ON THE RIEMANN HYPOTHESIS - PART 2

JOHN NUTTALL

May 2009

1. INTRODUCTION

This report contains some speculative remarks about several topics that may have relevance to a proof of the Riemann Hypothesis (RH) following the approach laid out in my previous report, "On the Riemann Hypothesis - Part 1". That approach is based on the study of the sequence of coefficients $\{a_n, n = 0,1,\ldots\}$ of the power series $F(z)$, where the entire function $F(z)$ as defined by Csordas et al [Cs86] has only real negative zeros if the RH is correct. The zeros of $F(z)$ occur at $z = -4\varsigma_i^2$, $i = 1,2,\ldots$, where ς_i, $i = 1,2\ldots$ are the imaginary parts of the zeros of the Riemann ς-function (we call them the Riemann zeros for short).

As described in Part 1, Sec. 1, Karlin [Ka68] Chap. 8 shows that a necessary and sufficient condition for the RH to be true is that the semi-infinite matrix A, defined by

$$A_{m,n} = a_{n-m},\ n \ge m;\quad = 0,\ n < m,\qquad m,n = 0,1,2,\ldots, \tag{1.1}$$

must be totally positive (TP). This means that the determinant of each of the sub-matrices $K(n,\lambda)$ of order λ must be positive, where

$$K(n,\lambda)_{i,j} = A_{i,j+n},\quad i,j = 1,2,\ldots,\lambda, \tag{1.2}$$

for $n = 0,1,2,\ldots,\quad \lambda = 1,2,\ldots.$

2. DETERMINANTAL RELATIONS

1. In this section we consider general sequences $\{a_n\}$, not just those with matrix A that is TP. Numerical experiments and the study of special cases have led to the suggestion that several interesting relations may hold. Some or all may already exist in the literature, and indeed we found that (2.2) is given in [Po76], p. 94.

2. Consider the case of $\det K(1,\lambda)$, $\lambda = 1,2,\ldots$, and write

$$\det K(1,\lambda) = a_0^{\lambda+1} d_\lambda, \quad \lambda = 1,2,\ldots . \tag{2.1}$$

An application of the LU decomposition method [Pr92] Chap. 2, shows that

$$\sum_{i=0}^{j} (-1)^i a_i d_{j-i} = 0, \quad j = 1,2,\ldots, \tag{2.2}$$

with d_0 defined as $1/a_0$. The quantities $\gamma_j(1) = a_0 d_j / d_{j-1} = \det K(1, j) / \det K(1, j-1)$ appear naturally in the LU method. Similarly we may define

$$\gamma_j(n) = \det K(n, j) / \det K(n, j-1). \tag{2.3}$$

If we define $G_1(z)$ by

$$G_1(z) = \frac{1}{a_0} \sum_{j=0}^{\infty} \det K(1, j) \left[z / a_0 \right]^j = \sum_{j=0}^{\infty} d_j z^j \tag{2.4}$$

then (2.2) is equivalent to

$$F(-z) G_1(z) = 1, \tag{2.5}$$

a relation studied in [Ka68], p. 394. In particular, it is noted by Karlin that, if the matrix A is TP, then so is the corresponding matrix D derived from the coefficients $\{d_i\}$ of $G(z)$ in the same manner as in (1.1). From D we may construct a family of sub-matrices $L(n,\lambda), \quad n = 0,1,\ldots; \quad \lambda = 1,2,\ldots,$ analogous to $K(n,\lambda)$.

3. It appears that

$$\det L(n,\lambda) = \det K(\lambda, n) \tag{2.6}$$

for all appropriate values of n, λ.

4. For all relevant values of n, λ it appears that

$$\begin{aligned} \det K(n-1,\lambda) \det K(n+1,\lambda) &= \left[\det K(n,\lambda) \right]^2 - \det K(n,\lambda-1) \det K(n,\lambda+1) \\ &= \left[\gamma_\lambda(n) - \gamma_{\lambda+1}(n) \right] \det K(n,\lambda-1) \det K(n,\lambda). \end{aligned} \tag{2.7}$$

There is an equivalent relation for $\det L(n,\lambda)$.

3. RELATION OF DETERMINANTS TO RIEMANN ZEROS

1. The function $F(z)$ is entire, so that (2.5) shows that $G_1(z)$ is meromorphic with poles at $z = 4\varsigma_i^2$, $i = 1,2,\ldots$. The quantity γ defined by $\gamma(1) = \lim(\gamma_j(1)), \quad j \to \infty$, is related to the first Riemann zero by

$$\gamma(1) = a_0\left(4\varsigma_1^2\right)^{-1} \tag{3.1}$$

Our numerical calculations confirm this relation to high precision.

2. We do not know of any previous information about the corresponding limits $\gamma(n)$ of $\gamma_j(n) = \det K(n, j)/\det K(n, j-1)$ for values of $n > 1$, but, for some low values of n, we also find excellent apparent convergence for the ratio as $j \to \infty$. This allows us to speculate that $G_n(z) = \dfrac{1}{a_0}\sum_{j=0}^{\infty} \det K(n, j)\left[z/a_0\right]^j$ is a meromorphic function. The numerical results provide strong evidence that the pole of $G_n(z)$ nearest to the origin occurs at

$$z = a_0\gamma(n)^{-1} = \prod_{i=1}^{n}\left[4\varsigma_i^2\right]. \tag{3.2}$$

3. By further analyzing the limiting behavior of $\gamma_j(n)$, we may also deduce information about the next nearest pole of $G_n(z)$. For the case $n = 1$, the location is $z = 4\varsigma_2^2$ as expected. For higher values of n, it appears that the second nearest pole is at

$$z = \left[\prod_{i=1}^{n-1} 4\varsigma_i^2\right] 4\varsigma_{n+1}^2 . \tag{3.3}$$

More analysis of the numerical results would provide information about some higher poles.

4. SUCCESSIVE APPROXIMATION AND $n = 1$ DETERMINANTS

1. In this section we provide an account of an analysis of the determinants of the sub-matrices $K(1, \lambda)$, $\lambda = 1,2,\ldots$, using both mathematical and computational information. In this particular case the form of the equations for the LU decomposition method [Pr92] is equivalent to a scheme for the successive approximation of the quantity γ defined above by $\gamma(1) = \lim(\gamma_j(1))$, $j \to \infty$. Because we are restricting attention to the case $n = 1$ we

will here denote $\gamma_j(1)$ by γ_j, and similarly for γ. The LU method produces the following equations

$$\gamma_j = a_1 - \frac{a_0}{\gamma_{j-1}} a_2 + \frac{a_0^2}{\gamma_{j-1}\gamma_{j-2}} a_3 - \frac{a_0^3}{\gamma_{j-1}\gamma_{j-2}\gamma_{j-3}} a_4 \ldots \pm \frac{a_0^{j-1}}{\gamma_{j-1}\gamma_{j-2}\cdots\gamma_1} a_j, \quad j = 1,2,\ldots \quad (4.1)$$

which start with $\gamma_1 = a_1$.

As implied in Sec. 3,

$$\det K(1,\lambda) = \prod_{i=1}^{\lambda} \gamma_j, \quad \lambda = 1,2,\ldots, \quad (4.2)$$

and we know from the analytic properties of $G_1(z)$ that

$\lim(\gamma_j) = \gamma = a_0(4\varsigma_1^2)^{-1}$, $j \to \infty$. Let us set $x(j) = a_0/\gamma_j$, $x = a_0/\gamma = 4\varsigma_1^2$, so that (4.1) becomes

$$\frac{1}{x(j)} = a_1 - a_2 x(j-1) + a_3 x(j-1)x(j-2) \ldots \pm a_j \left[x(j-1)x(j-2)\ldots x(1)\right], \ldots j = 1,2,\ldots. \quad (4.3)$$

We know from computations that $\{x(j)\}$ is a sequence of positive numbers that increase monotonically to the limit x. For large j the higher terms of (4.3) become negligible because of the fast decline of a_i, $i \to \infty$. Also the remaining terms on the right approach $\pm a_i x^{i-1}$, so (4.3) is close to the form

$$\frac{1}{x} = \sum_{i=1} (-1)^{i-1} a_i x^{i-1} \quad \text{or} \quad \sum_{i=0} (-1)^i a_i x^i = 0, \quad (4.4)$$

which is equivalent to $F(-x) = 0$ since the higher terms in (4.4) contribute little. This is consistent with the fact that $x = 1/\gamma = 4\varsigma_1^2$.

To summarize, we are pointing out that (4.3) is a (as yet unproven) method of calculating the value of the first Riemann zero by successive approximation. Similarly, for $n > 1$ the LU method should lead to a way of calculating higher Riemann zeros by successive approximation, based on (3.2). This is what our computations have done.

2. We can go further in this direction. If we define $y(j) = x(j) - x(j-1)$, then it is possible to show that (4.3) is equivalent to

$$y(j) = \sum_{i=1}^{j-1} M(j-1,i)\, y(j-i), \tag{4.5}$$

where

$$M(j-1,i) = \frac{1}{g(j)g(j-1)} \sum_{k=i+1}^{j} (-1)^k a_k X(j-1,k-1). \tag{4.6}$$

Here we have defined

$$X(j,k) = \prod_{i=1}^{k-1} x(j-i) \tag{4.7}$$

and

$$g(j) = \sum_{k=1}^{j} (-1)^{k+1} a_k X(j,k). \tag{4.8}$$

Again suppose that j is adequately large, so that in the important earlier terms in the sums in (4.6), (4.8) $x(i) \approx x$. (Remember that the decrease of a_k for large k makes the higher terms unimportant.) Thus, as in (4.4) we have $g(j) \approx x^{-1}$ and

$$M(j-1,i) \approx \sum_{k=i+1} (-1)^k a_k x^k. \tag{4.9}$$

Thus (4.5) may be written

$$y(j) = \sum_{i=1} m(i)\, y(j-i) \tag{4.10}$$

where $m(i) = M(j-1,i)$ is approximately independent of j.

The solution of the recurrence relation (4.10) for $y(j)$ is a linear combination of terms μ^j, where μ is one of the N_1 solutions of

$$\mu^j = \sum_{i=1}^{N_1} m(i)\mu^{j-i} \quad \text{or} \quad 1 = \sum_{i=1}^{N_1} m(i)\mu^{-i}, \tag{4.11}$$

and N_1 is the number of significant terms chosen from (4.10).

Because of the definition of *y*(*j*) we know that $y(j) \to 0,\ j \to \infty$, so that we require $|\mu| < 1$. Of the different solutions of (4.11) satisfying that condition, we choose the one (assuming it to be unique) that has largest $|\mu|$, for that one will eventually dominate *y*(*j*).

Suppose that $N_2 >> N_1$ is the number of terms kept in (4.9). Substitute (4.9) in (4.10) and set $y(j) = \mu^j$, where μ is the chosen solution of (4.11). After some rearrangement, we find that (4.10) becomes

$$1 = a_2 x^2(\mu^{-1}) - a_3 x^3(\mu^{-1} + \mu^{-2}) + a_4 x^4(\mu^{-1} + \mu^{-2} + \mu^{-3}) \ldots \qquad (4.12)$$

Taking (4.4) into account, this leads after some manipulations to

$$1 - a_1 q + a_2 q^2 - a_3 q^3 + a_4 q^4 \ldots \approx 0 \qquad (4.13)$$

where $q = x/\mu$.

The series in (4.13) is finite but the equation tells us that $z = -q$ is an approximate solution of $F(z) = 0$. There is a solution of (4.13) with $\mu = 1$, but, on account of our restrictions on μ, it must be that $x/\mu = 4\varsigma_2^2$ so that $\mu = x/(4\varsigma_2^2) = \varsigma_1^2 / \varsigma_2^2$. This 'prediction' is in close agreement with numerical calculations of the limit of $y(j)/y(j-1)$ for large j.

3. The foregoing analysis leads to a question that I am not prepared to answer at this time. Suppose for the case $n = 1$ we numerically calculated $x(j)$ and related derived quantities for $j = 1, \ldots, J$, where *J* is chosen to be suitably large. With the help of information about the operation of the software used, error bounds on the quantities should be calculated and made appropriately small.

QUESTION Is it possible to produce a rigorous proof that $x(j) > 0$ for all $j > J$ using the above numerical data and an analysis based on the ideas discussed above?

We remark that, even if the answer were affirmative, and the same process could be applied to $\det K(n, \lambda),\ \lambda = 1, 2, \ldots$ for $n > 1$, it would not constitute a proof of the RH, because we cannot perform an infinite number of calculations. Nevertheless, an affirmative result might still be useful.

4. We believe that a more fruitful approach will be an extension of the technique described in Part 1. There $\det K(n, \lambda)$ was given as a $\lambda -$ dimensional integral. In the case of Part 1, it was shown that, for large *n*, fixed λ, the integral was dominated by the contribution from a small part of the space, and the structure of the integrand led to relations that showed the integral was positive, as desired for the RH to be true. In the

situation discussed above where *n* is small, that is no longer the case. Nevertheless it is worth exploring whether cancellations among contributions from different parts of the space can be controlled to produce a positive result.

5. The weight function $\Phi(t)$ that determines the coefficients $\{a_n\}$ (see Part 1, [Cs86] and [Ed74] is closely related to a type of degenerate theta function $\psi(x)$, where

$$\psi(x) = \sum_{m=1}^{\infty} \exp(-m^2 \pi x). \qquad (5.1)$$

The function $\Phi(t)$ contains a sum related to that in (5.1). For all *t* in the range of integration it has been shown [Cs86] that the first term in the sum is greater than 200 times the remaining contributions. In the case covered in Part 1, only the first term was significant. However, we have shown, not surprisingly, that there are values of (n, λ) where the second and higher terms are important. For example, If we remove secondary terms from the definition of $\Phi(t)$ when calculating some low order coefficients, we find that $\det K(2,18)$ becomes negative. In this and many other cases the RH truth depends critically on the subsidiary terms in $\Phi(t)$.

The need to include higher terms in the $\Phi(t)$ sum could lead to significant complications when studying the multiple integral for the determinant. However, there is a possibility that the different contributions to $\Phi(t)$ are dominant in non-overlapping parts of the space of integration. If we can find an example of this behavior for low enough λ we might be able to study the question numerically.

REFERENCES

Cs86 G. Csordas, T.S. Norfolk and R. S.Varga, "The Riemann Hypothesis and the Turan Inequalities," Trans. Amer.Math. Soc., **296**, 521 (1986)

Ed74 H. M. Edwards, "Riemann's Zeta Function," New York: Academic Press (1974)

Ka68 S. Karlin, "Total Positivity," Stanford: Stanford University Press (1968)

Po76 G. Polya and G. Szego, "Problems and Theorems in Analysis," Vol II, Springer (1976)

Pr92 William H. Press et al,. "Numerical Recipes in Fortran 77," Cambridge University Press (1992)