A general theory for infectious disease dynamics.

Giuseppe Carbone,* Ilario De Vincenzo,

[1]Department of Mechanics, Mathematics and Management (DMMM),
Polytechnic University of Bari, Via E. Orabona n-4, 70125 Bari, Italy

*giuseppe.carbone@poliba.it.

September 14, 2022

We present a general theory of infection spreading, which directly follows from conservation laws and takes as inputs the probability density functions of latent times. The derivation of the theory substantially differs from Kermack and McKendrick (1927) argument, which instead was based on the concept of removal rates. We demonstrate the formal equivalence of the two approaches, but our theory provides a clear interpretation of the kernels of the integro-differential equations governing the infection spreading in terms of survival function of the latent times distributions. This aspect was never captured before. Real distributions of latent times can be, then, employed, thus overcoming the limitations of standard SIR, SEIR and other similar models, which implicitly make use of exponential or exponential-related distributions. SIR and SEIR-type models are, in fact, a subclass of the theory here presented. We show that beside the infection rate ν, the joint probability density function $p_{\mathrm{EI}}(\tau, \tau_1)$ of latent times in the exposed and infectious compartments governs

the infection spreading. Assuming that the number of infected individuals is negligible compare to the entire population, we were able to study the stability of the dynamical system and provide the general solution of equations in terms characteristic functions of the probability distribution of latent times. We present asymptotic solutions for the case $R_0 = 1$ and demonstrate that the moments of the latent times distribution govern the rate of disease spreading in this case. The present theory is employed to simulate the diffusion of COVID-19 infection in Italy during the first 120 days. The estimated value of the basic reproduction number is $R_0 \approx 3.5$, in very good agreement with existing data.

Introduction

Mathematical modeling has proven to be an extremely powerful tool to explain and predict the course of an epidemic, by helping to plan effective policies to control the infection spreading (*1–8*).

The most common framework for modeling infectious disease dynamics involves dividing the population into compartments and using ordinary differential equations (ODEs) to specify flows between them. For diseases conferring permanent immunity, the SIR (susceptible - infectious - recovered) model is usually employed (*1,9,10*). A simple enhancement of the SIR model is SEIR one (susceptible - exposed - infectious - recovered), where the compartment of exposed (i.e. infected) individuals but not yet infectious is introduced. These models assume that the time an individual spends in each compartment is exponentially distributed (*2–5*). While such an assumption could sound reasonable in some circumstances, it is often unrealistic (*11–15*), and usually leads to an underestimation of the basic reproduction number and to overoptimistic predictions about the efficacy of epidemic control measures (*16*). The assumption of exponentially distributed latent times, may also overestimate the numbers of individuals whose infection

period is much shorter or longer than its average (*13, 14*). Several investigations have, indeed, revealed that the shape of the probability density function of latent times may have a huge impact on the predicted infection dynamics (*14, 16–30*). Aware of these limits, we present a theory of infection spread, where the transition from the susceptible compartment to the exposed one is still governed by a reaction-like equation, but the flow of individuals to any other subsequent compartment (e.g., from exposed to infectious or from infectious to removed) is rather governed by integral equations that express the fact that individuals, who enter a certain compartment at a given time will move to the subsequent one after a generic stochastically distributed latent time. The present theory does not pose any restriction on the functional form of the latent times probability density functions. Thus, real distributions, obtained by tracing real data, can be easily utilized. The present theory is employed to describe the evolution of COVID-19 infection (*31–35*) in Italy during the first 120 day.

1 The Model

We assume homogenous mixing of the population. Accordingly, the effect of the topological structure of the individuals contact network is not considered in this study. However, the theory can be easily generalized to the case of disease spreading on networks or in the case of heterogeneous mixing (*36–38*).

Our model consists of six compartments: (i) susceptible, (ii) exposed, (iii) infectious, (iv) removed, (v) healed, (vi) deceased individuals as shown in Fig. 1. Let be N the total and constant number of individuals in the population, $s(t)$ the number of susceptible individuals, $e(t)$ the number of exposed individual, $i(t)$ the number of infectious individuals, $r(t)$ the number of removed (isolated) individuals, $h(t)$ the number of recovered (healed) individuals, and $d(t)$ the number of deceased individuals. We also define $n(t) = N - s(t)$ the cumulative number of cases, i.e. the total number of individual that have been infected. Mass conservation

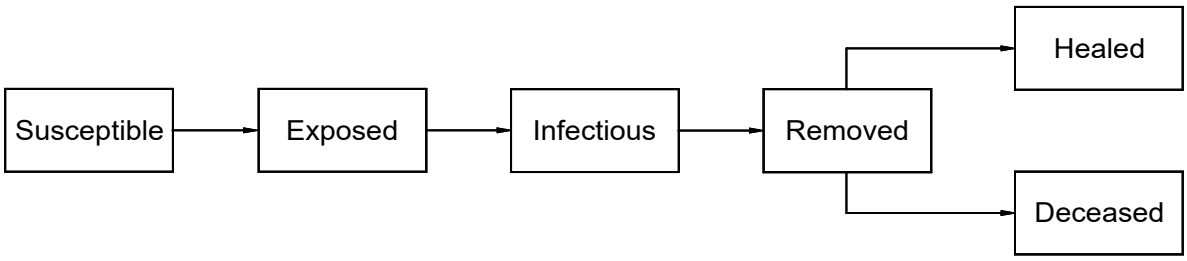


Figure 1: The compartments of the proposed model.

requires $s(t) + e(t) + i(t) + r(t) + h(t) + d(t) = N$, i.e.

$$e(t) + i(t) + r(t) + h(t) + d(t) = n(t) \tag{1}$$

We assume that all the above defined quantities are zero for negative times, i.e. we assume that $n = e = i = r = h = d = 0$ for $t < 0$. Moreover at $t = 0$ only $n(t)$, $s(t)$ and $e(t)$ undergoes a step change so that $n(0^+) = e(0^+) = n_0$ and $s(0^+) = s_0 = N - n_0$. By defining the Heaviside step function $\mathcal{H}(t)$ as $\mathcal{H}(t \leq 0) = 0$ and $\mathcal{H}(t > 0) = 1$, we can write that the time derivative $\dot{n}(t) = n_0\delta(t) + \mathcal{H}(t)\dot{n}(t)$, and analogously $\dot{e}(t) = n_0\delta(t) + \mathcal{H}(t)\dot{e}(t)$, $\dot{s}(t) = s_0\delta(t) + \mathcal{H}(t)\dot{s}(t)$, where $\delta(t)$ is the Dirac delta function. Now we assume that transition from the susceptible compartment to the exposed one is governed by the standard reaction-like equation

$$\dot{s}(t) = -\nu i(t)\frac{s(t)}{N} + s_0\delta(t) \tag{2}$$

Note that the forcing term $s_0\delta(t)$ is included to account for the condition $s(0^+) = s_0$. In fact integrating both hand sides from $t = 0^-$ to $t = 0^+$ one gets $s(0^+) - s(0^-) = s_0$ and recalling that $s(0^-) = 0$ we retrieve the initial condition $s(0^+) = s_0$. In Eq. (2) the infection rate ν is the average number of contacts per person per unit time, multiplied by the probability of disease transmission in a single contact between a susceptible and an infectious individual. Here we assume, since the beginning, that the infection rate of each individual does not depend on the time when the he/she got infected. However the extension to this case is straightforward.

Incidentally we notice that in general Kermack and McKendrick (*9*) model this assumption was enforced at the beginning of their argument. Eq. (2) can be rephrased as

$$\dot{n} = \nu \left(1 - \frac{n}{N}\right) i\left(t\right) + n_0 \delta\left(t\right) \tag{3}$$

Now consider a single individual that has been infected at time t. Such an individual will become infectious after a certain randomly distributed latent period $\tau_k \geq 0$ and therefore will move from the exposed to the infectious compartment at time $t + \tau_k$. Let us call $\delta m_k\left(t\right)$ the number of individuals in the time interval δt, having latent time τ_k, which move out from the exposed compartment at time t. Of course $\delta m_k\left(t\right)$ equals the number of individuals $\delta n_k\left(t - \tau_k\right)$ that have been infected (during the same time step δt) at time $t - \tau_k$. Hence, the increase $\delta e\left(t\right)$ of individuals in the exposed compartment at time t is just given by the number $\delta n\left(t\right)$ of newly infected individuals at time t diminished by the number of individuals $\delta m\left(t\right) = \sum_k \delta m_k\left(t\right) = \sum_k \delta n_k\left(t - \tau_k\right)$, i.e.

$$\delta e\left(t\right) = \delta n\left(t\right) - \sum_k \delta n_k\left(t - \tau_k\right) \tag{4}$$

or dividing by the infinitesimal time step δt

$$\dot{e}\left(t\right) = \dot{n}\left(t\right) - \sum_k \dot{n}_k\left(t - \tau_k\right) \tag{5}$$

Recalling that the latent time τ individuals spend in the exposed compartment is a continuous stochastic variable, one can write that the fraction of individuals that have latent time ranging in the interval $\tau_k \leq \tau < \tau_k + \Delta\tau$ is

$$\frac{\delta n_k}{\delta n} = \frac{\dot{n}_k}{\dot{n}} = p_{\mathrm{E}}\left(\tau_k\right) \Delta\tau \tag{6}$$

where $p_{\mathrm{E}}\left(\tau\right)$ is the density probability function of latent times $\tau > 0$. We assume that $p_{\mathrm{E}}\left(\tau\right)$ is independent of time t. This yields

$$\dot{e}\left(t\right) = \dot{n}\left(t\right) - \sum_k p_{\mathrm{E}}\left(\tau_k\right) \dot{n}\left(t - \tau_k\right) \Delta\tau \tag{7}$$

and by letting $\Delta\tau \to 0$ we get

$$\dot{e}(t) = \dot{n}(t) - \int_{0^-}^{+\infty} p_{\mathrm{E}}(\tau)\, \dot{n}(t-\tau)\, d\tau \tag{8}$$

Eq. (8) can be easily t-integrated to get

$$e(t) = n(t) - \int_{0^-}^{+\infty} p_{\mathrm{E}}(\tau)\, n(t-\tau)\, d\tau \tag{9}$$

Observe that Eq. (9) is Volterra equation. A similar argument applies to the number $i(t)$ of individuals in the infectious compartment (see Sec. A) to give

$$i(t) = \int_{0}^{+\infty} p_{\mathrm{E}}(\tau)\, n(t-\tau)\, d\tau - \iint d\tau d\tau_1 p_{\mathrm{EI}}(\tau,\tau_1)\, n(t-\tau-\tau_1) \tag{10}$$

where $p_{\mathrm{EI}}(\tau,\tau_1)$ is the joint probability density function of the latent times τ and τ_1 in the exposed and infectious compartments respectively (τ_1 is the time the single individual spends into the infectious compartment). Note that to simplify the notation we have used the symbol $\int$ in place of $\int_{0^-}^{+\infty}$. Consider that the lower limit of integration 0^-can be replaced with $-\infty$ since the probability density functions must vanish when at least one of the arguments is negative.

Using the same arguments for the case of individuals in the removed compartment, and defining the latent time τ_2 as the time a removed individual takes to recover and τ_3 the latent time a removed individual takes to enter the deceased compartment we get

$$\begin{aligned} r(t) &= n(t) - e(t) - i(t) \\ &\quad -\gamma \iiint d\tau d\tau_1 d\tau_2 p_{\mathrm{EIH}}(\tau,\tau_1,\tau_2)\, n(t-\tau-\tau_1-\tau_2) \\ &\quad -\beta \iiint d\tau d\tau_1 d\tau_3 p_{\mathrm{EID}}(\tau,\tau_1,\tau_3)\, n(t-\tau-\tau_1-\tau_3) \end{aligned} \tag{11}$$

and

$$h(t) = \gamma \iiint d\tau d\tau_1 d\tau_2 p_{\mathrm{EIH}}(\tau,\tau_1,\tau_2)\, n(t-\tau-\tau_1-\tau_2) \tag{12}$$

$$d(t) = \beta \iiint d\tau d\tau_1 d\tau_3 p_{\mathrm{EID}}(\tau,\tau_1,\tau_3)\, n(t-\tau-\tau_1-\tau_3) \tag{13}$$

where $p_{\mathrm{EIH}}(\tau, \tau_1, \tau_2)$ is the joint probability density function of latent times τ, τ_1 and τ_2, and $p_{\mathrm{EID}}(\tau, \tau_1, \tau_3)$ is the joint probability density function of latent times τ, τ_1 and τ_3. The quantity γ is the total fraction of infected people that has recovered in the long term (i.e. for $t \to +\infty$) and $\beta = 1 - \gamma$ is the fraction of infected people that, for $t \to +\infty$, has entered the deceased compartment. Also note that summing up Eq.(9-13) yields, as expected, Eq. (1). Note that in order to derive the dynamic equations of the infection diffusion we never had the need to refer to the concept of rate of removal, which was a crucial point in the Kermack and McKendrick argument (*9*). Evolution equations are, instead, naturally derived by enforcing mass conservation laws, once known the probability density functions of latent times. The present theory does not pose any restriction to the functional form of the probability density functions $p_{\mathrm{E}}(\tau)$, $p_{\mathrm{EI}}(\tau, \tau_1)$, $p_{\mathrm{EIH}}(\tau, \tau_1, \tau_2)$, $p_{\mathrm{EID}}(\tau, \tau_1, \tau_3)$. This is a great advantage as it is widely recognized that the such probability distributions very often do not obey the exponential distribution law. Attempts, indeed have been made to overcome this limitation by resorting to gamma-distributed or to the sub-class of the Erlang-distributed latent times, which allow the splitting of the problem in a sum/sequence of compartments each with a specific exponential distribution of latent times (*39–41*). However, many real latent time distributions cannot be accurately described by means of the aforementioned distributions, which that cannot either account for the statistical dependence of latent times. Recalling that $p_{\mathrm{E}}(\tau) = \int d\tau_1 p_{\mathrm{EI}}(\tau, \tau_1)$, Eqs. (3, 9, 10) show that, beside the quantity $\nu(t)$, the infection spreading is governed by the joint probability density function $p_{\mathrm{EI}}(\tau, \tau_1)$, which needs to be accurately measured to guarantee reliable predictions and suggest effective epidemic mitigation policies. Note that the dynamics of the removed, healed and deceased compartments can be solved a posteriori, once known the distributions $p_{\mathrm{EIH}}(\tau, \tau_1, \tau_2)$, $p_{\mathrm{EID}}(\tau, \tau_1, \tau_3)$.

2 The large class of SIR/SEIR-type models as a particular case of the present theory

In this section we show that the present theory contains as a subclass the classical SIR/SEIR and similar compartmental models, which implicitly assume that latent times are exponentially distributed. So let us make the hypothesis that $p_{\mathrm{E}}(\tau) = \mu \exp(-\mu\tau)$, replacing in Eq.(9) and using integration by parts we obtain

$$e(t) = \int_0^{+\infty} \exp(-\mu\tau)\, \dot{n}(t-\tau)\, d\tau \tag{14}$$

Noting that

$$\int_0^{+\infty} f(\tau)\, g(t-\tau)\, d\tau = \int_{-\infty}^{t} f(t-\tau)\, g(\tau)\, d\tau \tag{15}$$

yields

$$e(t) = \exp(-\mu t) \int_{-\infty}^{t} \exp(\mu\tau)\, \dot{n}(\tau)\, d\tau \tag{16}$$

taking the time derivative and recalling Eq. (3)

$$\dot{e}(t) = -\mu e(t) + \dot{n}(t) = -\mu e(t) + \nu(t)\, i(t)\, \frac{s(t)}{N} \tag{17}$$

which is the classical equation for the exposed compartment in the classical SIR/SEIR model. Now let us assume that the latent times τ and τ_1 are statistically independent i.e.

$$p_{\mathrm{EI}}(\tau, \tau_1) = p_{\mathrm{E}}(\tau)\, p_{\mathrm{I}}(\tau_1) \tag{18}$$

and the τ_1 is exponentially distributed, then we have

$$p_{\mathrm{I}}(\tau_1) = \rho \exp(-\rho\tau_1) \tag{19}$$

$$i(t) = n(t) - e(t) - \int d\tau p_{\mathrm{E}}(\tau) \int d\tau_1 \rho \exp(-\rho\tau_1)\, n(t-\tau-\tau_1) \tag{20}$$

recalling Eq. (15), integration by parts yields

$$i\left(t\right)=\exp\left(-\rho t\right)\int_{0}^{+\infty}d\tau p_{\mathrm{E}}\left(\tau\right)\exp\left(\rho\tau\right)\int_{-\infty}^{t-\tau}d\tau_{1}\exp\left(\rho\tau_{1}\right)\dot{n}\left(\tau_{1}\right) \tag{21}$$

Taking the time derivative of both hand sides and recalling Eqs. (8, 17) lead to

$$\frac{di\left(t\right)}{dt}=-\rho i\left(t\right)+\dot{n}\left(t\right)-\dot{e}\left(t\right)=-\rho i\left(t\right)+\mu e\left(t\right) \tag{22}$$

which is the classical SEIR equation for the infectious compartment. Following the same path, we now take

$$p_{\mathrm{EIH}}\left(\tau,\tau_{1},\tau_{2}\right)=p_{\mathrm{EI}}\left(\tau,\tau_{1}\right)p_{\mathrm{H}}\left(\tau_{2}\right)$$

$$p_{\mathrm{EID}}\left(\tau,\tau_{1},\tau_{2}\right)=p_{\mathrm{EI}}\left(\tau,\tau_{1}\right)p_{\mathrm{D}}\left(\tau_{3}\right)$$

with

$$p_{\mathrm{H}}\left(\tau_{2}\right)=\lambda\exp\left(-\lambda\tau_{2}\right)$$

$$p_{\mathrm{D}}\left(\tau_{3}\right)=\eta\exp\left(-\eta\tau_{3}\right)$$

replacing in Eq. (11), using integration by parts and recalling Eq. (15) yields

$$\begin{aligned}r\left(t\right)=&\iint d\tau d\tau_{1}p_{\mathrm{EI}}\left(\tau,\tau_{1}\right)\\&\times\left\{\gamma\exp\left[-\lambda\left(t-\tau-\tau_{1}\right)\right]\int_{-\infty}^{t-\tau-\tau_{1}}d\tau_{2}\exp\left(\lambda\tau_{2}\right)\dot{n}\left(\tau_{2}\right)+\right.\\&\left.\beta\exp\left[-\eta\left(t-\tau-\tau_{1}\right)\right]\int_{-\infty}^{t-\tau-\tau_{1}}d\tau_{3}\exp\left(\eta\tau_{3}\right)\dot{n}\left(\tau_{3}\right)\right\}\end{aligned} \tag{23}$$

Let us introduce the auxiliary variables

$$r_{1}\left(t\right)=\iint d\tau d\tau_{1}p_{\mathrm{EI}}\left(\tau,\tau_{1}\right)\exp\left[-\lambda\left(t-\tau-\tau_{1}\right)\right]\int_{-\infty}^{t-\tau-\tau_{1}}d\tau_{2}\exp\left(\lambda\tau_{2}\right)\dot{n}\left(\tau_{2}\right)$$

$$r_{2}\left(t\right)=\iint d\tau d\tau_{1}p_{\mathrm{EI}}\left(\tau,\tau_{1}\right)\exp\left[-\eta\left(t-\tau-\tau_{1}\right)\right]\int_{-\infty}^{t-\tau-\tau_{1}}d\tau_{3}\exp\left(\eta\tau_{3}\right)\dot{n}\left(\tau_{3}\right)$$

so that $r = \gamma r_1 + \beta r_2$. Taking the time derivative and recalling Eqs. (10, 11) we get

$$\dot{r}_1(t) = -\lambda r_1(t) + \rho i(t) \tag{24}$$

$$\dot{r}_2(t) = -\eta r_2(t) + \rho i(t) \tag{25}$$

$$\dot{r}(t) = -\lambda\gamma r_1(t) - \beta\eta r_2(t) + \rho i(t) \tag{26}$$

and similarly

$$\dot{h}(t) = \lambda\gamma r_1(t) \tag{27}$$

$$\dot{d}(t) = \beta\eta r_2(t) \tag{28}$$

We then have shown that the entire class SIR/SEIR and similar models are a particular case of the present theory.

3 The basic and the effective reproduction numbers

Now let us calculate the number of secondary infections $\chi(\tau_1)$, at the beginning of infection out-break, caused by a newly infectious individual (with infectious latent time τ_1) in a totally susceptible population. This number is $\chi(\tau_1) = \int_0^{\tau_1} \nu(t)\, dt$, where we have assumed that the infection rate depends on time t. Then, the average number of secondary infections caused by a single newly infectious individual, i.e. the so called basic reproduction numbers R_0, is

$$R_0 = \int_0^{+\infty} d\tau_1 p_{\mathrm{I}}(\tau_1)\, \chi(\tau_1) = \int_0^{+\infty} d\tau_1 p_{\mathrm{I}}(\tau_1) \int_0^{\tau_1} \nu(t)\, dt = \langle \tau_1 \nu_m(\tau_1) \rangle \tag{29}$$

where $\nu_m(\tau_1) = \tau_1^{-1} \int_0^{\tau_1} v_m(t)\, dt$ and $p_{\mathrm{I}}(\tau_1) = \int d\tau p_{\mathrm{EI}}(\tau, \tau_1)$ is the density probability function of latent times τ_1 in the infectious compartment. Note that the above definition is consistent with the usual definition in terms of survival function $\mathcal{F}_{\mathrm{I}}(\tau_1) = P(t \geq \tau_1)$, where $P(t \geq \tau_1)$ is the probability for a newly infectious individual to be infectious up to time τ_1. In fact, let us observe that $\mathcal{F}_{\mathrm{I}}(\tau_1) = 1 - \int_{-\infty}^{\tau_1} p_{\mathrm{I}}(\theta)\, d\theta$, which yields

$$\frac{d\mathcal{F}_{\mathrm{I}}}{d\tau_1} = \mathcal{F}_{\mathrm{I}}'(\tau_1) = \delta(\tau_1) - p_{\mathrm{I}}(\tau_1) \tag{30}$$

Replacing Eq. (30) in Eq. (29) we get

$$R_0 = \int_0^{+\infty} d\tau_1 \nu\left(\tau_1\right) \mathcal{F}_{\mathrm{I}}\left(\tau_1\right) > 0 \tag{31}$$

where we used integration by parts. Eq. (31) is the usual first principle formulation of the basic reproduction number (*42, 43*).

In order to monitor the evolution of infection spreading it could be useful to define an instantaneous effective reproduction number as

$$R_{\mathrm{t}}\left(t\right) = \int_0^{+\infty} d\tau_1 p_1\left(\tau_1\right) \int_t^{t+\tau_1} \sigma\left(\zeta\right) \nu\left(\zeta\right) d\zeta \tag{32}$$

where $\sigma = s/N = \left(1 - n/N\right)$ is the instantaneous fraction of susceptible population. Using Eq. (30) and integrating by parts one obtains

$$R_{\mathrm{t}}\left(t\right) = \int_0^{+\infty} d\tau_1 \sigma\left(t + \tau_1\right) \nu\left(t + \tau_1\right) \mathcal{F}_{\mathrm{I}}\left(\tau_1\right) \tag{33}$$

The reproduction number defined as in Eq. (33) can be also defined as the cohort reproduction number (*44*).

4 Deriving the Kermack and McKendrick integro-differential equations

Here we show that the present theory allows for an clear physical interpretation of the kernels of the integro-differential equations presented by Kermack and McKendrick (*9*). To this end note that Eqs. (9, 10, 11, 12, 13) can be rephrased as (see Sec. B for the detailed derivation)

$$e\left(t\right) = \int d\tau \mathcal{F}_{\mathrm{E}}\left(\tau\right) \dot{n}\left(t - \tau\right) \tag{34}$$

where the survival function is $\mathcal{F}_{\mathrm{E}}\left(\tau\right) = 1 - \int_{-\infty}^{\tau} p_{\mathrm{E}}\left(t\right) dt$ so that $\mathcal{F}_{\mathrm{E}}'\left(\tau\right) = \delta\left(\tau\right) - p_{\mathrm{E}}\left(\tau\right)$, and we have used integration by parts. Following the same argument Eq. (10), can be recasted as

$$i\left(t\right) = \int d\tau \left\{\mathcal{F}_1\left(\tau\right) - \mathcal{F}_{\mathrm{E}}\left(\tau\right)\right\} \dot{n}\left(t - \tau\right) \tag{35}$$

where we have changed the integration variable from τ_1 to $\theta_1 = \tau + \tau_1$ and used that the probability density function $p_1(\theta_1)$ of the quantity θ_1 is $p_1(\theta_1) = \int d\tau p_{\mathrm{EI}}(\tau, \theta_1 - \tau)$. The corresponding survival function is $\mathcal{F}_1(\tau) = 1 - \int_{-\infty}^{\tau} p_1(t)\, dt$. Note that θ_1 is the stochastic time span from the occurrence of infection to the time of removal. Analogously we get

$$r(t) = \int d\tau \left\{\gamma \mathcal{F}_2(\tau) + \beta \mathcal{F}_3(\tau) - \mathcal{F}_1(\tau)\right\} \dot{n}(t - \tau) \tag{36}$$

where $\mathcal{F}_2(\theta_2)$ is the survival function of the time variable $\theta_2 = \tau + \tau_1 + \tau_2$ and $\mathcal{F}_3(\theta_3)$ is the survival function of $\theta_3 = \tau + \tau_1 + \tau_3$. Now replacing Eq. (10) in Eq. (2) we get for $t > 0$

$$\dot{n}(t) = \nu \left(1 - \frac{n}{N}\right) \int d\tau \left\{\mathcal{F}_1(\tau) - \mathcal{F}_{\mathrm{E}}(\tau)\right\} \dot{n}(t - \tau) + n_0 \delta(t) \tag{37}$$

Eqs. (35, 36, 37) are formally equivalent to Eqs. (15, 14, 13) reported by Kermack and McKendrick in their paper (1927) (*9*). So we can interpret the Kernels of the Kermack and McKendrick model in terms of survival functions of latent times. Such distribution are crucial quantities which govern the infection spreading and can be measured from observed data (see for example Ref. (*45*)). Note also that for the healed and deceased compartments the following equations hold true

$$h(t) = \gamma n(t) - \gamma \int \mathcal{F}_2(\tau)\, \dot{n}(t - \tau)\, d\tau \tag{38}$$

$$d(t) = \beta n(t) - \beta \int \mathcal{F}_3(\tau)\, \dot{n}(t - \tau)\, d\tau \tag{39}$$

It is worth noticing that all the equations are time-invariant linear equations except for Eq. (37) which is non linear.

5 Linearized equations and stability analysis

In order to study the stability of the infection spreading it is enough to look at the Eq. (37) under the assumption that the ratio $n/N \ll 1$. This allows to look at the linearized the equation to get

$$\dot{n}(t) = \nu \int d\tau \left\{\mathcal{F}_1(\tau) - \mathcal{F}_{\mathrm{E}}(\tau)\right\} \dot{n}(t - \tau) + n_0 \delta(t) \tag{40}$$

Note that $\mathcal{F}_1(\tau) > \mathcal{F}_{\mathrm{E}}(\tau)$, $\lim_{\tau\to 0^+}\{\mathcal{F}_1(\tau) - \mathcal{F}_{\mathrm{E}}(\tau)\} = 0$, $\lim_{\tau\to +\infty}\{\mathcal{F}_1(\tau) - \mathcal{F}_{\mathrm{E}}(\tau)\} = 0$, and $\int d\tau\{\mathcal{F}_1(\tau) - \mathcal{F}_{\mathrm{E}}(\tau)\} = \langle\theta_1\rangle - \langle\tau\rangle = \langle\tau_1\rangle$. These conditions allow us to define the probability density function $q(\zeta)$ of the positive stochastic time variable $\zeta > 0$, as

$$q(\zeta) = \frac{\mathcal{F}_1(\zeta) - \mathcal{F}_{\mathrm{E}}(\zeta)}{\langle\tau_1\rangle} > 0 \tag{41}$$

Note in fact that $\int q(\zeta)\, d\zeta = 1$. Also observe that $q(\zeta = 0) = 0$ and $q(\zeta \to +\infty) = 0$. The moment m_k of $q(\tau)$ are

$$\begin{aligned} m_k &= \int d\zeta q(\zeta)\zeta^k = \frac{1}{\langle\tau_1\rangle}\frac{1}{k+1}\int d\zeta\{p_1(\zeta) - p_{\mathrm{E}}(\zeta)\}\zeta^{k+1} \\ &= \frac{1}{\langle\tau_1\rangle}\frac{\langle\theta_1^{k+1}\rangle - \langle\tau^{k+1}\rangle}{k+1} > 0; \qquad k = 0, 1, 2, ... \end{aligned} \tag{42}$$

where we used integration by parts and considered that the survival function exponentially decreases at large ζ. The positiveness of m_k follows directly from the binomial theorem and considering that both τ and τ_1 are larger than zero. Note in fact that

$$\langle\theta_1^n\rangle - \langle\tau^n\rangle =_{h=1}^{n} \frac{n!}{h!\,(n-h)!}\langle\tau^{n-h}\tau_1^h\rangle > 0$$

Eq. (40) can be, then, recasted in the following form

$$\dot{n}(t) = R_0\int d\tau q(\tau)\,\dot{n}(t-\tau) + n_0\delta(t) \tag{43}$$

where we used that $R_0 = \nu\langle\tau_1\rangle$ or equivalently as

$$n_0\delta(t) = (1 - R_0)\,\dot{n}(t) + R_0\int d\tau\,[\delta(t) - q(\tau)]\,\dot{n}(t-\tau)$$

Then integrating over time t and recalling that $\mathcal{F}_q'(\tau) = \delta(\tau) - q(\tau)$,

$$n_0\mathcal{H}(t) = (1 - R_0)\,n(t) + R_0\int d\tau\mathcal{F}_q'(\tau)\,n(t-\tau)$$

Integrating by parts and using the commutativity of the convolution product we get

$$n_0\mathcal{H}(t) = \int d\tau G_{\mathrm{R}_0}(t)\,\dot{n}(\tau) \tag{44}$$

where

$$G_{\mathrm{R}_0}(t) = (1 - R_0)\,\mathcal{H}(t) + R_0 \mathcal{F}_q(t) \tag{45}$$

Being $\mathcal{F}_{\mathrm{q}}(t < 0) = 0$, i.e. $\mathcal{F}_{\mathrm{q}}(t)$ satisfies the causality principle, and considering that $\mathcal{F}_{\mathrm{q}}(\tau \to +\infty) = 0$, $\mathcal{F}_q(\tau)$ can be re-interpreted, in mechanical terms, as the relaxation function of linear fluid with memory. The term $(1 - R_0)\,\mathcal{H}(\tau)$ represents instead a linear spring with stiffness $1 - R_0$. So let us consider the linear system

$$F(t) = \int d\tau G_{\mathrm{R}_0}(t - \tau)\,\dot{u}(\tau) \tag{46}$$

where the input is $u(t) = n(t)$ can be interpreted as a displacement, and the output $F(t)$ can be interpreted as a force. By taking the Laplace transform of both hand sides we have

$$F(s) = K_{\mathrm{R}_0}(s)\,u(s) \tag{47}$$

where $K_{R_0}(s)$ is complex stiffness of the system, i.e.

$$K_{\mathrm{R}_0}(s) = sG_{\mathrm{R}_0}(s) = (1 - R_0) + R_0 s \mathcal{F}_{\mathrm{q}}(s) \tag{48}$$

We also have

$$u(s) = C_{\mathrm{R}_0}(s)\,F(s) \tag{49}$$

where the complex compliance is

$$C_{\mathrm{R}_0}(s) = \frac{1}{1 - R_0 + R_0 s \mathcal{F}_{\mathrm{q}}(s)} \tag{50}$$

Focusing on Eq. (47) stability requires that the poles of the complex stiffness have all strictly negative real parts. This is indeed the case for the quantity $s\mathcal{F}_{\mathrm{q}}(s)$. In fact recalling that $\mathcal{F}_{\mathrm{q}}(t)$ satisfies causality principle, i.e., $\mathcal{F}_{\mathrm{q}}(t < 0) = 0$, it follows that the Laplace transform of $\mathcal{F}_{\mathrm{q}}(t)$ i.e. $\mathcal{F}_{\mathrm{q}}(s) = \mathcal{L}[\mathcal{F}_{\mathrm{q}}(t)](s)$, is analytic in the right half of the complex plane, i.e. all the poles of $\mathcal{F}_{\mathrm{q}}(s)$ have strictly negative real parts. Regarding the spring term $(1 - R_0)$ in Eq. (48) stability

also requires that the stiffness of the spring is positive, that is $R_0 < 1$. One concludes that, when the controlled variable is the displacement $u(t)$, the linear system defined by the equation $F(s) = K_{\mathrm{R}_0}(s)\, u(s)$ is stable provided that $R_0 < 1$. In such a case it is easily shown that also the quantity

$$C_{\mathrm{R}_0}(t) = \mathcal{L}^{-1}\left[C_{\mathrm{R}_0}(s)\right](t) = \frac{1}{2\pi j}\int_{0^+-i\infty}^{0^++i\infty} ds C_{\mathrm{R}_0}(s)\, e^{st} \tag{51}$$

(where $j = \sqrt{-1}$) satisfies the causality principle, so that $C_{\mathrm{R}_0}(s)$ is analytic in the right half of the complex plane. Note that in terms of $C_{\mathrm{R}_0}(t)$ we can write

$$u(t) = \int d\tau C_{\mathrm{R}_0}(t-\tau)\, F(\tau) \tag{52}$$

The stability of the system is then only governed by the basic reproduction number R_0 and is guaranteed for $0 < R_0 < 1$. In such conditions being $F(t) = n_0 \mathcal{H}(t)$, i.e. $F(s) = n_0/s$, the epidemic will vanish in the long term with a total number of infected individuals equal to

$$n_\infty = \lim_{t\to+\infty} n(t) = \lim_{s\to 0} s n(s) = \lim_{s\to 0} s C_{\mathrm{R}_0}(s)\, F(s) = n_0 \lim_{s\to 0} C_{\mathrm{R}_0}(s) = \frac{n_0}{1-R_0} \tag{53}$$

We will also have $i_\infty = e_\infty = r_\infty = 0$ and $h_\infty = \gamma n_\infty$ and $d_\infty = \beta n_\infty$.

6 The case $R_0 = 1$

In this section we focus on what happens when $R_0 = \nu\langle\tau_1\rangle = 1$ under the hypothesis that $n(t)/N \ll 1$. In such a case we have

$$K_{\mathrm{R}_0=1}(s) = K_1(s) = s\mathcal{F}_{\mathrm{q}}(s) \tag{54}$$

$$C_{\mathrm{R}_0=1}(s) = C_1(s) = \frac{1}{s\mathcal{F}_{\mathrm{q}}(s)} \tag{55}$$

Therefore the system will behave as linear fluid with memory. Being $\lim_{s\to 0} s\mathcal{F}_{\mathrm{q}}(s) = \mathcal{F}_{\mathrm{q}}(t\to+\infty) = 0$, the complex compliance $C_1(s)$ has a pole at $s = 0$, all the others being in the left-half of the

complex plane. Now, recalling that $s\mathcal{F}_{\mathrm{q}}(s) = 1 - q(s)$ and considering that

$$q(s) = \int d\tau q(t) e^{-st} = \langle \exp(-st) \rangle = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} m_k s^k = 1 - m_1 s + \frac{1}{2} m_2 s^2 - ...$$

for $s \to 0$ at first order in s we have $s\mathcal{F}_{\mathrm{q}}(s) = 1 - q(s) = m_1 s$. Then, $C_0(s) = 1/(m_1 s)$ and $u(s) = F(s)/(m_1 s)$. Recalling that, in our case, $F(t) = n_0 \mathcal{H}(t)$, i.e. $F(s) = n_0/s$, we get that in the long term the number of infected people $n(t)$ will increase linearly with a rate

$$\dot{n}_\infty = \lim_{s \to 0^+} s^2 u(s) = s^2 C_0(s) \frac{n_0}{s} = \frac{n_0}{m_1} = \frac{2 \langle \tau_1 \rangle n_0}{\langle \theta_1^2 \rangle - \langle \tau^2 \rangle} \tag{56}$$

Eq. (56) shows that the first moment m_1 of the probability density function $q(\tau)$ governs the velocity of spreading when $R_0 = 1$. In the long term we also have

$$e_\infty = \int d\tau \mathcal{F}_{\mathrm{E}}(\tau) \dot{n}_\infty = \langle \tau \rangle \dot{n}_\infty \tag{57}$$

Similarly

$$i_\infty = \int d\tau \left\{ \mathcal{F}_1(\tau) - \mathcal{F}_{\mathrm{E}}(\tau) \right\} \dot{n}_\infty = \langle \tau_1 \rangle \dot{n}_\infty \tag{58}$$

and

$$r_\infty = \int d\tau \left\{ \gamma \mathcal{F}_2(\tau) + \beta \mathcal{F}_3(\tau) - \mathcal{F}_1(\tau) \right\} \dot{n}_\infty = \langle \gamma \tau_2 + \beta \tau_3 \rangle \dot{n}_\infty \tag{59}$$

Regarding the healed individuals, taking the time derivative of

$$h(t) = \gamma n(t) - \gamma \int \mathcal{F}_2(\tau) \dot{n}(t - \tau) d\tau \tag{60}$$

and observing that in the long term $\ddot{n}_\infty = 0$ yields

$$\dot{h}_\infty = \gamma \dot{n}_\infty \tag{61}$$

Similarly for the deceased people

$$\dot{d}_\infty = \beta \dot{n}_\infty$$

7 Case study: the COVID-19 epidemic in Italy

In this Section we use the proposed model to simulate the spread of COVID-19 disease in Italy during the first 120 days from the first COVID-19 outbreak in the country. As already observed, the knowledge of the infection rate $\nu(t)$ and joint probability density function $p_{\mathrm{EI}}(\tau, \tau_1)$ is crucial to described the spreading of the disease. To estimate $p_{\mathrm{E}}(\tau)$ we use data reported in Ref. (*45*) (for the China case), which allow an estimation of the following probability density functions: (i) $q_{\mathrm{IN}}(t_1)$ where t_1 is the incubation period (i.e., the time from infection to illness onset), and (ii) $q_{\mathrm{IR}}(t_2)$ with t_2 being the time interval from illness onset to first medical visit. In particular, assuming that on the average an individual becomes infectious λ days before the appearance of the symptoms, we can estimate the distribution $p_{\mathrm{E}}(\tau)$ from $q_{\mathrm{IN}}(t_1)$ by simply scaling the time axis of the quantity $a = \langle t_1 \rangle^{-1} (\langle t_1 \rangle - \lambda)$, where $\langle t_1 \rangle = \int dt_1 t_1 q_{\mathrm{IN}}(t_1)$, to get

$$p_{\mathrm{E}}(\tau) = a^{-1} q_{\mathrm{IN}}\left(a^{-1}\tau\right) \tag{62}$$

Now note that the sum of the incubation period t_1 plus the period t_2 from illness to first medical visit (we assume that at the first medical visit the infected individual is isolated) obeys the relation $t_1 + t_2 = \tau + \tau_1$ (where τ is the latent time in the exposed compartment and τ_1 the latent time in the infectious compartment), then we have that

$$p_1(\theta_1) = \int dt_1 q_{\mathrm{IN}}(t_1)\, q_{\mathrm{IR}}(\theta_1 - t_1) \tag{63}$$

Fig. 2(a) shows the probability density function $p_{\mathrm{E}}(\tau)$ of the latent time τ in the exposed compartment for $\lambda = 2$ days (note that researchers estimate that people who get infected with the coronavirus are most contagious 2 days before the illness onset). Fig. 2(b) reports the probability density function $p_1(\theta_1)$ of the latent time $\theta_1 = \tau + \tau_1$ obtained using Eq. (63).

Now let us focus on the latent times τ_2 (i.e. the sojourn time the individuals spend into the removed compartment before healing) and τ_3 (i.e. the time period spent in the removed

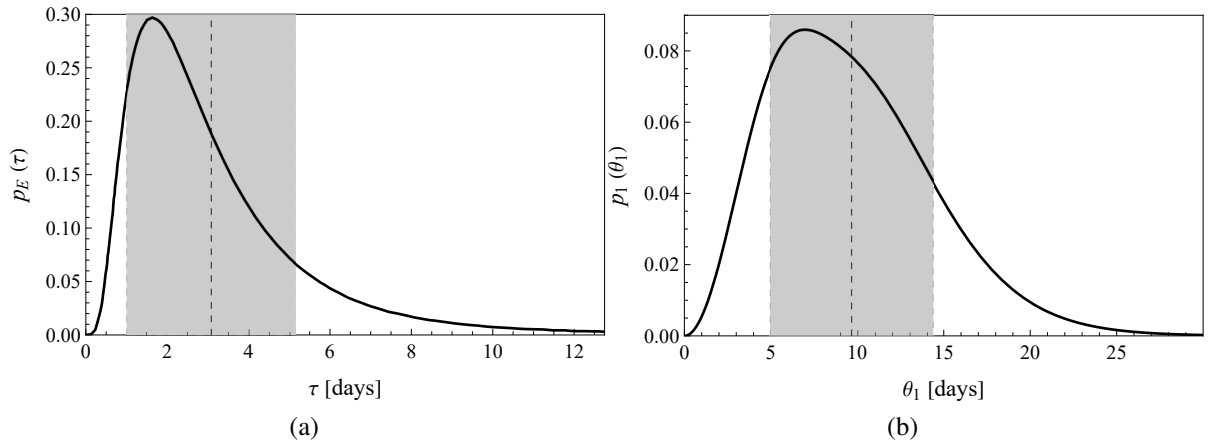


Figure 2: The probability density function $p_{\mathrm{E}}(\tau)$ of the latent times τ (sojourn time in the exposed compartment) for $\lambda = 2$ days (a); the probability density function $p_{\mathrm{I}}(\tau_1)$ of the latent times τ_1 (sojourn time in the infectious compartment).

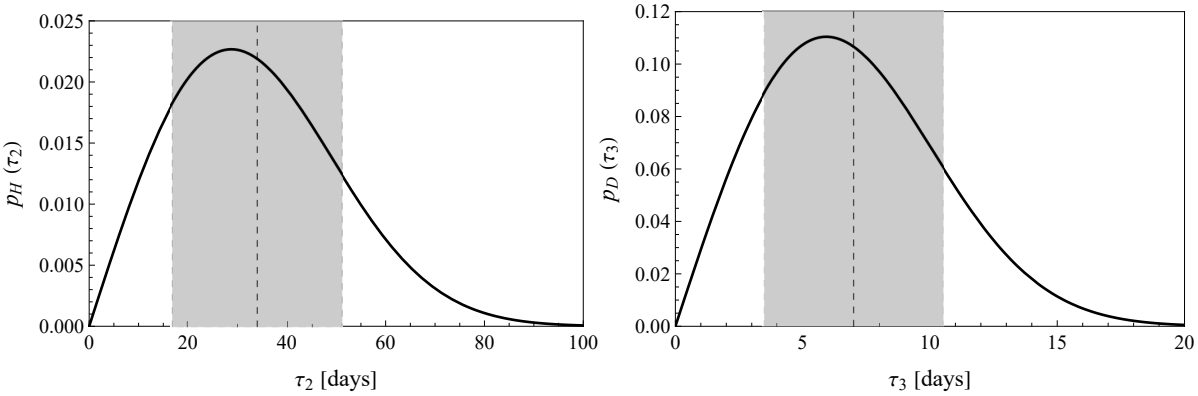
compartment before death). In this case the density probability functions $p_{\mathrm{H}}(\tau_2)$ and $p_{\mathrm{D}}(\tau_3)$ have been estimated from the data available from Refs. (*46–48*). In particular the distribution of latent time frequencies have been fitted through a double Gaussian distribution

$$p(\tau) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{erf\left(\mu/\sqrt{2\sigma^2}\right)} \left\{ \exp\left[-\frac{(\tau-\mu)^2}{2\sigma^2}\right] - \exp\left[-\frac{(\tau+\mu)^2}{2\sigma^2}\right] \right\} \qquad \tau \geq 0; \quad (64)$$

The fit parameters are μ and σ. Fig. 3(a) shows the probability density function $p_{\mathrm{H}}(\tau_2)$ of the latent time τ_2 from isolation to recovery ($\mu_{\mathrm{H}} = 34, \sigma_{\mathrm{H}} = 17$), whereas Fig. 2(b) reports the probability density function $p_{\mathrm{D}}(\tau_3)$ of the latent time τ_3 from isolation to death ($\mu_{\mathrm{D}} = 7, \sigma_{\mathrm{D}} = 3.5$). We also report the trend of the probability density functions $p_2(\theta_2)$ of $\theta_2 = \theta_1 + \tau_2$ [see Fig. 4(a)] and $p_3(\theta_3)$ where $\theta_3 = \theta_1 + \tau_3$ [see Fig. 4(b)]. Table 65 reports the average values and the standard deviation of the probability density function of times τ, θ_1, τ_2, θ_2, τ_3, θ_3.

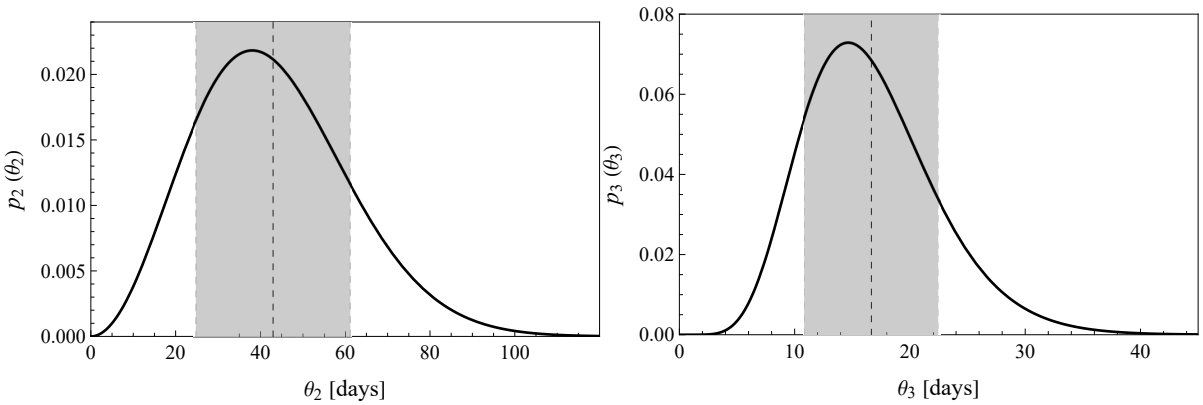
Prob. dens. function	average (days)	standard deviation (days)
$p_{\mathrm{E}}(\tau)$	3.08	2.07
$p_1(\theta_1)$	9.66	4.68
$p_{\mathrm{H}}(\tau_2)$	34.0	17.0
$p_{\mathrm{D}}(\tau_3)$	7.00	3.50
$p_2(\theta_2)$	42.99	18.14
$p_3(\theta_3)$	16.67	5.77

(65)



(a) (b)

Figure 3: The probability density function $p_{\mathrm{H}}(\tau_2)$ of the latent times τ_2 (latency time from isolation to healing) (a); the probability density function $p_{\mathrm{D}}(\tau_3)$ of the latent times τ_3 (latency time from isolation to death), (b).



(a) (b)

Figure 4: The probability density function $p_2(\theta_2)$ of $\theta_2 = \tau + \tau_1 + \tau_2$ (latency time from infection to healing) (a); the probability density function $p_3(\theta_3)$ of $\theta_3 = \tau + \tau_1 + \tau_3$ (latency time from infection to death), (b).

In order to simulate the effect of the control measures undertaken by the government (e.g. lockdown, quarantine and mask obligation) to limit the diffusion of the infection, we implement an exponential decrease of the infection rate $\nu = \nu(t)$ (*49, 50*), which takes place starting from the time T_{Q} at which control measure are enforced, i.e.

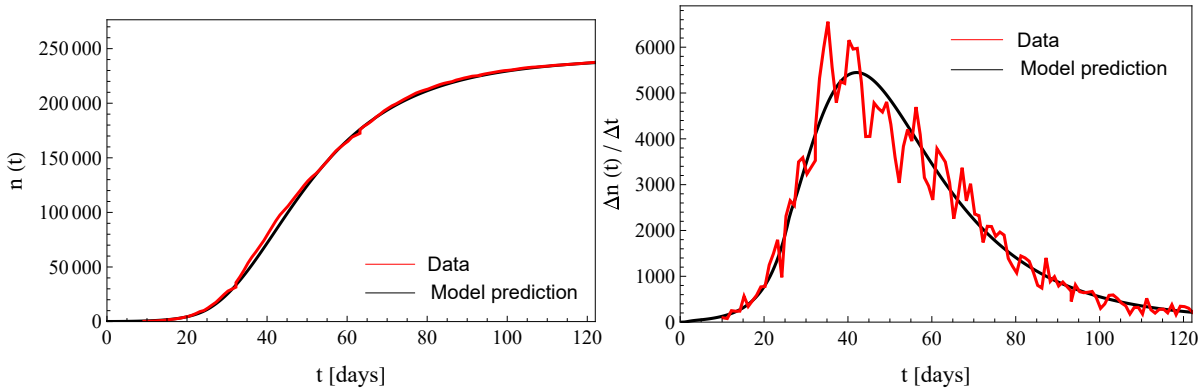
$$\begin{aligned} &\nu(t) = \nu_0; \quad t \leq T_{\mathrm{Q}} \\ &\nu(t) = \nu_0 (1-\alpha) \exp\left(-\frac{t - T_{\mathrm{Q}}}{\tau_e}\right) + \nu_0 \alpha; \quad t > T_{\mathrm{Q}} \end{aligned} \tag{66}$$

where τ_e is a time constant and $\nu_0 \alpha$ is the latent infection rate originating from imperfect social distancing.

The set of parameters used to run the model for the Italy case is $n_0 = n(t=0) = 117$, $\nu_0 = 0.503$, $T_{\mathrm{Q}} = 25.7$ days, $\tau_e = 10.41$ days, $\alpha = 0.167$, $\gamma = 0.86$, $\beta = 0.14$. Figure 5(a) shows the time evolution of the cumulative number $n(t)$ of infected cases during the first 120 days from the first COVID-19 outbreak in Italy, whereas Fig. 5(b) shows the daily new infected cases $\Delta n / \Delta t \approx \dot{n}(t)$ during the same period of time. Model predictions are in very good agreement with the real data, this can be accessed at the URL: https://opendatadpc.maps.arcgis.com/apps/dashboards/b0c68bce2cce478eaac82fe38d4138b1 Let us estimate the basic reproduction number R_0. To this end we first consider that $p_{\mathrm{I}}(\tau_1)$ is vanishing small for latent times $\tau_1 > T_{\mathrm{Q}}$ therefore the contribution to the τ_1-integral in Eq. (29) is limited to times $\tau_1 < T_{\mathrm{Q}}$. Thus, considering that $\nu(t) = \nu_0$ for $t \leq T_{\mathrm{Q}}$ we get

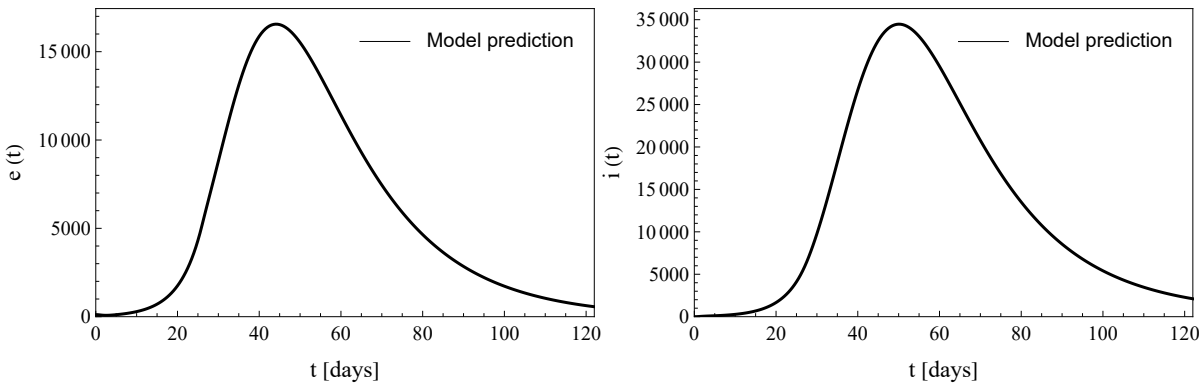
$$R_0 \approx \nu_0 \int_0^{\infty} d\tau_1 p_{\mathrm{I}}(\tau_1) \tau_1 \approx \nu_0 \langle \tau_1 \rangle$$

Recalling that $\langle \tau_1 \rangle \approx 7$ days, we get $R_0 \approx 3.5$ which is close the value found in other studies (*51, 52*). Figure 6(a) shows the time evolution of the cumulative number $e(t)$ of exposed cases, whereas Fig. 6(b) shows the cumulative number $i(t)$ of individuals in the infectious compartment. Figure 7(a) shows the time evolution of the number $r(t)$ of individuals in the removed



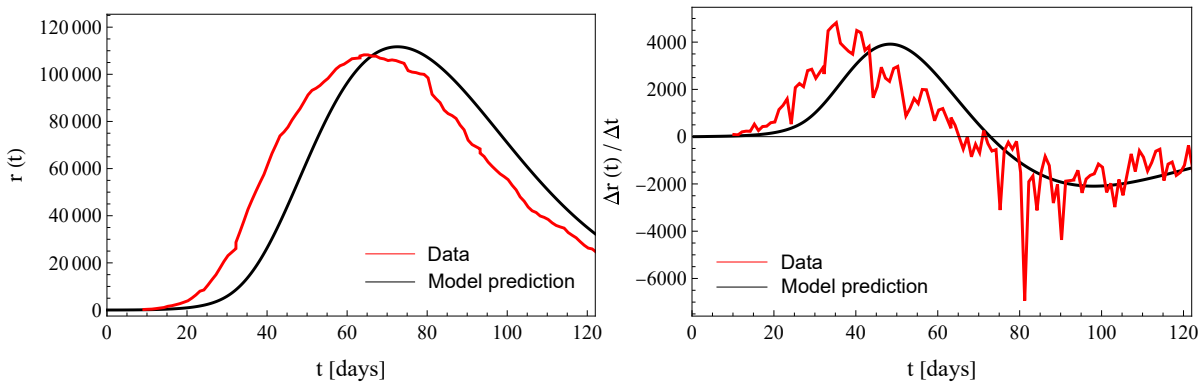
(a) (b)

Figure 5: The time evolution of the cumulative number of infected cases during the first 120 days of the first COVID outbreak in Italy, (a); the daily new infected cases during the same period of time, (b).



(a) (b)

Figure 6: The time evolution of the cumulative number of exposed cases, (a); the cumulative number of individuals in the infectious compartment, (b).



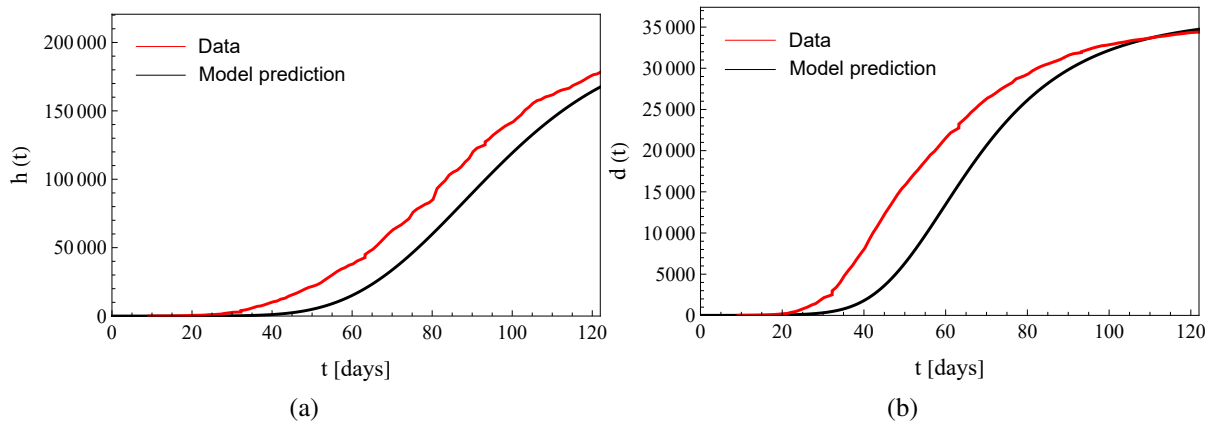
(a) (b)

Figure 7: The time evolution of individuals in the removed compartment, (a); the evolution of the daily removed individuals, (b). Model predictions are compared with real data. A certain time-lag is present between model prediction and the real data.

compartment and Figure 7(b) the evolution of the daily removed individuals $\Delta r/\Delta t \approx \dot{r}(t)$. Model predictions are compared with real data. We notice a certain time-lag between the model predictions and the real data set for the removed compartment, the theoretical and real curve run, indeed, parallel to each-other. The reason for such a difference does not have a clear interpretation. However, we believe the discrepancy should be related to an asynchronous recording of the number of infected individual and the number removed ones. Figure 8(a) shows the time evolution of the number $h(t)$ of recovered (healed) individuals and Figure 8(b) the number $d(t)$ of the deceased individuals. Also in this case a time-lag is observed similar to the one observed for the compartment of removed individuals.

8 Conclusion

We present a general theory of infection dynamics, leading to Volterra like equations where the kernels have a clear physical interpretation in terms of probability density function or in terms of survival function of latent time distributions, thus overcoming the main difficulty of Kermack and McKendrick model (1927), where the interpretation of the physical meaning of



(a) (b)

Figure 8: The time evolution of healed individuals, (a); the evolution of deaths, (b). Model predictions are compared with real data. A certain time-lag between model prediction and the real data set, which is the same as in the case of the removed individuals.

the kernels was really unclear and difficult to interpret in terms of real data. Our theory contains as a particular case the whole class of SIR, SEIR. Beside the infection rate $\nu(t)$, a very central role is played by the joint probability density function $p_{\mathrm{EI}}(\tau, \tau_1)$ of the latent times τ and τ_1 the individuals spend in the exposed and infectious compartments respectively. The quantity $p_{\mathrm{EI}}(\tau, \tau_1)$ is then a crucial information needed to predict how the infection will develop in time. The theory does not pose any restriction to the functional form of the probability density functions of compartmental latent times, which usually do not obey the exponential distribution, as, instead, implicitly assumed in classical SIR and SEIR-type models. When the number of infections is negligible with the entire population, the equations can be solved in general and their stability studied by resorting their representation in the Laplace domain. The particular case of $R_0 = 1$ is also investigated to show that the moments of the probability density functions of latent times, govern the asymptotic infection rate in this case. The present theory has been employed to simulate the spreading of COVID-19 infection in Italy during the first 120 days subsequent to the identification of the first cases. We found a very good agreement between theoretical predictions and real data. Our estimation of the basic reproduction number is $R_0 \approx$

3.5 in agreement with other studies.

A Compartmental dynamic equations

Consider the number δm_k of exposed individuals with latent time τ_k which enters, during a time interval δt, the infectious compartment at time t. Of such newly infectious individuals a certain number δl_{kh}, will leave the infectious department after a latent time τ'_h, the fraction $\delta l_{kh}/\delta m_k$ is

$$\frac{\delta l_{kh}}{\delta m_k} = p_{\mathrm{I}}\left(\tau'_h|\tau_k\right)\Delta\tau' \tag{67}$$

where $p_{\mathrm{I}}\left(\tau'|\tau\right)$ is the conditional probability density function of the infectious latent times τ' given the exposed latent time τ. Therefore the increase δi_k of infectious individuals at time t due to the contribution of exposed individuals δm_k with latent time τ_k is

$$\begin{aligned}\delta i_k &= \delta m_k - \sum_h \delta l_{kh} = \delta m_k - \sum_h p_{\mathrm{I}}\left(\tau'_h|\tau_k\right)\delta m_k\left(t-\tau'_h\right)\Delta\tau' \\ &= \delta n_k\left(t-\tau_k\right) - \sum_h p_{\mathrm{I}}\left(\tau'_h|\tau_k\right)\delta n_k\left(t-\tau'_h-\tau_k\right)\Delta\tau' \\ &= p_{\mathrm{E}}\left(\tau_k\right)\delta n\left(t-\tau_k\right)\Delta\tau - \sum_h p_{\mathrm{I}}\left(\tau'_h|\tau_k\right)p_{\mathrm{E}}\left(\tau_k\right)\delta n\left(t-\tau'_h-\tau_k\right)\Delta\tau'\Delta\tau\end{aligned} \tag{68}$$

and summing up over the index k

$$\delta i\left(t\right) = \sum_k \delta i_k = \sum_k p_{\mathrm{E}}\left(\tau_k\right)\delta n\left(t-\tau_k\right)\Delta\tau - \sum_{kh} p_{\mathrm{I}}\left(\tau'_h|\tau_k\right)p_{\mathrm{E}}\left(\tau_k\right)\delta n\left(t-\tau'_h-\tau_k\right)\Delta\tau'\Delta\tau \tag{69}$$

or

$$\frac{di\left(t\right)}{dt} = \sum_k \frac{di_k}{dt} = \sum_k p_{\mathrm{E}}\left(\tau_k\right)\dot{n}\left(t-\tau_k\right)\Delta\tau - \sum_{kh} p_{\mathrm{I}}\left(\tau'_h|\tau_k\right)p_{\mathrm{E}}\left(\tau_k\right)\dot{n}\left(t-\tau'_h-\tau_k\right)\Delta\tau'\Delta\tau \tag{70}$$

Taking the limit of vanishing $\Delta\tau$ and $\Delta\tau'$ gives

$$\frac{di\left(t\right)}{dt} = \int d\tau p_{\mathrm{E}}\left(\tau\right)\dot{n}\left(t-\tau\right) - \iint d\tau d\tau' p_{\mathrm{EI}}\left(\tau,\tau'\right)\dot{n}\left(t-\tau-\tau'\right) \tag{71}$$

where we have used that $p_{\mathrm{I}}\left(\tau'|\tau\right)p_{\mathrm{E}}\left(\tau\right) = p_{\mathrm{EI}}\left(\tau,\tau'\right)$, with τ and τ' both greater or equal to zero. Integrating over time t and renaming $\tau' \to \tau_1$ give

$$i\left(t\right) = \int d\tau p_{\mathrm{E}}\left(\tau\right) n\left(t-\tau\right) - \iint d\tau d\tau_1 p_{\mathrm{EI}}\left(\tau,\tau_1\right) n\left(t-\tau-\tau_1\right) \tag{72}$$

Recalling Eq. (9) we then get Eq. (3).

A similar argument applies to the removed compartment. Let us consider δl_{kh} individuals with latent time τ_k and τ'_h entering, during the time step δt, the removed compartment at time t. Of such individuals a number $\delta q'_{khs}$ will recover after a latent time τ''_s and another number $\delta q''_{khp}$ with latent time τ'''_p will enter the deceased compartment. We now consider that in the long term a fraction γ of the individuals will recover and the fraction $\beta = 1-\gamma$ will move into the deceased compartment. The we can write

$$\frac{\delta q'_{khs}}{\gamma \delta l_{kh}} = p_{\mathrm{H}}\left(\tau''_s|\tau'_h,\tau_k\right)\Delta\tau'' \tag{73}$$

$$\frac{\delta q''_{khp}}{\beta \delta l_{kh}} = p_{\mathrm{D}}\left(\tau'''_p|\tau'_h,\tau_k\right)\Delta\tau''' \tag{74}$$

therefore the increase δr_{kh} of individuals in the removed compartment at time t due to the contribution of infectious individuals δl_{kh} with latent times τ_k and τ'_h is

$$\begin{aligned}\delta r_{kh} &= \delta l_{kh} - \sum_s \delta q'_{khs} - \sum_p \delta q''_{khp} \qquad (75)\\ &= \delta l_{kh} - \gamma \sum_s p_{\mathrm{H}}\left(\tau''_s|\tau'_h,\tau_k\right)\delta l_{kh}\left(t-\tau''_s\right)\Delta\tau'' - \beta\sum_p p_{\mathrm{D}}\left(\tau'''_p|\tau'_h,\tau_k\right)\delta l_{kh}\left(t-\tau'''_p\right)\Delta\tau'''\end{aligned}$$

consider now that $\delta l_{kh}\left(t\right) = p_{\mathrm{I}}\left(\tau'_h|\tau_k\right)\Delta\tau'\delta m_k\left(t-\tau'_h\right)$ and $\delta m_k\left(t-\tau'_h\right) = p_{\mathrm{E}}\left(\tau_k\right)\Delta\tau\delta n\left(t-\tau'_h-\tau_k\right)$. Then, recalling that $p_{\mathrm{EI}}\left(\tau_k,\tau'_h\right) = p_{\mathrm{I}}\left(\tau'_h|\tau_k\right)p_{\mathrm{E}}\left(\tau_k\right)$ we have $\delta l_{kh}\left(t\right) = p_{\mathrm{EI}}\left(\tau_k,\tau'_h\right)\delta n\left(t-\tau'_h-\tau_k\right)\Delta\tau\Delta\tau'$, and replacing into the above equation we get

$$\begin{aligned}\delta r_{kh} &= p_{\mathrm{EI}}\left(\tau_k,\tau'_h\right)\delta n\left(t-\tau'_h-\tau_k\right)\Delta\tau\Delta\tau' \qquad (76)\\ &\quad -\gamma\sum_s p_{\mathrm{H}}\left(\tau''_s|\tau'_h,\tau_k\right)p_{\mathrm{EI}}\left(\tau_k,\tau'_h\right)\delta n\left(t-\tau'_h-\tau_k-\tau''_s\right)\Delta\tau\Delta\tau'\Delta\tau''\\ &\quad -\beta\sum_p p_{\mathrm{D}}\left(\tau'''_p|\tau'_h,\tau_k\right)p_{\mathrm{EI}}\left(\tau_k,\tau'_h\right)\delta n\left(t-\tau'_h-\tau_k-\tau'''_p\right)\Delta\tau\Delta\tau'\Delta\tau'''\end{aligned}$$

Also recalling that $p_{\mathrm{H}}\left(\tau_s''|\tau_h',\tau_k\right)p_{\mathrm{EI}}\left(\tau_k,\tau_h'\right)=p_{\mathrm{EIH}}\left(\tau_k,\tau_h',\tau_s''\right)$ and $p_{\mathrm{D}}\left(\tau_p'''|\tau_h',\tau_k\right)p_{\mathrm{EI}}\left(\tau_k,\tau_h'\right)=p_{\mathrm{EID}}\left(\tau_k,\tau_h',\tau_p'''\right)$ we obtain

$$\begin{aligned}\delta r_{kh} &= p_{\mathrm{EI}}\left(\tau_k,\tau_h'\right)\delta n\left(t-\tau_h'-\tau_k\right)\Delta\tau\Delta\tau' \\ &\quad -\gamma\sum_s p_{\mathrm{EIH}}\left(\tau_k,\tau_h',\tau_s''\right)\delta n\left(t-\tau_h'-\tau_k-\tau_s''\right)\Delta\tau\Delta\tau'\Delta\tau'' \\ &\quad -\beta\sum_p p_{\mathrm{EID}}\left(\tau_k,\tau_h',\tau_p'''\right)\delta n\left(t-\tau_h'-\tau_k-\tau_p'''\right)\Delta\tau\Delta\tau'\Delta\tau'''\end{aligned} \tag{77}$$

then by taking the sum over the indexes k and h and in the limit of vanishing $\Delta\tau$, $\Delta\tau'$, $\Delta\tau''$, $\Delta\tau'''$ one obtains

$$\begin{aligned}\dot{r} &= \iint d\tau d\tau_1 p_{\mathrm{EI}}\left(\tau,\tau_1\right)\dot{n}\left(t-\tau-\tau_1\right) \\ &\quad -\gamma\iiint d\tau d\tau_1 d\tau_2 p_{\mathrm{EIH}}\left(\tau,\tau_1,\tau_2\right)\dot{n}\left(t-\tau-\tau_1-\tau_2\right) \\ &\quad -\beta\iiint d\tau d\tau_1 d\tau_3 p_{\mathrm{EID}}\left(\tau,\tau_1,\tau_3\right)\dot{n}\left(t-\tau-\tau_1-\tau_3\right)\end{aligned} \tag{78}$$

Recalling Eq. (10) we get Eq. (11).

B Rephrasing the equations

Consider Eq. (9), which describes the evolution of the number of individuals in the exposed compartment. It can be rephrased as

$$e\left(t\right)=\int\left\{\delta\left(\tau\right)-p_{\mathrm{E}}\left(\tau\right)\right\}n\left(t-\tau\right)d\tau \tag{79}$$

and recalling that $\mathcal{F}_{\mathrm{E}}'\left(\tau\right)=\delta\left(\tau\right)-p_{\mathrm{E}}\left(\tau\right)$ we also have

$$e\left(t\right)=\int\mathcal{F}_{\mathrm{E}}'\left(\tau\right)n\left(t-\tau\right)d\tau=\int\mathcal{F}_{\mathrm{E}}\left(\tau\right)\dot{n}\left(t-\tau\right)d\tau$$

where we use integration by parts and considered that the survival function $\mathcal{F}_{\mathrm{E}}\left(\tau\right)$ is zero for $\tau<0$ and decreases exponentially as $\tau\to+\infty$. Now consider Eq (10)

$$i\left(t\right)=\int p_{\mathrm{E}}\left(\tau\right)n\left(t-\tau\right)d\tau-\iint d\tau d\tau_1 p_{\mathrm{EI}}\left(\tau,\tau_1\right)n\left(t-\tau-\tau_1\right) \tag{80}$$

By intoducing the new integration variable θ_1 such that $\theta_1 = \tau + \tau_1$, the double integral at the right hand side of the equation can be rephrased as $\iint d\tau d\theta_1 p_{\mathrm{EI}}(\tau, \theta_1 - \tau)\, n(t - \theta_1)$ and recalling that $\int d\tau p_{\mathrm{EI}}(\tau, \theta_1 - \tau) = p_1(\theta_1)$ one gets

$$i(t) = \int [p_{\mathrm{E}}(\tau) - p_1(\tau)]\, n(t - \tau)\, d\tau \tag{81}$$

Moving to the survival functions $\mathcal{F}_1(\tau)$ and $\mathcal{F}_{\mathrm{E}}(\tau)$ of the probability density functions $p_{\mathrm{E}}(\tau)$ and $p_1(\tau)$ respectively we get

$$i(t) = \int [\mathcal{F}_1'(\tau) - \mathcal{F}_{\mathrm{E}}'(\tau)]\, n(t - \tau)\, d\tau \tag{82}$$

and integrating by parts

$$i(t) = \int_0^{+\infty} [\mathcal{F}_1(\tau) - \mathcal{F}_{\mathrm{E}}(\tau)]\, \dot{n}(t - \tau)\, d\tau \tag{83}$$

Then Eq. (3) can be rephrases as

$$\dot{n} - \nu(t)\left(1 - \frac{n}{N}\right)\int_0^{+\infty} [\mathcal{F}_1(\tau) - \mathcal{F}_{\mathrm{E}}(\tau)]\, \dot{n}(t - \tau)\, d\tau = n_0 \delta(t)\,; \tag{84}$$

Using the same argument and setting $\theta_2 = \tau + \tau_1 + \tau_2$ we also have

$$h(t) = \gamma \int d\theta_2 n(t - \theta_2)\, d\tau d\tau_1 p_{\mathrm{EIH}}(\tau, \tau_1, \theta_2 - \tau - \tau_1) \tag{85}$$

and recalling that

$$p_2(\theta_2) = d\tau d\tau_1 p_{\mathrm{EIH}}(\tau, \tau_1, \theta_2 - \tau - \tau_1) \tag{86}$$

we get

$$\begin{aligned} h(t) &= \gamma \int d\tau \{\delta(\tau) - \mathcal{F}_2'(\tau)\}\, n(t - \tau) \\ &= \gamma n(t) - \gamma \int d\tau \mathcal{F}_2(\tau)\, \dot{n}(t - \tau) \end{aligned} \tag{87}$$

where $\mathcal{F}_2(\theta_2)$ is the survival function of $p_2(\theta_2)$. Analogously we have

$$\begin{aligned} d(t) &= \beta \int d\tau \left\{ \delta(\tau) - \mathcal{F}_3'(\tau) \right\} n(t-\tau) \\ &= \beta n(t) - \beta \int d\tau \mathcal{F}_3(\tau) \dot{n}(t-\tau) \end{aligned} \tag{88}$$

where $\mathcal{F}_3(\theta_3)$ is the survival function of $p_3(\theta_3)$, where $\theta_3 = \tau + \tau_1 + \tau_3$.

References

1. R. M. Anderson, R. M. May, *Infectious diseases of humans: dynamics and control* (Oxford university press, 1992).

2. H. W. Hethcote, *SIAM review* **42**, 599 (2000).

3. D. Earn, *Mathematical biology* **14**, 151 (2009).

4. O. Diekmann, J. A. P. Heesterbeek, *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*, vol. 5 (John Wiley & Sons, 2000).

5. F. Brauer, C. Castillo-Chavez, C. Castillo-Chavez, *Mathematical models in population biology and epidemiology*, vol. 2 (Springer, 2012).

6. B. Ivorra, B. Martínez-López, J. M. Sánchez-Vizcaíno, Á. M. Ramos, *Annals of operations research* **219**, 25 (2014).

7. B. Martínez-López, B. Ivorra, A. Ramos, J. M. Sánchez-Vizcaíno, *Veterinary microbiology* **147**, 300 (2011).

8. H. R. Thieme, *Mathematics in population biology*, vol. 12 (Princeton University Press, 2018).

9. W. O. Kermack, A. G. McKendrick, *Proc. R. Soc. Lond.* **115**, 700 (1927).

10. W. O. Kermack, A. G. McKendrick, *Bulletin of mathematical biology* **53**, 33 (1991).

11. E. B. Wilson, C. Bennett, M. Allen, J. Worcester, *Proceedings of the American Philosophical Society* pp. 357–476 (1939).

12. R. H. Simpson, *et al.*, *Lancet* pp. 549–54 (1952).

13. N. T. Bailey, *et al.*, *The mathematical theory of infectious diseases and its applications* (Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975).

14. K. Gough, *Biometrika* **64**, 559 (1977).

15. P. E. Sartwell, *et al.*, *American Journal of Hygiene* **51**, 310 (1950).

16. H. J. Wearing, P. Rohani, M. J. Keeling, *PLoS Med* **2**, e174 (2005).

17. A. Lloyd, *Proc. R. Soc. Lond. B* **268**, 985 (2001).

18. A. L. Lloyd, *Theoretical population biology* **60**, 59 (2001).

19. M. J. Keeling, B. T. Grenfell, *Proceedings of the Royal Society of London. Series B: Biological Sciences* **269**, 335 (2002).

20. H. T. Nguyen, P. Rohani, *Journal of The Royal Society Interface* **5**, 403 (2008).

21. A. J. Black, A. J. McKane, *Journal of the royal society interface* **7**, 1219 (2010).

22. A. J. Conlan, P. Rohani, A. L. Lloyd, M. Keeling, B. T. Grenfell, *Journal of the Royal Society Interface* **7**, 623 (2010).

23. N. T. Bailey, *Biometrika* pp. 15–22 (1956).

24. N. T. Bailey, *Biometrika* pp. 322–331 (1956).

25. H. Nishiura, M. Eichner, *Epidemiology & Infection* **135**, 1145 (2007).

26. H. Nishiura, *Emerging themes in epidemiology* **4**, 1 (2007).

27. M. Eichner, K. Dietz, *American Journal of epidemiology* **158**, 110 (2003).

28. N. T. Bailey, *Applied Statistics* pp. 9–19 (1964).

29. D. Anderson, R. Watson, *Biometrika* **67**, 191 (1980).

30. J. Ma, D. J. Earn, *Bulletin of mathematical biology* **68**, 679 (2006).

31. C. Wang, P. W. Horby, F. G. Hayden, G. F. Gao, *The lancet* **395**, 470 (2020).

32. A. E. Gorbalenya, *et al.*, *BioRxiv* (2020).

33. W. H. Organization, W. H. Organization, *et al.*, Naming the coronavirus disease (covid-19) and the virus that causes it (2020).

34. Y. Wang, Y. Wang, Y. Chen, Q. Qin, *Journal of medical virology* **92**, 568 (2020).

35. W. H. Organization, W. H. Organization, *et al.*, Statement on the second meeting of the international health regulations emergency committee regarding the outbreak of novel coronavirus (covid-19) (2020).

36. M. J. Keeling, P. Rohani, *Modeling infectious diseases in humans and animals* (Princeton university press, 2011).

37. S. Bansal, B. T. Grenfell, L. A. Meyers, *Journal of the Royal Society Interface* **4**, 879 (2007).

38. T. L. Burr, G. Chowell, *Mathematical and Computer Modelling* **48**, 122 (2008).

39. D. Anderson, R. Watson, *Biometrika* **67**, 191 (1980).

40. O. Krylova, D. J. D. Earn, *J. R. Soc. Interface* **10**, 20130098 (2013).

41. A. L. Lloyd, *Proc. R. Soc. Lond. B* **268**, 985–993 (2001).

42. L. I. Dublin, A. J. Lotka, *Journal of the American Statistical Association* **20**, 305 (1925).

43. W. L. Heffernan JM, Smith RJ, *J R Soc Interface* **2**, 281 (2005).

44. H. Nishiura, G. Chowell, *Mathematical and statistical estimation approaches in epidemiology* (Springer, 2009), pp. 103–121.

45. Q. Li, *et al.*, *New England Journal of Medicine* **382**, 1199 (2020).

46. EpiCentro, Caratteristiche dei pazienti deceduti positivi allapos;infezione da SARS-CoV-2 in Italia — epicentro.iss.it, https://www.epicentro.iss.it/coronavirus/sars-cov-2-decessi-italia. [Accessed 04-Aug-2022].

47. barbara meoni, Il rischio di decesso tra i malati di CoViD-19 x — ars.toscana.it, https://www.ars.toscana.it/2-articoli/4340-rischio-morte-tra-malati-covid-19-nuovo-coronavirus.html. [Accessed 04-Aug-2022].

48. C. Mamo, *et al.*, *Bollettino epidemiologico nazionale* (2021).

49. B. Ivorra, M. R. Ferrández, M. Vela-Pérez, A. Ramos, *Communications in nonlinear science and numerical simulation* **88**, 105303 (2020).

50. P. E. Lekone, B. F. Finkenstädt, *Biometrics* **62**, 1170 (2006).

51. M. Gatto, *et al.*, *Proceedings of the National Academy of Sciences* **117**, 10484 (2020).

52. M. Allieta, A. Allieta, D. Rossi Sebastiano, *Journal of Public Health* pp. 1–9 (2021).