# Context-free languages in Algebraic Geometry and Number Theory

José Manuel Rodríguez Caballero

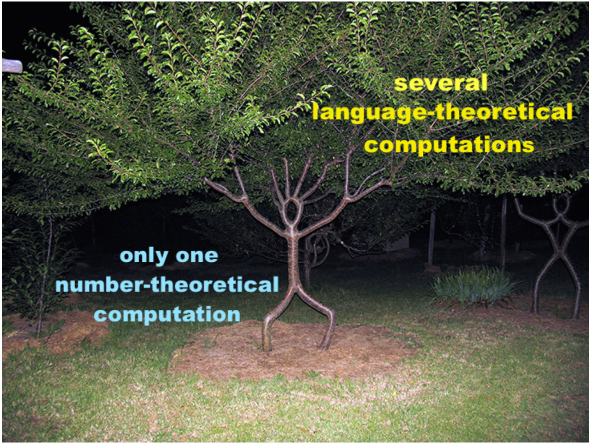
Ph.D. student
Laboratoire de combinatoire et d'informatique mathématique (LaCIM)
Université du Québec à Montréal (UQÀM)

Presentation at the LFANT Seminar
Institut de Mathématiques de Bordeaux (IMB)
Université de Bordeaux

October 24, 2017
work-in-progress

# Motivation

To simultaneously compute several arithmetical functions by means of several language-theoretical computations, but just doing one number-theoretical computation.



# Part I

## Kassel-Reutenauer polynomials

# Computational definition of $P_n(q)$

## Definition

For any integer $n \geq 1$, we define the $n$th *Kassel-Reutenauer polynomial* as follows :

$$P_n(q) := a_{n,0} + \sum_{k=1}^{n-1} a_{n,k} \left( q^{n-1+k} + q^{n-1-k} \right),$$

where $a_{n,k} := \# \left\{ d|n : \quad d - 2\frac{n}{d} \leq 2k < 2d - \frac{n}{d} \right\}$. Furthermore, we define $C_n(q) := (q-1)^2 \, P_n(q)$.

# Generating function

**Theorem**

$$\prod_{m=1}^{\infty} \frac{(1-t^m)^2}{(1-qt^m)(1-q^{-1}t^m)} = 1 + \sum_{n=1}^{\infty} \frac{C_n(q)}{q^n} t^n.$$

This result is essentially identity (9.2) in Nathan Jacob Fine, “Basic hypergeometric series and applications”, No. 27, American Mathematical Soc., 1988.

# Complex geometric interpretation of $C_n(q)$

Consider the Hilbert scheme $H_{\mathbb{C}}^n := \mathrm{Hilb}^n\left(\left(\mathbb{A}_{\mathbb{C}}^1\backslash\{0\}\right) \times \left(\mathbb{A}_{\mathbb{C}}^1\backslash\{0\}\right)\right)$ of $n$ points on the bidimensional complex torus.

**Theorem (Hausel, Letellier and Rodriguez-Villegas, 2011)**

*For each $n \geq 1$,the E-polynomial of $H_{\mathbb{C}}^n$ is $C_n(q)$.*

# Complex geometric interpretation of $P_n(q)$

Notice that $(\mathbb{C}\backslash\{0\}) \times (\mathbb{C}\backslash\{0\})$ acts naturally on

$$\left(\mathbb{A}^1_{\mathbb{C}}\backslash\{0\}\right) \times \left(\mathbb{A}^1_{\mathbb{C}}\backslash\{0\}\right).$$

This action induces an action of $(\mathbb{C}\backslash\{0\}) \times (\mathbb{C}\backslash\{0\})$ on $H^n_{\mathbb{C}}$. Define the geometric quotient $\widetilde{H}^n_{\mathbb{C}} := H^n_{\mathbb{C}} // \left((\mathbb{C}\backslash\{0\}) \times (\mathbb{C}\backslash\{0\})\right)$.

**Theorem (Hausel, Letellier and Rodriguez-Villegas, 2011)**

*For each $n \geq 1$,the $E$-polynomial of $\widetilde{H}^n_{\mathbb{C}}$ is $P_n(q)$.*

# Evaluation at roots of unity

**Theorem (Kassel and Reutenauer, 2016)**

*For any integer $n \geq 1$,*

*(i)* $P_n(1)$ *is the sum of divisors of* $n$,

*(ii)* $C_n(-1)$ *is the number of integer solutions to the equation* $x^2 + y^2 = n$,

*(iii)* $\left| C_n\left(\sqrt{-1}\right) \right|$ *is the number of integer solutions to the equation* $x^2 + 2y^2 = n$,

*(iv)* $6 Re\, P_n\left(\frac{-1+\sqrt{-3}}{2}\right)$ *is the number of integer solutions to the equation* $x^2 + xy + y^2 = n$.

# Finite fields geometric interpretation of $C_n(q)$

Consider the Hilbert scheme

$$H^n_{\mathbb{F}_q} := \mathrm{Hilb}^n\left(\left(\mathbb{A}^1_{\mathbb{F}_q}\backslash\{0\}\right)\times\left(\mathbb{A}^1_{\mathbb{F}_q}\backslash\{0\}\right)\right)$$

of $n$ points on the bidimensional $\mathbb{F}_q$-torus.

**Theorem (Kassel and Reutenauer, 2015)**

*For each $n \geq 1$,*

$$\sum_{m=1}^{\infty} C_n\left(q^m\right) t^m = \frac{t\frac{d}{dt} Z_{H^n_{\mathbb{F}_q}}(t)}{Z_{H^n_{\mathbb{F}_q}}(t)},$$

*where $Z_{H^n_{\mathbb{F}_q}}(t)$ is the local zeta function of $H^n_{\mathbb{F}_q}$.*

# Evaluation at prime powers

**Theorem (Kassel and Reutenauer, 2015)**

*For any prime power $q$ and any integer $n \geq 1$, $C_n(q)$ is the number of ideals $I$ of the group algebra $\mathbb{F}_q[\mathbb{Z} \oplus \mathbb{Z}]$ such that $\mathbb{F}_q[\mathbb{Z} \oplus \mathbb{Z}]/I$ is a vector space of dimension $n$ over $\mathbb{F}_q$.*

# The coefficients of $P_n(q)$

**Theorem (J. M. R. C., 2017)**

*For any integer $n \geq 1$,*

*(i) the largest coefficient of $P_n(q)$ is $F(n)$, where $F(n)$ is the Erdös-Nicolas function*[a]*, i.e.*

$$F(n) := \max_{t>0} \# \{d|n : \quad t < d \leq 2t\}.$$

*(ii) the polynomial $P_n(q)$ has a coefficient greater than 1 if and only if $2n$ is the perimeter*[b] *of a Pythagorean triangle,*

*(iii) all the coefficients of $P_n(q)$ are non-zero if and only if $n$ is 2-densely divisible*[c]*.*

---

*a*. Paul Erdös, Jean-Louis Nicolas. Méthodes probabilistes et combinatoires en théorie des nombres. Bull. SC. Math **2** (1976) : 301–320.

*b*. The perimeter of a Pythagorean triangle is always an even integer.

*c*. i.e. the quotient of two consecutive divisors of $n$ is less than or equal to 2. Densely divisible numbers were introduced by the international team *polymath8*, led by Terence Tao, in order to improve Zhang's bounded gaps between primes.

# Odd-trapezoidal numbers

An integer $n \geq 1$ is an *odd-trapezoidal number* if for each pair of integers $a \geq 1$ and $k \geq 1$, the equality

$$n = a + (a+1) + (a+2) + ... + (a+k-1)$$

implies that $k$ is odd.

# Odd-trapezoidal numbers

**Theorem (J. M. R. C., 2017)**

*For any integer $n \geq 1$, we have that $n$ is odd-trapezoidal if and only if*

$$a_{n,0} \geq a_{n,1} \geq a_{n,2} \geq \ldots \geq a_{n,n-1},$$

*where $a_{n,0}, a_{n,1}, a_{n,2}, \ldots, a_{n,n-1}$ are the coefficients in the computational definition of $P_n(q)$.*

# Conclusion of Part I

From the polynomial $P_n(q)$ it is computationally easy to derive the following information about $n$ :

(i) Whether or not $2n$ is the perimeter of a Pythagorean triangle.

(ii) Whether or not $n$ is 2-densely divisible.

(iii) Whether or not $n$ is odd-trapezoidal.

(iv) The number of middle divisors [1] of $n$.

(v) The number of integer solutions to the equations $x^2 + y^2 = n$, $x^2 + 2y^2 = n$ and $x^2 + xy + y^2 = n$.

(vi) The number of ideals $I$ of the group algebra $\mathbb{F}_q [\mathbb{Z} \oplus \mathbb{Z}]$ such that $\mathbb{F}_q [\mathbb{Z} \oplus \mathbb{Z}] / I$ is a vector space of dimension $n$ over $\mathbb{F}_q$.

(vii) The value of Erdös-Nicolas function at $n$.

(viii) The sum of divisors of $n$.

(ix) Topological information about $H^n_{\mathbb{C}}$ and $\widetilde{H}^n_{\mathbb{C}}$.

1. i.e the divisors $d$ satisfying $\sqrt{\frac{n}{2}} < d \leq \sqrt{2n}$.

# Part II
## Language Theory

(i) We will encode part of the information from Kassel-Reutenauer polynomials into formal words.

(ii) We will translate some of the properties satisfied by Kassel-Reutenauer polynomials into language-theoretical statements.

# The non-zero coefficients of $C_n(q)$

We can express $C_n(q) \in \mathbb{Z}[q]$, in a unique way, as follows[2],

$$C_n(q) = \eta_0 q^{e_0} + \eta_1 q^{e_1} + \dots + \eta_k q^{e_k},$$

for some $\eta_0, \eta_1, \dots, \eta_k \in \{+1, -1\}$ and $e_0, e_1, \dots, e_k \in \mathbb{Z}$ satisfying :

(i) $e_0 \geq e_1 \geq \dots \geq e_k \geq 0$,

(ii) for any $0 \leq j \leq k-1$, if $e_j = e_{j+1}$, then $\eta_j = \eta_{j+1}$.

So, the vector $\mathrm{KR}(n) := (\eta_0, \eta_1, \dots, \eta_k) \in \{+1, -1\}^{k+1}$ is well-defined. By abuse of notation, we will write $\mathrm{KR}(n)$ as a word over the alphabet $\{+, -\}$, identifying $+ \leftrightarrow +1$ and $- \leftrightarrow -1$.

2. Notice that each positive coefficient of $C_n(q)$ corresponds to the multiplicity of a pole of $Z_{H^n_{\mathbb{F}_q}}(t)$. Similarly, each negative coefficient of $C_n(q)$ corresponds to the multiplicity of a zero of the same rational function.

# Example

For $n = 6$,

$$\begin{aligned} C_6(q) &= q^{12} - q^{11} + q^7 - 2q^6 + q^5 - q + 1 \\ &= +q^{12} - q^{11} + q^7 - q^6 - q^6 + q^5 - q + 1. \end{aligned}$$

Therefore,

$$\mathrm{KR}(6) = + - + - - + - + \,.$$

# A hidden pattern

Notice that

$$\mathrm{KR}(75) = + - - + + - + - - + - + + - + - - + - + + - - +$$

can be obtained from the well-matched parentheses

$$\color{red}{()(())(())()}$$

by means of the letter-by-letter substitution [3]

$$\begin{array}{rcl} \mu : \{\color{red}{(}, \color{red}{)}\}^* & \longrightarrow & \{+, -\}^*, \\ \color{red}{(} & \mapsto & +-, \\ \color{red}{)} & \mapsto & -+. \end{array}$$

$$\mathrm{KR}(75) = \underbrace{+-}_{\color{red}{(}}\underbrace{-+}_{\color{red}{)}}\underbrace{+-}_{\color{red}{(}}\underbrace{+-}_{\color{red}{(}}\underbrace{-+}_{\color{red}{)}}\underbrace{-+}_{\color{red}{)}}\underbrace{+-}_{\color{red}{(}}\underbrace{+-}_{\color{red}{(}}\underbrace{-+}_{\color{red}{)}}\underbrace{-+}_{\color{red}{)}}\underbrace{+-}_{\color{red}{(}}\underbrace{-+}_{\color{red}{)}}$$

3. i.e. morphism of free monoids. Here, $\Sigma^*$ denotes the free monoid over the alphabet $\Sigma$.

# Well-matched parentheses

It follows from the computational definition of $C_n(q)$ that $\mathrm{KR}(n)$ is palindromic. The following property is less obvious.

**Theorem (J. M. R. C., 2017)**

*For each integer $n \geq 1$, $\mathrm{KR}(n) = \mu(w_n)$, for some well-matched parentheses $w_n$.*

Define the function $\delta : \mathbb{Z}_{\geq 1} \longrightarrow \{(,)\}^*$ by $\delta(n) := w_n$.

# A direct computation of $\delta(n)$

The following result can be interpreted as a language-theoretical version of a formula for the coefficients of $C_n(q)$ due to Kassel and Reutenauer.

**Theorem (J. M. R. C., 2017)**

*Let $n \geq 1$ be an integer. Denote $D_n$ the set of divisors of $n$. Define $2D_n := \{2d : \quad d \in D_n\}$. Let $\tau_1 < \tau_2 < ... < \tau_k$ be the elements of $D_n \triangle 2D_n$ written in increasing order. Then,*

$$\delta(n) = t_1\, t_2 \ldots t_k,$$

*where*

$$t_i := \begin{cases} ( & \textit{if } \tau_i \in D_n \backslash (2D_n)\,, \\ ) & \textit{if } \tau_i \in (2D_n)\backslash D_n. \end{cases}$$

# Sets and languages

## Definition

Let $\Sigma$ be a finite alphabet. Given a set $S \subseteq \mathbb{Z}_{\geq 1}$, we say that $S$ is *rational* (*context-free*) with respect to a function $f : \mathbb{Z}_{\geq 1} \longrightarrow \Sigma^*$, if

$$S = f^{-1}(L)$$

for some rational (context-free) language $L \subseteq \Sigma^*$.

# Rational sets with respect to $\delta$

**Theorem (J. M. R. C., 2017)**

*The following sets are rational with respect to $\delta$,*

*(i) the empty set of integers,*

*(ii) all the integers,*

*(iii) powers of 2,*

*(iv) semi-perimeters[a] of Pythagorean triangles.*

---

a. The semi-perimeter is a half of the perimeter.

# Blocks

The *number of blocks* of an integer $n \geq 1$ is defined as the number of connected components of the topological space

$$\bigcup_{d|n} [d, 2d].$$

Notice that $n$ is 2-densely divisible if and only if $n$ has only one block.

# Context-free sets with respect to $\delta$

**Theorem (J. M. R. C., 2017)**

*The following sets are context-free with respect to $\delta$,*

- *(i) integer having exactly $k$ blocks, for any fixed $k \geq 1$,*
- *(ii) numbers $n$ satisfying $F(n) \geq h$, for any fixed integer $h \geq 1$, where $F(n)$ is the Erdös-Nicolas function.*

# Definition of $\widehat{\delta}$

For all $n \geq 1$,

$$\widehat{\delta}(n) := \psi\,\delta\,(n)\,,$$

where $\psi : \{((, )), (), )(\}^* \longrightarrow \{A, B, C, D\}^*$ satisfies, for all $w \in \{((, )), (), )(\}^*$,

$$\begin{aligned}
\psi\varepsilon &:= \varepsilon, \\
\psi(w) &:= A\psi w, \\
\psi)w( &:= B\psi w, \\
\psi(w( &:= C\psi w, \\
\psi)w) &:= D\psi w.
\end{aligned}$$

# Hirschhorn function

We define the *Hirschhorn function* [4], $\mathrm{H} : \mathbb{Z}_{\geq 1} \times \{0, 1\} \longrightarrow \mathbb{Z}_{\geq 0}$ by means of the expression

$$\mathrm{H}(n, b) := \# \{(a, k) \in \Pi_n : \quad k \equiv b \pmod 2\},$$

where $\Pi_n$ is the set of pairs $(a, k) \in (\mathbb{Z}_{\geq 1})^2$ such that

$$n = a + (a + 1) + (a + 2) + \ldots + (a + k - 1).$$

Notice that $\mathrm{H}(n, b) = 0$ if and only if $n$ is odd-trapezoidal.

4. M. D. Hirschhorn and P. M. Hirschhorn. "Partitions into consecutive parts." (2009).

# Rational sets with respect to $\hat{\delta}$

**Theorem (J. M. R. C., 2017)**

*Given $k \in \mathbb{Z}_{\geq 0}$ and $b \in \{0, 1\}$, the set of integers $n \geq 1$ satisfying $\mathrm{H}(n, b) \geq k$ is rational with respect to $\hat{\delta}$.*

# Context-free sets with respect to $\hat{\delta}$

**Theorem (J. M. R. C., 2017)**

*For each integer $k \geq 1$, the set of numbers $n$ having exactly $k$ middle divisors is context-free with respect to $\hat{\delta}$.*

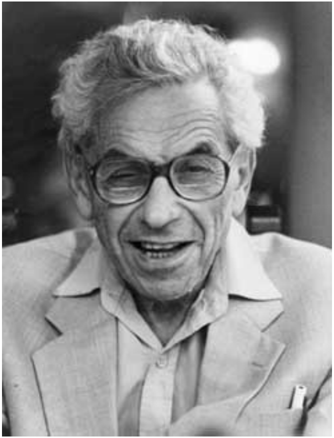
# Conclusions of Part II

Languages-theoretical algorithms can be used in order to compute the evaluation at $n$ of several nontrivial arithmetical functions (including characteristic functions) just from the information provided by $\delta(n)$.

# Part III

## Arithmetic theorems having language-theoretic proofs

# “God has the Big Book, the beautiful proofs of mathematical theorems are listed here”



Paul Erdös

## “Don’t come to me with your pretty proofs. We don’t bother with that baby stuff around here!”

Solomon Lefschetz

# Arithmetical theorems proved by Language Theory

**Theorem (J. M. R. C., 2017)**

*For all integers $n \geq 2$, if $n$ is not a power of 2 and $n$ is odd-trapezoidal, then $2n$ is the perimeter of a Pythagorean triangle.*

# Arithmetical theorems proved by Language Theory

**Theorem (J. M. R. C., 2017)**

*For all integers $n \geq 1$, if $2n$ is the perimeter of a Pythagorean triangle, then $n$ has at least two different prime divisors.*

# Arithmetical theorems proved by Language Theory

**Theorem (J. M. R. C., 2017)**

*For all integers $n \geq 2$, if $n$ is 2-densely divisible and $2n$ is not the perimeter of a Pythagorean triangle then $n$ is a power of 2.*

The following result is Theorem 3 in Hartmut F. W. Höft, *On the Symmetric Spectrum of Odd Divisors of a Number*, preprint on-line available at `https://oeis.org/A241561/a241561.pdf`

**Theorem (Höft, 2015)**

*For all $n \geq 1$, there exists at least a middle divisor of $n$ if and only if the number of blocks of $n$ is odd.*

# Conclusions of Part III

Using language-theoretical relationships, **nontrivial** elementary number-theoretical results can be derived via $\delta(n)$.

# An open question

A formal language is *decidable* if there exists a total Turing machine[5] that, when given a finite sequence of symbols as input, accepts it if it belongs to the language and rejects it otherwise.

Is the language $\delta\left(\mathbb{Z}_{\geq 1}\right)$ decidable ?

5. A total Turing machine is a Turing machine that halts for every given input.

# References


Tamás Hausel, Emmanuel Letellier and Fernando Rodriguez-Villegas. *Arithmetic harmonic analysis on character and quiver varieties*, Duke Mathematical Journal **160**.2 (2011) : 323-400.

Tamás Hausel, Emmanuel Letellier and Fernando Rodriguez-Villegas. *Arithmetic harmonic analysis on character and quiver varieties II*, Advances in Mathematics **234** (2013) : 85-128.

Christian Kassel and Christophe Reutenauer, *The Fourier expansion of* $\eta(z)\ \eta(2z)\ \eta(3z)\ /\ \eta(6z)$, Archiv der Mathematik **108**.5 (2017) : 453-463.

Christian Kassel and Christophe Reutenauer, *Counting the ideals of given codimension of the algebra of Laurent polynomials in two variables*, Michigan J Maths (to appear).

Christian Kassel and Christophe Reutenauer, *Complete determination of the zeta function of the Hilbert scheme of n points on a two-dimensional torus*, Ramanujan Journal (to appear).

José Manuel Rodríguez Caballero, *Symmetric Dyck Paths and Hooley's* $\Delta$*-function*, Combinatorics on Words. Springer International Publishing AG. 2017.

# The End