

A short history of multivariate Wronskians

Alin Bostan, Philippe Dumas

To cite this version:

Alin Bostan, Philippe Dumas. A short history of multivariate Wronskians. 2026. ⟨hal-05839886⟩

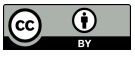
HAL Id: hal-05839886

https://hal.science/hal-05839886v1

Preprint submitted on 26 Oct 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

A short history of multivariate Wronskians

Alin Bostan
Inria, Sorbonne Université, LIP6, Paris, France
alin.bostan@inria.fr
ORCID: 0000-0003-3798-9281

Philippe Dumas
Inria Saclay, Palaiseau, France
philippe.dumas@inria.fr
ORCID: 0000-0001-9360-3844

Abstract

The Wronskian is the classical tool for testing the linear independence of functions of one variable: a family of analytic functions is linearly dependent if and only if the determinant built on their successive derivatives vanishes identically. For functions of several variables, there is no canonical choice of partial derivatives, and multivariate generalizations of the Wronskian have been proposed, forgotten and rediscovered many times since the 1870s, each time in the service of a different application: invariant theory, nomography, projective differential geometry, Diophantine approximation, differential algebra, the arithmetic of curves over finite fields, symmetry methods for differential equations, and computer algebra. We retrace this history, from Rosanes and Pasch to the present day, relying on the original sources. We point out an early erroneous criterion, due to Pasta, together with the counterexample that refutes it, given by Kellogg in 1912. We observe that most of the correct proofs, from Ostrowski's in 1919 to recent ones, rest on a single idea: restricting the functions to a sufficiently generic curve, and using the classical Wronskian along this curve. Finally, we describe how the question of which multivariate Wronskians are really needed was only settled in the twenty-first century.

Keywords: Wronskian; generalized Wronskian; linear independence; functions of several variables; Diophantine approximation; history of determinants.

1 Introduction

The notion of multivariate Wronskian is unsettling, for several reasons. Contrary to the univariate case, a finite family of functions admits many multivariate Wronskians, and many of them may vanish even when the family is linearly independent. Moreover, multivariate Wronskians have been defined in various

disguises, forgotten and rediscovered many times, and often studied merely as a technical tool rather than for their own sake. Their history spans more than a century. The qualitative criterion—a family of analytic functions is linearly dependent if and only if all determinants of a suitable kind vanish—was settled between 1912 and 1919; but a precise understanding of *which* determinants are really needed only dates from the early 21st century.

The reader is certainly familiar with the Wronskian through its use in the theory of linear differential equations. For a family of sufficiently regular functions $f_1, \ldots, f_n$ of one variable, the Wronskian of the family is the determinant $\det(f_j^{(k)})_{0\le k<n, 1\le j\le n}$ built on the successive derivatives of the functions. For functions of several variables, one can similarly build determinants on partial derivatives, but there is no canonical choice of the partial derivatives to be used.

This note follows on from our Monthly article (Bostan and Dumas 2010), which gave a short proof of the univariate criterion and of its multivariate extension. Shortly after its publication, Peter Olver drew our attention to his paper with Kamran and Milson (2000), which contains much more detailed information on the connection between Wronskians and the linear independence of functions of one and several variables, together with applications to nonlinear differential equations. Trying to situate these results with respect to earlier ones led us to read all the papers on multivariate Wronskians that we could find. The present note reports on this reading and completes its bibliography.

Two observations organize what follows. First, the earliest attempts restricted the functions to straight lines, and this does not work: the counterexample was given by Kellogg as early as 1912 (Section 2). Second, most correct proofs, from Ostrowski's in 1919 to recent ones, follow one pattern: restrict the functions to a curve that is generic enough for linear independence to survive, and use the univariate Wronskian along this curve (Section 6). Sections 3 to 5 follow the later history: polynomials in Diophantine approximation, the reduction of the number of determinants, and variations on the theme. We close with a remark on why the notion was rediscovered so often (Section 7).

2 The forerunners

Rosanes – 1872. Thomas Muir spent about fifty years collecting all the mathematical articles concerning the theory of determinants. The result is a monumental work in five volumes (Muir 1906, 1911, 1920, 1923, 1930); Wronskians are treated in Volumes III and IV, and in a whole chapter of the last volume, devoted to the years 1891 to 1919 (Muir 1930, Chap. IX). According to his Volume III, pages 481–484, Jakob Rosanes was the first to consider a functional determinant whose entries are partial derivatives, in a note dated Breslau, April 1872, and published in 1873 (Rosanes 1873). Moreover, he considered the more general derivation $\delta f = y_1 \partial f/\partial x_1 + y_2 \partial f/\partial x_2$ and the analogue $\det(\delta^i f_j)_{0\le i\le n-1, 1\le j\le n}$ of the Wronskian, where the f_j's are binary forms, that is, homogeneous polynomials in two variables. Rosanes's article, however, deals

with binary forms only, and is not concerned with the linear independence of general functions of several variables.

Pasch – 1874. Moritz Pasch, who is best remembered for his axiom of plane geometry, had been a close friend of Rosanes since their student years at the University of Breslau (now Wrocław, Poland); he spent his whole career at the University of Giessen. In a note signed in Giessen in December 1874 and published in 1875 (Pasch 1875), he starts from the classical theorem, established independently of the theory of differential equations by Christoffel and Frobenius, that λ functions f_1, …, f_λ of one variable x are linearly dependent if and only if their Wronskian vanishes identically. He reformulates it as follows: all the determinants of size λ extracted from the array

$$\begin{pmatrix} f_1 & \frac{df_1}{dx} & \frac{d^2f_1}{dx^2} & \frac{d^3f_1}{dx^3} & \cdots \\ f_2 & \frac{df_2}{dx} & \frac{d^2f_2}{dx^2} & \frac{d^3f_2}{dx^3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ f_\lambda & \frac{df_\lambda}{dx} & \frac{d^2f_\lambda}{dx^2} & \frac{d^3f_\lambda}{dx^3} & \cdots \end{pmatrix} \tag{1}$$

vanish as soon as the determinant made of the first λ columns vanishes; such a statement should then have a purely algebraic proof, which he derives from formulas of Christoffel and Frobenius. He adds that one easily sees that the same proofs hold for functions of several variables, if one sets

$$\delta f = y_1 \frac{\partial f}{\partial x_1} + y_2 \frac{\partial f}{\partial x_2} + \cdots \tag{2}$$

where y_1, y_2, … are arbitrary quantities: all the determinants of size λ extracted from the array

$$\begin{pmatrix} f_1 & \delta f_1 & \delta^2 f_1 & \delta^3 f_1 & \cdots \\ f_2 & \delta f_2 & \delta^2 f_2 & \delta^3 f_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ f_\lambda & \delta f_\lambda & \delta^2 f_\lambda & \delta^3 f_\lambda & \cdots \end{pmatrix} \tag{3}$$

vanish as soon as the determinant made of the first λ columns vanishes. For binary forms, he deduces that λ forms of the same degree m are linearly dependent if and only if the determinant $\det(\delta^i f_j)$ vanishes identically; for $\lambda = m+1$, this determinant reduces, up to a monomial factor, to

$$\mathfrak{D}(f_1, \ldots, f_{m+1}) = \det \left(\frac{\partial^m f_i}{\partial x_1^{m-j}\, \partial x_2^j} \right)_{1 \le i \le m+1, 0 \le j \le m} \tag{4}$$

which is essentially the determinant of the coefficients of the forms. Pasch is also clear about the limits of his method. For forms in more than two variables, he writes, it does not lead in general to consequences of the same scope: for three

forms, the vanishing of $\det(\delta^i f_j)$ only yields a relation $g_1 f_1 + g_2 f_2 + g_3 f_3 = 0$ whose coefficients g_i are linear forms in the variables, not constants. In one variable, on the other hand, Pasch's lemma had a long life: Bôcher used a version of it in his 1901 paper on the cases where the vanishing of the Wronskian implies linear dependence (Bôcher 1901), (Muir 1930, p. 278), the very paper on which Green would build his own theorem.

Pasta – 1908. Giuseppe Pasta, an Italian mathematician, published in 1908 a three-page note (Pasta 1908), dated Palermo, 10 January 1908, in the journal *Rivista di fisica, matematica e scienze naturali* of Cardinal Pietro Maffi. Apart from a one-sentence mention by Muir, who relates it to Pasch's paper (Muir 1930, p. 282), this note seems to have been overlooked, and it remained without issue.

Pasta looks, "in an elementary way", for necessary and sufficient conditions for n functions $f_1, \ldots, f_n$ of n variables, with continuous partial derivatives up to order $n-1$, to satisfy a linear relation with constant coefficients. He fixes a point a and considers the functions $F_i(t) = f_i(a + t(x-a))$ of the single variable t, whose successive derivatives are, symbolically, $F_i^{(k)} = (X_1 \partial/\partial x_1 + \cdots + X_n \partial/\partial x_n)^k f_i$ with $X_j = x_j - a_j$. He then takes for granted that the identical vanishing of the Wronskian of $F_1, \ldots, F_n$ is necessary and sufficient for the existence of a linear relation between the f_i's. At $t = 0$, this Wronskian is a polynomial in $X_1, \ldots, X_n$; expanding it by multilinearity, Pasta splits it into $\binom{n}{1}\binom{n+1}{2}\cdots\binom{k+n-1}{k}\cdots\binom{2n-2}{n-1}$ determinants which, up to monomials in the X_j's and numerical factors, read in modern notation $\det(\partial^{\alpha_i} f_j/\partial x^{\alpha_i})_{0\le i\le n-1, 1\le j\le n}$, where the multi-index α_i has length $|\alpha_i| = i$. He concludes that the family $f_1, \ldots, f_n$ is linearly dependent if and only if all these determinants vanish. Moreover, he introduces the *rectangular* matrix of size $n \times \binom{2n-1}{n}$ whose columns are the $\partial^\alpha f_i/\partial x^\alpha$ for the multi-indices α such that $|\alpha| < n$. The columns are sorted by the length of the multi-indices, then lexicographically among multi-indices of a given length, and the matrix is divided by vertical lines into blocks of columns sharing the same order of derivation:

$$\left(\begin{array}{c|ccc|ccc|ccc} f_1 & \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} & \cdots & \dfrac{\partial^{n-1} f_1}{\partial x_1^{n-1}} & \cdots \dfrac{\partial^{n-1} f_1}{\partial x^\alpha} & \cdots & \dfrac{\partial^{n-1} f_1}{\partial x_n^{n-1}} \\ f_2 & \dfrac{\partial f_2}{\partial x_1} & \cdots & \dfrac{\partial f_2}{\partial x_n} & \cdots & \dfrac{\partial^{n-1} f_2}{\partial x_1^{n-1}} & \cdots \dfrac{\partial^{n-1} f_2}{\partial x^\alpha} & \cdots & \dfrac{\partial^{n-1} f_2}{\partial x_n^{n-1}} \\ & & & & & & & & \\ f_n & \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} & \cdots & \dfrac{\partial^{n-1} f_n}{\partial x_1^{n-1}} & \cdots \dfrac{\partial^{n-1} f_n}{\partial x^\alpha} & \cdots & \dfrac{\partial^{n-1} f_n}{\partial x_n^{n-1}} \end{array}\right). \quad (5)$$

Pasta observes that the determinants above are obtained from matrix (5) by picking one single column in each block, in all possible ways.

Pasta's criterion is wrong, and not only because it ignores Peano's warnings (Peano 1889a,b) about the univariate Wronskian of non-analytic functions: it fails for polynomials. Consider $f_1 = 1$, $f_2 = x_1$, $f_3 = x_2$ as functions of

(x_1, x_2, x_3). These three functions are linearly independent, yet each of Pasta's determinants contains a row of second-order derivatives, which vanishes identically. Geometrically, restricted to any line $t \mapsto a + t(x - a)$, the three functions only span the space of affine functions of t, which has dimension two. The flaw lies in the restriction to straight lines, which produces, in the row of index i, derivatives of order *exactly* i; we shall see that the correct statement, made explicit by Roth, involves derivatives of order *at most* i (Sections 3 and 6). Four years later, Kellogg raised exactly this objection against Pasch's determinants.

Kellogg – 1912. The study of multivariate Wronskians in a general framework, in relation with linear independence, really begins with a communication of Oliver Kellogg to the Société mathématique de France, on 18 December 1912, whose two-page summary appeared in the *Bulletin* (Kellogg 1913). Muir, who records it, presents it as the third attempt, after Pasch's of 1874 and Pasta's of 1908, at an analogue of the Wronskian test for several variables (Muir 1930, p. 287). Kellogg observes that the conditions for the linear independence of functions of several variables do not seem to have been made precise, at least not as differential conditions. Gram's conditions, based on the matrix of the inner products of the functions, decide the question by means of definite integrals. Kellogg proposes instead conditions analogous to those given by Wronski's determinants in one variable.

He considers functions $f_1, \ldots, f_m$ in the variables $x = (x_1, \ldots, x_n)$ over a domain Ω, in which any two points can be joined by a continuous path of interior points. For simplicity of writing, he restricts himself to two variables, but claims that the results are general. In the same way as Pasta, but probably unaware of his work, Kellogg introduces a *rectangular* matrix K whose columns are the $\partial^\alpha f_i / \partial x^\alpha$ for the multi-indices α such that $|\alpha| < m$. For analytic functions, he states that a nontrivial linear relation with constant coefficients exists if and only if the rank of K is everywhere less than m; more precisely, if r denotes the maximal rank of the evaluations of K at the points of Ω—its *generic rank*—then all the f_i's are linear combinations with constant coefficients of r of them, and of no fewer. In other words, the rank of the family (f_i) is equal to the generic rank of K. For non-analytic functions, he states a result *à la* Frobenius and Bôcher: if K has generic rank $m - 1$ while the matrix associated with $f_1, \ldots, f_{m-1}$ has rank $m - 1$ at every point of Ω, then f_m is a linear combination with constant coefficients of $f_1, \ldots, f_{m-1}$. No proof is given.

Kellogg closes with two remarks. First, the vanishing of the Wronskian with respect to x and of the Wronskian with respect to y does not imply linear dependence, as shown by the three functions $x - 1$, $y - 1$, $xy - 1$. Second, the vanishing of Pasch's determinant $\det(\delta^i f_j)$, where $\delta f = h\, \partial f / \partial x + k\, \partial f / \partial y$, does not imply linear dependence either, as shown by the three functions x, y, 1—essentially the counterexample that refutes Pasta's criterion.

Kellogg's interest in linear dependence came from his work in nomography (Kellogg 1915), (Evesham 2010, Sec. 4.3); his paper on alignment nomograms is dated February 1913, two months after the communication. Nomography is the art of representing a functional relation between variables by a graphical

chart. The basic case is the following. A functional relation $F(u, v, w) = 0$ is represented by an alignment nomogram in three variables. The chart is made of three curves, parametrized respectively by u, v, and w. Given the values of u and v, the user puts a ruler on the points with parameters u and v on their respective curves, and reads off the value w where the ruler meets the third curve.

The theoretical question is to write, if possible, a given equation $F(u, v, w) = 0$ in a form that expresses the alignment of the three points, namely

$$F(u,v,w) = \begin{vmatrix} x_1(u) & y_1(u) & z_1(u) \\ x_2(v) & y_2(v) & z_2(v) \\ x_3(w) & y_3(w) & z_3(w) \end{vmatrix} = 0, \tag{6}$$

where $(x_1(u), y_1(u), z_1(u))$ is a parametrization of the first curve in projective coordinates, and similarly for the other ones. Since $F(u, v, w)$ is then a linear combination of the three functions $x_1(u)$, $y_1(u)$, $z_1(u)$, it satisfies a linear differential equation with respect to u of order at most 3, obtained as the univariate Wronskian in u of $F(u, v, w)$, $x_1(u)$, $y_1(u)$, $z_1(u)$. Expanding this Wronskian provides a linear relation between F and its partial derivatives F_u, F_{uu}, F_{uuu}, with coefficients depending only on u. As a consequence, Kellogg's 4×10 matrix of these four functions, with respect to the variables v and w, must have rank less than 4. This is a first necessary condition for the equation $F(u, v, w) = 0$ to admit a nomographic representation of the form (6). An easily accessible account of Kellogg's work can be found in Epstein (1958, Chap. VII, Sec. 7.3).

Green – 1916. Gabriel Marcus Green, at Harvard, announced his results in a note of the *Proceedings of the National Academy of Sciences* (Green 1916a), and developed them in a memoir of the *Transactions* (Green 1916b). His starting point is the theory of linear ordinary differential equations, where the Wronskian arises naturally because every solution of an equation of order n is a linear combination of a fundamental system of solutions. He generalizes the theory of linear dependence to functions of several variables, and the theory of linear differential equations to a class of systems of linear partial differential equations, with one unknown function and a finite-dimensional space of solutions.

His first theorem is *à la* Frobenius and Bôcher, whom he cites (Bôcher 1901). Consider n functions $y_1, \ldots, y_n$ of p variables and an $(n-1) \times n$ matrix whose rows are made of arbitrary partial derivatives $y^{(0)} = y$, $y^{(1)}, \ldots, y^{(n-2)}$ of the functions, and assume that the minor built on $y_1, \ldots, y_{n-1}$ vanishes nowhere in a connected region. Adjoin to this matrix all the first derivatives of its entries. If all the $n \times n$ minors of the new matrix that contain the previous minor vanish identically, then y_n is a linear combination of $y_1, \ldots, y_{n-1}$ with constant coefficients. Green is explicit on the main difficulty: *for functions of several variables it is not possible to define a single determinant which may properly be called a Wronskian.* He defines one only relative to a completely integrable system $\partial y^{(j)}/\partial u_k = \sum_i a_i^{(j,k)} y^{(i)}$, whose *primary derivatives* y, $y^{(1)}$, $\ldots$, $y^{(n-1)}$ form a *normal set*: each of them, except y, is obtained by a single

differentiation from another one, of order one less. The Wronskian of n solutions is the determinant of their primary derivatives, and it satisfies an analogue of Abel's formula. Green's third theorem further assumes that, together with each primary derivative, all the derivatives of lower order from which it can be obtained are primary derivatives. In modern terms, the multi-indices of the primary derivatives then form a lower set of $\mathbb{N}^p$; this is exactly the structure that Walker identified ninety years later as the relevant one (Section 4). Green, a projective differential geometer, points out that his third theorem gives a mechanical way to compute what Wilczynski called the seminvariants of such systems.

Ostrowski – 1919. Independently, Alexander Ostrowski, then in Göttingen, characterizes the linear independence of functions f_1, …, f_m of $x_1, \dots, x_n$ through a *square* matrix of size m (Ostrowski 1919a). To this end, he defines the linear operator

$$\Pi = \frac{\partial}{\partial u_0} + u_2 \frac{\partial}{\partial u_1} + u_3 \frac{\partial}{\partial u_2} + u_4 \frac{\partial}{\partial u_3} + \cdots = \frac{\partial}{\partial u_0} + \sum_{i=1}^{+\infty} u_{i+1} \frac{\partial}{\partial u_i}. \qquad (7)$$

To each renaming $x_i = u_{\alpha_i}$, in which the integers α_i are mutually distinct, he associates the determinant

$$\Delta_\Pi(f_1, \dots, f_m) = \begin{vmatrix} f_1(u) & f_2(u) & \dots & f_m(u) \\ \Pi f_1(u) & \Pi f_2(u) & \dots & \Pi f_m(u) \\ \vdots & \vdots & & \vdots \\ \Pi^{m-1} f_1(u) & \Pi^{m-1} f_2(u) & \dots & \Pi^{m-1} f_m(u) \end{vmatrix},$$

which Siegel would soon call the *Ostrowski determinant* (Siegel 1921b, p. 86), and which we call the *Ostrowskian*. Each renaming splits the indeterminates u into two blocks: on the one side, the x_i's, where $x = (x_i)$ varies in Ω; on the other side, the u_j's which are not assigned to any x_i, and which remain arbitrary. The definition of Π is more subtle than it may seem at first glance. For instance, for a function f of x_1 and x_2 and the renaming $x_1 = u_3$, $x_2 = u_5$, the first iterates of Π are

$$\Pi f(x) = u_4 \frac{\partial f}{\partial x_1}(x) + u_6 \frac{\partial f}{\partial x_2}(x),$$

$$\Pi^2 f(x) = x_2 \frac{\partial f}{\partial x_1}(x) + u_4^2 \frac{\partial^2 f}{\partial x_1^2}(x) + 2u_4 u_6 \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) + u_6^2 \frac{\partial^2 f}{\partial x_2^2}(x) + u_7 \frac{\partial f}{\partial x_2}(x).$$

Ostrowski proves that, for analytic functions, the nullity of Δ_Π, viewed as a polynomial in the u_j's, is equivalent to the linear dependence of the family. Since this result does not depend on the renaming, the condition can be given various forms. For instance, for three functions f_1, f_2, f_3 of two variables x_1, x_2, with the renaming $x_1 = u_0$, $x_2 = u_1$, the determinant Δ_Π is a polynomial in u_2 and u_3, and its nullity amounts to the vanishing of five 3×3 determinants.

Their first columns, that is those associated with f_1, read

$$\text{(i)} \begin{pmatrix} f_1 \\ \dfrac{\partial f_1}{\partial x_1} \\ \dfrac{\partial^2 f_1}{\partial x_1^2} \end{pmatrix}, \quad \text{(ii)} \; 2 \begin{pmatrix} f_1 \\ \dfrac{\partial f_1}{\partial x_1} \\ \dfrac{\partial^2 f_1}{\partial x_1 \partial x_2} \end{pmatrix} + \begin{pmatrix} f_1 \\ \dfrac{\partial f_1}{\partial x_2} \\ \dfrac{\partial^2 f_1}{\partial x_1^2} \end{pmatrix}, \quad \text{(iii)} \begin{pmatrix} f_1 \\ \dfrac{\partial f_1}{\partial x_1} \\ \dfrac{\partial f_1}{\partial x_2} \end{pmatrix},$$

plus two more obtained by exchanging the roles of x_1 and x_2 in (i) and (ii). Ostrowski does not restrict himself to the analytic case. For analytic functions, he implicitly uses the principle of analytic continuation; for the other cases, he refers to the works of Peano, Bôcher, and Curtiss (1908) on univariate functions, and claims that their sufficient conditions for the nullity of the Wronskian to imply dependence carry over (Ostrowski 1919a, p. 228). Where the operator Π comes from is explained in Section 6: it is the total derivative on the space of jets of functions of one variable, and Δ_Π is an ordinary Wronskian along a suitable curve.

Ostrowski gives no hint about the reason why he was interested in multivariate Wronskians. Walter Gautschi, who had been his doctoral student in Basel and who wrote an account of his life and work (Gautschi 2010), did not know either; he pointed out that Ostrowski was usually very conscientious in giving credit to those who had influenced his thinking, which suggests that the idea was his own. Gautschi advised us to look at the papers on multivariable algebra that Ostrowski wrote at about the same time (Ostrowski 1917, 1919c,b) (e-mail communication, 9 October 2013). We note that one of them deals with discriminants and resultants of binary forms (Ostrowski 1919c), the very objects of Rosanes and Pasch. As Gautschi writes in his account of Ostrowski's work (Gautschi 2010): *Notable are also his frequent attempts to establish results, even entirely classical ones, under the weakest assumptions possible, and his delight in finding proofs that are short and succinct.*

By the end of 1919, the qualitative question was therefore settled: Kellogg had stated the criterion in terms of the generic rank of his rectangular matrix, and Ostrowski had proved it with a square determinant depending on auxiliary indeterminates.

Muir's chapter on the Wronskians of the years 1891 to 1919 (Muir 1930, Chap. IX) shows how marginal these results were at the time. Muir notes that the writings of this period on the Wronskian exceed those of the previous twenty years by a half, and that many of them circle round the test for linear independence (Muir 1930, p. 275): most of his thirty or so entries deal with the univariate question raised by Peano. In several variables, he records Pasta's note and Kellogg's communication; Green appears with a note on functions of one variable (Muir 1930, pp. 290–291), his multivariate work only through two short items in the closing list of auxiliary, mainly applicational, writings (Muir 1930, p. 293); and Ostrowski's article, which settled the question, is not mentioned at all.

3 Wronskians of polynomials in Diophantine approximation

The particular case of multivariate Wronskians of polynomials occurs as a technical tool in number theory, more precisely in the study of irrationality measures. For a real number α and a positive exponent μ, consider the rational solutions p/q (with p, q integers, $q > 0$) of the inequality $|\alpha - p/q| < 1/q^\mu$. For $\mu = 1$, there are infinitely many of them. The larger μ, the more constraining the inequality, and there is a threshold $\mu(\alpha)$ beyond which the number of solutions is finite: this is the *irrationality exponent* of α. For a rational number α, one has $\mu(\alpha) = 1$. For an irrational number α, one has $\mu(\alpha) \geq 2$, as can be seen from the continued fraction expansion of α, which provides an infinite sequence $(p_n/q_n)_{n\geq 0}$, with $(q_n)_{n\geq 0}$ increasing, such that $|\alpha - p_n/q_n| < 1/q_n^2$.

For an irrational algebraic number α of degree $d \geq 2$, Liouville's famous result (Liouville 1844, 1851) asserts that every rational number p/q satisfies $|\alpha - p/q| > C/q^d$ for some constant $C > 0$ that only depends on α; hence $\mu(\alpha) \leq d$. The first significant improvement was made by Axel Thue (1909), who proved $\mu(\alpha) \leq d/2 + 1$.

From Thue onwards, the proofs follow a common scheme, the method of the auxiliary polynomial. Assume that α has many good rational approximations. One first constructs a polynomial $P(x_1, \ldots, x_k)$ with integer coefficients of controlled size, vanishing to a high order at $(\alpha, \ldots, \alpha)$; its existence follows from a counting argument on the coefficients, a version of the pigeonhole principle nowadays known as Siegel's lemma. One then picks k good approximations p_1/q_1, …, p_k/q_k of α, and looks for a derivative of P of small order that does not vanish at $(p_1/q_1, \ldots, p_k/q_k)$. This value is a nonzero rational number with controlled denominator, hence it cannot be too small; but Taylor's formula at $(\alpha, \ldots, \alpha)$ shows that it is very small. The contradiction bounds the irrationality exponent. The delicate step is the non-vanishing one, and this is where Wronskians enter.

Siegel – 1920. In his *Inauguraldissertation* presented at Göttingen in 1920, under the direction of Edmund Landau, and published in Siegel (1921a), Carl Siegel proves the inequality $\mu(\alpha) \leq 2\sqrt{d}$ for an irrational algebraic number of degree d, with an auxiliary polynomial $R(x, y)$ in two variables; the non-vanishing step rests on the classical Wronskian of polynomials in one variable, which Siegel proves again, for completeness, in the arithmetic form he needs (Siegel 1921a, Hilfssatz IV). (Siegel's proof is presented in Dickson 1929, chap. X.)

Siegel – 1921. In a second article (Siegel 1921b), received by the *Mathematische Annalen* on 2 January 1921, Siegel generalizes his thesis by using auxiliary polynomials in k variables. For a real algebraic number ξ of degree $d \geq 2$, let $a = e(\log d + 1/(2\log d))$. Siegel proves that the rational solutions p/q of $|\xi - p/q| \leq 1/q^a$, sorted by increasing denominators $q_1 \leq q_2 \leq \cdots$, are either finite in number, or so sparse that $\limsup_{\nu\to\infty} \log q_{\nu+1}/\log q_\nu = \infty$ (Siegel 1921b, Satz 3); as a consequence, infinitely many convergents p_λ/q_λ of the continued fraction expansion of ξ satisfy $|\xi - p_\lambda/q_\lambda| > 1/q_\lambda^a$ (Siegel 1921b, Satz 5). The exponent comes from the inequality $\min_{1\leq k\leq d} k\, d^{1/k} < e(\log d +$

$1/(2\log d)$) (Siegel 1921b, Hilfssatz 7), where k is the number of variables; $k=2$ gives back the exponent $2\sqrt{d}$ of the thesis. More generally, Siegel considers the approximation of ξ by the elements of a fixed number field and by algebraic numbers of bounded degree (Siegel 1921b, Satz 1 and 2).

With more than two variables, the non-vanishing step requires a criterion of linear independence for polynomials of several variables. This is Siegel's Hilfssatz 3, which is nothing but Kellogg's criterion for polynomials: *the number of linearly independent polynomials among g_1, …, g_p is equal to the rank of the matrix whose rows are the $(\partial^\nu g_1, \dots, \partial^\nu g_p)$ for $|\nu| < p$*, that is, of Kellogg's matrix K up to transposition. Siegel does not seem to be aware of Kellogg's note, but he cites Ostrowski's article, which had just appeared, and his proof takes half a page. He introduces Ostrowski's operator

$$\Pi = x_2 \frac{\partial}{\partial x_1} + x_3 \frac{\partial}{\partial x_2} + \dots + x_{k+1} \frac{\partial}{\partial x_k} + \dots ,$$

which corresponds to the renaming $x_i = u_i$, $1 \le i \le k$, and takes for granted that the number of linearly independent g_i's is equal to the rank of the matrix $\Omega = (\Pi^\kappa g_\lambda)_{0\le\kappa<p, 1\le\lambda\le p}$—a straightforward extension of Ostrowski's theorem from the determinant to the rank. If this rank is r, there exists a nonzero minor of Ω of size r; expanding it with respect to the products of powers of x_2, x_3, …, one gets the minors of K as coefficients, so that at least one minor of K of size r is nonzero. Conversely, any $r+1$ of the g_i's satisfy a linear relation with constant coefficients, hence so do their derivatives, and every minor of K of size $r+1$ vanishes. Therefore K has rank exactly r.

Siegel's Hilfssatz 4, the non-vanishing lemma, uses Hilfssatz 3 in an inductive elimination. The auxiliary polynomial $R(x_1, \dots, x_k)$ is written as $\sum_\lambda S_\lambda(x_1, \dots, x_{k-1})\,\varphi_\lambda(x_k)$ with linearly independent S_λ's. Hilfssatz 3 then provides a nonzero determinant $\Delta^{[1]}$ built on derivatives of the S_λ's with respect to x_1, …, x_{k-1}—a generalized Wronskian, in later terminology—and Cramer's rule yields a linear combination of derivatives of R equal to $\Delta^{[1]}(x_1, \dots, x_{k-1})\,\varphi_0(x_k)$. Repeating the process on $\Delta^{[1]}$ with respect to x_{k-1}, and so on, Siegel obtains a linear combination of derivatives of R equal to a product of polynomials in one variable each (Siegel 1921b, eq. (28)). Gauss's lemma then shows that none of the factors vanishes at the corresponding approximation, provided the heights of the successive approximations grow fast enough. This is, in germ, the inductive mechanism of what is now called Roth's lemma. The price to pay is that the heights of the successive approximations must lie in prescribed windows, with both lower and upper bounds on the ratios of their logarithms (Siegel 1921b, (44)–(45)); this is why Siegel's conclusion takes the form of an alternative between finitely many solutions and extremely sparse ones.

The two articles of Siegel thus show why multivariate Wronskians enter the theory: with two variables, one-variable arguments suffice; as soon as the number of variables exceeds two, which is how Siegel improves on the exponent $2\sqrt{d}$ for large d, a criterion of linear independence in several variables is needed.

Schneider – 1936. Theodor Schneider sharpened Siegel's theorem: for every exponent $\mu > 2$, the rational solutions p/q of $|\xi - p/q| \le 1/q^\mu$ are either finite

in number or so sparse that $\limsup_{\nu\to\infty} \log q_{\nu+1}/\log q_\nu = \infty$ (Schneider 1936, Satz 1). As Lang (1971, §8) observes, this is the first appearance of the right exponent, but only for approximations whose denominators grow extremely fast. Schneider works with polynomials in many variables, as Siegel did, but his non-vanishing lemma splits the auxiliary polynomial the other way round: writing it as $\sum_\beta g_\beta(z_1)\, U_\beta(z_2, \ldots, z_k)$, he only needs the ordinary Wronskian of the polynomials g_β in one variable, and he iterates (Schneider 1936, Hilfssatz 3). Siegel's multivariate criterion thus disappeared from Diophantine approximation for more than thirty years: Landau's textbook, Schneider and Dyson only need the criterion for polynomials in one variable.

Dyson – 1947. Freeman Dyson (1947) lowered Siegel's exponent to $\sqrt{2d}$; as Roth notes, Gelfond apparently obtained the same result independently (Roth 1955, p. 1). Dyson's contribution is a lemma that improves Siegel's estimates, and he still works with polynomials $R(x, y)$ in two variables, written as $\sum_j f_j(x)\, g_j(y)$ with as few terms as possible, for which univariate Wronskians are enough. His article has the great merit of being pedagogical: it clearly separates the proof of his lemma from an exposition of Siegel's method, which, he writes, runs closely parallel to Landau's account of Siegel's theorem in the third volume of his *Vorlesungen über Zahlentheorie* (Landau 1927, pp. 41–55); Landau follows a short paper of Siegel (1921c), and only uses Wronskians in one variable. Dyson attributes the square root in his bound to the two variables of his polynomial, and observes that his paper *probably represents the limit of what can be done with two variables only*; further progress, he adds, requires a study of polynomials in a larger number of variables (Dyson 1947, p. 239).

Roth – 1955. More than thirty years after Siegel, Klaus Roth (1955) obtained the optimal exponent: $\mu(\alpha) = 2$ for every irrational algebraic number α. (See Atiyah and Iagolnitzer 2003, p. 68–74 for a historical account by Roth himself.) The key point is the adaptation of Dyson's proof, based on bivariate polynomials and univariate Wronskians, to polynomials in many variables and generalized Wronskians. The non-vanishing step, now called Roth's lemma, is proved by induction on the number of variables, as Siegel's Hilfssatz 4, but it only requires lower bounds on the growth of the successive denominators, which can always be achieved by extracting a subsequence from an infinite sequence of approximations. In a footnote, Roth adds that, since writing his paper, he found that generalized Wronskians had been used by Siegel *in a similar connection*, and that Kellogg had stated the main result, his Lemma 1, without proof (Roth 1955, p. 3): Roth's generalized Wronskians are thus one more independent rediscovery. His original manuscript also used a combinatorial lemma of Schneider, which Davenport replaced by a simpler one (Roth 1955, p. 3).

Roth defines generalized Wronskians by means of the differential operators

$$\Delta = \frac{1}{i_1! \cdots i_n!} \left(\frac{\partial}{\partial x_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{i_n}$$

of order $i_1 + \cdots + i_n$. Following Roth's own indexing, a generalized Wronskian of $f_1, \ldots, f_m$ is the determinant of a matrix $(\Delta_\mu f_\nu)_{0\le\mu\le m-1,\, 1\le\nu\le m}$, where each Δ_μ

is an operator of the form above of order at most μ: the first row is made of the functions themselves, the row of index μ involves derivatives of order at most μ. In particular, a generalized Wronskian is, up to a constant, a minor of Kellogg's matrix, but some minors are not generalized Wronskians in Roth's sense. Roth shows that if m polynomials $f_1, \ldots, f_m$ in the n variables $x_1, \ldots, x_n$ are linearly independent, then at least one of their generalized Wronskians is nonzero. The proof relies on the substitution $x_1 = t$, $x_2 = t^b$, $x_3 = t^{b^2}$, $\ldots$, $x_n = t^{b^{n-1}}$, where the integer b is larger than all the partial degrees of the polynomials: since the exponents are written in base b, distinct monomials are mapped to distinct powers of t, and the problem is reduced to univariate Wronskians. This proof is reproduced by LeVeque (2002, Theorem 4.7(b)); other textbooks give inductive proofs (Cassels 1957; Schmidt 1980), (Hindry and Silverman 2000, Sec. D.6). In his 1959 revision of Landau's lectures, Walfisz replaced Siegel's theorem by Roth's, and with it Landau's one-variable Wronskians by Roth's generalized ones (Landau 1959, Hilfssatz 6).

Diophantine approximation later moved away from Wronskians. Observing that Dyson's and Bombieri's (1982) proofs of Dyson's lemma rest on *the theory of generalized Wronskians*, Viola (1985) replaced it by the theory of singular points of plane algebraic curves, while Esnault and Viehweg (1984) extended Dyson's lemma to several variables by methods of algebraic geometry.

4 Reductions

The number of generalized Wronskians grows very fast with the number of functions and of variables. Several authors remark that it is possible to lower the number of Wronskians whose nullity implies linear dependence.

Roth – 1955. Roth himself adds in a footnote (Roth 1955, p. 3): *It should perhaps be remarked (though it is immaterial to our argument) that the generalized Wronskians and their derivatives may satisfy identities, by virtue of which the vanishing of some of the generalized Wronskians implies the vanishing of the others.*

Berenstein, Chang, Li – 1993. In a note on entire functions of several complex variables (Berenstein et al. 1993), Carlos Berenstein, Der-Chen Chang and Bao Qin Li prove Roth's criterion (Berenstein et al. 1993, Cor. 2.2), together with a reduction *à la* Bôcher: if a generalized Wronskian of $f_1, \ldots, f_{m-1}$ does not vanish identically, the vanishing of $n(m-1)$ suitable Wronskians of $f_1, \ldots, f_m$ implies linear dependence (Berenstein et al. 1993, Th. 2.1). Their motivation is an application, a version in several variables of Borel's theorem on linear relations between exponentials, proved with Nevanlinna theory. They add that, after their note was written, they learned of Roth's work, through Cassels's book, where the same criterion had been found by a completely different method (Berenstein et al. 1993, Remark (5)).

Walker – 2005. Ronald Walker (2005) is the first author to determine a minimal set of generalized Wronskians needed to detect linear dependence; his historical

note mentions Kellogg, Siegel, Roth, Wolsson, and Berenstein, Chang and Li. He characterizes the *minimal* finite families of multi-indices whose generalized Wronskians suffice for meromorphic functions: they are multidimensional analogues of Young diagrams, which we call *Young-like sets*. A Young-like set is a lower set of $\mathbb{N}^n$: if it contains α, it contains every multi-index $\beta \leq \alpha$ (componentwise). This is exactly the condition of Green's third theorem.

Walker's Theorem 3.4 explains the role of these sets. If $A = \{\alpha^1, \ldots, \alpha^m\}$ is a Young-like set, the monomials $x^{\alpha^1}, \ldots, x^{\alpha^m}$ are linearly independent, and their generalized Wronskian associated with A does not vanish. If now $\beta = (b_1, \ldots, b_n)$ is a multi-index outside A, then for every $\alpha = (a_1, \ldots, a_n)$ in A there is an index j such that $a_j < b_j$, since A is a lower set; hence $D^\beta x^{\alpha^i} = 0$ for all i, and a row indexed by β in a generalized Wronskian of these monomials is zero. Thus no generalized Wronskian involving β can witness the linear independence of this monomial family, which shows that the generalized Wronskians associated with Young-like sets cannot be dispensed with in an (effective) criterion of linear dependence. In a later article (Walker 2011), Walker presents the same ideas in the more general framework of differential algebra (see Section 5), better suited to computer algebra: the data is a differential field $\mathbb{K}$ equipped with n commuting derivations, and its subfield of constants $\mathbb{C}$.

Walker first restates Roth's result. The derivatives in Kellogg's matrix correspond to the multi-indices of the simplex $|\alpha| < m$, and a generalized Wronskian $W_A(f)$ is associated with a subset A of size m of this simplex and with a family $f = (f_1, \ldots, f_m)$. Roth's Wronskians are associated with the sets G such that, for $1 \leq j \leq m$, G contains at least j multi-indices α with $|\alpha| < j$. Denoting by $\mathcal{G}_m$ the collection of these sets, Roth's result reads: the entries of f are linearly dependent over $\mathbb{C}$ if and only if $W_G(f) = 0$ for all $G \in \mathcal{G}_m$. Walker proves the same result with $\mathcal{G}_m$ replaced by the collection $\mathcal{Y}_m \subseteq \mathcal{G}_m$ of Young-like sets of size m, and he proves that every collection with this property contains $\mathcal{Y}_m$. The gain is significant. For $n = 2$, a Young-like set of size m is the Ferrers diagram of a partition of m, so that $\mathcal{Y}_m$ has $p(m) = \exp\big((\pi\sqrt{2/3} + o(1))\sqrt{m}\big)$ elements, whereas choosing one multi-index of each order $0, \ldots, m-1$ already provides $m!$ elements of $\mathcal{G}_m$. For fixed m, $|\mathcal{G}_m|$ grows like $n^{\binom{m}{2}}$, while $|\mathcal{Y}_m| = O(n^{m-1})$, since a Young-like set of size m involves at most $m-1$ variables. In terms of derivatives to be computed, Walker (2011, p. 606) estimates their number to be of order m^n for Roth's Wronskians, $m(\log m)^{n-1}$ for the Young-like ones, and m for his adaptive algorithm described below.

Walker then turns to the linear relations over $\mathbb{C}$ between the members of the family f. He shows that a certain Young-like set Y, of size at most m, determines these relations, and he designs an algorithm which takes f as input, computes Y, and returns a basis of the linear relations of the family with coefficients in the field $\mathbb{C}$ of constants. The computation is carried out by linear algebra over $\mathbb{K}$, which turns out to provide a result in $\mathbb{C}$. The set Y is such that the family of derivatives $(D^\alpha f)_{\alpha \in Y}$ is a basis over $\mathbb{K}$ of the space spanned by all the derivatives of f. Walker proves, by combinatorial arguments, that the number of derivatives visited by the algorithm is at most of order m.

As an example, with $\mathbb{K} = \mathbb{Q}(x, y)$ and the family

$$f = \left(\frac{x}{1-x}, \frac{4x-3y-xy}{(x-1)(y-1)}, \frac{y}{1-y}, \frac{y-x}{(x-1)(y-1)}, \frac{x+y-2xy}{(x-1)(y-1)} \right),$$

both $\partial_x f$ and $\partial_y f$ are linearly independent of f over $\mathbb{K}$; for a term order in which ∂_y comes before ∂_x, the associated Young-like set is $Y = \{(0,0), (0,1)\}$, hence the matrix

$$F = \begin{pmatrix} \frac{x}{1-x} & \frac{4x-3y-xy}{(x-1)(y-1)} & \frac{y}{1-y} & \frac{y-x}{(x-1)(y-1)} & \frac{x+y-2xy}{(x-1)(y-1)} \\ 0 & \frac{-3}{(y-1)^2} & \frac{1}{(y-1)^2} & \frac{1}{(y-1)^2} & \frac{1}{(y-1)^2} \end{pmatrix}$$

whose rows are f and $\partial_y f$. The linear relations over $\mathbb{K}$ between the columns of F are given by the rows of the matrix

$$R = \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{7}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix},$$

which have constant entries: they are the linear relations over $\mathbb{C}$ between the entries of f.

Etesse – 2023. In complex geometry, Brotbek (2017), studying the hyperbolicity of general hypersurfaces, evaluated the Wronskian of sections $s_0, \ldots, s_m$ of a line bundle on germs of curves γ, that is, the Wronskian of $s_0 \circ \gamma, \ldots, s_m \circ \gamma$ in local trivializations, in order to construct global jet differentials. To extend this construction from germs of curves to germs of p-dimensional submanifolds, Antoine Etesse (2023) needed generalized Wronskians compatible with changes of trivialization, so that they glue into global objects. He shows that the generalized Wronskians attached to sets of derivatives closed under taking sub-derivatives have this property, since they satisfy $W(gf_0, \ldots, gf_m) = g^{m+1} W(f_0, \ldots, f_m)$, and that they suffice to detect linear dependence; his proof follows the non-inductive scheme of our Monthly note. These sets are exactly Walker's Young-like sets: Etesse's theorem is Walker's sufficiency result, which he does not cite, found again for a geometric reason. He applies it to the algebraic degeneracy of holomorphic maps from $\mathbb{C}^p$ to Fermat hypersurfaces, and to foliations. The lower-set condition has thus appeared three times, independently: in Green's third theorem, in Walker's minimality result, and in Etesse's geometric construction.

5 Variations

So far, we have mostly considered analytic functions. The notion of (generalized) Wronskian can be defined and studied in other contexts, for instance formally in the setting of differential algebra, or for smooth real-valued functions.

Kolchin – 1941–1973. In the framework of *differential algebra*, Ellis Kolchin proves (Kolchin 1941, Appendix, Lemma 2*) that a finite family of elements of a differential field is linearly dependent over the field of constants if and only if its Wronskian vanishes. His proof mimics the classical proof by induction *à la* Bôcher (Bôcher 1900). Joseph Ritt (1950, Chap. 2, Item 22) gives a more conceptual proof in the same spirit.

Kolchin (1952) extends the notion of Wronskian to the case of several commuting derivations, in his Picard–Vessiot theory of partial differential fields, and again proves the equivalence between linear dependence over the constants and the vanishing of the generalized Wronskians. He gives two different proofs. The first one (Kolchin 1952, Lemma 1) reduces to the case of a single derivation. The second one (Kolchin 1973, Chap. II, Th. 1) proves a stronger result, valid for vector-valued functions and in arbitrary characteristic, by generalizing Ritt's proof with the help of a Kellogg-type matrix. Another proof of a result in the same vein, for a differential ring with a field of constants, is given in Di Vizio (2001, Appendix B). Kolchin later studied linear dependence over a projective variety (Kolchin 1974), a question going back to a remark of Ritt; Freitag, León Sánchez and Simmons (2016) extended it to complete differential varieties, from the standpoint of model theory, and recovered on the way the criterion by generalized Wronskians and Walker's minimal family.

The differential-algebraic statement contains the analytic one. On a connected domain, analytic functions form an integral domain, by the principle of analytic continuation, and their field of fractions is a differential field for the derivations $\partial/\partial x_1, \ldots, \partial/\partial x_n$, whose constants are the constant functions. Kolchin's theorem therefore yields the criterion for analytic functions at once, and the same holds for formal power series. As Mark van Hoeij remarked to us about our Monthly note (Bostan and Dumas 2010), this point of view also explains Peano's classical counterexample x^2, $x|x|$, two functions that are linearly independent on $\mathbb{R}$ although their Wronskian vanishes identically: no differential field of functions on $\mathbb{R}$ contains both, since $(x|x| - x^2)(x|x| + x^2) = 0$ while neither factor is zero. In the univariate case, the criterion is a standard result of textbooks on differential algebra (Kaplansky 1976, Th. 3.7), (Magid 1994).

Difference operators fit into the same picture, with the same caveat. For a single pseudo-linear operator θ, which covers derivations and automorphisms alike, Bronstein and Petkovšek (1996) define the rectangular "generalized Wronskian" $(\theta^i y_j)$, whose leading minor is the Wronskian or the Casoratian. Li, Wu and Zheng (2007) consider several commuting operators, each of them a derivation or an automorphism, and prove that invertible hyperexponential elements h_1, $\ldots$, h_n are linearly dependent over the constants if and only if the rectangular matrix $(\theta(h_i))$, where θ runs over the compositions of at most $n-1$ of these operators, has rank less than n (Li et al. 2007, Prop. 3.1): this is Kellogg's matrix, with mixed operators. They also observe that such criteria fail in general rings, already for the Casoratian: in the ring of complex sequences modulo the sequences of finite support, the sequences equal to 1 at the indices congruent to 0, respectively to 2, modulo 4, and to 0 elsewhere, are linearly independent over $\mathbb{C}$, but their Casoratian vanishes. As in van Hoeij's remark, zero divisors

break the criterion.

Hasse, Schmidt – 1937–1939. In characteristic $p > 0$, the criterion has to be reformulated. With ordinary derivations, the constants of $\mathbb{K}(x_1, \ldots, x_n)$ contain the x_i^p, so that the vanishing of the Wronskians only expresses linear dependence over $\mathbb{K}(x_1^p, \ldots, x_n^p)$: the polynomials 1 and x^p are linearly independent over $\mathbb{K}$, but their Wronskian vanishes. This is the situation covered by Kolchin's theorem (Kolchin 1973, Chap. II, Th. 1), and, in one variable, by Kaplansky (1976, Th. 3.7). To detect linear dependence over $\mathbb{K}$ itself, one uses the higher derivatives introduced by Hasse and F. K. Schmidt (1937), defined on monomials by $D^{(k)}x^j = \binom{j}{k}x^{j-k}$; in characteristic zero, $D^{(k)} = (1/k!)\, d^k/dx^k$, which is exactly Roth's normalization. In one variable, Schmidt (1939a) showed that elements f_1, …, f_m of an algebraic function field are linearly independent over the constants if and only if $\det(D^{(\varepsilon_i)}f_j) \neq 0$ for some integers $0 = \varepsilon_0 < \varepsilon_1 < \cdots < \varepsilon_{m-1}$. These orders may exceed $m-1$: for 1 and x^p, one needs $\varepsilon = (0, p)$. Schmidt's motivation was the theory of Weierstrass points of curves in positive characteristic (Schmidt 1939b). Stöhr and Voloch later recast his approach geometrically, and used such generalized Wronskians, with minimal orders ε_i, to bound the number of rational points of curves over finite fields, which yields a new proof of the Riemann hypothesis for these curves (Stöhr and Voloch 1986); for the role of Wronskians and of Schmidt's generalized Wronskians in algebraic geometry, see the survey of Gatto and Scherbak (2012). Garcia and Voloch (1987) showed that linear independence over $\mathbb{K}(x^{p^s})$ is detected by orders less than p^s. In several variables, Hsia and Wang (2004, Lemma 2) proved the analogue: functions that are linearly independent over the functions of $x_1^{p^s}$, …, $x_n^{p^s}$ admit a nonzero generalized Wronskian built on Hasse derivatives, with multi-indices of length at most $i\, p^{s-1}$ in the row of index i instead of Roth's bound i; their own motivation was, once again, an application, namely the *abc* theorem for function fields of higher dimension (see also Cherry and Toropu 2009, Th. 3.5). More recently, in order to decode multivariate multiplicity codes, Bhandari, Harsha, Kumar and Sudan (2024) needed the criterion in Roth's form, with Hasse derivatives, for polynomials whose partial degrees are less than the characteristic. Observing that the usual proofs extend to this setting, they nevertheless give a new one (Bhandari et al. 2024, Th. 5 and App. II), in order to extend an argument of Guruswami and Kopparty (2016) from one variable to several.

Wolsson – 1989. Unlike his predecessors, Kenneth Wolsson (1989b) does not confine himself to the *generic rank* of a family of functions f_1, …, f_m. He calls *critical point* of the family a point at which all the generalized Wronskians vanish, and *order* of a point the largest size of a subfamily admitting a generalized Wronskian that is nonzero at this point. His generalized Wronskians are Roth's, which he attributes to the books of LeVeque and Schmidt. He proves that on a connected open set where the order is constant and equal to r, the rank of the family is r (Wolsson 1989b, Th. 2), and he states, with proofs adapted from his univariate paper (Wolsson 1989a, Th. 4), that the order is a lower semicontinuous function (Wolsson 1989b, Th. 3). He then gives a necessary and sufficient condition for the linear dependence of functions with continuous partial

derivatives up to order $m-1$ (Wolsson 1989b, Th. 5). He cites LeVeque (2002) for the polynomial case and Wolfgang Schmidt (1980) for the rational case, but does not mention the analytic case. His approach is topological; in the analytic case, the topological assumption disappears, thanks to the principle of analytic continuation.

Ovsiannikov – 1978. The *method of prolongation*, designed by Élie Cartan in his study of partial differential equations (Chern and Chevalley 1952), led to the notion of *jet* developed by Charles Ehresmann (Ehresmann 1951a,b,c, 1952a,b), which formalizes the idea of a truncated Taylor expansion (see the introduction of Krasil'shchik and Verbovetsky (2011) for a historical account of the use of jet spaces in the study of differential equations). Lev Ovsiannikov's book on the group analysis of differential equations (Ovsiannikov 1982) works with prolongations in coordinates, without the intrinsic language of jet bundles, and Kellogg's matrix appears there in disguise. To compute the number of essential parameters of a family of mappings $\varphi(x,a)$, Ovsiannikov prolongs φ to the derivatives of order k and computes the generic rank ρ_k^* of the derivative of this prolongation with respect to the parameters a. These ranks form a nondecreasing sequence, and the number of essential parameters is the first ρ_k^* such that $\rho_k^* = \rho_{k+1}^*$ (Ovsiannikov 1982, §22.6). For $\varphi(x,a) = a_1 f_1(x) + \cdots + a_m f_m(x)$, the derivative with respect to a of the prolongation of order k is exactly Kellogg's matrix truncated at order k, and the number of essential parameters is the rank r of the family. Hence the generic ranks of the successive truncations of Kellogg's matrix increase strictly, then become constant from an index at most $r-1$; this can also be read off Walker's Young-like set, whose elements have order at most $r-1$. Peter Olver (2000) made this link explicit in the language of prolonged group actions: the Lie matrix of the abelian group $(x,u) \mapsto (x, u+t_1 f_1(x)+\cdots+t_m f_m(x))$ is Kellogg's matrix, and Ovsiannikov's stabilization theorem, of which Olver gives a corrected statement for smooth actions, then yields a multivariate Wronskian lemma for vector-valued functions (Olver 2000, Th. 6.19); since this group does not pseudo-stabilize, the rank stabilizes at order $m-1$ at most. Niky Kamran, Robert Milson and Peter Olver (2000) revisit this topic in the setting of prolonged group actions, and derive from generalized Wronskian matrices and their ranks algorithms that determine the annihilator of a finite-dimensional space of analytic functions; their motivation is the reduction of nonlinear partial differential equations to dynamical systems. Relying on Olver's results, they state that the generic ranks increase strictly until they stabilize at r (Kamran et al. 2000, Th. 3.2), and they observe that this fails at a given point x: the ranks $\rho_k(x)$ eventually reach r, but they may stall before; for the functions 1 and e^{x^k} at $x=0$, the rank remains equal to 1 up to order $k-1$, so that the order needed at a point can be arbitrarily large (Kamran et al. 2000, Th. 3.4).

6 Back to one variable

Looking back, most proofs of the multivariate criterion follow the same pattern: restrict the functions to a curve, chosen generic enough for linear independence to survive, and use the univariate Wronskian along this curve. The multivariate determinants only appear when this univariate Wronskian is expanded by the chain rule. Let us make this precise.

Let $t \mapsto \gamma(t)$ be a smooth curve in Ω. By Faà di Bruno's formula, for $k \geq 1$,

$$\frac{d^k}{dt^k} f(\gamma(t)) = \sum_{1 \leq |\alpha| \leq k} \frac{\partial^\alpha f}{\partial x^\alpha}(\gamma(t))\, P_{k,\alpha}\big(\gamma'(t), \dots, \gamma^{(k)}(t)\big), \tag{8}$$

where the $P_{k,\alpha}$ are universal polynomials. For $|\alpha| = k$, one has $P_{k,\alpha} = (k!/\alpha!)\,\gamma'(t)^\alpha$, while the terms with $|\alpha| < k$ involve the higher derivatives of γ. Consider the Wronskian of $f_1 \circ \gamma$, …, $f_m \circ \gamma$, whose row of index k is given by (8). Expanding it by multilinearity, one obtains a linear combination of determinants in which the row of index k is made of derivatives of order at most k; these determinants are precisely Roth's generalized Wronskians. Therefore, if the restrictions $f_1 \circ \gamma$, …, $f_m \circ \gamma$ are linearly independent for one curve γ, then (in the analytic case) their univariate Wronskian is nonzero, and so is at least one generalized Wronskian of f_1, …, f_m. The whole difficulty lies in the choice of γ.

Straight lines – Pasch, Pasta. For a line, $\gamma'' = 0$, and only the terms with $|\alpha| = k$ survive in (8): one gets Pasch's operator δ and the determinants of Pasta, with derivatives of order exactly k in the row of index k. But restriction to lines does not preserve linear independence, as Kellogg's example x, y, 1 shows. Pasch had noticed this weakness himself for forms in more than two variables; Pasta did not, and it is the flaw of his criterion.

Generic jets – Ostrowski. Ostrowski's operator Π becomes transparent if one reads u_0 as a parameter t, u_1 as a function $y(t)$, and u_{i+1} as the derivative of u_i for $i \geq 1$: then $(u_0, u_1, u_2, \dots) = (t, y, y', y'', \dots)$ are the coordinates on the space J^∞ of infinite jets of functions of one variable, and

$$\Pi = \frac{\partial}{\partial t} + y' \frac{\partial}{\partial y} + y'' \frac{\partial}{\partial y'} + \cdots$$

is the *total derivative* D_t: for a function g of finitely many of the u_i's,

$$(\Pi g)\big(t, y(t), y'(t), \dots\big) = \frac{d}{dt}\, g\big(t, y(t), y'(t), \dots\big).$$

A renaming $x_k = u_{\alpha_k}$ amounts to restricting the functions f_i to the curve

$$\gamma(t) = \big(y^{(\alpha_1 - 1)}(t), \dots, y^{(\alpha_n - 1)}(t)\big), \qquad \text{with the convention } y^{(-1)}(t) = t,$$

and the Ostrowskian $\Delta_\Pi(f_1, \dots, f_m)$ is the Wronskian of $f_1 \circ \gamma$, …, $f_m \circ \gamma$, in which the derivatives of y that are not used as coordinates remain free indeterminates. For instance, the renaming $x_1 = u_0$, $x_2 = u_1$ restricts the

functions to graphs $t \mapsto (t, y(t))$, and Siegel's renaming $x_i = u_i$ restricts them to curves $t \mapsto (y(t), y'(t), \dots, y^{(k-1)}(t))$. The key step of Ostrowski's proof, the fact that $\Pi g = 0$ only for functions g that are constant with respect to the u_i's, expresses that the total derivative has no nonconstant first integral on J^∞; in geometric language, the Cartan distribution of the infinite jet space, spanned here by D_t, is one-dimensional (Krasil'shchik and Verbovetsky 2011). Ostrowski himself notes that Π can be replaced by $c_1 u_1 \, \partial/\partial u_0 + c_2 u_2 \, \partial/\partial u_1 + \cdots$ with nonzero constants c_i, or by $\varphi(u_1) \, \partial/\partial u_0 + \varphi(u_2) \, \partial/\partial u_1 + \cdots$ with a function φ which is nowhere locally constant: only this property of Π matters (Ostrowski 1919a). Since all the jets of the curve are free, nothing is lost in the restriction.

Monomial curves – Roth, Bostan–Dumas, Ball–Parks–Pong. Along a curve $x_j = t^{w_j}$, a monomial x^α becomes $t^{w \cdot \alpha}$, with $w \cdot \alpha = w_1\alpha_1 + \cdots + w_n\alpha_n$. A family of polynomials or power series remains linearly independent as soon as $w \cdot \alpha \neq w \cdot \beta$ for all distinct exponents α, β occurring in the family, the univariate Wronskian of monomials with distinct exponents being, up to a monomial factor, a nonzero Vandermonde determinant (Bostan and Dumas 2010, Lemma 1). Roth's substitution is the case $w_j = b^{j-1}$, which separates monomials of bounded partial degrees thanks to the expansion of integers in base b; it works for polynomials only. The proof of Theorem 3 in Bostan and Dumas (2010) reduces the problem to leading monomials, and then uses the Vandermonde determinant of the linear forms $u \cdot \alpha^i$ in new indeterminates u_1, …, u_n: in effect, the weights w_j are generic, and the argument works for formal power series. Ball, Parks and Pong (unpublished preprint, 2016, revised 2018, https://arxiv.org/abs/1611.00248) consider power series in countably many variables X_p, indexed by the prime numbers, and map X_p to $p^{-s} = e^{-s \log p}$. A monomial $\prod_p X_p^{v_p}$ then becomes the Dirichlet term N^{-s}, with $N = \prod_p p^{v_p}$, on which the derivation $-d/ds$ acts by multiplication by $\log N$. The weights $\log p$ are linearly independent over $\mathbb{Q}$, by unique factorization, so that distinct monomials stay distinct, and the multivariate criterion follows from the Wronskian with respect to this single derivation.

A single derivation – Kolchin. In the same spirit, the first proof of Kolchin (Kolchin 1952, Lemma 1) replaces the derivations δ_1, …, δ_n by the single derivation $u_1\delta_1 + \cdots + u_n\delta_n$, where the u_i's are new elements, differentially algebraically independent. As with Ostrowski's operator, and unlike Pasch's δ, the coefficients are not constants: the direction of differentiation is allowed to turn. Kolchin's proof that this derivation creates no new constants is Ostrowski's argument again: consider the highest derivative of the u_i's that actually occurs (Kolchin 1952, §1).

Points instead of curves – Lee, Micchelli. Lee and Micchelli (2013) restrict the functions to finite sets of points rather than to curves. By the Cauchy–Binet formula, the determinant $\det(f_j(x_i))$ of the collocation matrix at m points x_1, …, x_m of $\mathbb{R}^d$ near a base point expands into a series of multivariate Wronskians at the base point, multiplied by generalized Vandermonde determinants in the points, so that the Wronskians govern the collocation determinants when the points coalesce (Lee and Micchelli 2013, Sec. 4). Their motivation is the "flat limit"

of interpolation by radial basis functions. Olver (2006) develops a multivariate divided-difference calculus, in which the "multivariate Wronskian matrix", that is, Kellogg's matrix of a vector-valued function, governs the limit of divided differences when the interpolation points coalesce. In algebraic complexity theory, Bhargava, Saraf and Volkovich (2026) also prefer points to derivatives. Generalized Wronskians, they observe, form an exponentially large collection of determinants of which only one is promised to be nonzero, which makes them hard to use efficiently, and Wronskians fail in small characteristic. They work instead with alternants, whose rows are evaluations at independent generic points, together with a "small witness" lemma: if k polynomials are linearly independent over a field $\mathbb{K}$, they remain linearly independent over the field of rational functions in all but at most $k-1$ of the variables. In characteristic zero, this lemma also follows from Walker's theorem: the Young-like set of k linearly independent functions has k elements, hence involves at most $k-1$ variables, and the corresponding nonzero generalized Wronskian only differentiates with respect to these variables, which forces linear independence over the functions of the other variables.

In all these proofs, the multivariate statement follows from the univariate one; the various classes of multivariate determinants reflect the various choices of curves. Straight lines give derivatives of exact order and an incorrect criterion; curved ones give Roth's generalized Wronskians and correct criteria; Walker's Young-like sets measure which of these determinants are really needed.

7 Conclusion: why so many rediscoveries?

Why was the same notion invented so many times, over a period of more than a century? A striking feature of this history is that almost every author met multivariate Wronskians on the way to something else. Rosanes and Pasch were interested in binary forms and invariant theory; Kellogg in nomograms; Green in completely integrable systems and Wilczynski's projective differential geometry; Siegel, Dyson and Roth in the approximation of algebraic numbers; F. K. Schmidt, and Stöhr and Voloch, in Weierstrass points and curves over finite fields; Hsia and Wang in the arithmetic of function fields; Kolchin in the Picard–Vessiot theory of partial differential fields; Ovsiannikov, and Kamran, Milson and Olver, in the symmetries and reductions of partial differential equations; Berenstein, Chang and Li in Nevanlinna theory; Walker in algorithms; Lee and Micchelli in interpolation by radial basis functions; Brotbek and Etesse in complex hyperbolicity and foliation theory; Kiselev in mathematical physics, where multivariate Wronskians serve as brackets of homotopy Lie algebras (Kiselev 2007); Bourgain, Demeter and Guo in harmonic analysis (Bourgain et al. 2017); Kübler and Polemarchakis in econometrics (Kübler and Polemarchakis 2017); Bhargava, Saraf and Volkovich in algebraic complexity theory; Bhandari, Harsha, Kumar and Sudan in coding theory; Ball, Parks and Pong in the algebraic independence of arithmetic functions. Only Pasta, Wolsson and perhaps Ostrowski seem to have looked at the question for its own sake.

For most of these authors, the multivariate Wronskian was a tool rather than an object, needed for one lemma and abandoned afterwards. The tool was rebuilt within each community, and the citations seldom crossed the boundaries: Kellogg cites Pasch but not Pasta; Green's note cites Bôcher, not Kellogg; Siegel cites Ostrowski, not Kellogg; even Muir, the most thorough bibliographer of determinants, does not record Ostrowski's article, and Lee and Micchelli, Ball, Parks and Pong, and Etesse are among the very few later authors who credit him (Lee and Micchelli 2013; Etesse 2023). Roth, and later Berenstein, Chang and Li, found out only after writing their papers that their criterion was already known. The condition that the derivatives form a lower set was found three times, by Green, Walker and Etesse, each time for a different reason. Our own Monthly note (Bostan and Dumas 2010) gave yet another proof of the multivariate criterion before we knew most of the papers discussed here; since then, it has become, for several recent authors in complexity theory, coding theory, harmonic analysis and economics, the only reference given for this criterion. The mathematical core, on the other hand, hardly changed: as shown in Section 6, the criterion keeps coming back to the univariate Wronskian along a well-chosen curve.

There is no reason to believe that this story is over: in 2025, Kiselev was still studying the identities satisfied by "Wronskians over multidimension" (unpublished preprints, https://arxiv.org/abs/2510.02145 and https://arxiv.org/abs/2511.03848), a distant echo of Roth's footnote on the identities between generalized Wronskians. The next rediscovery will presumably come, once again, from an application. We hope that this note will at least spare its author the trouble of looking for the proof.

Final words and acknowledgements

This work was begun in 2011 and, after several interruptions, completed in 2026.

We thank Peter Olver, whose message about his paper with Kamran and Milson (Kamran et al. 2000), shortly after the publication of our Monthly note (Bostan and Dumas 2010), was the starting point of this work; Mark van Hoeij for his remark on the differential-algebraic point of view; Jean-Paul Vincent for his questions and references on the case of positive characteristic; and Walter Gautschi for his advice on Ostrowski's papers. We also thank the readers of that note who sent us comments and references.

Statements and Declarations

Funding. This work was supported by the French–Austrian project EAGLES, funded by the Agence Nationale de la Recherche (ANR-22-CE91-0007) and the Austrian Science Fund (FWF, I6130-N).

Competing interests. The authors have no competing interests to declare that are relevant to the content of this article.

Data availability. Data sharing is not applicable to this article, as no datasets were generated or analysed during the current study.

Use of AI tools. A large language model (Claude, Anthropic) was used to assist in the final revision of this manuscript: checking statements against the primary sources, restructuring, and drafting some revised passages. All statements were checked by the authors against the original sources, and the authors take full responsibility for the content of this article.

References

Michael Atiyah and Daniel Iagolnitzer, editors. *Fields medallists' lectures*, volume 9 of *World Scientific Series in 20th Century Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, second edition, 2003.

Carlos Berenstein, Der-Chen Chang, and Bao Qin Li. A note on Wronskians and linear dependence of entire functions in $\mathbb{C}^n$. *Complex Variables Theory Appl.*, 24(1-2):131–144, 1993.

Siddharth Bhandari, Prahladh Harsha, Mrinal Kumar, and Madhu Sudan. Decoding multivariate multiplicity codes on product sets. *IEEE Trans. Inform. Theory*, 70(1):154–169, 2024.

Vishwas Bhargava, Shubhangi Saraf, and Ilya Volkovich. Linear independence, alternants and applications. *SIAM J. Comput.*, 55(1):28–64, 2026.

Maxime Bôcher. On linear dependence of functions of one variable. *Bull. Amer. Math. Soc.*, 7(3):120–121, 1900.

Maxime Bôcher. Certain cases in which the vanishing of the Wronskian is a sufficient condition for linear dependence. *Trans. Amer. Math. Soc.*, 2(2): 139–149, 1901.

Enrico Bombieri. On the Thue-Siegel-Dyson theorem. *Acta Math.*, 148:255–296, 1982. https://doi.org/10.1007/BF02392731.

Alin Bostan and Philippe Dumas. Wronskians and linear independence. *American Mathematical Monthly*, 117(8):722–727, 2010. https://doi.org/10.4169/000298910X515785.

Jean Bourgain, Ciprian Demeter, and Shaoming Guo. Sharp bounds for the cubic Parsell–Vinogradov system in two dimensions. *Adv. Math.*, 320:827–875, 2017.

Manuel Bronstein and Marko Petkovšek. An introduction to pseudo-linear algebra. *Theoret. Comput. Sci.*, 157:3–33, 1996. https://doi.org/10.1016/0304-3975(95)00173-5.

Damian Brotbek. On the hyperbolicity of general hypersurfaces. *Publ. Math. Inst. Hautes Études Sci.*, 126:1–34, 2017.

J. W. S. Cassels. *An introduction to Diophantine approximation*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 45. Cambridge University Press, New York, 1957.

Shiing-Shen Chern and Claude Chevalley. Obituary: Élie Cartan and his mathematical work. *Bull. Amer. Math. Soc.*, 58:217–250, 1952.

William Cherry and Cristina Toropu. Generalized ABC theorems for non-Archimedean entire functions of several variables in arbitrary characteristic. *Acta Arith.*, 136:351–384, 2009. https://doi.org/10.4064/aa136-4-4.

D. R. Curtiss. The vanishing of the Wronskian and the problem of linear dependence. *Math. Ann.*, 65(2):282–298, 1908.

Lucia Di Vizio. Sur la théorie géométrique des G-fonctions. Le théorème de Chudnovsky à plusieurs variables. *Math. Ann.*, 319(1):181–213, 2001. https://doi.org/10.1007/PL00004428.

Leonard E. Dickson. *Introduction to the theory of numbers*. University of Chicago Press, 1929.

Freeman J. Dyson. The approximation to algebraic numbers by rationals. *Acta Math.*, 79:225–240, 1947.

Charles Ehresmann. Les prolongements d'une variété différentiable. I. Calcul des jets, prolongement principal. *C. R. Acad. Sci. Paris*, 233:598–600, 1951a.

Charles Ehresmann. Les prolongements d'une variété différentiable. II. L'espace des jets d'ordre r de V_n dans V_m. *C. R. Acad. Sci. Paris*, 233:777–779, 1951b.

Charles Ehresmann. Les prolongements d'une variété différentiable. III. Transitivité des prolongements. *C. R. Acad. Sci. Paris*, 233:1081–1083, 1951c.

Charles Ehresmann. Les prolongements d'une variété différentiable. IV. Éléments de contact et éléments d'enveloppe. *C. R. Acad. Sci. Paris*, 234:1028–1030, 1952a.

Charles Ehresmann. Les prolongements d'une variété différentiable. V. Covariants différentiels et prolongements d'une structure infinitésimale. *C. R. Acad. Sci. Paris*, 234:1424–1425, 1952b.

L. Ivan Epstein. *Nomography*. Interscience Publishers, Inc., New York, 1958.

Hélène Esnault and Eckart Viehweg. Dyson's lemma for polynomials in several variables (and the theorem of Roth). *Invent. Math.*, 78:445–490, 1984.

Antoine Etesse. Geometric generalized Wronskians: applications in intermediate hyperbolicity and foliation theory. *Int. Math. Res. Not. IMRN*, 2023(10): 8251–8310, 2023.

Harold A. Evesham. *The history and development of nomography*. Docent Press, Boston, MA, 2010. With a preface by Scott Guthery.

James Freitag, Omar León Sánchez, and William Simmons. On linear dependence over complete differential algebraic varieties. *Comm. Algebra*, 44(6):2645–2669, 2016.

Arnaldo Garcia and José Felipe Voloch. Wronskians and linear independence in fields of prime characteristic. *Manuscripta Math.*, 59:457–469, 1987.

Letterio Gatto and Inna Scherbak. On generalized Wrońskians. In *Contributions to Algebraic Geometry. Impanga Lecture Notes*, pages 257–295. European Mathematical Society, Zürich, 2012.

Walter Gautschi. Alexander M. Ostrowski (1893–1986): His life, work, and students. In Bruno Colbois, Christine Riedtmann, and Viktor Schroeder, editors, *math.ch/100*, pages 257–278. European Mathematical Society, 2010.

Gabriel M. Green. On the linear dependence of functions of several variables and certain completely integrable systems of partial differential equations. *Proceedings of the National Academy of Sciences*, 2(4):209–214, April 1916a.

Gabriel M. Green. The linear dependence of functions of several variables, and completely integrable systems of homogeneous linear partial differential equations. *Trans. Amer. Math. Soc.*, 17(4):483–516, October 1916b. https://doi.org/10.2307/1988834.

Venkatesan Guruswami and Swastik Kopparty. Explicit subspace designs. *Combinatorica*, 36(2):161–185, 2016.

Helmut Hasse and Friedrich Karl Schmidt. Noch eine Begründung der Theorie der höheren Differentialquotienten in einem algebraischen Funktionenkörper einer Unbestimmten. *J. Reine Angew. Math.*, 177:215–237, 1937.

Marc Hindry and Joseph H. Silverman. *Diophantine geometry: an introduction*, volume 201 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.

Liang-Chung Hsia and Julie Tzu-Yueh Wang. The ABC theorem for higher-dimensional function fields. *Trans. Amer. Math. Soc.*, 356:2871–2887, 2004.

Niky Kamran, Robert Milson, and Peter J. Olver. Invariant modules and the reduction of nonlinear partial differential equations to dynamical systems. *Adv. Math.*, 156(2):286–319, 2000.

Irving Kaplansky. *An introduction to differential algebra*. Hermann, Paris, second edition, 1976. Actualités Scientifiques et Industrielles, No. 1251, Publications de l'Institut de Mathématique de l'Université de Nancago, No. V.

Oliver D. Kellogg. Sur l'indépendance linéaire des fonctions de plusieurs variables. *Bull. Soc. Math. France*, 41:19–21, 1913. Comptes rendus des séances, séance du 18 décembre 1912.

Oliver D. Kellogg. Nomograms with points in alignment. *Zeitschrift für Mathematik und Physik*, 63:159–173, 1915. Dated Feb. 23rd, 1913.

A. V. Kiselev. Associative homotopy Lie algebras and Wronskians. *J. Math. Sci.*, 141(1):1016–1030, 2007.

Ellis R. Kolchin. On the exponents of differential ideals. *Ann. of Math. (2)*, 42: 740–777, 1941.

Ellis R. Kolchin. Picard-Vessiot theory of partial differential fields. *Proc. Amer. Math. Soc.*, 3:596–603, 1952.

Ellis R. Kolchin. *Differential algebra and algebraic groups*. Academic Press, New York, 1973. Pure and Applied Mathematics, Vol. 54.

Ellis R. Kolchin. Differential equations in a projective space and linear dependence over a projective variety. In *Contributions to Analysis (A Collection of Papers Dedicated to Lipman Bers)*, pages 195–214. Academic Press, New York, 1974.

Joseph Krasil'shchik and Alexander Verbovetsky. Geometry of jet spaces and integrable systems. *J. Geom. Phys.*, 61(9):1633–1674, 2011. https://doi.org/10.1016/j.geomphys.2010.10.012.

Felix Kübler and Herakles Polemarchakis. The identification of beliefs from asset demand. *Econometrica*, 85(4):1219–1238, 2017.

Edmund Landau. *Vorlesungen über Zahlentheorie. Dritter Band: Aus der algebraischen Zahlentheorie und über die Fermatsche Vermutung*. S. Hirzel, Leipzig, 1927.

Edmund Landau. *Diophantische Gleichungen mit endlich vielen Lösungen*. Neu herausgegeben von A. Walfisz. Hochschulbücher für Mathematik, Bd. 44. VEB Deutscher Verlag der Wissenschaften, Berlin, 1959.

Serge Lang. Transcendental numbers and diophantine approximations. *Bull. Amer. Math. Soc.*, 77:635–677, 1971.

Yeon Ju Lee and Charles A. Micchelli. On collocation matrices for interpolation and approximation. *J. Approx. Theory*, 174:148–181, 2013. https://doi.org/10.1016/j.jat.2013.06.007.

William J. LeVeque. *Topics in number theory. Vol. I, II.* Dover Publications Inc., Mineola, NY, 2002. Reprint of the 1956 original [Addison-Wesley Publishing Co., Inc., Reading, Mass.], with separate errata list for this edition by the author.

Ziming Li, Min Wu, and Dabin Zheng. Testing linear dependence of hyperexponential elements. *ACM Commun. Comput. Algebra*, 41(1):3–11, 2007. https://doi.org/10.1145/1296772.1296773.

Joseph Liouville. Remarques relatives à des classes très-étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques; Nouvelle démonstration d'un théorème sur les irrationnelles algébriques. *Comptes rendus de l'Académie des Sciences*, 18:883–885, 910–911, 1844.

Joseph Liouville. Sur des classes très étendues de quantités dont la valeur n'est ni algébrique ni même réductible à des irrationnelles algébriques. *Journal de Mathématiques Pures et Appliquées*, 16:133–142, 1851.

Andy R. Magid. *Lectures on differential Galois theory*, volume 7 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1994.

Thomas Muir. *The theory of determinants in the historical order of development.* Macmillan and Co., limited, 1906, 1911, 1920, 1923.

Thomas Muir. *Contributions to the history of determinants 1900–1920.* Blackie and son, limited, 1930. Digitized by TIB Hannover, https://doi.org/10.14463/GBV:1017741204.

Peter J. Olver. Moving frames and singularities of prolonged group actions. *Selecta Math. (N.S.)*, 6(1):41–77, 2000.

Peter J. Olver. On multivariate interpolation. *Stud. Appl. Math.*, 116(2):201–240, 2006.

Alexander Ostrowski. Über die Existenz einer endlichen Basis bei gewissen Funktionensystemen. *Math. Ann.*, 78(1):94–119, 1917. https://doi.org/10.1007/BF01457091.

Alexander Ostrowski. Über ein Analogon der Wronskischen Determinante bei Funktionen mehrerer Veränderlicher. *Math. Z.*, 4(3-4):223–230, 1919a.

Alexander Ostrowski. Beweis der Irreduzibilität der Diskriminante einer algebraischen Form von mehreren Variablen. *Math. Z.*, 4(3-4):314–319, 1919b. https://doi.org/10.1007/BF01203019.

Alexander Ostrowski. Über eine neue Eigenschaft der Diskriminanten und Resultanten binärer Formen. *Math. Ann.*, 79(4):360–387, 1919c. https://doi.org/10.1007/BF01498416.

Lev Vasil'evich Ovsiannikov. *Group analysis of differential equations*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1982. Translated from the 1978 Russian original by Y. Chapovsky, translation edited by William F. Ames.

Moritz Pasch. Note über die Determinanten, welche aus Functionen und deren Differentialen gebildet werden. *Journal für die reine und angewandte Mathematik*, 80:177–182, 1875. Dated Giessen, December 1874.

Giuseppe Pasta. Sulla dipendenza lineare di n funzioni ad n variabili indipendenti. *Rivista di fisica, matematica e scienze naturali*, 17:71–73, 1908.

Giuseppe Peano. Sur le déterminant Wronskien. *Mathesis*, 9:75–76, 1889a.

Giuseppe Peano. Sur les Wronskiens. *Mathesis*, 9:110–112, 1889b.

Joseph F. Ritt. *Differential Algebra*. American Mathematical Society Colloquium Publications, Vol. XXXIII. American Mathematical Society, New York, N. Y., 1950.

Jakob Rosanes. Über Functionen, welche ein den Functionaldeterminanten analoges Verhalten zeigen. *Journal für die reine und angewandte Mathematik*, 75:166–171, 1873. Dated Breslau, April 1872.

Klaus F. Roth. Rational approximations to algebraic numbers. *Mathematika*, 2: 1–20; corrigendum, 168, 1955.

Friedrich Karl Schmidt. Die Wronskische Determinante in beliebigen differenzierbaren Funktionenkörpern. *Math. Z.*, 45(1):62–74, 1939a. https://doi.org/10.1007/BF01580273.

Friedrich Karl Schmidt. Zur arithmetischen Theorie der algebraischen Funktionen II. Allgemeine Theorie der Weierstraßpunkte. *Math. Z.*, 45:75–96, 1939b.

Wolfgang M. Schmidt. *Diophantine approximation*, volume 785 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980.

Theodor Schneider. Über die Approximation algebraischer Zahlen. *Journal für die reine und angewandte Mathematik*, 175:182–192, 1936.

Carl Siegel. Approximation algebraischer Zahlen. *Math. Z.*, 10(3-4):173–213, 1921a. https://doi.org/10.1007/BF01211608.

Carl Siegel. Über Näherungswerte algebraischer Zahlen. *Math. Ann.*, 84(1-2): 80–99, 1921b.

Carl Siegel. Über den Thueschen Satz. *Skrifter udgivne af Videnskabs-Selskabet i Christiania, I. Mat.-naturv. Klasse*, 1921(16):1–12, 1921c.

Karl-Otto Stöhr and José Felipe Voloch. Weierstrass points and curves over finite fields. *Proc. London Math. Soc. (3)*, 52:1–19, 1986.

Axel Thue. Über Annäherungswerte algebraischer Zahlen. *Journal für die reine und angewandte Mathematik*, 135:284–305, 1909.

Carlo Viola. On Dyson's lemma. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 12 (1):105–135, 1985.

Ronald A. Walker. Linear dependence of quotients of analytic functions of several variables with the least subcollection of generalized Wronskians. *Linear Algebra Appl.*, 408:151–160, 2005.

Ronald A. Walker. Regarding an adaptive algorithm for testing multivariate linear dependence. *Linear Algebra Appl.*, 434(2):605–613, 2011. https://doi.org/10.1016/j.laa.2010.09.024.

Kenneth Wolsson. A condition equivalent to linear dependence for functions with vanishing Wronskian. *Linear Algebra Appl.*, 116:1–8, 1989a.

Kenneth Wolsson. Linear dependence of a function set of m variables with vanishing generalized Wronskians. *Linear Algebra Appl.*, 117:73–80, 1989b.