process. But let us first discuss the structure of this process (in the spirit of A.G. McKendrick, who introduced the concept of the kinetic theory in a continuous state space for the life sciences in 1914 [10]). The individual moves in a state space

$$\{1,\dots,m\} \times \mathbb{R}_+ \times \{0,1\} \times L_1(\delta, \bar{C}) =: Y \times L_1(\delta, \bar{C}),$$

where the second coordinate is $(t-t_n)_+$, the third is $\chi_{t>t_n}$, a phase parameter, and the last is the choice α_n, the virus production as a function of $t-t_n$ for the individual. The first coordinate is the number $i(n) = n \mod (m_1)$.

Finally it should be noted that the whole process is unchanged whether one assigns $\alpha_{m_1(r-1)+i}$ randomly at $t=t_0$ or at $t = t_{m_1(r-1)+i} + \delta/2$ and the probability distribution $P_i(\alpha)$ is stationary under that process, by our assumption.

To calculate the transition probability is equivalent to let the process run $\bar{N}$ times independently and let $\bar{N} \to \infty$, with a fixed initial datum.

It is convenient for the 'household' as well as for the meeting places, to partition the sets $\Omega = \left(\mathbb{R}_0^+\right)^m$, $m = m_1$ or $m_2 - 1$, into 'barycentric' subdivisions of their facets. These are

$$\Omega_{\Pi,k} = \left\{\vec{a} \in \Omega \mid a_{\Pi(1)} > \ldots > a_{\Pi(k)} > a_{\Pi(k+1)} = \ldots = a_{\Pi(m)} = 0\right\},$$

where Π is any permutation of $\{1,\dots,m\}$. The process defines an empirical measure with a density ρ as the large N limit. For ρ we have the equation

$$\partial_t \rho + \sum_{j=1}^{k} \partial_{a_{\Pi(j)}} \rho = -\sum_{j=k+1}^{m} s_{i(\Pi(j))} f\left(\sum_{j=1}^{m} \alpha_j(a_j)\right) \rho \quad \text{on } \Omega_{\Pi,k}$$

and at the influx boundary $\Omega_{\Pi,k}$ of $\Omega_{\Pi,k+1}$ we have

$$\frac{d(\rho|_{\Omega_{\Pi,k+1}})}{da_{\Pi(k+1)}} = s_{i(\Pi(k+1))} f\left(\sum_{j=1}^{m} \alpha_j(a_j)\right) \rho\,.$$

Here $\frac{d}{da_{\Pi(k+1)}}$ denotes the Radon-Nikodym derivative. If the latency period δ is larger than $t_{2\nu+1}-t_{2\nu}$ and $t_{2\nu+2}-t_{2\nu+1}$, then $f(\sum_{j=1}^m \alpha_j(a_j))$ depends only on a_j at the starting time, and there is no need to calculate transition probabilities in this step. The equations are reformulated in the following Lemma. Here we assume the evolution $\alpha_j(a)$ to be given. So the only random variable is $a = (t-t_j)_+$ or the t_j equivalents. Since $P_j(\alpha)$ is tight and independent of the infection time, this poses no restriction.

Lemma 3.1. *Suppose m functions α_j, $j = 1, \ldots, m$ are given in $L_1(\mathbb{R})$, with $0 \le \alpha_j \le \chi_{[\delta,\bar{C}]}$. Define a process on the state space $(\mathbb{R}_0^+)^m \times \{0,1\}^m$*

$$\partial_t a_j = \sigma_j \ , \quad \text{where } \sigma_j(t) = H(t-t_j)$$

and H is the Heaviside function. Further let

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon} P\Big(\sigma_j(t) = 1 \ \Big|\ \sigma_j(t-\varepsilon) = 0\Big) = s_j X(t),$$

$$X(t) = f\left(\sum_j \alpha_j(a_j(t))\right).$$

Then the transition probability $\rho(\vec{a}, \vec{\sigma}, \vec{\alpha}, \vec{a}_0, \vec{\sigma}_0, t)$ satisfies the kinetic equation

$$\partial_t \rho + div(\vec{\sigma}\rho) = 0 \ , \ \rho(\cdot,\cdot,\cdot,\vec{a}_0,\vec{\sigma}_0,0) = \delta_{\vec{a}_0,\vec{\sigma}_0} \ ,$$

and for each $\Pi \in \sum_m$, the symmetric group on $\{1,\ldots,m\}$, $k \in \{1,\ldots,m\}$,

$$\Omega_{\Pi,k} := \big\{a \in (\mathbb{R}_0^+)^m \mid a_{\Pi(1)} > a_{\Pi(2)} > \ldots > a_{\Pi(k+1)} = \ldots = a_{\Pi(m)} = 0\big\}$$

we have the influx boundary condition

$$\frac{d(\rho\big|_{\Omega_{\Pi,k}})}{da_{\Pi(k)}} = s_{i(\Pi(k))} f\left(\sum_{j=1}^m \alpha_j(a_j)\right) \cdot \rho, \quad \text{on} \quad \Omega_{\Pi,k-1}.$$

The equations are to be understood in the weak sense. ρ remains singular with respect to $\{a_j, \sigma_0(j) = 1\}$ and remains zero on the faces $\{a_j = 0 \mid \sigma_0(j) = 1\}$.

The kinetic equation for the transition probability ρ is just a convenient notation for the generator of the process on the test functions - which are of course random variables. They are solved using the method of characteristics. The solutions of the resulting delay differential equations have a continuous dependence on the data. That means if $\vec{a}_{0,1}$, $\vec{a}_{0,2}$ are initial data in the same $\Omega_{\Pi,k}$ then we have the following norm inequality

$$\begin{aligned}\|\rho(\cdot,\cdot,\cdot,\vec{a}_{0,1},\vec{\sigma}_{0,1},\cdot) - \rho(\cdot+\vec{a}_{0,1}-\vec{a}_{0,2},\cdot,\cdot,\vec{a}_{0,2},\vec{\sigma}_{0,2},\cdot)\| \\ \le C\left(|\vec{a}_{0,1}-\vec{a}_{0,2}| + |\vec{\sigma}_{0,1}-\vec{\sigma}_{0,2}|\right) \ .\end{aligned}$$

Similarly if α_j and $\tilde{\alpha}_j$ are different functions, then

$$\|\rho(\cdot,\cdot,\vec{\alpha},\vec{a}_{0,1},\vec{\sigma}_{0,1},\cdot)-\rho(\cdot,\cdot,\vec{\tilde{\alpha}},\vec{a}_{0,1},\vec{\sigma}_{0,1},\cdot)\|\le C\sum\int|\alpha_j-\tilde{\alpha}_j|\,.$$

The dependence on $\vec{\sigma}_0$ is important only for non-generic data with $\sigma_j=1$, $a_j=0$. These estimates also hold for the transition probability in path space, since the value of $\vec{a}$ at a time t determines the times when $\vec{\sigma}$ has jumped and so the path in state space, up to time t.

I.e. if $\rho(\vec{a},\vec{\sigma},t,\vec{\tilde{a}},\vec{\tilde{\sigma}},\tilde{t})$ gives the transition probability for the process starting at $\vec{\tilde{a}}$ in $\tilde{t}$ and reaching $\vec{a}$ at time t, then the distribution of $X(s)$ is the distribution of $f\left(\sum_{j=1}^m\alpha_j(a_j)\right)$ under $\rho(\vec{a},\vec{\sigma},s,\cdot,\cdot,\cdot)$.

The empirical distribution of the mean of M_1 copies of the process running independently as mentioned before, converges in the large M_1 limit to $\rho(\vec{a},\vec{\sigma},t,\vec{\tilde{a}},\vec{\tilde{\sigma}},\tilde{t})$. So if the empirical measure on initial data at time $\tilde{t}$ converges weakly to a measure μ, then in the large M_1 limit the empirical distribution of $X(s)$ will converge to the value

$$E\left(X(s)\right)=\int f\left(\sum_{j=1}^m\alpha_j(a_j)\right)d\rho(\vec{a},\vec{\sigma},s,\vec{\tilde{a}},\vec{\tilde{\sigma}},\tau)\;d\mu(\vec{\tilde{a}},\vec{\tilde{\sigma}},\tau)\,.$$

Lemma 3.1 can be applied with $m=m_1$ and $m=m_2-1$ respectively, and allows to conclude the following theorem

Theorem 3.2. *Let the infection process be given by 3 consecutive steps for a population consisting of M_1 groups of m_1 individuals of different types, numbered $i=1,\dots,m_1$. Individuals get infected only once, at times t_n, and afterwards produce the infectious agent at rate $\alpha_n((t-t_n)_+)$. The processes α_n are independent and identically distributed for each type of individual with a law $P_{i(n)}(d\alpha)$, $i(n)=n \mod (m_1)$, satisfying $0\le\alpha\le C\chi_{[\delta,\bar{C}]}$ and a Fréchet Kolmogoroff condition*

$$E\left(\int_0^\infty|\alpha*\varphi_\varepsilon-\alpha|\right)<\omega(\varepsilon)$$

for a Dirac sequence φ_ε and a modulus of continuity ω.

• *The first step of the process is an independent infection process in the groups, numbered by $r=1,\dots,M_1$ with rates $s_iX_r(t)$, s_i the susceptibility of type i and $X_r(t)$ the force of infection*

$$X_r(t)=f\left(\sum_{i=1}^{m_1}\alpha_{m_1(r-1)+i}((t-t_{m_1(r-1)+i})_+)\right)$$

for $t_{2\nu}\le t\le t_{2\nu+1}$. f is monotone bounded Lipschitz, describing a dose response curve for in group infections.

• *The second step is a consecutive filling up of M_2 meeting places of capacity m_2, numbered $l = 1, \dots, M_2$, where $M_2 m_2 = M_1 m_1$, with equal probability for each individual.*

• *The third step is again an independent process of infection at the meeting places with rates $s_{i(n)} \bar{X}_{l(n,\nu)}(t)$, where $l(n,\nu)$ is the number l assigned to individual n at time $t_{2\nu+1}$.*

$$\bar{X}_l(t) = \bar{f}\left(\sum_{l(n,\nu)=l} \alpha_n((t-t_n)_+)\right)$$

for $t_{2\nu+1} \leqq t \leqq t_{2\nu+2}$ where $\bar{f}$, monotone bounded Lipschitz, is the dose response curve at the meeting places.

In the large M_1 limit the law of infection of a susceptible individual of type i in a group is given by the rates

$$s_i X(t) = s_i f\left(\sum_{j=1}^{m_1} \alpha_j((t-t_j)_+)\right) \quad \text{for} \quad t_{2\nu} < t < t_{2\nu+1}$$

and α_j is distributed according to $P_j(\alpha)$ and

$$s_i \bar{X}(t) = s_i \bar{f}\left(\sum_{j=1}^{m_2-1} \alpha_j((t-t_j)_+)\right) \quad \text{for} \quad t_{2\nu+1} < t < t_{2\nu}$$

where α_j and $(t_{2\nu+1} - t_j)_+ =: a_{j,\nu}$ and i are independently chosen according to the equal distribution of types, the distribution of α according to P_i and the distribution of a_i given as the marginal of a distribution in $(a_1, \dots, a_{m_1})$ with densities $\rho_{\Pi,k}$ on $\Omega_{\Pi,k}$ the barycentric subdivision of the facets on $(\mathbb{R}_0^+)^m$.

$$\Omega_{\Pi,k} := \{\vec{a} \mid a_{\Pi(1)} > \ldots > a_{\Pi(k+1)} = a_{\Pi(k+2)} = \ldots = 0\}$$

$\rho_{\Pi,k}$ satisfies the (deterministic) kinetic system

$$\partial_t \rho + div(\vec{\sigma}\rho) = 0\ ,\ \vec{\sigma} = (\chi_{a_i>0})_{i=1}^{m_1-1}$$

with the influx boundary conditions

$$\frac{d(\rho|_{\Omega_{\Pi,k+1}})}{da_{\Pi(k+1)}} = s_{i(\Pi(k+1))} \int A(\vec{a},t)\, d\rho(\vec{a}, t_{2\nu}) \text{ for } t_{2\nu} < t < t_{2\nu+1} \text{ in } \Omega_{\Pi,k}$$

$$\frac{d(\rho|_{\Omega_{\Pi,k+1}})}{da_{\Pi(k+1)}} = s_{i(\Pi(k+1))} \cdot m_1^{1-m_2} \sum_{I \in \{1,\dots,m_1\}^{m_2-1}} \int B_I(a_1, \dots, a_{m_2-1}, t - t_{2\nu+1}) \cdot \prod_{j=1}^{m_2-1} d\rho_{I(j)}(a_j, t_{2\nu+1}) \text{ for } t_{2\nu+1} < t < t_{2\nu+2} \text{ in } \Omega_{\Pi,k}.$$

The functions A and B are given by

$$A(a_1,\dots,a_{m_1},t)=\int f\left(\sum_{i=1}^{m_1}\alpha_i(b_i)\right)d\rho(\vec{b},\vec{\alpha},\vec{a},t)\prod_{i=1}^{m_1}dP_i(\alpha_i)$$

$$B_I(a_1,\dots,a_{m_2-1},t)=\int \bar{f}\left(\sum_{j=1}^{m_2-1}\alpha_j(b_j)\right)d\rho(\vec{b},\vec{\alpha},\vec{a},t)\prod_{j=1}^{m_2-1}dP_{i(I(j))}(\alpha_j)$$

with the transition probabilities $\rho(\vec{b},\vec{\alpha},\vec{a},t)$ given by Lemma 3.1.

Here we assume that the empirical distribution of the initial data in α and a converge to the product measure of $dP_i(\alpha)$ and $\rho(\vec{a})\ \prod_{i=1}^{m_1}da_i$. The dependence of the transition probability on $\vec{\sigma}$ can be dropped since the probability of $t_n=0$ is assumed to be zero.

4 Modeling competition between virus variants

Here we modify the model to take different variants of the virus, and therefore reinfection into account.

If competition between different, here two, variants of the virus have to be modeled, each individual can be infected twice, $t_{n,\varepsilon}$ denoting the time of infection of variant $\varepsilon=1,2$. Also in this case a temporary reduction of the susceptibility, $s_{\varepsilon,i}$ has to be taken into account. Denote the susceptibility of an individual with respect to the variants as $s_{\varepsilon,n}$. We assume that

$$\begin{aligned} s_{1,n}(t) &= s_{1,i(n)}\left(1-\delta_1((t-t_{n,2})_+)\right)\\ s_{2,n}(t) &= s_{2,i(n)}\left(1-\delta_2((t-t_{n,1})_+)\right)\end{aligned}$$

where $\chi_{(0,L]}\le\delta_1,\delta_2\le\chi_{(0,\bar{L}]}$. Further we will assume $\bar{L}>\bar{C}$ so that there is no simultaneously active infection with both variants. The state space for the individual will become

$$Y_0\times\left(L_1(\delta,\bar{C})\right)^2\times(L_1(0,\bar{L}))^2$$

with $Y_0=\{1,\dots,m_1\}\times(\mathbb{R}_0^+)^2\times(\{0,1\})^2$. In the groups as well as in the meeting places we will number the individuals by $\{1,\dots,m\}\times\{1,2\}$. As before we will partition $(\mathbb{R}_0^+)^{2m}$ into sets

$$\Omega_{\Pi,k}=\left\{\vec{a}\in(\mathbb{R}_0^+)^{2m}\mid a_{\Pi(1)}>a_{\Pi(2)}>\ldots>a_{\Pi(k+1)}=\ldots=a_{\Pi(2m)}=0\right\},$$

where Π is any bijective map from $\{1,\dots,2m\}$ into $\{1,\dots,m\}\times\{1,2\}$. We will denote by $\bar{\Pi}=(Id\times\tau)\circ\Pi$ the interchange of 1 and 2 in the 'permutation'. So $s_{\Pi(l)}$ will depend on $a_{\bar{\Pi}(l)}$ and the formula for the transition probability for the two competing virus variants will become

Lemma 4.1. *Let the functions $\alpha_{j,1}, s_{j,1}, \alpha_{j,2}, s_{j,2}$, $j = 1, \dots, m$ be given and let the infection rates for the susceptible individuals for variants $\varepsilon = 1, 2$ be given by*

$$X_{j,\varepsilon} = s_{j,\varepsilon}(a_{j,\tau(\varepsilon)}) f\left(\sum_{j=1}^{m} \alpha_{j,\varepsilon}(a_{j,\varepsilon})\right),$$

where $a_{j,\varepsilon} = (t - t_{j,\varepsilon})_+$ and where $t_{j,\varepsilon}$ is the time of infection of an individual j with virus variant ε. Then the transition probability of the process in $(\mathbb{R}_0^+)^{2m}$ satisfies the kinetic equation

$$\partial_t \rho + div(\vec{\sigma}\rho) = 0 \quad , \quad \vec{\sigma} = \chi_{(a_{j,\varepsilon}>0)} \text{ in } \Omega = (\mathbb{R}_0^+)^{2m}$$

with influx boundary condition

$$\frac{d(\rho|_{\Omega_{\Pi,k+1}})}{da_{\Pi(k+1)}} = s_{\Pi(k+1)}(a_{\bar{\Pi}(k+1)}) f_{(\Pi(k+1))_2}\left(\sum_{j=1}^{m} \alpha_{j,(\Pi(k+1))_2}(a_{j,(\Pi(k+1))_2})\right)\rho$$

in $\Omega_{\Pi,k}$ where $\Omega_{\Pi,k}$ is defined as above, $\frac{d}{da}$ denotes the Radon-Nikodym derivative, and $(\Pi(k+1))_2$ is the second component of $\Pi(k+1)$. The initial condition is $\rho(0) = \delta_{\vec{a}_0,\vec{\sigma}_0}$, $\vec{a}_0 \in (\mathbb{R}_0^+)^{2m}$, $\vec{\sigma}_0 \in \{0,1\}^{2m}$.

For the process defined as in Lemma 4.1 for M_1 independent groups in the time intervals $(t_{2\nu}, t_{2\nu+1})$, a consecutive random filling up of M_2 meeting places of capacity m_2, $M_2 m_2 = M_1 m_1$, at times $t_{2\nu+1}$, and again an independent infection process at the meeting places in the time intervals $(t_{2\nu+1}, t_{2\nu+2})$, in the large M_1 limit, the density in the phase space will satisfy

Theorem 4.2. *The limit of the empirical density $\frac{1}{M_1}\sum_{l=1}^{M_1} \delta_{\vec{a}_l(t)}$ in $(\mathbb{R}_0^+)^{2m}$ will satisfy the kinetic equation*

$$\partial_t \rho + div(\vec{\sigma}\rho) = 0 \quad , \quad \vec{\sigma}(\vec{a}) = (\chi_{a_j>0})_{j\in\{1,\dots,m_1\}\times\{1,2\}}$$

and

$$\frac{d(\rho|_{\Omega_{\Pi,k+1}})}{da_{\Pi(k+1)}} = A_{\Pi,k+1}(t - t_{2\nu}, \vec{a})\rho \quad in \quad (t_{2\nu}, t_{2\nu+1}) \times \Omega_{\Pi,k}$$

$$\frac{d(\rho|_{\Omega_{\Pi,k+1}})}{da_{\Pi(k+1)}} = m_1^{1-m_2} \sum_{I\in\{1,\dots,m_1\}^{m_2-1}} \left[\int_{(\mathbb{R}_0^+)^{2m_2-1}} B_I(t - t_{2\nu+1}, a_{\bar{\Pi}(k+1)}, \vec{b}) \cdot \prod_{j=1}^{m_2-1} \rho_{i(I(j))}(b_{j,1}, b_{j,2}, t_{2\nu+1})\, db_{j,1}\, db_{j,2}\right]\rho$$
$$in \quad (t_{2\nu+1}, t_{2\nu+2}) \times \Omega_{\Pi,k}$$

$$A_{\Pi,k+1}(\tau,\vec{a}) =$$
$$E\left(s_{\Pi(k+1)}(a_{\bar{\Pi}(k+1)})\int f_{\Pi(k+1)}\left(\sum_{i=1}^{m_1}\alpha_{i,(\Pi(k+1))_2}(\beta_{i,(\Pi(k+1))_2})\right)\,d\rho(\vec{\beta},\tau,\vec{a},\vec{s},\vec{\alpha})\right),$$

where the expected value is taken w.r.t. the product measure of $dP_i(s_1,s_2,\,\alpha_1,\alpha_2)$*, and* $\rho(\vec{\beta},\tau,\vec{a},\vec{s},\vec{\alpha})$ *is the transition probability calculated in Lemma 4.1.*

$$B_I(\tau,a_{\bar{\Pi}(k+1)},\vec{b}) =$$
$$E\left(s_0(a_{\bar{\Pi}(k+1)})\int \bar{f}_{(\Pi(k+1))_2}\left(\sum_{j=1}^{m_2}\alpha_{j,(\Pi(k+1))_2}\left(\beta_{j,(\Pi(k+1))_2}\right)\right)\,d\rho(\vec{\beta},\tau,\vec{b},\vec{s},\vec{\alpha})\right),$$

where $\rho(\vec{\beta},\tau,\vec{b},\vec{s},\vec{\alpha})$ *is the transition probability calculated in Lemma 4.1 and the expectation in* $s_0,(s_{j,1},s_{j,2},\alpha_{j,1},\alpha_{j,2})_{j=1}^{m_2-1}$ *is taken w.r.t. the product measure* $dP_{i(\Pi(k+1))}\times\prod_{j=1}^{m_2-1}dP_{i(I(j))}$.

Here we assumed that the empirical measure for the initial data s_1, s_2, α_1, α_2 for the individuals of type i with one or two infections before the initial time converges to the product measure $\prod_i dP_i(s_1,s_2,\alpha_1,\alpha_2)$, and the limit of the empirical measure of the time since infection at time zero has a density

5 Discussion

The type of convergence in the large N limit we prove is very much akin to two scale convergence in homogenization, or for that matter in lattice Hamiltonians [6]. The local process is itself a kinetic or Boltzmann type process as was shown in the lemmata, driven by the population evolution. In that respect the result is very similar to [4] only that for the McKendrick processes the driver is not an ODE but a PDE.

We formulated conditions on the virus production of infected individuals, which are in fact conditions on the large N limit of the empirical distribution. The independence assumption is very natural. It is also realistic not to assume the virus production α to be itself a process in $\mathbb{R}$. α is just a random variable depending on a hidden Markov process about which we have basically no data. But it is of course quite possible that our distribution P_i of α depends on t or even the virus exposition at the infection time. In principle that can be checked and the model be modified accordingly.

The shape of the nonlinearity - a concave polynomial of degree m_2 - is a consequence of the limited capacity m_2 of the meeting places. Roughly speaking this is the evaluation of the dose response curve w.r.t. to a multinomial distribution, see [7], Chapter 5. For large meeting places other limits, e.g. $m_2\approx N^\beta$ would be appropriate. And then $\bar{I}:=m_2^{-1}\sum_{j=1}^{m_2}\alpha(a_j)$

itself would become deterministic. This would lead to a nonlinearity of type $\rho\bar{f}(m_1^{-1}\sum_{i=1}^{m_1}\alpha_i(a_i)\rho)$ for the kinetic equation discussed e.g. in [8] for groups of two individuals.

The mixing process we have been looking at here, is a very particular example adapted e.g. to a metro filling up during the rush hour. Other processes have to be considered. The shape of the nonlinearity should remain the same [8].

The modification of the standard Kermack-McKendrick model we have proposed in this paper allows to separate the contagion mechanism in groups with reoccuring contacts from the one at random encounters. Further it allows for other infection laws than the standard mass action type of law. Both, we think, are necessary to realistically model not only Covid, but also other aerosol transmitted diseases.

For Covid we have data which allow to estimate the true extent of (asymptomatic) infection in the affected populations. And here we can see how limited the effect of lockdown measures has been. We also see, that stability w.r.t. the introduction of a small percentage of infections does not mean stability w.r.t. to a moderate percentage of imported infections. Both phenomena can be modeled with the type of system we have introduced here. And this system naturally arises from a stochastic model.

References

[1] Bundesamt für Gesundheit (BAG), Schweiz. https://www.bag.admin.ch/.

[2] A. Chandrashekar *et al.* SARS-CoV-2 infection protects against rechallenge in rhesus macaques. *Science*, 369(6505):812–817, 2020.

[3] Charité Research Organization (CRO). Studie zur Untersuchung des Corona-Infektionsrisikos im öffentlichen Personen-Nahverkehr. Ministry of Traffic of the German federal state of Baden Württemberg. https://vm.baden-wuerttemberg.de/fileadmin/redaktion/m-mvi/intern/Dateien/PDF/PM_Anhang/210510_%C3%96PNV_Studie_zum_Corona-Infektionsrisiko_im_%C3%96PNV.pdf, 2021.

[4] R. Forien and E. Pardoux. Household epidemic models and McKean-Vlasov Poisson driven stochastic differential equations. *Ann. Appl. Probab.*, 32(2):1210–1233, 2022.

[5] W. O. Kermack and A. G. McKendrick. A Contribution to the Mathematical Theory of Epidemics. *Proc. R. Soc. Lond. A*, 115(772):700–721, 1927.

[6] R. Kotecký and S. Luckhaus. Nonlinear elastic free energies and gradient Young-Gibbs measures. *Commun. Math. Phys.*, 326(3):887–917, 2014.

[7] S. Luckhaus. Corona, Mathematical Epidemiology, Herd Immunity, and Data. Preprint 105, MPI for Mathematics in the Sciences, Leipzig, 2020.

[8] S. Luckhaus and A. Stevens. A free boundary problem - in time - for the spread of Covid-19. *J. Math. Biol.*, 86(45):17pp., 2023. https://rdcu.be/c5JdO.

[9] S. Luckhaus and A. Stevens. Kermack and McKendrick models on a two-scale network and connections to the Boltzmann equations. In J.-M. Morel and B. Teissier, editors, *Mathematics Going Forward. Collected Mathematical Brushstrokes*, volume 2313 of *Lect. Notes Math.*, pages 417–427. Springer International Publishing, Cham, 2023.

[10] A. G. McKendrick. Studies on the Theory of Continuous Probabilities, with Special Reference to its Bearing on Natural Phenomena of a Progressive Nature. *Proc. London Math. Soc.*, 13:401–416, 1914.

[11] A. G. McKendrick. Application of Mathematics to Medical Problems. *Proc. R. Soc. Edinb.*, 44:98–130, 1925.

[12] National Health Service (NHS). https://www.england.nhs.uk.

[13] S. Stringhini *et al.* Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. *The Lancet*, 396:313–319, 2020.

Stephan Luckhaus ✉
Mathematical Institute
University of Leipzig
Leipzig, Germany
Stephan.Luckhaus@math.uni-leipzig.de

Angela Stevens
Institute for Analysis and Numerics
University of Münster
Münster, Germany
angela.stevens@wwu.de
https://www.uni-muenster.de/AMM/stevens