Spatio-temporal simulation of Covid-19 propagation via continuous automata

Vũ Ngọc San*

WARNING - UNFINISHED DRAFT – see disclaimer below

Abstract

We present a new continuous automata based computer simulation of virus propagation in human populations, and apply it to the Covid-19 outbreak, in various scales and situations. We also take the opportunity to propose various mathematical questions, and ask about their biological relevance.

Modelling the evolution of epidemic outbreaks has become an important tool for all politicians to take appropriate measures for fighting the disease. However, one has to be aware that all simulations are very crude, and their scientific grounds will *never* ensure that their predictions have anything to do with reality. The reason is that they are based on oversimplifications of models and parameters, choosing unrealistic rules simply because they lead to tractable mathematical equations. This is the case for the celebrated SIR model, which assumes that all individuals behave in the same way, and any infected person will possibly infect any healthy individual, regardless of their position in society, or even their geographical localization. The beauty of mathematics — and of simple models — is that they can still *mean* something. When researchers have succeeded in singling out the most important behaviour, neglecting all secondary effects, they might come up with a result that, qualitatively, show global trends that *can* be observed in reality, statistically. The temptation is great, then, to tune all possible parameters to try and make it fit the real curves (of infected people, of deaths, etc. — assuming that we can trust these numbers), in order to finally make predictions. This is of course very dangerous.

*Univ Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France

The goal of our work is to implement a "microscopic" simulation *at the level of individuals*, without attempting to obtain solvable differential equations, based as much as possible on the virus transmissions mechanisms that are observed in reality, in order to then scale the simulation to a whole population. Thus, this work belongs to the general class of "Individual-Based Models" (IBM), see [8].

A physics analogy would be to check whether the celebrated Ohm Law $V = RI$ relating electric current to electric potential, which is a purely phenomenological law, can be justified by understanding the microscopic behaviour of atoms and electrons in the metal. (Such a rigorous justification, by the way, is still an open question.) For us, atoms are replaced by individuals, and electron transfer is replaced by viral contamination; instead of observing global trends at the macroscopic level, we equip individuals with interaction laws, and let them evolve. Of course, a crucial difference between Ohm's law and the virus problem is that the former has been amply verified to high degree of precision by many experiments, while we can of course not perform such experiments with human populations.

Even though we believe that IBMs can implement much more realistic parameters than in the usual compartmental models like SIR or SEIR, we can't claim that the final result gives a better prediction; this remains a crude simulation and should be taken with all care. Moreover, our automata approach is by nature more adapted to a closed cohort situation rather than a whole country. It may, nevertheless, be useful in discovering new qualitative scenarios that can happen in response to various political decisions like closing schools, locking the population down, etc.

Disclaimer. The initial goal of this paper was mainly to help the author (who is a mathematician) understand virus modelling. The same incentive explains why I have decided to program the model "from scratch", without relying on existing libraries. However, I now hope that this can be helpful to others, and while I don't claim originality, I believe that it contains some new remarks of interest, be it merely the sometimes unusual viewpoint.

This work was done during the first pandemic lockdown of France in Spring 2020. I initially thought I would continue it afterwards, but I was unable to find the necessary time to do it. My excuse for putting it online in this draft form is twofold. First I think that some ideas are interesting. Second, I understand that, in such a case like the Covid-19 pandemic, researchers are expected to quickly post their findings, in a global effort to find the best solutions. This contradicts the natural desire to check everything in

details before release. This is my second reason. As a corollary, I'm happy to receive any feedback.

This version: June 13, 2020.

1 Continuous automata

The principle of the simulation is very simple. It can be compared to a computer game with many characters. Each individual has a set of specifications — age, gender, geographic position, viral load, immunity, etc. — and even a personal motion specification (stay at home, moves randomly, commutes to work, etc.). This automaton is called *continuous*, as opposed to the better known "cellular" automata, simply because its location is not restricted to a discrete grid, but can be any floating-point number in the 2-dimensional plane $\mathbb{R}^2$.

Then, each individual has an evolving state (healthy, infected, sick, etc. — we shall discuss the precise meaning of these states below). The state can evolve simply in time (a sick individual can naturally recover after a couple of weeks), or can switch to "infected" by proximity to other infected persons. One of the most important parts of our simulation is to compute the *transmission probability* between two persons.

The program is implemented in the OCaml language. This choice is motivated by a high-level functional language, which makes it simple to implement rigorous rules, and makes it hard to produce a code that would be obviously wrong, and by the fact that it is a fast compiled language, which we need when growing the number of automata to several thousands.

We propose two simulation modes. The *graphical* mode shows, on the computer's screen, all persons, moving and getting infected, on "real (accelerated) time". This is useful for small scale experiments (less than 1000 individuals) and for demonstrations and public outreach. In the *batch* mode, no graphics is produced, but the logs are saved in a file and can be used to show statistics. The batch mode is of course much faster and can handle a larger population, depending on the computer power and memory.

It would be very interesting to try and *parallelize* the algorithm; this is not currently done.

2 Individuals' states

Our choice of "states" is quite similar to the usual "compartments" used in many virus simulations. In order to simplify the reasoning — and the program — we have chosen *injective* states, meaning that an individual can only belong to one and only one state. We believe that the choice of states is very important and debatable. Our states are defined in the `Stats` module, as follows. For each state, we indicate the possible evolution, which is summarized in Figure 1.

1. `Healthy` : Has never been infected, or was `Latent` with viral load close to zero. No evolution when isolated. Will become `Latent` when infected.

2. `Latent` : Is infected but unnoticeable: is not symptomatic and cannot transmit the virus. Was either `Healthy` or `Recovered` before. Will becore either `Carrier` or `Healthy` (very rare).

3. `Carrier` : Carries the virus but not symptomatic. Was `Latent` before. Will become `Sick` or `Latent` (if the person can fight the virus away before being symptomatic).

4. `Sick` : Has symptoms (should stay in bed). Was `Carrier` or `Critical` before. Will become either `Critical` or `Recovered`.

5. `Critical` : Must/should be taken to Critical Care Unit. Was `Sick` before. Will become either `Dead` or `Sick` again.

6. `Recovered` : Same viral load as `Carrier` but only happens after being `Sick`. No symptoms anymore, cannot get sick again but can infect others. Was `Sick` before. Will become `Dormant`.

7. `Dormant` : Same as `Latent` but can be reinfected again.

8. `Dead` : Was `Critical` before.

In addition to these states, we allow any individual to be temporarily "removed" from our spatial domain, which means that they continue their self-evolution, but are isolated from the rest of the population. This can be used to model people going to sleep at night, for instance.

Question 2.1 Do all arrows between the states in Figure 1 correspond to existing biological or medical processes? ★

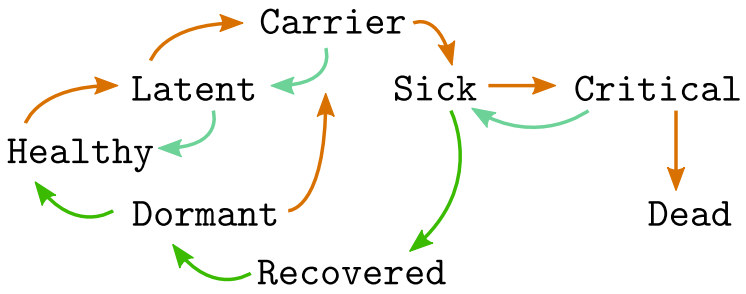


Figure 1: Individuals' states and their evolution

3 Viral load

One of the main novelty of this work, as far as we can tell, it to endow each individual with a specific *viral load function*, upon which all evolution between one state to another — and the virus transmission probability as well — will be essentially based.

Of course, the name "viral load" suggests that we are inspired here by the amount of viruses present in the body. However, the analogy with the biological viral load should not be take too literally. Our viral load should be taken as a phenomenological mean to conveniently summarize what happens to the person: will she or he become infected, infectious, or recover, or die, etc.

Any viral load can be tested with our program; in principle, it should depend on the virus under study, and on each individual. For the SARS-CoV-2, our choice for the viral load curve will be based on the reference curve shown in Figure 2; it is similar to what can be found in the literature (see for instance [4]), and obtained as the graph of the following function of time (d is expressed in days):

$$\text{vload}(d) = 1.1e^{\frac{-15}{(d+1)^2}}(b + e^{\frac{-80}{(28-d)^2}})$$

where the limit as $d \to \infty$ is given by b, which we take as $b = 0.05$. (For simplicity and speed, we assume that the limit is reached when $d > 25$). The precise mathematical expression is not important — we could have taken a mixture of Weibull shapes, but we wanted to implement the following features: first, after a short latency period, we have a steep exponential growth (virus replication inside the body), then a large peak, and finally a slower exponential decay (when the immune system succeeds in fighting the virus).

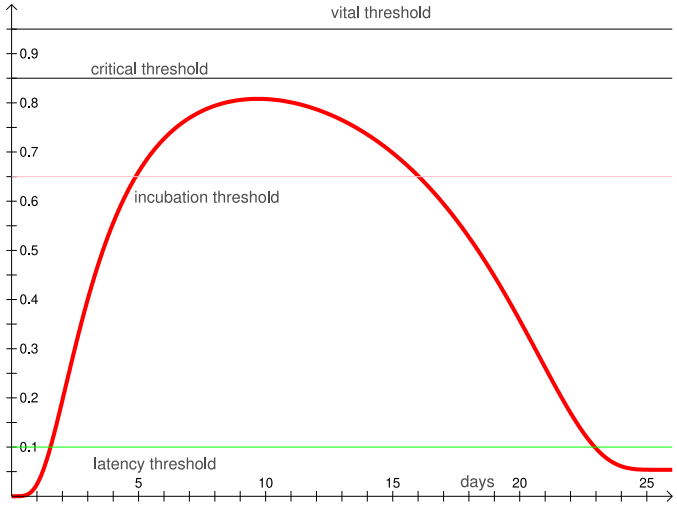


Figure 2: In red, the reference viral load function; the horizontal lines represent the various thresholds.

The limit b is taken non zero to account for the fact that most infected persons will keep a small portion of the virus in their body forever. However, we will show (??) that this does not have a big influence on simulations.

As soon as a simulated person is infected, we assign them the viral load curve, and let it evolve under (shifted and renormalized) time. In order to decide of the evolution of the disease, we introduce four *thresholds*.

1. *Latency threshold* : The virus is not transmissible if the viral load is below it.

2. *Incubation threshold* : The person is symptomatic only if the viral load is above it.

3. *Critical threshold* : When the viral load goes above it, the individual needs intensive care.

4. *Vital threshold* : A viral load above it leads to death.

Not everybody reacts in the same way to a viral attack. For the Covid-19, its was shown, to a sufficient degree of certitude(???ref), that children are mildly affected, while symptoms become more severe with increasing age. Hence, instead of remodelling the viral load curve for each individual (which, in fact, is left as a possibility by the program), we adjust the various thresholds according to the age of the person. For instance, the curve represented

in Figure 2 will be assigned to persons of age 40 (with random small variations). Such a person will be contagious after day 1.5, become sick at day 5, stay so until day 16, and finally recover. After day 23, he or she will not be contagious anymore. These numbers have been chosen to represent current knowledge, but are highly debatable.

Question 3.1 Are these numbers correctly following the true Covid-19 evolution, on average? Should we allow various time stretching depending on individuals? ★

The variation of thresholds with age is also a highly unclear issue. Here we consider that people of age > 70 will likely have to go to intensive care, and hence set their critical threshold to 0.70 instead of 0.85. In fact, we will raise of lower the four thresholds by the same amount. Likewise, for most children under 10, a SARS-CoV-2 infection will not develop the disease; hence we set their incubation threshold to 0.80. Of course, they can be carriers.

Question 3.2 What do we know precisely concerning statistics of Covid-19 evolution with respect to the age? Are the above numbers realistic? ★

4 Immunity

At the time of writing, one of the biggest mysteries about the Covid-19 disease is the conferred immunity. Can you become sick again one month after recovering from the disease? Or one year? or never? There is no firm answer. This explains why most governments are hesitant about the best strategy to remove lock-downs. Of course, many of us hope that, as is the case for Influenza, the virus should confer a immunity that should last *at least* a couple of months. It this were the case, and if severe cases were less frequent than what we currently observe, a reasonable strategy would be to let most of the population get infected and reach a mass immunity that would prevent the virus from spreading further. Wether this option is politically reasonable highly depends on the mortality rate, which is unknown as well, but strongly suspected to be high enough to rule this option out. Other, more pessimistic researchers remark that a few well known coronaviruses responsible for common cold confer very little immunity. As everybody knows, it is not unusual to catch several colds in a row in the same month.

In this study we take a semi-optimistic viewpoint. During the sickness and recovery period, we assume that temporary cellular response will be enough to prevent overinfection. Thus, only `Healthy` and `Dormant` states can become infected (hence changing their state to `Latent`). A long-term

humoral response (antibodies), which starts about 7-10 days after symptoms, should be strong after a couple a weeks, and remains present hopefully several months.

Hence, each person, after recovery, once the viral load has dropped down below the latency threshold, will acquire a certain immunity. We can devise a different immunity for each person. Moreover, we don't simply employ the naive, binary option "may I or may I not get infected again"; instead, as we did for the viral load, we will model this by the "immunity function", which depends on time. The effect of the immunity curve is to reduce the person's own viral load. Precisely, the immunity function takes values in the interval $[0, 1]$, and the person's *effective viral load* will be computed by multiplying the viral load by one minus the immunity:

$$\text{vload}_{\text{eff}}(d) = \text{vload}(d)(1 - \text{immunity}(d)). \tag{1}$$

The function we chose for modelling the long-term immunity is inspired by [2], and represented in Figure 3. You can see that this optimistic curve grants a fairly good immunity after 7-8 days, and on, forever. However, note that the limit immunity is not 1, which makes it possible to have a new infection, but with relatively low probability (see next section).

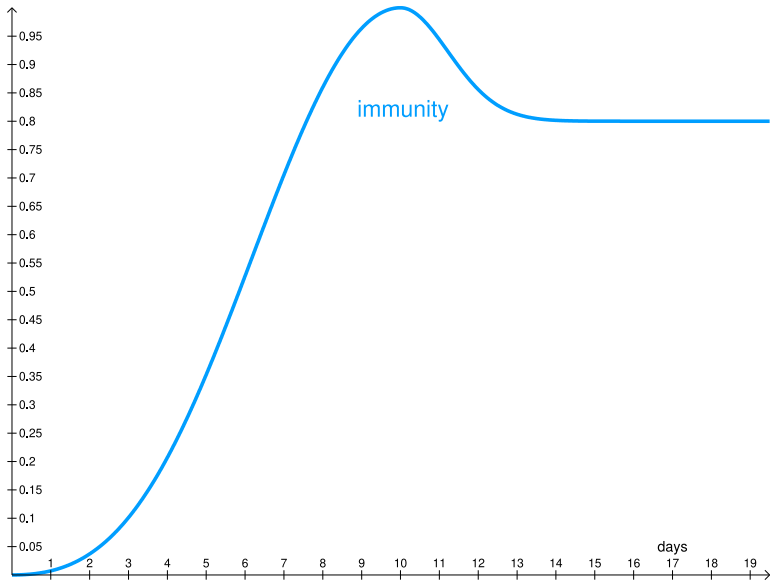


Figure 3: The immunity curve

Question 4.1 What do we really know about Covid-19 immunity? Is the

curve of Figure 3 realistic? ★

5 Transmission probability

The crucial ingredient of the simulation is, of course, how we decide that one individual A will infect another one B. Here again, we try to rely as much as possible on what is currently known about the mechanisms of transmission. At the time of writing, the main vector for the virus seems to be exhaled breath, and hence is highly dependent on the mutual distance between A and B. People are encouraged to respect *social distancing*, and a distance of 3 meters seems to be safe of infection risk — except when sneezing or coughing.

How to model this? It is an interesting question. Should we consider a contamination rate or a contamination probability?

In the first case, the contaminated person may be strongly or weakly contaminated, and this will influence the development of the disease. In the second case, a person is either contaminated or not.

Although the first option seems reasonable, the highly exponential rate of virus reproduction in the body could lead to conclude that the difference is only a matter of shifting the curve by a few hours, which is negligible.

On the other hand, a low virus quantity in the first few hours could be enough for the body to efficiently fight against the virus, before it gets overwhelmed. (This would explain why the biological viral load resembles the SIR curve, but at the scale of the body instead of that of a population).

Most research models think in terms of "transmission probability per contact per unit time", but the precise meaning of this is far from clear. If you stay close to an infected person during one minute or one hour, clearly your chances of being infected are not identical; but how many "contacts" did you have? Physically speaking, you might have never touched the person at all, but still have a high probability of getting infected by droplets carried by the breath.

Question 5.1 What is the good notion of transmission probability between two people, that would correspond to a solid biological ground? (see also Section 5.2 below). ★

5.1 Instantaneous infectiousness

In this work, we introduce a contamination probability $p_{A\to B}$ "per unit of time", but *not* per contact. Instead, the probability will strongly depend on

the *distance* AB between A and B. Thus, if they are far apart, the probability will be zero, and this amounts to considering that there was *no contact*. On the contrary, if the distance is zero, or close to zero, for a period of time T, then the contact is maximal, and it makes sense to introduce a probability $p_{A\to B}(T)$ that depends on T. How does $p_{A\to B}(T)$ depend on T? Suppose that A produces infected droplets at a constant rate. Then, for a very short period of time δt, you inhale a certain quantity $m\delta t$ of these droplets. This will trigger immediate response from your immune system, which is able to kill the virus at the rate M per second. If $M < m$, the virus will be able to replicate, leading to exponential growth, and you will be infected. On the contrary, if $M > m$, you will not let the virus replicate, and we consider that you are not infected. Of course, this description is simplistic, and we don't know m and M, but these numbers are reasonably accounted for by the "safe critical distance" beyond which you stay safe. Indeed, the number m clearly decreases with the distance AB (for at least two reasons: one is that they are emitted in every direction and hence the concentration per unit air volume decreases as the inverse distance squared; the second one is that droplets will eventually fall on the ground, so the concentration is actually even smaller). If AB is close to zero, with this reasoning the infection is certain, *i.e.* the probability of being infected in one second is 1. This is not realistic, first because the rate of expelled infected droplets may vary, especially if δt is very small, and only the *time average of* m is relevant; likewise, the immune system is not so steady, and may have time variations. We can change the reasoning as follows. Assume that the average number of "contaminating droplets" inhaled by B per unit of time is $\nu(AB)$ (taking into account the possible immune response). From the Poisson law, the probability of *zero* contaminating droplets during a time δt is $e^{-\nu(AB)\delta t}$. Hence, a first candidate would be

$$p_{A\to B}(\delta t) = 1 - e^{-\nu(AB)\delta t}$$

Thus, as soon as $\nu(AB) \neq 0$, we see that the probability of getting infected tends to 1 as $\delta t \to \infty$, which is consistent with the crudest analysis above.

Let us now compare to the experiment on influenza outbreaks conducted by [7]. They assumed $\nu(AB) = 0$ if $AB \geqslant 3\text{m}$, and $\nu(AB) = \nu_0$ to be constant otherwise. Based on the Moser study [6], they estimated the probability $p_{A\to B}(20\text{s}) = 0.003$, which means $\nu_0 = -\ln 0.997/20 \simeq 1.5 \times 10^{-4}\text{s}^{-1}$. With these figures, the probability of getting infected is 50% after one hour and 17 minutes. Most studies agree on the fact that the SARS-CoV-2 is more contagious that influenza, but we could not find precise experiments. For instance, we could decide that the infection probability would be 50% after

30 minutes close to a sick individual, which means $\nu_0 = 3.8 \times 10^{-4}\text{s}^{-1}$. Of course, ν_0 has an enormous influence on the development of the epidemics, see for instance (14), (15); in the absence of direct experiments, one could try to calibrate it *a posteriori*, but this may falsely compensate for other hidden mechanisms. ??? étudier la variation de ν_0

However, the infection probability $p_{A\to B}$ depends actually on many more parameters. It varies in a essential way with the distance $r = AB$: it is clearly much more dangerous to be at distance 1cm rather than one meter away of an infected individual A. Moreover, studies (???ref) also suggest that a severely ill individual will be more contagious than an asymptomatic one. Hence, $p_{A\to B}$ should depend on the viral load of A. Finally, there remains an important, external factor: prophylaxis (preventive healthcare) of both A and B.

Given our ignorance of a precise study concerning the dependence of $p_{A\to B}$ on the distance $r = AB$, we propose the following argument. Assume that A produces on average N infected droplets per second. These droplets will travel all around him or her. To simplify, we assume a purely horizontal motion at constant speed in the disc of radius $r_{\max} = 3$ meters, and then a quick fall down. Hence the droplets initially produced in a small disc of radius r_0 around A will move, at time t, into an annulus bounded by the circles of radii r and $r + r_0$, which has area $\pi r_0(2r + r_0)$. Therefore, if $r = AB$, the average number of infected droplets that can be inhaled by B per second is

$$\nu(r) = \frac{Nr_0}{2r + r_0},$$

Of course, a more realistic version would take into account droplets regularly falling down on the ground; moreover, what happens at a larger distance $r > r_{\max}$ is absolutely unclear. In order to take these effects into account, we simply multiply by an affine cut-off $\max(0, c(r_{\max} - r))$. Seeing that the average distance to the origin in the disc of radius $r_{\max}$ (assuming uniform 2D distribution) is $\frac{2r_{\max}}{3} = 2$, we wish to enforce $\nu(2) = \nu_0$, hence finally

$$\nu(r) = \frac{\nu_0(4 + r_0)(r_{\max} - r)_+}{2r + r_0}. \tag{2}$$

We will choose $r_0 = 15$cm.

How does the viral load of A contribute to $p_{A\to B}$? (We are talking here about the *effective* viral load, which takes into account A's immunity, see Section 4). To our knowledge, nothing more precise than "if the viral load is important, A should be more contagious" is known. In our model, a person is

not contagious when the viral load goes below the latency threshold. Hence it is reasonable to take into account only the proportion of viral load *above* the latency threshold τ_{lat}:

$$v_{\text{cont}} := \max\left(0, \quad \frac{\text{vload}_{\text{eff}} - \tau_{\text{lat}}}{v_{\max} - \tau_{\text{lat}}}\right)$$

where $v_{\max} \simeq 0.8$ is the maximum of the viral load curve, see Figure 2.

In order to model *prophylaxis*, we will endow each individual with a protection factor between 0 (not protected) and 1 (fully protected). For simplicity, we don't distinguish between *outgoing* protection (not infecting others) and *ingoing* protection (not infecting oneself). For instance, if someone wears a protective face mask, we assume it will work in both directions (which is not technically correct, but probably OK as a first approximation). If needed, it would be easy to modify the program to implement two protection factors per individual.

Summarizing, the final formula for $p_{A\to B}$ is the following (assuming that all characteristics of A and B don't vary within the time δt):

$$p_{A\to B}(\delta t) = 1 - e^{-\nu_{A\to B}\delta t}, \tag{3}$$

where $\nu_{A\to B}$ is the average amount of contamination per second, which we take as

$$\nu_{A\to B} = \nu_0 s(AB) v_{\text{cont}}(A)(1 - \text{prophylaxis}(A))(1 - \text{prophylaxis}(B)), \tag{4}$$

where $s(AB)$ is the excretion spread factor at distance $r = AB$ defined as (see (2))

$$s(r) = \frac{(4 + r_0)(r_{\max} - r)_+}{2r + r_0} \tag{5}$$

with $r_{\max} = 3\text{m}$, $r_0 = 0.15\text{m}$, and $\nu_0 = 3.8 \times 10^{-4}\text{s}^{-1}$. Thus, if we stay one meter away from a fully infected person during half an hour, without protection, the probability of getting infected is 93% (Figure 4).

Question 5.2 Is the infection probability represented in Figure 4 realistic? ★

5.2 Effective infectiousness and reproduction number R

In many studies, for which the time scale is much larger (the unit being often one day), and for which all constants are considered on a population average, the answer to Question 5.1 is of course easier. The "infectiousness" [3] of

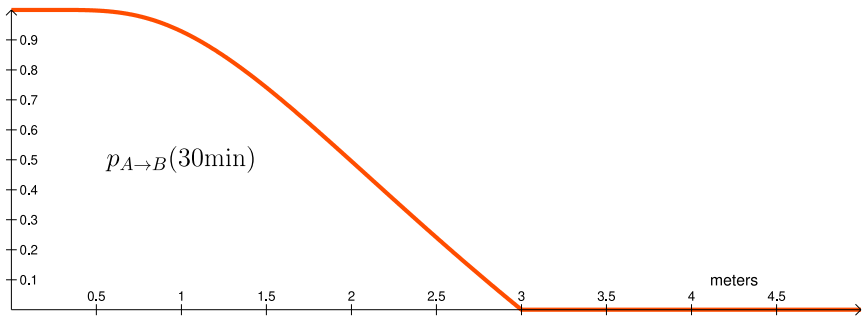


Figure 4: Infection probability after 30min as a function of the distance between A and B. Person A is supposed to have maximum viral load, and neither A nor B take any protective measure. When all parameters are randomized, the particular shape of this curve is not important, but the integral $\int_0^{r_{\max}} rs(s)\,\mathrm{d}r$ is crucial, see Section 8.1.

an "individual" (meaning an average individual) is the average number of persons who where infected by this individual during a day.

Then the "reproduction number" R (or R_0, if one discard long-term immunity effects and re-infections) is the cumulative infectiousness of an average individual during a complete infectious period.

How can one relate these quantities to our instantaneous infectiousness $p_{A\to B}(\delta t)$ defined in (3)? Of course we cannot expect a general formula, because it depends on the random properties of the motion and characteristics of all individuals.

For a given person A, one may reasonably assume that the random events "B is infected by A" and "B' is infected by A" are independent, whenever $B \neq B'$. Hence the expectation of the random variable "X_A :=number of persons infected by A" during the time δt is

$$E_A(\delta t) = \sum_{B\neq A} p_{A\to B}(\delta t). \tag{6}$$

However, to obtain the total expected number if infected people, one should *not* sum over all A's, because the random variables X_A and $X_{A'}$, when A and A' are different persons, are not independent: if A infects B, then B cannot be overinfected by A' anymore. This problem can be ignored if we know that the probability that three persons get close to each other is very small. In a densely populated environment, this probability is certainly not negligible. Another mathematical issue to be aware of is the fact that (3) is in principle not valid if A or B moves during the time δt. So, one can

certainly not use the formula for $\delta t =$ one day. Instead, one should consider a non-homogeneous Poisson distribution, which amounts to replacing $\nu_{A\to B}\delta t$ in (3) by $\int_{t_0}^{t_0+\delta t} \nu_{A(t)\to B(t)}\,\mathrm{d}t$. In view of (4), if we assume that the viral load of A does not significantly vary in the time δt (more than one day would not be realistic), and that prophylaxis stays constant, we see that this amounts to replacing $\nu_{A\to B}$ by the effective value

$$\nu_{\text{eff}} := \nu_0 v_{\text{cont}}(A)(1-\text{prophylaxis}(A))(1-\text{prophylaxis}(B))$$

and to replacing δt by the actual time T_{eff} passed in the vicinity of A, *weighted* by the spread factor $s(r)$:

$$T_{\text{eff}} := \int_{t_0}^{t_0+\delta t} s(AB(t))\,\mathrm{d}t.$$

As a first approximation, when $s(AB) \neq 0$ we may replace AB by its mean value $AB = 2$, which gives $s(AB) = 1$ (see (5)), thus T_{eff} is simply the time spent within distance $r_{\max} = 3\text{m}$ of the infected person A.

In our program, time has to be discretized. In most applications, we choose a constant time step δt (although the program accepts variable time steps as well), and at each step we apply the `infect_world` operator. This operator considers all pairs (A, B) and performs the infection conditionally to the computed probability. (In view of the remark above, this gives an advantage to the first considered A, since this A will be able to infect more people than if it were selected later.) We take advantage of the time discretization to compute, for each individual A, the expected value of the number E_A of persons infected by A according to (6), which gives:

$$E_A = \sum_{\text{steps } i} \sum_{\substack{B\neq A \\ \text{at step } i}} p_{A\to B}(\delta t),$$

which we can compare to the *actual* number of persons infected by A in the simulation

$$R_A = \sum_{\text{steps } i} \sum_{\substack{B \text{ infected by } A \\ \text{at step } i}} 1.$$

The number R_A is nothing but the *reproduction number* of A. We define the average reproduction number in a given period $[T_0, T_1]$ to be

$$R = \frac{1}{|N_I(T_0)|} \sum_{A\in N_I(T_0)} R_A(T_1) \tag{7}$$

where $N_I(T_0)$ is the population of infectious individuals at time T_0, and $R_A(T_1)$ is the number of persons infected by A as of time T_1. Notice that, with this definition, the average reproduction number can grow with the event of multiple infections of the same individual.

5.3 Infectiousness and immunity

Recall from Section 4 that only `Healthy` and `Dormant` states can become infected, so the probability of infection will not be applied to them. For the rest, how to model the relation between infectiousness and immunity? A natural option would be to make the probability $p_{A\to B}$ actually depend on B's immunity. However, this is not what we have chosen. Recall that B's *effective* viral load is reduced by her immunity (1). If B's immunity is strong (close to one), then, even if B is infected and becomes `Latent`, she may never develop the disease. We believe that applying twice the role of immunity, once in reducing the infectiousness, and once again in reducing the viral load, would be redundant, from a biological viewpoint. This is probably debatable.

However, this procedure impacts the way of counting infected people and the reproduction number. If B is very mildly infected and never becomes symptomatic, should it be counted as a new infected person? This "subtelty", which is not considered within the SIR categories, because a `Latent` person is neither susceptible nor infectious, has a noticeable effect on the computation of the effective reproduction number R. Formula (7) counts all infections, including those who never be symptomatic, and hence will give a value that can sometimes be surprisingly high.

In future works it could be interesting to keep track of the full *infection graph*, or *transmission network*, only linking individuals that have effectively been contaminated to the point of being infective themselves.

6 Spatial motion

An interesting feature of our simulation is to endow each person with an individual *motion* function. This function takes an arbitrary (*i.e.* floating-point) time as a parameter, and returns the 2D position of the character. Thus, mathematically, it is nothing but a time parametrization of the motion. Motions can be deterministic or stochastic (using the computer's random number generator). Currently, our spatial domain Ω is a rectangle of arbitrary size. It would be interesting to limit allowed positions by using geographical data like mountains, roads, buildings, or population density maps, etc.

The following question could be important for governments to decide what strategy to adopt to limit the epidemic: strict or partial lockdowns, closing schools, etc.

Question 6.1 Without information on the health status of individuals, what characteristics of people's motion influence virus spreading in the population? ★

In order to exemplify our discussion, we shall consider the following collective motions.

- **Static motion.** No motion; the population is randomly, uniformly distributed. We can only vary the density.
- **Chaotic motion.** A piece-wise affine random walk: each individual walks straight during a random time, then makes a random turn, and continues; this models a chaotic motion with diffusive properties. Although not particularly realistic, this kind of motion is often used to justify approximations of simple situations (typically, SIR-like models, and Fisher-KPP equations) by partial differential equations in the limit of a very dense population.
- **Bound motion.** A circular motion, with random initial angle, radius, and speed. This may model a confined motion (no diffusion).

We shall also be interested in irregular distribution of motion. For instance, what happens if the population is fully locked down (no motion), except for a few individuals with a chaotic motion (we called this the "outlaw" experiment). Details of the experiments can be found in the appendix.

6.1 Contact surface

In view of the transmission probability discussed in Section 5, we are tempted to introduce the notion of "contact surface per person" ς, as follows. Take a number of persons N, all equipped with the same motion (with possible randomized parameters), and randomly distributed in our domain, thus with density $\bar{\rho} = N/S$, where S is the area of the domain Ω (in m^2). For each person A, compute the average duration of "contacts" C_A, *i.e.* the time spent within a distance less than $r_{\max}$ to another individual, divided by the total time of the simulation. The contact duration is counted "with multiplicity" and hence can be larger than 1: for instance if two persons are close to A

during time δt, then C_A includes $2\delta t$. Then define

$$\gamma_N := \frac{1}{N} \sum_A C_A.$$

More precisely, since the right-hand side is a random variable, one should define γ_N as its expected value. Thus, γ_N is the average percentage of time that each person spends close to another person (with multiplicity). There are several ways to compute or interpret this quantity.

First, let us introduce the symmetric matrix $\mathcal{C} = (C_{A,B})$ where each entry $0 \leqslant C_{A,B} \leqslant 1$ with $A \neq B$ is the relative amount of time when $AB \leqslant r_{\max}$ (AB is the distance between A and B). By convention we take $C_{A,A} = 0$. Then C_A is the A'th coordinate of the vector $C := \mathcal{C} \cdot (1, 1 \ldots, 1)$, and γ_N is related to the L^1-norm of the matrix $\mathcal{C}$:

$$\gamma_N = \frac{1}{N} \sum_{A,B} C_{A,B} \tag{8}$$

and hence $0 \leqslant \gamma_N \leqslant N - 1$. The extreme case $\gamma_N = N - 1$ corresponds to the situation where all persons are constantly in contact with all others.

If we define $\delta_{A,B}$ to be the characteristic function $1_{AB \leqslant r_{\max}}(A, B)$ (*i.e.* $\delta_{A,B} = 0$ if the distance $AB > r_{\max}$ and 1 otherwise), then, for a simulation of time T,

$$C_{A,B} = \frac{1}{T} \int_0^T \delta_{A(t),B(t)} \, \mathrm{d}t,$$

and hence (by the Fubini Theorem, *i.e.* just swapping sum and integral),

$$C_A = \frac{1}{T} \int_0^T \sum_{B;B \neq A} \delta_{A(t),B(t)} \, \mathrm{d}t.$$

If the population is static (or if T is small), we see that γ_N is simply the number of pairwise contacts between two different persons. In general, it can grow with N, but it is reasonable to infer that γ_N is controlled by the density of the population. We already see a first indication of this with the inequality $\gamma_N/\bar{\rho} \leqslant (N-1)/\bar{\rho} \leqslant S\frac{N-1}{N} < S$. This leads to the following definition of "contact surface per person" ς:

$$\varsigma(N) := \frac{S\gamma_N}{N}, \qquad 0 \leqslant \varsigma < S.$$

Coming back to γ_N itself, we see that it is natural to express it in terms of population density. Let us define the local density:

$$\rho(A, r) := \frac{\#\{\text{persons within distance } r \text{ of } A\}}{\text{area of the disc } B(A, r)},$$

in other terms:

$$\rho(A,r) = \frac{1}{\pi r^2} \sum_B \delta_{A,B}, \quad C_A = \frac{\pi r_{\max}^2}{T} \int_0^T \rho(A(t), r_{\max})\, \mathrm{d}t - 1.$$

We now make our first assumption on the motion: namely that for each t, the points $B(t)$ are independent, identically distributed in the domain, with uniform distribution. Hence, by Monte-Carlo approximation,

$$\frac{S}{N-1} \sum_{B;B\neq A} \delta_{A(t),B(t)} \simeq \int_\Omega 1_{d(A(t),x)\leqslant r_{\max}}\, \mathrm{d}x,$$

where $\mathrm{d}x$ is the surface element in $\mathbb{R}^2$. If we neglect the boundary terms, which will contribute to a lower order when $\Omega \gg r_{\max}^2$, $\int_\Omega 1_{d(A(t),x)\leqslant r_{\max}}\, \mathrm{d}x \simeq \pi r_{\max}^2$, and hence

$$C_A \simeq \frac{1}{T} \int_0^T \frac{N-1}{S} \pi r_{\max}^2\, \mathrm{d}t = \frac{N-1}{S} \pi r_{\max}^2;$$

which does not depend on A anymore (again, up to our approximations, of course). Hence

$$\gamma_N = \frac{1}{N} \sum_A C_A \simeq \frac{N-1}{S} \pi r_{\max}^2,$$

which gives, very roughly,

$$\varsigma \simeq \frac{N-1}{N} \pi r_{\max}^2 \simeq \pi r_{\max}^2,$$

where N is large.

The first conclusion of this small computation is that, at first order, the "contact surface" ς is not a very interesting dynamical invariant, at least in the case where the population distribution is uniform, and if the collective motion of individuals does not destroy this uniformity. Although the Fubini argument looks like ergodicity, notice that we would have obtained the same result, obviously, if there were no motion at all. The approximation $\varsigma \simeq \pi r_{\max}^2$ should hold for all three examples above: static, chaotic or bound motions, provided the initial population distribution is uniform. A second conclusion, which may seem evident, it that before looking at fine characteristics of motion, the main factor governing the number of contacts is the population *density*; indeed, $C_A \sim \bar{\rho} \pi r_{\max}^2$. Finally, this suggests that motion will play a more interesting role in situations where the population is *not*

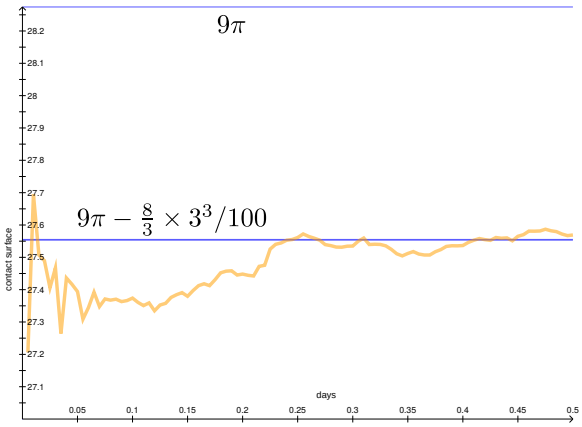


Figure 5: The orange graph is the numerical computations of the contact surface $\varsigma(T)$, for a population initially randomly uniformly distributed, and evolving from $T = 0$ to $T = 0.5$ day. Here the domain Ω is a square of size 100×100, and $r_{\max} = 3$. The whole population moves with a "chaotic" motion. In our discussion in Section 6.1 we have used the rough approximation $\varsigma \simeq \pi r_{\max}^2$ (top horizontal line in the figure). But, because of the boundary of the domain Ω, the correct value should be strictly less. One can directly approximate the sum (8) by the quadruple integral $\frac{N-1}{S^2} \int_{\Omega\times\Omega} 1_{d(x,y)\leqslant r}\, \mathrm{d}x\, \mathrm{d}y$. In the case of a square domain of area S, one has $\varsigma \simeq \frac{1}{S} \int_{\Omega\times\Omega} 1_{d(x,y)\leqslant r}\, \mathrm{d}x\, \mathrm{d}y = \pi r^2 - \frac{8}{3} r^3/S^{1/2} + \mathcal{O}(r^4/S)$. This gives the second horizontal line in the picture.

uniformly distributed on the territory. A contact surface $\varsigma > \pi r_{\max}^2$ suggests a population with locally high density. If, when time advances, ς gets closer to $\pi r_{\max}^2$, we have a sign that the motion scatters the population evenly. We shall study further such diffusive effects in the next section.

6.2 Social curve

For each individual A, we now define $F_A(t)$, the number of *distinct* persons met by A during the time interval $[0, t]$, where by 'met' we mean 'at distance less than $r_{\max}$. Like the contact surface defined in Section 6.1, this quantity is purely spatial and dynamical; it does not depend on any medical parameter. Then we consider the average

$$F(t) := \frac{1}{N} \sum_A F_A(t), \tag{9}$$

where N is the total population. Contrary to the contact surface, it is clear that a static population has a much smaller social number F than a fully

chaotic curve: in the static case, assuming that the population is uniformly distributed, we have (neglecting boundary effects)

$$\forall A, \qquad F_{A,\text{static}}(t) \simeq F(t) \simeq N\frac{\pi r_{\max}^2}{S} \text{ for all } t.$$

On the other hand, in the chaotic case, the whole population will be visited by each individual A:

$$\forall A, \qquad \lim_{t\to\infty} F_{A,\text{chaotic}}(t) = N - 1.$$

It would be very interesting to have a good mathematical understanding of the curve $t \mapsto F(t)$. Let us briefly discuss how to compute $F_A(t)$ when all persons have a chaotic motion, except for A, who stays still. It does not seem to us completely obvious to say how the asymptotic growth of $F_A(t)$ depends on the position of A inside the domain Ω, although, intuitively, a position close to the boundary should induce a lower increase rate, because people trying to "get out" of Ω are being reflected by the boundary, and hence A is more likely to meet them several times, instead of meeting new unknown persons. It turns out that this can be made explicit by a spectral analysis.

At each time t, we have to monitor the flux of persons entering and leaving the disc of radius $r_{\max}$ centered at A. All persons entering this area become "met by A" (one could say "infected", assuming in this case that the probability of infection is one in the disc). The increase of $F_A(t)$ equals the number of persons *not previously met by* A entering the disc. Inside the disc, all persons are already met by A, and hence their density ρ_A is equal to the global, uniform density $\bar{\rho} = N/S$ (strictly speaking, we should take $\frac{N-1}{S}$ since A itself should be excluded). Hence, the outgoing flux of persons out of the disc is constant, and easy to estimate. On the contrary, the crucial quantity, which is the incoming flux of "not already met by A" persons, is not obvious to determine. It depends on the spatial variation of the density ρ_A at the boundary of the disc, and we believe that this is not a local quantity: it depends on how persons are able to "escape" away from the disc $B = B(A, r_{\max})$. Initially, at $t = 0$, $\rho_A = \bar{\rho} > 0$ in B, while $\rho_A = 0$ outside of B. However, as soon as $t > 0$, and people start to escape from B, people outside of B but very close to it should enjoy the same density of population: ρ_A is now continuous. Therefore, the incoming flux of persons "not met by A" cannot depend only on the value of ρ_A on ∂B, since the net flux (incoming minus outgoing) should be high for small t, and tend to zero when $t \to \infty$ (when everybody in our domain was already met by A). A PDE (partial differential equation) description of this diffusion problem is a good way to

understand these issues. Let $\check{\Omega} = \Omega \setminus B$. Let $\rho_A(t,x)$ be the density of persons met by A at position x. Thus the total number of persons met by A outside the disc B is

$$\check{F}_A(t) = \int_{\check{\Omega}} \rho_A(t,x)\,\mathrm{d}x = F_A(t) + 1 - \bar{\rho}\pi r_{\max}^2,$$

where $\bar{\rho} = N/S$, and the (negligible) "+1" is to account for the individual A itself. The density $\rho_A(t,x)$ is subject to the following equation:

$$\begin{aligned} \partial_t \rho_A - D\Delta_x \rho_A = 0, &\quad (t,x) \in (0,\infty) \times \check{\Omega} \\ \rho_A(0,x) = 0, &\quad x \in \check{\Omega} \\ \rho_A = \bar{\rho}, &\quad (t,x) \in (0\infty) \times \partial B \\ \partial_n \rho_A = 0, &\quad (t,x) \in (0\infty) \times \partial\Omega. \end{aligned} \tag{10}$$

The first one is the standard diffusion (or heat) equation with diffusion coefficient D, which depends on the parameters of our random walk. The second line is the initial condition at $t = 0$. The third line is the boundary condition on ∂B, while the last one, where ∂_n denotes the outgoing normal derivative, expresses that no individual can escape the domain Ω. If we replace ρ_A by $\tilde{\rho}_A(t,x) := \rho_A(t,x) - \bar{\rho}$ we obtain a standard heat equation with homogeneous Dirichlet and Neumann boundary conditions, and non-homogeneous initial condition $-\bar{\rho}$. A common way of solving this problem is to consider the Laplace operator $\Delta_{\check{\Omega}}$ on the domain $\check{\Omega}$ with the mixed Dirichlet and Neumann boundary conditions on the two boundaries ∂B and $\partial\Omega$. Since, one each boundary, one of $\tilde{\rho}_A$ and $\partial_n \tilde{\rho}_A$ must vanish, $\Delta_{\check{\Omega}}$ is self-adjoint and non-negative. By ellipticity, the spectrum of $\Delta_{\check{\Omega}}$ is discrete and consist of a non-decreasing sequence of eigenvalues $(\lambda_j)_{j\in\mathbb{N}}$ tending to infinity. Hence we can, in principle, find a complete orthonormal basis of eigenfunctions u_j, $j \in \mathbb{N}$ in the appropriate Hilbert space, which is dense in $L^2(\check{\Omega})$. Hence we may express the initial condition as a sum

$$-\bar{\rho} = \sum_j a_j u_j(x), \quad a_j := -\rho\langle 1, u_j\rangle$$

and finally obtain

$$\tilde{\rho}_A(x,t) = \sum a_j e^{-D\lambda_j t} u_j(x).$$

Therefore

$$\begin{aligned} F_A(t) &= \int_\Omega \rho_A(t,x)\,\mathrm{d}x - 1 = N - 1 + \int_{\check{\Omega}} \tilde{\rho}_A(t,x)\,\mathrm{d}x \\ &= N - 1 + \sum a_j e^{-D\lambda_j t} \int_{\check{\Omega}} u_j(x)\,\mathrm{d}x, \\ &= N - 1 + \sum_j a_j e^{-D\lambda_j t} \int_{\check{\Omega}} u_j(x)\,\mathrm{d}x, \\ &= N - 1 - \bar{\rho} \sum_j e^{-D\lambda_j t} \left(\int_{\check{\Omega}} u_j(x)\,\mathrm{d}x \right)^2, \end{aligned} \tag{11}$$

which can be written equivalently as

$$F_A(t) = \bar{\rho}\pi r_{\max}^2 - 1 + \bar{\rho} \sum_j (1 - e^{-D\lambda_j t}) \left(\int_{\check{\Omega}} u_j(x)\,\mathrm{d}x \right)^2$$

To conclude this mathematical apparté we see that the increase of $F(t)$ up to the maximal value $N-1$ does not follow an simple exponential law, but a sum of exponentials. For large t, we should be able to get a good approximation of the increase rate by selecting the smallest eigenvalue $\lambda_0 > 0$ (note that $\lambda_0 \neq 0$ because the only constant solution to (10) is zero). Moreover, by Courant's theorem, the corresponding eigenfunction u_0 does not change sign, which implies that the constant $c_0 = \langle 1, u_0 \rangle = \int_{\check{\Omega}} u_0(x)\,\mathrm{d}x$ does not vanish. Hence we may conjecture the following asymptotic behaviour, for large t:

$$F_A(t) \sim N - 1 - \frac{N c_0^2}{S} e^{-D\lambda_0 t}. \tag{12}$$

Question 6.2 Prove (12) (when N and t are large). ★

We can now compare the asymptotic behaviours of the "social number of A", $F_A(t)$, for different initial positions of A. We used for this a numerical solver (FreeFEM++) to compute λ_0 in the following situations. Ω is a rectangle of size 100×100 with lower left corner at position $(0,0) \in \mathbb{R}^2$, and $r_{\max} = 3$. It is obvious from the table in Figure 6 that the position of A greatly influences the value of the lowest eigenvalue λ_0. If A is close to a corner, the growth of $F_A(t)$ will be much slower; the maximum growth rate seems to be at the center. The numbers are confirmed by the population simulations, see Figure 7. From the epidemiological viewpoint, we see that infected persons should not stand at the center of a crowd, because they will likely meet much more people than if they we standing along a wall, or

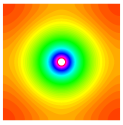
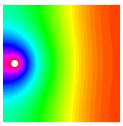
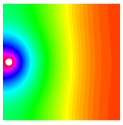
Position of A	λ_0	c_0	Eigenfunction
$(50, 50)$	0.0002771325491	98.45152348	
$(10, 50)$	0.0001657424295	95.96970218	
$(5, 50)$	0.0001427185441	96.82151423	
$(10, 10)$	0.0001072342178	96.75283201	
$(5, 5)$	8.261371309e-05	97.86664575	

Figure 6: First eigenvalue of the Laplace problem associated with (10), numerically computed with FreeFEM++.

even better, a corner of the room, and this holds even if other persons are "bouncing" on the walls.

Coming back to the question of the incoming flux of "infected" individuals, we see that we have to integrate $\partial_n \rho_A$ on ∂B, that is to say we want to recover the Neumann boundary condition from the Dirichlet condition that was imposed in the global Cauchy problem (10) (although the fact that we only need the integral might simplify the problem). Equation (11) suggests that the incoming flux for $t > 0$ is

$$\frac{\mathrm{d}F_A}{\mathrm{d}t}(t) \simeq \bar{\rho} D \sum_j \lambda_j e^{-D\lambda_j t} \left(\int_{\check{\Omega}} u_j(x)\, \mathrm{d}x \right)^2,$$

which is not well-defined at $t = 0$, as expected (for small t, one can ignore the boundary $\partial\Omega$ and solve the heat equation on $\mathbb{R}^2 \setminus B$, which gives a singularity of the flux of type $1/\sqrt{t}$, see [cf H. S. Carslaw, J. C. Jaeger, Conduction of Heat in Solids [2 ed., 1959] p335-336])

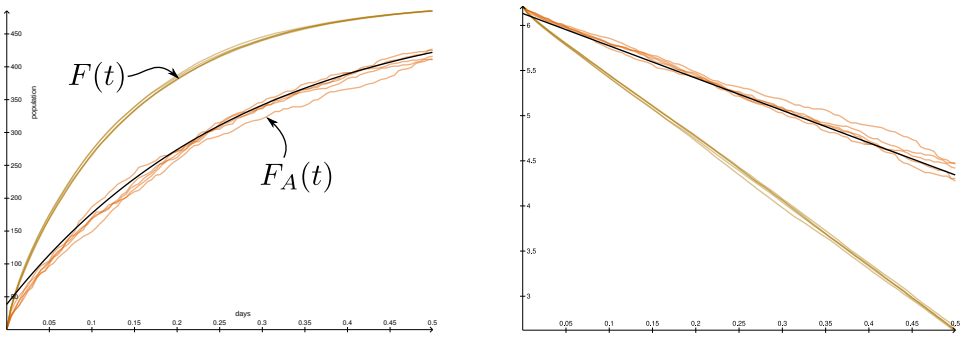


Figure 7: Five superposed simulations of the Social Number $F_A(t)$ (Section 6.2), with 500 persons in a square of size 100×100, when A is static at position $(10, 50)$, and all other persons move according to a random walk with diffusion coefficient $D = 1/4$. The average $F(t)$ is defined in (9). The conjectural theoretical approximation for large times, defined in (12), is displayed here in black, with $\lambda_0 \simeq 1.66 \times 10^{-4}$ (see Figure 6). The figure on the right presents the same data in a logarithmic scale (more precisely, mapped by the function $x \to \ln(N - 1 - x)$). Currently, we don't know how to express the growth rate of $F(t)$, see Question 6.4

Remark 6.3 Equation (10) is exactly the heat equation governing the temperature in a 2D rectangular plate (the domain Ω) with a circular hole in the middle (the disc B), insulated at $\partial\Omega$, and with imposed temperature at ∂B. $\triangle$

Question 6.4 What is the correct growth rate for $F(t)$ or $F_A(t)$ when A moves (*i.e.* performs the same random walk as the rest of the population)? Can one find a diffusion equation similar to (10)? The difficulty seems the be that we have to consider a moving boundary. However, considering the average $F(t)$ instead of $F_A(t)$ might simplify the problem. Numerics show that the growth rate is significantly higher than in the case of static A, see Figure 7. ★

Question 6.5 The numerical evidence that the asymptotics of $F_A(t)$ are governed by the eigenvalues of the Laplace operator (*eg.* Figure 7) was obtained when the population (except A) moves according to a simple random walk, where at each time step δt, the new position is $p'_A = p_A + \vec{r}$, where $\vec{r}$ has random direction and fixed size r. This walk is approximated by a Brownian motion (or diffusion process) of diffusion coefficient $D = r^2/4\delta t$. If we use the more general motion described above, which is a jump process, where we move on a line segment of (possibly random) size r for a (possibly random) time δt before performing a random turn, then there is a noticeable

discrepancy with the solution to (10) with $D = \langle r \rangle^2 / 4\langle \delta t \rangle$. How to explain this? ★

We don't know how to theoretically handle the intermediate cases like the "bound" motion; we will instead rely on the numerical results. In order to compute $F_A(t)$, one can simply use the main program, setting the initial population to `Healthy`, and "infect" any `Healthy` person with probability one as soon as they are within distance $r_{\max}$ of A. The infected person gets a new state, for instance `Sick`, in order not to be counted twice.

However, this strategy does not handle the computation of the average $F(t)$, because a person "met" by A should remain available for being later "met" by another person A'. Thus, we need to equip each person with a new data, the set of all encountered persons, which we update at each time step. This was used to produce Figure 7.

6.3 Population conductance

We consider here the spatial propagation speed, a dynamical quantity that now depends on the complete model, with infectiousness. It is sometimes called the "speed of spreading" of the epidemic wave, see for instance the recent article [1], which studies a SIR-like PDE model with spatial diffusion. The unusual name "population conductance" is inspired by electric conductance, and Ohm's law alluded to in the introduction. Take a domain which is a thin rectangle: $\Omega = [0, L] \times [0, \epsilon]$, where $L >> \epsilon$. Suppose that the whole population is `Healthy` at $t = 0$; add a small number N_0 of infected individuals on the left hand-side of the rectangle ($x \simeq 0$). What will happen? If the virus spreading is much faster than recovery, the whole rectangle will soon be infected. But if recovery is also fast, or if L is very large, we will observe a *front* of infected people (like a wave front), because as the head of the front becomes infected, the tail will recover. If recovery is faster than infection, the epidemic front should disappear. Using our program, one can numerically test these output parameters: speed and width of the wave front with various collective motions.

7 Operators

The program runs as a discrete dynamical system acting on the total population. Since we wish here to put a strong emphasis on spatial properties, it is natural to consider the population as a spatial distribution $\Psi(x, y)$ of persons. Of course, this distribution is discrete (sum of Dirac functions) but

when the total number of persons is large, it is useful to consider a continuous, approximate population density.

Our implementation consists in applying successively three operators.

1. A point-wise, time evolution operator $\mathcal{T}$. It acts separately on each person, considered in isolation, and update the temporal parameters: evolution of the viral load and immunity, and consequently performs the passage from one state (or compartment) to another. For instance, if a person is infected, then the viral load increases, and after crossing thresholds, the person becomes `Sick`, or `Critical`, etc. See Section 2. It is of course a *local* operator, because it does not change the position of each individual, and the evolution of the individual's parameters does not depend on other individuals.

2. A point-wise spatial propagation operator $\mathcal{P}$. This operator performs the motion of all individuals, but does not change any medical parameter. It a non-local operator, but acts point-wise: in our simulation, the motion of an individual is not influenced by the other persons. Hence it should not be thought of as a global transformation (diffeomorphism, etc.); on the contrary the trajectory of different persons may cross in arbitrary ways. This operator can be used to perform special time-dependent actions like 'going to bed', 'going to work', etc.

3. A non-local, infection operator $\mathcal{I}$. For each individual A, this operator considers all neighboring persons, and perform conditional infection, as described in Section 5. If the domain is large enough, this operator can be considered as "pseudo-local", because only persons B at distance $d(A, B) \leqslant r_{\max}$, which is small compared to the size of the domain, are involved in the conditional infection of A. In many cases this operator will be translation invariant (it will not depend on the position of A, but rather on the relative positions $B - A$); hence it can be viewed as a convolution operator, if we ignore boundary effects.

Note that the order of application of $\mathcal{T}$ and $\mathcal{P}$ is irrelevant: $[\mathcal{T}, \mathcal{P}] = 0$. However, neither $\mathcal{P}$ nor $\mathcal{T}$ commutes with $\mathcal{I}$.

8 Results

By varying population density, age distribution, motions of individuals, etc. one can simulate very different situations, including governments decisions of full of partial lock downs, school closing or opening again, etc. In this article,

we only start a few experiments, while more sophisticated ones are still under investigation.

8.1 Virus spreading in a random crowd

We simulate a crowd of random people, walking erratically in random directions, without stopping, for a few hours. This could represent a concert, a shopping center, a street demonstration, etc.

The domain is a square of 100m by 100m, and we vary the population from 10 to 10000 persons. Age is chosen randomly according to France's 2020 population distribution [5]. Motion is a random walk (see also Section 6.3), where each individual walks straight during 5 to 15 seconds, then makes a random turn, and continues. The distance of the straight line segment and the speed is fixed for each person (from 0 to 2 meters per second), but varies randomly from one person to another.

The time step for the simulation is $\delta t = 1$s, and we simulate 3 hours spent in the crowd. Because the whole setting is random and *iid*, the geographical map is not so interesting, so we shall only give here the time curves of each health state. At the start of the simulation, everybody is `Healthy`, except for 10% that are infected at a random stage, and placed randomly in the square.

Under appropriate assumptions on randomness, It is easy (and well known) to mathematically predict what will happen. Indeed, during the short time of the simulation, infected people are not infectious themselves. So, the number N_I of infectious persons stays constant. Moreover, we may assume that the viral charge of all individuals can be considered constant, and hence the whole evolution becomes a Markov chain. Let us compute $E_A([0,t])$, the number of persons contaminated by $A \in N_I$ (we use the same notation N_I for the set of infectious persons). By (6) and (3), the infinitesimal variation of E_A is

$$\frac{\mathrm{d}}{\mathrm{d}t} E_A([0,t]) = \lim_{\delta t \to 0} \frac{E_A(\delta t)}{\delta t} = \sum_{B \neq A} \nu_{A \to B}$$

where the sum is taken over the set of susceptible persons B. Let us assume for simplicity that the prophylaxis factor is zero for everyone. Then, using (4) we get

$$\frac{\mathrm{d}}{\mathrm{d}t} E_A([0,t]) = \nu_0 v_{\mathrm{cont}}(A) \sum_{B \neq A} s(AB).$$

Now the analysis is similar to Section 6.1, taking into account that the set where B varies has cardinal $N_S(t) := N - I(t)$, where $I(t)$ is the number of infected (but not yet infectious) individuals, and N is the number of initially

susceptible people, *i.e.* $I(0) = 0$. If we assume that, at each time t, the susceptible persons $B \in N_S$ are randomly uniformly distributed in the domaine Ω (this may hold only if the mixing strength of the motion — or the diffusion coefficient — is high enough), then the Monte-Carlo scheme says that the expectation of $\frac{S}{N-I(t)} \sum_{B \in N_S(t)} s(AB)$ is the integral $\int_{x \in \Omega} s(d(A,x))\,\mathrm{d}x$, which, if A is not too close to the boundary, is equal to

$$\sigma_0 := \int_{\mathbb{R}^2} s(r) r \,\mathrm{d}r\,\mathrm{d}\theta = 2\pi \int_0^{r_{\max}} r s(r)\,\mathrm{d}r.$$

In the case of Section 6.1, the factor s was not taken into account, and we integrated the constant 1, yielding the surface of the contact ball $\pi r_{\max}^2$. Here instead, the area is weighted by s. With the formula chosen for s (5), the integral can be explicitly computed; an approximate value (with $r_{\max} = 3m$) is $\sigma_0 \simeq 50.44\mathtt{m}^2$. We now have

$$\frac{\mathrm{d}}{\mathrm{d}t} E_A([0,t]) = \nu_0 v_{\mathrm{cont}}(A) \frac{\sigma_0}{S}(N - I(t)). \tag{13}$$

Since $I(t) = \sum_{A \in N_I} E_a([0,t])$, we can let

$$\beta := \nu_0 \frac{\sigma_0}{S} \sum_{A \in N_I} v_{\mathrm{cont}}(A), \tag{14}$$

and we obtain the simple ODE Cauchy problem

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} I(t) = \beta(N - I(t)) \\ I(0) = 0 \end{cases}$$

which gives

$$I(t) = N(1 - e^{-\beta t}). \tag{15}$$

So, in this regime, our model merely reproduces the simplest possible model with exponential behaviour. Of course, the factor β is primarily important; depending on it, after the simulation has finished, either a small fraction, or the quasi-totality of the population will be infected.

For instance, we have simulated a crowd of 3000 people with $N = 2700$ and $N_I = 300$, where the 300 infective individuals were all `Carrier` but with slightly different viral charges, leading to a theoretical $\beta = 18.937\mathrm{day}^{-1}$ using (14). Figure 8 shows that the simulated curve of $I(t)$ is in agreement with the "theoretical" exponential curve (15). In this case, after 3 hours (0.125 day), 2451 persons have been contaminated, *i.e.* 91% of the initially

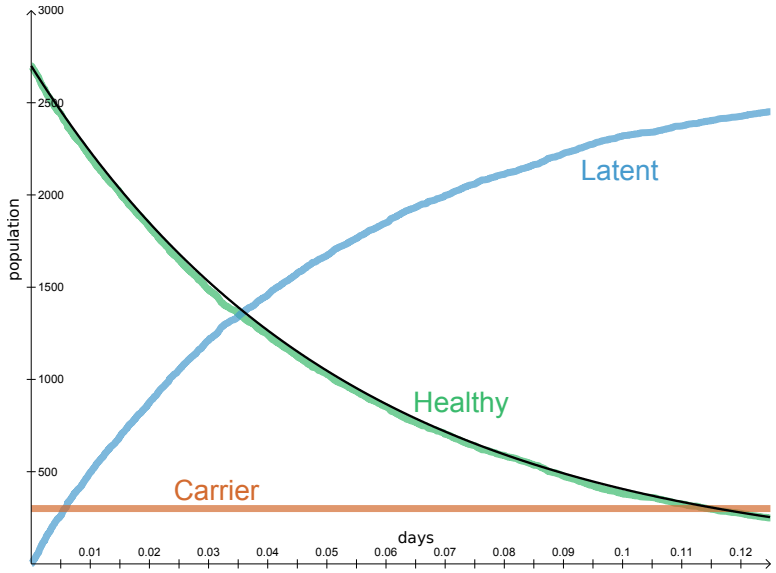


Figure 8: Randomly moving crowd simulation, see Section 8.1. 3000 persons in $10000\mathtt{m}^2$ during 3 hours. The color curves are the result of the experiment. The black curve is the theoretical prediction $Ne^{-\beta t}$ (14) (15).

healthy population. If we assume that the percentage of infective individuals is roughly constant in the population (here, 10%), then β is proportional to N. Hence if $N = 270$, we obtain a proportion of $1 - e^{-0.125\beta/10} \simeq 21\%$. We can have 3000 persons going to the shopping-center by successsive groups of 300, this will lead to $10 \times 270 \times (1 - e^{-0.125\beta/10}) \simeq 569$ contaminations, which of course much better than the 2451 contaminations if the 3000 persons go at the same time.

Remark 8.1 Had we instead imposed a *confined* motion (like the bound motion introduced above), the situation would be quite different. For a confined motion, at $t = 0$, formula (13) is still valid, so the initial decay rate βN is the same as in the chaotic case. But for $t > 0$ there are two main difficulties. First, each infective individual A can only reach a fraction F_A of the population, see Section 6.2. So the "limit" number of infected people (in this regime of small t, of course) is not N but is bounded by $\sum_{A \in N_I} F_A$. Second, we cannot assume that the distribution of susceptible persons B stays uniform. The susceptible population close to A decays faster, while distant persons will remain unaffected.

A mathematically tractable case is the *static* case (no motion), under the assumption that the few infectious persons A are far enough from each other,

so that we can separate the reachable populations: in this case, the decay of the susceptible population "controlled by A" (or rather, its expected value) depends only on the distance $r = AB$. Let us introduce $\rho_A(r,t)$, the *density* of persons infected by A at distance r; in other words,

$$E_A[(0,t)] = 2\pi \int_0^{r_{\max}} \rho_A(r,t) r \, \mathrm{d}r.$$

Restricting B to a thin annulus of radius r around A, and applying (6), we obtain, similarly to (13):

$$\partial_t \rho_A(r,t) = \nu_0 v_{\mathrm{cont}}(A) s(r) (\bar{\rho} - \rho_A(r,t))$$

where $\bar{\rho} = N/S$ is the initial density of susceptible individuals, which is uniform. Thus, we may introduce the radial exponential rate

$$\beta_A(r) := \nu_0 v_{\mathrm{cont}}(A) s(r)$$

and we have

$$\rho_A(r,t) = \frac{N}{S} \left(1 - e^{-\beta_A(r)t}\right).$$

Therefore,

$$E_A[(0,t)] = \frac{2\pi N}{S} \int_0^{r_{\max}} \left(1 - e^{-\beta_A(r)t}\right) r \, \mathrm{d}r. \tag{16}$$

(See Figure 9) Letting t be large, we recover that the maximum number of persons infected by A is $\frac{\pi r_{\max}^2}{S} N$, in accordance with Section 6.2. As a double check, we compute the initial slope $\frac{\mathrm{d}}{\mathrm{d}t} E_A([0,t])_{t=0} = \frac{2\pi N}{S} \int_0^{r_{\max}} r \beta_A(r) \, \mathrm{d}r = \nu_0 v_{\mathrm{cont}}(A) \sigma_0 N/S$, same as in (13). △

8.2 The aircraft carrier experiment

Suppose an aircraft carrier stops for a couple of days at a harbour, where soldiers can meet their family, and then leaves port for one month, with no contact with the outside world. After this period, at least 50% of the 2000 crew members are tested positive to the Covid-19. What happened? Such a situation is very interesting from the epidemiological viewpoint, because the ship is a relatively small, confined world, with homogeneous population, in complete isolation.

When a person becomes `Sick`, he or she stops moving, and resumes motion when becoming `Recovered`.

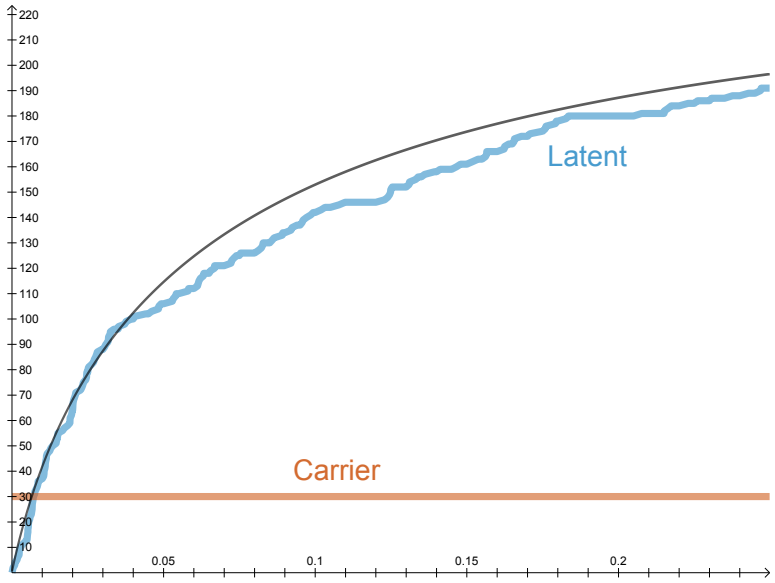


Figure 9: Infected curve for a static crowd. Simulation for a crowd of 3000 people including 30 infective persons. The curve first follows closely the theoretical prediction (16), but at some point the rate becomes slightly slower. This can be explained by the fact that the 30 infective persons have a random initial position, and several of them have overlapping contact surface (see Figure 12), while the theoretical curve assumes that the population groups reached by each infected individual are pairwise distinct.

8.3 Subway train

8.4 University classroom

8.5 Small city

We try here (very crudely) to implement a simulation of a city being infected by a few people, with and without lock-down.

References

[1] H. Berestycki, J.-M. Roquejoffre, and L. Rossi. Propagation of epidemics along lines with fast diffusion, 2020.

[2] A. Camacho and B. Cazelles. Does homologous reinfection drive multiple-wave influenza outbreaks? accounting for immunodynamics in epidemiological models. *Epidemics*, 5:187–196, 12 2013.

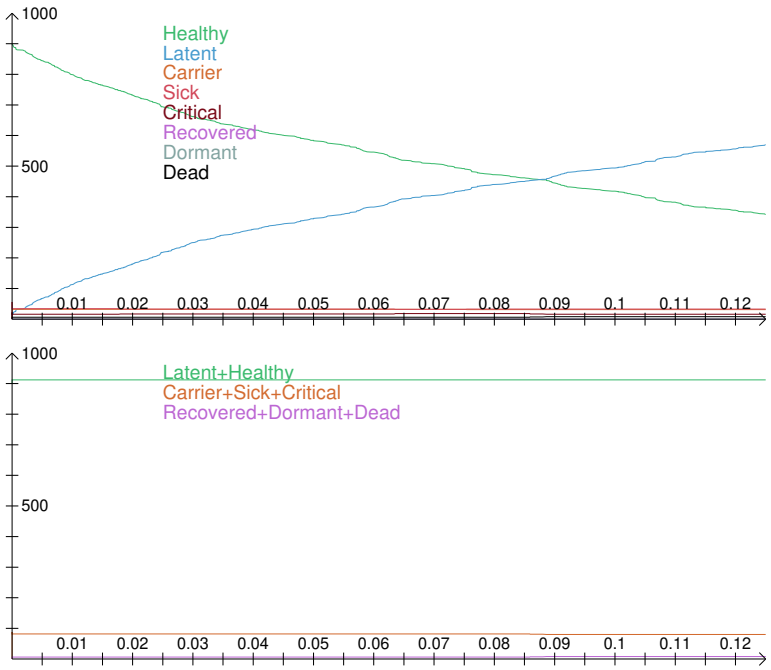


Figure 10: The crowd simulation: time plots. The top plot contains the population curves for all 8 states: number of `Healthy`, of `Latent`, etc. The bottom curves has cumulative states, as indicated, closer to the traditional SIR categories (susceptible, infected, removed). (But the "removed" category is here not the usual one, since `Recovered` people can still infect others, and `Dormant` ones may be re-infected.) We see that in this unrealistic very dense situation, the whole population is infected after about 3 days. After day 23, some people are completely recovered and hence become `Healthy` again — and susceptible: a few of them become infected for a second time, but due to the strong immunity that we have assumed, all of them will heal soon.

[3] N. C. Grassly and C. Fraser. Mathematical models of infectious disease transmission. *Nature Reviews Microbiology*, pages 477–487, 2008.

[4] C. Hadjichrysanthou, E. Cauët, E. Lawrence, C. Vegvari, F. de Wolf, and R. M. Anderson. Understanding the within-host dynamics of influenza a virus: from theory to clinical implications. *Journal of The Royal Society Interface*, 13(119):20160289, 2016.

[5] INSEE. Pyramide des âges en 2020. https://www.insee.fr/fr/statistiques/fichier/2381472/demo-pop-pyram.xls.

[6] M. R. Moser, T. R. Bender, H. S. Margolis, G. Noble, A. P. Kendal, and D. G. Ritter. An outbreak of influenza aboard a commercial airliner.

American journal of epidemiology, 110 1:1–6, 1979.

[7] M. Salathé, M. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman, and J. H. Jones. A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences*, 107(51):22020–22025, 2010.

[8] L. Willem, F. Verelst, J. Bilcke, N. Hens, and P. Beutels. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015). *BMC Infect Dis.*, 17(1), 2017.

A Crowd

B Population conductance: experiments

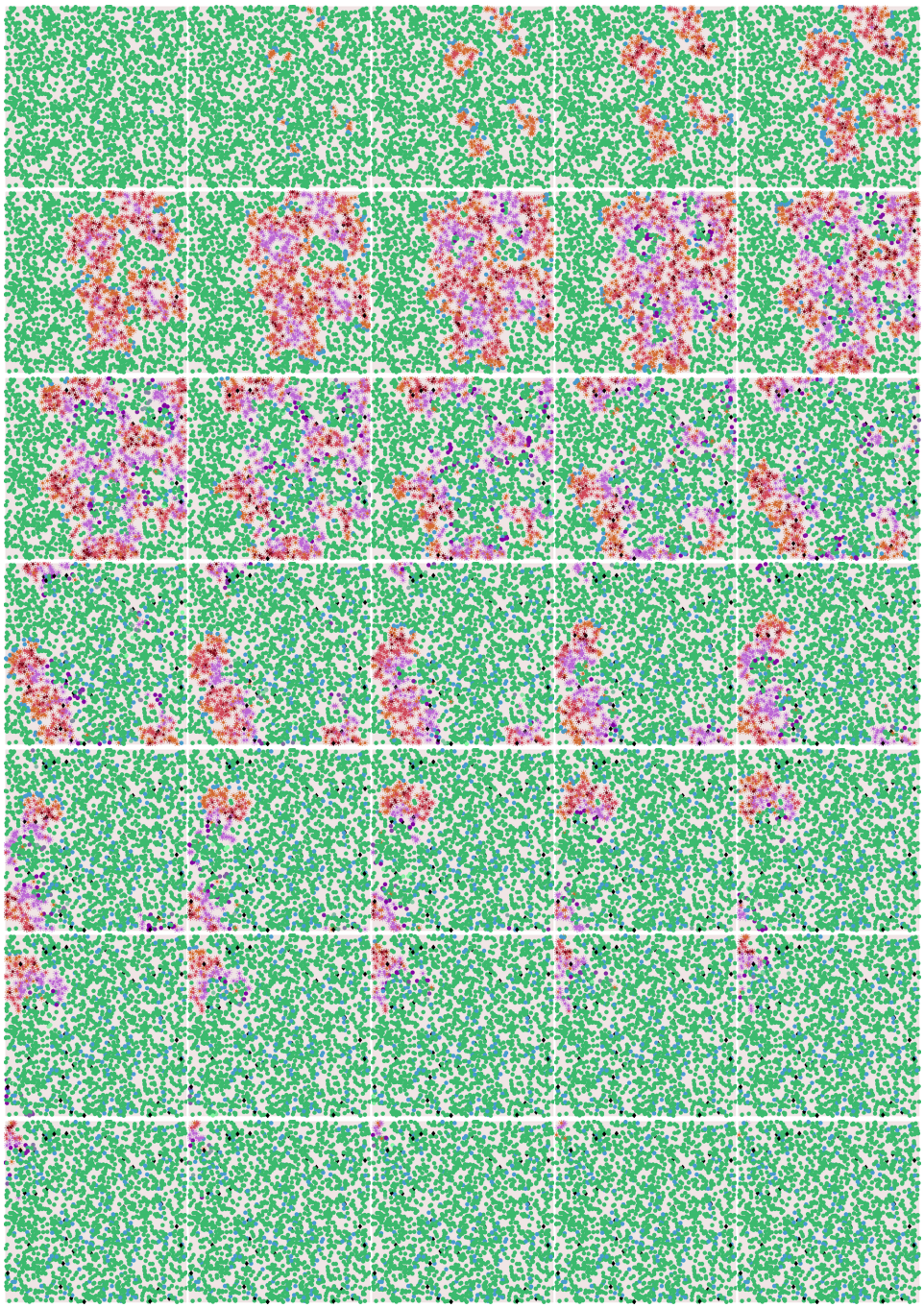


Figure 11: Static simulation. 2000 persons stay still in a box of size 100m×100m, and 10 of them are infected at $t = 0$. Time lapse between 2 images is 3.75 days.

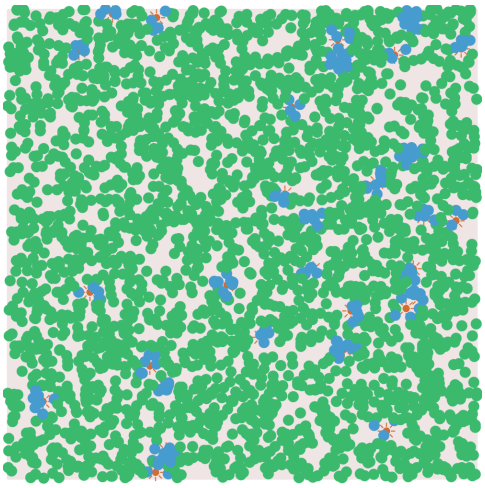


Figure 12: Static crowd simulation. 3000 static persons including 30 infective ones, after 6 hours. One can see that the 30 infective individuals have sometimes overlapping contact surfaces, leading to the curve in Figure 9.

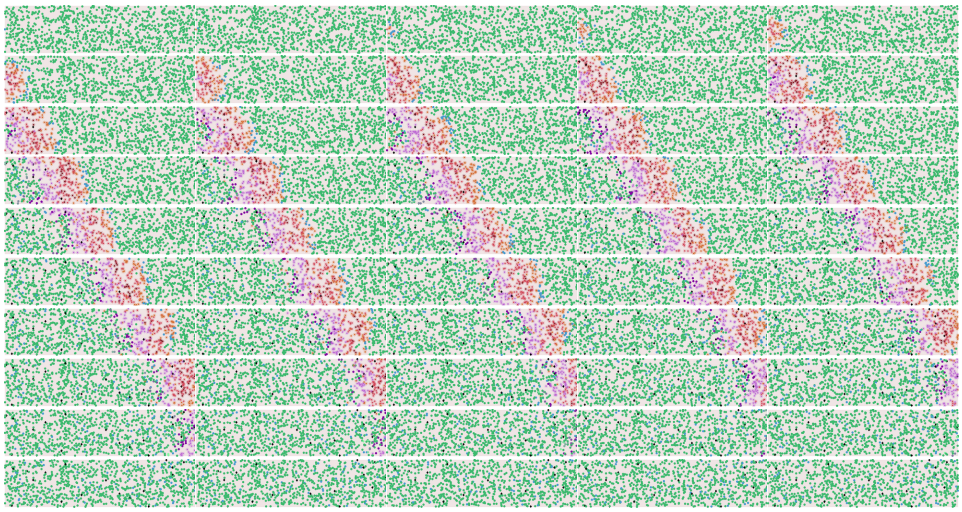


Figure 13: Bound conductance experiment. We have simulated 1000 persons in a rectangular domain of size 200m$\times$ 50m; 999 of them are initially `Healthy`, and one has just been infected, and is located on the left border, mid-height. Everybody moves with the "bound motion" (small stationary circles). The picture shows 50 images, each image is a screenshot of the simulation taken every day. Day 0 on top left, Day 49 on bottom right. After formation of the full epidemic front (about Day 30), it takes about 100 days for the front to move across the whole domain. See also Figure 14.

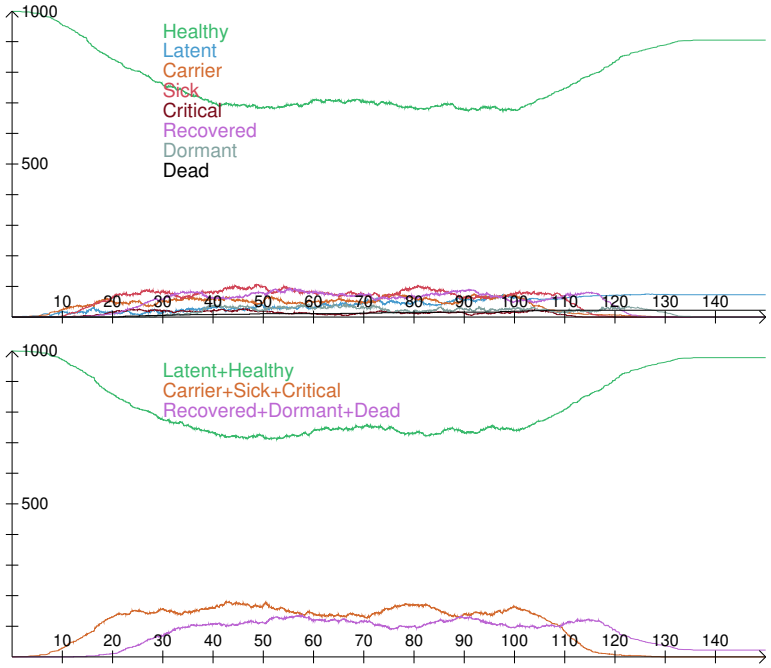


Figure 14: Bound conductance experiment. Population vs time in days. From this graphics it seems that the situation is stable between days 30 to 100, but Figure 13 shows that the cluster of infected people constantly moves.