ON SARS SPREAD IN 2020// POSITION NON-PAPER

SERGEI YAKOVENKO

Abstract. This is a position paper written in a completely non-academic way, not claiming any originality, nor citing any references. It will never be submitted (at least in its present form) to any journal, yet I wish to draw attention to it of as many professionals as possible. If found meaningful, it will be endlessly edited and rewritten.

1. SIR model and its drawbacks

The standard basic model describing the spread of an infectious disease is a system of nonlinear ODEs: the entire population is grouped into three (or more) cohorts, S (susceptible), I (infected) and R (recovered). The dynamics is modelled after equations traditional for chemical kinetics, where the speed of a reaction is proportional to the product of concentration of reagents. In the simplest case the "infection" equation takes the form

$$\dot{I} = kI \cdot \frac{S}{S+R} - \dot{R}, \tag{1}$$

where k is a constant characterizing the intensity and the fraction $S/(S+R)$ is the concentration of the susceptible cohort in the entire population and α is the recovery rate (we assume for simplicity that there is no significant mortality so that $S+I+R$ remains constant).

There is a number of assumptions behind (1), which may (or may not) be true for specific diseases. The primary assumption is that the cohort R is immune. The constant k is to a large extent a parameter that can be controlled. For instance, if one half of S is locked down, this is roughly equivalent to replacing k by $k/2$ with all implications. Yet the effective value of this parameter is a big unknown.

The simplest versions of (1) admit an easy study. On the initial stages of the epidemics the term $S/(S+R)$ is very close to one, thus I grows exponentially. This fits well the statistical data: the cumulative

Date: April 12, 2020.

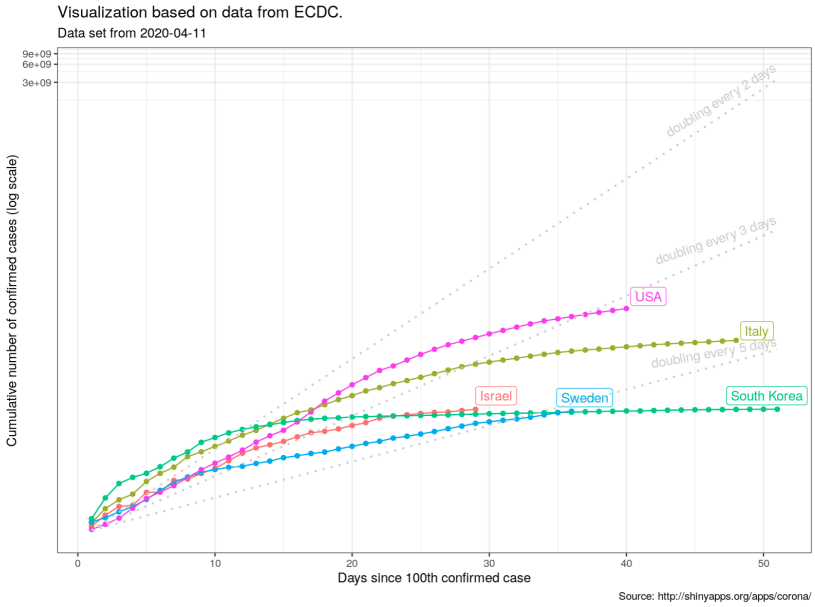

FIGURE 1. Cumulative number of SARS cases

number of infected people

$$\int_0^t I(\tau)\,\mathrm{d}\tau$$

grows exponentially at the rate somewhere between 25% and 33% (for different countries), see Fig. 1. This growth rate can be used to estimate the "unimpeded value" k for the "ignorant society" taking neither mitigation measures like lock down, nor proactive measures (blanket testing of suspicious people regardless of appearance of the symptoms).

As the concentration of R grows, the exponential growth of I slows down and there is a stable equilibrium in the phase state of the model, which corresponds to the so called *herd immunity*: once the concentration of R exceeds the corresponding limit, the virus cannot reproduce effectively (the term $S/(S+R)$ becomes small relative to the second term).

The problem with this description (at least when looking at the SARS data) is that the equilibrium corresponds to a very high level of $R/(S+R)$ of the immune population: quick estimates (depending on the estimates on k) suggest that the immune cohort R should be between 15% and 50% (and even higher). Although this is still an unknown number (one has to take into account the "invisible" parts of

I and R that passed through the cycle without exhibiting any symptoms), neither country seems to approach these values.

Example 1. In Italy today (Apr. 12) the cumulative number of infected stays at 150,000 out of the total population of 60,000,000. Even assuming that the invisible part is of the same size (a long shot from accepted estimates), we get the ration 1:200=0.5%. Yet the curve of new cases is clearly slowing down from the pure exponential growth, indicating that the saturation is not very far away. The Korean numbers (10,000 and 50,000,000) give even smaller percentage while the spread of virus is practically stopped.

This raises doubts about adequacy of the SIR model and its progeny (taking into account mortality, isolation of the sick, ...). Even a more realistic model with delay (see Appendix) cannot change the situation.

2. Very low concentration: "quantum" micromodels

It makes sense to analyze the mechanisms leading to the equation (1) on the micro-level, where everything becomes discrete rather than continuous.

Each virus carrier (whether sick or asymptomatic) spreads the virions (units of the live virus): originally contained in the body liquids, they become airborne with breath, speach, coughing, etc., and can be inhaled by a susceptible victim. The alternative route is through intermediate carriers (surfaces, buttons of elevators and cash dispensers, rails e.a.), on which the virus can survive for some time and then pass to another person on unwashed hands, food, ...

However, it would be wrong to think that any single contact or microdrop of a saliva leads to a transmission. The virologists agree (to the best of my understanding) that each viral or bacterial disease has its own "threshold of virality". In plain words, how many units of live virus (or bacteria for that sake) lead to the infection of the acceptor with probability one half. This number can be terribly small: I read somewhere, that just a single bacteria of the bubonic plague was able to do the job. Yet fortunately for most other diseases there is a non-negligible threshold, called ID50, which we will for simplicity call a **dose**. As an oversimplified scenario, for deterministic models we can assume the following dichotomy: to get infected, one has to absorb *two or more doses of virions*, while one *dose of virions* has no impact[1].

[1]Let us ignore for the moment, what happens with a person who got between one and two doses: the question will become irrelevant after a proper probabilistic modification of the model.

Unfortunately, we have no direct way to determine the size of the dose and how much is transmitted in a single contact. An indirect way is to accept the universal recommendation that "15 minutes at a distance of 2 meters or more is safe".

From the immunological point of view this "dosed approach" means that we have non-SARS-specific immune protection from invading viruses which may play the same role in the area of low concentrations as the touted "herd immunity" plays for high concentrations.

The most important question which can be answered only by the immunological research, is **whether the consecutive doses accumulate**. Quite obviously, two doses received within a short interval of time should be added to each other, guaranteeing the spread of virus within our simplified model. On the other hand, two doses received a week apart should be considered as independent events and should not lead to the infection. *What happens with two doses received overnight*? I will assume that the life cycle of a single virion of SARS (estimated at 30 min to one hour by different sources) means that 24 hours is long enough period to erase all memory of past dosage.

Remark 1. The data plotted on Fig. 1 correspond to the cumulative number of cases since the "zero day" which is placed at the moment the first 100 cases were registered. The data related to earlier stages (can also be retrieved from the same site) are much more erratic for a large number of reasons, but may give some indication of where the "low concentration threshold" could be, several hundred to single thousands of active carriers in a local region around the entry point of the first infected patient (Wuhan in China, Lombardy in Italy, New York in US).

3. Some back-of-the-envelope calculations

3.1. **How to estimate virality.** The exponential growth of 33% a day means, roughly, doubling of the infected people each three days. In other words, in an unimpeded settings every SARS carrier succeeds to infect in one day "1/3 of a victim" (with the incubation period of about one week this would correspond to $R_0 = 2.3$ or so).

Each person usually meets no more some 30-50 "contacts" a day, i.e., meeting people at a distance less than 2 m and for longer than 15 min (absent the lock-down measures). This number can be estimated from above by counting, how many people enter your 2-meter "personal sphere" simultaneously and for how long you are socially active (16 hours per day?). In fact, to estimate the virality, we need a *lower* bound which is probably around 10-20 daily "contacts".

This means that during each contact with a carrier, a susceptible *innocent victim gets about 1/30 of a dose* on average. Of course, in our model this means that in the "quantum model" nobody would get infected after a single contact!

This false conclusion is explained by the non-even distribution of the dosage between contacts, especially if there is more than one carrier you meet the same day. Here the probability of large deviations necessarily must come into play, but this is not exactly my cup of tea. However, one thing is clear,—the spread of the disease in the micro-model is controlled by the statistics of large deviations which is by far more merciful towards the victims than the additive count of probabilities.

4. An informal discussion

The difference between the macro-approach (1) and the micro-approach that I tried to sketch, can be mathematically described as follows. In the macro-approach we add infinitely small probabilities of infinitely rare events (a single act of transmission against the background of tens of millions of "particles"). In the micro-approach we add *integer parts* (doses, quants), which (put very roughly) corresponds to counting only integer parts (see the above remark about large deviations if we do, as should be done, study the honest probabilistic models).

As a result, instead of the "herd immunity" the main factor stopping the spread of the epidemics is our background non-specific immunity that is always "on duty" regardless of the specific immunity acquired after getting through the sickness bed (or skipping it). This background immunity, albeit very weak, may indeed play the key role in the area of low and very low concentration of active virus spreaders.

If we look at today's picture in Israel, then, apart from certain disaster areas which, fortunately, can be efficiently isolated and treated with all available resources, we have the residual daily growth rate of new cases at the level of 1%–2%, very close to the Korean success rate. If this exponent will be reduced by extra lock-down days to a number well below 1%, then the concentration of active carriers will become so small, that they will be unable (except for single exceptions that must be watched) to spread the virus at all.

On this optimistic note I wish you health and welcome any criticism from any side.

Appendix

There is some confusion between different parameters and indicators describing the exponential growth, especially in models with delay.

Here is an attempt to introduce some uniformity using the toy model (discrete time, difference equation).

The (absolutely) simplest model takes the form

$$X(t+T) = R_0 X(t), \qquad t = 0, T, 2T, 3T, \dots, \tag{2}$$

where T is the length of the full cycle, from infection to recovery, and $X(t)$ is the number of active carriers of virus. The value R_0 (historic notation) stands for the number of new people who are infected by one active virus carrier during the full cycle, to compare virality of different viruses. In a more convenient form where the time t is measured in days (sometimes weeks), a rescaled equation is used,

$$X(t+1) = (R_0/T)X(t), \qquad t = 0, 1, 2, 3, \dots, \tag{3}$$

although it is morally slightly incorrect (the proper constant should be $(R_0)^{1/T}$).

To relate this model to the statistical data, published daily, we have to use the sum $I(t) = \sum_{\tau=0}^{t} X(\tau)$. If X grows exponentially with $R_0 > 0$, so does (up to a constant term) the cumulative number $I(t)$.

If we want to address the delay, instead of just one variable $X(t)$ we need $n = T$ new variables $x_k(t)$, $k = 0, \dots, n$ denoting by them the number of people who are (active) carriers carrying the virus for exactly k days since their infection. The system of difference equations describing this model is as follows,

$$\begin{aligned} x_k(t+1) &= x_{k-1}(t), \qquad k = 0, 1, 2, \dots, n-1, \\ x_0(t+1) &= a_1 x_1(t) + \dots + a_n x_n(t), \\ & \qquad t = 1, 2, 3, 4, \dots \end{aligned} \tag{4}$$

Here the nonnegative coefficients $a_1, \dots, a_n \geq 0$ denote the number of new people who are infected in one day by one person carrying the virus for k days. These numbers need not to be equal, since on the initial stages the sick people can be less contagious than on the later stages.

Remark 2. The same system can be used to describe the spread of the "invisible" part of the infection, if n stands for the (average) number of days until development of clear symptoms that allow to place the virus carrier to isolation, thus plucking him/her from the virus spread business.

The parameter R_0 if literally translated into the delayed equation becomes the sum,

$$R_0 = a_1 + \dots + a_n,$$

and still has some meaning, though not as obvious as before if the profile $\{a_k\}_1^n$ is non-even.

Theorem 1 (obvious)**.** *The system* (4) *has an exponential solution of the form* $\lambda^t \cdot (x_0^*, x_1^*, \dots, x_n^*)$ *if and only if* λ *is the root of the characteristic equation*

$$\lambda^n = a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n, \qquad \lambda \in \mathbb{C}. \tag{5}$$

Example 2. If all a_k are zeros except for $a_n = R_0$, i.e., the carrier is contagious only on the last day before developing the symptoms (see Remark 2), then the equation takes the form $\lambda = a_n = R_0$ which brings us back to square one (2).

Example 3. In the opposite case where only $a_1 = R_0$ (the rest are zeros), the equation becomes $\lambda^n = R_0$ and $\lambda = R_0^{1/n}$, see above.

Remark 3. If we consider all equations (5) of degree n with nonnegative coefficients a_k normalized by the condition $\sum_1^n a_k = R_0 > 1$, then their biggest positive root $\rho = \rho(a, n)$ varies between some limits depending on n and R_0. It is not at all clear, which of these values has to be used for the parameter k in the continuous limit (1).

Remark 4. The equation (5) besides an obvious maximal real root ρ has lots of non-real roots (which differ by roots of unity in Example 3). These roots correspond to short-periodic oscillations which are the artefact of the model (unless understood otherwise).

The natural mathematical language to describe the phenomena with delay is that of integro-differential equations with deviating argument. However, such equations are infinite-dimensional and thus extra technical problems will appear. Perhaps, one can obtain meaningful conclusions passing to the limit $n \to \infty$ in the discrete system (4), but one should take special normalizing steps and I don't want to explore this direction right now.