Chapter 1

A chink in the Gates: Exponential tails of Covid-19 are modulated by daylight hours

In this first chapter, we are going to explain the significant reasons that motivated us to lead this research work, work that culminated with the submission of a conference abstract on the **4th of June** to the Thermodynamics 2.0 Conference (https://www.youtube.com/channel/UCJC98pmGjTo-xYP99MkLtrg).

Since the beginning of the COVID-19 crisis, we, the authors, were very concerned with the situation that our fellow citizens were experiencing, and not only in health terms but also with respect to the consequences for Western culture of curtailing free assembly. The economic aspects speak for themselves.

From our experience as neuroscientists we were extremely sceptical about the capacity of establishment institutional science to handle this crisis. For example we were acutely conscious of the Henry Markram debacle in Europe where the European Commission effectively wasted 1 billion of taxpayers' money on a failed Project ironically undertaken in Switzerland which is not an EU country. Conversely Sean worked at UC Berkeley with perhaps the greatest neuroscientist in history Walter Freeman III who was unfunded for almost two decades at the end of his life. Indeed a lot of the the time series analysis techniques we use are techniques that Sean learnt from Walter

Richard Strohman was another dear colleague of Sean's at UC Berkeley and together from 2005 from 2008 they taught a seminar on what Richard saw as a forthcoming revolution in biology. It is important that readers realise these great minds were marginalized by the American scientific establishment. Indeed it could be said they were sent off to die at the east end of campus in Donner lab which was owned by the Department of Energy not the university. next door to us was Professor Glaser a Nobel laureate who had been similarly marginalized for switching to neuroscience where he was investigating how noise can amplify small signals in a phenomenon called stochastic resonance.

To complete the motley crew upstairs was Peter Duesburg who was correctly pointing out that cancer is a dysfunction at the level of the chromosome not individual genes. He is of course more famous for questioning the causal role of HIV in AIDs. He pointed to particular phenomena like the fact people with HIV from blood transfusions tend to live longer and that HIV becomes proliferate only toward the end of the disease. We will allude later to the issue of HIV and whether it was used in in what may have been the engineering of covid-19. However it will not be a main concern here.

Sean had no career ambitions in working with these greats. The work was clearly interesting and important. What he didn't know is that he was unconsciously putting together the components with which we could solve the covid-19 problem.

Richard was also marginalized. In particular he correctly predicted that the Human Genome Project would be able to solve only a tiny fraction of disease etiologies. Specifically only 2% of diseases have have an explanation in terms of a single Gene. Thermodynamics exemplified as metabolism has a very important causal role. Indeed every organism can be seen as a solution to maximization of what it's called free energy as much as that organism can be seen as as a Victor in survival of the fittest. It is our hope that this book will bring attention to Richard's work now that he is no longer with us.

David is a physics graduate from the University of Barcelona who went on on to do a PhD in neuroscience in Heidelberg. We both have a strong background in statistical mechanics. We felt that we did not have a right to continue to teach at our foundations of mind summer school unless we could solve what look like a much simpler problem than the brain. We began exchanging emails about the topic daily and talking by phone every weekend about plausible ways of counterattacking the crisis.

After some talks back and forth suggesting several analyses, we started modeling the data at the end of April-beginning of May. Over a month, we took publicly available data from Ireland and Spain and fitted a plethora of mathematical models.

This chapter deals with how we initially approached the problem. The first step was to use publicly available data in which we expected to see a signature of the decay of the virus in daylight. That would exemplify itself in the decrease of the number of cases from peak. This methodology turned out to be powerful and to give unexpectedly significant results.

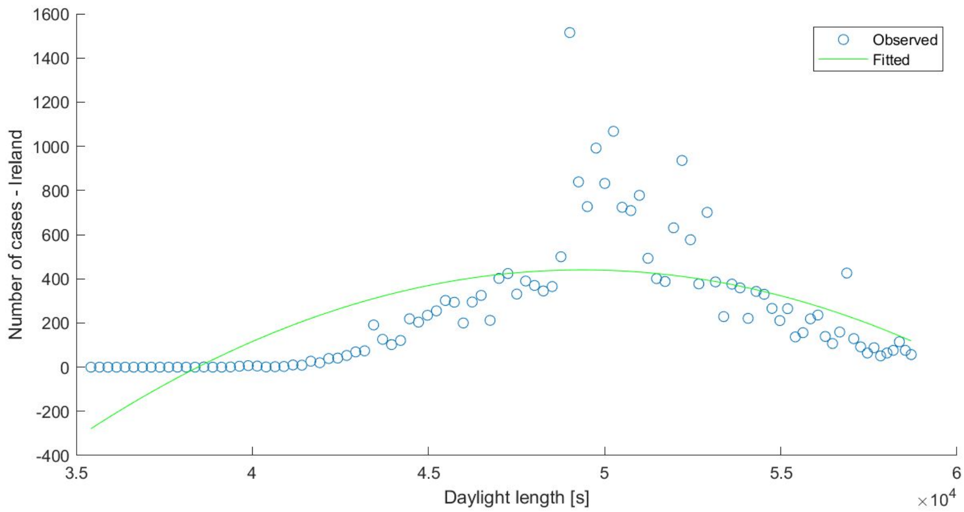
On the **27th of May**, we ran the following "non-significant" analyses that were the prelude for the forthcoming ones. It was clear that analysing from the start of the epidemic rather than from the peak gave non-significant results. We would need to to make the model more complex in order to handle the entire trajectory.

In these diagrams the y-axis has the number of cases. The x-axis has the number of seconds in units of ten thousand which corresponds to dates. Each hour has 3,600 seconds so as the date approaches 21 June in Ireland we get to days of 60,000 seconds or so.

What we lthen ooked for was a formula that would it relate y and x. This formula uses constants a and b. For non-specialist and specialist alike you can see how a and b converge over the two countries over time. Exp indicates that we found an exponent or power law and believe that one of the main problems with the establishment handling of covid-19 has been the use of conventional parametric statistics. For the specialist we have included other details like the correlation coefficient.

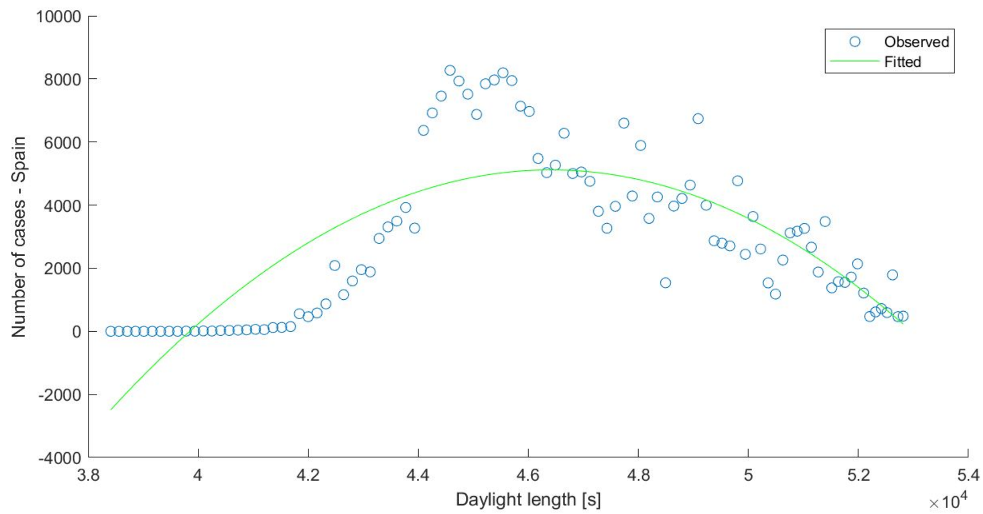
For Ireland:

Making the math model clearer by allowing negative numbers of cases;



Observed
Fitted
Number of cases - Ireland
Daylight length [s]
×10^4
1600
1400
1200
1000
800
600
400
200
0
-200
-400
3.5
4
4.5
5
5.5
6

And Spain:

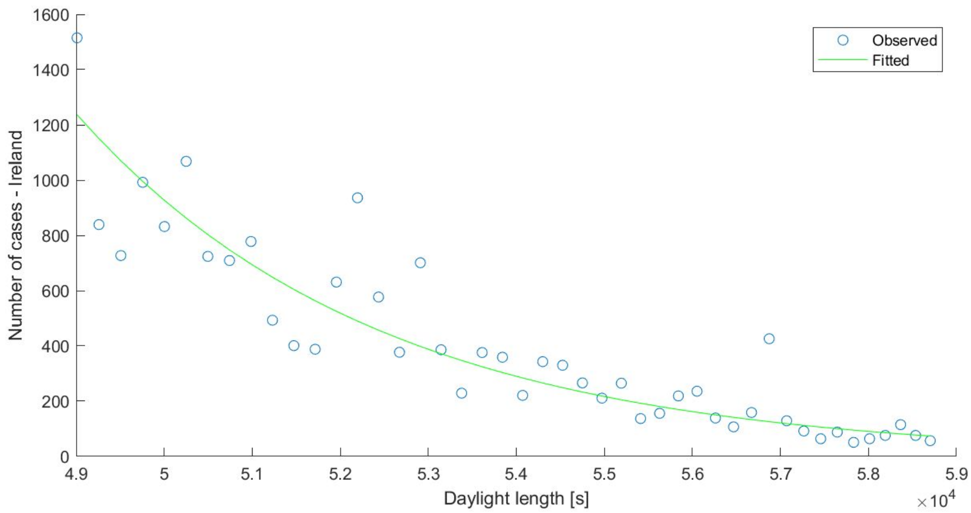


Observed
Fitted
Number of cases - Spain
Daylight length [s]
×10^4
10000
8000
6000
4000
2000
0
-2000
-4000
3.8
4
4.2
4.4
4.6
4.8
5
5.2
5.4

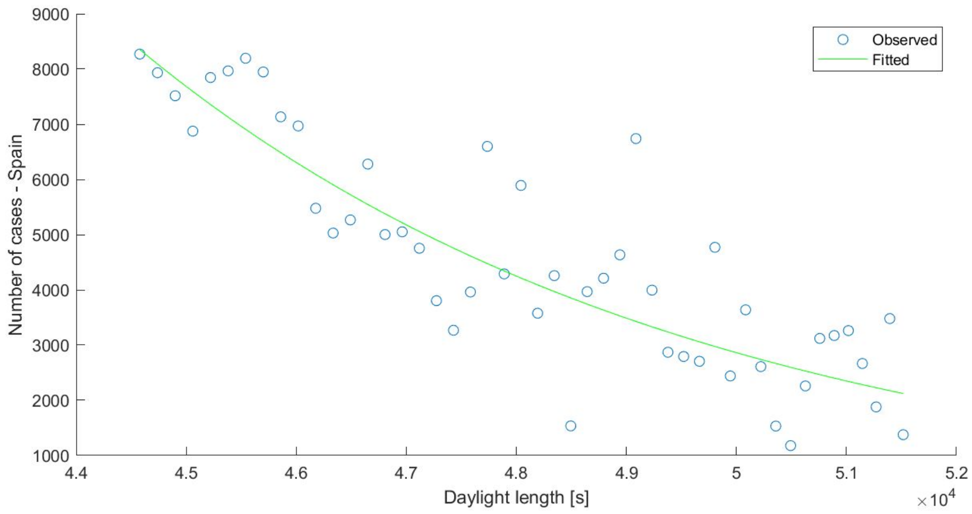
Therefore we plotted the number of cases vs. daylight hours and looked for models that could explain the data. We did not find any relevant models. However, in the next few days, we looked at the figures since every country reached the maximum peak and fitted exponential curves.

On the **31st of May**, we ran the following analyses:

For Ireland:



For Spain:



In summary, we obtained the following mathematical expressions and summary statistics for Ireland and Spain, respectively:

For Ireland:

Y = exp(a + b*X), with a = 21,3638 and b = -0,00029063.

Correlation coefficient = -0,928493,

R-square = 86,21%, and

- R-square (adjusted for degrees of freedom) = 85,8893%.

For Spain:

Y = exp(a + b*X), with a = 17,8342 and b = -0,000197494.

Correlation coefficient = -0,811314,

R-square = 65,823%, and

R-square (adjusted for d.f.) = 65,0801%.

The model itself explained about 90% of the variability in the number of cases of Ireland, a model consisting of a single variable: daylight hours.

Interestingly, this statistical significance seemed to be consistent in Ireland and Spain, countries with different climates, with the exponential coefficient converging to the same value.

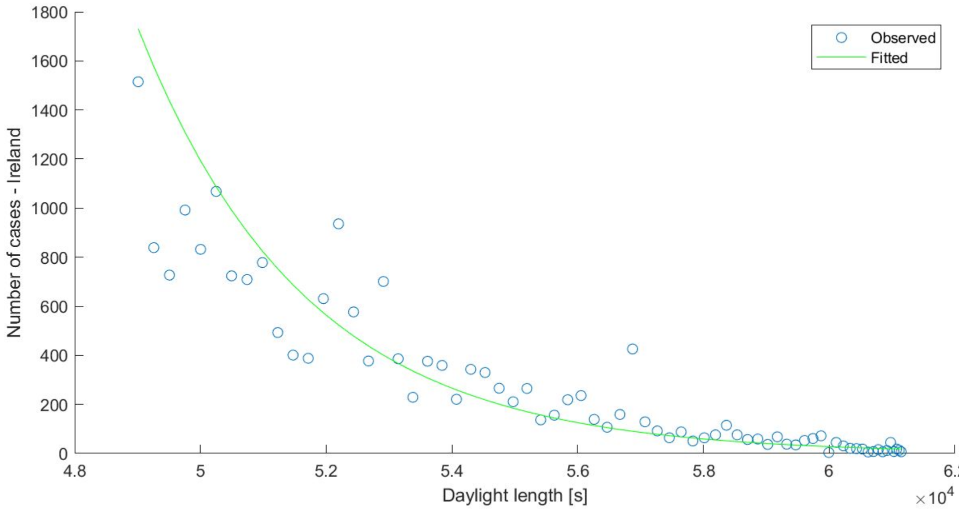
At this point, we decided to make public this research by submitting the abstract on the **4^{th} of June**. The conference would take place on the 24^{th} of June, so, in the meantime, we worked on the PowerPoint and oral presentation, and updated the analyses daily.

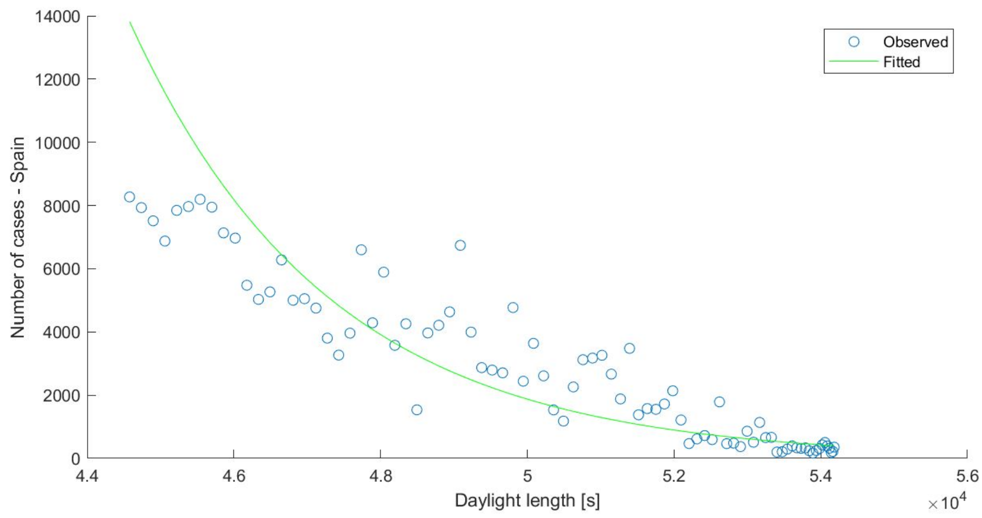
Importantly, the exponents were converging to the same value as we predicted *a priori*; we finalized this research on the **<u>19th of June</u>**, a few days before the conference.

The conclusion for us was clear: daylight hours modulates significantly the number of cases.

The updated results obtained on the **<u>19th of June</u>** were the following:

For Ireland;





14000
12000
10000
8000
6000
4000
2000
0
Number of cases - Spain
4.4
4.6
4.8
5
5.2
5.4
5.6
Daylight length [s]
$\times 10^4$
Observed
Fitted

As a result of this first part of analyses, we obtained the following mathematical expressions and summary statistics for Ireland and Spain, respectively:

For Ireland:

Y = exp(a + b*X), with a = 25,8428 and b = -0,000375156.

Correlation coefficient = -0.930879,

R-square = 86,6535 %, and

R-square (adjusted for d.f.) = 86,4543 %.

For Spain:

Y = exp(a + b*X), with a = 25,9625 and b = -0,000368539.

Correlation coefficient = -0,912269,

R-square = 83,2235%, and

R-square (adjusted for d.f.) = 83,0189 %.

In summary, the 2 free parameters: intercepts and exponents, converged.

While findings at that time were preliminary, the variability explained by daylight hours was enormous. The model itself explains about 90% of the variability in the number of cases in Ireland.

Interestingly, this statistical significance was consistent in Ireland and Spain, two European countries with different climates, with the exponential coefficient converging to the same value.

Importantly, the fitted responses were independent of the number of inhabitants of each country but showed a significant relation with daylight length.

*As said, this work finalized with the submission of the next conference abstract:

Exponential tails of Covid-19 cases are modulated by daylight hours

D. Bernal-Casas, S. Ó'Nualláin (corresponding author)

Video: https://youtu.be/qiqiNz53YyY