JOURNAL
de Théorie des Nombres
de BORDEAUX
anciennement Séminaire de Théorie des Nombres de Bordeaux

Henri COHEN

Computing *L*-Functions: A Survey

Tome 27, n° 3 (2015), p. 699-726.

<http://jtnb.cedram.org/item?id=JTNB_2015__27_3_699_0>

Journal de Théorie des Nombres
de Bordeaux **27** (2015), 699–726

Computing *L*-Functions: A Survey

par Henri COHEN

Résumé. Nous donnons un certain nombre de methodes pour le calcul de fonctions L, y compris de degré strictement plus grand que 2. Nous décrivons le calcul des coefficients de Dirichlet en utilisant une grande variété de méthodes, par exemple l'utilisation de la formule de Gross–Koblitz p-adique, et le calcul de transformées de Mellin inverse et de fonctions gamma incomplètes généralisées. Nous mentionnons ensuite l'utilisation des équations fonctionnelles approchées lissées, et les "formules explicites". Enfin, nous donnons un aperçu de la théorie relativement récente des motifs hypergéométriques, qui permettent de créer des fonctions L de grand degré de manière élémentaire.

Nous donnons aussi une brève liste des logiciels disponibles, y compris certains qui sont capable de détecter heuristiquement l'existence même de fonctions L ne connaissant que leurs facteurs gamma et leur conducteur. Comme applications, nous mentionnons en particulier la conjecture paramodulaire de Brumer–Kramer, et les gros calculs de formes de Maass sur $\mathrm{SL}_n(\mathbb{Z})$ pour $n = 2$, 3 et 4 effectués par Farmer et al.

Abstract. We survey a number of techniques for computing L-functions, including those of degree larger than 2. We discuss the computation of the Dirichlet coefficients using quite a variety of methods, for instance using the p-adic Gross–Koblitz formula, and the computation of inverse Mellin transforms and of generalized incomplete gamma functions. We then explain the use of smoothed approximate functional equations and of the so-called "explicit formulas". Finally, we discuss the recent and exciting topic of hypergeometric motives, which allows to create L-functions of high degree in an elementary way.

We also mention the available software, including some which can detect heuristically the sheer existence of L-functions knowing only their gamma factors and conductor. As applications, we mention in particular the paramodular conjecture of Brumer–Kramer,

Manuscrit reçu le 19 janvier 2014, révisé le 14 novembre 2014, accepté le 21 novembre 2014.
Mathematics Subject Classification. 11-04, 11Fxx, 11G40, 11Y35.
Mots-clefs. L-functions, inverse Mellin transforms, approximate functional equation, explicit formulas, Gross–Koblitz formula.

and the large scale computations of Maass cusp forms for $\mathrm{SL}_n(\mathbb{Z})$ for $n = 2$, 3, and 4 done by Farmer et al.

1. Definitions, Basic Examples, and Goals

The goal of this paper is to survey the remarkable progress done in the past ten years on the computational aspects of L-functions, in particular in higher degree.

1.1. Definitions. To set the stage, we need of course first to *define* the L-functions that we study.

As a first approximation, we could say that an L-function is a Dirichlet series $\sum_{n\geq 1} a(n)n^{-s}$ which converges for $\Re(s)$ sufficiently large, which can be extended to the whole complex plane into a meromorphic function having a finite number of poles with a functional equation of a specific type, and which has an Euler product.

In fact, A. Selberg introduced a class which gives precise conditions to have a nice theory. Modifying and restricting Selberg's definition we set the following. First we set

$$\Gamma_{\mathbb{R}}(s) = \pi^{-s/2}\Gamma(s/2)\ , \quad \Gamma_{\mathbb{C}}(s) = \Gamma_{\mathbb{R}}(s)\Gamma_{\mathbb{R}}(s+1) = 2\cdot(2\pi)^{-s}\Gamma(s)\ ,$$

and we make the following definition:

Definition. Let d be a nonnegative integer. We say that a Dirichlet series $L(s) = \sum_{n\geq 1} a(n)n^{-s}$ with $a(1) = 1$ is an L-function of *degree* d if the following conditions are satisfied:

(1) (Ramanujan bound): we have $a(n) = O(n^{\varepsilon})$ for all $\varepsilon > 0$, so that in particular the Dirichlet series converges absolutely and uniformly in any half plane $\Re(s) \geq \sigma > 1$.

(2) (Meromorphy and Functional equation): The function $L(s)$ can be extended to $\mathbb{C}$ to a meromorphic function of order 1 having a finite number of poles; furthermore there exist complex numbers λ_i with nonnegative real part and a positive real number N (usually an integer) such that if we set

$$\gamma(s) = N^{s/2}\prod_{1\leq i\leq d}\Gamma_{\mathbb{R}}(s+\lambda_i) \quad \text{and} \quad \Lambda(s) = \gamma(s)L(s)\ ,$$

we have the *functional equation*

$$\Lambda(s) = \omega\overline{\Lambda(1-\overline{s})}$$

for some complex number ω, called the *root number*, which will necessarily be of modulus 1.

(3) (Euler Product): For $\Re(s) > 1$ we have $L(s) = \prod_p L_p(s)$, where the product is over all prime numbers, with $L_p(s) = \prod_{1\leq j\leq d}(1-\alpha_{p,j}p^{-s})^{-1}$, where some (or all) the $\alpha_{p,j}$ may be zero; the $\alpha_{p,j}$ are

called the *Satake parameters*, and one requires the additional technical but necessary condition that there exists $\theta < 1/2$ such that $|\alpha_{p,j}| = O(p^{\theta})$.

Remarks.

(1) Since we want conjecturally our L-functions to satisfy the Riemann hypothesis, it is easy to show that one must impose conditions on the λ_i and the Satake parameters, for instance the ones given.
(2) We restrict to gamma factors of the above form, but more generally the Selberg class allows $\Gamma(\mu_i s+\lambda_i)$ with μ_i positive real in the gamma factors.
(3) Note that d is *both* the number of $\Gamma_{\mathbb{R}}$ factors, *and* the degree in p^{-s} of the Euler factors $L_p(s)^{-1}$, at least generically (the degree may decrease for "bad" primes p).
(4) Most L-function have functional equations of the form

$$\Lambda(s) = \omega\overline{\Lambda(k-\overline{s})}$$

for some $k > 0$, usually an integer. The above normalization, which is standard in analytic number theory, amounts to replacing $a(n)$ by $a(n)/n^{(k-1)/2}$, or equivalently $L(s)$ by $L_1(s) = L(s+(k-1)/2)$ which does satisfy $\Lambda_1(s) = \omega\overline{\Lambda_1}(1-\overline{s})$.
(5) In *practice* N will be an integer, called the *conductor*, and the Satake parameters will satisfy $|\alpha_{p,j}| = 1$ for $p \nmid N$, and $\alpha_{p,j} = 0$ or $|\alpha_{p,j}| = p^{-m/2}$ for some nonnegative integer m when $p \mid N$.

1.2. Examples. The only L-function of degree 0 is the constant 1. In what follows, we give the most important examples, with mention of their computability. • The basic example of an L-function is of course Riemann's zeta function $\zeta(s) = \sum_{n\geq 1} n^{-s}$, with $d = 1$, $\lambda_1 = 0$, $N = 1$, $\omega = 1$. Extremely easy to compute.

• More generally we have the Dirichlet L-series $L(\chi, s) = \sum_{n\geq 1} \chi(n)n^{-s}$, where χ is a Dirichlet character. Also extremely easy to compute. One can show that these are the only L-functions of degree 1.

• If E is an elliptic curve, the L-function $L(E, s)$ has degree 2. More generally, if f is a Hecke eigenform of integral weight k the L-function $L(f, s)$ also has degree 2 and satisfies a functional equation $s \mapsto k - s$. One can do the same for *Maass forms*. The L-functions of hololomorphic modular forms are very easy to compute, those of Maass forms a little more difficult because of the necessity of also computing the *spectral parameters* λ_i.

• More generally, one can associate L-functions to higher degree modular forms (for example Siegel modular forms), and even more generally to *automorphic representations*. Although explicit, these are *much more* difficult to compute. A very special (but easily computable) case is that of

symmetric powers of modular forms, for which many computations have been done.

• To any algebraic variety defined, say, over $\mathbb{Q}$, one associates local (rational) L-factors à la Weil, and a global L-function by replacing the formal variable T by p^{-s} and taking the product over all p. In practice, one only takes parts of the cohomology of the variety, thus obtaining the L-function of a *motive*. The Euler factors at the "good" primes are relatively easy to compute, but this is not at all the case for the bad primes, and similarly the exponents of the bad primes in the "conductor", essentially the constant N, are not easy to compute.

• N. Katz and more recently F. Rodriguez-Villegas introduced the notion of *hypergeometric motive*, which is a special case of the above, but with the advantage that, at least again the Euler factors at the good primes are now extremely easy to compute. This is the subject of active study, and we will say more about this subject later.

1.3. Goals. We must now define what we mean by "computing L-functions". This involves many different aspects, all interesting in their own right.

We first assume that we are "given" the L-function, in other words that we are given an "efficient" algorithm to compute the $a(n)$ (or the Euler factors), and that we know the gamma factor $\gamma(s)$. The main computational goals are then the following:

(1) Compute $L(s)$ for "reasonable" s: example, compute $\zeta(3)$. More sophisticated, but much more interesting: compute special values of symmetric powers L-functions of modular forms, and check numerically the conjectures of Deligne on the subject.
(2) Check the numerical validity of the functional equation, and in passing, if unknown, compute the numerical value of the *root number* ω occurring in the functional equation.
(3) Compute $L(s)$ for $s = 1/2 + it$ for rather large real values of t, and/or make a plot of the corresponding Z function (see below).
(4) Compute all the zeros of $L(s)$ on the critical line up to a given height, and check the Riemann hypothesis.
(5) Compute the residue of $L(s)$ at $s = 1$ (typically): for instance if L is the Dedekind zeta function of a number field, this gives the product hR.
(6) Compute the *order* of the zero of $L(s)$ at $s = 1/2$ (if it has one), and the leading term in the Taylor expansion: for instance this gives the *analytic rank* of an elliptic curve, together with the Birch and Swinnerton-Dyer data.

Unfortunately, we are not always given an L-function completely explicitly. We can lack more or less partial information on the L-function:

(1) One of the most frequent situations is that one knows the Euler factors for the "good" primes, as well as the corresponding part of the conductor, and that one is lacking both the Euler factors for the bad primes and the bad part of the conductor. The goal is then to find numerically the missing factors and missing parts.

(2) A more difficult but much more interesting problem is when essentially nothing is known on the L-function except $\gamma(s)$, in other words the $\Gamma_{\mathbb{R}}$ factors and the constant N, essentially equal to the conductor. It is quite amazing that nonetheless one can quite often tell whether an L-function with the given data can exist, and give some of the initial Dirichlet coefficients (even when several L-functions may be possible).

(3) Even more difficult is when essentially nothing is known except the degree d and the constant N, and one looks for possible $\Gamma_{\mathbb{R}}$ factors: this is the case in the search for Maass forms over $\mathrm{SL}_n(\mathbb{Z})$, which has been conducted very successfully for $n = 2$, 3, and 4.

1.4. Software. Many people working on the subject have their own software. I mention the available public data.[1]

• M. Rubinstein's `C++` program `lcalc`, which can compute values of L-functions, make large tables of zeros, and so on. The program uses `C++` language `double`, so is limited to 15 decimal digits, but is highly optimized, hence very fast, and used in most situations. Also optimized for large values of the imaginary part using the Riemann–Siegel formula. Available in `Sage`.

• T. Dokshitser's program `computel`, initally written in `GP/Pari`, rewritten for `magma`, and also available in `Sage`. Similar to Rubinstein's, but allows arbitrary precision, hence slower, and has no built-in zero finder, although this is not too difficult to write. Not optimized for large imaginary parts. The details of this package are explained in his paper [6].

• Last but not least, not a program but a huge *database* of L-functions, modular forms, number fields, etc. . . , which is the result of a collaborative effort of approximately 30 to 40 people headed by D. Farmer. This database can of course be queried in many different ways, it is possible and useful to navigate between related pages, and it also contains `knowls`, bits of knowledge which give the main definitions. In addition to the stored data, the site can compute on the fly (using the software mentioned above, essentially `Pari`, `Sage`, and `lcalc`) additional required information. Available at `http://www.lmfdb.org`

[1] Note added in proof: thanks to work of B. Allombert, K. Belabas, P. Molin, and the author, in the development version of `Pari/GP` there is now available a large and efficient package for L-function computation

2. Elementary Results, Sophisticated Analysis

2.1. More Examples. Evidently, many interesting results have and continue to be found. Before mentioning new results, let me explain some results that can be obtained using more sophisticated programs, but using *only* analytical methods. The problem is this: given a product of $\Gamma_{\mathbb{R}}$ factors $G(s) = \prod_{1\le j\le d}\Gamma_{\mathbb{R}}(s+\lambda_j)$, find *integers* N such that $\gamma(s) = N^{s/2}G(s)$ is the gamma factor of an L-function with *integral* coefficients (note that this is not part of the definition of an L-function). If the search limit for N is small enough (say $N \le 200$) and d is small, this can usually be done.

Note that if $L_1(s)$ and $L_2(s)$ are L-functions with corresponding gamma factors $G_i(s)$, the function $L_1(s)L_2(s)$ is also an L-function with gamma factor $G_1(s)G_2(s)$. An L-function which cannot be written nontrivially as L_1L_2 will be called *primitive*, and evidently we may restrict to primitive L-functions.

In what follows, we give a few examples of *primitive* L-functions with *integer* coefficients.

(1) $G(s) = \Gamma_{\mathbb{R}}(s)$. Possible N, and at most unique: $N = 1$, 5, 8, 12, 13, 17, etc... These exist and correspond to the L-function associated to an *even primitive real* Dirichlet character. Equivalently, the nonprimitive L-function $\zeta(s)L(s)$ is the Dedekind zeta function of a real quadratic field when $N > 1$.

(2) $G(s) = \Gamma_{\mathbb{R}}(s+1)$. Possible N, and at most unique: $N = 3$, 4, 7, 8, 11, 15, etc... These exist and correspond to the L-function associated to an *odd primitive real* Dirichlet character. Equivalently, the nonprimitive L-function $\zeta(s)L(s)$ is the Dedekind zeta function of an imaginary quadratic field.

(3) $G(s) = \Gamma_{\mathbb{R}}(s)^2$. Possible N: 49, 81, 148, 169, etc... These correspond to Artin L-functions $\zeta_K(s)/\zeta(s)$, where $\zeta_K(s)$ is the Dedekind zeta function of a totally real cubic field.

(4) $G(s) = \Gamma_{\mathbb{R}}(s)\Gamma_{\mathbb{R}}(s+1) = \Gamma_{\mathbb{C}}(s)$. Possible N: $N = 23$, 31, 44, 59, 76, etc... These correspond to to Artin L-functions $\zeta_K(s)/\zeta(s)$, where $\zeta_K(s)$ is the Dedekind zeta function of a complex cubic field. Other values of N are $N = 39$, 47 (two L-functions), 55, etc..., and correspond to binary theta series which are Hecke eigenforms.

(5) $G(s) = \Gamma_{\mathbb{R}}(s+1/2)\Gamma_{\mathbb{R}}(s+3/2) = \Gamma_{\mathbb{C}}(s+1/2)$. Possible N: $N = 11$, 14, 15, 17, 19, 20, 21, 24, 26 (two L-functions), 27, 30, etc... These L-functions do exist, and correspond to elliptic curves defined over $\mathbb{Q}$, or equivalently by Wiles, to modular eigenforms of weight 2.

More generally, $G(s) = \Gamma_{\mathbb{R}}(s+(k-1)/2)\Gamma_{\mathbb{R}}(s+(k+1)/2) = \Gamma_{\mathbb{C}}(s+(k-1)/2)$ correspond to modular eigenforms of weight k.

(6) $G(s) = \Gamma_{\mathbb{R}}(s)\Gamma_{\mathbb{R}}(s+1)^2$. These correspond to symmetric squares of modular forms of weight 1, which in small levels are binary theta series.

(7) $G(s) = \Gamma_{\mathbb{R}}(s+1)^2\Gamma_{\mathbb{R}}(s+2)$. These correspond to symmetric squares of modular forms of weight 2.

(8) $G(s) = \Gamma_{\mathbb{R}}(s+1/2)^2\Gamma_{\mathbb{R}}(s+3/2)^2 = \Gamma_{\mathbb{C}}(s+1/2)^2$. These correspond conjecturally to the global L-function of curves of genus 2, their jacobians, and more generally to abelian surfaces defined over $\mathbb{Q}$.

(9) $G(s) = \Gamma_{\mathbb{R}}(s+1/2)\Gamma_{\mathbb{R}}(s+3/2)\Gamma_{\mathbb{R}}(s+k-3/2)\Gamma_{\mathbb{R}}(s+k-1/2) = \Gamma_{\mathbb{C}}(s+1/2)\Gamma_{\mathbb{C}}(s+k-3/2)$. These correspond to the *spinor L-functions* of Siegel modular forms of degree 2 and weight k. In particular, for $k = 3$ this is the same gamma factor as the one of hypergeometric motives of degree 4 and motivic weight 2, a special case coming from part of the cohomology of the Dwork quintic pencil.

We could of course continue at will. Evidently none of the above results are new, but it is interesting to note that the possible L-functions for given gamma factors can be found experimentally by using (rather sophisticated) purely analytic techniques, with no input whatsoever from number theory or algebraic geometry.

2.2. New Results: I, Maass forms. I now mention what I consider to be the two most spectacular results (more precisely conjectures) that have been obtained using these methods, and a third, less spectacular but still quite interesting.

I heartily thank David Farmer for helping me write this section, which is essentially a paraphrase of a text that he sent me. I refer to [7] for technical details.

The computation of Maass forms for $\mathrm{SL}_2(\mathbb{Z})$ and subgroups is a priori not easy. However, using analytic (as opposed to algebraic or geometric) methods, H. Stark showed that it could in fact be done rather easily, and this was pursued very successfully by D. Hejhal and others.

On the other hand, very few results were known for subgroups of $\mathrm{SL}_J(\mathbb{Z})$ for $J \geq 3$. Farmer, Koutsoliotas, and Lemurell have obtained remarkable (numerical) results for the case $J = 3$ and $J = 4$, more precisely for small index subgroups of $\mathrm{SL}_3(\mathbb{Z})$ and for $\mathrm{SL}_4(\mathbb{Z})$ and $\mathrm{Sp}_4(\mathbb{Z})$. Let me briefly explain their results and sketch their methods.

Their approach (which in fact is the main theme of this paper) is to completely ignore the underlying Maass form and focus on the L-functions.

These L-functions satisfy a functional equation of the form

$$\Lambda(s) := N^{s/2} \prod_{j=1}^{J} \Gamma_{\mathbb{R}}(s+\delta_j+ir_j)L(s) = \overline{\Lambda}(1-s)\,, \tag{2.1}$$

where $\delta_j = 0$ or 1, $\sum_{1\le j\le J} r_j = 0$, and $J = 3$ or 4. The analogue of the Selberg eigenvalue conjecture is the assertion that the r_j are real, which is assumed in their calculations. The integer N is the "level", which is 1 when dealing with $\mathrm{SL}_3(\mathbb{Z})$ or $\mathrm{SL}_4(\mathbb{Z})$; if $N > 1$ we are dealing with a congruence subgroup. The case of $\mathrm{Sp}_4(\mathbb{Z})$ refers to $J = 4$ where the spectral parameters come in conjugate pairs: $\{r_1, r_2, r_3, r_4\} = \{\lambda_1, -\lambda_1, \lambda_2, -\lambda_2\}$, and the Dirichlet coefficients are also real.

Thus, once the degree J, the level N, and the shifts δ_j are chosen, one must find parameters r_j and Dirichlet coefficients such that the L-function satisfies (2.1). This is of course much too weak a condition, so as explained above, one must also add the conditions that the L-function must have an Euler product, and the Dirichlet coefficients must satisfy the Ramanujan bound.

The method, which is briefly described in the discussion of the approximate functional equation in Section 6, involves choosing two test functions in the approximate functional equation. The equality of the (purported) L-values is interpreted as a linear equation in the unknown Dirichlet series coefficients. By making several such choices one obtains a linear system of equations, which the coefficients of the L-function (if it exists) must satisfy. One then adds to this linear system nonlinear equations coming from the existence of the Euler product, such as $a(6) = a(2)a(3)$. Even so, the resulting system will generally still have a large number of solutions, most of which are extraneous. However these extraneous solutions can usually be eliminated because they do not satisfy the Ramanujan bound.

Note that an extra difficulty in this problem (as was the case for $\mathrm{SL}_2(\mathbb{Z})$) is that one must not only search for the Dirichlet coefficients, but also for the spectral parameters r_j.

Using this method approximately 2000 L-functions of Maass forms have been found for $\mathrm{SL}_3(\mathbb{Z})$, with a few dozen on subgroups up to level 9. A few dozen have also been found for $\mathrm{SL}_4(\mathbb{Z})$, and a few hundred for $\mathrm{Sp}_4(\mathbb{Z})$. These data are available in the LMFDB mentioned above.

It is of course highly plausible that corresponding Maass forms exist in all the cases found, but as far as I know no proofs have been given for $J \ge 3$.

2.3. New Results: II, the paramodular conjecture. Since L-functions seem to be quite well determined (at least up to finite dimensionality) by their gamma factor $\gamma(s) = N^{s/2}G(s)$, in other words by the gamma product $G(s)$ and the conductor N, it is very tempting to guess, or even conjecture, that if two L-functions have the same $\gamma(s)$, and perhaps also the first few Dirichlet coefficients, then they will in fact be equal. Let us pretend that we are ignorant for the moment.

We have seen that the (normalized) gamma factor of a modular eigenform of weight 2 on $\Gamma_0(N)$ is $N^{s/2}\Gamma_{\mathbb{R}}(s+1/2)\Gamma_{\mathbb{R}}(s+3/2) = N^{s/2}\Gamma_{\mathbb{C}}(s+1/2)$. On the other hand, if E is an elliptic curve defined over $\mathbb{Q}$ of conductor N with complex multiplication, its L-function is a product of two Hecke L-functions and it is easy to see that it will have the same gamma factor. This leads us to believe that it is also the L-function of a modular eigenform of weight 2 on $\Gamma_0(N)$. This was indeed proved by Shimura, and this observation led Taniyama–Shimura–Weil to formulate the conjecture, finally proved by Wiles and successors, that it should hold true for *all* elliptic curves over $\mathbb{Q}$, not only those with complex multiplication.

We can now do the same for abelian surfaces defined over $\mathbb{Q}$. Among the uninteresting ones are products of two elliptic curves, which by the modularity theorem thus have an L-function with gamma factor $N^{s/2}\Gamma_{\mathbb{C}}(s+1/2)^2$, where N is the conductor of the abelian surface, in this particular case equal to the product of the conductors of the elliptic curve factors. As for elliptic curves, it is reasonable to expect that for *all* abelian surfaces defined over $\mathbb{Q}$ the L-function extends to a holomorphic function on the whole plane with functional equation having as gamma factor $N^{s/2}\Gamma_{\mathbb{C}}(s+1/2)^2$. This is what we have indicated in the table given in Section 2.1. Now if we look at the end of this table, we note that the spinor L-function of Siegel modular forms of degree 2 and weight k have $G(s) = \Gamma_{\mathbb{C}}(s+(k-1)/2)^2$: this coincides with the one for abelian surfaces when $k = 2$. We can thus suspect that there is a conjectural correspondence between such surfaces and Siegel modular forms of degree 2 and weight 2, modular with respect to a suitable subgroup of $\mathrm{Sp}_4(\mathbb{R})$ linked with the conductor N of the abelian surface.

In a remarkable work, Brumer–Kramer, helped by essential computations of Poor–Yuen, have made the above observation into a very precise conjecture, which generalizes the Taniyama–Shimura–Weil conjecture. It would take a complete talk in itself to explain in detail the conjecture and the numerical verifications that have been done, but I will try to give some insight. I heartily thank Armand Brumer for help in writing this section.

First, on which subgroup of $\mathrm{Sp}_4(\mathbb{R})$ should the forms be modular? Let us look again at the simpler case of elliptic curves. The subgroup $\Gamma_0(N)$ enters the picture naturally because $Y_0(N) = \mathbb{H}/\Gamma_0(N)$ is the moduli space of elliptic curves together with a cyclic subgroup of order N, and the modularity conjecture is equivalent to the existence of a morphism from the compactification $X_0(N)$ to the elliptic curve.

For abelian surfaces, there exists a group $K(N)$ called the *paramodular group* of level N, such that $\mathbb{H}_2/K(N)$ is the moduli space of abelian surfaces with polarization $(1, N)$, where as usual $\mathbb{H}_2$ is the Siegel upper half space. This group can of course be explicitly described: it is *not* a subgroup of

$\mathrm{Sp}_4(\mathbb{Z})$, but of $\mathrm{Sp}_4(\mathbb{Q})$ (exactly one coefficient out of 16 may be nonintegral). One defines $K(N) = \gamma M_4(\mathbb{Z})\gamma^{-1} \cap \mathrm{Sp}_4(\mathbb{Q})$, where γ is the diagonal matrix with diagonal coefficients $(1, 1, N, 1)$. It is immediate to show that

$$K(N) = \left\{ g \in \mathrm{Sp}_4(\mathbb{Q}),\ g = \begin{pmatrix} * & * & */N & * \\ N* & * & * & * \\ N* & N* & * & N* \\ N* & * & * & * \end{pmatrix} \right\},$$

where the $*$ denote integers.

Second, to obtain an exact bijection between the two sides, one needs on both sides to exclude certain elements. On the abelian surface side, we must restrict to (isogeny classes of) abelian surfaces A with trivial endomorphism group, i.e., $\mathrm{End}_{\mathbb{Q}}(A) = \mathbb{Z}$. On the modular side, we must exclude certain Siegel modular forms on $K(N)$ which can be obtained by lifting certain Jacobi forms by using what is called a *Gritsenko lift*. This is not difficult to define, but I will not do it here. The paramodular conjecture of Brumer–Kramer is then as follows (see [3]):

Paramodular Conjecture. Paramodular Conjecture. There is a bijection between on the one hand:

- Isogeny classes of abelian surfaces A defined over $\mathbb{Q}$ with conductor N and $\mathrm{End}_{\mathbb{Q}}(A) = \mathbb{Z}$, and on the other hand:
- Lines of non-lift weight 2 Hecke cusp newforms F on $K(N)$ with rational eigenvalues.

In this correspondence, the Hasse–Weil L function of the abelian surface A should be equal to the spinor L-function attached to the Siegel cusp form F, and the ℓ-adic Galois representation attached to F should be isomorphic to that on the Tate module $T_\ell(A)$.

Note that Hecke cusp newforms on $K(N)$ cannot be normalized by setting $a(1) = 1$ as one does for ordinary newforms, and this is why one must consider "lines" of such newforms, i.e., up to a multiplicative constant.

Note also that any abelian surface A defined over $\mathbb{Q}$ and *not* covered by the above conjecture is of "GL_2-type", and so analogously to the case of elliptic curves with CM its L-series is a product of L-series attached to classical elliptic modular forms. This can be shown using work of Shimura, Ribet, and Khare–Wintemberger.

To be able to test this conjecture, it is of course necessary to gather evidence on *both* sides, and both are difficult.

• On the abelian side, Brumer and Kramer developped sophisticated methods originating with Fontaine and Schoof, using group schemes and class field theory to exclude certain conductors. On the other hand, they

had to construct sufficiently many examples of abelian surfaces of small conductor, not necessarily isogenous to jacobians.

The upshot was that for prime conductors $p \leq 600$, there exists an abelian surface of conductor p if and only if

$$p \in \{277, 349, 353, 389, 461, 523, 587\} .$$

• The modular side, due to Poor–Yuen [10] is also quite hard. The problem is that the explicit computation of $S_2(K(N))$ is very difficult because 2 is a "small weight". This is more or less analogous to the difficulty of computing ordinary modular cusp forms of weight 1. Instead, they begin by computing in $S_4(K(N))$ which is quite explicit, and using quite a number of tricks they manage at least to construct a non-lift for $p = 277$, and to check that the corresponding first Fourier coefficients agree. Since their initial paper, other methods have been used, for instance using Borcherds products, to construct non-lifts, so it can be hoped that the conjecture can be checked for many more N and a large number of Euler factors.

As a final note, the paramodular conjecture has been proved by Johnson–Leung and Roberts in the special case where $A/\mathbb{Q}$ is the Weil restriction of a modular elliptic curve E/k not isogenous to its conjugate, where k is a real quadratic field, using a lift from Hilbert modular forms to paramodular forms.

2.4. New Results: III, Hypergeometric Motives. We will mention hypergeometric motives later. The recent impetus on the subject started with the study by P. Candelas, X. de la Ossa, F. Rodriguez–Villegas, the author, and others, of the L-function associated to part of the cohomology of the *Dwork quintic pencil*

$$x_1^5 + x_2^5 + x_3^5 + x_4^5 + x_5^5 - 5\psi x_1 x_2 x_3 x_4 x_5 = 0$$

for $\psi \neq 0, 1, \infty$.

As for all hypergeometric motives, the computation of the Euler factors at the good primes presents no difficulty and can be done using a recipe due to Katz and Rodriguez–Villegas. Furthermore, since one easily computes the Hodge numbers, the gamma factor is given by a recipe of Serre. Using the functional equation that the L-function must satisfy, we have been able to find (initially conjecturally) recipes for all the Euler factors (thus at all the bad primes), for the conductor, and for the root number. I have been told, but cannot give a reference, that thanks to recent work these experimental observations are now in principle *proved*.

Once again, a comparison of the gamma factors shows that the L-functions should be the spinor L-function of a Siegel modular form of degree 2 and weight 3. The smallest possible conductor corresponding to the quintic is 525, which seems "small", but nonetheless Siegel form experts consider

that proving the existence of a form of degree 2, weight 3, and level 525 is completely out of reach for the moment.

3. Number-Theoretical Tools

To achieve the above goals, we need some tools which, although for the most part technical in nature, are of course absolutely essential. The first kind of tools are number-theoretical, for the computation of $a(n)$, and the second kind are purely (real or complex) analytic. To compute $a(n)$, we have at our disposal (at least) four rather different methods:

(1) The "elementary" methods: naive or baby-step giant step point counting on varieties, or direct computation of the coefficients of q-expansions for modular forms for instance.
(2) Expression of $a(n)$ in terms of *Gauss sums* or *Jacobi sums* and the use of theta functions to compute them.
(3) ℓ-adic methods: this is linked to *étale cohomology* and is the basic idea behind *Schoof's algorithm* for counting points on elliptic curves over finite fields of large characteristic, and much more recently on the Edixhoven–Couveignes method for computing coefficients of cusp forms.
(4) p-adic methods: these come in several flavors, more or less linked to *crystalline cohomology*: the Saito–Mestre methods for counting points in small characteristic, Kedlaya's algorithm using Monsky–Washnitzer cohomology for counting points on hyperelliptic curves in small characteristic, the use of Morita's p-adic gamma function and the Gross–Koblitz formula for computing $a(n)$ when expressed as coefficients of a *hypergeometric motive*.

To illustrate, let us take the following specific example: the function $f(\tau) = \eta(2\tau)^2\eta(10\tau)^2$ is a modular Hecke eigenform of weight 2 on $\Gamma_0(20)$ with trivial character. Writing $f(\tau) = \sum_{n\geq 1} a(n)q^n$ with as usual $q = e^{2\pi i\tau}$, we have $L(f,s) = \sum_{n\geq 1} a(n)n^{-s}$, and we need to compute $a(n)$. In fact, since f is an eigenform, it is only necessary to compute $a(p)$, since $a(p^k)$ is given by the recursion $a(p^k) = a(p)a(p^{k-1}) - \chi(p)pa(p^{k-2})$ with $\chi(p) = 1$ unless $p = 2$ or $p = 5$, in which case $\chi(p) = 0$, and $a(mn) = a(m)a(n)$ whenever m and n are coprime.

Thus, let us look at the different methods available to us. Note that we may have two different goals in mind, in particular if there are storage problems: first compute a large *table* of $a(n)$ up to some large bound, and second compute *individual values* of $a(n)$, for n relatively large.

(1) Computing directly from the q-expansion. Recall that $\eta(\tau) = 1 + \sum_{m\geq 1}(-1)^m(q^{(3m^2-m)/2} + q^{(3m^2+m)/2})$ has a completely explicit expansion. Thus, by using FFT multiplication techniques, performing

two multiplications of power series, we can compute the power series expansion of $f(\tau)$ to X terms in time $\widetilde{O}(X)$, where $\widetilde{O}(X^\alpha)$ means $O(X^{\alpha+\varepsilon})$ for all $\varepsilon > 0$. This is good, since it gives an average of $\widetilde{O}(1)$ per coefficient, and one cannot hope to do much better than that. However this is very specific to such products, and more general modular forms may not be computable in this way. In fact we can do better if we notice that $\eta(\tau)\eta(5\tau)$ is the modular form of weight 1 associated to a Hecke character, hence with explicit coefficients, so that only one FFT multiplication will be needed.

(2) A more sophisticated but efficient method, but this time to compute individual values, is to use the Eichler–Selberg *trace formula*. In our case, we have $\dim(S_2(\Gamma_0(20))) = 1$, so $a(p)$ is simply equal to the trace of $T(p)$ acting on that space. Explicitly, we find that

$$a(p) = p - 5 - \frac{1}{2}\sum_{|t|<p^{1/2}} \left(1 + \left(\frac{t^2-p}{5}\right)\right) \times \sum_{f^2|(4(t^2-p))} \alpha(p,f,t)\beta(\gcd(20,f))h'\left(\frac{4(t^2-p)}{f^2}\right),$$

where the first sum is over all positive and negative t (such that $(\frac{t^2-p}{5}) \neq -1$, and one can regroup t and $-t$ when $t \neq 0$), the second sum is over all $f > 0$ such that $4(t^2-p)/f^2 \equiv 0$ or 1 modulo 4, $h'(D) = h(D)/(w(D)/2)$ is the modified class number of the quadratic order of (negative) discriminant D, except that we set $h'(-3) = 1/3$ and $h'(-4) = 1/2$, $\alpha(p,f,t)$ is the number (equal to 0, 1, or 2) of $\varepsilon = \pm 1$ such that $p + 1 + 2\varepsilon t \equiv 0 \pmod{4\gcd(4,f)}$, and finally $\beta(g) = g\prod_{p=2,5,\ p\mid 20/g}(1+1/p)$.

Even though the formula looks complicated, it is immediate to program, and since one can compute class numbers in time $\widetilde{O}(1)$, this gives a $\widetilde{O}(p^{1/2})$ method for computing $a(p)$. Note that this method is much less specific, and can be applied to any modular form of any reasonable weight and level. For instance, to my knowledge it gives the fastest *practical* way to compute individual values of $\tau(p)$, where τ is the Ramanujan tau function. For instance, if we define $H(N) = \sum_{f^2|N} h'(-N/f^2)$, the Hurwitz class number, applying the trace formula to $\mathrm{PSL}_2(\mathbb{Z})$ gives

$$\tau(p) = 42p^6 - 90p^4 - 75p^3 - 35p^2 - 9p - 1 - \sum_{0<s<2p^{1/2}} s^{10}H(4p-s^2),$$

and even better, if we apply the trace formula to $\Gamma_0(2)$ and set $H_2(N) = H(N) + 2H(N/4)$ (where $H(N)$ must be computed from

$H(N/4)$ when both terms occur), we have the faster formula

$$\tau(p) = 28p^6 - 28p^5 - 90p^4 - 35p^3 - 1 - 128 \sum_{0<s<p^{1/2}} s^6(4s^4 - 9ps^2 + 7p^2)H_2(4(p-s^2))$$

(for instance, 8 minutes for $p = 10^{12} + 39$ on a quadcore laptop).

(3) As a special case of *hypergeometric motives* (see below), it is easy to show that for $p \geq 5$ we have

$$a(p) = \frac{1}{1-p}\left(1 + \sum_{\chi \neq \chi_0} \chi(2)J(\chi,\chi,\chi)\right) ,$$

where χ runs over all nontrivial characters of $\mathbb{F}_p^*$ and J is the standard triple Jacobi sum. This expression can be used in many different ways:

• Directly: we compute each individual term and sum over all characters, for a total of approximately p^2 operations. Using the symmetry $\chi \mapsto \overline{\chi}$ reduces this trivially to $p^2/2$ operations.

• Working in $\mathbb{Z}[X]/(X^{p-1} - 1)$. Let g be a primitive root modulo p, and set

$$P_a(X) = \sum_{1 \leq n \leq p-2} X^{(n + a\log_g(1-g^n)) \bmod p-1} \quad \text{and}$$
$$P(X) = P_1(X)P_2(X) \pmod{X^{p-1-1}} .$$

It is not difficult to show that if we write $P(X) = \sum_{0 \leq n \leq p-2} c(n)X^n$ and if we set $\ell = -\log_g(2) \bmod p-1$ with $0 \leq \ell \leq p-2$ we have

$$a(p) = p - 3 - (-1)^{(p^2-1)/8} - c(\ell) .$$

This gives a much faster $O(p^{1+\varepsilon})$ method for computing $a(p)$. The main disadvantage of this method is that we also need $O(p)$ storage, which can become prohibitive.

• Using theta functions: if χ^3 is nontrivial we have $J(\chi,\chi,\chi) = \mathfrak{g}(\chi)^3/\mathfrak{g}(\chi^3)$, where $\mathfrak{g}(\chi)$ is a Gauss sum (and when χ^3 is trivial but χ is not we have $J(\chi,\chi,\chi) = -\mathfrak{g}(\chi)^3/p$). Even though this looks like we are complicating matters, we can compute $\mathfrak{g}(\chi)$ in $O(p^{1/2})$ operations using an idea of Louboutin: indeed, if χ is an even character, the functional equation of the theta function implies that

$$\mathfrak{g}(\chi) = p^{1/2}\frac{\sum_{m \geq 1} \chi(m)\exp(-\pi m^2/p)}{\sum_{m \geq 1} \chi^{-1}(m)\exp(-\pi m^2/p)} ,$$

with a similar formula for χ odd, and since the series converge very fast and $\mathfrak{g}(\chi)$ needs to be known to only a reasonable accuracy (recall that $a(p)$ is an integer), this indeed leads to a $\widetilde{O}(p^{1/2})$ algorithm for

computing $\mathfrak{g}(\chi)$, hence a $\widetilde{O}(p^{3/2})$ algorithm for $a(p)$, so slower than the preceding one, but not requiring much storage.

Note that if the denominator vanishes there exist similar fast formulas. Louboutin conjectures that this never happens, but note that similar quantities for *nonprime* p can vanish, see [5].

• Using a congruence: it is easy to show that

$$a(p) \equiv \sum_{0 \le n \le (p-1)/3} \frac{(3n)!}{2^n n!^3} \pmod{p} ,$$

and on the other hand, the Hecke or Hasse–Weil bounds imply that $|a(p)| \le 2p^{1/2}$, so the congruence determines $a(p)$ as soon as $p \ge 17$. Since the summands can be computed recursively, this again gives a $O(p^{1+\varepsilon})$ method for computing $a(p)$, this time with no storage.

• Using Morita's p-adic gamma function and the Gross–Koblitz formula: (in fancy language, this is a crystalline method). In the present very simple case, it does not bring us much, but in the general case it is the most powerful method available. As for many "natural" congruences, the congruence that we have just given is only the first level of a p-adic *equality* for $a(p)$. Indeed, as before the Jacobi sums which enter into the formula for $a(p)$ can be expressed in terms of Gauss sums, and it is a remarkable result of Gross and Koblitz that *all* Gauss sums over finite fields can be expressed as simple products of values of *Morita's p-adic gamma function* at rational arguments. There is no need for us to define it, simply note that it is very easy to compute, and that, as any p-adic function, it gives congruences modulo any power of p, so that if $a(p) \bmod p$ was not sufficient to determine $a(p)$, we could for instance immediately determine $a(p) \bmod p^2$ for instance. In our example (where we do *not* need the congruence modulo p^2 since $|a(p)| < 2p^{1/2}$) we have for $p \ge 5$:

$$a(p) \equiv \sum_{0 \le n \le (p-1)/3} \frac{(3n)!}{2^{pn} n!^3} (1 + 3pn(H_{3n} - H_n)) - p \sum_{(p-1)/3 < n \le 2(p-1)/3} \frac{(3n-(p-1))!}{2^n n!^3} \pmod{p^2} ,$$

where $H_n = \sum_{1 \le j \le n} 1/j$ is the harmonic sum.

(4) Using elliptic curves: since $f(\tau)$ is a modular form of weight 2 on $\Gamma_0(20)$ with trivial character, it corresponds to an elliptic curve of conductor 20, and in fact up to isogeny there is only one such curve E, with minimal equation $y^2 = x^3 + (x+2)^2$ (an isogenous curve is $y^2 = x^3 + x^2 - x$), so for $p \ge 7$ we have $a(p) = p + 1 - |E(\mathbb{F}_p)|$. Using a remarkable algorithm due to R. Schoof, we can compute $a(p)$ in time

polynomial in $\log(p)$, hence extremely fast. This algorithm, now called the SEA algorithm (S=Schoof, E=Elkies, A=Atkin) is implemented in many packages and is very efficient. It consists in computing $a(p)$ modulo ℓ for sufficiently many small primes ℓ and reconstructing $a(p)$ using the Chinese remainder theorem. This is an ℓ-adic method, as opposed to the p-adic methods seen above, and in our specific case is the fastest available.

For higher weight modular forms, for instance for $\Delta(\tau)$, a similar algorithm has been developed by Edixhoven, Couveignes, et. al., but for now it does not seem to be very practical, although it is a theorem that it has a polynomial running time in $\log(p)$.

4. Inverse Mellin Transforms

We also need a number of more or less sophisticated *analytic* tools. Two types of formulas that are of constant use are on the one hand the *approximate functional equation* (a misnomer since it is not approximate but exact), and *explicit formulas*, a vague name but which refers more precisely to the link via Fourier transforms between primes and zeros of L-functions, as developed by Weil, Stark, Odlyzko, Poitou, Serre, etc... We will state these formulas below, but for now note that for the approximate functional equation we will in particular need to compute *inverse Mellin transforms*.

4.1. Definitions. Recall that if $f(t)$ is a reasonably behaved function, its *Mellin transform* $\mathcal{M}(f)(s)$ is defined by

$$\mathcal{M}(f)(s) = \int_0^\infty t^s f(t)\,\frac{dt}{t}\ .$$

Although by far not the weakest possible conditions, we will require that f be a C^∞ function on $]0,\infty[$, tending exponentially to 0 as $t \to \infty$ in the sense that $|f(t)| = O(e^{-at^b})$ for some $a > 0$ and $b > 0$, and that around 0 we have $f(t) = O(t^{-\varepsilon})$ for all $\varepsilon > 0$. Under these conditions, the integral converges absolutely for all s such that $\Re(s) > 0$, uniformly for $\Re(s) \geq \delta > 0$, so it defines a holomorphic function in $\Re(s) > 0$.

The main result that we will need is the *Mellin inversion formula* (in fact a variant of Fourier inversion), which says the following: if $g(s) = \mathcal{M}(f)(s)$ then for all $\sigma > 0$ and $t > 0$ we have

$$f(t) = \frac{1}{2\pi i}\int_{\Re(s)=\sigma} t^{-s} g(s)\,ds\ .$$

An auxiliary but sometimes useful result is the *convolution formula*: if $g_i(s) = \mathcal{M}(f_i)(s)$ for $i = 1$ and 2, then $g_1 g_2(s) = \mathcal{M}(f_1 * f_2)(s)$, where the convolution $f_1 * f_2$ is given by

$$(f_1 * f_2)(t) = \int_0^\infty f_1(t) f_2(1/t)\,\frac{dt}{t} = \int_1^\infty (f_1(t)f_2(1/t) + f_1(1/t)f_2(t))\,\frac{dt}{t}\ .$$

In view of the approximate functional equation, it will be necessary to compute *many* (sometimes millions) of values of $f(t) = \mathcal{M}^{-1}(g)(t)$, in fact at equally spaced points. A large number of methods have been suggested to do this computation, and we review a few, with no pretense at being exhaustive.

4.2. Simple Cases. First note that in the simplest cases $f(t)$ will be an explicit and easily computable function. Here is a simple table, including some elementary transformations

$g(s)$	$f(t) = \mathcal{M}^{-1}(g)(t)$
$\Gamma(s/2)$	$2\exp(-t^2)$
$\Gamma((s+1)/2)$	$2t\exp(-t^2)$
$\Gamma(s)$	$\exp(-t)$
$\Gamma(s/2)\Gamma((s+1)/2)$	$2\exp(-2t)$
$\Gamma(s/2)^2$	$4K_0(2t)$
$\Gamma(s/2)\Gamma((s+2)/2)$	$4tK_1(2t)$
$\Gamma(s)^2$	$2K_0(2\sqrt{t})$
$\Gamma(s)\Gamma(s+1)$	$2\sqrt{t}K_1(2\sqrt{t})$
$m^s g(s)$	$f(t/m)$
$sg(s)$	$-tf'(t)$
$g_1(s)g_2(s)$	$(f_1 * f_2)(t)$

4.3. Computation for small *t*. All these inverse Mellin transforms have something in common: their computation for small t and large t will be very different. Take the simplest case where $f(t) = \exp(-t)$. When t is not too large we can use the power series $f(t) = \sum_{n\geq 0}(-1)^n t^n/n!$. However when t becomes large enormous cancellation occurs, and it is *not* advised to compute $\exp(-t)$ by its power series, although it is of course possible since it has infinite radius of convergence.

Thus we first consider the computation of $f(t)$ for "small" t. As for $\exp(-t)$, in every case, this leads to a power series (with possible occurences of powers of $\log(t)$) with infinite radius of convergence: we simply shift the line of integration in the formula for the inverse Mellin transform towards $-\infty$ and keep track of the residues that we catch; it is then a simple matter of bookkeeping.

However, the analysis does not stop here. Indeed, we then need to *evaluate* this power series at t (and in fact for a huge number of different t): this seems very straightforward (apply some sort of Horner scheme for instance), but can in fact be considerably improved by using *continued fractions*, and we will see that this is essential for large t.

For now, consider the following simplest example:

Let $f(t) = \exp(-t)$. We compute the power series expansion S of f to 35 terms, say, so that $1/35! < 10^{-40}$, my computer working at a default precision of 38 decimal digits. Note that inverse Mellin transforms will be used *additively* so we need *absolute* and not relative accuracy. Using the truncated power series S for $t = 2$ gives only 30 decimal digits ($2^{35}/35! \approx 10^{-30}$). On the other hand, transforming *formally* S into a continued fraction and evaluating at $t = 2$ gives perfect (i.e., less than 10^{-38}) accuracy. Similarly, for $t = 4$ one obtains respectively 19 and 29 digits of accuracy. In the present example, the continued fraction is completely explicit, and these results can be *proved*. We leave to the reader to prove that

$$\exp(-t) = 1+t/(-1+t/(-2+t/(3+t/(2+t/(-5+t/(-2+t/(7+\cdots)))))))\ ,$$

where the sequence of odd terms is -1, 3, -5, 7, etc..., and that of even terms is -2, 2, -2, 2, ...

4.4. Computation for large *t*. As already mentioned, we can use the generalized power series expansion for any value of t, since the radius of convergence is infinite. However when t is at all large, this is not the best method by far. It is easy to show that all the inverse Mellin transforms that we consider have an explicit *asymptotic expansion* as $t \to \infty$, which is nonconvergent in general, see [4] and [2].

As an example, consider $f(t) = K_0(t)$, the K-Bessel function which occurred in some of the above examples. Its (nowhere convergent) asymptotic expansion is as follows:

$$K_0(t) = \sqrt{\frac{\pi}{2t}} e^{-t} \sum_{n \geq 0} (-1)^n \frac{(2n!)^3}{2^{5n} n!^3} t^{-n}\ .$$

The error is smaller than the absolute value of the first neglected term, and the smallest term in the expansion is for n close to $2t$, which gives an absolute error of approximately $e^{-3t}/(t\sqrt{2})$. Thus, if we want 38 decimal digits, this is obtained only for $t > 28$, corresponding to $n = 56$ for instance. Thus assume that we choose $n = 50$, and that we evaluate both by the truncated series and by the corresponding continued fraction. For $t = 28$ the results are of course perfect for both methods. On the other hand, for $t = 20$, 15, 10, 5, 2, 1 we get 28, 19, 8, 0, 0, 0 correct decimals by using the series, but ≥ 38, ≥ 38, 36, 26, 17, 12 correct decimals by using the continued fraction. It would seem in fact that the continued fraction converges for all t, and rather fast.

Thus, the method suggested by T. Dokshitser is the following: both for small t and large t, after removing the log terms (for small t) or the exponential and power factors (for large t), which are always known explicitly, convert the power series and the asymptotic series into a continued fraction,

and when evaluating $f(t)$ use one or the other continued fraction depending on the size of t.

There are many problems to be solved before using this method: first, for a given accuracy, one must compute the number of terms to be taken in the power series or the asymptotic expansion (which will essentially correspond to the number of terms in the continued fraction), and second one has to determine the threshold which will say when t is "small" or "large". The program `computel` written by T. Dokshitser does this quite well, but not perfectly. As mentioned above, it is available in `Pari/GP`, `Sage`, and `magma`.

Probably the main drawback of this method is that very little can be *proved*, so for now the method is heuristic. However I will venture a guess: since this method is so useful, I believe that it should not be difficult to prove that the method works, to find the exact rate of convergence of the continued fraction, explicit upper bounds for the error, and a good estimate of the threshold, but to my knowledge nobody has done so in the decade since the publication of Dokshitser's paper. The only case where it has been proved (apart from the elementary cases of the exponential and related functions) is for the K-Bessel functions, where the continued fraction is sufficiently regular to be analyzed.

5. Generalized Incomplete Gamma Functions

Another tool that we need is the computation of what one can call a *generalized incomplete gamma function*: as above let $g(s)$ be a product of $\Gamma_{\mathbb{R}}$ factors, and let $f(t)$ be its inverse Mellin transform, so that $g(s) = \int_0^\infty t^s f(t)\, dt/t$. We define the generalized incomplete gamma function associated to g as the two-variable function

$$g(s,x) = \int_x^\infty t^s f(t)\, \frac{dt}{t}\ ,$$

where usually $x \in \mathbb{R}_{\geq 0}$, so that $g(s,0) = g(s)$. We can also define the complementary function as the integral from 0 to x, and if needed we will use both.

The simplest case is $f(t) = e^{-t}$, $g(s) = \Gamma(s)$, and the function $g(s,x)$ is then simply called the incomplete gamma function and denoted $\Gamma(s,x)$.

We can use similar tools to those used for computing inverse Mellin transforms, but here with the added complication that we must deal with *two* variables s and x.

5.1. The Incomplete Gamma Function. Consider first the incomplete gamma function $\Gamma(s,x)$, in other words

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

When x is close to 0, the best is to use power series expansions. There are in fact *two* such expansions:

$$\Gamma(s,x) = \Gamma(s) - x^s \sum_{n\geq 0} (-1)^n \frac{x^n}{(n+s)n!}$$
$$= \Gamma(s) - x^s e^{-x} \sum_{n\geq 0} \frac{x^n}{s(s+1)\cdots(s+n)} .$$

It is in fact easy to show that the second expansion is slightly better.

As for inverse Mellin transforms, these series have infinite radius of convergence, so can be used for all x, but when x is a little large there will be enormous cancellation and they should not be used.

Again as for inverse Mellin, there are asymptotic formulas as $x \to \infty$, which are nonconvergent unless s is a positive integer:

$$\Gamma(s,x) = x^{s-1}e^{-x} \sum_{n\geq 0} (s-1)(s-2)\cdots(s-n)x^{-n} .$$

Once again, we can convert this asymptotic series into a continued fraction, but in this simplest case we obtain a very simple continued fraction, and it is easy to prove that it converges for all x and to give its rate of convergence:

$$\Gamma(s,x) = \cfrac{x^s e^{-x}}{x+1-s-\cfrac{1(1-s)}{x+3-s-\cfrac{2(2-s)}{x+5-s-\ddots}}} ,$$

and if we denote by p_n/q_n the nth convergent, we have

$$\Gamma(s,x) - \frac{p_n}{q_n} \sim \frac{2\pi}{\Gamma(1-s)} e^{-4\sqrt{nx}} .$$

Using all this precise information, it is not difficult to decide which formula to use depending on s and x, and how many terms to take.

Note that when $|s|$ is large, in particular if $s = 1/2+it$ with t large in the context of looking for zeros of L-functions, we encounter a new difficulty when x is close to $|s|$: in that case we have to use other kinds of formulas called *uniform asymptotic expansions*, whose study was initiated by Temme.

5.2. The General Case. For general incomplete gamma functions the situation will be similar: for small x use a generalized power series expansion, for large x use an asymptotic expansion after converting it into a continued fraction. This continued fraction will *conjecturally* converge for all x, and it is possible to give a good estimate of the rate of convergence. However, to my knowledge, nothing is proved except in the simplest cases such as the incomplete gamma function itself.

6. Smoothed Approximate Functional Equations

As mentioned at the beginning, to perform computations on L-functions we have at our disposal two main analytic tools, which are the smoothed approximate functional equations, and the explicit formulas. Although the L-functions we consider have at the same time a functional equation *and* an Euler product, note that the smoothed approximate functional equations need *only* the existence of a functional equation, (the explicit formulas usually need both, although for certain applications which only use the Hadamard product, this is not necessary).

We have the following theorem, easily proved by complex integration, here taken from Rubinstein [11]:

Theorem 6.1. *Assume that* $L(s) = \sum_{n\geq 1} a(n)n^{-s}$ *satisfies the assumptions of Definition 1.1, and in particular that we have a functional equation* $\Lambda(s) = \omega\overline{\Lambda(1-\overline{s})}$, *with* $\Lambda(s) = \gamma(s)L(s)$ *and* $\gamma(s) = N^{s/2}\prod_{1\leq i\leq d}\Gamma_{\mathbb{R}}(s+\lambda_i)$. *For simplicity of exposition, assume that* $L(s)$ *has no poles in* $\mathbb{C}$. *Let* $g(s)$ *be an entire function such that for fixed* s *we have* $|\Lambda(z+s)g(z+s)/z| \to 0$ *as* $\Im(z) \to \infty$ *in any bounded strip* $|\Re(z)| \leq \alpha$. *We have*

$$\Lambda(s)g(s) = \sum_{n\geq 1}\frac{a(n)}{n^s}f_1(s,n) + \omega\sum_{n\geq 1}\frac{\overline{a(n)}}{n^{1-s}}f_2(1-s,n)\ ,$$

where

$$f_1(s,x) = x^s\int_{\sigma-i\infty}^{\sigma+i\infty}\frac{\gamma(z)g(z)x^{-z}}{z-s}\,dz \quad \textit{and}$$
$$f_2(s,x) = x^s\int_{\sigma-i\infty}^{\sigma+i\infty}\frac{\gamma(z)\overline{g(1-\overline{z})}x^{-z}}{z-s}\,dz\ ,$$

and where σ *is any real number greater than the real parts of all the poles of* $\gamma(z)$ *and than* $\Re(s)$.

Several comments are in order concerning this theorem:

(1) It is very easy to modify the formula to take into account possible poles of $L(s)$, see [11] once again.
(2) The functions $f_i(s,x)$ are exponentially decreasing as $x \to \infty$, and in fact one can give a rather precise formula giving their behavior. Thus this gives fast formulas for computing values of $L(s)$ for reasonable values of s. The very simplest case of this approximate functional equation, even simpler than the Riemann zeta function, is for the computation of the value at $s = 1$ of the L-function of an *elliptic curve* E: if the sign of its functional equation is equal to $+1$ (otherwise

$L(E,1)=0$), we have

$$L(E,1)=2\sum_{n\geq 1}\frac{a(n)}{n}e^{-2\pi n/N^{1/2}}\ ,$$

where N is the conductor of the curve, a formula which is immediate from the (unsmoothed) approximate equation, corresponding to $g(s)=1$ in the theorem.

(3) It is not difficult to show that as $n\to\infty$ we have

$$f_i(s,n)\sim C\cdot t^a e^{-\pi d(n/N^{1/2})^{2/d}}$$

for some explicit constants a and C (in the preceding example $d=2$). Thus, we have the so-called $N^{1/2}$-*paradigm*: the series of the theorem indeed converge exponentially fast, but we need at least $\widetilde{O}(N^{1/2})$ terms to obtain any accuracy at all. This is an extremely serious limitation, and probably the most important question in this field of computational number theory: is it possible to do any better? In particular cases (such as $L(E,1)$ above, or some other special values), there are often other methods using the deeper structure of the problem to find the result, but for instance I am pretty sure that nobody has a method faster than $\widetilde{O}(N^{1/2})$ to compute $L(E,\pi)$, say (not that this number has any interest).

Note, however, that quite surprisingly there are some apparent counterexamples to this paradigm: for instance, in [8], Hiary has shown that if the conductor N is far from squarefree, for instance if $N=m^3$, at least in the case of Dirichlet L-functions the computation can be done in time $\widetilde{O}(m)=\widetilde{O}(N^{1/3})$.

(4) The theorem can be used with $g(s)=1$ to compute values of $L(s)$ for "reasonable" values of s. When s is unreasonable, for instance when $s=1/2+iT$ with T large (to check the Riemann hypothesis for instance), one chooses other functions $g(s)$ adapted to the computation to be done, such as $g(s)=e^{is\theta}$ or $g(s)=e^{-a(s-s_0)^2}$; I refer to Rubinstein's paper for detailed examples.

(5) By choosing two very simple functions $g(s)$ such as a^s for two different values of a close to 1, one can compute numerically the value of the root number ω if it is unknown. In a similar manner, if the $a(n)$ are known but not ω nor the conductor N, by choosing a few easy functions $g(s)$ one can find them. But much more surprisingly, as mentioned at the beginning of this paper, if almost nothing is known apart from the gamma factors and N, say, by cleverly choosing a number of functions $g(s)$ and applying techniques from numerical analysis one can prove or disprove (numerically of course) the existence of an L-function having the given gamma factors and conductor, and find

its first few Fourier coefficients if they exist. This method has been used extensively by D. Farmer in his search for $\mathrm{GL}_3(\mathbb{Z})$ and $\mathrm{GL}_4(\mathbb{Z})$ Maass forms, but also by Poor and Yuen, in computations related to the paramodular conjecture of Brumer–Kramer and abelian surfaces, and by A. Mellit in the search of L-functions of degree 4 with integer coefficients and small conductor.

Note that very recent unpublished work of P. Molin (personal communication) shows that to compute values of $L(s)$ one can dispense completely with the approximate functional equation, hence with generalized incomplete gamma functions. As mentioned in a note added in the beginning, his method is now available as part of a large `Pari/GP` package included in the development version.

7. Explicit Formulas

The term "explicit formula" has been used at least since Weil to denote an exact equality linking sums over primes to sums over zeros of L-functions. Of course this basic idea is due to Riemann. Here, we usually assume that our L-function has both a functional equation and an Euler product, although in some cases (but not for the statement of the theorem), the Euler product is not necessary.

Let us consider the simplest example of this, known to Riemann. It is not difficult to show that for all $s \in \mathbb{C}$ we have the identity

$$s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s) = \prod_{\rho}(1 - s/\rho) \ ,$$

where ρ runs over the nontrivial zeros of $\zeta(s)$, and where it is understood that ρ and $1-\rho$ are grouped together (otherwise the product does not converge). On the other hand, if $\Re(s) > 1$ we have the Euler product $\zeta(s) = \prod_p (1-p^{-s})^{-1}$. Thus, taking the logarithmic derivative and expanding we obtain for $\Re(s) > 1$:

$$\frac{1}{s} + \frac{1}{s-1} + \frac{\psi(s/2)}{2} - \frac{\log(\pi)}{2} - \sum_{m \geq 1,\ p} \frac{\log p}{p^{ms}} = \sum_{\rho} \frac{1}{s-\rho} \ ,$$

which is indeed a formula of the indicated type, where $\psi(s) = \Gamma'(s)/\Gamma(s)$ is the logarithmic derivative of the gamma function. In itself this formula is not very interesting, although by *not* expanding the Euler product (which is of course contrary to the spirit of explicit formulas), one can compute explicitly sums and products linked to the nontrivial zeros such as $\sum_\rho \rho^{-k}$ for $k \geq 1$, $\prod_\rho (1 - 1/\rho^2)$, or $\prod_\rho (1 - 4/\rho^2)$.

The proof of the above identity is easily done by integrating $\Lambda'(s)/\Lambda(s)$ around a suitable contour and applying the residue theorem, thus catching the poles ρ of Λ'/Λ. The idea of explicit formulas is, instead of integrating

Λ'/Λ, to choose a suitable test function Φ, and integrate $\Phi(s)\Lambda'(s)/\Lambda(s)$. We will thus essentially obtain a sum of $\Phi(\rho)$ (plus some other explicit expressions), and using the Euler product of ζ, this will be equal as above to a sum over prime powers of an *integral transform* of Φ, essentially its Fourier, Laplace, or Mellin transform, depending on the normalization. The computations are very easy, the only slight difficulty in the proof lying in their justifications.

The reader can find a general version in Mestre's paper [9], we simply give without technical details the formula for Dirichlet L-series, the general case being very similar:

Theorem 7.1. *Let χ be an* even *primitive Dirichlet character of conductor N, and let F be a real function satisfying a number of easy technical conditions. We have the explicit formula:*

$$\begin{aligned}\sum_\rho \Phi(\rho) = &- \sum_{p,k\ge 1} \frac{\log(p)}{p^{k/2}}(\chi^k(p)F(k\log(p)) + \overline{\chi^k(p)}F(-k\log(p)))\\ &+ 2\delta_{N,1}\int_{-\infty}^{\infty} F(x)\cosh(x/2)\,dx + F(0)\log(N/\pi)\\ &+ \int_0^\infty \left(\frac{e^{-x}}{x}F(0) - \frac{e^{-x/4}}{1-e^{-x}}\frac{F(x/2)+F(-x/2)}{2}\right)dx\ ,\end{aligned}$$

where $\delta_{N,1}$ is the Kronecker symbol and where we set

$$\Phi(s) = \int_{-\infty}^{\infty} F(x)e^{(s-1/2)x}\,dx\ ,$$

and as above the sum on ρ is a sum over all the nontrivial zeros of $L(\chi,s)$ taken symmetrically ($\sum_\rho = \lim_{T\to\infty}\sum_{|\Im(\rho)|\le T}$).

Remarks.

(1) Write $\rho = 1/2 + i\gamma$ (if the GRH is true all γ are real, but even without GRH we can always write this). Then

$$\Phi(\rho) = \int_{-\infty}^{\infty} F(x)e^{i\gamma x}\,dx = \widehat{F}(\gamma)$$

is simply the value at γ of the *Fourier transform* $\widehat{F}$ of F.

(2) It is immediate to generalize to odd χ or to more general L-functions, see [9].

This theorem can be used in several different directions, and has been an extremely valuable tool in analytic number theory. Just to mention a few:

(1) Since the conductor N occurs, we can obtain *lower bounds* on N, possibly assuming certain conjectures such as the generalized Riemann hypothesis. For instance, this is how Stark–Odlyzko–Poitou–Serre find

discriminant lower bounds for number fields. This is also how Mestre finds lower bounds for conductors of abelian varieties, and so on.

(2) When the L-function has a zero at its central point (here of course it usually does not, but for more general L-functions it is important), this can give good upper bounds for the order of the zero. For instance, using this formula, J. Bober et al. have proved that the elliptic curves of very high rank ≥ 20 found by a number of people do have exactly the claimed rank (and no larger), assuming both the GRH and the Birch and Swinnerton-Dyer conjecture.

(3) More generally, suitable choices of the test functions can give information on the nontrivial zeros ρ of small imaginary part.

(4) In [1], A. Booker explains in great detail how to use, among other tools, the explicit formula both to verify Artin's conjecture on the holomorphy of L-functions, and the Generalized Riemann Hypothesis to moderately large heights. In particular, he gives a completely rigorous method to compute (to moderate accuracy) *all* the zeros of an L-function on the critical line, without missing any.

8. Hypergeometric Motives

We finish this paper by describing in some detail the notion of *hypergeometric motives*, initially introduced by N. Katz, but put in the present completely explicit form by F. Rodriguez–Villegas, and which is currently the subject of active research. Note that thanks to the work of M. Watkins, there is now a remarkable `magma` package for working with hypergeometric motives.

To any algebraic variety defined over $\mathbb{Q}$, say, one can associate a local zeta function, or better local L-functions depending on the splitting of the cohomology of the variety, at least for all primes except the "bad" ones, finite in number. These are (for smooth, projective varieties at least) rational functions, and replacing the formal variable T by p^{-s} and taking the product leads to *global* L-functions attached to the variety or parts of its cohomology. Ignoring the fact that completing these L-functions at the bad primes is often difficult, this leads to interesting L-functions of higher degree, but which may not be easy to compute explicitly.

An example where this *can* easily be done is in the case of dimension 0: this amounts to studying the *Dedekind zeta function* of number fields, and there are very efficient methods for doing this, and for computing it, although as far as I am aware the only available implementation is due to my ex-student E. Tollis in `Pari/GP` (hence in `Sage`) more than 12 years ago, and is far from being optimal, but nobody has taken the time to improve it since.

The idea of Katz and Rodriguez-Villegas is instead to introduce *directly* the local and global L-functions (at least for the good primes), *without* explicit reference to the underlying variety (which of course exists: we call the process of finding equations for that variety, which is *not* necessary, *reverse engineering*). This is done by using a finite field variant of *hypergeometric functions*.

Let me give a short and unmotivated introduction to this: let $\gamma(T) = \sum_{n\geq 1}\gamma_n T^n \in \mathbb{Z}[T]$ be a polynomial satisfying the conditions $\gamma(0) = 0$ and $\gamma'(1) = 0$ (in other words $\gamma_0 = 0$ and $\sum_n n\gamma_n = 0$). For any finite field $\mathbb{F}_q$ with $q = p^f$ and any character χ of $\mathbb{F}_q^*$, recall that the Gauss sum $\mathfrak{g}(\chi)$ is defined by

$$\mathfrak{g}(\chi) = \sum_{x\in\mathbb{F}_q^*} \chi(x)\exp(2\pi i \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)/p) \ .$$

We set

$$Q_q(\gamma;\chi) = \prod_{n\geq 1} \mathfrak{g}(\chi^n)^{\gamma_n}$$

and for any $t \in \mathbb{F}_q \setminus \{0,1\}$

$$a_q(\gamma;t) = \frac{q^d}{1-q}\left(1 + \sum_{\chi\neq\chi_0} \chi(Mt)Q_q(\gamma;\chi)\right) \ ,$$

where as usual χ_0 is the trivial character and d is an integer and M a nonzero rational number which can easily be given explicitly (M is simply a normalization parameter, since one could change Mt into t). Then the "theorem" of Katz is that for $t \neq 0, 1$ the quantity $a_q(\gamma;t)$ is the trace of Frobenius on some *motive defined over* $\mathbb{Q}$ (I put theorem in quotes because it is not completely clear what the status of the proof is, although there is no doubt that it is true). In the language of L-functions this means the following: define à la Weil the formal power series

$$L_p(\gamma;t;T) = \exp\left(\sum_{f\geq 1} a_{p^f}(\gamma;t)\frac{T^f}{f}\right) \ .$$

Then L_p is a rational function of T, satisfies the local Riemann hypothesis, and if we set

$$L(\gamma;t;s) = \prod_p L_p(\gamma;t;p^{-s})^{-1} \ ,$$

then L once completed at the "bad" primes and the prime at infinity (i.e., suitable gamma factors) should satisfy a functional equation of the standard type.

The work being done on this subject goes in a number of different directions.

• Find a suitable recipe for the Euler factors and the conductor at the bad primes. This can be done experimentally using the methods described above, and is very successful. However one then needs theoretical methods to prove that the experimental guesses are correct, and this is being done in particular by D. Roberts and F. Rodriguez-Villegas. Note that some of the bad primes are "tame" and quite well understood, but some are wild and much less well understood.

• Perform the reverse engineering for many interesting examples, in other words find an equation of an algebraic variety to which the motive corresponds. This has in particular been done by F. Beukers, M. Watkins, and the author. As an example, we have found that the Artin motive corresponding to $W(F_4)$, the Weyl group of F_4, with $\gamma(T) = T^{12}-T^6-T^4-T^3+T$ and parameter t, comes from counting points on the affine elliptic surface $y^2 = x^3 + x^2 + cz^3 + z$, where $c = -27t/4$.

• Link the L-functions of hypergeometric motives with L-functions of automorphic forms. Many such links clearly exist, but almost none are proved. The link is easy to guess simply by looking at the gamma factors, which are given for hypergeometric motive through its Hodge numbers, a conjecture due to Corti and Golyshev, and thanks to the well-known recipe of Serre. We thus find ordinary elliptic modular forms (equivalently, elliptic curves over $\mathbb{Q}$), symmetric squares of them, Siegel modular forms, and so on.

• Classify hypergeometric motives corresponding to *Calabi–Yau* manifolds. Note for instance that the whole subject was given a new impetus after the work of Candelas–De la Ossa and Rodriguez-Villegas on the Calabi–Yau quintic $x_1^5 + x_2^5 + x_3^5 + x_4^5 + x_5^5 - 5\psi x_1x_2x_3x_4x_5 = 0$, since the interesting part of its cohomology corresponds to the hypergeometric motive with $\gamma(T) = T^5 - 5T$. The L-function of the corresponding motives should correspond to Siegel modular forms of degree 2 and weight 3, but as already mentioned, the smallest conductor of such a motive (525) is larger than what experts in Siegel modular forms can tabulate. However, I have been told, but have no reference, that such Calabi–Yau manifolds have indeed been proved to be modular, so in this case the existence of the corresponding Siegel modular form is a theorem.

References

[1] A. R. Booker, "Artin's conjecture, Turing's method, and the Riemann hypothesis", *Experiment. Math.* **15** (2006), no. 4, p. 385-407.

[2] B. L. J. Braaksma, "Asymptotic expansions and analytic continuations for a class of Barnes-integrals", *Compositio Math.* **15** (1964), p. 239-341 (1964).

[3] A. Brumer & K. Kramer, "Paramodular abelian varieties of odd conductor", `http://arxiv.org/abs/1004.4699v3`.

[4] H. COHEN, *Advanced topics in computational number theory*, Graduate Texts in Mathematics, vol. 193, Springer-Verlag, New York, 2000, xvi+578 pages.
[5] H. COHEN & D. ZAGIER, "Vanishing and non-vanishing theta values", *Ann. Math. Qué.* **37** (2013), no. 1, p. 45-61.
[6] T. DOKCHITSER, "Computing special values of motivic L-functions", *Experiment. Math.* **13** (2004), no. 2, p. 137-149.
[7] D. FARMER, S. KOUTSIOLIOTAS & S. LEMURELL, "Maass forms on GL(3) and GL(4)", `http://arxiv.org/abs/1212.4545`.
[8] G. HIARY, "Computing Dirichlet character sums to a power-full modulus", `http://arxiv.org/abs/1205.4687`.
[9] J.-F. MESTRE, "Formules explicites et minorations de conducteurs de variétés algébriques", *Compositio Math.* **58** (1986), no. 2, p. 209-232.
[10] C. POOR & D. YUEN, "Paramodular cusp forms", `http://arxiv.org/abs/0912.0049v1`.
[11] M. RUBINSTEIN, "Computational methods and experiments in analytic number theory", in *Recent perspectives in random matrix theory and number theory*, London Math. Soc. Lecture Note Ser., vol. 322, Cambridge Univ. Press, Cambridge, 2005, p. 425-506.

Henri COHEN
Université de Bordeaux,
Institut de Mathématiques de Bordeaux,
UMR 5251 du CNRS, 351 Cours de la Libération,
33405 TALENCE Cedex
FRANCE
E-mail: `cohen@math.u-bordeaux.fr`
URL: `http://www.math.u-bordeaux.fr/~hecohen/`