Mathematics of the Pandemic

M. Ram Murty and V. Kumar Murty

1 Introduction

Mathematics offers a framework for tackling the coronavirus pandemic that is now confronting the human race. We will show below how basic mathematics already gives us a viable method to control the spread of the virus. More advanced techniques can be used to make predictions and plot future trajectories. This, of course, does not preclude the essential aspect of pharmaceutical research needed to develop a vaccine. It however offers us a method of social behaviour essential to thwart the spread of the disease until a vaccine is developed.

This paper is largely expository and can be dubbed "basic epidemiology for (pure) mathematicians." Our main references are the book by Bailey [1] and Chapter 21 of [10]. Though this paper is a condensed exegesis of known results, we do however make some new remarks concerning the approximate solution by Kermack and McKendrick [16] of their "SIR" model to study epidemics. A study of this approximation and how it falls short of the reality has been discussed in several papers such as [15] and [12]. One of our main results is that a small correction term in the Kermack–McKendrick solution improves the error term in their model.

Daniel Bernoulli [5] was the first to propose a mathematical model to study the spread of smallpox back in 1760. But the idea of using mathematics to study

Research of both authors is partially supported by NSERC Discovery grants.

M. Ram Murty (✉)
Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada
e-mail: murty@queensu.ca

V. Kumar Murty
Fields Institute for Research in Mathematical Sciences, Toronto, ON, Canada
e-mail: director@fields.utoronto.ca; murty@math.toronto.edu

V. K. Murty, J. Wu (eds.), *Mathematics of Public Health*, Fields Institute Communications 85, https://doi.org/10.1007/978-3-030-85053-1_16

the spread of diseases did not quite catch on (pardon the pun). The simplest mathematical model to study the transmission of disease is called the SI model, where there is no recovery. It was proposed by Hamer [13] in 1906 and then expanded by Ross [21] in 1915. Later, Kermack and McKendrick [16] in 1927 proposed the celebrated SIR model which accomodates recovery from the infection. (Incidentally, Ross's work on malaria won him the 1902 Nobel Prize and a knighthood in 1911.) By the middle of the twentieth century, the mathematics of epidemics evolved into a serious subject and the book by Bailey [1] is considered a primary source.

In this paper, we will examine the SIR model. After surveying the paper by Kermack and McKendrick, we examine the question of error in the approximate solution in [16]. Then, we discuss the topic of "exact solutions" mainly from a combinatorial perspective via an application of the Lagrange inversion theorem. This leads to the introduction of the important role played by the Lambert W-function. We derive at the end an assortment of results that are also of number theoretic interest. In particular, we give two new integral expressions for the Lambert function for positive real arguments.

In the simplest model, we have a *branching process*. This means that each infected individual meets k other people and infects every individual he meets with probability p. Thus, the number of new cases generated by a single individual is $R_0 = pk$, which is called *basic reproductive number*. The infected people go on to infect other people in the same way and we can model this using graph theory by a *tree* structure. It is then clear that if we let q_n be the number of infected people at the n-th level of this tree, then $q_n = (pk)^n$. Since we want q_n to go to zero, we derive our first theorem.

Theorem 1 *The number of infected people will eventually tend to zero if* $pk < 1$.

This simple result already gives us a powerful method to deal with the pandemic. First, we need to know p. That is obtained by collecting data, which is very important and this means extensive testing of the population. The probability p is approximated by the ratio of the number of infected people by the total number of people tested. Of course, a total lockdown of the entire population will substantially diminish the spread of the disease. However, this is not practical in the long run and Theorem 1 tells us that we can relax this and curtail the number of people an individual can come into contact with. That number is $[1/p]$, where $[\cdot]$ is the greatest integer function. We record this as:

Corollary 2 *The number of people each individual can come into contact with should be less than* $[1/p]$ *in order to contain the pandemic.*

Our simple model, though expedient in giving us a quick understanding of the pandemic and its spread, is not an accurate reflection of the real state of things. For instance, there is no single value of p that can be ascribed to an entire nation since often, the value of p changes from region to region. Therefore, it is necessary to view p as a function of two co-ordinates x, y giving the geographical position. The same can be said of the quantity k. Thus, if $\mathcal{R}$ is the region we are interested in, the number of people infected at the n-th level in the region $\mathcal{R}$ is

$$\iint_{\mathcal{R}} p(x, y)^n k(x, y)^n dxdy.$$

Suppose that

$$f(x, y) = \lim_{n\to\infty} p(x, y)^n k(x, y)^n$$

exists almost everywhere. This is a reasonable assumption since our functions are locally constant functions. As the sequence of functions

$$f_n(x, y) = p(x, y)^n k(x, y)^n$$

is bounded by the total population of the region, an application of Lebesgue's dominated convergence theorem (see page 26 of [22]) shows that if we want

$$\lim_{n\to\infty} \iint_{\mathcal{R}} p(x, y)^n k(x, y)^n dxdy = 0,$$

then $f(x, y) = 0$ almost everywhere. In other words, we must have $p(x, y)k(x, y) < 1$ almost everywhere. We thus arrive at our second theorem.

Theorem 3 *The pandemic is contained if* $p(x, y)k(x, y) < 1$ *almost everywhere in every part of the region.*

In order to formulate a practical public policy, it is then prudent to consider these localized probabilities and functions.

There are several models that describe the spread of epidemics. Broadly, they can be grouped into two categories: discrete and continuous with respect to the time parameter. Our theorems above deal with the discrete time model. Much of this paper is devoted to the continuous time model which uses the theory of differential equations. By contrast, the discrete model uses graph theory. In this model, each person of the population is represented by a vertex. A vertex can be in one of two states: susceptible or infected. A directed edge from node i to node j means that i can infect j. The rate of infection β is attached to each edge and the rate of recovery γ is attached to each infected node. This results in a graph G and one defines the epidemic threshold of the graph G as the value τ such that if $\beta/\gamma < \tau$ then the epidemic dies over time and if $\beta/\gamma > \tau$, the epidemic spreads over time. In [25], the authors show that if λ_1 is the largest eigenvalue of the adjacency matrix of G, then the epidemc threshold of G is $1/\lambda_1$. This gives an interesting connection between spectral graph theory and the study of epidemics. We refer the reader to [9, 25] and the survey [6] for further details.

2 The SI Model

The SI model to study epidemics is discussed in Chapter 5 of Bailey [1]. We give a brief exposition. At time $t = 0$, we consider a population of size $N + a$, with N "susceptible" and a persons "infected". Let $S(t)$ be the number of susceptible people and $I(t)$ the number of infected people at time t respectively. We assume that the infection rate is β and that

$$\frac{dS(t)}{dt} = -\beta S(t)I(t).$$

As we assume the population is constant throughout the epidemic, we have $S(t) + I(t) = N + a$ so that our differential equation reduces to

$$\frac{dS(t)}{dt} = -\beta S(t)(N + a - S(t)),$$

which is easily solved using basic calculus. We have

$$S(t) = \frac{N(N + a)}{N + ae^{(N+a)\beta t}}.$$

Thus, as $S(t) + I(t) = N + a$, we have

$$\frac{dI(t)}{dt} = -\frac{dS(t)}{dt} = \frac{aN(N + a)^2 e^{(N+a)\beta t}}{(N + ae^{(N+a)\beta t})^2},$$

often called the *epidemic curve* since it gives the rate at which infections occur. It is not difficult to see that this attains its maximum at time

$$t = \frac{\log(N + a)}{\beta(N + a)}.$$

If β is very large, the peak will be reached very early in the time period. Though this model is simple, it leads to some interesting probability theory for which we refer the reader to Chapter 5 of [1].

3 The SIR Model

To understand the spread of a virus in a community, Kermack and McKendrick formulated in 1927 the so-called SIR model. (See in particular equation (29) of [16].) This is one of the basic models that mathematical biology uses to study epidemics. Other models are variations on this theme. There are three quantities that need to be studied as a function of time. The first is the number of "susceptible" people who can contract the disease. This is denoted $S(t)$ as a function of the time parameter t. The next is the number of infected people at time t, denoted $I(t)$. The third is the number of "recovered" people $R(t)$. This explains the acronym 'SIR.'

These quantities are inter-related. We make the following assumptions.

$$S(t) + I(t) + R(t) = N, \tag{1}$$

where N is the total population of the region under consideration. In other words, we assume our population of the region is constant and does not increase with time t. We assume that a proportion γ of the infected people will recover. (Sadly, the word 'recover' may also euphemistically include deaths.) Thus,

$$\frac{dR(t)}{dt} = \gamma\, I(t). \tag{2}$$

The number of new infections is proportional to the number of interactions between susceptible people and the infected people and so, taking into account the recovery rate, we see that the differential equation encoding this fact is

$$\frac{dI(t)}{dt} = \beta\, S(t)I(t) - \gamma\, I(t), \tag{3}$$

where β is the infection rate. Finally, the number of susceptible people satisfies

$$\frac{dS(t)}{dt} = -\beta S(t)I(t). \tag{4}$$

These equations together with

$$\frac{dS(t)}{dt} + \frac{dI(t)}{dt} + \frac{dR(t)}{dt} = 0,$$

obtained by differentiating (1) give us a system of three differential equations govering the spread of the pandemic. This can also be deduced simply by adding up (4), (3) and (2). In essence, we actually have two differential equations. Let us record this in the following system describing the SIR model:

$$\begin{cases} \dfrac{dS(t)}{dt} = & -\beta S(t)I(t) \\[2ex] \dfrac{dI(t)}{dt} = & \beta\, S(t)I(t) - \gamma\, I(t) \\[2ex] \dfrac{dR(t)}{dt} = & \gamma\, I(t). \end{cases} \tag{5}$$

Fortunately, this system can be simplified as follows. Dividing (4) by (2) and using the chain rule, we get

$$\frac{dS}{dR} = -\frac{\beta}{\gamma} S,$$

so that

$$-\frac{\beta}{\gamma} R(t) = \log(S(t)/S(0))$$

since at time $t = 0$, we have $R(0) = 0$. In other words, $S(t) = S(0)\exp(-\frac{\beta}{\gamma} R(t))$. Inserting this back into (2) we obtain

$$\frac{dR}{dt} = \gamma \left(N - R(t) - S(0)\exp(-\frac{\beta}{\gamma} R(t)) \right),$$

which is the basic differential equation governing the behaviour of the epidemic.

The ratio β/γ is often denoted R_0 in the literature[1] and is called the *basic reproduction number* or *basic reproduction ratio* and it is not unrelated to the concept introduced in the earlier section. The reciprocal $\rho = 1/R_0$ is called the *removal rate* by Bailey [1]. The rate at which infections are increasing is given by (3) and we can re-write it as

$$\frac{dI(t)}{dt} = \gamma[R_0 S(t) - 1]I(t),$$

from which we note that as $S(t) \le S(0)$, the function $I'(t)$ is decreasing if $R_0 S(0) < 1$. We therefore see that the epidemic can start only when $S(0) > 1/R_0 = \rho$. Thus, the reproduction ratio measures the spread of the epidemic.

Using the notation of R_0, our differential equation for R now becomes

$$\frac{dR}{dt} = \gamma \left(N - R - S(0)e^{-R_0 R} \right). \tag{6}$$

This differential equation has no closed form solution for R and one approximates the exponential function by its Taylor series. Thus, the equation studied becomes

$$\frac{dR}{dt} = \gamma \left(N - S(0) + (R_0 S(0) - 1)R - (R_0^2 S(0)/2)R^2 \right). \tag{7}$$

This is a special case of the generalized Riccati differrential equation.

[1] This is unfortunate notation since R_0 has nothing to do with the R function. Bailey [1] denotes this as $1/\rho$ and this is a better notation which we also use sporadically, whenever convenient.

We could have equivalently divided (4) by (3) and obtained a relation between $I(t)$ and $S(t)$ but this arrangement is simpler and follows [16]. For the sake of clarity, we rewrite (7) as

$$\frac{dR}{dt} = \gamma(A + BR + CR^2), \quad \text{with} \quad \begin{cases} A = N - S(0), \\ B = R_0 S(0) - 1, \\ C = -R_0^2 S(0)/2. \end{cases} \tag{8}$$

Let us observe that the discriminant of this quadratic is

$$(R_0 S(0) - 1)^2 + 2(N - S(0))R_0^2 S(0) > 0. \tag{9}$$

Now we insert the following lemma from first year calculus (which we leave as an exercise to the reader):

Lemma 4 *Let* $\Delta = B^2 - 4AC > 0$. *Then*

$$\int \frac{dR}{A + BR + CR^2} = -\frac{2}{\sqrt{\Delta}} \tanh^{-1}\left(\frac{2CR + B}{\sqrt{\Delta}}\right).$$

Thus, (8) is solved by

$$\frac{2}{\sqrt{\Delta}} \tanh^{-1}\left(\frac{2CR + B}{\sqrt{\Delta}}\right) = -\gamma t + \phi \tag{10}$$

where Δ is given by (9), and

$$\phi = \frac{2}{\sqrt{\Delta}} \tanh^{-1}\left(\frac{B}{\sqrt{\Delta}}\right),$$

is the integration constant determined by $R(0) = 0$. Hence, we deduce that

$$R(t) = \frac{-B + \sqrt{\Delta}\tanh\left(-\frac{\sqrt{\Delta}}{2}(\gamma t - \phi)\right)}{2C}.$$

As tanh is an odd function, we can simplify this to

$$R(t) = \frac{R_0 S(0) - 1 + \sqrt{\Delta}\tanh\left(\frac{\sqrt{\Delta}}{2}(\gamma t - \phi)\right)}{R_0^2 S(0)},$$

which agrees with formula (30) in the Kermack–McKendrick paper [16].

We record the solution of this in the following theorem and interpret the results.

Theorem 5 *The solution of (8) is given by*

$$R(t) = \frac{R_0 S(0) - 1 + \sqrt{\Delta}\tanh\left(\frac{\sqrt{\Delta}}{2}(\gamma t - \phi)\right)}{R_0^2 S(0)},$$

where

$$\phi = \frac{2}{\sqrt{\Delta}} \tanh^{-1}\left(\frac{R_0 S(0) - 1}{\sqrt{\Delta}}\right),$$

and

$$\Delta = (R_0 S(0) - 1)^2 + 2(N - S(0)) R_0^2 N^2 S(0) > 0.$$

Since we are interested in the infection rate which is the derivative of $R(t)$, we find easily: (there is a typo in [16], $\sqrt{-q}$ in formula (31) should be $-q$. This seems to have been corrected in the book by Bailey [1] on page 83.)

Theorem 6

$$I(t) = \frac{\Delta}{2R_0^2 S(0)} sech^{\,2}\left(\frac{\sqrt{\Delta}}{2}(\gamma t - \phi)\right).$$

The function appearing in the above theorem is often called the *epidemic curve*. It is a symmetrical bell-shaped curve that reflects what is often seen in epidemics, namely that new cases continue to rise until they hit a certain peak point and then slowly reduce in number. Let us note that this model does not take into account variable rates of infection or removal. One can, for instance, drastically reduce the rate of infection by following the suggested guidelines of personal hygiene, mask wearing and social distancing with or without a vaccine.

To determine the maximum number of infected people, we can differentiate this function representing $I(t)$ and note that it attains its maximum value of

$$\frac{\Delta}{2R_0^2 S(0)} \qquad \text{at} \qquad t_0 = \frac{\phi}{\gamma}.$$

We also see from Theorem 5 that

$$R(\infty) = \frac{\rho^2}{S(0)}\left(S(0) R_0 - 1 + \sqrt{\Delta}\right).$$

At the beginning of the outbreak, we expect N and $S(0)$ to be very close, and so we can approximate $\sqrt{\Delta}$ by $S(0)R_0 - 1$. Thus,

$$R(\infty) \approx 2\rho\left(1 - \frac{\rho}{S(0)}\right).$$

How good is this approximation? Using Taylor's theorem with error term, it is possible to write down an approximation of the error incurred by using only the quadratic term. We can also analyse if the approximation has any serious repercussions with respect to making predictions. This we do in the next section.

4 Error Terms

We will now analyze the error terms incurred in the use of the quadratic approximation of the exponential function in the previous section. This will also give us insight into an understanding of using higher polynomial approximations and the errors that arise in that context.

Let us write

$$e^{-R_0 R} = 1 - R_0 R + \frac{1}{2}R_0^2 R^2 + E$$

so that in fact, the difference

$$\int_0^t \frac{dR}{(N - R - S(0)e^{-R_0 R})} - \int_0^t \frac{dR}{(N - S(0) + (R_0 S(0) - 1)R - \frac{1}{2}R_0^2 S(0) R^2)}$$

is equal to

$$\int_0^t \frac{S(0)E\,dR}{(N - R - S(0)e^{-R_0 R})(N - S(0) + (R_0 S(0) - 1)R - \frac{1}{2}R_0^2 S(0) R^2)}. \tag{11}$$

By Taylor's theorem with approximation (see for example, Theorem 1.18 on pages 36–37 of [17]), we have

$$|E| = R_0^3 e^{-R_0 \xi} \frac{R^3}{3!},$$

for some $\xi \in [0, t]$. As the function $e^{-R_0 r}$ is decreasing, we have

$$|E| \leq R_0^3 \frac{R^3}{3!}.$$

Thus, the error (11) is bounded by $O(\log t)$ where the implied constant depends on N and $S(0)$.

If on the other hand, we use a cubic polynomial approximation to $e^{-R_0 R}$, then the resulting error integral is of the form

$$\int_0^t \frac{E_3(R)dR}{f(R)}$$

where $E_3(R) = O(R^4)$ and $f(R)$ is a polynomial in R of degree 6. Thus, the integral is convergent and we can write it as

$$\int_0^\infty \frac{E_3(R)dR}{f(R)} - \int_t^\infty \frac{E_3(R)dR}{f(R)}$$

which is of the form

$$\text{constant} \quad + O\left(\frac{1}{t}\right).$$

One can also use a quartic approximation to $e^{-R_0 R}$ and this leads to a similar conclusion with the error being of the form

$$\text{constant} \quad + O\left(\frac{1}{t^2}\right).$$

In the last two cases, therefore, the error is "negligible" in the sense that as t goes to infinity, there is no significant increase in the term we would obtain if we use either a cubic or quartic approximation to the exponential function.

Since

$$\int_0^t \frac{dR}{A + BR + CR^2} - \int_0^t \frac{dR}{A + BR + CR^2 + DR^3} =$$
$$\int_0^\infty \frac{DR^3 dR}{(A + BR + CR^2)(A + BR + CR^2 + DR^3)} + O\left(\frac{1}{t}\right),$$

the integral on the right hand side is a convergent integral. In other words, (10) changes to

$$\frac{2}{\sqrt{\Delta}} \tanh^{-1}\left(\frac{2CR + B}{\sqrt{\Delta}}\right) = -\gamma t + \phi + \phi' + O\left(\frac{1}{t}\right),$$

where ϕ' is a suitable constant alluded to above. We state this formally as:

Theorem 7 *The exact solution for $R(t)$ in the SIR model (5) satisfies*

$$R(t) = \frac{R_0 S(0) - 1 + \sqrt{\Delta}\tanh\left(\frac{\sqrt{\Delta}}{2}(\gamma t - \phi - \phi' + O\left(\frac{1}{t}\right))\right)}{R_0^2 S(0)},$$

where

$$\phi = \frac{2}{\sqrt{\Delta}}\tanh^{-1}\left(\frac{R_0 S(0) - 1}{\sqrt{\Delta}}\right),$$

and

$$\Delta = (R_0 S(0) - 1)^2 + 2(N - S(0))R_0^2 N^2 S(0) > 0.$$

5 Kendall's Exact Solution

The implication of the deliberations of the previous section to the result of Kermack and McKendrick (sometimes called K and K in the literature) is that one need only adjust the constant ϕ to correct the error incurred in using the quadratic approximation to $e^{-R_0 R}$. We underline this observation in light of a paper written by Kendall [15] in 1956, where he shows that the use of the quadratic approximation "consistently underestimates the infection rate" (see page 151 of [15]). Referring to the approximation we have recorded in Theorem 6 above, he adds that "it is curious that the K and K approximation should have been accepted without comment for nearly 30 years; the exact solution is easily obtained and the difference between the two can be of practical significance."

Kendall uses a calculus of variations argument to show that the K and K approximation underestimates the rate of infection β. For the sake of clarity, we outline his argument following [15] and [1]. Suppose that β is a function of R. Then, proceeding as we did before we can solve for $S(t)$:

$$S(t) = S(0)\exp\left(-\frac{1}{\gamma}\int_0^R \beta(r)dr\right),$$

which leads to as earlier

$$\frac{dR(t)}{dt} = \gamma\left(N - R - S(0)\exp\left(-\frac{1}{\gamma}\int_0^R \beta(r)dr\right)\right).$$

The function

$$\beta(R) = \frac{2\beta}{(1 - R_0R) + (1 - R_0R)^{-1}}$$

leads to the K and K approximation. Thus, $\beta(0) = \beta$ and $\beta(R) < \beta$ for $0 < R < 1/R_0$ and this is the basis of Kendall's statement that the K and K model underestimates the infection rate β. He adds that for $R > 1/R_0$, the model gives a negative infection rate. Because of these objections, he suggests a renormalization using the exact solution.

The "exact" solution alluded to is simply the transcendental function obtained by integrating (6). It is surprising that it was (perhaps independently) re-discovered by the authors in [14] in 2014. Neither Kendall nor the classic text book by Bailey which discusses Kendall's work are mentioned in the references of [14].

Kendall's approach can be described as follows. Using basic calculus, it is easily seen

$$N - R - S(0)e^{-R_0R} = 0$$

has exactly two real roots, one negative and one positive, denoted $-\eta_1$ and η_2 (using the notation of page 85 of [1]). Thus,

$$\gamma t = \int_0^R \frac{dr}{N - r - S(0)e^{-R_0r}}, \qquad 0 < R < \eta_2. \tag{12}$$

The integral diverges for $R \to \eta_2$ and so, $R(\infty) = \eta_2$. But the integral also diverges if $S(0) = N$ which suggests that there is an infinite amount of time before the epidemic starts, which is absurd. This absurdity is resolved by changing the origin to the point where $S = 1/R_0$ which is referred to as the center of the epidemic. The peak of the epidemic occurs when $I(t)$ reaches its maximum, which is when $I'(t) = 0$. Since

$$\gamma I'(t) = \frac{d}{dt}\left(\frac{dR}{dt}\right) = \gamma^2 I(t)\left(R_0S(t) - 1\right),$$

we see that the peak of the epidemic curve occurs when $S(t) = 1/R_0$ which is the center of the epidemic defined above. From (5), the maximum number of infections also occurs at the same point. This leads to the parametric solution described in the following 1956 theorem due to Kendall [15].

Theorem 8 *The parametric solution of the SIR model (5) is:*

$$t = \frac{1}{\gamma}\int_0^R \frac{dr}{I(0) - r + R_0^{-1}(1 - e^{-rR_0})},$$

$$\frac{dR}{dt} = \gamma(I(0) - r + R_0^{-1}(1 - e^{-R_0 r})),$$

where $-\infty < t < \infty$ *and* $-\zeta_1 < r < \zeta_2$*, with* $-\zeta_1$ *and* ζ_2 *being the unique negative and positive roots of*

$$I(0) - \zeta + R_0^{-1}(1 - e^{-R_0\zeta}) = 0. \tag{13}$$

The meaning of the roots $-\zeta_1$ and ζ_2 is that $\zeta_1 + \zeta_2$ is the total number of recovered people during the entire pandemic. In fact, ζ_1 is the number before the peak and ζ_2 is the number after the peak. It is therefore of some interest to determine these roots with some accuracy. This we do in the next section using the Lagrange inversion formula and the Lambert W-function. We study roots of exponential equations such as (13) from a general perspective.

6 An Application of the Lagrange Inversion Formula

If we want to determine the extremal points of $R(t)$, we need only look at when the right hand side of Eq. (6) vanishes. In other words, we want to solve for R in the equation

$$N - R - S(0)e^{-R_0 R} = 0.$$

Lagrange's inversion formula states that if we have a functional relation $f(r) = t$ with f being analytic at 0 and $f'(0) \neq 0$, then we can invert and write r as a power series in t. The precise formula is

$$r(t) = r(0) + \sum_{n=1}^{\infty} c_n t^n$$

where

$$c_n = \frac{1}{n!} D^{n-1}\left(\phi(t)^n\right)\Big|_{t=0}, \tag{14}$$

where $D = d/dt$ and $\phi(t) = 1/r'(t)$.

We give a simple proof of the Lagrange inversion formula using the Cauchy residue theorem, and then apply it to derive the needed facts about the Lambert W-function.

Given a formal power series

$$f(t) = \sum_{n=-\infty}^{\infty} c_n t^n,$$

we use the notation

$$[t^n](f(t)) := c_n.$$

Thus, for instance $[t^n](f(t)) = [t^{n+m}](t^m f(t)]$, a fact we will use later in the proof of the proposition below.

Proposition 9 *(Lagrange Inversion Formula) Suppose $f(z)$ is analytic in a neighborhood of $z = 0$ with $f(0) = 0$ and $f'(0) \neq 0$. Then f^{-1} is analytic in a neighborhood of $z = 0$ and*

$$[z^n](f^{-1}(z)) = [z^{n-1}]\left(\frac{z^n}{nf(z)^n}\right).$$

Proof Since $f'(0) \neq 0$, we have by the inverse function theorem (see page 62 of [17]) that $f^{-1}(z)$ is well-defined and analytic in a neighborhood of $f(0) = 0$. Consequently, it has a power series expansion in a neighborhood of zero. By the Cauchy residue theorem,

$$[z^n](f^{-1}(z)) = \frac{1}{2\pi i}\int_C \frac{f^{-1}(w)dw}{w^{n+1}},$$

with C being a sufficiently small circle centered at zero. We change variables in the integral by setting $w = f(v)$ which is a conformal map if C is of sufficiently small radius. Since $f^{-1}(f(v)) = v$, we have

$$[z^n](f^{-1}(z)) = \frac{1}{2\pi i}\int_{C'} \frac{vf'(v)dv}{f(v)^{n+1}},$$

where C' is the closed contour image of C under our conformal mapping. Our integral can be re-written as

$$-\frac{1}{2\pi in}\int_{C'} vd\left(\frac{1}{f(v)^n}\right) = \frac{1}{2\pi in}\int_{C'}\left(\frac{1}{f(v)^n}\right)dv,$$

on integrating by parts and noting that the residue of $d(v/f(v)^n)$ at $v = 0$ is zero. Thus, by the Cauchy residue theorem, we conclude that

$$[z^n](f^{-1}(z)) = [z^{-1}]\left(\frac{1}{nf(z)^n}\right) = [z^{n-1}]\left(\frac{z^n}{nf(z)^n}\right),$$

the last equality being clear by shifting the power series appropriately. □

Remark 10 This elegant proof is due to Whittaker and Watson [26]. The analyticity assumption in the proposition is a red herring and can be dispensed with using the theory of formal power series. A proof centered on these ideas can be found in Chapter 5 of [24]. It is straightforward to see that Proposition 9 gives (14).

It is not difficult to see that the above proof can be suitably modified to yield the following more general result.

Proposition 11 *Suppose $f(z)$ is analytic in a neighborhood of $z = 0$ with $f(0) = 0$ and $f'(0) \neq 0$. Then f^{-1} is analytic in a neighborhood of $z = 0$ and for each $1 \leq k \leq n$, we have*

$$[z^n](f^{-1}(z))^k = [z^{n-k}]\left(\frac{kz^n}{nf(z)^n}\right).$$

Another version of the same theorem is often useful in applications. To this end, we introduce the *Bell polynomials* which are defined as follows. Suppose that

$$f(z) = \sum_{n=1}^{\infty} c_n \frac{z^n}{n!}$$

Then

$$\exp(f(z)) = \sum_{n=0}^{\infty} B_n(c_1, c_2, ..., c_n)\frac{z^n}{n!}. \tag{15}$$

Thus, the so-called n-th *complete Bell polynomial* is given by

$$B_n(c_1, ..., c_n) = \frac{\partial^n}{\partial z^n} \exp\left(\sum_{n=1}^{\infty} c_n \frac{z^n}{n!}\right)\bigg|_{z=0}.$$

The Lagrange inversion formula can also be stated in terms of the Bell polynomials (see page 151 of [7]).

One more variation of the Lagrange inversion formula is given in [19]. To present it, we let

$$a_1 z + a_2 z^2 + \cdots = w$$

be a power series which converges in a neighborhood of $z = 0$. If $a_1 \neq 0$, this gives a conformal mapping of a sufficiently small disc centered at $z = 0$ onto a neighborhood of $w = 0$. Thus, we can write

$$z = b_1 w + b_2 w^2 + \cdots .$$

If we let

$$\phi(z) = \left(a_1 + a_2 z + a_3 z^2 + \cdots\right)^{-1}$$

then $w = z/\phi(z)$.

Proposition 12 *Let f be any analytic function. With ϕ and w as above, we have*

$$\frac{f(z)}{1 - w\phi'(z)} = \sum_{n=0}^{\infty} \frac{w^n}{n!} \frac{d^n}{dx^n} \left(f(x)\phi(x)^n\right)\Big|_{x=0}.$$

Proof For C a circle enclosing $\zeta = 0$ and oriented counterclockwise, we have

$$\frac{1}{n!} \frac{d^n}{dx^n} \left(f(x)\phi(x)^n\right)\Big|_{x=0} = \frac{1}{2\pi i} \int_C \frac{f(\zeta)\phi(\zeta)^n d\zeta}{\zeta^{n+1}},$$

by Cauchy's formula. Thus, for $|w\phi(\zeta)/\zeta| < 1$, we have

$$\sum_{n=0}^{\infty} \frac{w^n}{n!} \frac{d^n}{dx^n} \left(f(x)\phi(x)^n\right)\Big|_{x=0} = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta} \left(\sum_{n=0}^{\infty} (w\phi(\zeta)/\zeta)^n\right) d\zeta$$

which is

$$\frac{1}{2\pi i} \int_C \frac{f(\zeta) d\zeta}{\zeta - w\phi(\zeta)}.$$

Since $|w\phi(\zeta)| < |\zeta|$, we deduce from Rouché's theorem that ζ and $\zeta - w\phi(\zeta)$ have the same number of zeros in C which means the zero is unique, simple and equal to z. Thus the integral is

$$\frac{f(z)}{1 - w\phi'(z)}.$$

This completes the proof. □

We apply the Lagrange inversion theorem to study the inverse function of the map $w \mapsto we^w$. This map sends zero to zero and satisfies the hypothesis of Proposition 9. Its inverse function, called the Lambert W-function is sometimes denoted as W_0 and we can compute its power series easily using the Lagrange inversion formula. We find

$$W_0(z) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} z^n. \tag{16}$$

Using Stirling's formula, it is easy to see that this power series converges absolutely for $|z| < 1/e$.

It seems that the series (16) was independently re-discovered by Ramanujan where Question 738 on page 332 of [20] asks one to show that for $0 < x < 1$,

$$x = \sum_{n=1}^{\infty} \frac{n^{n-1}}{n!} x^n e^{-nx}. \tag{17}$$

It is not clear what proof Ramanujan may have had, but it follows immediately from our derivation above. Indeed, from (16), we see that

$$W_0(-xe^{-x}) = -\sum_{n=1}^{\infty} \frac{n^{n-1}}{n!} x^n e^{-nx},$$

so that the right hand side of (17) is $-W_0(-xe^{-x}) = x$, since W is the inverse function of the map $z \mapsto ze^z$. According to Berndt [4], Ramanujan had a more general result of which (17) is a special case and proofs of the related entries that appear in Ramanujan's notebooks are supplied on page 70 of [4]. It appears that Ramanujan re-discovered Lambert's function. Lambert himself introduced his function in 1758 to solve trinomial equations and Ramanujan's "Entry 14" in his famous notebooks does just that. Euler [11] extended Lambert's work but it seems to be E.M. Wright [27] who recognized the importance of the function in solving certain transcendental equations.

Proposition 11 gives immediately the Taylor series around $z = 0$ for all powers of $W(x)$. Indeed,

$$W_0(x)^k = \sum_{n=k}^{\infty} \frac{-k(-n)^{n-k-1}}{(n-k)!} x^n,$$

which can be re-written as

$$\left(\frac{W_0(x)}{x}\right)^k = \sum_{n=0}^{\infty} \frac{k(n+k)^{n-1}}{n!} (-x)^n. \tag{18}$$

This series is valid for $|x| < 1/e$.

In combinatorics, there is a cognate function called the Tree function, denoted $T(x)$ which is the inverse function of the map $w \mapsto we^{-w}$. Using the Lagrange inversion theorem, it is not difficult to see that

$$T(x) = \sum_{n=1}^{\infty} \frac{n^{n-1}}{n!} x^n,$$

which again converges for $|x| < 1/e$. It is so-called because n^{n-1} is the number of labelled trees having n vertices. Clearly $T(x) = -W_0(-x)$. The analog of (18) is

$$T(x)^k = \sum_{n=0}^{\infty} \frac{k(n+k)^{n-1} x^{n+k}}{n!}.$$

If in Proposition 12, we let $\phi(z) = e^z$ and $f(z) = e^{az}$, we obtain

$$\sum_{n=0}^{\infty} \frac{(n+a)^n}{n!} w^n = \frac{e^{az}}{1 - we^z} = \frac{e^{az}}{1-z}.$$

In particular, setting $a = 0$, we deduce

$$\frac{1}{1 - T(x)} = \sum_{n=0}^{\infty} \frac{n^n x^n}{n!}.$$

But W_0 given by (16) is only a (tiny) piece of the inverse function of the complex map $w \mapsto we^w$. We can plot the map $w \mapsto we^w$ for w real and see that its global minimum occurs at $w = -1$. From the graph, it is evident that the inverse function has two real branches, one branch whose range includes $(-1, \infty)$ and another whose range includes $(-\infty, -1]$. These branches are denoted W_0 and W_{-1} respectively. In fact, in analogy with the logarithm function, the inverse function of $w \mapsto we^w$ has countably many branches. More precisely, $W_0(z)$ extends to an analytic function on $\mathbb{C}\backslash(-\infty, -1]$ with -1 as a branch point. The countably many branches are denoted W_k, $k \in \mathbb{Z}$. We refer the reader to [8] for details. In the next section, we discuss the analytic continuation of $W_0(x)$ and give two integral formulas for positive real values of x (Fig. 1).

7 Integral Formulas for $W_0(x)$

This function $W_0(x)$ and its other branches are as ubiquitous in mathematics as the logarithm function and its range of wide applicability is discussed in [8]. Comtet (see pages 228–229 of [7]) derived the asymptotic behaviour of the principal branch

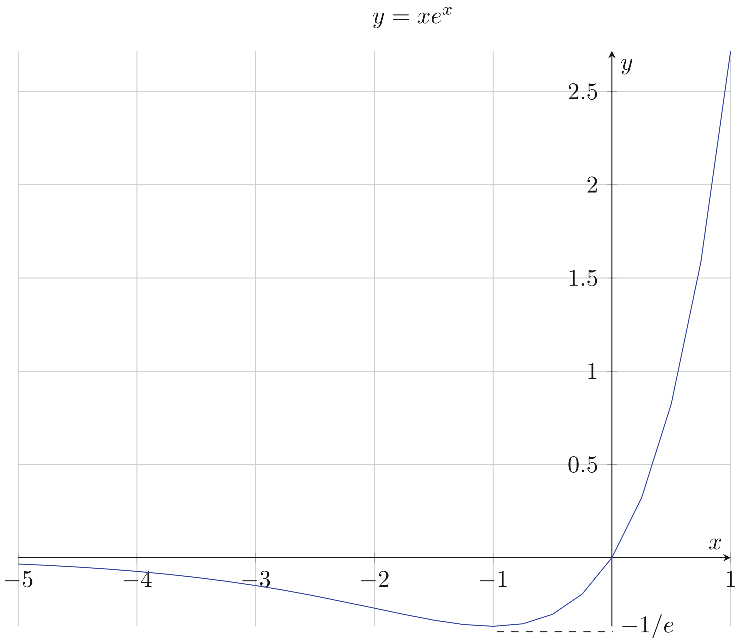

Fig. 1 $y = xe^x$

of $W_0(x)$. He proved that

$$W_0(x) = \log x - \log\log x - \sum_{n=1}^{\infty} \frac{(-1)^n}{(\log x)^n} \sum_{m=1}^{n} s(n, n-m+1) \frac{(\log\log x)^m}{m!},$$

where $s(n, k)$ is the Stirling number of the first kind. As $x \to \infty$, this shows that

$$W_0(x) = \log x - \log\log x + o(1).$$

It is possible to derive a similar expansion for the other branches as in [8]. Introducing the notation

$$\log_k z := \log z + 2\pi i k, \tag{19}$$

as in [8], we have

$$W_k(x) = \log_k x - \log\log_k x - \sum_{n=1}^{\infty} \frac{(-1)^n}{(\log_k x)^n} \sum_{m=1}^{n} s(n, n-m+1) \frac{(\log\log_k x)^m}{m!}.$$

(The "outer" log is the principal branch of the logarithm.) Here, we briefly discuss methods of extending the domain of the Lambert W-function. In this context, it may help the reader to recall how the logarithm function is analytically continued to the cut complex plane $\mathbb{C}\backslash(-\infty, 0]$. On this region, the principal branch of the logarithm is defined

$$\log z = \log|z| + i \arg z. \qquad -\pi < \arg z < \pi.$$

The branches of the logarithm are then given simply by (19). Following [8], we can give an analogous description of the branches of the W-function. But first, we review some basic information concerning the branches of W. Let us write $z = we^w$ with $z = x + iy$ and $w = u + iv$, and u, v, x, y real. Then

$$x = e^u(u \cos v - v \sin v)$$
$$y = e^u(v \cos v + u \sin v)$$

Under this mapping, the image of the x-axis (that is, the curve $y = 0$) consists of the curves

$$v = 0 \quad \text{or} \quad u = -v \cot v$$

in the (u, v)-plane. This means that

$$x = -e^u \frac{v}{\sin v} < 0 \iff (2k-1)\pi < v < (2k+1)\pi, \qquad k \in \mathbb{Z}.$$

The curve which separates the principal branch W_0 from W_1 and W_{-1} is

$$-v \cot v + vi : \quad -\pi < v < \pi.$$

We give an analytic integral formula for $W_0(x)$ for $x > 0$ in the following theorem.

Theorem 13 *For $x > 0$,*

$$\frac{1}{x(1 + W_0(1/x))} = \int_{-\infty}^{\infty} \frac{du}{(e^u - xu)^2 + \pi^2 x^2}$$

Proof Let R_T be the closed rectangular contour with vertices

$$(-T, -\pi), \quad (T, -\pi), \quad (T, \pi), \quad (-T, \pi)$$

oriented counterclockwise. Consider the integral

$$I_T := \int_{R_T} \frac{dw}{e^w + xw}.$$

Since

$$\frac{1}{|e^w + xw|} \le \frac{1}{||e^w| - |x||w||} \to 0, \quad \text{as} \quad |w| \to \infty,$$

we see that the vertical integrals in the contour integral tend to zero as $T \to \infty$. Thus,

$$\lim_{T\to\infty} I_T = \int_{-\infty}^{\infty} \frac{du}{e^{u-\pi i} + x(u - \pi i)} - \int_{-\infty}^{\infty} \frac{dxu}{e^{u+\pi i} + x(u + \pi i)}.$$

Therefore,

$$\lim_{T\to\infty} I_T = \int_{-\infty}^{\infty} \frac{2\pi i x du}{(-e^u + xu)^2 + \pi^2 x^2}$$

Let us examine the integrand of I_T. Writing $w = u + iv$, and keeping in mind that $x > 0$, we see that $e^w + xw = 0$ if and only if

$$e^u \cos v = -xu, \qquad e^u \sin v = -xv.$$

In the region R_T, the second equation $e^u \sin v = -xv$ has a solution only when $v = 0$ in which case the first equation $e^u \cos v = -xu$ becomes $1/x = (-u)e^{-u}$ and we see the solution to this is $-u = W_0(1/x)$. This is the unique singularity of our integrand in the region, for T sufficiently large. If we denote by α this solution, we have

$$\lim_{T\to\infty} \frac{I_T}{2\pi i} = \frac{1}{e^\alpha + x}.$$

Since $e^\alpha = -x\alpha$, we find

$$\int_{-\infty}^{\infty} \frac{du}{(e^u - xu)^2 + \pi^2 x^2} = \frac{1}{x(1 + W_0(1/x))},$$

as claimed. □

The number $W_0(1)$ is often called the "omega constant" and denoted as Ω in the literature. Our integral formula in this special case yields the elegant formula

$$\frac{1}{1+\Omega} = \int_{-\infty}^{\infty} \frac{dx}{(e^x - x)^2 + \pi^2}$$

usually attributed to Adamchik, though there doesn't seem to be any published paper of his on the topic. Our formula can be viewed as a generalization of his result.

8 Zeros of Exponential Polynomials

Equation (13) in Kendall's parametric solution of the SIR model encoded in Theorem 8 is a special case of an exponential polynomial. In this section, we make some number-theoretic remarks concerning the arithmetic nature of roots of such polynomials. This will also lead to some interesting results about special values of the Lambert function.

An exponential polynomial has the general form

$$\sum_{i=1}^{n} p_i(z)e^{\alpha_i z},$$

where $p_i(z)$ is a polynomial for $1 \leq i \leq n$ and $\alpha_1, ..., \alpha_n$ are distinct complex numbers. We want to study zeros of such polynomials when $\alpha_1, ..., \alpha_n$ are algebraic numbers and $p_i(z)$ are non-zero polynomials with algebraic coefficients.

Theorem 14 *If $\alpha_1, ..., \alpha_n$ are distinct algebraic numbers, and $p_1(z), ..., p_n(z)$ are polynomials with algebraic coefficients, then any non-zero root z_0 of*

$$\sum_{i=1}^{n} p_i(z)e^{\alpha_i z} \tag{20}$$

is transcendental unless z_0 is a common root of all the polynomials $p_i(z)$ for $1 \leq z \leq n$.

Proof We recall the following version of a theorem of Lindemann and Weierstrass proved in 1885 (see Theorem 4.1 of [18] on page 15): if $\alpha_1, ..., \alpha_n$ are distinct algebraic numbers, then $e^{\alpha_1}, ..., e^{\alpha_n}$ are linearly independent over the field of algebraic numbers $\overline{\mathbb{Q}}$. So, if z_0 is a non-zero algebraic root of (20), then

$$\sum_{i=1}^{n} p_i(z_0)e^{\alpha_i z_0} = 0.$$

Since $\alpha_1 z_0, ..., \alpha_n z_0$ are then distinct algebraic numbers, and $p_i(z_0)$ are algebraic numbers, this would contradict the Lindemann-Weierstrass theorem, unless z_0 is a common root of all the polynomials $p_i(z)$, $1 \leq i \leq n$. □

Corollary 15 *Any root of (13) is transcendental.*

Proof We need only observe that in Kendall's parametric solution of the SIR model, Eq. (13) is an exponential equation and as all the constants appearing there are rational numbers, an immediate application of the theorem gives the result. □

Corollary 16 *If α is algebraic and lies in $\mathbb{C}\backslash(-\infty, 0]$, then $W_k(\alpha)$ is transcendental.*

Proof $W_k(\alpha)$ is the root of the exponential polynomial equation $ze^z - \alpha = 0$. Since $\alpha \neq 0$, we see that $W_k(\alpha) \neq 0$. The result is now evident. □

A minor variation of this result leads to a small generalization. The previous corollary is the case with $n = 1$ of the following.

Corollary 17 *If $\alpha_1, ..., \alpha_n$ are non-zero algebraic numbers and $c_1, ..., c_n$ are rational numbers, then at least one of the following is true:*

(a) $c_1 W(\alpha_1) + \cdots + c_n W(\alpha_n)$ is either zero or transcendental;

(b) $W(\alpha_1)^{c_1} \cdots W(\alpha_n)^{c_n}$ is non-zero and transcendental,

where W is any branch of the Lambert function.

Proof We have

$$W(\alpha_1)^{c_1} \cdots W(\alpha_n)^{c_n} \exp(c_1 W(\alpha_1) + \cdots + c_n W(\alpha_n)) = \alpha_1^{c_1} \cdots \alpha_n^{c_n},$$

Suppose that $W(\alpha_1)^{c_1} \cdots W(\alpha_n)^{c_n}$ is non-zero and algebraic. Then

$$\exp(c_1 W(\alpha_1) + \cdots c_n W(\alpha_n))$$

is algebraic. But this contradicts the Lindemann-Weierstrass theorem unless $c_1 W(\alpha_1) + \cdots + c_n W(\alpha_n) = 0$. □

As a consequence of these comments on the transcendence of $W(z)$ for algebraic z, we deduce immediately from Theorem 13:

Corollary 18 *For any positive algebraic x, the integrals*

$$\int_{-\infty}^{\infty} \frac{du}{(e^u - xu)^2 + \pi^2 x^2}$$

are transcendental numbers.

A celebrated conjecture of Schanuel (see page 17 of [18]) predicts that if $x_1, ..., x_n$ are linearly independent over the rational number field, then the transcendence degree of

$$\mathbb{Q}(x_1, ..., x_n, e^{x_1}, ..., e^{x_n})$$

is at least n. Thus, if $\alpha_1, ..., \alpha_n$ are algebraic numbers such that $W(\alpha_1), ..., W(\alpha_n)$ are linearly independent over $\mathbb{Q}$, then applying the functional relation of the Lambert function leads to the prediction that the transcendence degree of

$$\mathbb{Q}(W(\alpha_1), ..., W(\alpha_n))$$

is at least n. In other words, Schanuel's conjecture predicts that if $\alpha_1, ..., \alpha_n$ are algebraic numbers such that $W(\alpha_1), ..., W(\alpha_n)$ are linearly independent over $\mathbb{Q}$, then they are algebraically independent. In particular, this would suggest that $W(\alpha_1), ..., W(\alpha_n)$ are linearly independent over $\overline{\mathbb{Q}}$. Perhaps this last implication can be proved using existing techniques from transcendental number theory.

9 The Mellin Transform of $W(x)$

The Mellin inversion formula allows us to derive another integral formula for $W(x)$ for $x > 0$. As noted in [8], the Mellin transform of $W(x)$ is

$$\int_0^\infty W(x)x^{s-1}dx = \int_0^\infty e^{-W(x)}x^s dx,$$

since $W(x)/x = e^{-W(x)}$. Setting $W(x) = u$ so that $x = W^{-1}(u) = ue^u$, and using the fact that

$$dx = e^W(1 + W)dW,$$

we get

$$\int_0^\infty W(x)x^{s-1}dx = \int_0^\infty (ue^u)^s(u + 1)du.$$

Putting $us = -t$, we obtain

$$\int_0^\infty W(x)x^{s-1}dx = -\int_0^\infty e^{-t}(-t/s)^s(-t/s + 1)dt/s.$$

The right hand side is

$$-(-s)^{-s}\int_0^\infty e^{-t}t^s(1 - t/s)dt/s = -(-s)^{-s}\left\{\Gamma(s) - \frac{\Gamma(s+2)}{s^2}\right\}.$$

Using the functional relation $\Gamma(s + 2) = (s + 1)s\Gamma(s)$, this simplifies to

$$\int_0^\infty W(x)x^{s-1}dx = (-s)^{-s}\frac{\Gamma(s)}{s}.$$

Since the logarithm is analytic in $\mathbb{C}\backslash(-\infty, 0]$ and the Γ-function has simple poles only at negative integers, we see that the right hand side is analytic in the region $-1 < \mathrm{Re}(s) < 0$.

For ease of reference, we recall the Mellin inversion formula (see p. 273 of [23]):

Proposition 19 *Suppose that $F(s)$ is a function of the complex variable $\sigma+it$ which is regular in the infinite strip $S=\{s : a<\sigma<b\}$ and for arbitrary small positive number ϵ, $F(s)$ tends to zero uniformly as $|t|\to\infty$ in the strip $a+\epsilon\le\sigma\le b-\epsilon$. Then, if the integral*

$$\int_{-\infty}^{\infty} F(\sigma+it)dt$$

is absolutely convergent for each value of σ in the open interval (a,b) and if for positive real values of x and a fixed $c\in(a,b)$ we define

$$f(x)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{-s}F(s)ds,$$

then in the strip S, we have

$$F(s)=\int_0^{\infty} x^{s-1}f(x)dx.$$

We can apply the Mellin inversion formula to deduce:

Theorem 20 *Let $-1<c<0$. For $x>0$ we have*

$$W(x)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{-s}(-s)^{-s}\frac{\Gamma(s)}{s}ds. \tag{21}$$

Proof To apply Proposition 19, we verify the growth conditions on $(-s)^{-s}\Gamma(s)/s$. To this end, let us recall Stirling's approximation: for $s=\sigma+it$ with $\sigma,t\in\mathbb{R}$, we have

$$|\Gamma(s)|=\sqrt{2\pi}\,|t|^{\sigma-1/2}e^{-\pi|t|/2}(1+O(|t|^{-1}))$$

for $|t|$ sufficiently large. Moreover, writing $\log(-s)=\log|s|+i\arg(-s)$, we see that

$$|(-s)^{-s}\Gamma(s)/s|=|s|^{-\sigma}e^{t\arg(-s)}|t|^{\sigma-3/2}e^{-\pi|t|/2}$$

satisfies the condition of the theorem in that it tends to zero uniformly as $|t|\to\infty$ in the strip $-1+\epsilon\le\sigma\le-\epsilon$. We need only note that in this region $|\arg(-s)|\le\pi/2$. □

Incidentally, the above theorem can be used to derive the power series representation of $W(x)$ for $|x|<1/e$ obtained earlier using the Lagrange inversion theorem. To see this, we truncate the integral of the theorem and view it as

$$\lim_{T\to\infty}\frac{1}{2\pi i}\int_{c-iT}^{c+iT} x^{-s}(-s)^{-s}\frac{\Gamma(s)}{s}ds.$$

Having done this, we move the contour to the left and note that the singularities of the integrand occur at negative integers. The residue is easily calculated. At $s = -n$, it is

$$-x^n n^{n-1}(-1)^n/n!$$

and both the horizontal and vertical integrals tend to zero since $|x| < 1/e$. Summing this over $n \geq 1$ gives (16). (We leave the details to the reader as an exercise in complex analysis.)

Our integral formulas are valid only for $x > 0$ and it would seem natural to inquire if there is an analytic power series that gives the analytic continuation of the principal branch of W. This was done recently by Beardon [2]. Using fairly elementary complex analysis, he shows that for $z \in \mathbb{C}\backslash(-\infty, -1/e]$ we have the following series for the principal branch:

$$W_0(z) = \sum_{m=1}^{\infty} a_m \left(\frac{\sqrt{ez+1}-1}{\sqrt{ez+1}+1}\right)^m, \qquad a_m = \sum_{n=1}^{m} \frac{(-n)^{n-1}}{n!}\left(\frac{4}{e}\right)^n \binom{m+n-1}{m-n}.$$

He also indicates how one may derive similar series representations of the other branches.

After this digression into the theory of the Lambert function, we describe how it relates to an exact solution of the SIR model in the next section.

10 Solving $A + Br + e^{Cr} = 0$ and Kendall's Arrangement

One can solve for roots of exponential equations using the W-function. Indeed, suppose we have an equation of the form

$$A + Br + e^{Cr} = 0, \tag{22}$$

with A, B, C given, and we want to solve for r. Thus, $(A + Br)e^{-Cr} = -1$ and we can re-write this as

$$\frac{B}{C}\left(-\frac{AC}{B} - Cr\right) e^{-Cr-AC/B} = e^{-AC/B}.$$

Thus, $-Cr - AC/B = W_k(Ce^{-AC/B}/B)$ so that all the roots are given by

$$-\frac{A}{B} - \frac{1}{C}W_k\left(\frac{C}{B}e^{-AC/B}\right).$$

For algebraic values A, B, C, the non-zero roots of (22) are transcendental numbers by the theorem of Hermite and Lindemann alluded to earlier. It is curious to note that, by contrast, $W_k\left(-\frac{AC}{B}e^{-AC/B}\right) = -AC/B$ is algebraic.

We apply this discussion to determine the roots of (13) appearing in Kendall's Theorem 8 which we re-write as

$$-[R_0 I(0) + 1] + R_0\zeta + e^{-R_0\zeta} = 0.$$

The positive root ζ_2 is

$$\zeta_2 = \frac{R_0 I(0) - 1}{R_0} + \frac{1}{R_0} W_0\left(-e^{1-R_0 I(0)}\right).$$

Since $R_0 > 1$ in the case of an epidemic, and we may suppose that there are initially at least two infected people, we see that $1 - R_0 I(0) < -1$ so that the value of this root can easily be determined using the power series (16) since the argument of the W-function lies in the domain of absolute convergence.

To find the negative root $-\zeta_1$, we have

$$-\zeta_1 = \frac{R_0 I(0) - 1}{R_0} + \frac{1}{R_0} W_{-1}\left(-e^{1-R_0 I(0)}\right).$$

Interestingly, by our earlier remark, both ζ_1 and ζ_2 are transcendental numbers. This should not be a cause for too much consternation since, after all, the SIR model is not an exact description of "reality" regarding the pandemic but only a mathematical tool to enable us to understand its behaviour.

According to Kendall's arrangement, the total number of people infected during the entire epidemic will be $\zeta_1 + \zeta_2$ which is

$$\frac{1}{R_0}\left(W_0\left(e^{1-R_0 I(0)}\right) - W_{-1}\left(e^{1-R_0 I(0)}\right)\right).$$

11 The Hadamard Product of $A + Bz + e^{Cz}$

The function $f(z) = A + Bz + e^{Cz}$ is of order 1 (in the sense of Hadamard) and as such, admits a factorization of the form

$$e^{c_0 z + c_1} \prod_{\omega} \left(1 - \frac{z}{\omega}\right) e^{z/\omega},$$

where the product is over the zeros of $f(z)$ and c_0, c_1 are appropriate constants. These are easily determined as follows. Setting $z = 0$ gives $e^{c_1} = A + 1$. We will assume $A + 1 \neq 0$. Taking the logarithmic derivative gives

$$\frac{f'(z)}{f(z)} = c_0 + \sum_{\omega} \left[\frac{1}{z - \omega} + \frac{1}{\omega} \right],$$

so by setting $z = 0$ we get $c_0 = f'(0)/f(0) = (B + C)/(A + 1)$. This proves:

Proposition 21

$$A + Bz + e^{Cz} = (A + 1)e^{(B+C)z/(A+1)} \prod_{\omega} \left(1 - \frac{z}{\omega}\right) e^{z/\omega},$$

where the product is over the zeros of $A + Bz + e^{Cz}$.

As we observed in the previous section all the roots of $f(z)$ can be written as

$$\omega_k := -\frac{A}{B} - \frac{1}{C} W_k \left(\frac{C}{B} e^{-AC/B} \right).$$

We also noted that the Hermite-Lindemann theorem shows that ω_k is transcendental whenever A, B, C are algebraic. We can deduce the following curious result, reminiscent of Euler's explicit evalutaion of $\zeta(2n)$.

Proposition 22 *For A, B, C algebraic, the sums*

$$\frac{\ell_{n+1}}{n!} := \sum_{\omega} \frac{1}{\omega^{n+1}}$$

are algebraic for $n \geq 1$.

Proof By Taylor's theorem, we have

$$\frac{1}{n!} \frac{d^n}{dz^n} \left(\frac{f'(z)}{f(z)} \right) \Big|_{z=0} = -\sum_{\omega} \frac{1}{\omega^{n+1}}.$$

The left hand side is an algebraic number. □

12 Kendall's Integral and Lagrange Inversion

The exact solution to the SIR model provided by Kendall [15] amounts to writing the time function as a function of R via an integral of the form

$$\int_0^R \frac{dr}{A + Br + e^{cr}},$$

with suitable values of A, B, C. Though perhaps not practical, we want to show that it is theoretically possible to invert this and write R as a function of the time parameter t. To do this, we first write the above integral as a power series in R as follows. Using our Hadamard factoriztion we see that

$$A + Br + e^{Cr} = (A+1)\exp\left(\frac{(B+C)r}{A+1} + \sum_{\omega}\left[\log\left(1-\frac{r}{\omega}\right)+\frac{r}{\omega}\right]\right).$$

Thus,

$$\frac{1}{A+Br+e^{Cr}} = (A+1)^{-1}\exp\left(-\frac{B+C}{A+1}r + \sum_{n=2}^{\infty}\sum_{\omega}\omega^{-n}\frac{r^n}{n}\right).$$

It seems convenient to define $\ell_1 = -(B+C)/(A+1)$ so that applying (15) gives

$$\frac{1}{A+Br+e^{Cr}} = (A+1)^{-1}\sum_{n=0}^{\infty}B_n(\ell_1, \ldots, \ell_n)\frac{r^n}{n!}.$$

We can now integrate term by term and deduce that

$$\int_0^R \frac{dr}{A+Br+e^{Cr}} = (A+1)^{-1}\sum_{n=0}^{\infty}B_n(\ell_1, \ldots, \ell_n)\frac{R^{n+1}}{(n+1)!}.$$

The right hand side is of the form $Rg(R)$ with $g(0) \neq 0$, as required by Proposition 9. We can then invert this power series using the Lagrange inversion formula.

13 Further Remarks

The approximation of the exponential function with a quadratic polynomial leads to the above analysis. In particular, it implies that there is a single maximum for the number of infected people. In practice however, there seem to be waves of the epidemic pointing to a series of local maxima. So perhaps, this model may have to be modifed by allowing β to be a function of time and introducing an oscillation factor.

It is evident from the discussion of the previous sections that by using higher order polynomial approximations of the exponential function, one is led to the study of integrals of the form

$$\int \frac{dr}{f(r)}$$

where $f(r)$ is a polynomial. One can evaluate such integrals using the method of partial fractions and we immediately see that r is an algebraic function over the field generated by e^t.

In the particular cases of cubic and quartic approximations to the exponential function, we end up with explicit formulas for $r(t)$ in terms of the exponential function since we have formulas for roots of the general cubic and quartic polynomials. The general quintic can also be solved using elliptic functions. Thus, it may be fruitful to push this analysis to at least these three levels of approximation and derive further local maxima of the $I(t)$ function. This would be only of theoretical interest since our analysis shows that the Kermack and McKendrick solution is very accurate if we introduce the constant ϕ' mentioned in Theorem 7.

There are efficient numerical methods that can be used to solve the original differential equation without using the polynomial approximations of the exponential function. Most notable in this array of methods is the Runge-Kutta method. The idea here is to use the differential equation for r as the starting point of developing its Taylor series. We will not go into details but refer the reader to section 14 of Chapter 8 of [3]. Since we are interested in knowing before hand the peaks of the number of infected people at any given time, it is the derivative of the function r that we will need to focus on and determine its zeros. This can be done numerically through these methods.

Acknowledgments We thank Siddhi Pathak, Troy Day and Jianhong Wu for their careful reading and helpful remarks on an earlier version of this paper.

References

1. N.T.J. Bailey, The Mathematical Theory of Infectious Diseases and its Application, Charles Griffin and Company, London, 1975.
2. A.F. Beardon, The principal branch of the Lambert W-function, Computational Methods and Function Theory. https://doi.org/10.1007/s40315-020-00329-6, published online 13 August 2020.
3. G. Birkhoff and G.-C. Rota, Ordinary differential equations, Blaisdell Publishing Company, Waltham, Massachusetts, 2nd Edition, 1969.
4. B. Berndt, Ramanujan's Notebooks, Part I, Springer-Verlag, New York, 1985.
5. D. Bernoulli, Essai d'une nouvelle analyse de la mortalité causée par la petite verole et des avantages de l'inoculation pour la prevenir, *Mem. Phys. Acad. Roy. Soc., Paris.* I, 1760.
6. S. Cioaba, Some applications of eigenvalues of graphs, in Structural Analysis of Complex Networks, edited by M. Dehmer, 357–379, (2011), Springer.
7. L. Comtet, Advanced Combinatorics, D. Reidel Publishing Company, Dodrecht, Holland, 1974.
8. R.M. Corless, G.H. Gonner, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth, On the Lambert W-function, Advances in Computational Mathematics, Vol. 5, (1996), 329–359.
9. M. Draief, A. Ganesh, L. Massoulié, Threshold for virus spread on networks, *Annals of Applied Probability, **18*** (2008), 359–378.
10. D. Easley and J. Kleinbert, Networks, Crowds and Markets: Reasoning about a highly connected world, Cambridge University Press, 2010.

11. L. Euler, De serie Lambertina Plurimisque eius insignibus proprietatibus, reprinted in Opera Omnia, Series Prima, Vol. 6: Commentationes Algebraicae. Leipzig, Germany: Teubner, pp. 350–369, 1921.
12. J. Gani and R.J. Swift, On the Kermack–McKendrick approximation for removals from an epidemic, *Math. Scientist*, ***35***, (2010), 29–36.
13. W.H. Hamer, Epidemic disease in England, *Lancet*, ***1*** (1906), 733–739.
14. T. Harko, F.S.N. Lobo, and M.K. Mak, Exact solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates, *Applied Mathematics and Computation*, ***236*** (2014), 184–194.
15. D.G. Kendall, Deterministic and stochastic epidemics in closed populations, *Proc. Third Berkeley Symp. Math. Statistics and Probability*, ***4*** (1956), 149–165, Berkeley and Los Angeles, University of California Press.
16. W.O. Kermack and A.G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. of the Royal Society A*, ***115*** (1927), 700–721.
17. M. Ram Murty, A Second Course in Analysis, Hindustan Book Agency, 2020.
18. M. Ram Murty and P. Rath, Transcendental Numbers, Springer, 2014.
19. G. Pólya and G. Szegö, Problems and Theorems in Analysis I, Springer-Verlag, New York, 1978.
20. S. Ramanujan, Collected Papers, edited by G.H. Hardy, P.V. Seshu Aiyar and B.M. Wilson, AMS Chelsea Publishing, American Mathematical Society, Providence, Rhode Island, 2013.
21. R. Ross, Some a priori pathometric equations, *British Medical Journal*, ***1*** (1915), 546–547.
22. W. Rudin, Real and Complex Analysis, Third Edition, McGraw-Hill, 1987.
23. I. Sneddon, The Use of Integral Transforms, McGraw-Hill, 1972.
24. R. Stanley, Enumerative Combinatorics, Volume 2, Cambridge University Press, 1999.
25. Y. Wang, D. Chakraborti, C. Wang, C. Faloutsos, Epidemic spreading in real networks: an eigenvalue viewpoint, in Proceedings of the Symposium on Reliable Distributed Computing, Florence, Italy, 25–34.
26. E.T. Whittaker and G.N. Watson, A Course in Modern Analysis, Cambridge University Press, 1927.
27. E.M. Wright, Solution of the equation $ze^z = a$, *Bulletin of the American Mathematical Society*, ***65*** (1959), 89–93.