Computing the Incomplete Gamma Function

Kimberly Spears

ARC/UC LEADS, Mathematics

University of California, Santa Barbara

Mentor: Jeffrey Stopple, Ph.D.

ABSTRACT. An algorithm to approximate the incomplete gamma function $G(s,x)$ is described. This approximation is then evaluated approximately one million times for various values of s. This algorithm proves to be a highly efficient method for evaluating the incomplete gamma function, and should provide assistance for further investigation of the Dirichlet L-Series.

Special thanks to Jeffrey Stopple.

1. Introdution

The incomplete gamma function is defined to be,

$$G(s,x) = \int_x^\infty e^{-t} t^s \frac{dt}{t}.$$

It appears in the evaluation of the Dirichlet L-Series,

$$L(s,\chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s}.$$

The Dirichlet L-series is a zeta function, and is thus related to the Riemann zeta function. This means the study of the former might lead to insight in proving the extended hypothesis of the latter. In the Riemann Hypothesis, Riemann conjectured that all zeroes of $L(s,\chi)$ lie on the line $Re(s) = \frac{1}{2}$. Though 150 years have passed, no one has been able to prove this conjecture. By analyzing the distribution of the zeros of the L-function, a series with similar properties, information might be revealed which can assist in this larger goal.

Specifically, we will analyze the spacing of the zeros of $L(s,\chi)$ and look for a correlation between them and the class number of binary quadratic forms of discriminant q. In order to do this we need an efficient way to evaulate the L-function many times. This is where the incomplete gamma function comes in.

As seen above, $L(s,\chi)$ is an infinite series which converges slowly on the line $Re(s) = \frac{1}{2}$. It is difficult to compute because of this. However, if we can relate $L(s,\chi)$ to a Θ function defined as

$$\Theta(t,\chi) = \sum_n n\chi(n)e^{\frac{-\pi n^2 t}{q}},$$

then our computation becomes easier and more efficient since $\Theta(t,\chi)$ converges rapidly due to exponential decay.

To relate $L(s,\chi)$ to $\Theta(t,\chi)$, we use a function $\Lambda(s,\chi)$, which is essentially $L(s,\chi)$ times a fudge factor. We then apply a Mellin transform to $\Theta(t,\chi)$, which and then show this is equal to $\Lambda(s,\chi)$. After much computation, we see in our expression the piece

$$G(s,x) = x^{\frac{-s-1}{2}} \int_x^\infty e^{-y} y^{\frac{s+1}{2}} \frac{dy}{y} \tag{1}$$

which is equivalent to $x^{\frac{-s-1}{2}} G(\frac{s+1}{2}, x)$. This explains the motivation to find an *efficient* method for computing $G(s,x)$ many times and for large values. [1]

Such an efficient method comes from combining three different numerical methods: Chebyshev polynomials, Pade approximants, and continued fractions.

[1] When we talk about computing the incomplete gamma function, we mean computing this function, $G(s,x)$.

Chebyshev polynomials are a moderately effective method for approximating functions. They revolve around the geometric idea and application of orthogonal projection. We have our original function, $G(s,x)$, and our approximation, call it $G(s,x)'$, and then use the orthogonal projection of $G(s,x)$ onto $G(s,x)'$ to minimize the distance between them. We used the Chebyshev polynomials because of their convienant properties from their orthogonality relation.

Our second tool is the Pade approximation method. Given a function, we determine its infinite power series and then create a rational function whose power series expansion agrees with our given power series to the highest possible order. Usually a power series refers to the typical Taylor Series polynomial. However, here we use a Chebyshev polynomial in its place **[1]**. For example, a term containing x^2 would be replaced by the second Chebyshev polynomial. These two concepts combined allow us to hybridize the ideas of orthogonality and power series. Using Pade approximation with Chebyshev polynomials, we are able to get an extremely accurate approximation to our incomplete gamma function, $G(s,x)$.

Lastly, while the approximation is quite accurate, the efficiency of the computation can be improved. That is, once completed, our approximation function, earlier referred to as $G(s,x)'$, is in the form of a rational polynomial function. For example, if our approximation function were

$$\frac{a_0T_n(x) + a_1T_{n-1}(x) + ... + a_nT_0(x)}{T_n(x) + b_1T_{n-1}(x) + ... + b_nT_0(x)},$$

then to compute this would require n multiplications and n additions on top and bottom. However, if converted to a continued fraction it would require only n divisions and $2n$ additions total. Thus, after finding our approximation function we convert it into continued fraction form.

Finally, to display the accuracy of our approximation function, we compare the difference between it and $G(s, x)$.

2. Chebyshev Approximation

2.1. Orthogonality. The Chebyshev polynomials are what we use to approximate the incomplete gamma function. The concept that makes this approximation desirable is orthogonality. This requires a brief background on the topic.

Take any two vectors, $\vec{v}$ and $\vec{w}$, and place them at the origin of the coordinate plane. Then the orthogonal projection of $\vec{v}$ onto $\vec{w}$ is a vector drawn from $\vec{v}$ to $\lambda\vec{w}$, for some scalar λ, such that $\lambda\vec{w}$ and this new vector, $p_{\vec{v}}$, are perpindicular to each other. Hence, the term *orthogonal projection*. From a geometric view, we see $p_{\vec{v}}$ is equal to

$\vec{v} - \lambda\vec{w}$, and the length of $p_{\vec{v}}$ is the *minimum distance* between the two vectors. This information is elementary, yet important since it is a good tool for creating approximation methods, as seen next.

Now, let $\vec{v}$ represent $G(s, x)$, our incomplete gamma function, and suppose $\vec{w}$ represents a ”nice function”, or rather, a function that is relatively easy to compute. $G(s, x)$, is then the ”nasty function” in comparison. As before, to approximate $G(s, x)$, we want to minimize the distance between $G(s, x)$ and $\vec{w}$. To do this, we must find a scalar, λ. From linear algebra, we know two vectors are perpindicular to each other if their inner product is equal to zero. Then for the two-dimensional case, we can just solve for λ, as shown,

$$\begin{aligned}
0 &= \langle \lambda\vec{w}, \vec{v} - \lambda\vec{w} \rangle \qquad (2)\\
&= \langle \lambda\vec{w}, \vec{v} \rangle + \langle \lambda\vec{w}, -\lambda\vec{w} \rangle \qquad \text{(by linearity)}\\
&= \lambda\langle \vec{w}, \vec{v} \rangle - \lambda^2\langle \vec{w}, \vec{w} \rangle\\
\Rightarrow \lambda &= \frac{\langle \vec{w}, \vec{v} \rangle}{\langle \vec{w}, \vec{w} \rangle}.
\end{aligned}$$

This idea can be extended to higher dimensions, and the definition of orthogonality on the vectors $\vec{v}$ and $\vec{w}$ can be redefined if we wish to use two *functions*, instead of vectors. This is because the idea of orthogonality stems from the inner product, so that two vectors or functions are orthogonal if their inner product is zero. The inner product is bilinear, symmetric, and positive. These properties, and particularly the linearity property, are what allow us to use inner product to determine and find an orthogonal projection. In two dimensions, the geometry of two perpendicular vectors leads to the algebra that their inner product must be zero. However, in terms of two functions, we want to define our inner product *first* and *then* see what geometry we get in return. For example, the definition of the inner product of two functions, $f(x)$ and $g(x)$, for $a \leq x \leq b$, is

$$\langle f, g \rangle = \int_a^b f(x)g(x)w(x)dx, \tag{3}$$

where $w(x)$ is a given *weight* of the function **[2]** . It follows that $f(x)$ is said to be *orthogonal* to $g(x)$ if

$$\langle f, g \rangle = 0.$$

In our case, as mentioned before, $\vec{v}$ is replaced by the incomplete gamma function, $G(s, x)$, and the "nice functions" we referred to earlier are the Chebyshev polynomials. The properties that Chebyshev polynomials have will allow us to find a "nice" orthogonality relation for them similar to (3).

2.2. Definition of a Chebyshev polynomial. A Chebyshev polynomial of the n$^{\text{th}}$ degree is defined to be,

$$T_n(x) = \cos(n \arccos x). \tag{4}$$

To make this look like a polynomial, we do a change of variables with $x = \cos\phi$ and $\phi = \cos^{-1} x$, which gives us,

$$T_n(x) = \cos(n\phi), \tag{5}$$

on a range $-1 \leq x \leq 1$. Then we use the trig identity, $\cos((n+1)\phi) + \cos((n-1)\phi) = 2\cos\phi\cos n\phi$, which gives us,

$$T_{n+1}(x) = 2T_1(x)T_n(x) - T_{n-1}(x). \tag{6}$$

Substituting $n = 0$ up to $n = 4$ into (6), we get the first four terms of the Chebyshev polynomial in ”polynomial” form,

$$
\begin{aligned}
(7) \qquad & T_0(x) = 1 \\
& T_1(x) = x \\
& T_2(x) = 2x^2 - 1 \\
& T_3(x) = 4x^3 - 3x \\
& T_4(x) = 8x^4 - 8x^2 + 1
\end{aligned}
$$

Now that we can see they are polynomials, the next step is to show that they have properties.

2.3. Properties of the Chebyshev polynomials. The Chebyshev polynomials are orthogonal in the interval $[-1, 1]$ over a weight $(1-x^2)^{\frac{-1}{2}}$, where the inner product is defined as

$$\langle T_i, T_j \rangle = \int_{-1}^{1} \frac{T_i(x)T_j(x)}{\sqrt{1-x^2}} dx,$$

since,

$$\int_{-1}^{1} \frac{T_i(x)T_j(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0, & i \neq j \\ \frac{\pi}{2}, & i = j \neq 0 \\ \pi, & i = j = 0 \end{cases} \tag{8}$$

Since $T_n(x)$ is a polynomial of the n$^{\text{th}}$ degree, we know there are n zeros. Solving for the zeros of $T_n(x)$, we see they are located at

$$x = \cos(\frac{\pi(k-\frac{1}{2})}{n}), \qquad k = 1, 2, \ldots, n. \tag{9}$$

Notice all the zeros lie in the interval, $[-1, 1]$, the same interval on which the orthogonality relation is defined. Similarly, there are $n+1$ extrema, which are also all on this interval, and located at the points,

$$x = \cos(\frac{\pi k}{n}), \qquad k = 0, 1, \ldots, n.$$

All the maxima of $T_n(x)$ occur at $T_n(x) = 1$, while all the minima occur at $T_n(x) = -1$. This property is one that makes the Chebyshev polynomials useful. See Appendix (A) for an explanation.

Our continuous orthogonality relation, however, is not so useful. When we find $G(s, x)$ in terms of a Chebyshev polynomial, and, after adjusting the range, try to plug it into (8), we get a brutal integral,

$$\int_{-1}^{1} T_k(x) \int_{w}^{\infty} e^{-y} y^{\frac{1}{2}+\imath(2N-1+x)} \frac{dy}{y} dx.$$

Fortunately, instead of using integrals, we can use sums by defining an inner product with a *discrete* orthogonality relation:

$$\sum_{k=1}^{m} T_i(x_k)T_j(x_k) = \begin{cases} 0, & i \neq j \\ \frac{m}{2}, & i = j \neq 0 \\ m, & i = j = 0 \end{cases} \tag{10}$$

This discrete orthogonality relation is our key; it allows us to find an orthogonal projection of $G(s,x)$ onto a subspace of Chebyshev polynomials. Just as we did for the two-dimensional case, recall in two dimensions, to find our vector $\lambda\vec{w}$, we simply were left with having to solve for λ with,

$$\lambda = \frac{\langle \vec{w}, \vec{v} \rangle}{\langle \vec{w}, \vec{w} \rangle}.$$

Analogously, to find the orthogonal projection, p_f, of $f = G(s,x)$ onto a Chebyshev polynomial of the i^{th} degree, T_i, we are left to solve

$$p_f = \sum_{i=0}^{m-1} \frac{\langle f, T_i \rangle}{\langle T_i, T_i \rangle} T_i,$$

using the orthogonality relation from Equation (10).

Now that we see the orthogonality relation of the Chebyshev polynomials, we are ready to see an approximation formula for an ”explicit” way to approximate $G(s,x)$. Using algebra, substitution, and equations, (4), (9), and (10), we get a theorem which states, THEOREM 1. *If f(x) is a function in the interval* $[-1, 1]$*, and if N coefficients are defined by,*

$$c_j = \frac{2}{N} \sum_{k=1}^{N} f(x_k) T_j(x_k) = \frac{2}{N} \sum_{k=1}^{N} f\left[\cos(\frac{\pi(k-\frac{1}{2})}{n})\right] \cos\left[\cos(\frac{\pi j(k-\frac{1}{2})}{n})\right], \tag{11}$$

then

$$f(x) = \left[\sum_{k=0}^{N-1} c_k T_k(x)\right] - \frac{1}{2} c_0$$

for x equal to all of the N zeros of $T_n(x)$*.*

Notice that this theorem gives an exact approximation of f(x) for a *discrete* number of points, those points being the zeros of the function, x_k for $k = 1, \cdots, N$. Hence, a question might arise: what is the value in this? We could find many approximation formulas who are exact for a certain number of discrete values that would give just as accurate an approximation. The value of the Chebyshev polynomials is that we can truncate our polynomial of N$^{\text{th}}$ degree to a lower degree m, $m \ll N$, and be left with an excellent approximation.

The idea is this: If we take a really big value for N, then the approximation to our function is virtually perfect, since our approximation is *exact* for all of the N zeros. Then, by computing the coefficients, using the formula for the c_j's in Theorem (1), we get a lot of "very accurate" coefficients, so to speak. But, instead of evaluating the approximation formula given in Theorem (1) up to $N-1$, we evaluate it up to our truncated polynomial degree, m, well, actually to m-1. Visually, we see it as,

$$f(x) \approx \left[\sum_{k=0}^{m-1} c_k T_k(x)\right] - \frac{1}{2}c_0. \tag{12}$$

This is where the convienence of the extrema of $T_n(x)$ comes in. Because $-1 \leq T_n(x) \leq 1$, we know the error when using (12) can be no more than the sum of the rest of the coefficients, $c_m + \ldots + c_{N-1}$. Furthermore, the c_k's are rapidly descreasing, which means the error is dominated by the ${c_m}^{\text{th}}$ term. This new formula is accurate *and* easy to compute.

In addition to this truncating tool, we are also given a change of variables formula to change the intervals of a function we want to use, into the required one, $[-1, 1]$.

Now that we have this approximation tool, which centers around the idea of orthogonality, we are ready to hybridize this method with one that centers around the idea of the power series.

3. Pade Approximants

The Chebyshev approximation method is useful for minimizing error on line segments because it uses orthogonality, as shown in the last section. However, Pade approximantion is a tool that relates a given power series to an approximating rational function. This would be advantageous to us since a rational function allows more flexibility than a polynomial.

If we consider the Chebyshev series to be our power series, then we may hybridize the Pade approximant method with the Chebyshev approximation method.

3.1. Hybridizing Chebyshev and Pade. Instead of using the Chebyshev expression we found in Equation (12), we want to make $f(x)$ equivalent to an *infinite* Chebyshev series. So now we have,

$$f(x) \equiv \sum_{n=0}^{\infty} \frac{\langle f, T_n \rangle}{\langle T_n, T_n \rangle} T_n(x). \tag{13}$$

The entire point of this new method is to make the approximation more accurate by *avoiding* truncating the power series of f(x), or in our case, the incomplete gamma function. Notice the $\equiv$ sign, instead of the $\approx$ sign. This is just like writing a function in terms of its equivalent power series, except we are writing it in terms of its equivalent Chebyshev series, if you will.

Now should come the question, but what will we do instead of truncating? The answer is to find a rational function whose Chebyshev series expansion agrees with (13) to the highest possible order, (which will be one of our choosing). In other words,

THEOREM 2. *Given a function, f(x), with*

$$f(x) \equiv \sum_{n=0}^{\infty} \frac{\langle f, T_n \rangle}{\langle T_n, T_n \rangle} T_n(x).$$

Then the Pade approximant to f(x) is given by,

$$R(x) \equiv \frac{\sum_{k=0}^{M} a_k T_k(x)}{1 + \sum_{k=0}^{N} b_k T_k(x)},$$

if

$$R(x) = \sum_{k=0}^{\infty} c_k T_k(x)$$

so that

$$c_k = \frac{\langle f, T_k \rangle}{\langle T_k, T_k \rangle}, \qquad k = 1, 2, \ldots, M + N.$$

This should make sense and appear rather straightforward. However, finding the actual coefficients of our approximating function, $R(x)$, is a little more gritty.

3.2. Finding the Pade Approximant. To determine $R(x)$, we must begin by deciding what degree each polynomial of our rational function should have. (See Theorem (2)). We choose $M = N = 3$ because the error for this choice is small.

From here, it is simply algebra. We begin by writing out our new rational function with its presently unknown coefficients, and set it approximately equal to the function we want to approximate, $f(x)$,

$$f(x) = \sum_{k=0}^{\infty} c_k T_k(x) \approx \frac{a_0 T_0 + a_1 T_1 + a_2 T_2 + a_3 T_3}{b_0 T_0 + b_1 T_1 + b_2 T_2 + T_3}.$$

Because of the recurrence relation of the Chebyshev polynomials, we multiply with the property, $T_i T_j = (T_{i+j} + T_{i-j})\frac{1}{2}$. Using this, the next step is to cross multiply and subtract,

$$(b_0 T_0 + b_1 T_1 + b_2 T_2 + T_3) \sum_{k=0}^{\infty} c_k T_k(x) - (a_0 T_0 + a_1 T_1 + a_2 T_2 + a_3 T_3) = 0.$$

The next step involves a lot of computation. From here, we multiply out a finite number of terms up to the T_9 th level. Finally, we are left with a system of three linear equations which we put into a matrix to solve for the b's. Then we are also left with a set of four more equations which we will plug the b's into to find the a's. (Remember that the c's are already known values.) Visually, the matrix and equations, respectively, are,

$$\begin{vmatrix} 2c_4 & c_3+c_5 & c_2+c_6 \\ 2c_5 & c_4+c_6 & c_3+c_7 \\ 2c_6 & c_5+c_7 & c_4+c_8 \end{vmatrix} \begin{vmatrix} b_0 \\ b_1 \\ b_2 \end{vmatrix} = \begin{vmatrix} -c_1-c_7 \\ -c_2-c_8 \\ -c_3-c_9 \end{vmatrix}, \tag{14}$$

and,

$$a_0 = (2b_0c_0 + b_1c_1 + b_2c_2 + c_3)/2$$

$$a_1 = (2b_0c_1 + 2b_1c_0 + b_1c_2 + b_2c_1 + c_2 + c_4 + b_2c_3)/2$$

$$a_2 = (2b_0c_2 + b_1c_1 + b_1c_3 + 2b_2c_0 + b_2c_4 + c_5 + c_1)/2$$

$$a_3 = (2b_0c_3 + b_1c_2 + b_1c_4 + b_2c_1 + b_2c_5 + 2c_0 + c_6)/2$$

We plugged (14) into Mathematica to compute these values.

While they give a very accurate approximation, (whose data will be displayed in the next section), the evaluation of them in their rational function form can be made more efficient. We will still use our hybridized Chebyshev and Pade approximation, but convert the final approximation function into continued fraction form.

4. Converting to Continued Fractions

Continued fractions is not a new method of approximation. It is just another way to represent a number or function. For a number, it is a process of subtracting off the number's integer part, taking the reciprocal of this, and repeating again and again. For a function it is not as straightforward. In our case we used an algorithm to convert our third degree Pade approximant into continued fraction form. The step to convert a first degree Chebyshev/Pade approximant into continued fraction form is shown:

$$
\begin{aligned}
&\frac{a_1 T_1 + a_0 T_0}{T_1 + b_0 T_0} \\
&= \frac{a_1 T_1 + a_1 b_0 T_0 - a_1 b_0 T_0 + a_0 T_0}{T_1 + b_0 T_0} \\
&= a_1 + \frac{a_0 - a_1 b_0}{T_1 + b_0}
\end{aligned}
\tag{15}
$$

The second and third degree conversions follow by induction and the recursion relation of the Chebyshev polynomials.

Now that we have our approximation function in its final form, we can compare it to the original incomplete gamma function.

5. Results

5.1. Adjustments. Remember, for the larger goal, we want to evaluate $L(s, \chi)$, which, as explained in section (1), involves evaluating the function, $\Lambda(s, \chi)$. However, evaluating $\Lambda(s, \chi)$ for a large discriminant involves a large amount of computation. To fix this, we use a Taylor series expansion. Instead of looking at each possible value for x in f, we are going to look at a collection of points with a power series expansion at each one.

The collection of points, called basepoints, are the x_0 values. They are calculated with the function,

$$x(j) = \frac{\left(\frac{e+1}{e-1}\right)^{j-1}}{2} + \frac{1}{2}.$$

The shift distance between these basepoints is named ks.

Then the new function, essentially $G(\frac{s+1}{2}, x)$ with some stuff added to it, is

$$G(\frac{3}{4} + \imath\frac{t+1}{2} + ks, x(j)\Delta) = x(j)^{-\frac{3}{4}-\imath\frac{t+1}{2}-ks}\Gamma(\frac{3}{4} + \imath\frac{t+1}{2} + ks, x(j)\Delta),$$

where Δ is a size adjusting function defined to be,

$$\Delta(j, ks) = \left(\frac{\frac{e+1}{e-1}^{j-1}}{2e}\right)^{ks}.$$

With G, we computed our approximation function as described in sections (2) thru (3).

5.2. Accuracy of Approximation. In order to compare the two functions, we computed $G - P_G$ with ks and j both evalated from

	A	B
0	-0.000168136 - 0.000072861 $\imath$	-0.00549956 - 0.11402 $\imath$
1	-0.584251 + 9.88927 $\imath$ + t	108.54780362440852 + 23.94864969654435 $\imath$
2	3.6913 + 18.7953 $\imath$	-174.187 - 7.25584 $\imath$
3	-1.49351 + 8.77926 $\imath$ + t	

TABLE 1. Values for Continued Fraction Form for $ks =$

$j = 0$

$1 - 10$. Here is the approximation function, P_G,

$$P_G = A_0 + \cfrac{B_1}{A_1 + t + \cfrac{B_2}{A_2 + \cfrac{B_3}{A_3 + t} + t}}.$$

As an example, the values for the A's and B'S for $ks = j = 0$ are,

Table 2 gives the first 20 error values.

As you can see in Table 2, the largest error is 1.3213654590131134^{-10} and the smallest error is 3.63178×10^{-228}. Both are excellent. This shows our approximation is suitable for use in the calculation of $L(s, \chi)$, as wanted.

Figures 1,2, and 3 are graphs of the error for $ks = 0, j = 0, ks = 4, j = 4$,and $ks = 6, j = 6$, respectively. (There are too many graphs

$G - P_G$	ks	$x(j)$
$1.3213654590131134\times10^{-10}$	1	1.
$1.0780712615495542\times10^{-11}$	1	1.5819767068693265
$1.5009813820008643\times10^{-13}$	1	2.8413471884155848
$7.129576684816231\times10^{-16}$	1	5.566566241119301
$8.606792222723689\times10^{-21}$	1	11.46381331340313
$2.26204650672183\times10^{-26}$	1	24.225181247131992
$1.6171117807222558\times10^{-38}$	1	51.84018695129955
$1.7814610888105603\times10^{-64}$	1	111.59777281524529
$8.810895685562281\times10^{-121}$	1	240.91040474231139
$2.966490838203838\times10^{-242}$	1	520.736916040416
$9.959524662413049\times10^{-11}$	2	1.
$1.285544268172432\times10^{-11}$	2	1.58197670
$2.656040051446266\times10^{-13}$	2	2.8413471884155848
$5.307416259662594\times10^{-16}$	2	5.566566241119301
$1.9020525855962412\times10^{-20}$	2	11.46381331340313
$7.771235466954762\times10^{-26}$	2	24.225181247131992
$2.351117221630487\times10^{-37}$	2	51.84018695129955
$5.521622124731802\times10^{-63}$	2	111.59777281524529

TABLE 2. First 20 Error Values for $G - P_G$

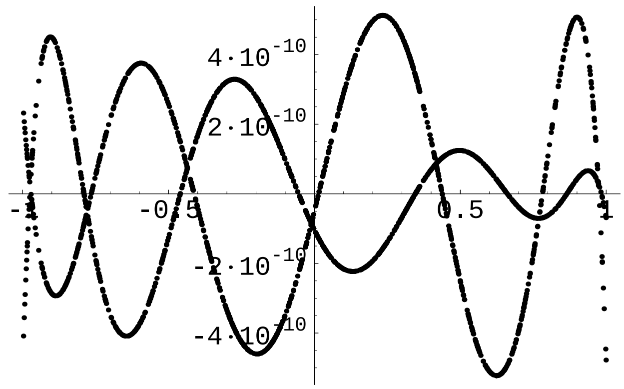

FIGURE 1. Graph of $G - P_G$ with $ks = 0, j = 0$

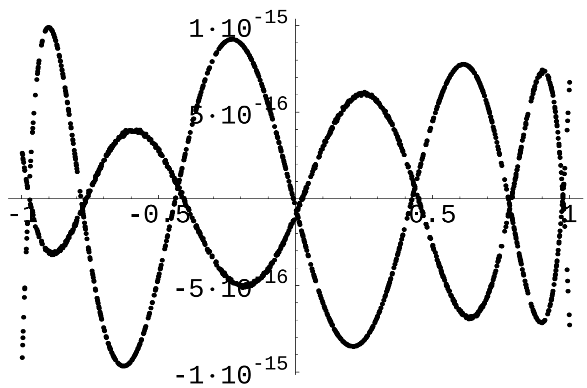

FIGURE 2. Graph of $G - P_G$ with $ks = 4, j = 4$

to include all of them). Notice the error gets smaller as ks and j increase.

Again, these show an extremely close approximation of G has been achieved. Thus, our method for approximation with Chebyshev polynomials, Pade approximants, and continued fractions has

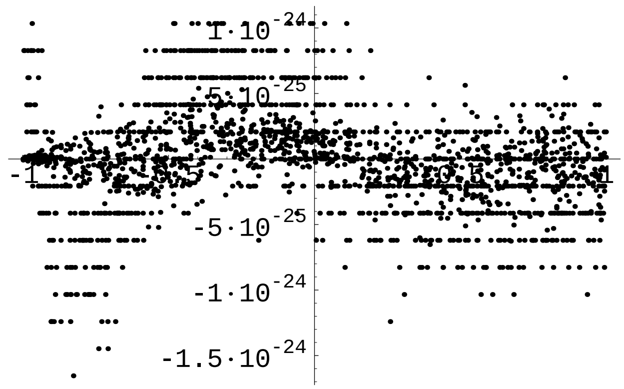

FIGURE 3. Graph of $G - P_G$ with $ks = 6, j = 6$

proved to be successful, and will be applied in calculations of the Dirichlet L-function.

References

[1] Martin Avery Snyder (1996), *Chebyshev Method in Numerical Approximation*, Prentice Hall, New Jersey.

[2] Sheldon Axler (1997), *Linear Algebra Done Right* Springer, New York.

[3] (1988-1992), *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, United Kingdom.

Appendix A. Advantage of Chebyshev extrema

The Chebyshev polynomials have *all* of their maxima at $T_n(x) = 1$ and *all* of their minima at $T_n(x) = -1$. This appendix will explain *why* this is convienant and makes Chebyshev polynomials an ideal tool for approximation.

An infinite number of Chebyshev polynomials means there is an infinite vector space in which they form a basis. Then if there is a function we want to approximate, call it $f(x)$, it can be defined as,

$$f(x) = \sum_{k=0}^{\infty} c_k T_k(x),$$

where the c_k's are the coefficients of the approximating Chebyshev polynomial with,

$$c_k = \frac{\langle f, T_k \rangle}{\langle T_k, T_k \rangle}.$$

With $m(x)$ as our approximating function, we see,

$$m(x) = \sum_{k=0}^{N} c_k T_k(x),$$

where the error in this approximationis the difference between the two series:

$$f(x) - m(x) = \sum_{k=N+1}^{\infty} c_k T_k(x).$$

Then this implies,

$$|f(x) - m(x)| \leq \sum_{k=N+1}^{\infty} |c_k||T_k(x)|.$$

But, since all Chebyshev polynomials, and in particular all $T_k(x)$ for $k = N+1$ up to ∞ are bounded by 1, this means,

$$|f(x) - m(x)| \leq \sum_{k=N+1}^{\infty} |c_k|,$$

where $\sum_{k=N+1}^{\infty} |c_k|$ is just a scalar. This makes Chebyshev polynomials efficient tools for approximating functions since it allows us to easily keep track of and record error in our approximations.