An expository paper on D-Finite Generating Functions

Natalia Kokoromyti

Euler Circle, Fall 2021

Generating Functions Class

* nataliakokoromyti@gmail.com

Abstract

The present paper provides an analysis on D-Finite Generating functions. It was completed in the context of Euler Circle Generating Functions class. Here we consider a class of power-series, which we call differentiably finite, whose coefficients can be quickly computed. At first we will delve into the theory of differential equations, by giving an overview over the main properties and the classical algorithms for D-finite functions. In the end we will examine how D-finite generating functions and asymptotic counting help us study the evolution of k-non crossing σ -canonical RNA structures.

1 Definitions

1.1 k-noncrossing

K-noncrossing is a set of distinct arcs $(i_1, j_1), (i_2, j_2), ..., (i_k, j_k)$, such that:
$i_1 < i_2 < ... < i_k < j_1 < j_2 < ... < j_k$. A graph without k-crossings is a k-noncrossing graph and a k-noncrossing graph without isolated points is called a k-noncrossing matching.

1.2 stack

Stack of size σ, $S^{\sigma}_{i,j}$, is the maximal sequence of "parallel" arcs,
$((i, j), (i + 1, j - 1), ..., (i + (\sigma - 1)..., (i + (\sigma - 1), j - (\sigma - 1)))$.
A stack of size σ is called a σ -stack.

1.3 k-noncrossing, σ -canonical structures

A k-noncrossing, σ -canonical structures is a k-noncrossing graph with minimum arc length, $\lambda \geq 2$, and minimum stack size σ. The nucleus is a k-noncrossing structure with minimum arc length $\lambda \geq 2$, where every stack has size 1.
V_k, which will appear often in this paper, is a k-noncrossing matching with stacks of size exactly 1.

2 Introduction

The purpose of this dissertation is to study the interplay between holonomic functions and the structure of RNA false nodes. Here we note that, that the functions of various types RNA structures will be considered known for the purpose of the paper, since they get extracted from methods exceed the fringes of mathematics. Before we move on with the main theorems let's start with some basic definitions.

3 D-Finiteness

A formal power series y ϵ $C\,[[x]]$ is called D-finite (differentially finite) if it satisfies a linear homogeneous differential equation with polynomial coefficients.
The importance of D-finiteness in the enumeration comes from the fact that a function generator is D-finite if and only if its coefficients are P-recursive. (polynomially recursive).

Examples of holonomic functions and sequences:

(1) all algebraic functions,

(2)all sine and cosine functions,

(3)exponential functions and logarithms,

(4)the generalized hypergeometric function,

(5)the error function,

(6)the Bessel function,

(7)the Airy functions,

(8)all constant-recursive sequences,

(9)the sequence of factorials n!,

(10)the Catalan numbers,

(11)the Motzkin numbers,

(12)the sequence of derangements,

3.1 P-recursiveness

P-recursive (polynomially recursive) is called a sequence if there exist polynomials $p_0(n), ..., p_m(n) \epsilon C[n]$, with $p_m(n) \neq 0$, such that for every n ϵ N: $p_m(n)f(n+m) + p_{m-1}(n)f(n+m-1) + ... + p_0(m)f(n) = 0$.
A power series is called algebraic if there exist polynomials $q_0(x), ..., q_m(x) \epsilon C[[x]]$, with $q_m(x) \neq 0$, such that $q_m F^m(x) + q_{m-1}F^{m-1}(x) + ... + q_1 F(x) + q_0(x) = 0$.

3.1.1 Theorem (Stanley 2012)

For the P-Recursive series, D-Finite and algebraic power series the following are true:

a. If f, g are P-Recursive then $f \cdot g$ is P-Recursive.

b. If F, G are D-Finite and a,b ϵC, then $aF + bG$ and FG are D-Finite.

c. If F is D-Finite and G is algebraic with $G(0) = 0$, then $F(G((x))$ is D-Finite.

Proof:
We will prove (c) first, which will be useful later on. Since G(0)=0, F(G(x)) is well-defined. Let K=F(G(x)). Then $K^{(i)}$ is a linear combination of $F(G(x)), F'(G(x)), ...,$ over $C[G, G', ...]$, that is the polynomial ring in G,G',... with complex coefficients. Let $G^{(i)} \epsilon C, i \geq 0$ hence $C[G, G', ...] \subset C(x, G)$, where C(x,G) denotes what is produced by x and G.
Since G is algebraic, it satisfies:
$q_d(x)G^d(x) + q_{d-1}(x)G^{d-1}(x) + ... + q_1(x)G(x) + q_0(x) = 0$, where $q_0(x), ..., q_d(x) \epsilon C[x], q_d(x) \neq 0$ and d is minimum, that is $(G^i(x))_{i=0}^{d-1}$ is linearly independent over $C[x]$.

Let $P(x, G) = q_d(x)G^d(x) + q_{d-1}(x)G^{d-1}(x) + ... + q_1(x)G(x) + q_0(x)$.

We differentiate both sides and we get:

$$0 = \frac{d}{dx}P(x, G) = \frac{\vartheta P(x,y)}{\vartheta x}\Big|_{y=G} + G' \frac{\vartheta P(x,y)}{\vartheta y}\Big|_{y=G}.$$

The degree of $\frac{\vartheta P(x,y)}{\vartheta y}|_{y=G}$ in G is smaller than d-1 and $q_d(x) \neq 0$, hence $\frac{\vartheta P(x,y)}{\vartheta y}|_{y=G} \neq 0$.
As a result:

$$G' = -\frac{\frac{\vartheta P(x,y)}{\vartheta x}|_{y=G}}{\frac{\vartheta P(x,y)}{\vartheta y}|_{y=G}} \epsilon C(x, G).$$

By repeating the aforementioned process, we can conclude that $G^{(i)} \epsilon C(x, G), i \geq 0$, thus $C[G, G', ...] \subset C(x, G)$, and our hypothesis is true. We denote by V the vector space in $C(x, G)$ that is generated from F(G(x)), F'(G(x)),...Since F is D-Finite, we know that $dim_{C(x)} \langle F, F', ...\rangle < \infty$, which implies that $dim_{C(G)} \langle F(G), F'(G), ...\rangle$ is finite. Knowing that $C(G) \subset C(x, G)$ we obtain the following: $dim_{C(x,G)} \langle F, F', ...\rangle < \infty$.
This allows us to conclude that $dim_{C(x,G)}V < \infty$ and $dim_{C(x)}C(x, G) < \infty$.
Hence, $dim_{C(x)}V = dim_{C(x,G)}V \cdot dim_{C(x)}C(x, G) < \infty$ and because of the fact that $K^{(i)} \epsilon V$, we conclude that $F(G(x))$ is D-finite.

4 The equation $F_k(z)$

4.1 Lemma

Let $I_r(2x) = \sum_{j\geq 0} \frac{x^{2j+r}}{j!(r+j)!}$ be a Bessel function of the first kind with degree r. The generating function of the k-noncrossing matchings is given from the formula below: $\sum_{n\geq 0} f_k(2n)\frac{x^{2n}}{(2n)!} = det\left[I_{i-j}(2x) - I_{i+j}(2x)\right]|_{i,j=1}^{k-1}$.
The lemma above is important because it implies that $H_k(z) = \sum_{n\geq 0} f_k(2n)\frac{z^{2n}}{(2n)!}$ is D-finite.
This observation will prove to be particularly significant later on.

Theorem 2.2.1 and 2.1.2 will be presented without a proof, since their proofs exceed the goal of this paper.

4.1.1 Theorem (Reidys 2011)

For an arbitrary $k\epsilon N, k \geq 2, arg(z) \neq \pm\frac{\pi}{2}$, the following is true:

$H_k(z) = \left[\prod_{i=1}^{k-1} \Gamma(i + 1 - \frac{1}{2}) \prod_{r=1}^{k-2} r!\right] (\frac{e^{2\pi}}{\pi})^{k-1} z^{-(k-1)^2 - \frac{k-1}{2}} (1 + O(|z|^{-1}))$, where $\Gamma(z)$ is the gamma function.

4.1.2 Theorem (Reidys 2011)

For an arbitrary $k\epsilon N, k \geq 2$, we have that $f_k(2n) = c_k n^{-((k-1)^2 + (k-1)/2)}$, where $c_k > 0$.

4.1.3 Lemma

The generating function of the combinatorial class of the k-noncrossing matchings of 2n vertices, $F_k(z) = \sum_{n\geq 0} f_k(2n)z^n$ is D-finite.

Proof: From lemma 2.1 we know the exponential generating function of of $f_k(2n)$, which is:

$\sum_{n\geq 0} f_k(2n)\frac{x^{2n}}{(2n)!} = det\left[I_{i-j}(2x) - I_{i+j}(2x)\right]|_{i,j=1}^{k-1}$, where $I_m(x)$ is a Bessel function of the first kind.
This function satisfies $I_n(x) = i^{-n}J_n(ix)$ with $J_n(x)$ being the solution of Bessel's functional equation: $x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - n^2)y = 0$. Notice that for every $n\epsilon N$, $J_n(x)$ is D-finite.

Let $G(x) = ix$.
Obviously, $G(x)\epsilon C_{alg}[x]$ and $G(0) = 0$, hence $J_n(ix)$ and $I_n(x)$ are D-finite according to (c) of theorem 1.1.1. Similarly, we prove that $I_n(2x)$ is D-finite for some constant n ϵ N.
From $\sum_{n\geq 0} f_k(2n)\frac{x^{2n}}{(2n)!} = det\left[I_{i-j}(2x) - I_{i+j}(2x)\right]|_{i,j=1}^{k-1}$ and (b) of theorem 1.1.1, we conclude that $H_k(x) = \sum_{n\geq 0} \frac{f_k(2n)}{(2n)!}x^{2n}$ is D-finite. In other words, the sequence $f(n) = \frac{f_k(2n)}{(2n)!}$ is P-recursive and $g(n) = (2n)!$ is also P-recursive, since $(2n + 1)(2n + 2)g(n) - g(n + 1) = 0$.

As a result, $f_k(2n) = f(n)g(n)$ is P-recursive, which proves that $F_k(z) = \sum_{n\geq 0} f_k(2n)z^n$ is D-finite.

Since, $F_k(z)$ is D-finite, there exists an e ϵ N, such that it satisfies an ordinary differentiable equation of the form: $q_{0,k}(z)\frac{d^e}{dz^e}F_k(z) + q_{1,k}(z)\frac{d^{e-1}}{dz^{e-1}}F_k(z) + ... + q_{e,k}(z)F_k(z) = 0$, where $q_{j,k}(z)$ are polynomials.

Knowledge of this particular ODE is particularly significant for two reasons:

(i)every dominant singularity of a solution belongs in the roots of $q_{0,k}(z)$.In other words, the ODE checks the dominant singularities that are important for the asymptotic counting.

(ii)under specific conditions, the singular expansion of $F_k(z)$ follows from the ODE. (lemma 2.1.4) Subsequently, we present without a proof (Reidys 2011) the ordinary differentiable equations of $F_k(z)$ for $2 \leq k \leq 9$, but also the singular expansion of $F_k(z)$.

4.1.4 Lemma

For $2 \leq k \leq 9$, we have:
$q_{0,2} = (4z - 1)z$,
$q_{0,3} = (16z - 1)z^2$,
$q_{0,4} = (144z^2 - 40z + 1)z^3$,
$q_{0,5} = (1024z^2 - 80z + 1)z^4$,
$q_{0,6} = (14400z^3 - 4144z^2 + 140z - 1)z^5$,
$q_{0,7} = (147456z^3 - 12544z^2 + 224z - 1)z^6$,
$q_{0,8} = (2822400z^4 - 826624z^3 + 31584z^2 - 336Z + 1)z^7$,
$q_{0,9} = (37748736z^4 - 3358720z^3 + 69888z^2 - 480Z + 1)z^8$.

The equations above, combined with theorem 2.1.2 show that for $2 \leq k \leq 9$ the only dominant singularity of $F_k(z)$ is given from ρ_k^2, where $\rho_k = \frac{1}{2(k-1)}$.

4.1.5 Lemma

For $2 \leq k \leq 9$, the singular expansion of $F_k(z)$ for $z \to \rho_k^2$ is given from the formula below:
$F_k(z) = P_k(z - \rho_k^2) + c'_k(z - \rho_k^2)^{((k-1)^2 + \frac{k-1}{2})^{-1}} log(z - \rho_k^2)(1 + o(1))$, if k is odd,
$F_k(z) = P_k(z - \rho_k^2) + c'_k(z - \rho_k^2)^{((k-1)^2 + \frac{k-1}{2})^{-1}}(1 + o(1))$, if k is even.
In addition, the terms of $P_k(z)$ are polynomials of such a degree that is not greater than $(k-1)^2 + \frac{k-1}{2} - 1$,$c'_k$ is a constant and $\rho_k = \frac{1}{2(k-1)}$.

5 Asymptotic counting

Now that we have the necessary tools, we can proceed with the asymptotic counting of various structures. Subsequently we will end up with the asymptotic counting of k-noncrossing sigma-canonical RNA sequences with minumum arc length of 4. The generating function $I_k(s, m)$ of k-noncrossing graphns, with length $2s$ and m 1-arcs, is given from:
$I_k(z, u) = \frac{1+z}{1+2z-zu}F_k(\frac{z(1+z)}{(1+2z-zu)^2})$.

5.0.1 Theorem (Reidys and Wang 2010)

For $2 \leq k \leq 9$, the number of V_k shapes of length 2s is given asymptotically from the formula:
$i_k(s) = c_k s^{-((k-1)^2+(k-1)/2)}(\mu_k^{-1})^s$, where μ_k is the only real solution to the equation $\frac{z}{1+z} = \rho_k^2$ and c_k is some positive constant.

Proof:
The generating function $F_k(z) = \sum_{n \geq 0} f_k(2n)z^n$ is D-finite and the internal function $\theta(z) = \frac{z}{1+z}$ is algebraic, satisfies the $\theta(0) = 0$ and is analytic for $|z| < 1$.Using the fact that all singularities of $F_k(z)$ are included in the roots of $q_{0,k}(z)$ we can confirm that $F_k(\theta(z))$ has the unique dominant real singularity $\mu_k < 1$, that satisfies $\theta(\mu_k) = \rho_k^2$ for $2 \leq k \leq 9$. Since $\theta'(\mu_k) \neq 0$, we know that $i_k(s) = c_k s^{-((k-1)^2+(k-1)/2)}(\mu_k^{-1})^s$.
qed.

We know that the generating function of the nuclei is given from:
$C_k(z) = \frac{1}{r(z)z^2-z+1}F_k((\frac{\sqrt{r(z)}z}{r(z)z^2-z+1})^2)$, with $r(z) = \frac{1}{1+z^2}$.

5.0.2 Theorem

For $k\epsilon N, k \geq 2$, we have that: $C_k(n) = c_k n^{-((k-1)^2)+\frac{(k-1)}{2})}(\frac{1}{\kappa_k})^n, k = 3,4,...,9$, where κ_k is the only positive real dominant singularity of $C_k(z)$ and the smallest positive real solution of the equation $\frac{\sqrt{r(x)}x}{r(x)x^2-x+1} = \rho_k^2$, for $k = 3,4,..,9$.
Proof: Prigsheim's Theorem (Titchmarsh 1939) guarantees that $C_k(z)$ has one positive real dominant singularity κ_k. We confirm that there exists a unique solution of $w(x) = (\frac{\sqrt{r(x)}x}{r(x)x^2-x+1})^2 = \rho_k^2$, for $k = 3,4,..,9$. This solution is equal to κ_k, the thus unique, real dominant singularity of $C_k(z)$. Also, since κ_k is strictly less than the singularity of $w(x)$, $w'(\kappa_k) \neq 0$, and $w(x)$ is algebraic we derive:
$C_k(n) = c_k n^{-((k-1)^2)+\frac{(k-1)}{2})}(\frac{1}{\kappa_k})^n$, for some $c_k > 0$.
qed

6 Final Remarks

The results that were presented in the previous sections, become particularly interesting because they showcase how we can use d-finite generating functions in order to study biological molecules and unveil surprising connections in the evolution of RNA false nodes: many of them still awaiting to be explored.

References

http://www.algebra.uni-linz.ac.at/people/mkauers/publications/kauers15n.pdf

https://arxiv.org/abs/2102.00766

https://math.mit.edu/ rstan/pubs/pubfiles/45.pdf

http://www.algebra.uni-linz.ac.at/people/mkauers/publications/kauers11c.pdf

https://www.emis.de/proceedings/MKM2001/printed/meunier.ps

https://arxiv.org/abs/0806.2414

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3892343/