



*Gen. Math. Notes, Vol. 19, No. 2, December, 2013, pp. 83-87*  
*ISSN 2219-7184; Copyright © ICSRS Publication, 2013*  
*www.i-csrs.org*  
*Available free online at <http://www.geman.in>*

# On the Mathematical Interpretation of Epidemics by Kermack and McKendrick

Raúl Isea<sup>1</sup> and Karl E. Lonngren<sup>2</sup>

<sup>1</sup>Fundación IDEA, Hoyo de la Puerta, Baruta, Venezuela  
E-mail: raul.isea@gmail.com

<sup>2</sup>Department of Electrical and Computer Engineering, University of Iowa  
Iowa City, IA 52246 – USA  
E-mail: lonngren@engineering.uiowa.edu

(Received: 12-5-13 / Accepted: 31-10-13)

## Abstract

*The mathematical details that allow one to understand and individually obtain the results appearing in the classical Kermack and McKendrick 1927 paper are presented. As these results are so important for understanding an individual epidemic, the mathematical details that are glossed over in that paper are brought forth for the ease of usage by the educator and researcher in the field.*

**Keywords:** *Epidemics, Kermack–McKendrick theory, Disease Modeling.*

## 1 Introduction

It is currently impossible to predict where and when an epidemic for any disease will occur without being cognizant of the initial infection. Unfortunately, only non-conclusive predictive capabilities are available at the present time. For that reason, it is important to teach and explain the mathematical details with some clarity of one of the earliest and most important models that can be used to

interpret disease characteristics in an epidemic that are described in the classic 1927 paper by Kermack and McKendrick [1].

Other models have been developed but they do have certain limitations [2] such as: (a) The model that assumes there is genetic uniformity and all people react to the disease identically. (b) The model that assumes everyone has the same transmission rate from the disease (c) The model that assumes the same proportion of deaths regardless of the geographical location of the disease. (d) The model that does not incorporate any effects of the localized weather conditions as was indicated in a recent study correlating the number of dengue cases in Venezuela with the El Niño conditions in the Pacific Ocean [3]. (e) The model must also incorporate the effects of certain members of a population having an innate immunity [4].

## 2 Kermack and McKendrick Equations

The Kermack and McKendrick subdivide the total population  $N$  that exists in the community into three separate distinct populations: the population  $S$  that is susceptible to the disease, the population  $I$  that is infected with the disease and the population  $R$  that has recovered from the disease. These populations are modeled with three coupled first-order nonlinear ordinary differential equations. This model has been given the acronym of the SIR model and it is described with the set of three equations with the independent variable being the time  $t$ . The coefficients  $\beta$  and  $\nu$  are constants that represent the rates of infection and recovery during the epidemic respectively.

$$\frac{dS}{dt} = -\beta SI \quad (1)$$

$$\frac{dI}{dt} = \beta SI - \nu I \quad (2)$$

$$\frac{dR}{dt} = \nu I \quad (3)$$

Although an analytical solution is presented in the original paper, several steps are missing in the seminal paper that led to the solution. It is the purpose of this note to fill in the missing steps to assist the reader in the understanding of this important topic. No new ground will be plowed in this note but hopefully the missing steps that are included here make the subject more transparent to the user. Initially, we will be able to follow in the footsteps of Kermack and McKendrick before embarking on the hidden path that will lead to the solution with the missing steps brought forth for the users' benefit.

Adding the three differential equations (1)-(3) together yields a first order differential equation which can be immediately integrated to yield

$$\frac{d(S + I + R)}{dt} = 0 \Rightarrow S + I + R = N \quad (4)$$

where  $N$  is a constant of integration that represents the total population of susceptible, infected and recovered populations during a particular epidemic in the community. This value may change for different epidemics but it is assumed that the population remains constant within a particular epidemic.

Dividing (1) by (3), we obtain a first-order ordinary differential equation that can be immediately integrated to yield

$$\frac{\frac{dS}{dt} = -\beta SI}{\frac{dR}{dt} = \nu I} \Rightarrow \frac{dS}{dR} = -\frac{\beta S}{\nu} \Rightarrow S = S_0 \exp\left(-\frac{\beta}{\nu} R\right) \quad (5)$$

where the constant integration is equal to 0. Incorporating the results of (4) and (5) in (3), we find that

$$\frac{dR}{dt} = \nu I \Rightarrow \frac{dR}{dt} = \nu(N - S - R) \Rightarrow \frac{dR}{dt} = \nu\left(N - S_0 \exp\left(-\frac{\beta}{\nu} R\right) - R\right) \quad (6)$$

If the ratio of  $\frac{\beta}{\nu}$  is small, the exponential term can be approximated with the first three terms of a mathematical expansion to yield

$$\frac{dR}{dt} = \nu\left(N - S_0\left(1 - \frac{\beta}{\nu} R + \frac{\beta^2}{2\nu^2} R^2\right) - R\right) \quad (7)$$

At this stage of the original manuscript, Kermack and McKendrick fail to provide the missing steps in the derivation and they immediately jump to the final result. It is the purpose of this note to provide the bridge in this classic work.

Equation (7) is a first order differential equation that can be written in the form

$$\frac{dR}{\nu\left((N - S_0) + \left(S_0 \frac{\beta}{\nu} - 1\right)R - \left(S_0 \frac{\beta^2}{2\nu^2}\right)R^2\right)} = dt \quad (8)$$

The integral of the right-hand side of (8) is equal to  $t + C$  where  $C$  is a constant of integration that will be subsequently determined.

The left-hand side of (8) can be put in the form of an equation that is found in a table of integrals, namely

$$\int \frac{dx}{a+bx+cx^2} = \frac{2 \tan^{-1} \left( \frac{b+2cx}{\sqrt{4ac-b^2}} \right)}{\sqrt{4ac-b^2}} \quad (9)$$

For simplicity, we will examine (9) before relating it with the proper constants given in (8). Equation (9) can be written in terms of the hyperbolic function  $\tanh(x)$  using the definition that

$$\tanh(x) = -\sqrt{-1} \tan(\sqrt{-1}x) \quad \& \quad \tan(\sqrt{-1}x) = -\frac{1}{\sqrt{-1}} \tanh(x) = \sqrt{-1} \tanh(x) \quad (10)$$

Equation (8) and (9) can be rewritten with the constant of integration  $C$  included as

$$\sqrt{-1} \tanh \left( \frac{\sqrt{4ac-b^2}}{2\sqrt{-1}} (t+C) \right) = \left( \frac{b+2cx}{\sqrt{4ac-b^2}} \right) \quad (11)$$

Rearranging terms in (11), we write

$$2cx = -b + \sqrt{-1} \sqrt{4ac-b^2} \tanh \left( \frac{\sqrt{4ac-b^2}}{2\sqrt{-1}} (t+C) \right) \quad (12)$$

The constant of integration  $C$  can be determined from the location  $x$  which will be taken to be equal to 0 at the time  $t=0$ .

$$C = -\frac{2}{\sqrt{-1} \sqrt{4ac-b^2}} \tanh^{-1} \left( \frac{b}{\sqrt{-1} \sqrt{4ac-b^2}} \right) \quad (13)$$

Substituting (13) into (12) leads to

$$2cx = -b + \sqrt{-1} \sqrt{4ac-b^2} \tanh \left( \frac{\sqrt{4ac-b^2}}{2\sqrt{-1}} \left( t - \frac{2}{\sqrt{-1} \sqrt{4ac-b^2}} \tanh^{-1} \left( \frac{b}{\sqrt{-1} \sqrt{4ac-b^2}} \right) \right) \right) \quad (14)$$

Comparing (8) and (9), we find that the constants have the following values

$$a = \nu(N - S_0) \quad b = \nu \left( S_0 \frac{\beta}{\nu} - 1 \right) \quad c = -\nu S_0 \frac{\beta^2}{2\nu^2} \quad x = R \quad (15)$$

Using these constants and the following definitions

$$q = \sqrt{-\left(2(N - S_0)\left(S_0 \frac{\beta^2}{v^2}\right) + \left(S_0 \frac{\beta^2}{2v^2}\right)^2\right)} \quad \& \quad \phi = \tanh^{-1}\left(\frac{\left(S_0 \frac{\beta}{v} - 1\right)}{\sqrt{-q}}\right) \quad (16)$$

We find that (14) can be written as

$$R = \frac{v^2}{\beta^2 S_0} \left( \frac{\beta}{v} S_0 - 1 + \sqrt{-q} \tanh\left(\frac{\sqrt{-q}}{2}(vt - \phi)\right) \right) \quad (17)$$

Equations (16) and (17) are the equations that appear in the original Kermack and McKendrick paper.

## 4 Conclusion

We believe that the detailed derivation which appears here will be useful for educators and practitioners in this field. It is also possible to expand on these results, for example by assuming that the total population  $N$  possessed a cyclical nature due to external conditions not included in this work.

## Acknowledgement

RI wants to express his sincere thanks go to my advisor and mentor, Dr. Ángel Vegas for the publication of this paper

## References

- [1] W.O. Kermack and A.G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. R. Soc. Lond.*, 115(1927), 700-721.
- [2] R. Isea, Modelos matemáticos requeridos en la predicción de brotes de enfermedades infecciosas entre los humanos, *Revista Electrónica Medicina, Salud y Sociedad*, 3(2) (2013), 1-12.
- [3] R. Isea, R. Mayo-García and K.E. Lonngren, A potential correlation between dengue and the temperature of the Pacific Ocean yielding an additional warning sign, *Environ Sci: A India J.*, 7(11) (2012), 395-396.
- [4] B.J. McMorran, V.M. Marshall, C. de Graaf, K.E. Drysdale and M. Shabbar et al., Platelets kill intraerythrocytic malarial parasites and mediate survival to infection, *Science*, 323(2009), 797-800.