Hypergeometric L-functions via Frobenius structures

Kiran S. Kedlaya

Department of Mathematics, University of California, San Diego
School of Mathematics, Institute for Advanced Study (2018–19)
kedlaya@ucsd.edu
http://kskedlaya.org/slides/

Arithmetic Geometry, Number Theory, and Computation
Massachusetts Institute of Technology
August 24, 2018

Work in progress; no preprint available yet, but see this Jupyter notebook for a previous lecture.
Kedlaya is supported by NSF (grants DMS-1501214, DMS-1802161), UC San Diego (Warschawski Professorship), and IAS during the 2018–19 academic year. Thanks also to the Simons Foundation for support via the Collaboration.

Contents

1. Motives and L-functions
2. Hypergeometric motives and their L-functions
3. Finite hypergeometric sums and trace formulas
4. Frobenius structures
5. GKZ A-hypergeometric systems

Zeta functions of schemes

For X a scheme of finite type over $\mathbb{Z}$, its **zeta function** is the Dirichlet series (absolutely convergent for $\mathrm{Re}(s) \gg 0$)

$$\zeta(X, s) = \prod_{x \in X^\circ} (1 - \#\kappa(x)^{-s})^{-1}$$

where X° denotes the set of closed points of X and $\kappa(x)$ the residue field of the closed point x.

- When $X = \operatorname{Spec} \mathbb{Z}$, $\zeta(X, s) = \zeta(s)$, the Riemann zeta function.
- When $X = \operatorname{Spec} \mathfrak{o}_K$ for K a number field, $\zeta(X, s) = \zeta_K(s)$, the Dedekind zeta function (without archimedean factors).
- When X lies over the finite field $\mathbb{F}_q$, $\zeta(X, s)$ is a rational function of q^{-s} (Dwork).

Zeta functions of schemes

For X a scheme of finite type over $\mathbb{Z}$, its **zeta function** is the Dirichlet series (absolutely convergent for $\mathrm{Re}(s) \gg 0$)

$$\zeta(X, s) = \prod_{x \in X^\circ} (1 - \#\kappa(x)^{-s})^{-1}$$

where X° denotes the set of closed points of X and $\kappa(x)$ the residue field of the closed point x.

- When $X = \mathrm{Spec}\,\mathbb{Z}$, $\zeta(X, s) = \zeta(s)$, the Riemann zeta function.
- When $X = \mathrm{Spec}\,\mathfrak{o}_K$ for K a number field, $\zeta(X, s) = \zeta_K(s)$, the Dedekind zeta function (without archimedean factors).
- When X lies over the finite field $\mathbb{F}_q$, $\zeta(X, s)$ is a rational function of q^{-s} (Dwork).

Zeta functions of schemes

For X a scheme of finite type over $\mathbb{Z}$, its **zeta function** is the Dirichlet series (absolutely convergent for $\mathrm{Re}(s) \gg 0$)

$$\zeta(X, s) = \prod_{x \in X^\circ} (1 - \#\kappa(x)^{-s})^{-1}$$

where X° denotes the set of closed points of X and $\kappa(x)$ the residue field of the closed point x.

- When $X = \operatorname{Spec} \mathbb{Z}$, $\zeta(X, s) = \zeta(s)$, the Riemann zeta function.
- When $X = \operatorname{Spec} \mathfrak{o}_K$ for K a number field, $\zeta(X, s) = \zeta_K(s)$, the Dedekind zeta function (without archimedean factors).
- When X lies over the finite field $\mathbb{F}_q$, $\zeta(X, s)$ is a rational function of q^{-s} (Dwork).

Zeta functions of schemes

For X a scheme of finite type over $\mathbb{Z}$, its **zeta function** is the Dirichlet series (absolutely convergent for $\mathrm{Re}(s) \gg 0$)

$$\zeta(X, s) = \prod_{x \in X^\circ} (1 - \#\kappa(x)^{-s})^{-1}$$

where X° denotes the set of closed points of X and $\kappa(x)$ the residue field of the closed point x.

- When $X = \operatorname{Spec} \mathbb{Z}$, $\zeta(X, s) = \zeta(s)$, the Riemann zeta function.
- When $X = \operatorname{Spec} \mathfrak{o}_K$ for K a number field, $\zeta(X, s) = \zeta_K(s)$, the Dedekind zeta function (without archimedean factors).
- When X lies over the finite field $\mathbb{F}_q$, $\zeta(X, s)$ is a rational function of q^{-s} (Dwork).

More on zeta functions over finite fields

For X smooth proper of dimension d over $\mathbb{F}_q$, we have

$$\zeta(X,s) = \frac{P_1(q^{-s}) \cdots P_{2d-1}(q^{-s})}{P_0(q^{-s}) \cdots P_{2d}(q^{-s})}$$

where $P_i(T) \in 1 + T\mathbb{Z}[T]$ has all roots on the circle $|T| = q^{-i/2}$. For example, if X is an elliptic curve, then

$$\zeta(X,s) = \frac{1 - a_q q^{-s} + q^{1-2s}}{(1-q^{-s})(1-q^{1-s})}, \quad a_q = q + 1 - \#X(\mathbb{F}_q)$$

and the condition on the roots of P_1 is equivalent to the Hasse bound $|a_q| \leq 2\sqrt{q}$.

L-functions

For X smooth proper of dimension d over $\mathbb{Q}$, for $i = 0, \ldots, 2d$ we define the L-**functions**

$$L_i(X, s) = \prod_p P_{i,p}(p^{-s})^{-1}$$

where for each prime p of good reduction, $P_{i,p}$ is the factor P_i of the zeta function of the reduction of X modulo p. (There are also Euler factors for bad primes and archimedean factors to clean up the functional equation.)

For example, if X is an elliptic curve, then $L_1(X, s)$ is the usual L-function (with bad Euler factors suppressed):

$$L_1(X, s) = \prod_p (1 - a_p p^{-s} + p^{1-2s})^{-1}.$$

For X over a number field K, one has a similar definition indexed by prime ideals of $\mathfrak{o}_K$. But we won't see any examples with $K \neq \mathbb{Q}$ in this lecture.

L-functions

For X smooth proper of dimension d over $\mathbb{Q}$, for $i = 0, \dots, 2d$ we define the L-**functions**

$$L_i(X, s) = \prod_p P_{i,p}(p^{-s})^{-1}$$

where for each prime p of good reduction, $P_{i,p}$ is the factor P_i of the zeta function of the reduction of X modulo p. (There are also Euler factors for bad primes and archimedean factors to clean up the functional equation.)

For example, if X is an elliptic curve, then $L_1(X, s)$ is the usual L-function (with bad Euler factors suppressed):

$$L_1(X, s) = \prod_p (1 - a_p p^{-s} + p^{1-2s})^{-1}.$$

For X over a number field K, one has a similar definition indexed by prime ideals of $\mathfrak{o}_K$. But we won't see any examples with $K \neq \mathbb{Q}$ in this lecture.

L-functions

For X smooth proper of dimension d over $\mathbb{Q}$, for $i = 0, \dots, 2d$ we define the L-**functions**

$$L_i(X, s) = \prod_p P_{i,p}(p^{-s})^{-1}$$

where for each prime p of good reduction, $P_{i,p}$ is the factor P_i of the zeta function of the reduction of X modulo p. (There are also Euler factors for bad primes and archimedean factors to clean up the functional equation.)

For example, if X is an elliptic curve, then $L_1(X, s)$ is the usual L-function (with bad Euler factors suppressed):

$$L_1(X, s) = \prod_p (1 - a_p p^{-s} + p^{1-2s})^{-1}.$$

For X over a number field K, one has a similar definition indexed by prime ideals of $\mathfrak{o}_K$. But we won't see any examples with $K \neq \mathbb{Q}$ in this lecture.

Motives

In some cases, L-functions of schemes factor for "geometric reasons". For example, if $\dim(X) = 1$, then one gets factorizations of $L_1(X, s)$ from isogeny decompositions of the Jacobian of X. (E.g., for modular curves these arise from the action of Hecke correspondences.)

For $i > 1$, there is no analogue of Jacobians to explain geometric factorizations. Instead, one defines into existence a category of **motives of weight** i in which X "decomposes" into "direct summands" corresponding to factors of the L-function. (One can be more precise about this, but it is a long and messy story...)

Motives

In some cases, L-functions of schemes factor for "geometric reasons". For example, if $\dim(X) = 1$, then one gets factorizations of $L_1(X, s)$ from isogeny decompositions of the Jacobian of X. (E.g., for modular curves these arise from the action of Hecke correspondences.)

For $i > 1$, there is no analogue of Jacobians to explain geometric factorizations. Instead, one defines into existence a category of **motives of weight** i in which X "decomposes" into "direct summands" corresponding to factors of the L-function. (One can be more precise about this, but it is a long and messy story...)

Motives

In some cases, L-functions of schemes factor for "geometric reasons". For example, if $\dim(X) = 1$, then one gets factorizations of $L_1(X, s)$ from isogeny decompositions of the Jacobian of X. (E.g., for modular curves these arise from the action of Hecke correspondences.)

For $i > 1$, there is no analogue of Jacobians to explain geometric factorizations. Instead, one defines into existence a category of **motives of weight** i in which X "decomposes" into "direct summands" corresponding to factors of the L-function. (One can be more precise about this, but it is a long and messy story...)

Contents

1. Motives and L-functions
2. Hypergeometric motives and their L-functions
3. Finite hypergeometric sums and trace formulas
4. Frobenius structures
5. GKZ A-hypergeometric systems

Hypergeometric data

A **hypergeometric (HG) datum** of degree n consists of two disjoint n-tuples $\underline{\alpha}, \underline{\beta} \in (\mathbb{Q} \cap [0,1))^n$. (Repeats within a tuple are allowed.) Such a datum is **Galois-stable** if the tuples $e^{2\pi i \underline{\alpha}}, e^{2\pi i \underline{\beta}}$ are $G_{\mathbb{Q}}$-stable; that is, any two $\frac{r}{s}, \frac{r'}{s} \in \mathbb{Q} \cap [0,1)$ in lowest terms have the same multiplicities. For example, one Galois-stable HG datum of degree 5 is

$$\underline{\alpha} = \left(\frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}\right), \qquad \underline{\beta} = \left(\frac{0}{1}, \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}\right).$$

For any given n, there are finitely many Galois-stable HG data of length n.

```
sage: from sage.modular.hypergeometric_motive import possible_hypergeometric_data as poss
sage: l = poss(5); len(l)
147
sage: l[56] # For brevity, Sage reports only one of (alpha, beta) and (beta, alpha)
Hypergeometric data for [1/8, 3/8, 1/2, 5/8, 7/8] and [0, 1/4, 1/3, 2/3, 3/4]
```

This example uses the Sage HG motives package by Chapoton-K; this is a partial port of the Magma package by Watkins (based on work of Cohen).

Hypergeometric data

A **hypergeometric (HG) datum** of degree n consists of two disjoint n-tuples $\underline{\alpha}, \underline{\beta} \in (\mathbb{Q} \cap [0,1))^n$. (Repeats within a tuple are allowed.) Such a datum is **Galois-stable** if the tuples $e^{2\pi i \underline{\alpha}}, e^{2\pi i \underline{\beta}}$ are $G_{\mathbb{Q}}$-stable; that is, any two $\frac{r}{s}, \frac{r'}{s} \in \mathbb{Q} \cap [0,1)$ in lowest terms have the same multiplicities. For example, one Galois-stable HG datum of degree 5 is

$$\underline{\alpha} = \left(\frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}\right), \qquad \underline{\beta} = \left(\frac{0}{1}, \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}\right).$$

For any given n, there are finitely many Galois-stable HG data of length n.

```
sage: from sage.modular.hypergeometric_motive import possible_hypergeometric_data as poss
sage: l = poss(5); len(l)
147
sage: l[56] # For brevity, Sage reports only one of (alpha, beta) and (beta, alpha)
Hypergeometric data for [1/8, 3/8, 1/2, 5/8, 7/8] and [0, 1/4, 1/3, 2/3, 3/4]
```

This example uses the Sage HG motives package by Chapoton-K; this is a partial port of the Magma package by Watkins (based on work of Cohen).

Hypergeometric data

A **hypergeometric (HG) datum** of degree n consists of two disjoint n-tuples $\underline{\alpha}, \underline{\beta} \in (\mathbb{Q} \cap [0,1))^n$. (Repeats within a tuple are allowed.) Such a datum is **Galois-stable** if the tuples $e^{2\pi i \underline{\alpha}}, e^{2\pi i \underline{\beta}}$ are $G_{\mathbb{Q}}$-stable; that is, any two $\frac{r}{s}, \frac{r'}{s} \in \mathbb{Q} \cap [0,1)$ in lowest terms have the same multiplicities. For example, one Galois-stable HG datum of degree 5 is

$$\underline{\alpha} = \left(\frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}\right), \qquad \underline{\beta} = \left(\frac{0}{1}, \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}\right).$$

For any given n, there are finitely many Galois-stable HG data of length n.

```
sage: from sage.modular.hypergeometric_motive import possible_hypergeometric_data as poss
sage: l = poss(5); len(l)
147
sage: l[56] # For brevity, Sage reports only one of (alpha, beta) and (beta, alpha)
Hypergeometric data for [1/8, 3/8, 1/2, 5/8, 7/8] and [0, 1/4, 1/3, 2/3, 3/4]
```

This example uses the Sage HG motives package by Chapoton-K; this is a partial port of the Magma package by Watkins (based on work of Cohen).

Hypergeometric motives (HGMs)

For any Galois-stable HG datum $(\underline{\alpha}, \underline{\beta})$, one gets a one-parameter family of motives indexed by $t \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$. For each $t \in \mathbb{Q} \setminus \{0, 1\}$, one gets a motive with an associated L-function of degree n.

The discrete data of these L-functions (motivic weight, Hodge numbers) are determined combinatorially by $\underline{\alpha}, \underline{\beta}$ (see below); these obey no obvious constraints (i.e., they are "diverse"). Moreover, the L-functions can be computed efficiently (see below). This makes HGMs a rich source of L-functions, and a prime candidate for inclusion in the LMFDB.

Warning: this construction is "folklore" by analogy with work of Katz on the ℓ-adic realizations of these motives (see *Exponential Sums and Differential Equations*), but AFAIK there is no available reference for the motivic construction. As a stopgap, see the Magma documentation.

Hypergeometric motives (HGMs)

For any Galois-stable HG datum $(\underline{\alpha}, \underline{\beta})$, one gets a one-parameter family of motives indexed by $t \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$. For each $t \in \mathbb{Q} \setminus \{0, 1\}$, one gets a motive with an associated L-function of degree n.

The discrete data of these L-functions (motivic weight, Hodge numbers) are determined combinatorially by $\underline{\alpha}, \underline{\beta}$ (see below); these obey no obvious constraints (i.e., they are "diverse"). Moreover, the L-functions can be computed efficiently (see below). This makes HGMs a rich source of L-functions, and a prime candidate for inclusion in the LMFDB.

Warning: this construction is "folklore" by analogy with work of Katz on the ℓ-adic realizations of these motives (see *Exponential Sums and Differential Equations*), but AFAIK there is no available reference for the motivic construction. As a stopgap, see the Magma documentation.

Hypergeometric motives (HGMs)

For any Galois-stable HG datum $(\underline{\alpha}, \underline{\beta})$, one gets a one-parameter family of motives indexed by $t \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$. For each $t \in \mathbb{Q} \setminus \{0, 1\}$, one gets a motive with an associated L-function of degree n.

The discrete data of these L-functions (motivic weight, Hodge numbers) are determined combinatorially by $\underline{\alpha}, \underline{\beta}$ (see below); these obey no obvious constraints (i.e., they are "diverse"). Moreover, the L-functions can be computed efficiently (see below). This makes HGMs a rich source of L-functions, and a prime candidate for inclusion in the LMFDB.

Warning: this construction is "folklore" by analogy with work of Katz on the ℓ-adic realizations of these motives (see *Exponential Sums and Differential Equations*), but AFAIK there is no available reference for the motivic construction. As a stopgap, see the Magma documentation.

Examples

In the case

$$n = 2, \qquad \underline{\alpha} = \left(\frac{1}{2}, \frac{1}{2}\right), \qquad \underline{\beta} = (0, 0),$$

the resulting family of HGMs is essentially the 1-motives of the Legendre family of elliptic curves.

Other particular families of HGMs are known (work of Naskręcki) to correspond to (particular families of) hyperelliptic curves, ruled/elliptic/K3 surfaces, Calabi–Yau threefolds, etc. For instance,

$$n = 3, \qquad \underline{\alpha} = \left(\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right), \qquad \underline{\beta} = (0, 0, 0)$$

gives a pure Chow motive occurring in the elliptic K3 surface

$$xyz(1 - (x + y + z)) = \frac{1}{256t}.$$

Examples

In the case

$$n = 2, \qquad \underline{\alpha} = \left(\frac{1}{2}, \frac{1}{2}\right), \qquad \underline{\beta} = (0, 0),$$

the resulting family of HGMs is essentially the 1-motives of the Legendre family of elliptic curves.

Other particular families of HGMs are known (work of Naskręcki) to correspond to (particular families of) hyperelliptic curves, ruled/elliptic/K3 surfaces, Calabi–Yau threefolds, etc. For instance,

$$n = 3, \qquad \underline{\alpha} = \left(\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right), \qquad \underline{\beta} = (0, 0, 0)$$

gives a pure Chow motive occurring in the elliptic K3 surface

$$xyz(1 - (x + y + z)) = \frac{1}{256t}.$$

Hodge numbers and motivic weight

For $x \in \mathbb{R}$, define the **zigzag function**

$$D(x) = \#\{i \in \{1, \ldots, n\} : \alpha_i \leq x\} - \#\{i \in \{1, \ldots, n\} : \beta_i \leq x\}.$$

To get the Hodge vector, compute the vector $(h_i)_{i \in \mathbb{Z}}$ where

$$h_i = \#\{j \in \{1, \ldots, n\} : D(\alpha_j) = i\},$$

shift it to be symmetric across 0, then drop leading and trailing zeroes. The motivic weight is one less than the length of the vector. These are invariant under interchanging $\underline{\alpha}$ and $\underline{\beta}$.

```
sage: from sage.modular.hypergeometric_motive import HypergeometricData as Hyp
sage: H = Hyp(alpha_beta=([1/4,1/3,1/2,2/3,3/4], [0,1/8,3/8,5/8,7/8]))
sage: alpha,beta = H.alpha_beta()
sage: [H.zigzag(a) for a in alpha]
[-1, 0, 0, 0, 1]
sage: H.hodge_numbers()
[1, 3, 1]
sage: H.weight()
2
```

Hodge numbers and motivic weight

For $x \in \mathbb{R}$, define the **zigzag function**

$$D(x) = \#\{i \in \{1, \ldots, n\} : \alpha_i \leq x\} - \#\{i \in \{1, \ldots, n\} : \beta_i \leq x\}.$$

To get the Hodge vector, compute the vector $(h_i)_{i \in \mathbb{Z}}$ where

$$h_i = \#\{j \in \{1, \ldots, n\} : D(\alpha_j) = i\},$$

shift it to be symmetric across 0, then drop leading and trailing zeroes. The motivic weight is one less than the length of the vector. These are invariant under interchanging $\underline{\alpha}$ and $\underline{\beta}$.

```
sage: from sage.modular.hypergeometric_motive import HypergeometricData as Hyp
sage: H = Hyp(alpha_beta=([1/4,1/3,1/2,2/3,3/4], [0,1/8,3/8,5/8,7/8]))
sage: alpha,beta = H.alpha_beta()
sage: [H.zigzag(a) for a in alpha]
[-1, 0, 0, 0, 1]
sage: H.hodge_numbers()
[1, 3, 1]
sage: H.weight()
2
```

Bad primes

For $t \in \mathbb{Q} \setminus \{0, 1\}$, the primes p of bad reduction of the associated HGM are of two types.

- **wild**: p divides the least common denominator of $\underline{\alpha} \cup \underline{\beta}$.
- **tame**: p is not wild and $t \equiv 0, 1, \infty \mod p$.

For some applications (e.g., computing special values), it is important to know (or guess well) the Euler factors for bad primes. For other applications (e.g., computing statistics over Euler factors as in the Sato–Tate or Lang–Trotter conjectures), it is less important. We will ignore bad primes hereafter.

Bad primes

For $t \in \mathbb{Q} \setminus \{0, 1\}$, the primes p of bad reduction of the associated HGM are of two types.

- **wild**: p divides the least common denominator of $\underline{\alpha} \cup \underline{\beta}$.
- **tame**: p is not wild and $t \equiv 0, 1, \infty \mod p$.

For some applications (e.g., computing special values), it is important to know (or guess well) the Euler factors for bad primes. For other applications (e.g., computing statistics over Euler factors as in the Sato–Tate or Lang–Trotter conjectures), it is less important. We will ignore bad primes hereafter.

Bad primes

For $t \in \mathbb{Q} \setminus \{0, 1\}$, the primes p of bad reduction of the associated HGM are of two types.

- **wild**: p divides the least common denominator of $\underline{\alpha} \cup \underline{\beta}$.
- **tame**: p is not wild and $t \equiv 0, 1, \infty \mod p$.

For some applications (e.g., computing special values), it is important to know (or guess well) the Euler factors for bad primes. For other applications (e.g., computing statistics over Euler factors as in the Sato–Tate or Lang–Trotter conjectures), it is less important. We will ignore bad primes hereafter.

Good Euler factors

For p a good prime, Magma and (somewhat less efficiently) Sage can compute the Euler factor at p of the associated HGM. (Warning: Magma's t is Sage's $1/t$; the Sage convention is more consistent with the literature.)

```
sage: H.euler_factor(5/7, 11)
-161051*T^5 - 9317*T^4 - 484*T^3 + 44*T^2 + 7*T + 1
sage: H.euler_factor(5/7, 7)
...
NotImplementedError: p is tame
sage: H.euler_factor(5/7, 3)
...
NotImplementedError: p is wild
```

In the remainder of the talk, we aim to answer two questions (with the broader goal of getting interesting L-functions into the LMFDB).

- How is this computation done currently?
- What alternatives may be more efficient in certain circumstances?

Good Euler factors

For p a good prime, Magma and (somewhat less efficiently) Sage can compute the Euler factor at p of the associated HGM. (Warning: Magma's t is Sage's $1/t$; the Sage convention is more consistent with the literature.)

```
sage: H.euler_factor(5/7, 11)
-161051*T^5 - 9317*T^4 - 484*T^3 + 44*T^2 + 7*T + 1
sage: H.euler_factor(5/7, 7)
...
NotImplementedError: p is tame
sage: H.euler_factor(5/7, 3)
...
NotImplementedError: p is wild
```

In the remainder of the talk, we aim to answer two questions (with the broader goal of getting interesting L-functions into the LMFDB).

- How is this computation done currently?
- What alternatives may be more efficient in certain circumstances?

Good Euler factors

For p a good prime, Magma and (somewhat less efficiently) Sage can compute the Euler factor at p of the associated HGM. (Warning: Magma's t is Sage's $1/t$; the Sage convention is more consistent with the literature.)

```
sage: H.euler_factor(5/7, 11)
-161051*T^5 - 9317*T^4 - 484*T^3 + 44*T^2 + 7*T + 1
sage: H.euler_factor(5/7, 7)
...
NotImplementedError: p is tame
sage: H.euler_factor(5/7, 3)
...
NotImplementedError: p is wild
```

In the remainder of the talk, we aim to answer two questions (with the broader goal of getting interesting L-functions into the LMFDB).

- How is this computation done currently?
- What alternatives may be more efficient in certain circumstances?

Good Euler factors

For p a good prime, Magma and (somewhat less efficiently) Sage can compute the Euler factor at p of the associated HGM. (Warning: Magma's t is Sage's $1/t$; the Sage convention is more consistent with the literature.)

```
sage: H.euler_factor(5/7, 11)
-161051*T^5 - 9317*T^4 - 484*T^3 + 44*T^2 + 7*T + 1
sage: H.euler_factor(5/7, 7)
...
NotImplementedError: p is tame
sage: H.euler_factor(5/7, 3)
...
NotImplementedError: p is wild
```

In the remainder of the talk, we aim to answer two questions (with the broader goal of getting interesting L-functions into the LMFDB).

- How is this computation done currently?
- What alternatives may be more efficient in certain circumstances?

Good Euler factors

For p a good prime, Magma and (somewhat less efficiently) Sage can compute the Euler factor at p of the associated HGM. (Warning: Magma's t is Sage's $1/t$; the Sage convention is more consistent with the literature.)

```
sage: H.euler_factor(5/7, 11)
-161051*T^5 - 9317*T^4 - 484*T^3 + 44*T^2 + 7*T + 1
sage: H.euler_factor(5/7, 7)
...
NotImplementedError: p is tame
sage: H.euler_factor(5/7, 3)
...
NotImplementedError: p is wild
```

In the remainder of the talk, we aim to answer two questions (with the broader goal of getting interesting L-functions into the LMFDB).

- How is this computation done currently?
- What alternatives may be more efficient in certain circumstances?

Contents

1. Motives and L-functions
2. Hypergeometric motives and their L-functions
3. Finite hypergeometric sums and trace formulas
4. Frobenius structures
5. GKZ A-hypergeometric systems

A hypergeometric trace formula

In this approach, we compute a good Euler factor of an HG motive by computing the trace of the q-Frobenius for each prime power q (on ℓ-adic cohomology for some prime ℓ not dividing q; the exact choice is irrelevant).

Building on work of Greene, Katz, McCarthy, Ono, ..., an explicit formula for this trace was given by Beukers–Cohen–Mellit. This involves Gauss sums, and so must be computed in a context where roots of unity are available. Working exactly in a large cyclotomic field is impractical. Using floating-point interval arithmetic is possible but less efficient than...

... using the Gross-Koblitz formula to efficiently compute Gauss sums in a p-adic field. This leads to the algorithm implemented in `Magma`/`Sage`.

A hypergeometric trace formula

In this approach, we compute a good Euler factor of an HG motive by computing the trace of the q-Frobenius for each prime power q (on ℓ-adic cohomology for some prime ℓ not dividing q; the exact choice is irrelevant).

Building on work of Greene, Katz, McCarthy, Ono, ..., an explicit formula for this trace was given by Beukers–Cohen–Mellit. This involves Gauss sums, and so must be computed in a context where roots of unity are available. Working exactly in a large cyclotomic field is impractical. Using floating-point interval arithmetic is possible but less efficient than...

... using the Gross-Koblitz formula to efficiently compute Gauss sums in a p-adic field. This leads to the algorithm implemented in `Magma`/`Sage`.

A hypergeometric trace formula

In this approach, we compute a good Euler factor of an HG motive by computing the trace of the q-Frobenius for each prime power q (on ℓ-adic cohomology for some prime ℓ not dividing q; the exact choice is irrelevant).

Building on work of Greene, Katz, McCarthy, Ono, ..., an explicit formula for this trace was given by Beukers–Cohen–Mellit. This involves Gauss sums, and so must be computed in a context where roots of unity are available. Working exactly in a large cyclotomic field is impractical. Using floating-point interval arithmetic is possible but less efficient than...

... using the Gross-Koblitz formula to efficiently compute Gauss sums in a p-adic field. This leads to the algorithm implemented in `Magma/Sage`.

The Beukers–Cohen–Mellit formula (part 1)

Define positive integers $p_1, \dots, p_r, q_1, \dots, q_s$ (and r and s) by

$$\prod_{i=1}^{n} \frac{x - e^{2\pi i \alpha_i}}{x - e^{2\pi i \beta_i}} = \frac{\prod_{j=1}^{r} (x^{p_j} - 1)}{\prod_{k=1}^{s} (x^{q_k} - 1)},$$

then set $M := \prod_{j=1}^{r} p_j^{p_j} \prod_{k=1}^{s} q_k^{-q_k}$. Fix a generator ω of the complex character group of $\mathbb{F}_q^\times$ and a nontrivial additive character ψ_q of $\mathbb{F}_q$. For $m \in \mathbb{Z}$, define the Gauss sum

$$g(m) := g(\omega^m) = \sum_{x \in \mathbb{F}_q^\times} \omega^m(x) \psi_q(x).$$

The Beukers–Cohen–Mellit formula (part 2)

Set

$$g(\mathbf{p}m, -\mathbf{q}m) := \prod_{j=1}^{r} g(p_j m) \prod_{k=1}^{s} g(-q_k m).$$

Put $\epsilon := (-1)^{\sum_k q_k}$ and let $s(m)$ denote the multiplicity of $e^{2\pi i m/(j-1)}$ as a root of $\gcd(\prod_{j=1}^{r}(x^{p_j} - 1), \prod_{k=1}^{s}(x^{q_k} - 1))$. The formula now states

$$H_q(\underline{\alpha}, \underline{\beta}|t) = \frac{(-1)^{r+s}}{1-q} \sum_{m=0}^{q-2} q^{-s(0)+s(m)} g(\mathbf{p}m, -\mathbf{q}m)\omega(\epsilon M^{-1} t)^m.$$

The Gross–Koblitz formula (part 1)

To compute the Beukers–Cohen–Mellit formula p-adically, one must replace ω with a p-adic character of $\mathbb{F}_q^\times$; it is natural to use the canonical one given by Teichmüller lifting. There is also a convenient choice of ψ_q, namely the composition of ω with the **Dwork exponential series**

$$\Theta_q(x) = \exp(\pi(x - x^q))$$

where $\pi \in \mathbb{C}_p$ satisfies $\pi^{p-1} = -p$. This series has radius of convergence > 1; remember this for later.

Exercises:

- Check that $\mathbb{Q}_p(\pi) = \mathbb{Q}_p(\zeta_p)$.
- Check that $\Theta_q(x_1 + x_2) = \Theta_q(x_1)\Theta_q(x_2)$ if $x_1^q = x_1, x_2^q = x_2$. (Hint: compute in the ring $\mathbb{Q}_p[\pi, x_1, x_2][[x]]/(x_1^q - x_1, x_2^q - x_2)$.)

The Gross–Koblitz formula (part 1)

To compute the Beukers–Cohen–Mellit formula p-adically, one must replace ω with a p-adic character of $\mathbb{F}_q^\times$; it is natural to use the canonical one given by Teichmüller lifting. There is also a convenient choice of ψ_q, namely the composition of ω with the **Dwork exponential series**

$$\Theta_q(x) = \exp(\pi(x - x^q))$$

where $\pi \in \mathbb{C}_p$ satisfies $\pi^{p-1} = -p$. This series has radius of convergence > 1; remember this for later.

Exercises:

- Check that $\mathbb{Q}_p(\pi) = \mathbb{Q}_p(\zeta_p)$.
- Check that $\Theta_q(x_1 + x_2) = \Theta_q(x_1)\Theta_q(x_2)$ if $x_1^q = x_1, x_2^q = x_2$. (Hint: compute in the ring $\mathbb{Q}_p[\pi, x_1, x_2][[x]]/(x_1^q - x_1, x_2^q - x_2)$.)

The Gross–Koblitz formula (part 1)

To compute the Beukers–Cohen–Mellit formula p-adically, one must replace ω with a p-adic character of $\mathbb{F}_q^\times$; it is natural to use the canonical one given by Teichmüller lifting. There is also a convenient choice of ψ_q, namely the composition of ω with the **Dwork exponential series**

$$\Theta_q(x) = \exp(\pi(x - x^q))$$

where $\pi \in \mathbb{C}_p$ satisfies $\pi^{p-1} = -p$. This series has radius of convergence > 1; remember this for later.

Exercises:

- Check that $\mathbb{Q}_p(\pi) = \mathbb{Q}_p(\zeta_p)$.
- Check that $\Theta_q(x_1 + x_2) = \Theta_q(x_1)\Theta_q(x_2)$ if $x_1^q = x_1, x_2^q = x_2$. (Hint: compute in the ring $\mathbb{Q}_p[\pi, x_1, x_2][[x]]/(x_1^q - x_1, x_2^q - x_2)$.)

The Gross–Koblitz formula (part 2)

The Morita p-adic Gamma function is the unique continuous function $\Gamma_p : \mathbb{Z}_p \to \mathbb{Z}_p$ satisfying

$$\begin{aligned}\Gamma_p(0) &= 1 \\ \Gamma_p(x+1) &= -x\Gamma_p(x) \qquad (x \not\equiv 0 \pmod{p}) \\ \Gamma_p(x+1) &= -\Gamma_p(x) \qquad (x \equiv 0 \pmod{p}).\end{aligned}$$

The Gross–Koblitz formula asserts that for $q = p^f$, for $m = 0, \dots, q-2$,

$$\sum_{x \in \mathbb{F}_q^\times} \omega^{-m}(x)\Theta_q(\omega(x)) = -\pi^{S_p(m)} \prod_{i=0}^{f-1} \Gamma_p\left(\frac{m^{(i)}}{q-1}\right)$$

where $S_p(m)$ denotes the sum of the p-adic digits of m and $m^{(i)}$ is the remainder of mp^i modulo $q-1$ (i.e., a cyclic digital shift of m).

The Gross–Koblitz formula (part 2)

The Morita p-adic Gamma function is the unique continuous function $\Gamma_p : \mathbb{Z}_p \to \mathbb{Z}_p$ satisfying

$$\begin{aligned} \Gamma_p(0) &= 1 \\ \Gamma_p(x+1) &= -x\Gamma_p(x) \qquad (x \not\equiv 0 \pmod{p}) \\ \Gamma_p(x+1) &= -\Gamma_p(x) \qquad (x \equiv 0 \pmod{p}). \end{aligned}$$

The Gross–Koblitz formula asserts that for $q = p^f$, for $m = 0, \dots, q-2$,

$$\sum_{x \in \mathbb{F}_q^\times} \omega^{-m}(x)\Theta_q(\omega(x)) = -\pi^{S_p(m)} \prod_{i=0}^{f-1} \Gamma_p\left(\frac{m^{(i)}}{q-1}\right)$$

where $S_p(m)$ denotes the sum of the p-adic digits of m and $m^{(i)}$ is the remainder of mp^i modulo $q-1$ (i.e., a cyclic digital shift of m).

Pros and cons

The resulting formula involves terms which are efficient to compute.

However, the q-trace is a sum over $q-1$ terms. For computing full Euler factors, this yields exponential dependence on the degree n: given the sign of the functional equation, for any given p we need to take $q = p^f$ for $f = 1, \dots, \lfloor \frac{n}{2} \rfloor$. (For computing N coefficients of the Dirichlet series, the degree is immaterial because one only needs to consider $q \leq N$.)

Moreover, there is no obvious way to convert a sum over $q-1$ terms into an **average polynomial time** in the sense of Harvey–Sutherland, in which one computes the first N coefficients of the Dirichlet series in time $O(N \operatorname{polylog} N)$.

Both issues can (potentially) be circumvented using Frobenius structures.

Pros and cons

The resulting formula involves terms which are efficient to compute.

However, the q-trace is a sum over $q-1$ terms. For computing full Euler factors, this yields exponential dependence on the degree n: given the sign of the functional equation, for any given p we need to take $q = p^f$ for $f = 1, \dots, \lfloor \frac{n}{2} \rfloor$. (For computing N coefficients of the Dirichlet series, the degree is immaterial because one only needs to consider $q \leq N$.)

Moreover, there is no obvious way to convert a sum over $q-1$ terms into an **average polynomial time** in the sense of Harvey–Sutherland, in which one computes the first N coefficients of the Dirichlet series in time $O(N \operatorname{polylog} N)$.

Both issues can (potentially) be circumvented using Frobenius structures.

Pros and cons

The resulting formula involves terms which are efficient to compute.

However, the q-trace is a sum over $q-1$ terms. For computing full Euler factors, this yields exponential dependence on the degree n: given the sign of the functional equation, for any given p we need to take $q = p^f$ for $f = 1, \dots, \lfloor \frac{n}{2} \rfloor$. (For computing N coefficients of the Dirichlet series, the degree is immaterial because one only needs to consider $q \leq N$.)

Moreover, there is no obvious way to convert a sum over $q-1$ terms into an **average polynomial time** in the sense of Harvey–Sutherland, in which one computes the first N coefficients of the Dirichlet series in time $O(N \operatorname{polylog} N)$.

Both issues can (potentially) be circumvented using Frobenius structures.

Pros and cons

The resulting formula involves terms which are efficient to compute.

However, the q-trace is a sum over $q-1$ terms. For computing full Euler factors, this yields exponential dependence on the degree n: given the sign of the functional equation, for any given p we need to take $q = p^f$ for $f = 1, \dots, \lfloor \frac{n}{2} \rfloor$. (For computing N coefficients of the Dirichlet series, the degree is immaterial because one only needs to consider $q \leq N$.)

Moreover, there is no obvious way to convert a sum over $q-1$ terms into an **average polynomial time** in the sense of Harvey–Sutherland, in which one computes the first N coefficients of the Dirichlet series in time $O(N \operatorname{polylog} N)$.

Both issues can (potentially) be circumvented using Frobenius structures.

Contents

1 Motives and L-functions

2 Hypergeometric motives and their L-functions

3 Finite hypergeometric sums and trace formulas

4 Frobenius structures

5 GKZ A-hypergeometric systems

Frobenius structures on differential equations

In the following discussion, let L be the completion of $\mathbb{Q}_p(z)$ for the Gauss norm. (This is called the **field of analytic elements** in z over $\mathbb{Q}_p$.)

Consider a first-order differential system $N\mathbf{v} + \frac{d}{dz}(\mathbf{v}) = 0$ where N is an $n \times n$ matrix over $\mathbb{Q}_p(z)$. (Conversion of linear ODEs to first-order systems left as an exercise.) A **Frobenius structure** with respect to the Frobenius lift $\sigma : z \mapsto z^p$ is a σ-linear endomorphism of L^n (identified with column vectors) given by an $n \times n$ matrix A over L (whose columns are the images of the standard basis vectors) satisfying

$$NA + \frac{d}{dz}(A) = pz^{p-1}AN.$$

In other words, a Frobenius structure is an isomorphism of the associated connection with its σ-pullback. (Here $pz^{p-1} = \frac{d\sigma(z)}{dz}$.)

Frobenius structures on differential equations

In the following discussion, let L be the completion of $\mathbb{Q}_p(z)$ for the Gauss norm. (This is called the **field of analytic elements** in z over $\mathbb{Q}_p$.)

Consider a first-order differential system $N\mathbf{v} + \frac{d}{dz}(\mathbf{v}) = 0$ where N is an $n \times n$ matrix over $\mathbb{Q}_p(z)$. (Conversion of linear ODEs to first-order systems left as an exercise.) A **Frobenius structure** with respect to the Frobenius lift $\sigma : z \mapsto z^p$ is a σ-linear endomorphism of L^n (identified with column vectors) given by an $n \times n$ matrix A over L (whose columns are the images of the standard basis vectors) satisfying

$$NA + \frac{d}{dz}(A) = pz^{p-1}AN.$$

In other words, a Frobenius structure is an isomorphism of the associated connection with its σ-pullback. (Here $pz^{p-1} = \frac{d\sigma(z)}{dz}$.)

Frobenius structures on differential equations

In the following discussion, let L be the completion of $\mathbb{Q}_p(z)$ for the Gauss norm. (This is called the **field of analytic elements** in z over $\mathbb{Q}_p$.)

Consider a first-order differential system $N\mathbf{v} + \frac{d}{dz}(\mathbf{v}) = 0$ where N is an $n \times n$ matrix over $\mathbb{Q}_p(z)$. (Conversion of linear ODEs to first-order systems left as an exercise.) A **Frobenius structure** with respect to the Frobenius lift $\sigma : z \mapsto z^p$ is a σ-linear endomorphism of L^n (identified with column vectors) given by an $n \times n$ matrix A over L (whose columns are the images of the standard basis vectors) satisfying

$$NA + \frac{d}{dz}(A) = pz^{p-1}AN.$$

In other words, a Frobenius structure is an isomorphism of the associated connection with its σ-pullback. (Here $pz^{p-1} = \frac{d\sigma(z)}{dz}$.)

The HG equation and its Frobenius structure

The hypergeometric equation (with $D = z\frac{d}{dz}$) is

$$(z(D+\alpha_1)\cdots(D+\alpha_n) - (D+\beta_1-1)\cdots(D+\beta_n-1))(y) = 0;$$

this is a Picard–Fuchs equation associated to the family of HGMs. It admits a Frobenius structure F which is meromorphic on $|z-1| > 1-\epsilon$ for some $\epsilon > 0$ with singularities only at $z = 0, \infty$. (This follows from general nonsense about crystalline cohomology, but we will use a more concrete construction to be sketched later.)

For $t \in \mathbb{Q}$ for which p is good, the Euler factor at p of the associated HGM is $\det(1 - p^{-s}F(\omega(\overline{t})))$, where $\overline{t} \in \mathbb{F}_p \setminus \{0,1\}$ is the reduction of t mod p.

The HG equation and its Frobenius structure

The hypergeometric equation (with $D = z\frac{d}{dz}$) is

$$(z(D+\alpha_1)\cdots(D+\alpha_n) - (D+\beta_1-1)\cdots(D+\beta_n-1))\,(y) = 0;$$

this is a Picard–Fuchs equation associated to the family of HGMs. It admits a Frobenius structure F which is meromorphic on $|z-1| > 1-\epsilon$ for some $\epsilon > 0$ with singularities only at $z = 0, \infty$. (This follows from general nonsense about crystalline cohomology, but we will use a more concrete construction to be sketched later.)

For $t \in \mathbb{Q}$ for which p is good, the Euler factor at p of the associated HGM is $\det(1 - p^{-s}F(\omega(\overline{t})))$, where $\overline{t} \in \mathbb{F}_p \setminus \{0,1\}$ is the reduction of t mod p.

Series solutions of the HG equation

In terms of the rising Pochhammer symbol

$$(\alpha)_n := \alpha(\alpha+1)\cdots(\alpha+n-1)$$

we define the **Clausen–Thomae hypergeometric series**

$$ {}_nF_{n-1}\left(\left.\begin{matrix}\alpha_1,\dots,\alpha_n\\ \beta_1,\dots,\beta_{n-1}\end{matrix}\right| z\right) := \sum_{k=0}^{\infty} \frac{(\alpha_1)_k\cdots(\alpha_n)_k}{(\beta_1)_k\cdots(\beta_{n-1})_k}\frac{z^k}{k!}$$

If $\beta_1,\dots,\beta_n$ are pairwise distinct, then the HG equation has a $\mathbb{C}$-basis of solutions in the Puiseux field $\bigcup_{m=1}^{\infty}\mathbb{C}((z^{1/m}))$ given by (for $i=1,\dots,n$)

$$z^{1-\beta_i}{}_nF_{n-1}\left(\left.\begin{matrix}\alpha_1-\beta_i+1,\dots,\alpha_n-\beta_i+1\\ \beta_1-\beta_i+1,\dots,\widehat{\beta_i-\beta_i+1},\dots,\beta_n-\beta_i+1\end{matrix}\right| z\right).$$

An algorithmic approach

The compatibility between the matrices A and N implies that one can compute the Frobenius structure using formal solutions, as long as one knows the "initial condition" at $z = 0$. That turns out to be given by some explicit combination of p-adic Gamma functions (more on this below).

What one gets from the previous paragraph is a representation of F using Puiseux series (truncated in both the p-adic and z-adic directions; some care is needed here). One can then recognize the resulting series as representations of elements of L (approximated p-adically by elements of $\mathbb{Q}_p(z)$), which one can then evaluate at $\omega(\bar{t})$.

I have a toy `Sage` implementation of this. It is slower than `Magma` for n small, but for $n = 10$ it is significantly faster! (This is expected because we have eliminated the exponential dependence on n.)

An algorithmic approach

The compatibility between the matrices A and N implies that one can compute the Frobenius structure using formal solutions, as long as one knows the "initial condition" at $z = 0$. That turns out to be given by some explicit combination of p-adic Gamma functions (more on this below).

What one gets from the previous paragraph is a representation of F using Puiseux series (truncated in both the p-adic and z-adic directions; some care is needed here). One can then recognize the resulting series as representations of elements of L (approximated p-adically by elements of $\mathbb{Q}_p(z)$), which one can then evaluate at $\omega(\bar{t})$.

I have a toy Sage implementation of this. It is slower than Magma for n small, but for $n = 10$ it is significantly faster! (This is expected because we have eliminated the exponential dependence on n.)

An algorithmic approach

The compatibility between the matrices A and N implies that one can compute the Frobenius structure using formal solutions, as long as one knows the "initial condition" at $z = 0$. That turns out to be given by some explicit combination of p-adic Gamma functions (more on this below).

What one gets from the previous paragraph is a representation of F using Puiseux series (truncated in both the p-adic and z-adic directions; some care is needed here). One can then recognize the resulting series as representations of elements of L (approximated p-adically by elements of $\mathbb{Q}_p(z)$), which one can then evaluate at $\omega(\bar{t})$.

I have a toy `Sage` implementation of this. It is slower than `Magma` for n small, but for $n = 10$ it is significantly faster! (This is expected because we have eliminated the exponential dependence on n.)

About that initial condition...

To study the Frobenius structure at $z = 0$, I first did some numerical experiments (which I won't describe here) to get a candidate formula. I then tried to extract information from Dwork's *Generalized Hypergeometric Functions*. (Don't try this at home!)

I am now pretty sure (but have not finished checking) that Dwork's methods do suffice to prove my guessed formula. In fact, this yields a formula that makes sense for any pairwise distinct $\beta_1, \ldots, \beta_n \in \mathbb{Z}_p$ (and even without pairwise distinctness, under a suitable interpretation).

However, one also obtains an alternative algorithm for computing the Frobenius structure directly. This approach turns out to have surprising links to previously studied algorithms...

About that initial condition...

To study the Frobenius structure at $z = 0$, I first did some numerical experiments (which I won't describe here) to get a candidate formula. I then tried to extract information from Dwork's *Generalized Hypergeometric Functions*. (Don't try this at home!)

I am now pretty sure (but have not finished checking) that Dwork's methods do suffice to prove my guessed formula. In fact, this yields a formula that makes sense for any pairwise distinct $\beta_1, \dots, \beta_n \in \mathbb{Z}_p$ (and even without pairwise distinctness, under a suitable interpretation).

However, one also obtains an alternative algorithm for computing the Frobenius structure directly. This approach turns out to have surprising links to previously studied algorithms...

About that initial condition...

To study the Frobenius structure at $z = 0$, I first did some numerical experiments (which I won't describe here) to get a candidate formula. I then tried to extract information from Dwork's *Generalized Hypergeometric Functions*. (Don't try this at home!)

I am now pretty sure (but have not finished checking) that Dwork's methods do suffice to prove my guessed formula. In fact, this yields a formula that makes sense for any pairwise distinct $\beta_1, \dots, \beta_n \in \mathbb{Z}_p$ (and even without pairwise distinctness, under a suitable interpretation).

However, one also obtains an alternative algorithm for computing the Frobenius structure directly. This approach turns out to have surprising links to previously studied algorithms...

Contents

1 Motives and L-functions

2 Hypergeometric motives and their L-functions

3 Finite hypergeometric sums and trace formulas

4 Frobenius structures

5 GKZ A-hypergeometric systems

A PDE associated to the HG equation

A function $\Phi(\underline{x} := x_1, \ldots, x_n; \underline{y} := y_1, \ldots, y_n)$ is killed by **Euler operators**

$$x_j \frac{\partial}{\partial x_j} + y_k \frac{\partial}{\partial y_k} + \alpha_j - \beta_k + 1 \qquad (j, k = 1, \ldots, n)$$

if and only if there exists a univariate function $f(z)$ such that

$$\Phi(\underline{x}, \underline{y}) = x_1^{-\alpha_1} \cdots x_n^{-\alpha_n} y_1^{\beta_1 - 1} \cdots y_n^{\beta_n - 1} f((-1)^n x_1^{-1} \cdots x_n^{-1} y_1 \cdots y_n).$$

In this case, Φ is also killed by the **toric operator**

$$\prod_{j=1}^{n} \frac{\partial}{\partial x_j} - \prod_{j=1}^{n} \frac{\partial}{\partial y_j}.$$

if and only if $y = f(z)$ is a solution of the HG equation

$$(z(D + \alpha_1) \cdots (D + \alpha_n) - (D + \beta_1 - 1) \cdots (D + \beta_n - 1))(y) = 0.$$

A PDE associated to the HG equation

A function $\Phi(\underline{x} := x_1, \dots, x_n; \underline{y} := y_1, \dots, y_n)$ is killed by **Euler operators**

$$x_j \frac{\partial}{\partial x_j} + y_k \frac{\partial}{\partial y_k} + \alpha_j - \beta_k + 1 \qquad (j, k = 1, \dots, n)$$

if and only if there exists a univariate function $f(z)$ such that

$$\Phi(\underline{x}, \underline{y}) = x_1^{-\alpha_1} \cdots x_n^{-\alpha_n} y_1^{\beta_1 - 1} \cdots y_n^{\beta_n - 1} f((-1)^n x_1^{-1} \cdots x_n^{-1} y_1 \cdots y_n).$$

In this case, Φ is also killed by the **toric operator**

$$\prod_{j=1}^{n} \frac{\partial}{\partial x_j} - \prod_{j=1}^{n} \frac{\partial}{\partial y_j}.$$

if and only if $y = f(z)$ is a solution of the HG equation

$$(z(D + \alpha_1) \cdots (D + \alpha_n) - (D + \beta_1 - 1) \cdots (D + \beta_n - 1))(y) = 0.$$

GKZ A-hypergeometric systems

Define the Weyl algebra

$$W_m := \mathbb{C}\left\langle x_1, \dots, x_m, \partial_1 := \frac{\partial}{\partial x_1}, \dots, \partial_m := \frac{\partial}{\partial x_m} \right\rangle.$$

Let A be a $d \times m$ matrix over $\mathbb{Z}$. The associated **toric ideal** in W_m is the left ideal generated by

$$\{\partial_1^{u_1} \cdots \partial_m^{u_m} - \partial_1^{v_m} \cdots \partial_1^{v_m} : u, v \in \mathbb{Z}_{\geq 0}^m, A(u - v) = 0\}.$$

For $\delta \in \mathbb{C}^d$, add to this left ideal the **Euler operators**

$$A_{i1} x_1 \partial_1 + \cdots A_{im} x_m \partial_m - \delta_m \qquad (i = 1, \dots, d)$$

to obtain the **GKZ ideal** associated to A, δ. (GKZ is short for Gelfand–Kapranov–Zelevinsky; independently discovered by Dwork.)

GKZ A-hypergeometric systems

Define the Weyl algebra

$$W_m := \mathbb{C}\left\langle x_1, \dots, x_m, \partial_1 := \frac{\partial}{\partial x_1}, \dots, \partial_m := \frac{\partial}{\partial x_m} \right\rangle.$$

Let A be a $d \times m$ matrix over $\mathbb{Z}$. The associated **toric ideal** in W_m is the left ideal generated by

$$\{\partial_1^{u_1} \cdots \partial_m^{u_m} - \partial_1^{v_m} \cdots \partial_1^{v_m} : u, v \in \mathbb{Z}_{\geq 0}^m, A(u-v) = 0\}.$$

For $\delta \in \mathbb{C}^d$, add to this left ideal the **Euler operators**

$$A_{i1}x_1\partial_1 + \cdots A_{im}x_m\partial_m - \delta_m \qquad (i = 1, \dots, d)$$

to obtain the **GKZ ideal** associated to A, δ. (GKZ is short for Gelfand–Kapranov–Zelevinsky; independently discovered by Dwork.)

GKZ A-hypergeometric systems

Define the Weyl algebra

$$W_m := \mathbb{C}\left\langle x_1, \dots, x_m, \partial_1 := \frac{\partial}{\partial x_1}, \dots, \partial_m := \frac{\partial}{\partial x_m} \right\rangle.$$

Let A be a $d \times m$ matrix over $\mathbb{Z}$. The associated **toric ideal** in W_m is the left ideal generated by

$$\{\partial_1^{u_1} \cdots \partial_m^{u_m} - \partial_1^{v_m} \cdots \partial_1^{v_m} : u, v \in \mathbb{Z}^m_{\geq 0}, A(u-v) = 0\}.$$

For $\delta \in \mathbb{C}^d$, add to this left ideal the **Euler operators**

$$A_{i1}x_1\partial_1 + \cdots A_{im}x_m\partial_m - \delta_m \qquad (i = 1, \dots, d)$$

to obtain the **GKZ ideal** associated to A, δ. (GKZ is short for Gelfand–Kapranov–Zelevinsky; independently discovered by Dwork.)

A dual construction

Let R be the $\mathbb{C}$-subalgebra of $\mathbb{C}[X_1^{\pm}, \dots, X_n^{\pm}]$ generated by

$$X^{(j)} := X_1^{A_{1j}} \cdots X_d^{A_{dj}} \qquad (j = 1, \dots, m).$$

Fix $\pi \in \mathbb{C}^\times$ (take it to be 1 for the moment). View $R[x] = R[x_1, \dots, x_m]$ as a left W_m-module by adding $\pi X^{(j)}$ to the "natural" ∂_j-action.

Theorem (Dwork)

As a left W_m-module, the quotient $R[x]/\sum_{i=1}^d D_{\delta,i} R[x]$, where

$$g := \pi^{-1} \sum_{j=1}^m x_j X^{(j)} \in R[x],$$

$$D_{\delta,i} = X_i \frac{\partial}{\partial X_i} + \delta_i + X_i \frac{\partial g}{\partial X_i},$$

is isomorphic to the quotient of W_m by the GKZ ideal.

A dual construction

Let R be the $\mathbb{C}$-subalgebra of $\mathbb{C}[X_1^{\pm}, \dots, X_n^{\pm}]$ generated by

$$X^{(j)} := X_1^{A_{1j}} \cdots X_d^{A_{dj}} \qquad (j = 1, \dots, m).$$

Fix $\pi \in \mathbb{C}^\times$ (take it to be 1 for the moment). View $R[x] = R[x_1, \dots, x_m]$ as a left W_m-module by adding $\pi X^{(j)}$ to the "natural" ∂_j-action.

Theorem (Dwork)

As a left W_m-module, the quotient $R[x]/\sum_{i=1}^d D_{\delta,i} R[x]$, where

$$g := \pi^{-1} \sum_{j=1}^m x_j X^{(j)} \in R[x],$$

$$D_{\delta,i} = X_i \frac{\partial}{\partial X_i} + \delta_i + X_i \frac{\partial g}{\partial X_i},$$

is isomorphic to the quotient of W_m by the GKZ ideal.

A dual construction

Let R be the $\mathbb{C}$-subalgebra of $\mathbb{C}[X_1^{\pm}, \dots, X_n^{\pm}]$ generated by

$$X^{(j)} := X_1^{A_{1j}} \cdots X_d^{A_{dj}} \qquad (j = 1, \dots, m).$$

Fix $\pi \in \mathbb{C}^\times$ (take it to be 1 for the moment). View $R[x] = R[x_1, \dots, x_m]$ as a left W_m-module by adding $\pi X^{(j)}$ to the "natural" ∂_j-action.

Theorem (Dwork)

As a left W_m-module, the quotient $R[x]/\sum_{i=1}^d D_{\delta,i} R[x]$, where

$$g := \pi^{-1} \sum_{j=1}^m x_j X^{(j)} \in R[x],$$

$$D_{\delta,i} = X_i \frac{\partial}{\partial X_i} + \delta_i + X_i \frac{\partial g}{\partial X_i},$$

is isomorphic to the quotient of W_m by the GKZ ideal.

The dual construction of Frobenius structures (part 1)

Using the previous construction, Dwork obtains Frobenius structures on GKZ systems (and by restriction on HG equations) as follows. Let $\varphi : R[x] \to R[x]$ be the $\mathbb{C}$-linear substitution

$$x_i \mapsto x_i^p, \qquad X_j \mapsto X_j^p.$$

Since

$$D_{p\delta,i} \circ \varphi = p\varphi \circ D_{\delta,i},$$

φ induces a self-map on $R[x]/\sum_{i=1}^d D_{\delta,i}R[x]$. This becomes compatible with W_m-actions if we define the action of ∂_i on the target with respect to pA rather than A.

The dual construction of Frobenius structures (part 1)

Using the previous construction, Dwork obtains Frobenius structures on GKZ systems (and by restriction on HG equations) as follows. Let $\varphi : R[x] \to R[x]$ be the $\mathbb{C}$-linear substitution

$$x_i \mapsto x_i^p, \qquad X_j \mapsto X_j^p.$$

Since

$$D_{p\delta,i} \circ \varphi = p\varphi \circ D_{\delta,i},$$

φ induces a self-map on $R[x]/\sum_{i=1}^d D_{\delta,i}R[x]$. This becomes compatible with W_m-actions if we define the action of ∂_i on the target with respect to pA rather than A.

The dual construction of Frobenius structures (part 2)

Let $\tilde{R}$ be the ring obtained by $R[x]$ by adjoining symbols

$$E_j := \exp(\pi(X^{(j)} - (X^{(j)})^p)) \qquad (j = 1, \dots, m).$$

We may extend ∂_i and $\frac{\partial}{\partial X_j}$ to $\tilde{R}$ in the obvious fashion.

Taking the map induced by

$$f \mapsto E_1 \cdots E_m \varphi(f),$$

then dividing the action of ∂_i by p, yields a W_m-equivariant map $\tilde{R}/\sum_{i=1}^{d} D_{\delta,i}\tilde{R}$.

If we replace polynomials in the x_i over $\mathbb{C}$ with analytic functions over $\mathbb{Q}_p$, then the symbols E_j are already present thanks to the Dwork exponential series. We thus obtain the desired Frobenius structure.

The dual construction of Frobenius structures (part 2)

Let $\tilde{R}$ be the ring obtained by $R[x]$ by adjoining symbols

$$E_j := \exp(\pi(X^{(j)} - (X^{(j)})^p)) \qquad (j = 1, \dots, m).$$

We may extend ∂_i and $\frac{\partial}{\partial X_j}$ to $\tilde{R}$ in the obvious fashion.

Taking the map induced by

$$f \mapsto E_1 \cdots E_m \varphi(f),$$

then dividing the action of ∂_i by p, yields a W_m-equivariant map $\tilde{R}/\sum_{i=1}^d D_{\delta,i}\tilde{R}$.

If we replace polynomials in the x_i over $\mathbb{C}$ with analytic functions over $\mathbb{Q}_p$, then the symbols E_j are already present thanks to the Dwork exponential series. We thus obtain the desired Frobenius structure.

The dual construction of Frobenius structures (part 2)

Let $\tilde{R}$ be the ring obtained by $R[x]$ by adjoining symbols

$$E_j := \exp(\pi(X^{(j)} - (X^{(j)})^p)) \qquad (j = 1, \dots, m).$$

We may extend ∂_i and $\frac{\partial}{\partial X_j}$ to $\tilde{R}$ in the obvious fashion.

Taking the map induced by

$$f \mapsto E_1 \cdots E_m \varphi(f),$$

then dividing the action of ∂_i by p, yields a W_m-equivariant map $\tilde{R} / \sum_{i=1}^{d} D_{\delta,i} \tilde{R}$.

If we replace polynomials in the x_i over $\mathbb{C}$ with analytic functions over $\mathbb{Q}_p$, then the symbols E_j are already present thanks to the Dwork exponential series. We thus obtain the desired Frobenius structure.

Reduction in the quotient

Computing the Frobenius structure in this fashion requires:

- finding a basis for $R[x]/\sum_{i=1}^{d} D_{\delta,i}R[x]$;
- reducing elements of $R[x]$ to basis combinations in the quotient.

(These steps commute with specialization of $x_1, \ldots, x_m$.)

Let $\Delta \subseteq \mathbb{R}^d$ be the convex hull of the origin plus the column vectors of A. Define a filtration on $R[x]$ so that $R[x]_e$ consists of the $\mathbb{C}[x_1, \ldots, x_m]$-multiples of those monomials $X_1^{c_1} \cdots X_d^{c_d}$ with $(c_1, \ldots, c_d) \in e\Delta$. (In general, e may be fractional.)

On the associated graded ring, $D_{\delta,i}$ acts by shifting degrees up by 1 and multiplying by $X_i \frac{\partial g}{\partial X_i}$. In particular, finding a basis for the above quotient amounts to computing the (toric) **Jacobian ring** of $g.x$

Reduction in the quotient

Computing the Frobenius structure in this fashion requires:

- finding a basis for $R[x]/\sum_{i=1}^{d} D_{\delta,i}R[x]$;
- reducing elements of $R[x]$ to basis combinations in the quotient.

(These steps commute with specialization of $x_1, \dots, x_m$.)

Let $\Delta \subseteq \mathbb{R}^d$ be the convex hull of the origin plus the column vectors of A. Define a filtration on $R[x]$ so that $R[x]_e$ consists of the $\mathbb{C}[x_1, \dots, x_m]$-multiples of those monomials $X_1^{c_1} \cdots X_d^{c_d}$ with $(c_1, \dots, c_d) \in e\Delta$. (In general, e may be fractional.)

On the associated graded ring, $D_{\delta,i}$ acts by shifting degrees up by 1 and multiplying by $X_i \frac{\partial g}{\partial X_i}$. In particular, finding a basis for the above quotient amounts to computing the (toric) **Jacobian ring** of g_x

Reduction in the quotient

Computing the Frobenius structure in this fashion requires:

- finding a basis for $R[x]/\sum_{i=1}^{d} D_{\delta,i} R[x]$;
- reducing elements of $R[x]$ to basis combinations in the quotient.

(These steps commute with specialization of $x_1, \dots, x_m$.)

Let $\Delta \subseteq \mathbb{R}^d$ be the convex hull of the origin plus the column vectors of A. Define a filtration on $R[x]$ so that $R[x]_e$ consists of the $\mathbb{C}[x_1, \dots, x_m]$-multiples of those monomials $X_1^{c_1} \cdots X_d^{c_d}$ with $(c_1, \dots, c_d) \in e\Delta$. (In general, e may be fractional.)

On the associated graded ring, $D_{\delta,i}$ acts by shifting degrees up by 1 and multiplying by $X_i \frac{\partial g}{\partial X_i}$. In particular, finding a basis for the above quotient amounts to computing the (toric) **Jacobian ring** of $g.x$

The punchline: controlled reduction

The resulting process strongly resembles the use of *controlled reduction* in the computation of the zeta function or the nondegenerate toric hypersurface $g = 0$ (for particular values of $x_1, \ldots, x_m$); see my ANTS-XIII paper with Costa and Harvey. That paper describes some worked examples where the ambient space has dimension ≤ 5. We have not yet attempted to adapt to this setting, but this should be straightforward.

The ANTS paper does not address average polynomial time, but this should be feasible. This may in turn facilitate developing an average polynomial time method for HGMs. (The approach via the initial condition can also be adapted to give average polynomial time.)

Side note: there is a related algorithm of Sperber–Voight for zeta functions of nondegenerate toric hypersurfaces which might also be adaptable to this setting, but I did not attempt to perform the adaptation.

The punchline: controlled reduction

The resulting process strongly resembles the use of *controlled reduction* in the computation of the zeta function or the nondegenerate toric hypersurface $g = 0$ (for particular values of $x_1, \dots, x_m$); see my ANTS-XIII paper with Costa and Harvey. That paper describes some worked examples where the ambient space has dimension ≤ 5. We have not yet attempted to adapt to this setting, but this should be straightforward.

The ANTS paper does not address average polynomial time, but this should be feasible. This may in turn facilitate developing an average polynomial time method for HGMs. (The approach via the initial condition can also be adapted to give average polynomial time.)

Side note: there is a related algorithm of Sperber–Voight for zeta functions of nondegenerate toric hypersurfaces which might also be adaptable to this setting, but I did not attempt to perform the adaptation.

The punchline: controlled reduction

The resulting process strongly resembles the use of *controlled reduction* in the computation of the zeta function or the nondegenerate toric hypersurface $g = 0$ (for particular values of $x_1, \dots, x_m$); see my ANTS-XIII paper with Costa and Harvey. That paper describes some worked examples where the ambient space has dimension ≤ 5. We have not yet attempted to adapt to this setting, but this should be straightforward.

The ANTS paper does not address average polynomial time, but this should be feasible. This may in turn facilitate developing an average polynomial time method for HGMs. (The approach via the initial condition can also be adapted to give average polynomial time.)

Side note: there is a related algorithm of Sperber–Voight for zeta functions of nondegenerate toric hypersurfaces which might also be adaptable to this setting, but I did not attempt to perform the adaptation.

Next steps: beyond HGM?

For arbitrary A, δ (satisfying some genericity conditions which I omit here), one can introduce the analogue of the Galois-stable condition: for any integer e coprime to the least common denominator of δ, the pairs A, δ and $A, e\delta$ define isomorphic GKZ systems (after inverting $x_1, \dots, x_m$).

Is there a family of motives in this situation? The de Rham realizations have been studied by Beukers and others; the ℓ-adic realizations were recently constructed by Lei Fu by analogy with Katz's construction, together with the analogue of finite hypergeometric sums and the trace formula. (This might be tricky to render in the motivic setting: one needs the motivic analogue of middle convolution of perverse sheaves over a higher-dimensional base.)

If so, this would give many new examples of L-functions. It might also provide new parametrizations giving access to objects of prior interest.

Next steps: beyond HGM?

For arbitrary A, δ (satisfying some genericity conditions which I omit here), one can introduce the analogue of the Galois-stable condition: for any integer e coprime to the least common denominator of δ, the pairs A, δ and $A, e\delta$ define isomorphic GKZ systems (after inverting $x_1, \dots, x_m$).

Is there a family of motives in this situation? The de Rham realizations have been studied by Beukers and others; the ℓ-adic realizations were recently constructed by Lei Fu by analogy with Katz's construction, together with the analogue of finite hypergeometric sums and the trace formula. (This might be tricky to render in the motivic setting: one needs the motivic analogue of middle convolution of perverse sheaves over a higher-dimensional base.)

If so, this would give many new examples of L-functions. It might also provide new parametrizations giving access to objects of prior interest.

Next steps: beyond HGM?

For arbitrary A, δ (satisfying some genericity conditions which I omit here), one can introduce the analogue of the Galois-stable condition: for any integer e coprime to the least common denominator of δ, the pairs A, δ and $A, e\delta$ define isomorphic GKZ systems (after inverting $x_1, \dots, x_m$).

Is there a family of motives in this situation? The de Rham realizations have been studied by Beukers and others; the ℓ-adic realizations were recently constructed by Lei Fu by analogy with Katz's construction, together with the analogue of finite hypergeometric sums and the trace formula. (This might be tricky to render in the motivic setting: one needs the motivic analogue of middle convolution of perverse sheaves over a higher-dimensional base.)

If so, this would give many new examples of L-functions. It might also provide new parametrizations giving access to objects of prior interest.

Fin

Thank you for your attention!