Algebraic Analysis of the Hypergeometric Function $\,_{1}F_{1}$ of a Matrix Argument

Anna-Laura Sattelberger (MPI MiS Leipzig)

Online Seminar on Computational Algebra

November 27, 2020
What is this talk about?

[3] P. Görlach, C. Lehn, and A.-L. S.: Algebraic Analysis of the Hypergeometric Function $\mathbf{1F}_1$ of a Matrix Argument. *Beitr. Algebra Geom.*, November 2020.

**Computational Algebraic Analysis**

- investigation of linear partial differential equations by algebraic methods
- tackle concrete problems in the sciences by computer-aided computations
- exploit and construct algorithms and software
1  Hypergeometric functions of a matrix argument
2  $D$-Modules behind
3  Characteristic variety and singular locus
Hypergeometric Functions of a Matrix Argument
Let $p, q \in \mathbb{N}$. The **hypergeometric series** $pF_q$ is

$$pF_q(a_1, \ldots, a_p; c_1, \ldots, c_q)(x) := \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n x^n}{(c_1)_n \cdots (c_q)_n n!},$$

where $(a)_n = a \cdot (a + 1) \cdots (a + n - 1)$ denotes the Pochhammer symbol.

- $p < q + 1$: entire function
- $p = q + 1$: convergent for $|x| < 1$, divergent for $|x| > 1$
- $p > q + 1$: divergent except at $x = 0$

omnipresent in Hodge Theory, Physics, Toric Geometry, and many more

Examples

- $0F_0(x) = \exp(x)$
- $2F_1(a_1, a_2; c)(x)$ Gauß’ hypergeometric function
- $1F_0(a; x) = (1 - x)^{-a}$
- $2F_2$ and $0F_1$ related to Bessel’s functions
Zonal polynomials

Let $m \in \mathbb{N}_{>0}$. Let $\lambda = (\lambda_1, \ldots, \lambda_m)$ be a partition of $d = |\lambda| = \lambda_1 + \cdots + \lambda_m$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$. The zonal polynomial $C_\lambda \in \mathbb{C}[x_1, \ldots, x_m]$ is a certain symmetric homogeneous polynomial.

Zonal polynomials of a matrix argument

Let $X \in \mathbb{C}^{m \times m}$ be a square matrix and $\lambda = (\lambda_1, \ldots, \lambda_m)$ a partition. One defines the zonal polynomial $C_\lambda(X)$ as

$$C_\lambda(X) := C_\lambda(x_1, \ldots, x_m),$$

where $x_1, \ldots, x_m$ are the eigenvalues of $X$. 
Let $p, q \in \mathbb{N}$. The **hypergeometric series** $pF_q$ of a matrix argument $X \in \mathbb{C}^{m \times m}$ is

$$pF_q(a_1, \ldots, a_p; c_1, \ldots, c_q)(X) := \sum_{n=0}^{\infty} \sum_{\lambda \vdash n} \frac{(a_1)_{\lambda} \cdots (a_p)_{\lambda} \cdot C_{\lambda}(X)^n}{(c_1)_{\lambda} \cdots (c_q)_{\lambda} \cdot n!},$$

where the $\lambda$ are partitions of $n$, $(\bullet)_{\lambda}$ denotes the **generalized Pochhammer symbol**

$$(a)_{\lambda} := \prod_{i=1}^{m} \left(a - \frac{i - 1}{2}\right)^{\lambda_i}.$$

- $1F_1$ related to distribution of largest eigenvalue of Wishart matrices
  - stated in Muirhead’s book [8]
  - holonomic gradient method in [4]
- $0F_1$ related to the Fisher distribution
  - holonomic gradient descent for the Fisher distribution on $SO(3)$ in [13]
  - further study of the equivariant $D$-module in [6]
  - $D$-ideal generalized to compact Lie groups other than $SO(n)$ in [1]
$D$-Modules behind
The Weyl algebra

The **Weyl algebra** is the non-commutative algebra

\[ D := \mathbb{C}[x_1, \ldots, x_n] \langle \partial_1, \ldots, \partial_n \rangle, \]

where the non-commutativity is given by Leibniz’ rule \([\partial_i, x_i] = 1, \ i = 1, \ldots, n\).

A **D-ideal** (resp. **D-module**) is a **left** **D-ideal** (resp. **D-module**).

**Some facts**

- Elements of \( D \) are linear differential operators with polynomial coefficients:

\[
D \ni P = \sum_{\alpha, \beta \in \mathbb{N}^n} c_{\alpha, \beta} x^\alpha \partial^\beta, \quad c_{\alpha, \beta} \in \mathbb{C}.
\]

- A **D-module** \( M \) is a natural generalization of linear PDEs.
- \( \text{Hom}_D(M, \mathcal{O}) \) is the space of holomorphic solutions to \( M \).
The characteristic variety of a $D_n$-ideal $I$ is the subscheme $\text{Char}(I)$ of $\mathbb{A}^{2n}$ determined by

$$\text{in}_{(0, 1)}(I) = \langle \text{in}_{(0, 1)}(P) \mid P \in I \rangle \triangleleft \mathbb{C}[x_1, \ldots, x_n, \xi_1, \ldots, \xi_n].$$

$I$ is holonomic if $\dim \text{Char}(I) = n$. The singular locus of $I$ is the set

$$\text{Sing}(I) := \bigcup_{Z \subseteq \text{Char}(I)} \overline{\pi_x(Z)} \subseteq \mathbb{A}^n,$$

where $Z$ runs over all irreducible components of $\text{Char}(I)$ distinct from the zero section $\{\xi_1 = \cdots = \xi_n = 0\}$ as sets.

**Theorem (Sato–Kawai–Kashiwara)**

Let $I$ be a holonomic $D_n$-ideal. Then every irreducible component $Z$ of $\text{Char}(I) \subseteq T^* \mathbb{A}^n = \mathbb{A}^n_x \times \mathbb{A}^n_\xi$ is a conormal variety to its projection to $\mathbb{A}^n_x$. In particular, $Z$ is Lagrangian.
Denote the **rational Weyl algebra** by

\[ R_n := \mathbb{C}(x_1, \ldots, x_n)\langle \partial_1, \ldots, \partial_n \rangle. \]

**Theorem (Cauchy–Kovalevski–Kashiwara)**

*Let I be a holonomic D-ideal. Outside \( \text{Sing}(I) \), the space of holomorphic solutions on a simply connected domain to I has dimension*

\[ \text{rank}(I) := \dim_{\mathbb{C}(x)}(R/RI). \]

**Computation of the singular locus**

- For a single \( P \in D \), the singular locus is easy to read.
- For a general D-ideal, computer algebra systems can make life easier in “small” examples.
- Implementations are available in Macaulay2 or Singular:Plural.
Holonomic functions

Many function spaces are $D$-modules in a natural way.

**Definition**

Let $M \in \text{Mod}(D)$. An element $f \in M$ is **holonomic**, if its annihilating $D$-ideal is holonomic.

**Facts & features**

- Holonomic functions are encoded by their annihilating $D$-ideal together with finitely many initial conditions.
- They possess good closure properties.
- Many special functions arising in the sciences are holonomic.
- Various holonomic functions are implemented in the Mathematica package `HolonomicFunctions`.
- Numerical evaluation (resp. local minimization) via the **holonomic gradient method** (resp. **holonomic gradient descent**).
Weyl closure

Definition
Let $I$ be a $D$-ideal. Its **Weyl closure** is the $D$-ideal $W(I) := RI \cap D$.

Some features
- The Weyl closure turns a $D$-ideal with finite holonomic rank into a holonomic $D$-ideal.
- The Weyl closure contains all annihilating operators of a holomorphic solution to $I$ at a generic point.
- Computationally expensive!
Muirhead’s $D$-ideal

\[ X = \text{diag}(x_1, \ldots, x_m) \in \mathbb{C}^{m \times m} \]

Annihilating $D$-ideal of $\,^1F_1$ [8]

The linear partial differential operators

\[ g_k := x_k \partial_k^2 + (c - x_k) \partial_k + \frac{1}{2} \left( \sum_{\ell \neq k} \frac{x_\ell}{x_k - x_\ell} (\partial_k - \partial_\ell) \right) - a \in R_m, \]

\( k = 1, \ldots, m, \) annihilate $\,^1F_1 (a; c) (X)$ wherever they are defined. Denote by $P_k \in D$ the differential operator obtained from $g_k$ by clearing the denominators. The **Muirhead ideal** is the $D$-ideal $I_m := \langle P_1, \ldots, P_m \rangle$.

[3, Proposition 5.6] (refining [8, Theorem 7.5.6])

Let $m \in \mathbb{N}_{>0}$, $a \in \mathbb{C}$ and $c \in \mathbb{C} \setminus \{ \frac{k}{2} \mid k \in \mathbb{Z}, \ k \leq m - 1 \}$. Then $\,^1F_1 (a; c)$ is the unique formal power series solution to $I_m$ around 0 with $\,^1F_1 (a; c)(0) = 1$. In particular, $\,^1F_1 (a; c)$ is the unique convergent power series solution to $I_m$ around 0 with $\,^1F_1 (a; c)(0) = 1$. 
Muirhead’s ideal cont’d

[4, Theorem 2]
For the graded lexicographic term order on $R_m$, a Gröbner basis of $R_m I_m$ is given by \( \{g_k = x_k \partial_k^2 + \text{l.o.t.} \mid k = 1, \ldots, m\} \).

[3, Corollary 4.4]
The holonomic rank of $I_m$ is given by $\text{rank}(I_m) = 2^m$. In particular, the Weyl closure $W(I_m)$ of $I_m$ and the function $\frac{1}{F_1}$ of a diagonal matrix are holonomic.

Some more properties of $I_m$
- $I_2$ holonomic, $I_4$ not holonomic
- $I_m \subsetneq W(I_m)$ already for $m = 2$
- computation of $W(I_m)$ not feasible for $m \geq 3$
- decomposition of $\text{Char}(I_m)$ computable for $m = 2, 3$
  - $m = 4$: fixed parameters $a, c$, finite field
Characteristic Variety and Singular Locus
Singular locus of Muirhead’s ideal

[3, Theorem 5.1]

Let $m \in \mathbb{N}_{>0}$, $a \in \mathbb{C}$ (and $c \in \mathbb{C} \setminus \{\frac{k}{2} \mid k \in \mathbb{Z}, k \leq m - 1\}$). Then the singular locus of $I_m$ agrees with the singular locus of $W(I_m)$. It is the hyperplane arrangement

$$A := \left\{ x \in \mathbb{C}^m \mid \prod_{i=1}^{m} x_i \cdot \prod_{k \neq \ell} (x_k - x_\ell) = 0 \right\}.$$

Sketch of proof

$\subseteq$ $\text{in}_{(0,1)}(I) \supseteq \langle \text{in}_{(0,1)}(P_1), \ldots, \text{in}_{(0,1)}(P_m) \rangle$

$\supseteq$

- Lemma: $p \in \mathbb{C}^m$ with distinct coordinates, one of which is zero. Then the space of formal power series solutions to $I_m$ centered at $p$ is of dimension at most $2^{m-1}$.

- Lemma: $p = (p_1, \ldots, p_m) \in (\mathbb{C}^*)^m$ with $\#\{p_1, \ldots, p_m\} = m - 1$. Then the space of formal power series solutions to $I_m$ centered at $p$ is of dimension at most $2^{m-2} \cdot 3$.

- Combine with the Theorem of Cauchy–Kowalevski–Kashiwara

$^1$The assumption on $c$ can actually be dropped.
Characteristics variety of $W(I_m)$

[3, Corollary 5.7]

*The characteristic variety of $W(I_m)$ contains the zero section and the conormal bundles of the irreducible components of $\mathcal{A}$, i.e.,*

$$\text{Char}(W(I_m)) \supseteq V(\xi_1, \ldots, \xi_m) \cup \bigcup_i V(x_i, \xi_1, \ldots, \hat{\xi}_i, \ldots, \xi_m)$$

$$\cup \bigcup_{i\neq j} V(x_i - x_j, \xi_i + \xi_j, \xi_1, \ldots, \hat{\xi}_i, \ldots, \hat{\xi}_j, \ldots, \xi_m).$$

**Proof**

Combining Theorem 5.1 and the theorem of Sato–Kawai–Kashiwara proves the claim.
Characteristic variety of $W(I_m)$

Let $J_0 | J_1 \ldots J_k$ denote a partition of $[m] = \{1, \ldots, m\}$, such that only $J_0$ may possibly be empty. Denote by $C_{J_0 | J_1 \ldots J_k}$ the linear subspace

$$V(\{x_j \mid j \in J_0\} \cup \{\sum_{i \in J_\ell} \xi_i \mid \ell = 1, \ldots, k\} \cup \bigcup_{\ell=1}^{k} \{x_i - x_j \mid i, j \in J_\ell\}) \subseteq \mathbb{A}^{2m}.$$

[3, Conjecture 6.2]

The (reduced) characteristic variety of $W(I_m)$ is the following arrangement of $m$-dimensional linear spaces:

$$\text{Char}(W(I_m))^{\text{red}} = \bigcup_{[m] = J_0 | J_1 \ldots J_k} C_{J_0 | J_1 \ldots J_k}. $$

In particular, it has $B_{m+1}$ many irreducible components, where $B_n$ denotes the $n$-th Bell number$^2$.

$^2(B_n)_{n \in \mathbb{N}} = 1, 1, 2, 5, 15, 52, 203, 877, 4140, \ldots$
Upper bound for $\text{Char}(I_m)$

For a partition $J_0|J_1 \ldots J_k$ as before, we denote by $\hat{C}_{J_0|J_1 \ldots J_k}$ the linear subspace

$$V(\{x_j \mid j \in J_0\} \cup \bigcup_{\ell=1}^{k}\{x_i - x_j \mid i, j \in J_\ell\} \cup \left\{\sum_{i \in J_\ell} \xi_i \mid \ell = 1, \ldots, k \text{ s.t. } |J_\ell| \leq 2\right\}).$$

Clearly, $\hat{C}_{J_0|J_1 \ldots J_k} \supseteq C_{J_0|J_1 \ldots J_k}$ with equality iff $|J_\ell| \leq 2$ for all $\ell \geq 1$.

[3, Proposition 6.3]

The (reduced) characteristic variety of $I_m$ is contained in the arrangement of the linear spaces $\hat{C}_{J_0|J_1 \ldots J_k}$, i.e.:

$$\text{Char}(I_m)^{\text{red}} \subseteq \bigcup_{[m] = J_0|J_1 \ldots J_k} \hat{C}_{J_0|J_1 \ldots J_k}.$$
Examples

Computation for $m = 2$ in $\mathbb{Q}(a, c)[x_1, x_2]\langle \partial_1, \partial_2 \rangle$

$\text{Char}(W(I_2))^{\text{red}} = V(x_1, x_2) \cup V(x_1, \xi_2) \cup V(\xi_1, x_2) \cup V(\xi_1, \xi_2) \cup V(\xi_1 + \xi_2, x_1 - x_2)$.

Computations for $m = 3$, generic $a, c$

$\text{Char}(I_3)^{\text{red}}$ decomposes into the $15 = B_4$ irreducible components

\[
V(x_1, x_2, x_3) \cup V(\xi_1, x_2, x_3) \cup V(x_1, \xi_2, x_3) \cup V(x_1, x_2, \xi_3) \\
\cup V(\xi_1, \xi_2, x_3) \cup V(\xi_1, x_2, \xi_3) \cup V(x_1, \xi_2, \xi_3) \cup V(\xi_1, \xi_2, \xi_3) \\
\cup V(x_1 - x_2, \xi_1 + \xi_2, x_3) \cup V(x_1 - x_3, \xi_1 + \xi_3, x_2) \cup V(x_2 - x_3, \xi_2 + \xi_3, x_1) \\
\cup V(x_1 - x_2, \xi_1 + \xi_2, \xi_3) \cup V(x_1 - x_3, \xi_1 + \xi_3, \xi_2) \cup V(x_2 - x_3, \xi_2 + \xi_3, \xi_1) \\
\cup V(x_1 - x_2, x_1 - x_3, \xi_1 + \xi_2 + \xi_3).
\]

Computations for $m = 4$ with fixed $a, c$ over a finite field

Computations suggest that $\text{Char}(I_4)^{\text{red}}$ decomposes into $51 = B_5 - 1$ irreducible components. One of them, namely $V(x_1 - x_2, x_1 - x_3, x_1 - x_4)$, is 5-dimensional.
Future work

Question

Is there an intrinsic description of Muirhead’s ideal using more advanced tools from the theory of $\mathcal{D}$-modules?

[3, Problem 6.7]

Compute the Weyl closure $W(I_m)$ of $I_m$ for any $m$.

[3, Problem 6.8]

Show that $\text{Char}(W(I_m))$ (and possibly $\text{Char}(I_m)$) are invariant under the action of $\mathbb{C}^* \times \mathbb{C}^*$ on $T^*\mathbb{A}^m = \mathbb{A}^m \times \mathbb{A}^m$ given by scalar multiplication on the factors.

[3, Problem 6.9]

Can the scaling relation $1F_1(a; c)(\frac{1}{\alpha}X)^a \overset{a \to \infty}{\longrightarrow} 0F_1(c)(X)$ be used to deduce a relation between $\text{Char}(I_m)$ and the characteristic variety of the corresponding ideal generated by the annihilating operators of $0F_1$?
References

[1] M. Adamer, András C. Lörincz, A.-L. S., and B. Sturmfels. Algebraic Analysis of Rotation Data. Alg. Stat. 11 (2020), 189–211.

[2] W. Decker, G.-M. Greuel, G. Pfister, and H. Schönemann. Singular 4-1-3—A computer algebra system for polynomial computations. http://www.singular.uni-kl.de, 2020.

[3] P. Görlich, C. Lehn, and A.-L. S. Algebraic Analysis of the Hypergeometric Function $\text{1F}_1$ of a Matrix Argument. Beitr. Algebra Geom., November 2020.

[4] H. Hashiguchi, Y. Numata, N. Takayama, and A. Takemura. The holonomic gradient method for the distribution function of the largest root of a Wishart matrix. J. Multivar. Anal. 117 (2013), 296–312.

[5] R. Hotta, K. Takeuchi, and T. Tanisaki. D-Modules, Perverse Sheaves, and Representation Theory. Progress in Mathematics 236, Birkhäuser, 2008.

[6] T. Koyama. The annihilating ideal of the Fisher integral. Kyushu J. Math. 74 (2020), 415–427.

[7] J. Martín-Morales and V. Levandovskyy. dmod_lib: A Singular:Plural library for algorithms for algebraic D-modules. https://www.singular.uni-kl.de/Manual/4-2-0/sing_537.htm

[8] R. J. Muirhead. Aspects of multivariate statistical theory. John Wiley & Sons, Inc., New York, 1982. Wiley Series in Probability and Mathematical Statistics.

[9] H. Nakayama, K. Nishiyama, M. Noro, K. Ohara, T. Sei, N. Takayama, and A. Takemura. Holonomic Gradient Descent and its Application to the Fisher–Bingham Integral. Adv. in Appl. Math. 47 (2011), 639–658.

[10] M. Noro. System of Partial Differential Equations for the Hypergeometric Function $\text{1F}_1$ of a Matrix Argument on Diagonal Regions. ISSAC ’16: Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation (2016), 381–388.

[11] M. Saito, B. Sturmfels, and N. Takayama. Gröbner Deformations of Hypergeometric Differential Equations. Algorithms and Computation in Mathematics 6, Springer, Heidelberg, 1999.

[12] A.-L. S. and B. Sturmfels. D-Modules and Holonomic Functions. arXiv:1910.01395 [math.AG], 2019. To appear in the proceedings of the MATH+ Fall School on Algebraic Geometry.

[13] T. Sei, H. Shibata, A. Takemura, K. Ohara, and N. Takayama. Properties and Applications of Fisher Distribution on the Rotation Group. J. Multivariate Anal. 116 (2013), 440–455.

[14] N. Takayama, T. Koyama, T. Sei, H. Nakayama, K. Nishiyama. hgm: Holonomic Gradient Method and Gradient Descent. https://CRAN.R-project.org/package=hgm

[15] D. Zeilberger. A Holonomic Systems Approach to Special Functions Identities. J. Comput. Appl. Math. 32 (1990), 321–368.
Thank you very much for your attention!