

# Solvable delay model for epidemic spreading: the case of Covid-19 in Italy

Luca Dell'Anna

*Dipartimento di Fisica e Astronomia "G. Galilei",  
Università degli Studi di Padova, via F. Marzolo 8, 35131 Padova, Italy*

We present a simple but realistic model for describing the diffusion of an infectious disease on a population of individuals. The dynamics is governed by a single functional delay differential equation, which, in the case of a large population, can be solved exactly, even in the presence of a time-dependent infection rate. This delay model has a higher degree of accuracy than the so-called SIR model, commonly used in epidemiology, which, instead, is formulated in terms of a set of three ordinary differential equations. We apply our model to describe the outbreak of the new virus COVID-19 in Italy, taking into account the containment measures implemented by the government in order to mitigate the spreading of the virus and the social costs for the population.

## INTRODUCTION

In a very few months a viral infection called Covid-19 (Coronavirus disease 19) originated in China, breaking through the borders of all the countries, rapidly spread all over the globalized world. Italy is one of the hardest hit countries suffering from the very dramatic consequences of this disease. The outbreak of the virus, the new coronavirus which caused the infection, seems out of our control. In the absence of a therapy and a vaccine, social distancing measures and a strict lockdown appear to be the most effective means to contain the growth of the infection. We should remind that there are places in the world where often infectious diseases, also those already defeated in the so-called more developed countries, can still cause very severe consequences among the local populations.

Even if we cannot answer the question why a virus starts spreading and which is its origin, we can still wonder how it diffuses. The aim of this work is, therefore, to provide a simple handy model for epidemic spreading, which could depend only on the couple of parameters which generally characterize an infectious disease: the infection rate and the infectiousness (or recovery) time. Both these quantities can be taken from the experience, therefore, we do not need further parameters to fit the data which could cause artificial predictions. We will show that the model we are presenting have the same, or even higher, predictive power than that of one of the most widely used technique in epidemiology, the SIR model [1–3]. This latter model requires the presence of a fictitious recovery rate related to the number of recovered persons, without considering that the new cases of recovery (and fatality) come from infected cases occurring a period of time earlier. The model we are proposing, instead, is based on the fact that the closed cases come from the infected ones after an average delay recovery time, therefore, contrary to the SIR model, formulated in terms of a set of three ordinary differential equations, it is described by just a single equation, a functional retarded differential equation, bringing predictions more under control. In this work we derive the exact analytical solution of this model in the limit of large population number, also in the presence of a time-dependent infection rate, which is the case when containment measures are implemented in order to reduce the spreading of the infection. Moreover, the definition of the so-called basic reproduction number  $\mathcal{R}_0$  (a parameter determining whether a infectious disease can spread or not) comes out naturally in our delay model.

We finally apply this technique to give a quantitative description of the diffusion of Covid-19 in Italy, showing the current scenario based on the actual situation and what would have happened without the containment measures. Generally it is quite difficult to give a reliable forecast on the fate of the epidemic spreading because it heavily depends on individual and social behaviors, on the effectiveness of the containment measures already implemented, or that will be taken, by the government and on the future political decisions. At the time being, even if the situation in Italy is improving, it seems that more efforts are needed in order to change course and rapidly stop the spreading of the disease. Further measures might be useful, like, for instance, i) running more diagnostic tests, at least, on all the doctors and medical workers who are in contact with many patients, ii) improving the food distribution to avoid the crowding in the food shops and to ensure subsistence goods also to those who need, iii) providing medical devices like surgical masks to all the population.

As last remark, we remind that the outbreak of Covid-19 has been declared a pandemic by the World Health Organization. Many countries are already heavily overwhelmed by this infection and by the risk for the public health, therefore, in a networked world we all have to behave and operate with an improved spirit of cooperation. The bitter lesson imparted by this tough situation is that we cannot save ourselves alone.

## THE MODEL

Let us first consider the case in which a population of individuals, subjected to an infection, is not too large or the infection is such that the recovery time for an infected person is sufficiently long. In this conditions one can expect that the epidemic diffusion is governed by the logistic equation (equivalent to the so-called SI model)

$$\frac{dF(t)}{dt} = r F(t) \left(1 - \frac{F(t)}{p}\right) \quad (1)$$

whose solution is simply given by

$$F(t) = \frac{p F_o e^{rt}}{p + F_o (e^{rt} - 1)} \quad (2)$$

where  $F_o = F(t = 0)$  is the number of initial infected persons,  $r$  the rate of the infection, namely the number of new infections from one infected person in unit of time (the number of new infections a day), and  $p$  the number of individuals of the population involved. The dynamics goes on until all the population  $p$  is infected. This model has a poor predictable power, however, if we have enough data about the diffusion of the epidemic disease, specially in the first stage of the spreading, in order to get a rough forecast of what can happen in the near future, one could use Eq. (1) to fit the data with  $F_o$ ,  $r$  and  $p$  as free parameters.

The main issue of Eq. (1) is that it does not contain the mechanism of reduction of the spreading and the desired end of an epidemic disease. We have therefore to take into account the number of closed cases (persons who recovered or died), which do not contribute to the infection anymore. The model we are going to consider includes, therefore, the total number of infected,  $F(t)$ , and the total number of recovered and deceased persons,  $R(t)$ , so that Eq. (2) becomes

$$\frac{dF(t)}{dt} = r (F(t) - R(t)) \left(1 - \frac{F(t)}{p}\right) \quad (3)$$

In principle also  $R(t)$  can follow another dynamical equation, however, generally, there is an average time of recovering  $\delta t$  so that the total number of cases at some time  $t$  becomes closed cases at later time  $t + \delta t$ , namely

$$R(t) \simeq F(t - \delta t) \quad (4)$$

This seems to be the case also for the new coronavirus spreading, by looking at some reported data for Covid-19 in Italy, shown in Fig. 1 (see also Ref. [4]). Eq. (4) allows us to write Eq. (3) in terms of only the function  $F(t)$ . Notice that, as we will discuss later, Eq. (3) can be derived also from the SIR model. If we consider the case where the population  $p$  is very large, as long as  $F(t) \ll p$ , we can neglect the logistic term,  $(1 - \frac{F(t)}{p}) \simeq 1$ , so to have

$$\frac{dF(t)}{dt} = r (F(t) - F(t - \delta t) \Theta(t - \delta t)) \quad (5)$$

where  $\Theta$  is the Heaviside theta function. Eq. (5) is a functional retarded differential equation.

## EXACT SOLUTION

Writing the time  $t$  as  $t = n \delta t + t'$ , where  $n = \lfloor \frac{t}{\delta t} \rfloor$  is the integer part of  $t/\delta t$ , the solution of Eq. (5) is given by

$$F(t) = F(n \delta t + t') = F_o \prod_{\ell=1}^n A_{\ell}(\delta t) A_{n+1}(t') \quad (6)$$

where the functions  $A_{\ell}$  fulfill the following iterative equation

$$A_{\ell}(t) = e^{rt} \left(1 - r A_{\ell-1}(\delta t)^{-1} \int_0^t dt' e^{-rt'} A_{\ell-1}(t')\right) \quad (7)$$

with  $A_0(t) = 0$  for any  $t < \delta t$  and  $A_0(\delta t) = 1$ , so that, for  $\ell = 1$ , we recover  $A_1(t) = e^{rt}$ . The full exact solution is, therefore, obtained by solving a cascade of  $n$  local integrals. The proof of Eqs. (6) and (7) is given in Appendix A.

At time  $t = n \delta t$ , from Eq. (7), performing the chain of integrals, and putting the results in Eq. (6), we get the following exact result,

$$F(n \delta t) = F_o \sum_{\ell=0}^{n-1} \frac{(-1)^\ell}{\ell!} ((n - \ell) r \delta t)^\ell e^{(n-\ell)r \delta t} \quad (8)$$

For instance, for  $n = 1$  and  $n = 2$ , namely up to twice the infectiousness period, the total number of cases is simply

$$F(n \delta t) = F_o \left( e^{nr \delta t} - (n - 1) r \delta t e^{(n-1)r \delta t} \right) \quad (n = 1, 2) \quad (9)$$

Surprisingly we find that Eq. (8) depends only on  $(r \delta t)$ , which is called basic reproduction number  $\mathcal{R}_0$  (see below). It is easy to check from Eq. (8) that, while for large  $\mathcal{R}_0 = r \delta t$ ,  $F(n \delta t)$  is dominated by an exponential behavior, for  $\mathcal{R}_0 = 1$ ,  $F(n \delta t)$  becomes linear in  $n$ . From Eq. (6) and Eq. (7), we can notice that the function  $F(t)$  depends on its past, therefore, it seems governed by a non-Markovian dynamics. Once we have the total number of infections  $F(t)$ , we can also calculate the number of persons who are still infected, at a given time  $t$ , which is defined by

$$I(t) = F(t) - R(t) \quad (10)$$

In our model  $I(t) = F(t) - F(t - \delta t)\Theta(t - \delta t)$ , so that Eq. (5), in terms of this quantity is simply  $\frac{dF(t)}{dt} = r I(t)$ . Before to proceed, a comment on the comparison with another model is in order. The so-called SIR model is one of the most used techniques for describing the spreading in time of an infection disease. According to this model the population is divided into three parts represented by the number of susceptible  $S(t)$ , infected  $I(t)$  and recovered  $R(t)$  individuals which vary over time (see Appendix B). This model, is almost equivalent to our simpler model, Eq. (3)-(4). However a criticism which can be raised against the SIR model is related to the fact that, being formulated in terms of ordinary differential equations, the model requires the presence of an effective recovery (and fatality) rate which might not correspond to the actual rate since the new cases of recovery (and fatality) come from infected cases occurring a few days earlier. For that reason, instead of writing the problem in terms of ordinary differential equations one has to do it in terms of functional differential equations.

### Basic reproduction number

Let us consider Eq. (5), for  $t > \delta t$ , in the following form

$$\frac{dF(t)}{dt} = \mathcal{R}_0 \frac{F(t) - F(t - \delta t)}{\delta t} \quad (11)$$

where we introduce and naturally identify  $\mathcal{R}_0$  as the so-called basic reproduction number

$$\mathcal{R}_0 = r \delta t \quad (12)$$

which is a widely used parameter for predicting whether the infectious disease will spread into a population or turns off, and represents the average number of cases originated by a single infectious case during the infectiousness period. Eq. (11) implies that the first derivative of  $F(t)$  is equal to its increment in a time interval  $\delta t$ , divided by  $\delta t$ , namely  $F(t)$  is linear in  $t$  if the rate is equal to the critical value  $r = \delta t^{-1}$  ( $\mathcal{R}_0 = 1$ ). For  $r > \delta t^{-1}$  ( $\mathcal{R}_0 > 1$ ), the function  $F(t)$  increases more than linearly, while for  $r < \delta t^{-1}$  ( $\mathcal{R}_0 < 1$ ),  $F(t)$  goes slower than linearly. If we let  $r$  vary in time, when  $r = \delta t^{-1}$  ( $\mathcal{R}_0 = 1$ ) the function  $F(t)$  has an inflection point, where it changes from being concave to convex or vice versa. Making a comparison with the SIR model, where  $\mathcal{R}_0 = r/\beta$ , one can identify  $\beta$ , the fictitious recovery rate (see Appendix B) with the inverse of the recovery time  $\beta \sim 1/\delta t$ .

### Time-dependent infection rate: exact solution

Let us now consider the possibility of having a time-dependent infection rate  $r(t)$  in the dynamical equation for the total number of infected persons

$$\frac{dF(t)}{dt} = r(t) \left( F(t) - F(t - \delta t)\Theta(t - \delta t) \right). \quad (13)$$

Also in this more general case the exact solution, valid for any profile of  $r(t)$ , can be written in the same form of Eq. (6), namely,  $F(t) = F(n\delta t + t') = F_0 \prod_{\ell=1}^n A_{\ell}(\delta t) A_{n+1}(t')$ , where now the functions  $A_{\ell}$  are given by

$$A_{\ell+1}(t) = e^{\int_{\ell\delta t}^{\ell\delta t+t'} r(t') dt'} \left( 1 - A_{\ell}(\delta t)^{-1} \int_0^t dt' r(\ell\delta t + t') e^{-\int_{\ell\delta t}^{\ell\delta t+t'} r(t'') dt''} A_{\ell}(t') \right). \quad (14)$$

For instance,  $A_1(t) = e^{\int_0^t r(t') dt'}$ ,  $A_2(t) = e^{\int_{\delta t}^{\delta t+t'} r(t') dt'} \left( 1 - e^{-\int_0^{\delta t} r(t') dt'} \int_0^t dt' r(\delta t + t') e^{-\int_{\delta t}^{\delta t+t'} r(t'') dt''} e^{\int_0^{t'} r(t'') dt''} \right)$ , and so on. For constant  $r$ , Eq. (14) reduces to Eq. (7). See Appendix A for more details about the derivation. The solution  $F(t)$  has therefore to fulfill the following recursive equation, after splitting the time in  $n$  intervals  $\delta t$  with the residual time  $t'$

$$F(t) = F(n\delta t + t') = e^{\int_{n\delta t}^{n\delta t+t'} r(s) ds} \left( F(n\delta t) - \int_0^{t'} ds r(n\delta t + s) e^{-\int_{n\delta t}^{n\delta t+s} r(t'') dt''} F((n-1)\delta t + s) \right) \quad (15)$$

This general result implies that if we knew the time dependence of the infection rate or if we could tailor its evolution by, for instance, containment measures, we can know the exact analytical expression of  $F(t)$ , the total number of infected persons, as a function of time, as long as  $F(t)$  is much smaller than  $p$ .

### COVID-19 IN ITALY

Let us consider the delay model in its general form Eqs. (3)-(4), where the infection rate  $r$  varies in time

$$\frac{dF(t)}{dt} = r(t) \left( F(t) - F(t - \delta t) \Theta(t - \delta t) \right) \left( 1 - \frac{F(t)}{p} \right) \quad (16)$$

as the effect of some containment measures taken in order to reduce the impact of an infection on the population. As

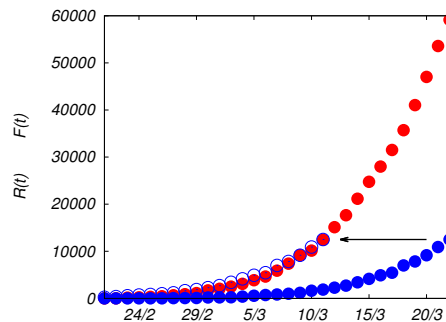


FIG. 1: Total number of confirmed cases of Covid-19 in Italy,  $F(t)$  (red dots), reported in Ref. [5], since 21th February to 22th March 2020, compared with the closed cases,  $R(t)$  (blue dots), in the same period of time. If the numbers of closed cases are shifted in time by  $\delta t \simeq 11$  days (blue circles) they fairly overlap with the total numbers of cases.

an example, let us suppose that  $r(t)$  is modified by social distancing measures, lockdown and the shutdown of many work activities, as in it is happening in Italy (and in many other countries) to mitigate and reduce the spreading of the new coronavirus, Covid-19, after two main decrees imposed by the Italian Prime Minister ordering the lockdown of the whole national territory, taken on March 11-th (lockdown and shutdown of many stores) and March 22-th 2020 (shutdown of many factories and strengthening of social distancing measures), after some other measures taken right before for local regions (e.g. the decree of March 8-th for the lockdown of Lombardy and other areas). As a result, we can imagine that  $r(t)$  decreases smoothly after those dates taking into account the adaptation time for the individuals to the new social behaviors and the period needed to complete the last activities before the blockade of the factories. Let us suppose, therefore, that  $r(t)$  can change in time according to a smooth step function as in Eq. (17),

$$r(t) = \left( \frac{r_1 - r_2}{1 + e^{(t-t_1)/\tau_1}} + r_2 - r_3 \right) \frac{1}{1 + e^{(t-t_2)/\tau_2}} + r_3 \quad (17)$$

where  $t_1$  and  $t_2$  are the times where the steps are located,  $\tau_1$  and  $\tau_2$  make the function to be smooth,  $r_1$  is the initial observed infection rate which causes the starting exponential growth of the epidemic disease,  $r_2$  the intermediate rate, which fits with the data, supposed to be reached after the first decree of lockdown, and  $r_3$  the supposed asymptotic infection rate after the second decree of lockdown. Fixing the average of recovery and fatality time  $\delta t$ , the basic reproduction number is also a function of time according to  $\mathcal{R}_0 = r(t)\delta t$ , with a profile shown in Fig. 2.

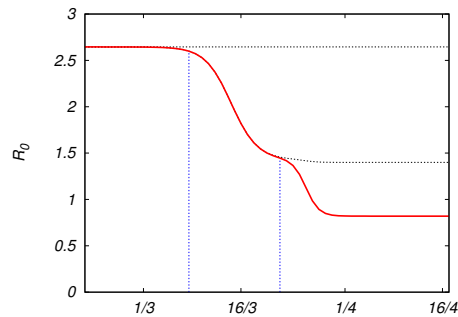


FIG. 2: Basic reproduction number  $\mathcal{R}_0 = r(t)\delta t$ , as a function of time, based on the profile for the infection rate described by Eq. (17). We take  $\delta t$  about 11 or 12 days [4],  $t_1$  between 13th and 14th March 2020,  $t_2$  on 26th March,  $\tau_1 \sim 2$  days,  $\tau_2 \sim 1$  day. The starting value is  $\mathcal{R}_0 = r_1\delta t \simeq 2.65$  and the intermediate value is  $\mathcal{R}_0 = r_2\delta t \simeq 1.45$ , and a current value,  $\mathcal{R}_0 = r_3\delta t \simeq 0.85$ . The vertical dotted blue lines point the dates of the main laws for the containment measures (8th-11th March and 22th March 2020). The gray dotted lines correspond to  $\mathcal{R}_0$  in the absence of the first and the second containment measures.

Solving Eq. (16), or, analogously, using the recursive relation in Eq. (15), with the time-dependent rate  $r(t)$  given by Eq. (17), fixed by the parameters producing the profile depicted in Fig. 2, we obtain the solution  $F(t)$  which slowly goes to saturation over time, in perfect agreement with the data for the total number of confirmed infected cases, as shown by Fig. 3, where the blue line is the expected curve, while the red points are the official data. The dotted gray lines in Fig. 3 represent  $F(t)$  if the containment measures had not been taken. As one can see from Figs. 2-3, only when  $\mathcal{R}_0$  becomes smaller than 1, the curve flattens allowing for a stop of the epidemic spreading, avoiding that a large part of the population gets infected. For  $\mathcal{R}_0 \simeq 1$ ,  $F(t)$  would increase linearly, and  $I(t)$  would become almost constant, meaning that the number of new infections would be equal to the number of closed cases. A reliable forecast has to take into account the fact that the official data of infectious cases are made by counting mostly the symptomatic cases, probably discarding other infectious cases which could transfer the virus even without or with mild symptoms. Moreover, the data of both the total number of infected persons and that of the recovered ones could be affected by the procedure, the realization times and the number of the diagnostic tests. However, since our model relies on the infectiousness time, it does not need a fitting of the data for recovered persons which may be affected by systematic errors. This uncertainty on the data for closed cases would compromise the result for the SIR model. On the contrary, our theoretical prediction based on the delay model agrees fairly well with the data-set for total infected cases, as shown in Fig. 3.

As a final remark we remind that most of the confirmed infected cases in Italy are counted after the appearance of the symptoms and the persons who exhibit severe ones are mostly hospitalized, and then counted as infected persons. Some of them, unfortunately, die approximately 4 days after (therefore after approximately 9 days from the onset of the first symptoms, as reported by the Istituto Superiore di Sanità, [www.iss.it](http://www.iss.it)). We observe that, splitting the closed cases between real recovered persons,  $R_R$ , and dead persons,  $R_D$ ,

$$R(t) = R_R(t) + R_D(t) \quad (18)$$

and since the confirmation of recovery needs extra diagnostic tests which are still not widely performed, the most reliable data is that of dead persons  $R_D(t)$ , which is found to be linked to the total number of confirmed infected cases,  $F(t)$ , in the following way

$$R_D(t) \simeq \gamma F(t - \delta t_d) \quad (19)$$

with  $\gamma = \frac{1}{7}$  and a delay time of  $\delta t_d = 4$  days, as show in Fig. 4. The number of victims follows the number of total confirmed cases and it is equal to 1/7 of its value four days before.

The fatality of the sick persons, those who exhibit some symptoms, is therefore quite high,  $\lim_{t \rightarrow \infty} \frac{R_D(t)}{F(t)} = \gamma \simeq 14\%$ .

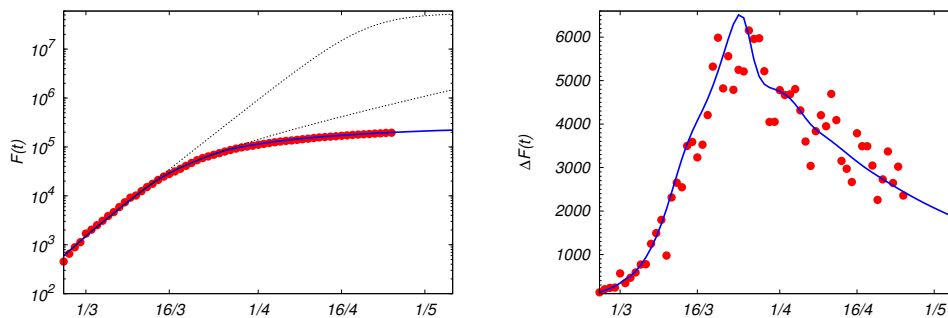


FIG. 3: (Left) Total number of infected persons over time,  $F(t)$  (red points), from official data for Covid-19 in Italy [5], where  $p = 6 \cdot 10^7$ , up to 14th April. The blue line is the theoretical prediction  $F(t)$  as solution of Eq. (16), with initial conditions, fixed at  $t = 0$  the 21th February,  $F_0 \simeq 150$ , and  $r = r_1 \simeq 0.23$  ( $\mathcal{R}_0 \simeq 2.65$ ), and using the profile for the infection rate given by Eq. (17), with the parameters reported in Fig. 2. The gray dotted lines are the expected curves for  $F(t)$  if the first and the second containment measures (8-11th March and 22th March) had not been taken. (Right) Daily number of infected persons,  $\Delta F(t)$ , compared with the theoretical  $\frac{dF(t)}{dt}$  from the solution of Eq. (16).

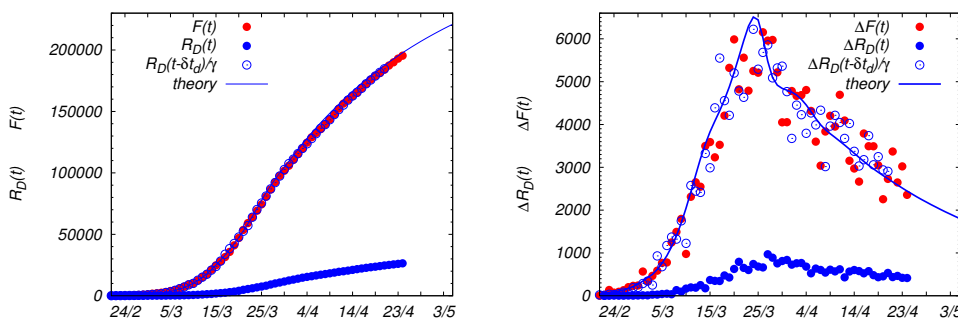


FIG. 4: (Left) Total number of confirmed cases of Covid-19 in Italy,  $F(t)$  (red dots), reported in Ref. [5], up to 25th April 2020, compared with the deceased cases,  $R_D(t)$  (blue dots), in the same period of time. If the numbers of died persons are shifted in time by  $\delta t_d \simeq 4$  days and rescaled by  $\gamma^{-1} = 7$  (blue circles) they perfectly overlap with the total numbers of cases. The blue solid line is theoretical prediction for  $F(t)$ , solution of Eq. (16). (Right) Daily number of infected persons,  $\Delta F(t)$  (red dots), and daily number of victims  $\Delta R_D(t)$  (blue dots). The blue circles are the daily number of victims after rescaling according to Eq. (19), with  $\delta t_d = 4$  days and  $\gamma = 1/7$ . The solid blue line is  $\frac{dF(t)}{dt}$  from the solution of Eq. (16).

## CONCLUSIONS

We present a simple but realistic model for describing epidemic spreading, based on the fact that the closed cases come from infected ones at an early time. This observation allows us to formulate the problem in terms of a single functional differential equation depending on two well defined clinically relevant parameters: the infection rate and the infectiousness time. We provide the exact analytical solution for such an equation, in the limit of large population number, finding that it depends exclusively on the basic reproduction number  $\mathcal{R}_0 = r\delta t$ , see Eq. (8). We derive the analytic solution also in the presence of a generic time-dependent infection rate. We apply our model to the case of the spreading of Covid-19 in Italy, allowing the infection rate to vary in time, as a result of some containment measures implemented by the government in order to mitigate the consequences of the infection on the population. We find perfect agreement between the official data and the expected theoretical results. In general terms, the basic reproduction number should be suppressed well below 1 in order to rapidly recover the initial condition. By a rough estimation, in order to have a decline of the infection as fast as its growth, containment measures or possible therapies should be so effective to reduce the basic reproduction number and reach the final value  $R_{0f}$  such that  $R_{0f} \simeq \frac{R_{0i}}{2R_{0i}-1}$ , starting from an initial value  $R_{0i}$ . In the case of Covid-19 in Italy, the initial value for the basic reproduction number was  $R_{0i} \simeq 2.6$ , while the current one seems to settle at  $R_{0f} \simeq 0.8$ , implying a rather slow decline of the infection. Finally we discussed the fatality rate, showing that the number of victims is exactly a fraction of the total number of cases few days before.



### Appendix A: Solution of the retarded differential equation

For  $t \leq \delta t$ , the solution of Eq. (5) is  $F(t) = F_o e^{rt}$ . Let us consider  $t = \delta t + dt$  with infinitesimal  $dt$ , from Eq. (5)

$$F(\delta t + dt) = F(\delta t) + dt r (F(\delta t) + F(0)) = F(\delta t) (1 + r dt) - F_o r dt = F_o e^{r\delta t} (1 + r dt) - F_o r dt \quad (20)$$

Using this result we can calculate

$$F(\delta t + 2dt) = F(\delta t + dt) + dt r (F(\delta t + dt) + F(dt)) = F_o e^{r\delta t} (1 + r dt)^2 - F_o r dt [(1 + r) + e^{r dt}] \quad (21)$$

Analogously, from that, we can proceed calculating

$$F(\delta t + 3dt) = F(\delta t + 2dt) + dt r (F(\delta t + 2dt) + F(2dt)) = F_o e^{r\delta t} (1 + r dt)^3 - F_o r dt [(1 + r dt)^2 + e^{r dt} (1 + r dt) + e^{2r dt}] \quad (22)$$

and going on by adding infinitesimal time steps, we find iteratively that

$$F(\delta t + m dt) = F_o e^{r\delta t} (1 + r dt)^m - F_o r dt \sum_{j=0}^{m-1} e^{jr dt} (1 + r dt)^{m-j-1} \equiv F_o A_1(\delta t) A_2(m dt) \quad (23)$$

$$= F(\delta t) A_2(m dt) \quad (24)$$

with  $A_1(\delta t) = e^{r\delta t}$  and defining

$$A_2(m dt) = (1 + r dt)^m - e^{-r\delta t} r dt \sum_{j=0}^{m-1} e^{jr dt} (1 + r dt)^{m-j-1}. \quad (25)$$

In particular, for  $m dt = \delta t$ , we have an expression for  $F(2\delta t)$  in terms of the function at early time,  $F(2\delta t) = F(\delta t) A_2(\delta t)$ . We can now start again with the iteration

$$F(2\delta t + dt) = F(2\delta t)(1 + r dt) - r dt F(\delta t) = F(\delta t) A_2(\delta t) (1 + r dt) - r dt F(\delta t) \quad (26)$$

One can proceed in the same way as before getting

$$F(2\delta t + m dt) = F(\delta t) \left[ (1 + r dt)^m A_2(\delta t) - r dt \sum_{j=0}^{m-1} A_2(j dt) (1 + r dt)^{m-j-1} \right] \quad (27)$$

which can be written as

$$F(2\delta t + m dt) = F(\delta t) A_2(\delta t) A_3(m dt) = F(2\delta t) A_3(m dt) \quad (28)$$

where

$$A_3(m dt) = (1 + r dt)^m - A_2(\delta t)^{-1} r dt \sum_{j=0}^{m-1} A_2(j dt) (1 + r dt)^{m-j-1}. \quad (29)$$

We can notice that at any step  $\delta t$  we can perform the same calculation since we can factorize the function  $F$  as

$$F(n \delta t + m dt) = F(n \delta t) A_{n+1}(m dt) \quad (30)$$

where, therefore,  $F(n \delta t) = F_o \prod_{\ell=1}^n A_\ell(\delta t)$  and

$$A_\ell(m dt) = (1 + r dt)^m \left[ 1 - A_{\ell-1}(\delta t)^{-1} r dt \sum_{j=0}^{m-1} \frac{A_{\ell-1}(j dt)}{(1 + r dt)^{j+1}} \right]. \quad (31)$$

In the continuum limit,  $dt \rightarrow 0$  and  $m \rightarrow \infty$ , keeping finite the time interval  $m dt = t$ , reminding that

$$\lim_{m \rightarrow \infty} \left( 1 + \frac{rt}{m} \right)^m = e^{rt} \quad (32)$$

we finally obtain the result reported Eq. (7).

In the presence of time dependent infection rate, splitting again the time in  $n$  intervals  $\delta t$  and the residual time in  $m$  infinitesimal intervals  $dt$ , we define

$$r(t) = r(n \delta t + m dt) \equiv r_m^{(n)}. \quad (33)$$

Proceeding iteratively as done for the constant rate case, but now taking trace of the different values of  $r$ ,

$$F(n \delta t + m dt) = F(n \delta t + (m-1)dt) \left(1 + r_m^{(n)} dt\right) - r_m^{(n)} dt F((n-1) \delta t + (m-1)dt) \quad (34)$$

after several steps, similar to those done previously, we find that Eq. (31) can be generalized in the following way

$$A_\ell(m dt) = \prod_{i=1}^m \left(1 + r_i^{(\ell-1)} dt\right) - A_{\ell-1}(\delta t)^{-1} dt \sum_{j=0}^{m-1} \left[ r_j^{(\ell-1)} A_{\ell-1}(j dt) \prod_{i=1}^j \left(1 + r_i^{(\ell-1)} dt\right)^{m-j-1} \right], \quad (35)$$

whose continuum limit is given in Eq. (14).

### Appendix B: Comparison with the SIR model

The most commonly used model for epidemic spreading is the so-called SIR model, which describes the dynamics of the number of susceptible,  $S(t)$ , infected,  $I(t)$  and recovered,  $R(t)$  persons, according to the following differential equations

$$\frac{dS(t)}{dt} = -\alpha \frac{S(t)}{p} I(t) \quad (36)$$

$$\frac{dI(t)}{dt} = \alpha \frac{S(t)}{p} I(t) - \beta I(t) \quad (37)$$

$$\frac{dR(t)}{dt} = \beta I(t) \quad (38)$$

with generally the initial condition  $S(0) \simeq p$ . The free parameters  $\alpha$ , the infection rate and  $\beta$ , the recovery rate, can be fixed by fitting the data sets. Defining

$$F(t) = I(t) + R(t) \quad (39)$$

and summing Eqs. (37)-(38) we get

$$\frac{dF(t)}{dt} = \alpha \frac{S(t)}{p} I(t) = \alpha \frac{S(t)}{p} (F(t) - R(t)). \quad (40)$$

Summing the three Eqs. (37)-(38) one gets, for any time  $t$ ,

$$\frac{d}{dt} (S(t) + I(t) + R(t)) = 0 \quad (41)$$

namely that, the sum of the three functions is constant and equal to the population  $p$  for any  $t$ , since at the beginning  $S(0) + I(0) = p$ , therefore,

$$\frac{S(t)}{p} = 1 - \frac{F(t)}{p} \quad (42)$$

which is nothing but the logistic term so that Eq. (40) is exactly equal to Eq. (3), where  $\alpha = r$ . This implies that one equation among Eqs. (36)-(38) is redundant, therefore, instead of considering three equations one can take just two. For instance, we can choose to express the time evolution in terms of  $F(t)$  and  $R(t)$ ,

$$\frac{dF(t)}{dt} = \alpha (F(t) - R(t)) \left(1 - \frac{F(t)}{p}\right) \quad (43)$$

$$\frac{dR(t)}{dt} = \beta (F(t) - R(t)) \quad (44)$$



In particular, as long as  $F(t) \ll p$  so that  $(1 - F(t)/p) \approx 1$ , from Eqs. (43) and (44), we have

$$\frac{dI(t)}{dt} = (\alpha - \beta) I(t) \quad (45)$$

meaning that only if  $\alpha$  is smaller than  $\beta$ ,  $\alpha < \beta$ , the infection shuts down. One can introduce the so-called basic reproduction number  $\mathcal{R}_0$  which predicts whether the infectious disease will spread into a population or die out, and represents the average number of cases originated by a single infectious case in a totally susceptible population during the infectiousness period. This quantity is defined as

$$\mathcal{R}_0 = \frac{\alpha}{\beta}, \quad (46)$$

and, from Eq. (45), one can see that only for  $\mathcal{R}_0 < 1$  the infection turns off.

We notice that Eq. (43) is equal to Eq. (3). The difference between the delay model, described by Eqs. (3)-(4), and the SIR model, described by Eqs. (43)-(44), is that, in our model, the number of closed cases  $R(t)$  is locked to be equal to the total cases at a former time  $F(t - \delta t)$ , an average recovery period  $\delta t$  before. In other words, in the delay model Eq. (44) is substituted by Eq. (4), which is equivalent to Eq. (44) with  $\beta = \alpha$  but with a different initial condition,  $R(\delta t) = F(0)$ . The delay time allows us to define the basic reproduction number more naturally as

$$\mathcal{R}_0 = \alpha \delta t, \quad (47)$$

the number of new infections during the infectiousness time  $\delta t$ . On the contrary, in the SIR model one has to artificially adjust the rate  $\beta$  to take into account the right initial condition.

- 
- [1] R. M. Anderson, B. Anderson, R. M. May, *Infectious diseases of humans: dynamics and control*, Oxford University Press (1992).
  - [2] M. J. Keeling, P. Rohani, *Modeling infectious diseases in humans and animals* (Princeton University Press, 2011).
  - [3] W.O. Kermack, A.G. McKendrick, *A Contribution to the Mathematical Theory of Epidemics*, Proceedings of the Royal Society A. 115 (772): 700-721 (1927)
  - [4] *Symptoms of Novel Coronavirus (2019-nCoV)*, CDC (Center for Disease Control and Prevention) [www.cdc.gov](http://www.cdc.gov), 10/02/2020, <https://www.cdc.gov/coronavirus/2019-ncov/about/symptoms.html>
  - [5] <https://lab.gedidigital.it/gedi-visual/2020/coronavirus-i-contagi-in-italia/>