Jacobi’s Bound *

Jacobi’s results translated in Kőnig’s, Egerváry’s and Ritt’s mathematical languages

François Ollivier (CNRS)

Dedicated to the memory of Harold W. Kuhn (1925-2014)

דער ניגון האָט באַקומען תּיקון	*The melody got out of Purgatory*
און דער למדן אויך!	*and the erudite too!*
י. ל. פּרץ	I. L. Peretz

רעזומע

יאַקאָביס רעזולטאַטען וועגן דעם חשבון פֿון דער אָרדענונג און די נאָרמאַלע פֿאָרמעס פֿון אַ דיפֿערענציאַלן סיסטעם ווערן דערוויזן אין די ראַמען פֿון דיפֿערענציאַלער אַלגעברע. מע גיט גאַנצע דערווייזונגען נאָך יאַקאָביס אַרגומענטען. דער עצם־טעאָרעם איז *יאַקאָביס גרענעץ*: די אָרדענונג פֿון אַ דיפֿערענציאַלן סיסטעם $P_1, \ldots, P_n$ איז ניט גרעסער ווי דער מאַקסימום $\mathcal{O}$ פֿון אַלע סך־הכּלען $\sum_{i=1}^n a_{i,\sigma(i)}$ פֿאַר אַלע אינדעקסען סובסטיטוציעס σ, ווּ $a_{i,j} := \mathrm{ord}_{x_{\sigma(i)}} P_i$, ד״ה *דעם טראָפּישן דעטערמינענט פֿון דער מאַטריצע* $(a_{i,j})$. די אָרדענונג איז פּונקט גלייך צו $\mathcal{O}$ אויב און נאָר אויב ס׳איז נישט נול *יאַקאָביס אָפּגעשניטן דעטערמינאַנט*.

יאַקאָבי האָט אויך געפֿינען אַ פּאָלינאָמישע קאָמפּליצירטקייט אַלגאָריטם צו חשבונען $\mathcal{O}$, ענלעך צו קונס „אונגערישן מעטאָד“ און אַ מין קירצערער וועג אַלגאָריטמען, שייך צו דעם חשבון פֿון די קלענסטע פּאָזיטיווע גאַנצע נומערן ℓ_i, אַזוי אַז מען קען רעכענען אַ נאָרמאַלע פֿאָרמע דיפֿערענצירנדיק די גלײַכונג P_i בלויז ℓ_i מאָל, על־פּי גענערישקייט היפּאָטעזן.

אַ פּאָר יסודותדיקע רעזולטאַטן וועגן סדר ענערונגען און די פֿאַרשיידענע נאָרמאַלע פֿאָרמעס וואָס אַ סיסטעם קען זיי נעמען זײַנען אויך דערציילט, אַרומנעמענדיק דיפֿערענציעלע רעזאָלווענטן.

Abstract

Jacobi’s results on the computation of the order and of the normal forms of a differential systems are expressed in the framework of differential algebra. We give complete proofs according to Jacobi’s arguments. The main result is *Jacobi’s bound*: the order of a differential system $P_1, \ldots, P_n$ is not greater than the maximum $\mathcal{O}$ of the sums $\sum_{i=1}^n a_{i,\sigma(i)}$, for all permutations σ of the indices, where $a_{i,j} := \mathrm{ord}_{x_{\sigma(i)}} P_i$, *viz.* the *tropical determinant of the matrix* $(a_{i,j})$. The order is precisely equal to $\mathcal{O}$ if and only if Jacobi’s *truncated determinant* does not vanish.

Jacobi also gave an algorithm to compute $\mathcal{O}$ in polynomial time, similar to Kuhn’s “Hungarian method” and some variants of shortest path algorithms, related to the computation of integers ℓ_i such that a normal form may be obtained, under genericity hypotheses, by differentiating ℓ_i times equation P_i.

Some fundamental results about changes of orderings and the various normal forms a system may have, including differential resolvents, are also provided.

François Ollivier (CNRS)
LIX UMR 7161
CNRS–École polytechnique
F-91128 Palaiseau CEDEX (France)
Email francois.ollivier@lix.polytechnique.fr

* With the support of Agence Nationale de la Recherche, project ANR-10-BLAN 0109 LÉDA (Logistic of Differential Algebraic Equations and Maths AM-Sud, project 13MATH-06 – SIMCA – Implicit Systems, Modeling and Automatic Control.

Introduction

History

In 1865 appeared in Crelle's journal a posthumous paper of Jacobi, edited by Borchardt [36] after a transcription by S. Cohn, followed by a second one in the volume *Vorlesungen über Dynamik*, edited by Clebsch in 1866 [37]. These two papers contain the following main result: *the order of an ordinary differential system of n equations P_i in n variables x_j is, at most, the maximum $\mathcal{O}$ of the "transversal" sums $\sum_{i=1}^{n} \mathrm{ord}_{x_{\sigma(i)}} P_i$ for all permutations $\sigma \in S_n$.* Known as *Jacobi's bound*, it mostly survived during the XX$^{\text{th}}$ century in the differential algebra community, thanks to J.F. Ritt, who gave a first complete proof of it in the linear case. It was extended by Kondratieva *et al.* [52, 53] to systems satisfying Johnson's regularity hypothesis [48]. But some important aspects were completely forgotten, such as a simplest normal form reduction, bounds on the order of derivations requested for computing normal forms, including differential resolvent, and a first polynomial time algorithm to solve the assignment problem, *i.e.* in our setting computing the bound faster than by trying the $n!$ permutations. A similar algorithm was rediscovered by Kuhn in 1955 [56]; Cohn [13] was the first, in 1983, to mention Jacobi's contribution. Jacobi's bound was rediscovered in 1960 by Volevitch [78] for differential operators and his simplest normal form reduction by Shaleninov in 1990 [72] and Pryce in 2001 [68] for the resolution of implicit DAE's. One may also mention that in modern vocabulary the expression of Jacobi's number $\mathcal{O}$ is known as the *tropical determinant* [60]. Two algorithms introduced by Jacobi to compute his *minimal canon* may be regarded as precursors of Dijstra's [19] and Bellman [3] shortest paths algorithms.

It is difficult to know precisely when were written the manuscripts related to Jacobi's bound. Jacobi did not use to date his writings. We know that it is a byproduct of his work on isoperimetric systems, evocated in a letter to his brother Moritz in 1836 [44]. The second part of his paper on the last multiplier [38], which appeared in 1845, contains a section devoted to these systems, where he promised to publish later his method for computing normal forms.

Proofs are often omitted in Jacobi's manuscripts. The style of some passages evocates a mathematical cookbook, providing computational methods without justifications, but no examples of precise differential systems are given, only general abstract families of systems like *isoperimetric equations*. It is clear that the efficiency is a constant preoccupation, even if it is not formalized. This work is closely related to Jacobi's interest in mechanics; there was at that time a strong need for fast computational tools, mostly for astronomical ephemerides [31][1].

[1]Jacobi himself had an experience in practical computing, on a smaller scale and in a different

In his 1840 letter to the Académie des Sciences de Paris [39], Jacobi said that he was working for some years on a publication that included his last multiplier method. One may guess that the various unpublished fragments were intended to take part in this never achieved ambitious project entitled *Phoronomia*. As Ritt guessed [70], the bound may have been suggested to him by his method for computing normal forms. We refer to our survey [67] for more historical details.

Aims of this paper

We present Jacobi's main results related to the order and normal forms of differential system, using the formalism of differential algebra. We prove them under hypotheses that could have been implicit in Jacobi's work and using, as far as possible, methods suggested in his work. Two particular aspects require attention.

Jacobi gives no detail about the nature of the functions he considers. Never does he describe the tools to be used to perform the requested eliminations. We restrict here to polynomial equations. It seems implicit that Jacobi's attention was foccussed on physical equations, generating prime differential ideals. However, we tried to consider the case of systems defining many components, whenever the extra work remained little. Jacobi's results related to *normal forms* of differential systems will be translated using *characteristic sets* of differential ideals.

Jacobi often considers implicit genericity conditions and will sometimes give first a "generic" theorem (*i.e.* a proposition that stands in some Zariski open set) followed by a second theorem describing the cases where the first assertion fails to be true. We will try to provide explicitely such conditions, most of the time expressed by the non vanishing of some Jacobian determinant.

Keeping in mind such particularities of the xx$^{\text{th}}$ century mathematical style, we recommend the reading of Jacobi's original papers, this text beeing only a partial commentary, completed with some technical parentheses devoted to contemporary developments.

The computation of the tropical determinant occcupies a large part of Jacobi's manuscripts and of this paper too. Contrary to those related to differential systems, Jacobi gave [36] very precise proofs of his combinatorial results. This may have dispense of longer comments, but a more carefull study shows that complexity issues are by no means obvious as well as the relations between Jacobi's *canons* and Egerváry's *covers*, a notion that allows us to make a link between Jacobi's shortest reduction and the choice of a ranking on derivatives used in dif-

field, when he published his *Canon arithmeticus* [43]. The revision of the half million numbers it contains requested the help of his friends and relatives [44], including Dirichlet's wife and mother!

ferential algebra algorithms. The implicit but fundamental role played by basic concepts and problems of graph theory must also be underlined.

Content

Section 1 introduces Jacobi's bound in the context of applying his last multiplier method to isoperimetrical equations. We limit ourselves here to an informal presentation of the genesis of the results. The next section 2 details Jacobi's algorithm, extended to rectangular matrices and studies its complexity.

It is followed by a short combinatorial parenthesis about the "strong bound" and reduction to order one, completed with algorithmic hints to get block decompositions. A second algebraic parenthesis is devoted to *quasi-regularity*, a key implicit assumption in Jacobi's proof and Lazard's lemma 5 contains some preliminary technical results of algebra related to *"Lazard's lemma"*, that will play a central part establishing the results on shortest reduction that characterizes some quasi-regular components.

Jacobi's bound is proved in section 6, together with the necessary and sufficient conditions for the bound to be reached, expressed by the system's *truncated determinant* ∇. The shortest normal form reduction is presented in section 7, followed in section 9 by a method for computing a characteristic set for some ordering, knowing one for some other ordering. The section 10 is devoted to resolvent computations and the last section 8 by a study of the various possible normal forms of a given system, including a complete description of the possible structures for zero dimensional linear ideals in two variables.

Notations

We will consider here equations $P_1, \ldots, P_m$ in the differential polynomials algebra $\mathcal{F}\{x_1, \ldots, x_n\}$, where $\mathcal{F}$ is a differential field of characteristic 0. The perfect differential ideal $\{P\}$ is denoted by $\mathcal{Q}$ and is equal to the intersection of prime components $\cap_{i=1}^{s} \mathcal{P}_i$. Jacobi's bound is denoted by $\mathcal{O}$, the notations Λ, λ_i $a_{i,j}$ α_i, β_j are introduced in definition 13, ∇, $\mathcal{J}^P$ in def. 65, $S_{s,n}$ in def. 1.

If $\mathcal{A}$ is the characteristic set of a differential (resp. algebraic) ideal, we denote by $H_{\mathcal{A}}$ the product of initials and separants (resp. initials only) of its elements.

By convention we write $F(n_1, \ldots, n_p) = O(G(n_1, \ldots, n_p))$, with $F, G : \mathbb{N}^p \mapsto \mathbb{N}$ if there exist constants A and B such that $F \leq AG + B$.

1 The last multiplier and isoperimetric systems

1.1 The last multiplier

The last multiplier method is first evocated by Jacobi in a short paper published in French in 1842, entitled "On a new principle of analytical mechanics" [40], followed by a second one in Italian in 1844: "On the principle of the last multiplier and its use as a new general principle of mechanics" [41]. It is not the place here to give details on the subject and we shall limit ourselves to a few hints in order to help understand the link with the genesis of Jacobi's bound. The reader will find illuminating illustrations on classical examples in Nucci and Leach's papers [64, 65].

Jacobi presents his last multiplier as a generalization of Euler's multiplier μ. If one has a Lagrange system in two variables:

$$\frac{\mathrm{d}x_1}{a_1(x_1,x_2)} = \frac{\mathrm{d}x_2}{a_2(x_1,x_2)}, \tag{1}$$

Euler's multiplier may be defined by the property $\mathrm{d}(\mu(1/a_1\mathrm{d}x_1 - a/_2\mathrm{d}x_2)) = 0$. Knowing the exact differential $\mu(1/a_1 dx_1 - 1/a_2 dx_2)$, finding a first integral for the system (1), which is a solution of $\frac{1}{a_1}\frac{\partial\omega}{\partial x_1} + \frac{1}{a_2}\frac{\partial\omega}{\partial x_2}$ is reduced to the computation of integrals $\int \frac{1}{a_1}\mathrm{d}x_1 = \omega + C_1(x_1)$ and $\int \frac{1}{a_2}\mathrm{d}x_2 = \omega + C_2(x_2)$.

In the case of a Lagrange system in n variables,

$$\frac{\mathrm{d}x_1}{a_1(x)} = \cdots = \frac{\mathrm{d}x_n}{a_n(x)}, \tag{2}$$

the last multiplier may be defined in the following way. Let ω_i, $1 \le i < n$, be first integrals for (2), any first integral ω is a solution of

$$\begin{vmatrix} \frac{\partial\omega}{\partial x_1} & \cdots & \frac{\partial\omega}{\partial x_n} \\ \frac{\partial\omega_1}{\partial x_1} & \cdots & \frac{\partial\omega_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial\omega_{n-1}}{\partial x_1} & \cdots & \frac{\partial\omega_{n-1}}{\partial x_n} \end{vmatrix} = 0.$$

Let us denote by D_i the Jacobian determinant

$$\begin{vmatrix} \frac{\partial\omega_1}{\partial x_1} & \cdots & \frac{\partial\omega_1}{\partial x_{i-1}} & \frac{\partial\omega_1}{\partial x_{i+1}} & \cdots & \frac{\partial\omega_1}{\partial x_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial\omega_{n-1}}{\partial x_1} & \cdots & \frac{\partial\omega_{n-1}}{\partial x_{i-1}} & \frac{\partial\omega_{n-1}}{\partial x_{i+1}} & \cdots & \frac{\partial\omega_{n-1}}{\partial x_n} \end{vmatrix} = 0.$$

The last multiplier μ is defined by

$$\mu \sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i} = \sum_{i=1}^{n} D_i \frac{\partial}{\partial x_i},$$

which is for $n = 2$ the definition of Euler multiplier.

Given any system of ordinary differential equations of order 1

$$x_i' = f_i(x), \quad 1 \le i \le n,$$

one may complete it with $t' = 1$ and associate to it the Langrange system

$$\frac{\mathrm{d}x_1}{f_1(x)} = \cdots = \frac{\mathrm{d}x_n}{f_n(x)} = \frac{\mathrm{d}t}{1}.$$

Jacobi's goal is explicitly exposed in 1842 [40]: having first remarked that for a system in two variables, the computation of solutions only requires integrations, he claims that his last multiplier method allows to generalize this result to any system of ordinary differential equations in n variables, provided that one already knows $n-1$ first integrals. This circumstance is of course very unlikely, but in 1840 he insisted on the inportance of a remark of Poisson [42] providing a method to compute a sequence of new first integrals, for any conservative mechanical system, that already possesses two first integrals, independently of energy.

The definition obviously depends of the choice of the first integrals, but also of the coordinate functions x_i. The multiplier is also given by the formula:

$$\mu = e^{-\int \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} \mathrm{d}t},$$

which may be interpreted as a Wrońskian, expressing the variation of a volume form along a trajectory, so that the last multiplier μ_2 associated to coordinates y_i must satisfy

$$\mu_2 = \left| \frac{\partial y_i}{\partial x_j} \right| \mu,$$

a formula that appears (up to a logarithm) in manuscript [II 23 a)] [37, formula 7)].

The application of the last multiplier method thus requires the knowledge of a normal form for a system of equations, and the result, after order 1 reduction, will depend on the chosen normal form, which may explain in part the interest of Jacobi in the various normal forms a given system may possess and differential elimination.

1.2 Isoperimetric equations

In 1844 and 1845, Jacobi published in two parts a 135 pages paper in latin, describing his last multiplier method for the integration of differential systems [38]. Among the examples of applications he gives, stands the *isoperimetric problem.*

"Let U be a given function of the independent variable t, the dependent ones x, y, z etc. and their derivatives x', x'', etc., y', y'', etc., z', z'', etc. etc. If we propose the problem of determining the functions x, y, z in such a way that the integral

$$\int U \mathrm{d}t$$

be *maximal or minimal* or more generaly that the differential of this integral vanish, it is known that the solution of the problem depends on the integration of the system of differential equations:

$$\begin{aligned} 0 &= \frac{\partial U}{\partial x} - \frac{d\frac{\partial U}{\partial x'}}{\mathrm{d}t} + \frac{d^2\frac{\partial U}{\partial x''}}{\mathrm{d}t^2} - \text{etc.,} \\ 0 &= \frac{\partial U}{\partial y} - \frac{d\frac{\partial U}{\partial y'}}{\mathrm{d}t} + \frac{d^2\frac{\partial U}{\partial y''}}{\mathrm{d}t^2} - \text{etc.,} \\ 0 &= \frac{\partial U}{\partial z} - \frac{d\frac{\partial U}{\partial z'}}{\mathrm{d}t} + \frac{d^2\frac{\partial U}{\partial z''}}{\mathrm{d}t^2} - \text{etc. etc.,} \end{aligned}$$

I will call these in the following *isoperimetric differential equations* ..."
[GW IV, p. 495]

For simplicity, we write x_1, ..., x_n, instead of x, y, z, etc. and denote by $P_i = 0$ the i^{th} isoperimetric equation. Jacobi noticed the difficulty of applying his last multiplier method if he could not first reduce the system to a normal form (see also [36, first section]). If the highest order derivative of x_i in U is $x_i^{(e_i)}$, the order of x_j in the i^{th} isoperimetric equation is at most $e_i + e_j$. If the e_i are not all equal to their maximum e, then we cannot compute a normal form without using *auxiliary equations* obtained by differentiating the i^{th} isoperimetric equation λ_i times, and a first problem is to determine minimal suitable values for the λ_i. In 1845, Jacobi had clearly in mind a thorough study of normal forms computation for he wrote: "I will expose in another paper the various ways by which this operation may be done, for this question requires many remarkable theorems that necessitate a longer exposition." [GW IV, p. 502]

Jacobi's method for computing a normal form may be sketched in the following way. Assume that the Hessian matrix $\left(\partial^2 U/\partial x_i^{(e_i)}\partial x_j^{(e_j)}\right)$ has a non zero determinant. We may further assume, up to a change of indices, that the sequence e_i is non decreasing and that the principal minors of the Hessian have full rank.

From the first isoperimetric equation P_1, as $\partial P_1/\partial x_1^{(2e_1)} = \pm\partial^2 U/\partial\left(x_1^{(e_1)}\right)^2 \neq 0$ one will deduce on some open set, using the implicit function theorem, an expression

$$x_1^{(2e_1)} := F_1(x_1, \ldots, x_1^{(2e_1-1)}, x_2, \ldots, x_2^{(e_1+e_2)}, \ldots, x_n, \ldots, x_n^{(e_1+e_n)}).$$

Using the first equation and its derivatives up to the order $e_2 - e_1$, together with the second equation, one may invoque again the implicit function theorem, using the fact that the Jacobian matrix of P_2 and $P_1^{(e_2-e_1)}$, with respect to the derivatives $x_1^{(e_1+e_2)}$ and $x_2^{(2e_2)}$, is equal to the second principal minor of the Hessian of U, which is assumed not to vanish. One deduces an expression

$$x_2^{(2e_2)} := F_2(x_1, \ldots, x_1^{(2e_1-1)}, x_2, \ldots, x_2^{(2e_2-1)}, x_3, \ldots, x_2^{(e_2+e_3)}, \ldots, x_n, \ldots, x_n^{(e_2+e_n)}).$$

Repeating the process, we get a last expression

$$x_n^{(2e_n)} := F_n(x_1, \ldots, x_1^{(2e_1-1)}, \ldots, x_n, \ldots, x_n^{(2e_n-1)}),$$

that may be obtained using each isoperimetric equation $P_i = 0$ and its derivatives up to order $\lambda_i := e_n - e_i$.

In this normal form, each variable x_i appears with the order $2e_i$, so that the order of the system is $2\sum_{i=1}^n e_i$. This appears to be both a special case of Jacobi's bound (see sec. 6) and of Jacobi's algorithm for computing normal forms (sec. 7), using the minimal number of derivatives of the initial equation, provided that the "system determinant" or "truncated determinant", here equal to the Hessian of U, does not vanish. In case of arbitrary equations P_i for which $a_{i,j} := \mathrm{ord}_{x_j} P_i$ can take any value, things become more complicated, starting with the computation of the bound $\max_\sigma \sum_{i=1}^n a_{i,\sigma(i)}$, that is the subject of our next section. But we easily understand how this particular simple example may have suggested the whole theory.

In section 2. of [36], we have restored a passage of [II/13 b), fo 2200] that quotes the isoperimetrical equations as an example for which all the transversal sums have the same value.

2 Computing the bound. Jacobi's algorithm

In algorithms, we will assume that matrices are represented by some array structure, so that one may get or change the value of some entry $a_{i,j}$ with constant cost.

The *assigment problem* has been first considered by Monge in 1781 [62], in the special case of the transportation problem (moving things from initial places to new places, minimizing the sum of the distances) and in a continuous setting (digging excavations somewhere in order to create some enbankment somewhere else). Such kind of problems reappeared in the middle of the xx$^{\text{th}}$ Century in the following form: n workers must be assigned to n tasks; assuming that the worker i has a productivity $a_{i,j}$ when affected at task j, how can we find an affectation $j = \sigma(i)$ that maximizes the sum of productivity indices?

At a meeting of the American Psychological Association in 1950, a participant described the following reaction: *"[he] said that from the point of view of a mathematician there was no problem. Since the number of permutations was finite, one had only to try them all and chose the best. [...] This is really cold confort for the psychologist, however, when one considers that only ten men and ten jobs mean over three and a half million of permutations."*[2] ([74] p. 8.) Jacobi did not consider the brute force method as a solution... and he gave a polynomial time algorithm!

The assignment problem also appears as a weighted generalization of the mariage or maximal bipartite matching problem: a graph describing couples of compatible boys and girls is represented by a $s \times n$ matrix of zeros and ones. The problem of computing the maximal number of compatible couples between these s boys and n girls amounts to computing a maximal transversal sum.

Kuhn's [56] and Jacobi's algorithms are quite similar. The main difference is the following. Jacobi remarks that if the columns of the matrix admit maxima placed in different rows, then their sum is the maximum to be found. He will then add minimal constants λ_i to the rows in order to get a matrix with this property. Kuhn considers integers α_i and β_j, with $\sum_{i=1}^{n} \alpha_i + \beta_i$ minimal, such that $a_{i,j} \leq \alpha_i + \beta_j$; this is called the *minimal cover*. He then uses Egerváry's theorem [23, 74]: $\sum_{i=1}^{n} (\alpha_i + \beta_i) = \max_{\sigma \in S_n} \sum_{i=1}^{n} a_{i,\sigma(i)}$. He will then look for the *minimal cover* α_i and β_j, adding constants to the rows and columns of the matrix. On this precise topic, I cannot do better than refeering to Kuhn's excellent—and moving—presentation [57].

Some of Jacobi's results could be extended with no extra work to the case of underdetermined systems. This is why we will expose his algorithm in the case of a $s \times n$ matrix A, with $s \leq n$. Entries are assumed to belong to an ordered additive commutative group, *i.e.* a commutative group with an order such that $x > y \Longleftrightarrow x - y > 0$. The special case of $-\infty$ entries will be considered in subsec. 4.1.

Definition 1. — *Let $s \leq n$ be two integers, we denote by $S_{s,n}$ the set of injections $\sigma : [1, s] \mapsto [1, n]$.*

[2]From some optimistic standpoint, it may have been a way to escape ethical issues raised by the use of psychology in management.

Let A be a $s \times n$ matrix of elements in $\mathbf{N} \cup \{-\infty\}$, *the* Jacobi number *of A is defined by the formula*

$$\max_{\sigma \in S_{s,n}} \sum_{i=1}^{s} a_{i,\sigma(i)}$$

and is denoted by $\mathcal{O}_A$.

Without further precision, a maximum $a_{i,j}$ *in A, is understood as being a maximal element in its column, i.e. such that $a_{i,j} \geq a_{i',j}$ $\forall 1 \leq i' \leq s$. We call* transversal maxima *a set of maxima placed in all different rows and columns. It is said to be a* maximal set of transversal maxima *if there is no set of transversal maxima with more elements in A.*

Let ℓ be a vector of s integers , we denote by $A + \ell$ the matrix $(a_{i,j} + \ell_i)$. We call a canon *a matrix $A + \ell$ that possesses s transversal maxima and also the vector ℓ itself.*

The partial order we will use on canons is defined by $\ell \leq \ell'$ if $\forall 1 \leq i \leq s$ $\ell_i \leq \ell'_i$.

Remark 2. — In the case of a square matrix, if ℓ is a canon and if $a_{i,\sigma(i)} + \ell_i$ is a maximal set of transversal maxima, then $\sum_{i=1}^{n} a_{i,\sigma(i)}$ is the *maximal transversal sum* we are looking for.

For $s < n$, one may compute $\mathcal{O} := \max_{\sigma \in S_n} \sum_{i=1}^{n} a_{i,\sigma(i)}$ by completing A with $n - s$ rows of zeros, which reduces the problem to the case of a square matrix.

Applied to a rectangular matrix, Jacobi's algorithm still returns the minimal canon λ[3], that will be used in section 7 to compute the shortest reduction in normal form, λ_i being the minimal number of times one needs to differentiate P_i in order to compute a normal form (under some genericity hypotheses). But the sum of the corresponding maxima, and so the order of this normal form, may fail to be equal to $\mathcal{O}$.

Example 3. — Consider the matrix

$$\begin{pmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & 2 & 0 \end{pmatrix}.$$

The minimal canon corresponds to $\lambda_1 = \lambda_2 = 0$; however the sum of the corresponding 3 pairs of transversal maxima are 2, 4 and 5 whereas the maximal transversal sum is 6. To find it, we may add two rows of zeros:

$$\begin{pmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

[3] Its existence is shown in prop. 4 below.

Then the minimal canon corresponds to $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 2$ and $\lambda_4 = 2$:

$$\begin{pmatrix} 1 & 0 & 3 & 4 \\ 1 & 2 & 3 & 1 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{pmatrix}.$$

The following proof of the unicity of the minimal (simplest) canon (assuming canons do exist) is due to Jacobi [36, th. IV, sec. 2].

PROPOSITION 4. — *Let A be a $s \times n$ matrix of elements in $\mathbf{N} \cup \{-\infty\}$, ℓ and ℓ' two canons:*

i) The s-uple $\ell'' := (\min(\ell_i, \ell'_i))$ is a canon for A.

ii) There exists a unique minimal canon for the ordering defined by $\ell \le \ell'$ if $\ell_i \le \ell'_i$ for $1 \le i \le s$.

PROOF. — Let $I := \{i \in [1, s] | \ell_i \ge \ell'_i\}$ and $I' := [1, s] \setminus I$. Let σ and σ' be the elements of $S_{s,n}$ corresponding to maximal sets of transversal elements for the canons $A + \ell$ and $A + \ell'$. We define $\sigma''(i) = \sigma(i)$ if $i \in I$ and $\sigma''(i) = \sigma'(i)$ if not, so that $\sigma''(i) \neq \sigma''(i')$ if $i \neq i'$ are both in I or both in I'. Furthermore, if $i \in I$ and $i' \in I'$, then $a_{i,\sigma(i')} + \ell_i \le a_{i',\sigma(i')} + \ell_{i'}$ (as $a_{i',\sigma(i')} + \ell_{i'}$ is maximal in $A + \ell$) and $a_{i',\sigma(i')} + \ell_{i'} < a_{i',\sigma(i')} + \ell'_{i'}$ (as $i' \in I'$), so that

$$a_{i,\sigma(i')} + \ell_i < a_{i',\sigma(i')} + \ell'_{i'}.$$

By construction, $a_{i,\sigma''(i)} + \ell''_i$ is maximal in its column. The inequality above implies then that $\sigma''(i) = \sigma(i) \neq \sigma'(i') = \sigma''(i')$: σ is an injection. This achieves the proof of i), of which ii) is a straightforward consequence. ■

2.1 Jacobi's algorithm

See [36, § 3] for Jacobi's proof of the algorithm and [37, § 1] for a detailed example.

Input: an $s \times n$ matrix A. The case $s = 1$ is trivial, so we assume $s \ge 2$.
Output: its minimal canon λ if it exists or "failed".

Step 1. **(Preparation process)** — Increase each row of the least integer such that one of its elements become maximal (in its column). This step requires $O(s^2)$ operations. It produces a new matrix $A' = A + \ell_0$ such that each row possesses a maximal element. The number of transversal maxima in A' is at least 2, that corresponds to the case where all elements in row i and column j are maximal (except perhaps the element $a'_{i,j}$). If $s = 2$, the problem is solved.

If $s > 2$, we enter step 2 with A', ℓ_0 and a set of exactly 2 transversal maxima.

Step 2. — **a)** For readability, we may reorder the rows and columns, so that the transversal maxima in A' are the elements $a'_{i,i}$ for $1 \leq i \leq r < s$. *Left* (resp. *right*) columns are columns $j \leq r$ (resp. $j > r$). *Upper* (resp. *lower*) rows are columns $j \leq r$ (resp. $j > r$), as bellow.

We define the starred elements of A' as being the transversal maxima $a'_{i,i}$[4].

$$
\begin{array}{r}
\\
\textit{upper}\left\{\vphantom{\begin{array}{c}*\\ \ddots\\ *\end{array}}\right.\\
\textit{lower}\left\{\vphantom{\begin{array}{c}\\ \\ \\ \end{array}}\right.
\end{array}
\begin{array}{c}
\overbrace{\hphantom{*\ \ \ddots\ \ *}}^{\textit{left}}\ \overbrace{\hphantom{*\ \ \ddots\ \ *}}^{\textit{right}}\\
\left(\begin{array}{ccc|c}
* & & & \\
 & \ddots & & \\
 & & * & \\ \hline
 & & & \\
 & & & \\
 & & &
\end{array}\right)
\end{array}
$$

b) Assume that there is a maximal element located in a right column and a lower row. We can add it to the set of transversal maxima. If it now contains s elements, the process is finished. If not, we repeat step 2.

c) We say that there is a *path*[5] from row i to row i' if there is a starred maximum in row i, equal to some element of row i' located in the same column, or recursively if there is a path from row i to row i'' and from row i'' to row i'. We also recursively define *first class rows* as being upper rows with at least a right maximal element, or rows to which there is a path from a first class row. The construction of the set of first class rows, together with paths to them from rows with a right maximalterm may be done in $O(sn)$ operations, using an array F of booleans with $F.i =$ true if row i belongs to the first class (we cannot afford looking into a list).

d) If there is no lower row of the first class, we go to step e).

Assume that there is a lower row of the first class, then there is a path to it from an upper row i_0 containing a right maximal element $a_{i_0,\beta}$. Let it be a path of lenght p, consisting of rows i_0 to i_p, so that for all $0 \leq \alpha < p$ the element $a_{i_{\alpha+1},i_\alpha}$ is equal to the starred element a_{i_α,i_α}. We can then construct a set of $p+1$ transversal

[4]Jacobi defined also the maximal elements in right columns as "starred"; we prefer to reserve this denomination to left transversal maxima to underline the specific roles played by these two sets of maxima in the algorithm.

[5]This notion is closely related to that of *increasing path*, as defined in [32], which explains the choice of that word in the translation of *transitum datur* in [36].

maxima by replacing a_{i_α,i_α} by $a_{i_{\alpha+1},i_\alpha}$, for $0 \leq i \leq p$, and adding $a_{i_0,\beta}$ to the list, as illustrated in the figure below, where the new increased set of transversal elements is written in **bold letters**.

$$\left(\begin{array}{ccccc|c}
a^*_{i_0,i_0} & \cdots & & & & \boldsymbol{a}_{i_0,\beta}\\
\| & & & & & \\
\boldsymbol{a}_{i_1,i_0} & \cdots & a^*_{i_1,i_1} & & & \\
& & \| & & & \\
& & \boldsymbol{a}_{i_2,i_1} & \ddots & & \\
& & & \vdots & & \\
& & & \ddots & a^*_{i_{p-1},i_{p-1}} & \\ \hline
& & & & \| & \\
& & & & \boldsymbol{a}_{i_p,i_{p-1}} &
\end{array}\right)$$

If $p+1 = n$, we have finished, if not, we repeat step 2 a), c) and d) untill no lower row of the first class is found.

The next lemma is given by Jacobi in [36, sec. 3].

Lemma 5. — *The maximal number of transversal maxima in A' is r iff there is no lower right maximum, nor lower row of the first class.*

Proof. — The algorithm above proves that the given condition is necessary. Let us assume that there is no lower right maximum, nor lower row of the first class, but that there exists a set T of $r' > r$ transversal maxima a_{μ_i,ν_i}. Some of these must belong to left columns and the others to upper rows. As there are only r left columns, $h \geq r' - r$ of them, say a_{μ_i,ν_i} $1 \leq i \leq r'$, are upper right maxima. From their first class rows, one can build paths as above, starting from rows μ_i, $1 \leq i \leq h$, and considering only maximal elements a_{μ_i,μ_i} and a_{μ_i,ν_i}, $1 \leq i \leq r'$. The sets of rows in such paths are disjoints, for T is a set of transversal maxima. As there is no lower first class rows, these h paths must end with some element $a_{j,j}$, such that there is no $\nu_i = j$, *i.e.* no maximum in T located in the same column. So that there are in T h right elements and at most $r-h$ left elements, a contradiction. ∎

Definition 6. — *We define the rows of the third class as being the lower rows and all the rows from which there is a path to a lower row. The rows not in the first or third class form the* second class.

We increase the third class rows by the smallest integer μ such that one of their elements become equal to a right maximum or a starred element located in some first class of second class row. This may be done in at most $O(sn)$ operations.

We then iterate step 2 with a new matrix $A' = A + \ell$ and a new vector ℓ.

If this element belongs to a second class row, this row will go to the third class and the cardinal of the second class will decrease. If it belongs to the first class, then at the next step there will be a lower right maximal element (if it is in a right column) or a first class lower row (if it belongs to a left one), so that the number of transversal elements will increase. Let p be the number of starred elements, there are at most $p-2$ second class rows, at least 1 first class and 1 upper third class row, so that we need at most $p-1$ iterations to exaust the second class and increase the number of transversal maxima, which can occur at most $s-2$ times. So step 2 is iterated at most $\sum_{p=2}^{s-1}(p-1) = (s-1)(s-2)/2$ times before the algorithm returns the requested result.

If the integer inlement of A are bounded by C, then each integer operation requires $O(\ln C)$ bit operations.

This leads for $s = n$ to a $O(n^4)$ complexity, which corresponds to that of Kuhn's original Hungarian algorithm [7, Ch. 4.1 p. 77].

Theorem 7. — *The above algorithm returns the minimal canon, if it exists, in at most $O(s^3 n)$ elementary operations. Assuming that the elements in the matrix are integers of size C, it requires at most $O(s^3 n \log C)$ bit operations.*

Proof. — The termination and complexity of the algorithm have already been proved. We only have to show that the obtained canon is the smallest.

The proof, that follows Jacobi's, relies on the following lemma.

Lemma 8. — *Let λ be the minimal canon for A', assume that $a_{i,i}$ $1 \le i \le r$ form the set of transversal maxima in $A' + \ell'$, $\ell' < \lambda$, with respect to which the classes are defined at step e) of the algorithm and that there is no lower right maxima nor first class lower row. Then there is no* unchanged *row of the third class in $A' + \ell'$, i.e. a third class row of index i_0 with $\lambda_{i_0} = \ell'_{i_0}$.*

Proof of the lemma. — We assume $a_{i,i}$ $1 \le i \le r$ to be a maximal set of transversal maxima in $A' + \ell'$. Let $a_{i,\sigma(i)} + \lambda_i$, $1 \le i \le s$ be a maximal set of transversal maxima in $A' + \lambda$.

If row i is an unchanged row of the third class, the element $a_{i,\sigma(i)}$ is maximal (in its column) in $A' + \lambda$, and so it is also maximal in $A' + \ell'$. It cannot be an upper right element, for then the row i would be of the first class, and it cannot be lower right, for third class rows are considered only if no lower right maximum is found a step 2. b). So, $1 \le \sigma(i) \le r$.

Let H denote the set of integers $1 \le i \le r$ such that row i is an unchanged row of the third class. For $i \in H$ the elements $a_{i,\sigma(i)} + \ell'_i$ and $a_{\sigma(i),\sigma(i)} + \ell'_{\sigma(i)}$ are both maximal elements of the column $\sigma(i)$. So, the row $\sigma(i)$ must be unchanged in $A' + \lambda$ and, as there is a path from it to row i, it belongs to the third class: $\sigma : H \mapsto H$ is a bijection. Hence, there is no unchanged lower row i' of the third class, for we would have $\sigma(i') \in H$ and $i' \notin H$.

Let the row i_0 be and unchanged row of the third class. Due to the third class definition, we can find a sequence of third class rows i_α, $0 \leq \alpha \leq p$, such that:
i) $a_{i_{\alpha+1}, i_\alpha} = a_{i_\alpha, i_\alpha}$;
ii) rows i_α, $1 \leq \alpha < p$ are upper rows;
iii) row i_p is lower.
The row i_0 is unchanged. Using i), we prove by recurrence that all rows i_α, $0 \leq \alpha \leq p$ are unchanged. As row i_p is lower, we arrive to a final contradiction, that concludes the proof of the *lemma* ∎

Each row of a canon must contain a maximal element. So $\lambda \geq \ell$, where ℓ is the vector produced by the preparation process. As there is no unchanged row of the third class, and as, during step 2) e) we increase third class rows by the minimal integer requested to change the class partition, the canon returned by the algorithm must be the minimal canon λ. This concludes the proof of the theorem. ∎

Remarks. — 9) Let the $n \times n$ integer matrix A be defined by $a_{i,j} = (n-1)^2 - (i-1)(j-1)$, one shall apply step 2 precisely $(n-1)(n-2)/2$ times. E.g. for $n = 4$, the matrix is:

$$\begin{array}{c} \mathrm{I}\\ \mathrm{III}\\ \mathrm{III}\\ \mathrm{III}\end{array}\begin{pmatrix} 9 & \mathbf{9} & 9 & 9\\ \mathbf{9} & 8 & 7 & 6\\ 9 & 7 & 5 & 3\\ 9 & 6 & 3 & 0\end{pmatrix},$$

where we have indicated the classes of the rows on he left, the starred maxima being in bold. Step 2 shall be applied 3 times and here is the sequence of matrices it produces, with the increment of each rows, the last being the canon.

$$\begin{array}{c} \mathrm{I}\\ \mathrm{II}\\ \mathrm{III}\\ \mathrm{III}\end{array}\begin{pmatrix} 9 & 9 & \mathbf{9} & 9\\ 10 & \mathbf{9} & 8 & 7\\ \mathbf{10} & 8 & 6 & 4\\ 10 & 7 & 4 & 1\end{pmatrix}\begin{array}{c} 0\\ 1\\ 1\\ 1\end{array}, \quad \begin{array}{c} \mathrm{I}\\ \mathrm{III}\\ \mathrm{III}\\ \mathrm{III}\end{array}\begin{pmatrix} 9 & 9 & \mathbf{9} & 9\\ 10 & \mathbf{9} & 8 & 7\\ \mathbf{11} & 9 & 7 & 5\\ 11 & 8 & 5 & 2\end{pmatrix}\begin{array}{c} 0\\ 1\\ 2\\ 2\end{array}, \quad \begin{pmatrix} 9 & 9 & 9 & \mathbf{9}\\ 11 & 10 & \mathbf{9} & 8\\ 12 & \mathbf{10} & 8 & 6\\ \mathbf{12} & 9 & 6 & 3\end{pmatrix}\begin{array}{c} 0\\ 2\\ 3\\ 3\end{array}$$

10) Jacobi gave the criterion of lemma 5 as a way to help finding a maximal set of transversal maxima, but seemed to assume that, most of the time, one will find them by inspection, as he does for the 10×10 matrix provided as an example in [37]. So, our presentation is a modern reinterpretation that does not fully reflect the spirit of a method intented for hand computation.

In his analysis of Jacobi's contribution [57], Kuhn made a distinction in the algorithm between a *Kőnig step*, *i.e.* finding the maximal number of transversal maxima, and an *Egerváry step*, *i.e.* increasing the number of transversal maxima. This underlines the deep similarity with the Hungarian method and is coherent with Jacobi's presentation.

Jacobi completed his work with a few more algorithms, allowing to compute the minimal canon, knowing an arbitrary canon or a maximal set of s transversal maxima ([53], lemma 3), which he did not use in his study of differential systems that will be exposed in the next section 3, together with some complements about algorithms and complexity. We will conclude this paragraph with the case of the maximal matching problem, followed by some properties of covers that will be needed in sections 6 and 7.

2.2 Maximal matching

Egerváry's results were influenced by the following theorem of Kőnig [54, 55, 73] (see also [53, lemma 2]), which was in turn inspired by previous works of Frobenius [25, 26] (one may refer to Schrijver [74] for historical details. It is an easy consequence of Jacobi's criterion for characterizing maximal transversal families of maxima (lemma 5).

THEOREM 11. — *Let A be a $s \times n$ matrix of zeros and ones, with $s \leq n$, m be the smallest integer such that the ones are all located in the union of p rows and $m-p$, then m is the maximal diagonal sum $\mathcal{O}$ in A.*

PROOF. — It si easily seen that $\mathcal{O} \leq m$: in any diagonal sum, at most p "ones" belong to these p rows, $m-p$ to these $m-p$ columns, and the sum is $m-p+p-q$, where q is the number of ones that belong both to these rows and these columns.

To prove $\mathcal{O} \geq m$, one may use Jacobi's construction. Assume that we have $r = \mathcal{O}$ diagonal starred ones, that we may assume to be $a_{1,1}$, ..., $a_{r,r}$. We can use lemma 5 with the following change : the ones are the maximal elements, the zeros the non-maximal elements. According to the lemma, there are no lower right ones. Let p be the number of first class rows, that we may assume to be rows 1 to p. Rows $p+1$ to r do not belong to the first class and so they contain no ones located in columns 1 to p nor $r+1$ to n. Rows $r+1$ to s belong to the third class and, in the same way cannot contain ones in columns 1 to p nor $r+1$ to n, which would contradict the minimality of r. So, all the ones belong to p rows and $r-p$ columns, as illustrated by the figure below.

$$
\begin{array}{ll}
\overbrace{\hphantom{* \quad \ddots \quad *}}^{r} & \\
\left(\begin{array}{c|c|c}
\multicolumn{3}{l}{* \quad \ddots \quad *} \\
\hline
 & * \quad \ddots \quad * & \\
0 & & 0 \\
 & &
\end{array}\right) & \left.\vphantom{\begin{array}{c}*\\ \ddots\\ *\end{array}}\right\} p \\
\hphantom{0 \quad} \underbrace{\hphantom{* \quad \ddots \quad *}}_{r-p} &
\end{array}
$$

This concludes the proof. ∎

2.2.1 A naïve algorithm

Algorithm 12. **Input data**: A $s \times n$ matrix of zeros and ones.
Output: A maximal transversal sum in A.

Classes of rows will be constructed here, not with respect to maximal elements, but with respect to “ones”.

To solve the problem, we only have to construct the set of first order rows, with a cost of $O(sn)$ operation, and to apply lemma 5, which may only occur $\mu - 1$ times, where μ is the size of the matching; hence a total cost of $O(s^2 n)$ operations for the whole algorithm. This is the complexity of an improved version of Jacobi algorithm (see below 3.2).

But it is posible to lower the complexity with a slight modification, due to Hopcroft and Karp[32]. See below 3.1.1.

2.3 Covers

If not stated overvise, we consider in this section only square matrices A. Covers do not appear in Jacobi’s paper and it is interesting to investigate their relations with canons.

Definition 13. — *We call a* cover *for A the data of two vectors of integers $(\mu_1,\ldots,\mu_n)$ and $(\nu_1,\ldots,\nu_n)$, such that $a_{i,j}\le\mu_i+\nu_j$. A cover μ, ν is a* minimal cover *if the sum $\sum_{i=1}^n\mu_i+\nu_i$ is minimal.*

Let μ, ν and μ', ν' be two covers for A, then we say that they are equivalent if there exists some integer γ such that $\mu'_i=\mu_i+\gamma$ and $\nu'_j=\nu_j-\gamma$.

Proposition 14. — *i) $\Longrightarrow$ A cover μ, ν of A is minimal iff there exists a permutation σ such that $a_{i,\sigma(i)}=\mu_i+\nu_{\sigma(i)}$.*

$\Longleftarrow$ Let us assume that there is no such permutation σ. Then, the entries $a_{i,j}$ with $a_{i,j}=\mu_i+\nu_j$ belong to p rows and $m-p$ columns, with $m<n$, that we may suppose to be rows $1,\ldots,p$ and columns $1,\ldots,m-p$. Let $e:=\min_{i=p}^n\min_{j=m-p}^n(\mu_i+\nu_j-a_{i,j})$, we define $\mu'_i:=\mu_i$ if $i\le i\le p$ and $\mu'_i:=\mu_i-e$ if $p<j\le n$, $\nu'_j:=\nu_j+e$ if $1\le j\le m-p$ and $\nu'_j=\nu_j$ if $m-p<j\le n$: μ', ν' is a cover, with $\sum_{i=1}(\mu'_i+\nu'_i)=\sum_{i=1}^n(\mu_i+\nu_i)-(n-m)e$.

ii) Let ℓ be a canon for A. We will denote by $L:=\max_{i=1}^s\ell_i$, $\mu_i=L-\ell_i$ and $\nu_j=\max_{i=1}^s a_{i,j}-\mu_i$. The vectors μ, ν form a minimal cover for A, that we define as the cover associated to the canon ℓ.

iii) Let μ, ν be a cover for A, the integers $\ell_i:=\max_k\mu_k-\mu_i$ form a canon for A, that will be called the canon associated to the cover μ, ν.

Proof. — i) We have, by hypothesis, $\sum_{i=1}^n\mu_i+\nu_i=\sum_{i=1}^n a_{i,\sigma(i)}$ and, by definition of a cover, $\sum_i\mu_i+\nu_i\ge\sum_i a_{i,\sigma(i)}$, hence the minimality of the cover μ, ν.

ii) By construction, $\mu_i+\nu_j\ge a_{i,j}$, so that α, β is a cover. Minimality is a consequence of i), remarking that if $a_{i,\sigma(i)}+\ell_i$ form a maximal transversal sum in $A+\ell$, then $a_{i,\sigma(i)}=\mu_i+\nu_i$.

iii) By i) there exists a permutation σ such that $a_{i,\sigma(i)}=\mu_i+\nu_{\sigma(i)}$, so that $a_{i,\sigma(i)}+\ell_i=\nu_{\sigma(i)}+L\ge a_{i',\sigma(i)}+\ell_{i'}$ for all $1\le i'\le n$, so that ℓ is a canon. ∎

Definition 15. — *The minimal cover associated to the minimal canon will the called the* Jacobi cover *or the* canonical cover.

Remarks. — 16) Knowing the minimal canon, we may obviously compute the Jacobi cover in $O(sn)$ operations.

17) If A is a matrix of non negative integers, then any cover is equivalent to a cover of non negative integers. It is easilly seen that $\min_i\mu_i+\min_j\nu_j\ge\min_{i,j}a_{i,j}$, so that one just has to define $\mu'_{i'}:=\mu_{i'}-\min_i\mu_i$ and $\nu'_{j'}:=\nu_{j'}-\min_i\mu_i$ to be sure that $\mu'_i\ge0$ and $\nu'_j\ge0$.

18) In our definition, covers are vectors of integers and not non negative integers as assumed by Egerváry [23]. It is easily seen that any matrix of non negative integers admits a minimal cover of non negative integers, but we cannot restrict to this case, even if differentiation orders are non negative, because we will need to consider in sec. 4.1 $-\infty$ entries.

The matrix

$$\begin{pmatrix} 0 & 2 \\ -1 & 1 \end{pmatrix}$$

admits no minimal cover of non negative integers: $\mu_1 + \nu_1 = 0$ would imply $\mu_1 = 0$, so that $\nu_2 \geq 2$, which contradicts $\mu_2 + \nu_2 = 1$.

To try to get an equivalent cover of non negative integers, we may add $\min_j \nu_j$ to all the μ_i and substract it from all the ν_j; doing so, one gets a new cover μ', ν' with $\min_j \nu'_j = \nu_{j_0} 0$ and $\min_i \mu'_i$ minimal. If some μ'_{i_0} remains negative, no equivalent cover of non negative integers exists, but then $a_{i_0,j_0} \leq \mu'_{i_0} + \nu'_{i_0} < 0$.

19) For any integer matrix A of zeros and ones, all minimal covers are equivalent to minimal covers that are vectors of zeros and ones, as well as their associated canon.

20) In the Jacobi cover, some $\alpha_{i_0} = 0$ must be 0.

If α, β is the Jacobi cover of A, β, α is not in general the canonical cover of A^{t}, and the canonical cover of A is not even in all cases equivalent to β, α. For A such that $a_{i,j} = \alpha_i + \beta_j$, the canonical covers of A (resp. A^{t}) will be equivalent to α, β (resp. β, α) and they will even be equal if some $\alpha_{i_0} = 0$ (resp. some $\beta_{j_0} = 0$).

But the canonical cover of the matrix

$$A := \begin{pmatrix} 1 & 1 & 1 & \mathbf{1} & 1 \\ 1 & 1 & \mathbf{1} & 1 & 1 \\ 1 & \mathbf{1} & 0 & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & \mathbf{0} \end{pmatrix}$$

is $(1,1,1,0,0)$, $(1,0,0,0,0)$, that of A^{t} is $(1,1,0,0,0)$, $(1,1,0,0,0)$ and the two covers have no obvious relation.

21) Assume that $a_{i,i}$ form a maximal transversal sum in the matrix A. For any cover μ, ν of A, there is an elementary path from line i_0 to line i_1 of the canon of A (sec. 2.1 step 2 c) iff $a_{i_1,i_0} = \mu_{i_1} + \nu_{i_0}$.

22) If A is a canon, then A^{t} is not necessarily a canon, but we can easily compute the associated cover μ_i, ν_j for A, then ν_j, μ_i will be a cover for A^{t} and the associated canon $\ell_i := \max_k \mu_k - \mu_i$ will be easily computed, *viz.* in $O(n^2)$ operations.

23) Let A be a matrix. We may compute its simplest canon B and the simplest canon C of B^{t}. The matrix C^{t} is the simplest matrix D that is a canon for A and such that D^{t} is a canon for A^{t}, meaning that minimal quantities are added to the rows and columns of A to obtain D.

24) Let A be a matrix of zeros and ones, and μ, ν a cover of A. The non zero elements of A are located in the rows i with $\mu_i \neq 0$ and columns j with $\nu_j \neq 0$. We recover Kőnig's theorem. Reciprocally, if R and C are two sets of rows and

columns containing all the ones appearing in A, with $\sharp R+\sharp C$ minimal, then $\mu_i = 1$ if $i \in R$ and $\nu_j = 1$ if $j \in C$ defines a minimal cover for A.

So, in this case, a cover is exactly equivalent to the associated minimal sets of rows and columns.

25) Let A be a matrix of non negative integers and $a_{i,\sigma(i)}$ be transversal maxima for A, then this matrix has at most

$$\prod_{i=1}^{n}(a_{i,\sigma(i)} + 1)$$

covers of non negative integers and this number is reached if all elements of A except these transversal values are 0 or less. Assume that the $a_{i,j}$ belong to some ordered group G where there is no infinite strictly decreasing sequence, then any square matrix admits a finite number of covers, up to equivalence.

On the other hand, if there are an infinite number of values $c \in G$ such that $\min_{j\neq\sigma(i_0)} a_{i_0,j} < c < a_{i_0,\sigma(i_0)}$, there are an infinite number of non equivalent covers $\mu'_{i_0} := c$, $\mu_{i\neq i_0} := \mu_i$; $\nu'_{\sigma(i_0)} := a_{i_0,\sigma(i_0)} - c$, $\nu'_{j\neq\sigma(i_0)} := \nu_j$.

The next proposition will help to clarify the situation and to compute, in case of need, non Jacobi covers and their canons.

Proposition 26. — *Let A be a matrix, we assume without loss of generality, that $a_{i,i}$ form a maximal transversal sum. Let μ, ν be a minimal cover for A. Using remark 21, we will use the reflexive transitive closure of the path relation $\prec$, defined on rows of the associated canon of A and the transposed relation $\prec^{\mathrm{t}}$ defined on the rows of the associated canon of A^{t}, i.e. the columns of A. Rows and colums will be denoted by their indexes. For convenience, we repeat the rules:*
i) $i_1 \prec i_2$ if $a_{i_2,i_1} = \mu_{i_2} + \nu_{i_1}$;
ii) $i_1 \prec^{\mathrm{t}} i_2$ if $a_{i_1,i_2} = \mu_{i_1} + \nu_{i_2}$.

i) For any integer i_0, the rules:

$$\mu'_i := \mu_i + e \quad \textit{and} \quad \nu'_i := \nu_i - e \quad \textit{if} \quad i_0 \prec i$$

and

$$\mu'_i := \mu_i \quad \textit{and} \quad \nu'_i := \nu_i \quad \textit{if not},$$

where

$$e \leq \min_{\substack{i_0 \prec i \\ i_0 \not\prec i'}} (\mu_{i'} + \nu_i - a_{i',i}), \tag{3}$$

define a minimal cover for A.

i′) For any integer i_0, the rules

$$\mu''_i := \mu_i - e \quad \textit{and} \quad \nu''_i := \nu_i + e \quad \textit{if} \quad i_0 \prec^{\mathrm{t}} i$$

and

$$\mu_i'' := \mu_i \quad \textit{and} \quad \nu_i'' := \nu_i \quad \textit{if not},$$

where

$$e \leq \min_{\substack{i_0 \prec^{\mathrm{t}} i \\ i_0 \not\prec^{\mathrm{t}} i'}} (\mu_i + \nu_{i'} - a_{i,i'}),$$

define a minimal cover for A.

ii) There exists no minimal cover μ', ν' for A such that $\mu'_{i_0} > \mu_{i_0}$, $\mu'_i < \mu_i + \mu'_{i_0} - \mu_{i_0}$ and $i_0 \prec i$.

ii′) There exists no minimal cover μ', ν' for A such that $\mu'_{i_0} < \mu_{i_0}$, $\mu'_i > \mu_i - \mu'_{i_0} + \mu_{i_0}$ and $i_0 \prec^{\mathrm{t}} i$.

Proof. — i) As the transversal sum is unchanged, the minimality is granted. We only have to prove that we obtain a cover. If $i_0 \prec i$ and and $i_0 \prec j$ or $i_0 \not\prec i$ and $i_0 \not\prec j$, then $a_{i,j} \leq \mu'_i + \nu'_i = \mu_i + \nu_j$. If $i_0 \prec i$ and $i_0 \not\prec j$, $a_{i,j} \leq \mu_i + \nu_j \leq \mu'_i + \nu'_i = \mu_i + \nu_i + e$. If $i_0 \not\prec i$ and $i_0 \prec j$, $a_{i,j} \leq \mu'_i + \nu'_j = \mu_i + \nu_j - e$ by (3).

The proof of i′) is similar.

ii) We may chose i with a shortest path from i_0 to i. Let i_r be the penultimate row of the path. Then $a_{i,i_r} = \mu_i + \nu_{i_r}$ by the path definition, $\mu'_{i_r} \geq \mu_{i_r} + \mu'_{i_0} - \mu_{i_0}$ by the path minimality hypothesis and $a_{i,i_r} \leq \mu'_i + \nu'_{i_r}$ by the cover definition. As the cover is minimal, we need to have for all k $\mu'_k + \nu'_k = \mu_k + \nu_k$. So

$$\mu'_i \geq a_{i,i_r} - \nu'_{i_r} = \mu_i + \nu_{i_r} - \nu'_{i_r} = \mu_i + \mu'_{i_r} - \nu_{i_r}.$$

A contradiction. The proof of ii′) is exactly similar. ∎

The following examples are easy illustrations of the use of this theorem.

Examples. — 27) A matrix A with $a_{i,j} = \mu_i + \nu_j$, admits a single class of minimal covers: that of μ, ν.

28) A triangular matrix of ones

$$\begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & & & \\ 0 & 0 & \ddots & & \vdots \\ \vdots & & \ddots & 1 & 1 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

has exactly $n+1$ minimal covers of non negative integers indexed by $0 \leq k \leq n$, defined by

$$\begin{array}{llllll} \alpha_i = 1 & \text{if} & i \leq k & \text{and} & \alpha_i = 0 & \text{if} \quad i > k \\ \beta_j = 0 & \text{if} & j \leq k & \text{and} & \beta_j = 1 & \text{if} \quad j > k. \end{array}$$

They belong to n classes, as the covers obtained for $k = 0$ and $k = n$ are obviously equivalent and correspond to the minimal canon.

Remark 29. — We could have defined covers for rectangular matrices, but if $s < n$, we need to impose some lower bound for the ν_j in order to ensure the existence of a minimal cover: if not the covers $\mu_i + c$ and $\nu_j - c$ will be equivalent, but

$$\sum_{i=1}^{s}(\mu_i + c) + \sum_{j=1}(\nu_j - c) = \left(\sum_{i=1}^{s}\mu_i + \sum_{j=1}^{n}\nu_j\right) - (n-s)c!$$

If we impose $\nu_j \geq 0$, for any matrix A of non negative elements (*viz.* $a_{i,j} \geq 0$), the minimum cover μ, ν is such that

$$\sum_{i=1}^{s}\mu_i + \sum_{j=1}^{n}\nu_j = \mathcal{O}.$$

so that Jacobi's bound (see def. 1) could also be handled using covers in a direct way, without adding $n - s$ lines of 0. But this trick is more general, as it works without any restriction on the entries in A.

Minimal covers in this setting are characterized, as one may easily check, by the fact that there exists an injection $\sigma \in S_{s,n}$ such that $\mu_i + \nu_{\sigma(i)} = a_{i,\sigma(i)}$ and $\nu_j = 0$ for all j that do not belong the the image of σ.

3 Related algorithms and deeper complexity analysis

Discussing the complexity of Jacobi's algorithm is an interesting subject, but we need to keep in mind that it is anachronical to do it in the setting of modern computation models, when Jacobi's concern was to spare the work of useless lines rewriting, in a time when paper and pen remained the main computation tools. We will now provide some improvements that lead to a better complexity, in our contemporary formalism.

3.1 Finding a maximal set of transversal maxima. The Bipartite matching problem

We have encountered with Jacobi's algorithm the following special problem of finding a maximal set of transversal maxima. This amounts to solving the assignment problem with a matrix of zeros and ones, using jacobi's characterization (see lemma 5). In what follows, as all maximal values will be 1, we will speak of *transversal ones*, *starred ones* instead of transversal or starred maxima. This is known as the maximal bipartite matching, or mariage problem.

We could change the data structure and use the graph of the relation $a_{i,j} = 1$, smaller that the full matrix, as initial data, but for the sake of clarity, we will stay here in the dense setting.

3.1.1 Hopcroft and Karp's algorithm in Jacobi's setting.

First, we may repeatedly look for lower right ones. This may be done in sequence, until no one is found with a total cost $O(sn)$. On may notice that Kőnig's theorem 11 implies that this first step already produces $\lceil \mathcal{O}/2 \rceil$ transversal ones.

The elementary relation "there is a path from row i to row j" can be constructed with cost $O(sn)$ and its graph has size at most s^2. The main idea is to build a maximal set (in the sense that it is not strictly included in another such set) of disjoints paths of minimal length leading to a lower first class row, before building a new path relation. So the main step of the algorithm is not to produce a single augmenting path, but, at each stage k, a maximal set of disjoint paths of the same length β_k.

ALGORITHM 30. *Length.* — **Input data**: a matrix A and a transversal set of "ones", given by an injection $\sigma : [1, r] \mapsto [1, s] \times [1, n]$, and that we assume here for convenience to be $a_{i,i}$, $1 \leq i \leq r$.
Outputs: the list of sets of rows L_i, $0 \leq i \leq k$, that may be reached from a first class row with a right "one" with paths of length at least i, and the minimal length k of a path from L_0 to a lower row, or *"failed"* if no such row exists.

We start with the upper lines with right "ones", that form the set L_0. Let $M := L_0$[6].

At step 1, we define L_1 to be the set of elements not in M such that there is a path to them from some element of L_0. We increase M with L_1. We then define L_2 to be the set of elements not in M to which there is a path from an element of L_1, etc.

We stop this process as soon as L_k is empty—and we return then *"failed"*—or contains a lower line, that will be by construction a first class lower line. The integer k will be the minimal length of a path leading to a lower first class row, that we return.

This process is achieved in $O(s^2)$ operations.

To find a maximal set of disjoint paths we may use for brevity the following recursive process (see [32] for a different more detailled presentation). The maximal number of disjoint paths is bounded by the cardinal of L_0 and these elements are the possible starting points of any of them. We define first a set F

[6] We assume it is possible to know with constant time if some integer belongs to any such subsets of $[1, s]$, by storing boolean values in some array.

of available rows, which is initialized with the set of rows not in L_0. We define first the following function, assuming that the length k has been computed as in algo. 30.

ALGORITHM 31. *Path.* **Input data**: an integer j and a row i in L_j.
Output: some list of rows $i_\alpha \notin F$ that forms a path of length $k-j$ leading from row i to a lower first class row, or *"failed"* if no such path exists.
Global variable: F, a set of elements to be used for building paths during the process.

Step 1) Let C be the set of elements of $L_{j+1} \cap F$ such that there is a path from i to them.
If $j = k-1$ and $C \neq \emptyset$, let $c \in C$, remove c from F and return c.
If $j < k-1$, remove from F the elements of C and go to step 2).

Step 2) For $c \in C$, if $\mathrm{Path}(c) = L$ is a path, then we put back in F the elements of C to which "Path" has not be applied and we return the list $[c, L]$.

This process clearly returns a path of length k to a lower first class row. The elements in that path are removed from F, as well as the elements of L_i from which no such path of length $k-i$ has been found. So repeated call to that function will produce disjoint paths and the function Path can be applied only once to a given row.

This implies that we can apply in sequence Path *to the elements of* L_0 *to get a maximal set of disjoint paths of minimal length k in* $O(s^2)$ *operations.*

ALGORITHM 32. *Increase.* — **Input data**: a list T of transversal maxima, and an "increasing" path $[j_1, \ldots, j_k]$ from a row with a right maxima to a lower first class row.
Output: an increased list of transversal maxima. We proceed as in 2.1 step 2) d) p. 12.

ALGORITHM 33. *Hopcroft–Karp.* — **Input data**: a matrix A of zeros and ones.
Output: the elements of a maximal transversal sum of A and a minimal cover.

Step 1) As stated above, we repeatedly look for lower right ones, producing first at least $\lceil \mathcal{O}/2 \rceil$ transversal ones $a_{\sigma(i)}$, $1 \leq i \leq r$.

Step 2) Let $k :=$*Length*(A, σ, τ). If$k =$*"failed"*, then return σ.
If not, for $i \in L_0$ do:
— $\mathcal{J} :=$ *Path*(k, i).
— If $\mathcal{J} \neq$*"failed"* then $\sigma :=$*Increase*$(\sigma, \mathcal{J})$.
— Repeat *Step 2)*.

The total cost of steps 1) or 2) is $O(sn)$*, so the key point in bounding the complexity is to evaluate how many times step 2) is performed, which is the goal of the next two lemmata 34 and 39.*

Lemma 34. — *Let β_k be the length of the paths used at stage k, then the sequence β_k is strictly increasing.*

Proof. — Assume it is not the case and $\beta_k \leq \beta_{k-1}$. We may assume that k is minimal with that property. We call a changed line, a line that has been used in some path at stage $k-1$. In any path used at stage k, there must be a changed line. If not, whever $\beta_k = \beta_{k-1}$ and this contradicts the fact that algorithm 31 produces a maximal set of disjoint paths of length β_{k-1}, whever $\beta_k < \beta_{k-1}$ and this contradicts the minimality of the length of paths produced by algorithm 30.

An injective function $\varphi : [1, s] \mapsto [1, s]$ defines a unique set of disjoint paths and loops, the union of which is equal to the union of its image and its definition domain. If $\exists r\ \varphi^r(i) = i$, then i belongs to a loop, if not let $r_0 := \max\{r | \varphi^{-r}(i)$ is defined$\}$ and $r_1 := \max\{r | \varphi^r(i)$is defined$\}$, then i belongs to the path $\varphi^{-r_0}(i)$, ..., $\varphi^{r_1}(i)$.

Let φ be the function defined by the τ paths of stage $k-1$ and i_0, ..., i_{β_k} be a path of stage k: it must have some rows in common with the paths of stage $k-1$. Let them be i_{h_1}, ..., i_{h_r}, $r \geq 1$. If $\varphi^{-1}(i_{h_\kappa})$ is defined, we replace in the graph of φ $(\varphi^{-1}(i_{h_\kappa}), i_{h_\kappa})$ with $(\varphi^{-1}(i_{h_\kappa}), i_{h_\kappa+1})$. We then add to the graph of φ the couples $(i_\zeta, i_{\zeta+1})$, $0 \leq \zeta < r$, $\zeta \notin \Im(h)$ and the couples $(i_{h_\kappa}, \varphi(i_{h_\kappa}))$, $1 \leq \kappa \leq r$, $\varphi(i_{h_\kappa})$. This construction is illustrated by the following figure.

Elementary path relations at stage $k-1$ are indicated by $\vdots$ and at stage k by $|$; after rearrangement by $[$.

The starred ones of stage $k-1$ by *1* and those of stage k by **1**.

$$\left(\begin{array}{ccccccl}
\mathit{1} & \cdots & & \cdots & & \cdots & \textit{row } \varphi^{-1}(i_{h_k}) \\
\vdots & & & & & & \\
\vdots & & & & \mathbf{1} & \cdots & \textit{row } i_{h_k-1} \\
\vdots & & & & | & & \\
\mathbf{1} & \cdots & \mathit{1} & \cdots & 1 & \cdots & \textit{row } i_{h_k} \\
| & & \vdots & & & & \\
| & & 1 & \cdots & & \cdots & \textit{row } i_{h_k+1} \\
| & & & & & & \\
1 & \cdots & & \cdots & & \cdots & \textit{row } \varphi(i_{h_k})
\end{array}\right)$$

This defines an injection to which is associated a new set of paths and (possibly) loops.

Then, the sum of their lengths is at most $\tau\beta_{k-1} + \beta_k - r$ (and strictly smaller iff loops do exists). So, as $\beta_k \leq \beta_{k-1}$ one path must be of length strictly smaller that β_{k-1}. This contradicts the minimality of β_{k-1}. ∎

Remark 35. — Our paths of length r correspond to paths of length $2r+1$ following the conventions of Hopcroft and Karp [32]. This is due to the fact that they define the path relations, not between rows but between the "ones" involved in the path relation. In their setting, starred elements appear with a minus sign and the others with a plus sign. So the process of reconstruction reduces to computing the sum of the two paths. E.g. denoting by $1_{i,j}$ a one placed in row i and column j we have:

$$\begin{aligned} & (+1_{1,1} - 1^*_{1,3} + 1_{2,3} - 1^*_{2,4} + 1_{4,4}) + (+1_{2,2} - 1^*_{2,3} + 1_{3,3}) \\ = \ & (+1_{1,1} - 1^*_{1,3} + 1_{3,3}) + (+1_{2,2} - 1^*_{2,4} + 1_{4,4}). \end{aligned}$$

The element $1_{2,3}$ that appeared two times has vanished and the two paths of lengths 5 and 3 are replaced by two paths of lengths 3 and 3.

Examples. — 36) In the following example, the stared ones of the first stage ($k-1$ in lemma lemma-decreasing) are *italicized* and those of the second stage (k in the lemma) **bold**. The first—increasing—path includes rows 1, 2 and 3. The second rows 3, 2, 1 and 4. Using Hopcroft and Karp's convention we have $(1_{1,3} - 1_{1,2} + 1_{2,2} - 1_{2,1} + 1_{3,1}) + (1_{3,4} - 1_{3,1} + 1_{2,1} - 1_{2,2} + 1_{1,2} - 1_{1,3} + 1_{4,3}) = 1_{3,4} + 1_{4,3}$. This means that we have, in Jacobi's setting, two paths of length zero, *viz.* lower right ones.

$$\begin{pmatrix} 0 & \mathit{1} & \mathbf{1} & 0 \\ \mathit{1} & \mathbf{1} & 0 & 0 \\ \mathbf{1} & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

37) Using the same conventions as above, the first path includes rows 1, 2 and 3, the second rows 3 and 4. In Hopcroft and Karp's convention: $(1_{1,3} - 1_{1,2} + 1_{2,2} - 1_{2,1} + 1_{3,1}) + (1_{3,4} - 1_{3,1} + 1_{4,1}) = (1_{1,3} - 1_{1,2} + 1_{2,2} - 1_{2,1} + 1_{4,1}) + 1_{3,4}$. In Jacobi's setting: one path of length 2, formed of rows 1, 2 and 4, and one of length 0, *viz.* a lower right "one": $1_{3,4}$.

$$\begin{pmatrix} 0 & \mathit{1} & \mathbf{1} & 0 \\ \mathit{1} & \mathbf{1} & 0 & 0 \\ \mathbf{1} & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

38) The first path includes rows 1, 2 and 3, the second rows 3, 1 and 4. In H & K's convention: $(1_{1,3} - 1_{1,2} + 1_{2,2} - 1_{2,1} + 1_{3,1}) + (1_{3,4} - 1_{3,1} + 1_{1,1} - 1_{1,3} + 1_{4,3}) = (1_{1,1} - 1_{1,2} + 1_{2,2} - 1_{2,1}) + 1_{3,4} + 1_{4,3}$. Jacobi's: one loop, formed of rows 1 and 2, and two lower right "ones": $1_{3,4}$ and $1_{4,3}$.

$$\begin{pmatrix} 1 & \mathit{1} & \mathbf{1} & 0 \\ \mathit{1} & \mathbf{1} & 0 & 0 \\ \mathbf{1} & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Lemma 39. — *Let $(a_{i,j})_{(i,j)\in G_1}$ and $(a_{i,j})_{(i,j)\in G_2}$, where G_1, G_2 are the graphs of two functions $[1,s] \mapsto [1,n]$, be two families of $r_1 := \sharp G_1$ and $r_2 := \sharp G_2$ transversal ones of A. We assume $r_2 > r_1$. Lower right "ones" in the family $(a_{i,j})_{(i,j)\in G_2}$, i.e. elements that are not placed in the same rows or columns as the elements of G_1, will be considered to be paths of length* 0.

Then, using only the starred "ones" in $(a_{i,j})_{(i,j)\in G_1}$ and the "ones" in $(a_{i,j})_{(i,j)\in G_2}$ placed in the same columns, we define a path relation such that there exists a path of length at most $\lfloor r_1/(r_2-r_1) \rfloor$.

Proof. — If lower right "ones" exist in $(a_{i,j})_{(i,j)\in G_2}$, then the result stands according to our convention. We obtain possibly loops (if G_1 and G_2 have a element (i,j) in common, then we consider it as a loop from row i to itself) and at least $r_2 - r_1$ open paths, as there are as many elements from G_1 and G_2 in loops.

As a path of length m involves m starred ones in G_1, the sum of the lengths of all paths is at most r_1 and there exists a path of length at most $\lfloor r_1/(r_2-r_1)\rfloor$. ■

THEOREM 40. — *Assume that the maximal number of transversal "ones" in A is s_0, then the algorithm requests at most $\sqrt{s_0}$ steps. Its complexity is $O(s_0^{1/2} sn)$.*

PROOF. — Let $k := \lfloor s_0 - \sqrt{s_0}\rfloor$ and ℓ be the number of steps. Let G_1 be set of starred ones at step $m \le \ell - k$ and G_2 the maximal set of starred ones at step ℓ. Using lemma 39, the length of a path at step m is a most $s_0/\sqrt{s_0}$, so that there are at most $\lfloor\sqrt{s_0}\rfloor$ steps before step $\ell - \lfloor s_0 - \sqrt{s_0}\rfloor$ and $\ell \le \lfloor\sqrt{s_0}\rfloor$. ■

This problem was first considered by Frobenius [25] in order to decide *a priori* if a matrix where non zero element can appear at known places has an identically vanishing determinant. It is a desappointing that the complexity of solving this problem is for the moment bigger than that of computing a numerical determinant. We can only achieve the exponent of matrix multiplication with probabilistic algorithms using random numerical values! See Ibarra and Moran [35]. One may notice that this method, when it succeeds, only gives a cover, but no maximal matching and its seems uneasy to compute it faster that the Hopcroft and karp method, even if a cover is *a priori* known.

3.2 A $O(s^2 n)$ version of Jacobi's algorithm

In order to improve the complexity of Jacobi's algorithm, we only have to remark that it is useless to reconstruct the whole path relation in order to reduce the number of second class rows or make some lower first class row appear, as the starred maxima will remain unchanged.

At step **2 c)** p. 12, we also have to define the set $C_{\mathrm{I,II}}$ of first and second class rows. For each $i \in C_{\mathrm{I,II}}$, we compute the minimal distance d_i between its starred maximum $a_{i,i}$ and some third class row element in the same column, or between some upper right maxima of a first class row i and some third class row in the same column.

All this is done with a cost at most $O(sn)$.

c′) We will then increase all third class row by $d_{i_0} := \min_{i\in C_{\mathrm{I,II}}} d_i$. If this creates a lower first row, step c) is finished. If not, we remove i_0 from the set $C_{\mathrm{I,II}}$, and add it to the third class. We redefine d_i to be the minimum of d_i and the distance between its starred maximum $a_{i,i}$ and $a_{i_0,i}$, or between some upper right maxima $a_{i,j}$ and $a_{i_0,j}$. We iterate step c′) with these new values.

The substep c′) is performed with a total cost $O(n)$ and will be interated at most $s-2$ times, until 2) is completed and we don't need step e) any more.

THEOREM 41. — *Using substep 2. c′), the complexity of Jacobi's algorithm is bounded by $O(s^2 n)$.*

This improved complexity $O(n^3)$ was first obtained for square matrices by Dinic and Kronrod [20] in 1969, rediscovered independently by Tomizawa [77] in 1971 and then by Edmonds and Karp [22] in 1972.

Remark 42. — As we have already seen (see rem. 9), we cannot escape, in some cases, to repeat at least $s-2$ times step 2 of Jacobi's algorithm. It could be possible to speed up the construction of elementary step relation, as in most cases they are unchanged or reversed. But it seems unavoidable to escape a $O(s^2)$ complexity when building the class partition. In this situation, we don't know how to construct in a single step a large set of augmenting paths, as we have been able to do for the maximal matching problem (see 2.2).

One may notice that pionnering aspects of Jacobi's work include reachability issues and computing the transitive closure of a directed graph. But this problem is not formalized and its solution is implicitely assumed to be achieved in some naïve way for small size data. However, some of his algorithms solve problems equivalent to some instances of the shortest path problem.

3.3 A canon being given, to find the minimal one

In order to solve this problem, Jacobi proposes ([36] VII) first to compute a maximal set of transversal maxima, which may be done using the method developped in 3.1 with complexity $O(n^{5/2})$ for a square matrix A. Knowing transversal maxima, we may use then the following method.

ALGORITHM 43. *Data:* a square matrix A of size n and a maximal system of transversal maxima for a canon of A, that we assume for simplicity to be $a_{i,i}+\ell_i$.

Step 1. We decrease all the ℓ_i by $\min_{i=1}^n \ell_i$, so that, at least one ℓ_i is 0.

Step 2. We build the path relation. Then, we establish the list L_1 of rows with $\ell_i=0$, or to which there is a path from a row with $\ell_i=0$, and the list L_2 of the remaining rows.

Lemma 44. — *If there is a path from all rows to a row i with $\ell_i=0$, then the canon is minimal.*

PROOF. — Assume it is possible to decrease some ℓ_{i_0} to a new value ℓ'_i. We may choose ℓ_{i_0} so that there is a path of minimal length from row i_0 to a row i with $\ell_i=0$. It means that there is a path from row i_0 to some row i_1 with ℓ_{i_1} unchanged, meaning that $a_{i_0,i_0}+\ell'_{i_0} < a_{i_1,i_0}+\ell_{i_1}$, a contradiction. ∎

By lemma 44, if $L_2=\emptyset$, we have finished. If not, we compute

$$b := \min(\min_{i\in L_2} \ell_i, \min_{i\in L_1} \min_{i'\in L_2} a_{i,i} - a_{i',i}),$$

that is the minimal distance between some starred element in L_1 and the elements in the same column in some row of L_2. Step 2) may be achieved with complexity $O(n^2)$.

Step 3. We decrease all the ℓ_i, $i \in L_2$ by b. In this way, some rows will go from L_1 to L_2. We then repeat step 2).

A naïve complexity analysis gives a $O(n^3)$ complexity for the whole process. It is possible to turn it to $O(n^2 \ln(n))$, using "balanced trees" or "AVL trees, from the name of their inventors Adelson-Velsky and Landis [1]. See also Knuth [50, 6.2.3 p. 451]. This tree structure allows to maintain dynamically an ordered list of p elements, allowing to insert, delete, search the order of an element or an element of a given order in $O(\ln p)$ operations.

ALGORITHM 45. We use the same input and data and convention as above (alg. 43).

Step 1) a) Decrease all the ℓ_i by $\min_{i=1}^n \ell_i$. Create a list L_1 of the rows i with $\ell_i = 0$ and a list L_2 of the remaining elements.

b) For $i \in L_1$, create a balanced tree T_i, containing for all the the rows i' of L_2 the pairs $(\min(\ell_i, a_{i',i} - a_{i,i}), i')$ of L_2, sorted by reversed lexicographical order. This may be achieved with total cost $O(n^2 \ln n)$.

Step 2) Compute the smallest pair $(\min(\ell_{i_0}, a_{i'_0,i_0} - a_{i_0,i_0}), i'_0)$ in the trees T. Decrease the ℓ_i of the value $\min(\ell_{i_0}, a_{i'_0,i_0} - a_{i_0,i_0})$. For all $i \in L_1$, suppress the pair $(\min(\ell'_{i_0}, a_{i'_0,i} - a_{i,i}), i'_0$ from T_i. Add row i'_0 to L_1 and suppress it from L_2. Create a balance tree $T_{i'_0}$ as above.

All this may be done with total cost $O(n \ln n)$.

If L_2 is empty, we have finished, if not we iterate step 2), which will be performed at most n times, providing a total cost $O(n^2 \ln n)$.

THEOREM 46. — *Knowing a canon for a square matrix of size $n \times n$ and a set of transversal maximal elements in this canon, one may compute the minimal canon with cost $O(n^2 \ln n)$.*

PROPOSITION 47. — *Let $B_{i,j}$ be a square matrix at least a transversal set of maxima, that we may assume to be: $b_{i,i}$. Then, the reflexive transitive closure of the path relation does not depend on the choice of this transversal set.*

PROOF. — Assume that $b_{i,i}$ and $b_{i,\sigma(i)}$ be two transversal sets of maxima. We denote by $\prec_1$ (resp. $\prec_2$) the path relation defined using the first family (resp. the second). Assume that there is a elementary path $i \prec_1 j$. Consider the cycle $i_0 = i$ and $i_{p+1} = \sigma(i_p)$. Let $i_r = i$. According to the path definition, there is an elementary path $i_p \prec_1 i_{p+1}$ and $i_{p+1} \prec_2 i_p$. Using the cycle, we have

$$i \prec_2 i_{r-1} \prec_2 \cdots \prec_2 i_1 = \sigma(i) \prec_2 j,$$

so that the reflexive transitive closure of $\prec_1$ and $\prec_2$ are the same. ∎

DEFINITION 48. — *We will denote by π_A the path relation associated with the minimal canon $A + \lambda$ of A.*

Remarks. — 49) Generically, *i.e.* if quantities $a_{i,j} - a_{i,i}$ are all different, π_A defines a forest of rooted trees with n labeled vertices, where the roots correspond to rows with $\lambda_i = 0$. By a variant of Cayley's formula there are $(n+1)^{n-1}$ possibilies, and as much formulas for the values of the minimal canon ℓ_i. If there is an elementary path from row i_0 to row i_1, then to row i_2 ...up to row i_r with $\lambda_{i_r} = 0$, then

$$\lambda_{i_0} = \sum_{k=1}^{r} a_{i_k, i_{k-1}} - a_{i_{k-1}, i_{k-1}}.$$

If a given row i_0 may be connected to two different "roots" (*i.e.* rows i with $\lambda_i = 0$, it is enough to consider one to fix the value of λ_{i_0}. This may be visualized using some mechanical construction (see below subsection 3.8).

50) In the case of a rectangular matrix, we cannot use this method. E.g. considering the matrix

$$\begin{pmatrix} \mathbf{2} & \mathit{2} & 1 \\ 1 & \mathbf{1} & \mathit{2} \end{pmatrix},$$

with a canon $\ell = (0, 1)^{\mathrm{t}}$ and the transversal elements in bold, we cannot decrease the canon without changing these transversal element. However, the minimal canon is 0 with the transversal elements in italics.

51) Assume we have computed Jacobi's bound $\mathcal{O}$ for some $s \times n$ matrix by adding $n - s$ lines of 0 and that $\mathcal{O} = \sum_{i=1}^{s} a_{i,i}$. Then, for any canon ℓ, the value of ℓ_{i_0} for $s < i_0 \leq n$ is $\max_{j=s+1}^{n} \max_{i=1}^{s} a_{i,j} + \ell_i$. This means that,if one needs to compute the minimal canon for such a matrix, one just has to compute the minimal canon of the square matrix $(a_{i,j})_{1 \leq i,j \leq s}$ in $O(s \ln s)$ operation, and then the common value of the ℓ_i, $s < i \leq n$ in $O(s(n-s))$ operations.

Assume now that we want to compute the minimal canon for a matrix A with $n - s$ columns of 0. We may assume that the transversal maximal elements are $a_{i,i}$, $1 \leq i \leq s$. Then, it is easily seen that for $i' > s$, the minimal canon λ is such that $\lambda_{i'} = \max_{i=1}^{s} \lambda_i$.

Before leaving this subject, we will emphasize the special case of matrices of zeros and ones, associated to maximum matching problems. For this, the algorithm 43 will run in $\mathcal{O}(n^2)$ steps, as we only need to apply step 2) one time. It is so efficient enough and will allow us to maximize the number of rows (or columns) in Kőnig's theorem 11.

PROPOSITION 52. — *Let A be some $s \times n$ matrix of zeros and ones (possibly horizontally or vertically rectangular), A' the $\max(s, n) \times \max(s, n)$ matrix obtained* *by adding $|n - s|$ columns or rows of* 0 *to* A*. Let* λ *be the minimal canon of* A'*,* μ *and* ν *the associated canonical cover and* $\mathcal{O}$ *the maximal transversal sum of* A'*.*

i) If some λ_i *is not* 0 *(or equivalently some* $\mu_i = 1 - \lambda_i$ *is not* 1*), then for any sets* R *of rows and* C *of columns containing all the* 1 *in* A*,* $i_0 \in R$ *implies* $\lambda_{i_0} = 0$*.*

ii) In Kőnig's theorem, there exists a unique couple of sets of rows R *and columns* C *with* R *maximal for inclusion (resp. with* C *maximal for inclusion).*

Proof. — The assertion i) is a straightforward consequence of the minimality of λ.

ii) The result is straightforward if the Jacobi number $\mathcal{O}$ of A is s. Then $R = \{1, \ldots, s\}$ is the maximal set of rows.

If some $\lambda_i = 1$, then the result is a direct consequence of i).

If all the λ_i are 0 and $\mathcal{O} < s$, one just has to consider the $(\max(s, n) + 1) \times (\max(s, n)+1)$ matrix A'' obtained by adding to A' a row of $\max(s, n)+1$ ones and a column of $\max(s, n)$ zeros. We get a maximal transversal sum of value $\mathcal{O}$ for A'' by completing one for A' with the 1 in column and line of index $\max(s, n)+1$. As there must be some zero in any maximal transversal sum, some λ_i must be 1 in the minimal canon of A'', so that we can now apply i).

The statement for columns is obtained by considering the transpose matrix A^{t}. ∎

Definition 53. — *We call this cover the* row maximal *(resp.* column maximal*) minimal cover.*

Algorithm 54. A $s \times n$ matrix A of zeros and ones being given, together with the elements of a maximal transversal sum of elements of A the following algorithm computes a row maximal minimal cover.

Step 1. Compute the Jacobi number $\mathcal{O}$ of A. If $\mathcal{O} = s$, then the s rows of A form the row maximal cover.

Step 2. Make a square matrix A' by adding to A $|n - s|$ rows or columns and add a row of $\max(s, n) + 1$ ones and a column of $\max(s, n)$ zeros to define a $(\max(s, n) + 1) \times (\max(s, n) + 1)$ matrix A'' as in the proof of prop. 52. Then compute the minimal cover of A'' using algorithm 43; the rows $1 \leq i \leq s$ with $\ell_i = 0$ form the row maximal minimal cover of A.

3.4 Transversal maximas being given, to find the minimal canon

If we don't have a canon but just know the place of transversal maxima in the matrix A, then we can proceed in the following way.

Algorithm 55. We assume that the transversal family is $a_{i,i}$. For $1 \leq i \leq n$, increase row i by $\max_{k=1}^{n} a_{k,i} - a_{i,i}$. This may be done in $O(n^2)$ operations.

Repeat the process untill all rows remain unchanged.

PROPOSITION 56. — *This algorithm produces the minimal canon in $O(n^3)$ operations.*

PROOF. — The process in the algorithm will be repeated at most n times, the exact number being, in the generic case the maximal distance from any row to a row with $\lambda_i = 0$, according to the path relation of def. 48. ∎

Remarks. — 57) This algorithm may be easily modified to compute the path relation forest. Given any transversal family $a_{i,\sigma(i)}$, it may be used to test if it corresponds to transversal maxima, the stoping of the algorithm after n step beeing a necessary and sufficient condition.

If the algorithm does not stop, it means that the path relation contains a loop τ (which may be tested before step n), so that $\sum a_{i,\sigma(i)} < \sum a_{i,\tau\circ\sigma(i)}$.

58) The last example of [36, § 3] is the transpose of a canon. Then, this transpose is not a canon, but the terms of a maximal transversal sum are known and we can apply the above method. We may also compute a cover and deduce of it a canon (see rem. 22), allowing to use the more efficient method of 3.3. In Jacobi's informal setting, the two methods have comparable complexities.

3.5 Tropical geometry

We will denote by $M \odot N$ the tropical matrix multiplication. One may wonder why the analogy with the determinant cannot be used in a straightforward way. One may remark first that the analogy suffers important limitations: the analog of addition is "$\max$" that has no inverse and the tropical determinant of a tropical product of matrices is not in general the sum of their tropical determinants. Such a property stands only in special situations, e.g. $|A \odot B|_T = |A|_T \odot |B|_T$ if B is a canon and A the transpose of a canon. Moreover, the tropical determinant is also the tropical permanent...

Assume that $a_{i,i}$ is a transversal family with a maximal sum. Then, reducing row i by $a_{i,i}$ we get a new matrix B with $b_{i,i} = 0$, the result of the last algorithm 55 is the tropical matrix product: $(a_{1,1}, \ldots, a_{n,n}) \odot B^n$. A $O(n^\alpha)$ algorithm for the tropical multiplication would produce a $O(n^\alpha \ln(n))$ algorithm for the problem of finding a minimal canon, knowing the elements of a maximal transversal sum.

3.6 Minimal canons subject to inequalities

PROPOSITION 59. — *Let A be a square $n \times n$ matrix, and c_i, $1 \leq i \leq n$ positive integers. Then there exists a unique minimal canon subject to the condition $\ell_i \geq c_i$.*

Proof. — Consider the new matrix $A' := A + c$. Then, ℓ is a canon of A, subject to $\ell_i \geq c_i$ iff $\ell - c$ is a canon of A', so that the unique minimal canon λ of A' is such that $\lambda + c$ is the unique minimal canon of A, subject to $\ell_i \geq c_i$. ∎

This proof obviously provides an easy algorithm to compute such minimal canons, that we will be used in sections 9 and 10 to bound the order of derivation of initial equations or normal foms necessary to perform change of orderings or resolvents.

3.7 Minimal canons and shortest paths

Let $A+\ell$ be a canon for A; assume that $a_{i,i}$ form a maximal transversal sum. Then, we define a weighted directed graph G on the set $\{0, 1, \ldots, n\}$, by associating the weight $w_{j,i} := a_{i,i} + \ell_i - a_{j,i} - \ell_j \geq 0$ to the ordered pair (j, i), and $w_{0,i} := \ell_i$ to the ordered pair $(0, i)$.

Reciprocally, we may associate to any such directed graph with positive weight a square matrix A and a canon $A + \ell$, defined by $a_{i,i} = C$, for $C \geq 2 \max_{i,j} w_{i,j}$, $\ell_i = w_{0,i}$ and $a_{i,j} := C - w_{j,i} + \ell_i - \ell_j$.

Proposition 60. — *The vector of integers λ is the minimal canon of A iff there exists in G a shortest path of length $\ell_i - \lambda_i$ from vertex 0 to vertex i.*

Proof. — It is enough to remark that there exists such a shortest path from vertex 0 to vertex i in G iff there exists a path, in the meaning of lemma 44, from row i to a row i_0 with $\lambda_{i_0} = 0$. So, according to this lemma, $A + \lambda_i$ is the minimal canon of A. ∎

In the same way, let A be a $n \times n$ square matrix. Define an oriented weighted graph on the set of vertices $\{0, 1, \ldots, n\}$ by setting $w_{0,i} := 0$ on edge $(0, i)$ and $w_{j,i} := a_{i,i} - a_{j,i}$ on edge (j, i). Reciprocally, define for any such weighted graph a matrix A with $a_{i,i} := C = max(0, \max_{i,j} w_{i,j})$ and $a_{i,j} := C - w_{j,i}$.

Proposition 61. — *i) The entries $a_{i,i}$ of A form a maximal transversal sum iff G admits no negative cycle.*

ii) Assuming the $a_{i,i}$ to form a maximal transversal sum, the vector λ is the minimal canon of A iff there is a shortest path of length $-\lambda_i$ from vertex 0 to vertex i in G.

Proof. — i) There is a cycle in G, with negative value γ iff there exists a permutation $\sigma : [1, n] \mapsto [1, n]$, so that $\sum_{i=1}^{n} a_{i,\sigma(i)} = -\gamma + \sum_{i=1}^{n} a_{i,i}$, so that the $a_{i,i}$ do no form a maximal sum.

ii) To see that λ is a canon, it is enough to remark that if $a_{i,i} + \lambda_i < a_{j,i} + \lambda_j$, then there is a path from 0 to i of length $-\lambda_j + w_{j,i} = -\lambda_j + a_{i,i} - a_{j,i} < -\lambda_i$. ∎

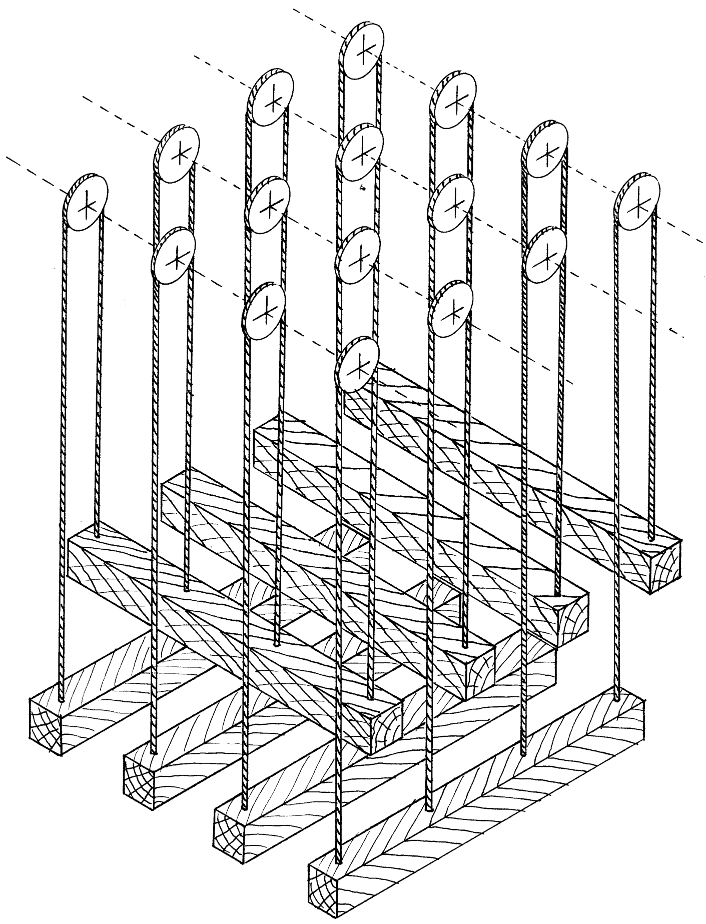
This means that the problems considered in sections 3.3 and 3.4 are equivalent to computing a shortest path, respectively for a directed graph with positive weights and a directed graph with arbitrary weights but no negative cycle. Although this contribution is non explicit, it seems that Jacobi deserves some mention of his pionnering contribution to graph theory. The complexity of Jacobi's original algorithm is $O(n^3)$, similar to that of Bellman [3] that computes minimal paths between *all* couples of vertices. See Schrijver's very interesting article for more details and references on the history of the shortest path problem [75].

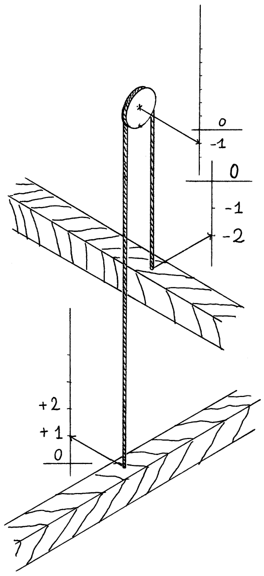
3.8 Physical analogies

It is always a greatest help for mathematical intuition to rest on physical models; one may easily design devices computing minimal covers.

3.8.1 Mechanical computation of a minimal cover

E.g., one may consider a mechanical system consisting of $2n$ horizontal rods, n standing for the rows and n standing for the columns, crossing at right angles.





At each crossing of two rods i and j, a cable passing to a pulley is attached to both of them, so that if the relative height μ_i and ν_j of the two rods, as well as the height $a_{i,j}$ of the pulley is defined to be 0 at rest, when the $a_{i,j}$ are increased to take new positive values, one has:

$$\frac{\mu_i + \nu_j}{2} \geq a_{i,j}, \text{ that becomes } \mu_i + \nu_j \geq a_{i,j},$$

by chosing a half scale for the pulley height. Under gravity, the total energy of the system, which for rods of equal masses is proportional to

$$\sum_{i=1}^{n} \mu_i + \nu_i,$$

will be minimal, so that this device will produce a minimal cover. Assuming that the weight of a rod is M, adding a little extra weight to those standing for the rows, say $0 < \varepsilon < M/n$, the equilibrium point will be unique and will correspond to minimal values for the μ_i, which corresponds to the minimal canon, provided that we impose $\mu_i \geq 0$, using some wedge.

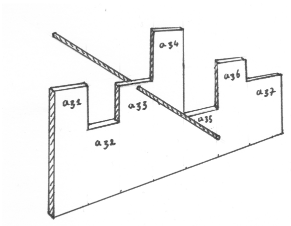
Using such a device, $-\infty$[7] entries can be modeled by suppressing the cable and pulley at some crossing. One can also allow rods to move down, so that negative values for the ν_j can be achieved too. If $\mathcal{O}$ is $-\infty$, then some rods will fall down... untill they are stopped by the finite length of the cables.

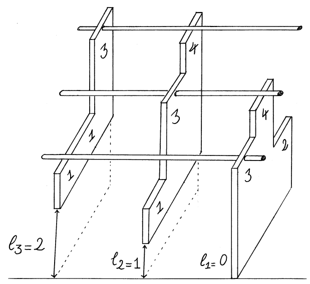
3.8.2 Materialization of the path relation

A second mechanical device may help visualize the graph of the path relation π_A (see def. 48) and rem. 49. Some vertical paterns reproduce the profile of each row of the matrix, e.g. below on the left row 3 of some 7×7 matrix. At the top of the part of each patern i correponding to $a_{i,i}$, an orthogonal rod is fixed. The paterns are assumed to be able to move vertically, so that if some $a_{j,i}$ is greater than $a_{i,i}$, the rod of patern i will rest on patern j. The lowest paterns rest on the floor corresponding to $\ell_i = 0$. The drawing below on the right corresponds to the minimal canon of

$$A = \begin{pmatrix} \mathbf{3} & 4 & 2 \\ 1 & \mathbf{3} & 4 \\ 1 & 1 & \mathbf{3} \end{pmatrix}, \quad \text{which is:} \quad A + \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \mathbf{3} & 4 & 2 \\ 2 & \mathbf{4} & 5 \\ 3 & 3 & \mathbf{5} \end{pmatrix}.$$

[7] See subsec. 4.1.

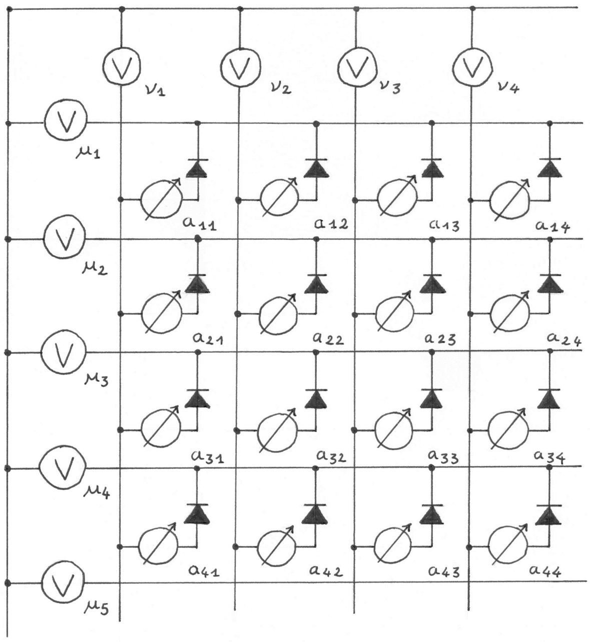




If we use some wedges to impose minimal values c_i for the height of some profiles, one will obtain the minimal canon subject to the condition $\ell_i \geq c_i$.

3.8.3 Electrical computation of a minimal cover

We finish with an electric circuit that may be used to compute a minimal cover. The voltmeters placed in the rows and columns of the circuit will mesure quantities corresponding to the covers μ_i and ν_j. Some adjustable voltage generators are connected at each crossing, providing a tension corresponding to $a_{i,j}$. The presence of a diode realizes the inequality: $\mu_i + \nu_j \geq a_{i,j}$. So, μ, ν is a cover. We cannnot with this device model entries $a_{i,j} < 0$; the absence of connection or a generator with a negative voltage are equivalent. If all the internal resistances of the voltmeters are equal to some value R, we need have $\sum_i \mu_i = \sum_j \nu_j$, as the intensity in and out of the circuit must be equal.



If we assume that the $a_{i,j}$ are 0, except for $a_{1,j} := 1$, $1 < j \leq n$ and $a_{i,1} := 1$,

$1 < i \leq n$, one finds the solution $\mu_1 = \nu_1 = (n-1)/n$ and that $\mu_i = \nu_j = 1/n$ for $1 < i, j \leq n$, which is not a minimal cover.

We need some extra assumption: let $|a| := \max_{i,j} a_{i,j}$ (assuming $|a| > 0$) it is then enough to replace the $a_{i,j}$ with $b_{i,j} := a_{i,j} + (n-1)|a|$ to have a working device.

Let indeed $b_{i,j} = \mu_i + \nu_i$ for r rows and s columns, with $r+s$ minimal. For simplicity, let these r rows be rows 1 to r and these s columns be columns 1 to s. Let I_1 be the electrical intensity from the first r rows to the last $n-s$ columns, I_2 from the last $n-r$ rows to the first s columns and $\mathcal{J}$ from the first r rows to the first s comlumns; by hypothesis, the intensity from the last $n-r$ rows to the last $n-s$ columns is 0. For $1 \leq i \leq r$, the minimal value of μ_i is at most $R(I_1 + \mathcal{J})/r$ and the minimal value of ν_j for $1 \leq j \leq s$ is $R(I_2 + \mathcal{J})/s$. In the same way, the minimal values μ_{i_0} of μ_i for $r < i \leq n$ and the minimal value ν_{i_0} of ν_j for $s < j \leq n$ are resp. at most $RI_1/(n-s)$ and $RI_2/(n-r)$.

As $\mu_{i_0} + \nu_{j_1} = b_{i_0,j_1}$ for $s < j \leq n$, we need have $R((I_1+\mathcal{J})/r+I_1/(n-s)) \leq n|a|$, so that

$$I_1 \leq \frac{r(n-s)}{n-s+r}\frac{n|a|}{R}.$$

In the same way, we have

$$I_2 \leq \frac{s(n-r)}{n-r+s}\frac{n|a|}{R}.$$

This implies that μ_{i_0} is at most $sn|a|/(n-r+s)$ and ν_{i_0} is at most $rn|a|/(n-s+r)$.

We will show that $r+s = n$, so that μ, ν is a minimal cover. If not, $r+s \leq n-1$, $\mu_{i_0} \leq snA/(2s+1)$ and $\nu_{i_0} \leq rnA/(2r+1)$ so that

$$\mu_{i_0} + \nu_{i_0} \leq \left(\frac{s}{n-r+s} + \frac{r}{n-s-r}\right) n|a| \leq (n-1)|a| \leq b_{i_0,j_0},$$

and we need have $\mu_{i_0} + \nu_{i_0} = b_{i_0,j_0}$. Equality is achived for some "lower right" element b_{i_0,j_0}, which contradicts minimality of $r+s$.

Precise computations of the values μ_i and ν_j would exceed the ambition of this example and are left to the reader. The conception of a better analog device for computing the tropical determinant may have some practical interest.

3.9 Conclusion of section 3

The best complexity bounds for the assigment problem rely on "scaling" methods, that is recursively replacing in A $a_{i,i}$ by $\lfloor a_{i,i}/2 \rfloor$ to obtain an approximate maximum, as in Gabow and Tarjan [27], where a $O\left(n^{5/2}\ln(nC)\right)$ complexity is achieved (with $a_{i,j} \leq C$). See Schrijver [74] or Burkard *et al.* [7] for more details.

The basic idea is to use Hopcroft and Karp algorithm, which is faster, to improve the approximation at each of the $\ln C$ steps; some extra $\ln n$ steps are required to check the last approximation and deduce from it the exact value.

We have seen that Jacobi's work contained the germs of important notions in combinatorial optimization and graph theory. The efficiency considerations in Jacobi's papers reflect his computational tools: pen and paper, but his algorithm for the assignment problem may be easily adapted to express improved complexity bounds obtained in the early seventies.

4 A differential parenthesis. Various forms of the bound.

4.1 Ritt's strong and weak bound

Jacobi did not mention what should be done if some variable x_j and its derivatives do not appear in some polynomial P_i. The easiest answer is to define it as 0, but a better choice in such a case is the convention introduced by Ritt [70] $\mathrm{ord}_{x_j} P_i = -\infty$. Lando [59] defined the first choice as the *weak bound*, and the second as the *strong bound*. Our definition also includes some minor modifications in order to extend the bound to underdetermined systems.

Definition 62. — *By convention,* $\mathrm{ord}_{x_j} P_i = -\infty$ *if* x_j *and its derivatives do not appear in* P_i. *Let* $a_{i,j} := \mathrm{ord}_{x_j} P_i$, *we define* $S_{s,n}$ *to be the set of injections* $[1, s] \mapsto [1, n]$. *We define* Jacobi's number *as* $\mathcal{O}_P := \mathcal{O}_A = \max_{\sigma \in S_{s,n}} \sum_{i=1}^{n} a_{i,\sigma(i)}$.

Remark 63. — An easy consequence of remark 49 is that, assuming that A is a matrix of non negative integers and $-\infty$ elements, with $\max_{i,j} a_{i,j} = C$ and $\mathcal{O}_A \in \mathbf{N}$, then, assuming that for the minimal canon λ, the sequence λ_i is non decreasing, $0 \leq \lambda_i \leq (i-1)C$, and the associated minimal cover α, β is such that $0 \leq \alpha_i \leq (n-i)C$, and $-(n-1)C \leq \beta_j \leq C$.

So, if A is a $n \times n$ matrix of integers and $-\infty$ values, we may use also any value smaller that $-n \max_{i,j} a_{i,j}$ instead of $-\infty$ to define a new matrix A' such that $\mathcal{O}(A') < 0 \Longleftrightarrow \mathcal{O}(A) = -\infty$.

If $s < n$ we can also complete the matrix $a_{i,j}$ of orders with $n - s$ rows of zeros, in order to get a square matrix A'. Jacobi's bound is then also equal to $\max_{\sigma \in S_n} \sum_{i=1}^{n} a'_{i,\sigma(i)}$, this equivalent definition allows to compute the bound, using Jacobi's algorithm.

Of course, we described this algorithm for matrices with coefficients in an ordered abelian group, and $\mathbf{Z} \cup \{-\infty\}$ with the convention $\mathbf{Z} - \infty = \{-\infty\}$ has not such a structure. So it is best to use the group $\{a\infty + b, (a, b) \in \mathbf{Z}^2\}$ with

$(a_1\infty + b_1) + (a_2\infty + b_2) = (a_1 + a_2)\infty + (b_1 + b_2)$ and $(a_1\infty + b_1) < (a_2\infty + b_2)$ $\Longleftrightarrow a_1 < a_2$ or $a_1 = a_2$ and $b_1 < b_2$.

Jacobi also introduces a determinant ∇, the non vanishing of which is a necessary and sufficient condition for the bound to be reached. In order to define it, he considers the matrix $(\partial P_i/\partial x_j^{(a_{i,j})})$ and forms its determinant. Then, he only keeps the products $\pm \prod_{i=1}^n \partial P_i/\partial x_{\sigma(i)}^{(a_{i,\sigma(i)})}$ such that $\sum_{i=1}^n a_{i,\sigma(i)} = \mathcal{O}$. this is why he calls this expression the *truncated determinant*[8] of the system. We may equivalently use the following definition.

Definition 64. — *Let A' be the order matrix of the system P, completed with $n-s$ rows of zeros and λ' be the minimal canon of A'. The minimal cover α', β' is defined as in definition 13: $\Lambda' := \max_{i=1}^s \lambda'_i$, $\alpha'_i = \Lambda' - \lambda'_i$ and $\beta'_j = \max_{i=1}^s a'_{i,j} - \alpha'_i$. We pose furthermore $B' := \max_{j=1}^n \beta'_j$ and $\mu'_j := B' - \beta'_j$.*

Let $Q \in k[x]$, we define $\mathrm{ord}_{x_j}^{\mathrm{J}} Q := \mathrm{ord}_{x_j} Q + \mu'_j$ and $\mathrm{ord}^{\mathrm{J}} Q = \max_{j=1}^n \mathrm{ord}_{x_j}^{\mathrm{J}} Q$.

Definition 65. — *We denote by $\mathfrak{J}_P$ the matrix $(\partial P_i/\partial x_j^{\alpha_i+\beta_j})$. If $s = n$, we call the* system determinant *and denote by ∇ the determinant of $\mathfrak{J}_P$. If $s < n$, ∇ will denote the set determinant of all $s \times s$ submatrices of $\mathfrak{J}_P$.*

It is straightforward that this definition of ∇ is equivalent to Jacobi's one, which is to keep in the Jacobiaan determinant only the terms corresponding to maximal sums in the order matrix. Partial derivatives $\partial P_i/\partial x_j^{\alpha_i+\beta_j}$ are in fact non zero iff $\mathrm{ord}^{\mathrm{J}} x_j^{a_{i,j}} = \mathrm{ord}^{\mathrm{J}} P_i$. Moreover, this result stands for any cover, not only Jacobi's cover.

Proposition 66. — *Let μ_i, ν_j be any cover for the matrix $A := (\mathrm{ord}_{x_j} P_i)$, $\nabla = |\partial P_i/\partial x_j^{(\mu_i+\nu_j)}|$.*

4.2 Reduction to order 1

We conclude with the well known reduction to first-order equations. Lando [59] did prove Jacobi's bound for order one systems, also considering underdetermined systems, but only with the weak bound. She remarks that the weak bound for the first order reduction may be greater than that of the original system, but that the strong bound remains the same. We can even prove that the truncated determinant is unchanged, up to sign.

We introduce new variables $u_{j,k}$ for $1 \le i \le n$ and $0 \le k < r_j := max_{i=1}^s \mathrm{ord}_{x_j} P_i$ and replace in the equations P_i $x_j^{(k)}$ by $u_{j,k}$ for $0 \le k < r_j$ and by u'_{j,r_j-1} for $k = r_j$, obtaining a new equation Q_i. We complete the new system with the equations $W_{j,k} := u'_{j,k-1} - u_{j,k}$ for $1 \le k < r_j$.

[8] *Determinans mancum*, or *Determinans mutilatum*.

Lemma 67. — Let P be a system of n differential polynomials in $\mathcal{F}\{x\}$, and Q, W the system of $n + \sum_{i=1}^{n} r_i$ equations in $\mathcal{F}\{y\}$ obtained by reduction to the first order, as defined above, the P_i and the then $\mathcal{O}_P = \mathcal{O}_{Q,W}$, $\nabla_P = \pm\nabla_Q$ and in the canon ℓ for the order matrix B for the system Q, W the integers ℓ_i, $1 \le i \le s$ are the same than the λ_i in the canon λ for the order matrix A of P.

Proof. — We choose to put first in the system the P_i, in the same order, and then the $W_{j,k}$ and use on the $u_{j,k}$ the order $u_{j,k} < u_{j',k'} \iff j < j'$ or $j = j'$ and $k > k'$. To build the order matrix A_Q, we take first u_{1,r_1-1}, ..., u_{n,r_n-1} and then the $W_{j,k}$ in the same order as the $u_{j,k}$. We will show that $\nabla_P = \pm\nabla_Q$; it is clear that the choice of a different ordering can only change the sign of ∇_Q.

The order matrix B of Q has the following shape: $(L_1 \cdots L_n)$ with $L_j :=$

$$\begin{pmatrix}
\genfrac{[}{]}{0pt}{}{1,j}{r_j-1} & \cdots & \genfrac{[}{]}{0pt}{}{1,j}{k} & \cdots & \genfrac{[}{]}{0pt}{}{1,j}{0} \\
\vdots & & \vdots & & \vdots \\
\genfrac{[}{]}{0pt}{}{i,j}{r_j-1} & \cdots & \boldsymbol{\genfrac{[}{]}{0pt}{}{i,j}{k}} & \cdots & \genfrac{[}{]}{0pt}{}{i,j}{0} \\
\vdots & & \vdots & & \vdots \\
\genfrac{[}{]}{0pt}{}{n,j}{r_j-1} & \cdots & \genfrac{[}{]}{0pt}{}{n,j}{k} & \cdots & \genfrac{[}{]}{0pt}{}{n,j}{0} \\
\multicolumn{5}{c}{\cdots \sum_{\jmath<j} r_\jmath \textit{ rows} \cdots} \\
\mathbf{0} \quad 1 & & & & \\
\quad \mathbf{0} \quad 1 & & & & \\
& \ddots \ \ddots & & & \\
& & \mathbf{0} \quad 1 & & \\
& & 0 \quad \mathbf{1} & & \\
& & \quad 0 \quad \mathbf{1} & & \\
& & & \ddots \ \ddots & \\
& & & & 0 \quad \mathbf{1} \\
\multicolumn{5}{c}{\cdots \sum_{\jmath>j} r_\jmath \textit{ rows} \cdots}
\end{pmatrix}.$$

For more readability, only terms possibly different from $-\infty$ are displayed. The terms $\genfrac{[}{]}{0pt}{}{i,j}{r_j-1}$ are 1 if $\mathrm{ord}_{x_j} P_i = r_j$, 0 if $\mathrm{ord}_{x_j} P_i = r_j - 1$ and $-\infty$ otherwise. The terms $\genfrac{[}{]}{0pt}{}{i,j}{k}$ for $0 \le k < r_j - 1$ are 0 if $x_j^{(k)}$ appear in P_i and $-\infty$ otherwise. It is easily seen that n transversal non elements in the first n rows may be completed in a maximal set of transversal non $-\infty$ elements in one and only one way. Indeed, once an element $\boldsymbol{\genfrac{[}{]}{0pt}{}{i,j}{k}}$ is chosen in L_j there is a unique choice of integer tranversal element in L_j, represented in bold above. And among them there are exactly k **1**, so that their sum is equal to k. These elements appear in **bold** in the figure above.

Hence, there is a one to one correspondence between maximal transversal sums in A and B, so that both matrices have the same tropical determinant.

The structure of the system determinant $\nabla(Q)$ is similar and an easy computation relying on the number of inversions shows that

$$\nabla(Q) = (-1)^{\sum_{j=1}^{n}(n-j+1)r_j}\nabla(P).$$

Indeed, for any term in a determinant that corresponds to the permutation σ, we may define an inversion as the case when $i > i'$ and $\sigma(i) < \sigma(i')$. So the number of inversions is $\sum_i \iota_i$ with $\iota_i := \sum_{i>i',\, \sigma(i)<\sigma(i')} 1$. Then, the signature $e(\sigma)$ of a permutation being determined by the number ρ of transpositions τ_j in any decomposition $\sigma = \tau_1 \circ \cdots \tau_\rho$, $e(\sigma) = \sum_i \iota_i$.

We have shown that for any permutation σ for A is associated to a single permutation in B. First, one sees that the inversion numbers ι_i for the first n lines are the same for A and B. Then, one only has to consider the remaining lines that are decomposes in blocks

$$M_j := \left(\begin{array}{ccccccccc}
 & \cdots & & \sum_{\hat{\jmath}<j} r_{\hat{\jmath}} \textit{ rows} & & & \cdots & & \\
\mathbf{1} & -1 & & & & & & & \\
 & \mathbf{1} & -1 & & & & & & \\
 & & \ddots & \ddots & & & & & \\
 & & & \mathbf{1} & -1 & & & & \\
 & & & & 1 & \mathbf{-1} & & & \\
 & & & & & 1 & \mathbf{-1} & & \\
 & & & & & & \ddots & \ddots & \\
 & & & & & & & 1 & \mathbf{-1} \\
 & \cdots & & \sum_{\hat{\jmath}>j} r_{\hat{\jmath}} \textit{ rows} & & & \cdots & &
\end{array}\right).$$

The presence of the -1 terms compensates the extra inversions, so that we may reduce to the diagonal of 1, for which the total inversion number in block M_j is $n-j+1$ for each term and that produces a total of $(n-j+1)r_j$.

In order to show that $\ell_i = \lambda_i\ 1 \leq i \leq s$, we first remark that if λ is a canon for B and $\sum_j a_{\sigma(j),j}$ a maximal transversal sum, then the rows of B corresponding to equations $W_{j,w}$, $w \geq a_{\sigma(j),j}$ must be respectively increased by $\max_{i|a_{i,j}>a_{\sigma(j),j}}(a_{i,j} - w + \lambda_i - 1$, provided that this quantity is positive. On the over hand, for $w < a_{\sigma(j),j}$, the rows must be increased by $\max_i \min_{k\leq w|a_{i,k}\neq-\infty} \lambda_i + k - w$.

The definition of a canon for the column of $u_{j,w}$ implies that we must have $\max_i a_{i,j} + \lambda_j \leq a_{\sigma(j),j} + \lambda_{\sigma(j))}$. Then, the λ_i of the minimal canon for B are the minimal integers with this property, which also characterizes the canon of A, so that $\lambda_i = \ell_i$, $1 \leq i \leq n$. ∎

Remark 68. — It may be difficult to model $-\infty$ entries for a matrix in some computer algebra system. An easy trick is to replace them by some suitable negative value, say $D := -(nC + 1)$ if A is matrix of non negative integers and $-\infty$ elements with $\max_{i,j} a_{i,j} = C$. Then, the tropical determinant of A is $\mathcal{O}(A) = -a\infty + b$ iff that of the new matrix A' is $\mathcal{O}(A') = aD + b$ with $-(a)D \leq \mathcal{O}(A') < -(a-1)D$.

4.3 Block decomposition

If the integer elements in the order matrix admit a minimal cover of a rows and b columns with $0 < a, b < n$, then the system P admits a non trivial triangular block decomposition. In the case where $s = n$ and $a + b = n$, one may look for such a block decomposition using the reflexive transitive closure of the elementary path relation, as defined in subsubsection 3.3. One gets so a partial preorder that defines equivalence classes of rows i, j with $i \prec j$ and $j \prec i$. Sorting the variables and equations according to this preorder produces a block decomposition, the block corresponding to these equivalence classes, that do not depend on the choice of a maximal transversal family by prop. 47.

In the same spirit, considering the reflexive, transitive and symmetric closure provides a diagonal block decomposition. We will not develop these easy results, but they can be very helpfull to clarify the structure of a system before any attempt to solve it, whenever its size makes difficult to find the requested form by simple inspection.

5 An algebraic parenthesis. Quasi-regularity and "Lazard's lemma"

Jacobi considers functions without any precision about their nature. One may present his results in the framework of *diffiety theory*, provided that the equations are defined by $\mathcal{C}^\infty$ functions, satisfying some natural regularity hypotheses (see [66]). We use here the formalism of Ritt's differential algebra, that allows effective computations. Here characteristics set will be used instead of Jacobi's "normal forms", and Lazard's lemma will take the place of the implicit function theorem.

5.1 Quasi-regularity

As we will see, quasi-regularity, although it remains an implicit hypothesis, plays a central role in Jacobi's proof of the bound. The informal meaning of this notion is that a differential system $P_i(x) = 0$, "behaves like" the linearized system $\mathrm{d}P_i = 0$, *viz.* $\sum_{j=1}^{s} \sum_{k=0}^{\infty} \partial P_i/\partial x_i^{(k)} \mathrm{d}x_1^{(k)} = 0$, in the neighborood of a generic point of some component of $\{P\}$.

This idea was formalized by Johnson [45, 46, 47] who used it to prove Janet's conjecture [48]. It is also the key of the first complete proof of Jacob's bound in the non linear case, given by Kondratieva *et al.* [53]. Ritt was able to prove the bound for general components, that is without the quasi-regularity hypothesis, but only for $s = n = 2$ [71, Chap. VII 6. p. 136].

We will provide here a more general definition than the one used in [53, 66], in order to underline that the property used is wider than the "independence" of Kähler differentials $\mathrm{d}P_i$ of which it is a consequence, in the spirit of the "regular" differential ideals, as defined by Johnson [47]. Quasi-regular was chosen because this property is shared by some components of a differential equations, that the classical theory considers as "singular" (See Houtain [33] or Hubert [34]).

5.1.1 Notations and definitions

In the following, $\mathcal{F}$ will denote a differential field of characteristic 0. We refer to Ritt [71] and Kolchin [51] for more details about diffferential algebra, and to Boulier [5] for characteristic sets. It is natural here to state the definition for an arbitrary differential field $\mathcal{F}_\Delta$, with a finite set $\Delta := \{\delta_1, \ldots, \delta_m\}$ of commuting derivations, possibly empty.

Definition 69. — *Let $\mathcal{G}/\mathcal{F}$ denote the differential field extension defined by $\mathcal{P} \subset \mathcal{F}\{x_1, \ldots, x_n\}$, Δ denote the set of derivations of the differential fields $\mathcal{F}$ and $\mathcal{G}$. We denote by $\mathcal{D} := \mathcal{G}[\Delta]$ the non commutative ring of differential operators and by $\mathcal{M}$ the module $\mathcal{G} \otimes_{\mathcal{F}\{x\}} \Omega_{\mathcal{F}\{x\}/\mathcal{F}}$; $\mathcal{M}$ is a $\mathcal{D}$ free module generated by $\mathrm{d}x_i$, $1 \leq i \leq n$. For any $Q \in \mathcal{F}\{x\}$, $\mathrm{d}Q \in \mathcal{M}$ denotes the differential of Q.*

Let P_i $1 \leq i \leq s$ be differential polynomials in $\mathcal{F}\{x_1, \ldots, x_n\}$, and $\{P\} = \cap_{j=1}^{r} \mathcal{P}_j$, where the $\mathcal{P}_j$ are prime differential ideals such that $\mathcal{P}_i \subset \mathcal{P}_j$ implies $i = j$. The prime ideals $\mathcal{P}_j$ are called the components *of $\{P\}$. We say that $\mathcal{P}_j$ is a quasi-regular component of the system $P = \{P_1, \ldots, P_s\}$ if $\mathrm{d}\mathcal{P} = \langle \mathrm{d}P \rangle_{\mathcal{M}}$.*

In the ordinary differential case, we say that $\mathcal{P}_j$ is strongly regular *if the family $(\delta^k \mathrm{d}P_i)$, $k \in \mathbf{N}$, $1 \leq i \leq s$, is linearly independent.* Obviously, strong regularity, that corresponds to the usual definition, implies regularity.

This property is very usefull, mostly combined with the following properties of $\mathrm{d}\mathcal{P}$.

Lemma 70. — *Let $\mathcal{P}$ be a quasi-regular component of P.*

i) A characteristic set $\mathcal{A}$ exists for $\mathcal{P}$ for some ordering $\prec$ on derivatives with main derivatives v_i, $1 \le i \le n$ iff a standard basis exists for the $\mathcal{D}$-module $\mathrm{d}\mathcal{P}$, for the ordering induced by $\prec$ on differentials $\mathrm{d}\mathrm{Y}$ (where Y denotes the set of dedivative of the x_i), with main derivatives $\mathrm{d}v_i$, $1 \le i \le$.

i′) The main component $\mathcal{P} = [\mathcal{A}] : H_{\mathcal{A}}^{\infty}$[9] of a system $\mathcal{A}$ that is a characteristic set of some prime differential ideal $\mathcal{P}$ is regular and strongly regular in the ordinary case.

ii) Let $Y \subset \{x_1, \ldots, x_n\}$, $\mathcal{P} \cap \mathcal{F}\{Y\} \neq (0)$ iff $(\mathrm{d}P) \cap (\mathrm{d}Y) \neq (0)$.

iii) The component $\mathcal{P}$ is strongly regular iff it is regular and of codimension s.

Proof. — i) $\Longrightarrow$. Assume that $\mathcal{A}$ is a characteristic set of $\mathcal{P}$ for some ordering $\prec$. Then any $Q \in \mathcal{P}$ is reducible by $\mathcal{A}$ so that $\mathrm{d}Q$ is also reducible by $\mathrm{d}\mathcal{A}$. It is easilly checked that (A_1, A_2) is reducible by $\mathcal{A}$ implies that $(\mathrm{d}A_1, \mathrm{d}A_2)$ is reducible by $\mathrm{d}\mathcal{A}$, so that $\mathrm{d}\mathcal{A}$ is a standard basis for the $\mathcal{D}$-module $\mathrm{d}\mathcal{P}$ and for the ordering induced by $\prec$.

$\Longleftarrow$. If G is a standard basis of $\mathrm{d}\mathcal{P}$ for some ordering, consider a characteristic set $\mathcal{A}$ of $\mathcal{P}$ for the corresponding ordering. By what preceeds, $\mathrm{d}\mathcal{A}$ is also a standard basis for the same ordering and G and $\mathrm{d}\mathcal{A}$ have the same leading terms, hence the result.

i′) It is a straightforward consequence of i).

ii) Using i), it is enough to consider an ordering $\prec$ that eliminates letters not in Y.

iii) The component $\mathcal{P}$ is of codimension s iff there is no non trivial relations between the $\mathrm{d}P_i$ and their derivatives. ∎

In the "algebraic" case, that is when $\Delta = \emptyset$, Lazard's lemma provides a simple criterion for quasi-regularity.

5.2 Lazard's lemma

Many proofs of this folkloric result are already available in the differential algebra literature (see *e.g.* Morrison [61] or Boulier *et al.*[5]). The interest of the following one is to make a link with the implicit function theorem by using Newton's method.

Theorem 71. — *Let $P_1, \ldots, P_s$ be polynomials in $k[x_1, \ldots, x_n]$ with $s \le n$ and $\mathcal{J} := (\partial P_i/\partial x_j | 1 \le i, j \le s)$. If $\mathcal{Q} := (P) : |\mathcal{J}|^{\infty} \neq (1)$, then*

[9] We denote by $H_{\mathcal{A}}$ the product of initials and separants of polynomials in $\mathcal{A}$.

i) Any component $\mathcal{P}$ of $\mathcal{Q}$ is a quasi-regular component of the system P and of codimension s;

ii) $\mathcal{Q} \cap k[x_{s+1}, \ldots, x_n] = (0)$*;*

iii) $\mathcal{Q}$ *is radical.*

Proof. — i) We notice that $\mathrm{d}P_i = \sum_{j=1}^{n} \partial P_i/\partial x_j \mathrm{d}x_j$, so that the differentials $\mathrm{d}P_i$, $1 \leq i \leq s$, are linearly independent. This means that the codimension of $\mathcal{P}$ is at least s. By the dimension theorem, it is at most s, so that it is s, a characteristic set of $\mathcal{P}$ as s elements and $\mathrm{d}\mathcal{P} = \langle \mathrm{d}P \rangle$.

ii) is a straightforward consequence of i) and lemma 70 ii).

iii) We will denote here by $\mathcal{D}$ the ring$\mathcal{F}\{x\}/\mathcal{Q}[\Delta]$ and by $\mathcal{M}$ the module $\mathcal{F}\{x\}/\mathcal{Q} \otimes_{\mathcal{F}\{x\}} \Omega_{\mathcal{F}\{x\}/\mathcal{F}}$; $\mathrm{d}Q$ will denote the differential of Q in $\mathcal{M}$ using this definition. Let G be a standard basis of $\sqrt{\mathcal{Q}}$, we will show that $G \subset \mathcal{Q}$. Let $\tilde{\mathcal{J}}$ denote the adjugate matrix of $\mathcal{J}$, then for any $Q \in G$,

$$|\mathcal{J}|Q = (\partial Q/\partial x_1, \ldots, \partial Q/\partial x_s)\tilde{\mathcal{J}}(P_1, \ldots, P_s)^{\mathrm{T}}[\sqrt{\mathcal{Q}}^2],$$

as $\mathrm{d}Q_2 = 0$. Assume that $\forall Q \in G$, $\mathcal{J}^\alpha Q = Q_p \in \mathcal{Q}[\sqrt{\mathcal{Q}}^{2^p}]$, then the elements of $\mathcal{Q}^{2^p}$ are linear combinations of monomials depending on the $Q \in G$. Applying this result recursively, we find $\mathcal{J}^{\alpha^{2^p}+\alpha}Q = Q_{p+1} \in \mathcal{Q}[\sqrt{\mathcal{Q}}^{2^{p+1}}]$. For p great enough, $\sqrt{\mathcal{Q}}^{2^p} \subset \mathcal{Q}$, hence the result. ∎

Corollary 72. — *Assume that* $P_{s_0+1}, \ldots, P_s$ *belong to* $k[x_{s_0+1}, \ldots, x_n]$.

i) For any prime component $\mathcal{P}$ *of* $(P) : |\mathcal{J}|^\infty$, $\mathcal{P} \cap k[x_{s_0+1}, \ldots, x_n]$ *is a prime component of* $(P_{s_0+1}, \ldots, P_s) : |\mathcal{J}_0|^\infty$, *where* $\mathcal{J}_0$ *is the jacobian matrix of the polynomials* $P_{s_0+1}, \ldots, P_s$ *with respect to the variables* $x_{s_0} + 1, \ldots, x_s$.

Proof. — i) Let $\mathcal{A}$ be a characteristic set of $\mathcal{P}$ for the ordering $x_1 > \cdots > x_n$. If $A_1 > \cdots > A_s$, the set $\{A_{s_0+1}, \ldots, A_s\}$ is a characteristic set of $\mathcal{P} \cap k[x_{s_0+1}, \ldots, x_n]$ ([71] § 17–19 p. 88–90). This ideal must be included in a prime component of $(P_{s_0+1}, \ldots, P_s) : |\mathcal{J}_0|^\infty$, which, according to the theorem, has the same codimension $s - s_0$, so that it must be equal to $\mathcal{P} \cap k[x_{s_0+1}, \ldots, x_n]$. ∎

6 Jacobi's proof of the bound

In manuscript [II-13 b)] (*cf.* [36, prop. 1 p. 16 and pro. 2 p. 17], Jacobi gives two different versions of this result. In the first, he writes that the order is H, but in the second, he claims that the order is H iff the truncated determinant vanishes. A modern reader may be surprised by this way of giving in a first theorem a generic result, and then describing more precisely, in a second theorem, the possible exceptions to the first one. However, such a style of presentation is, as we have seen, quite common in the set of manuscripts we consider here. Jacobi's proof also contains paradoxical arguments that led Ritt to conclude it was whimsical.

It seems however possible to save the proof and get the second version of Jacobi's theorem, more precise than that of Kondratieva *et al.* [53].

We first prove the theorem in its original form, assuming $s = n$.

THEOREM 73. — *Let $\mathcal{O}$ denote Jacobi's bound for the system P_i, $1 \leq i \leq n$ of differential polynonials in $k[x_1, \ldots, x_n]$ and $\mathcal{P}$ be a strongly quasi-regular component of P.*

i) The order of $\mathcal{P}_j$ is at most $\mathcal{O}$.

ii) The order of $\mathcal{P}_j$ is equal to $\mathcal{O}$ iff $\nabla \notin \mathcal{P}$.

PROOF. —

Before considering Jacobi's arguments, we need first the following lemma.

Lemma 74. — *If $\mathcal{P}$ is a strongly quasi-regular component of P, it is of differential dimension* 0 *and* $\mathcal{O} \in \mathbf{N}$[10].

PROOF. — The first part of the claim is lemma 70 iii). The variant of of Kőnig's theorem (see above th. 11), that is also stated by Cohn [13], shows that, if $\mathcal{O} = -\infty$, then one may find a rows and b columns in A', containing all the elements in $\mathbf{N}$, with $a + b < n$. So that $n - b > a$ equations in $\mathrm{d}P$ must depend of a differentials $\mathrm{d}x_j$, which contradict strong quasi-regularity. ∎

Remark 75. — Generalizing this lemma to an arbitrary components is related to a difficult conjecture: the *dimensional conjecture.* Cohn has shown that it would be implied by Jacobi's bound (even week) for arbitrary systems [13].

a) First argument *Linearization.* — Jacobi first claims that one may reduce the problem to the case of a linear system. This, of course, cannot stand in all cases: we needed the strong quasi-regularity hypothesis. We can assume that such assumptions were implicit in the physical situations that were considered by Jacobi: for proving that the order of P corresponds to the dimension of the space of solutions of $\mathrm{d}P$, Jacobi used the fact that, if the set of solutions of P depends of parameters γ_i, then $\partial P/\partial \gamma_i$ is a solution of the linearized system $\mathrm{d}P = 0$.

The order of the differential field extension $\mathcal{G}/\mathcal{F}$ is the dimension of the quotient module $\Omega_{\mathcal{G}/\mathcal{F}} = \mathcal{M}/(\mathrm{d}P)_{\mathcal{M}}$ by lemma 70 i).

b) Second argument *Stationnary systems* — Jacobi claims then that one can assume the linearized system $\mathrm{d}P$ to have constant coefficients. This affirmation seems really paradoxical, but it is dubious that he could have written it without a precise idea in mind. One may also notice that Cohn and Borchardt made no remark on that point. We have only the following indication, the start of an

[10] In the linear case the result has been proved by Ritt [70].

argument that have been ruled out by Jacobi: *In explogrando ordine systematis cum tantum altissima differentialia respiciuntur, in æquationibus differentialibus linearibus, ad quas proposita revocata sunt, supponere licet Coëfficientes esse constantes. Nam æquationibus 3) iteratis vicibus differentiatis, ut novæ obtinentur æquationes [...]* [11].

We propose the following argument, inspired by the theory of standard bases of $\mathcal{D}$-module (*cf* Catro-Jiménez [8], which agrees with Jacobi's idea of looking at highest derivatives in the linearized system $\mathrm{d}P$.

Definition 76. — *We denote by K the field $\mathcal{G}$ equipped with the derivation δ_0 with $\delta_0 c = 0$ $\forall c \in K$, and by $\mathcal{M}_0$ the free $K[\delta_0]$-module generated by the $\mathrm{d}x_i$.*

Let $m \in \mathcal{M}$ or $m \in \mathcal{M}_0$, $m = \sum_{v\in \Upsilon} c_v v$, where Υ denotes the set of derivatives of the x_i. We extend the definition of $\mathrm{ord}^{\mathrm{J}}$ *to $\mathcal{M}$ and $\mathcal{M}_0$ and define the* head *of m to be $\kappa\, m := \sum_{\mathrm{ord}^{\mathrm{J}} v = \mathrm{ord}^{\mathrm{J}} m} c_v v$, the sum of terms of greatest order.*

Lemma 77. — *If $\nabla \in \mathcal{P}$, then $\mathcal{P}$ is a strongly quasi-regular component of $\{P\}$ and $\kappa\,(\mathrm{d}P)_{\mathcal{M}} = \kappa\,(\kappa\,\mathrm{d}P)_{\mathcal{M}}$ and $\kappa\,(\mathrm{d}P)_{\mathcal{M}_0} = \kappa\,(\kappa\,\mathrm{d}P)_{\mathcal{M}_0}$.*

Proof. — If $\nabla \notin \mathcal{P}$, the matrix $\mathcal{J} = (\partial P_i/\partial x_j^{(\alpha_i+\beta_j)}$ is invertible in $\mathcal{G}$, so that the families $\kappa \mathrm{d}P_i = \sum_{j=1}^{n} \partial P_i/\partial x_j^{(\alpha_i+\beta_j)}$ and $\mathrm{d}P$ generate respectively free submodules of $\mathcal{M}_0$ and $\mathcal{M}$, with $\kappa(\mathrm{d}P) = \kappa(\kappa \mathrm{d}P)$. ∎

So, if $\nabla \notin \mathcal{P}$, it is indeed enough to prove the bound for some constant coefficient linear system. Assume that $\nabla \in \mathcal{P}$, then let i_0 be the smallest integer such that the first i_0 lines of $\mathcal{J}$ are dependent. We may find some $c_i \in \mathcal{G}$, $1 \le i \le i_0$ with $c_{i_0} \neq 0$ such that $\sum_{i=1}^{i_0} c_i \delta_0^{\lambda_i} \kappa\, \mathrm{d}P_i = 0$. So, $(\mathrm{d}P)$ is generated by the family $\mathrm{d}P_1, \ldots, \mathrm{d}P_{i_0-1}, \sum_{i=1}^{i_0} c_i \delta^{\lambda_i} \mathrm{d}P_i = 0, \mathrm{d}P_{i_0+1}, \ldots, \mathrm{d}P_n$. We may compute Jacobi's bound for this new linear system, that will be strictly smaller than $\mathcal{O}$.

We may iterate the process until we find a free linear system m_i, $1 \le i \le n$, generating $(\mathrm{d}P)$, with a non vanishing system determinant ∇. This must happen, for $\mathcal{O} \ge 0$ and if $\mathcal{O} = 0$, as m generates $(\mathrm{d}P)$, which is of differential dimension 0, ∇ cannot vanish.

iii) Third argument *Determinant degree.* Assume that we have a linear system with constant coefficient $m_i = 0$, $1 \le i \le n$. We may represent it as a matrix of differential operators $M(\delta_0)$ with $m_{i,j} = \sum_{p=0}^{a_{i,j}} c_{i,j,p} \delta_0^p$: $M(\delta_0)(\mathrm{d}x_1, \ldots, \mathrm{d}x_n)^{\mathrm{T}} = 0$. The number of independent solutions of such a system is the number of roots ξ of $|M(y)| = 0$. Jacobi did only consider the simple case of all different roots. The general situation was later investigated by Chrystal [11].

[11]Looking for the order of the system, as one only considers the highest derivatives in the linear equations to which the proposed ones are reduced, one may assume [their] coefficients to be constants. For differentiating the equations 3) iterated times in order to obtain new equations [...]

This equation has degree at most $\mathcal{O}_m := \max_{\sigma\in S_n} \sum_{i=1}^n a_{i,j}$, and the coefficient of $y^{\mathcal{O}_m}$ is equal to $\nabla_m = |c_{i,j,a_{i,j}}|$, so that the order of the system is exactly $\mathcal{O}$ if $\nabla \notin \mathcal{P}$. This concludes the proof of the theorem. ∎

It is now easy to extend Jacobi's bound to underdetermined systems. We need first to define the order of such a system, by analogy with the degree of an algebraic system, as done in [66].

Definition 78. — *Let $\mathcal{P}$ be a prime differential ideal of $\mathcal{F}\{x\}$ of differential dimension m. The order of $\mathcal{P}$ is the maximal order of quasi-regular components of differential dimension 0 of the ideals $\mathcal{P} + [L_1, \ldots, L_m]$, where the L_i are linear equations of order* 0*, with coefficients in $\mathcal{F}$.*

Corollary 79. — *For any strongly quasi-regular component $\mathcal{P}$ of P, the order of $\mathcal{P}$ is at most $\mathcal{O}$.*

The order is equal to $\mathcal{O}$ iff the matrix $\mathcal{J}_P$[12] has full rank in $\mathcal{G}$.

Proof. — We have seen that $\mathcal{O}$ is obtained by completing matrix A with $n-s$ lines of zeros that corespond to the orders of generic linear equations L_i. So, the theorem, applied to the system P, L implies that the order of $\mathcal{P}$ is bounded by $\mathcal{O}$.

We may find coefficients for the L_i such that $\nabla_{P,L}$ does not vanish iff $\mathcal{J}_P$ has full rank. So, using theorem again, the order is equal to $\mathcal{O}$ iff $\mathcal{J}_P$ has full rank. ∎

7 Shortest normal form reduction

We consider here one of Jacobi's results that may have the greatest consequences for improving the resolution of differential systems in most practical cases. Jacobi describes the method that, generically, *i.e.* when his system determinant ∇ does not vanish, allows to compute a normal form or a characteristic set, using as few derivatives as possible of the system equations: more precisely, it is enough to differentiate P_i up to order λ_i, where λ is the minimal canon and, generically, it is impossible to compute a normal form by differentiating one of the P_i a smallest number of times.

In fact, except for minimality, Jacobi's results stand for any canon. One may guess that Jacobi was aware of this fact although he did not state it explicitly. In [37, § 3], he claims indeed that, if a normal form can be computed using the equations P_1, ..., P_n and a minimal (for inclusion) set of derivatives of these equations, then there exist a permutation σ such that the highest derivative of x_j appearing in these equations, appear in in the highest derivative of $P_{\sigma(j)}$. Such a property is not general, but is valid if the truncated determinant does not vanish and the

[12]See def. 65 p. 39.

normal form is associated to a canon and what we have called Egerváry ordering on the derivatives.

This method may be suggested as a default strategy in computer algebra systems, when it is requested to compute a characteristic set without specifying a precise ordering. It may also be used as a first step in methods using a change of orderings, such as *Pardi !*, designed by Boulier *et al.* [4].

Shaleninov [72] and Pryce [68] proposed strategies for the integration of implicit DAE that turn to be equivalent to Jacobi's shortest reduction. It seems that in many practical situations ∇ actually does not vanish, so that this method can be efficiently used.

Jacobi only considers the case when there are as many equations as variables. The generalization to underdetermined systems is easy.

Definition 80. — *Let I and $\mathcal{J}$ be two ideals of some ring, we denote by $I : \mathcal{J}^{\infty}$ the ring $\{a | \forall b \in \mathcal{J}\, \exists n \in \mathbf{N}\ ab^n \in I\}$.*

It is well known, using a folkloric version of Rabinovich trick [69], that, if $I, \mathcal{J} \subset k[x]$ and Q_i $1 \le i \le s$ generate $\mathcal{J}$, then $I : \mathcal{J}^{\infty} = (I; \sum_{i=1}^{s} u_i Q_i - 1) \cap k[x]$. The following proposition is easily established.

Proposition 81. — *If $I, \mathcal{J} \subset k[x]$ and $\mathcal{J} = (Q_i | 1 \le i \le q)$, $\sqrt{I} : \mathcal{J}^{\infty} = \cap_{i=1}^{q} \sqrt{I} : Q_i^{\infty}$.*

Definition 82. — *Let A be an order matrix for a differential system, α, β a cover for A, we say that an ordering $\prec$ on derivatives is* an Egerváry ordering, *if $k - \beta_j < k' - \beta_{j'}$ implies that $x_j^{(k)} \prec x_{j'}^{(k')}$.*

An Egerváry ordering is the Jacobi ordering *if α, β is the canonical cover (as defined in prop. 14).*

Remark 83. — Considering these more general Egerváry ordering may prove to be usefull even if they require a greater number of derivation. E.g. if one needs a characteristic set that is by chance of this kind, then it should be easier to compute it directly than computing first a characteristic set for a Jacobi ordering and then a package such as Pardi. This of course requires experimentations. However, there is no extra work to expose this more general case.

Theorem 84. — *i) The ideal $\mathcal{Q} := [P] : (\nabla)^{\infty}$ is radical.*

ii) Let λ be a canon for the order matrix of P, μ, ν the associated cover and $\prec$ the corresponding Egerváry ordering on derivatives, as defined in def. 82. Let $\mathcal{I} := (P_i^{(k)} | 1 \le i \le s, 0 \le k \le \lambda_i) : (\nabla)^{\infty}$, we may find a decomposition $\mathcal{I} = \cap_{i=1}^{q} \mathfrak{I}_i$ such that $\mathfrak{I}_i$ is a radical ideal that admits an algebraic characteristic set $\mathcal{B}_i$ for $\prec$ with $\mathfrak{I}_i = (\mathcal{B}_i) : H_{\mathcal{B}_i}^{\infty}$. Let $\mathcal{A}_i$ be the minimal autoreduced set contained in $\mathcal{B}_i$, there exists a decomposition $\mathcal{Q} = \cap_{i=1}^{q} \mathcal{D}_i$, where the $\mathcal{D}_i$ are radical differential ideals with characteristic set $\mathcal{A}_i$ for $\prec$ and $\mathcal{D}_i = \mathcal{A}_i : H_{\mathcal{A}_i}^{\infty}$.

iii) Assume that the polynomials A_i in $\mathcal{A}_h$ are indexed by increasing order, then $\mathrm{ord}^{\mathcal{J}} A_i = \mu_i$.

*iv) Assume that the n matrices obtained by suppressing the row i_0 in $\mathcal{J}_P$ and the column j, $1 \le j \le n$, have all full rank in some prime component $\mathcal{P}$ of $\mathcal{Q}$. Take integers ℓ_i, $1 \le i \le s$ such that $\ell_{i_0} < \lambda_{i_0}$, where λ is the **minimal** canon. Let $\mathcal{A}$ be any characteristic set for any ranking of $\mathcal{P}$, there exists some element A in $\mathcal{A}$ that is not a zero divisor modulo $(P_i^{(k)}|1 \le i \le s, 0 \le k \le \ell_i) : (\nabla)^\infty$.*

Proof. — i) Let $\mathcal{P}$ be a prime component of $\mathcal{Q}$. We assume that the equations P_i are ordered according to the ordering induced by $\prec$ on derivatives. Let $x_{j_1}^{(\mu_1+\nu_{j_1})}$ be the greatest derivative according to $\prec$ such that $\partial P_1/\partial x_{j_1}^{(\mu_1+\nu_{j_1})} \notin \mathcal{P}$. We recursively define $x_{j_i}^{(\mu_i+\nu_{j_i})}$ to be the greatest derivative such that the determinant Δ_i of the minor of $\mathcal{J}_P$ contained in the first i rows and the columns j_1, ..., j_i does not belong to $\mathcal{P}$.

For $r \in \mathbf{N}$, we consider the algebraic system $P^{[r]}$ defined by the derivatives $P_i^{(k)}$, $1 \le i \le s$, $0 \le k \le \lambda_i + r$. Its Jacobian matrix with respect to the derivatives appearing in it contain a maximal minor corresponding to the derivatives $x_{j_i}^{(k)}$, $1 \le i \le s$ $\mu_i + \nu_{j_i} \le k \le \mu_0 + \nu_{j_i} + r$, the determinant of which is $\Delta^{[r]} := \Delta_s^{r+1} \prod_{i=1}^{s-1} \Delta_i^{\lambda_i - \lambda_{i+1}}$—with the convention $\lambda_0 = 0$! So, its determinant is not in $\mathcal{P}$ and we can apply Lazard's lemma to the ideal $(P^{[r]}) : \Delta_{[r]}^\infty$ and conclude that it is a non trivial radical ideal.

The ideal $\mathcal{Q}_\mathcal{P} := \cup_{r\in\mathbf{N}}(P^{[r]}) : \Delta_{[r]}^\infty$ is thus a radical differential ideal, contained in $\mathcal{Q}$ and containing $\mathcal{P}$. This proves that $\mathcal{Q}$ is equal to the intersection of radical ideals $\cap_\mathcal{P}\mathcal{Q}_\mathcal{P}$, and so is radical. This achieves the proof of i).

ii) We may find a decomposition of $(P^{[\Lambda]}) : \Delta_{[\Lambda]}^\infty$ as an intersection of radical differential ideals $\mathcal{I}_i$, with characteristic sets $\mathcal{B}_i$ for $\prec$ such that $\mathcal{I}_i = (\mathcal{B}_i) : H_{\mathcal{B}_i}^\infty$ (see [5]). By Lazard's lemma, the $\mathcal{I}_i$ contain no polynomial that do not depend of the derivatives $x_{j_i}^{(k)}$, $1 \le i \le s$ $\mu_i + \nu_{j_i} \le k \le \mu_0 + \nu_{j_i} + r$, so that these are precisely the leading derivatives of the polynomials in $\mathcal{B}_i$. As Δ_s does not vanish on $\mathcal{I}_i$, any generic zero of a component of $\mathcal{I}_i$ may be completed into a generic zero of $(P^{[r]}) : \Delta_{[r]}^\infty$ for any $r > \Lambda$. So, $(P^{[r]}) : \Delta_{[r]}^\infty \cap \mathcal{F}\{x\}_{[\Lambda]} = (P^{[\Lambda]}) : \Delta_{[\Lambda]}^\infty$, which implies that $[\mathcal{B}_i] : H_{\mathcal{B}_i}^\infty \cap \mathcal{F}\{x\}_{[\Lambda]} = (\mathcal{B}_i) : H_{\mathcal{B}_i}^\infty$. We may now extract from $\mathcal{B}_i$ a minimal autoreduced set $\mathcal{A}_i$. All the elements of $[\mathcal{B}_i] : H_{\mathcal{B}_i}^\infty$ are reduced to 0 by $\mathcal{A}_i$, and $[\mathcal{A}_i] : H_{\mathcal{A}_i}^\infty \subset [\mathcal{B}_i] : H_{\mathcal{B}_i}^\infty$, so that $\mathcal{A}_i$ is a characteristic set of $[\mathcal{A}_i] : H_{\mathcal{A}_i}^\infty = [\mathcal{B}_i] : H_{\mathcal{B}_i}^\infty$.

We conclude the proof by considering the union of all such decompositions, associated to all the possible sequences $\Delta^{[r]}$, constructed as in the proof of i), that do not vanish on some component $\mathcal{P}$ of $\mathcal{Q}$.

iii) As the main derivatives of the polynomials in $\mathcal{B}_h$ are the $x_{j_i}^{(k)}$, $1 \le i \le s$ $\mu_i + \nu_{j_i} \le k \le \mu_0 + \nu_{j_i} + r$, the main derivatives of the polynomials in $\mathcal{A}_h$ must be $x_{j_i}^{(\mu_i+\nu_{ij})}$, so that $\mathrm{ord}^\mathcal{J} A_i = \mu_i$.

iv) Here α, β denotes the canonical cover. Assume that $\mathcal{A} = \{A_1, \ldots, A_s\}$ and that all the A_i are zero divisors modulo $(P_i^{(k)}|1 \leq i \leq s, 0 \leq k \leq \ell_i) : (\nabla)^\infty$. Then, using prop. 81, there exists $\Delta \in \nabla$ such that $\mathcal{A}$ may be completed to form a characteristic set $\mathcal{B}$ of a prime component of the radical of the algebraic ideal $(P_i^{(k)}|1 \leq i \leq s, 0 \leq k \leq \ell_i) : (\Delta)^\infty$, for some ordering.

For all $1 \leq p \leq s$, there exists a polynomial S_p and an integer n_p such that $A_p S \Delta^{n_p} = \sum_{i=1}^{s} \sum_{k=0}^{\ell_i} N_{p,i,k} P_i^{(k)}$. Without loss of generality, we may assume that the ℓ_i are the minimal integers such that expressions of that kind exist, so that some elements of $\mathcal{A}$ actually depend of the $P_q^{(\ell_q)}$, such that the $\ell_q - \lambda_q$ are maximal, equal to r_0. Let A_1, ..., A_{s_0} be these elements.

The leading derivatives of these A_i must be some $x_j^{(\Lambda + r_0 + \beta_j)}$. First, we may show that some of these derivatives do appear in them. If not, we could chose $\mathcal{B}$ according to an ordering on the derivatives[13] such that these derivatives are greater than those present in $\mathcal{A}$. Cor. 72 implies that elements of $\mathcal{A}$ appear in a characteristic set of a component of $(P_i^{(k)}|1 \leq i \leq s, 0 \leq k \leq \ell'_i) : (\Delta)^\infty$, with $\ell'_\mu - \lambda_\mu < r$, which contradicts the minimality of ℓ. Assume now that the leading derivatives of A_1, ..., A_{s_0} are not among the $x_j^{(\Lambda + r_0 + \beta_j)}$. Then, these derivatives must be strict derivatives of the derivatives of the remaining elements $A_{s_0+1}, \ldots A_s$, so that they cannot appear in A_1, ..., A_{s_0}. A contradiction.

We may now assume than the leading derivative $v = x_{j_0}^{(\Lambda + r_0 + \beta_{j_0})}$ of A_1 is smaller than those of A_2, ..., A_{s_0}. So, v must be the only derivative among the $x_j^{(\Lambda + r_0 + \beta_j)}$ that do appear in A_1. But this is impossible, as the matrix obtained by suppressing row i_0 and column j_0 in $\mathcal{J}_P$ must have full rank modulo $\mathcal{P}$. ■

With more work, one should be able to prove when $s = n$ a similar result by only assuming that the square submatrices of $\mathcal{J}_P$ that possess non zero diagonal elements have a non zero determinant.

Examples. — 85) Consider the system $x_1^{(5)} + x_2'' + x_3''' = 0$, $x_2' = 0$, $x_1''' - x_3' = 0$. We have $\lambda = (0, 1, 2)$, $\alpha = (2, 1, 0)$ and $\beta = (3, 0, 1)$. We have two possible classes of characteristic sets that may be computed using the shortest reduction, *viz.* by derivating the second equation 1 time and the second 2 times: $\mathcal{A}_1 := \{x_1^{(5)}, x_2'', x_1''' - x_3'\}$ and $\mathcal{A}_1 := \{x_3''', x_2'', x_1''' - x_3'\}$.

86) The system $x_1^{(5)} + x_2'' + x_3''' = 0$, $x_2' + x_3'' = 0$, $x_3' = 0$ admits a single class of characteristic sets for the shortest reduction: $\mathcal{A} := \{x_1^{(5)}, x_2'', x_3'\}$. However, if we suppress the row 3 and the column 1 in $\mathcal{J}$, we get a matrix that is not of full rank, so that the condition of the theorem iv) is not satisfied. It is easily seen that, in order to compute $\mathcal{A}$, it is enough to differentiate the last equation $1 < \lambda_3 = 2$ time.

[13] This ordering does not need to be compatible with the derivation, as we consider here an *algebraic* ideal.

87) The system $x_1^{(5)} + x_2'' + x_3''' = 0$, $x_3'' = 0$, $x_2' + x_3' = 0$ admits a single class of char. set for Jacobi orderings, that may be computed using the shortest reduction, *viz.* by differentiating the second and the third equations only 1 time. It is represented by $\mathcal{A} := \{x_1^{(5)}, x_3'', x_2' + x_3'\}$. However, with the same derivatives, we may also compute the following characteristic set, that does not correspond to a Jacobi ordering, but to an Egerváry ordering: $\mathcal{B} := \{x_1^{(5)}, x_2'', x_3' + x_2'\}$. By chance, it may be computed with fewer derivatives than predicted by the bound.

Remark 88. — Jacobi [37, end of § 3 p. 58] claims that the number of possible normal forms of a system that one may find by the shortest reduction, is equal to the number of monomials in the truncated determinant, or equivalently to the number of transversal maximal sums in the order matrix. The last example has already produced a contradiction.

Restricting ourselves to normal forms, or classes of characteristic sets, associated to Jacobi orderings does not solve the problem. It is also easily seen that the number of normal forms may be smaller than $n!$ for systems such as $x + y + z = 0$, $x' + y' + 2z' = 0$, $x'' - y'' + z'' = 0$, for which all 6 possible monomials appear in ∇, only has 4 different normal forms: $x = -y - z, z' = 0, y'' = 0$; $y = -x - z, z' = 0, x'' = 0$; $z = -x - y, x' = -y', y'' = 0$ and $z = -x - y, y' = -x', x'' = 0$. Furthermore, a system such as $x + y = 0$, $x' = 0$ has only a single monomial in ∇ but two normal forms: $x = y, y' = 0, y = x, x' = 0$, for the Egerváry ordering associated to the canon $\lambda_1 = 1$, $\lambda_2 = 0$

The best bound I could find on the possible number of normal forms for Jacobi orderings is the following. But first, we need a new definition.

Definition 89. — *Let $\mathcal{A} = \{A_i | i \le i \le p$ be a characteristic set of a prime differential ideal in $\mathcal{F}\{x\}$. We assume that a reduction process using $\mathcal{A}$ has been chosen and recursively denote by $\widetilde{A}_i^{(k)}$ the reduction of $A_i^{(k)}$ by the $\widetilde{A}_{i'}^{(k')}$ with $i' \neq i$ or $k' < k$.*

Any finite subset $\mathcal{B}$ of $\{\widetilde{A}_i^{(k)} | 1 \le i \le p, k \in \mathbf{N}\}$ is the chacteristic set of the prime algebraic ideal $(cB) : H_{\mathcal{B}}^{\infty}$.

Proposition 90. — *With the notations of the theorem, let $\mathcal{P}$ be a prime component of $\mathcal{Q}$. Assume that the polynomial P_i are listed by decreasing ℓ_i. For any injection $\sigma : [1, s] \mapsto [1, n]$, the three following propositions are equivalent:*

i) there exists a characteristic set $\mathcal{A} = \{A_i | 1 \le i \le s\}$ of $\mathcal{P}$, such that the leading derivative of A_i is $x_{\sigma(i)}^{(\alpha_i + \beta_{\sigma(i)})}$;

ii) there exists a characteristic set $\mathcal{A} = \{A_i | 1 \le i \le s\}$ of $\mathcal{P}$, such that, denoting by $(i_h)_{0 \le h \le p}$ the increasing sequence such that $i_0 = 0$, $i_p = s$ and the remaining i_h are the integers satisfying $\ell_{i_h} < \ell_{i_h+1}$, the set $\{\widetilde{A}_i^{(k)} | 1 \le i \le i_h, 0 \le k \le \ell_i - \ell_{i_h}\}$ is a characteristic set of a prime component of the algebraic ideal $(P_i^{(k)} | 1 \le i \le i_h, 0 \le k \le \ell_i - \ell_{i_h}) : (\nabla)^{\infty}$;

iii) for all $1 \leq i \leq s$*, the determinant of the minor contained in the rows* 1 *to* i *and columns* $\sigma(1)$*, ...,* $\sigma(i)$ *of* $\mathcal{J}_P$ *is non zero modulo* $\mathcal{P}$*.*

Proof. — i)⇒ii). Up to a change of ordering, we may assume $\mathcal{A}$ to be such that A_i does not contain the derivatives $x_{\sigma(i')}^{(\alpha_{i'}+\beta_{\sigma(i')})}$ for $i' > i$. With this assumption, we shall prove, by induction on h, that the set $\mathcal{A}$ of i) also satisfies ii). The property is straightforward for $h = 0$.

Assume the result is true for $h' < h$. The system $P_{[h]} := \{P_i^{(k)} | 1 \leq i \leq i_h, 0 \leq k \leq \ell_i - \ell_{i_h}\}$ must be reduced to 0 by $\mathcal{A}$. If we reduce a element of $P_{[h]}$ using only the set $\mathcal{B}_{[h]} := \{tA_i^{(k)} | 1 \leq i \leq i_h, 0 \leq k \leq \ell_i - \ell_{i_h}\}$, we find a rest R. If R is not zero, the induction hypothesis implies that the reduction by $\{tA_i^{(k)} | 1 \leq i \leq i_{h-1}, 0 \leq k \leq \ell_i - \ell_{i_0}\}$ does not depend of derivatives of each x_j greater than $x_j^{(\alpha_{i_h}+\beta_j)}$. The reduction by $\{A_i | i_{h-1} < i \leq i_h\}$ cannot introduce derivatives of the $x_{\sigma(i)}$, $i > i_h$ being not less than $\alpha_i + \beta_{\sigma(i_h)}$, according to our assumption on $\mathcal{A}$. So, R is irreducible by $\mathcal{A}$, which is impossible.

By [71] § 17–19 p. 88–90, $\mathcal{B}_{[h]}$ is a characteristic set of a prime ideal that contains $P_{[h]}$. As $\mathcal{P}$ is a prime component of $\mathcal{Q}$, $\mathcal{B}_{[h]}$ must be the characteristic set of a prime component of $(P_{[h]}) : (\nabla)^\infty$, and so exactly of codimension equal to $\sharp\, P_{[h]} = \sharp\, \mathcal{B}_{[h]}$. So, $\mathcal{B}_{[h]}$ must be the char. set of a prime component of $(P_{[h]}) : (\nabla)^\infty$.

ii)⇒iii). As $\{dA_k | 1 \leq k \leq i\}$ and $\{\mathrm{d}P_k | 1 \leq k \leq i\}$ generate the same differential vector space, the determinant of one of these minor does not vanish modulo $\mathcal{P}$ if the corresponding minor of $\mathcal{J}_A$ does not, which is straightforward.

iii)⇒i). It is a consequence of the theorem. ∎

From a combinatorial standpoint, it means that the maximal number of such characteristic sets is, for a generic system, at most the number of permutations σ such that for all $1 \leq i \leq s$, the minor matrix contained in the i first lines and the rows $\sigma(1)$, ..., $\sigma(i)$ of $A_P + \ell$ possesses a set of i transversal maxima—but the set for i need not be included in that for $i + 1$!

The first difficulty in computing characteristic sets for non linear differential systems is to differentiate the equations. Using a classical representation of data, the sizes of the successive derivatives are exponential in the order and it is well possible to saturate the available memory before starting any actual elimination.

As this method reduces the number of requested differentiations to the minimum, it suggests many computational applications and easily implemented improvement of existing softwares such as Diffalg[4, 5].

8 The various normal forms of a system

Jacobi considers in [II/13 b)] [36, p. 9–14] and [II/23 a) f° 2217 seq.] (*cf.* [37, p. 37–43] the various normal forms that a given system may possess. Systems in normal form include those of the form:

$$x_i^{(\alpha_i} = f_i(x), \quad 1 \le i \le n,$$

with $\mathrm{ord}_{x_j} f_i < \alpha_j\, forall 1 \le i, j \le n$, which he calls *explicit normal form*. But Jacobi also includes in this category systems $A_i(x) = 0$, with $\mathrm{ord}_{x_j} A_i = \alpha_i$ and such that $|\partial A_i / \partial x_j^{(\alpha_j)}| \neq 0$, such as our characteristic sets enter in this category.

Jacobi claims that, if one cannot reduce a system to an equivalent one, with fewer equations than variables, that is, in our language, if the differential dimension is zero, then one can eliminate all dependent variables, except one, and get an equation of which the order is the order of the system. This is only generically true, and Jacobi was aware of it, for in [II/23 a), fo 2217, note], he introduces the order in some different way, claiming that the reduction to a simple equation was sometimes impossible, e.g. if each equation A_i depends only of x_i.

The order does not depend of the chosen explicit normal form and is equal to $\sum_{i=1}^n \alpha_i$. If we associate to the system a prime differential ideal $\mathcal{P}$, the order is the *algebraic* transcendence degree of the associated differential field extension $\mathcal{G}/\mathcal{F}$. At the time of Jacobi it was refered to as the *number of arbitrary constants appearing in a complete integration*, constants that could be, e.g., initial conditions. Jacobi claims that, in the generic case, the orders of the leading derivatives in a normal form may be arbitrarilly chosen, provided that their sum is equal to the order of the system.

Then, he considers systems possessing fewer possible normal forms, starting with the example of two equations in two variables.

Lemma 91. — *Let* $\{A_1, A_2\}$ *be a characteristic set of a prime differential ideal* $\mathcal{P} \in \mathcal{F}\{x_1, x_2\}$*, such that the main derivative of* A_i *is* x_i *and* $a_{i,j} := \mathrm{ord}_{x_j} A_i$*.*

i) If x_1 *appears in* A_2*,*

a) there exists a new characteristic set $\mathcal{B}$ *of* $\mathcal{P}$ *with* $\mathrm{ord}_{x_1} B_1 = a_{2,1}$*, and* $\mathrm{ord}_{x_2} B_2 = a_{2,2} + a_{1,1} - a_{2,1}$*;*

b) there is no characteristic set $\mathcal{B}$ *of* $\mathcal{P}$ *with* $a_{1,1} > \mathrm{ord}_{x_1} B_1 > a_{2,1}$*.*

ii) If x_1 *(resp.* x_2*) does not appear in* A_2 *(resp.* A_1*), then there exists no characteristic set* $\mathcal{B}$ *of* $\mathcal{P}$ *with* $\beta_1 < \alpha_1$ *(resp.* $\beta_2 < \alpha_2$*).*

Proof. — In this proof, one may consider by lemma 70 i′) that the system $A = 0$ is linear.

i) a) Consider a new order $\prec$ such that $x_1^{(a_{2,1})} \succ x_2^{(a_{2,2})}$ and $x_1^{(a_{2,1}-1)} \prec x_2^{(a_{2,2})}$. Then we may take $B_1 = A_2$ and for B_2 the reduction of A_1 by A_2, which depends on $A_2^{(a_{1,1}-a_{2,1})}$, so that it has order $a_{2,2} + a_{1,1} - a_{2,1}$ in x_2.

b) In such a case, B_1 is irreducible by A_1, we may reduce it using derivatives of A_2 of order at most $\mathrm{ord}_{x_2}B_1 - \mathrm{ord}_{x_2}A_2 < \mathrm{ord}_{x_1}A_1 - \mathrm{ord}_{x_1}B_1 < \mathrm{ord}_{x_1}A_1 - \mathrm{ord}_{x_1}A_2$. So, the order of the rest in x_1 is less than $\mathrm{ord}_{x_1}A_1$: the rest is unreducible and must be 0. This implies that B_1 depends only on A_2 and strict derivatives of A_2, so that A_2 cannot be expressed as a linear combinnation of B_1 and B_2, which is impossible.

ii) The proof is similar to i) b). ∎

Jacobi comes then to the case of an arbitrary number of variables and considers the problem of increasing the order of m variables x_1, ..., x_m in a normal form, when decreasing the order of variables x_{m+1}, ..., x_{2m}, the orders of the remaining variables staying unchanged. He gives then the following result.

Lemma 92. — *Let* $\mathcal{A} = \{A_1, \ldots, A_n\}$ *be a characteristic set of a prime differential ideal* $\mathcal{P} \subset \mathcal{F}\{x_1, \ldots, x_n\}$, *let* $x_i^{(\alpha_i)}$ *be the leading derivative of* A_i. *Let* $x_i^{(\beta_i)}$ *be the highest derivative of* x_i *appearing in the equations* A_{m+1}, ..., A_{2m}.

If $|\partial A_{m+i}/\partial x_j^{(\beta_j)}; 1 \le i,j \le m| \notin \mathcal{P}$, *then there exists a characteristic set* $\mathcal{B}$ *of* $\mathcal{P}$ *such that*

i) for $1 \le j \le m$, $\mathrm{ord}_{x_j}\mathcal{B} = \beta_j < \alpha_j$;

ii) for $m < j \le 2m$, $\mathrm{ord}_{x_j}\mathcal{B} > \alpha_j$;

iii) for $2m < j \le n$, $\mathrm{ord}_{x_j}\mathcal{B} = \alpha_j$;

Proof. — It is enough to chose an ordering $\prec$ on derivatives, such that $x_j^{(\beta_j)} \succ x_i^{(\alpha_i)}$, $1 \le j \le m$, $m < i \le 2m$; $x_j^{(\beta_j-1)} \prec x_i^{(\alpha_i)}$, $1 \le j \le m$, $m < i \le n$; $x_j^{(\alpha_j)} \succ x_j^{(\alpha_i)}$ and $x_j^{(\alpha_j-1)} \prec x_j^{(\alpha_i)}$, $2m < j \le n$, $1 \le i \le 2m$. ∎

What happens if $|\partial A_{m+i}/\partial x_j^{(\beta_j)}; 1 \le i,j \le m| \notin \mathcal{P}$? Jacobi concludes with these words: *"Such questions require then a deeper investigation, that I will expose in some other occasion"*. One may guess that Jacobi was thinking of applying his method for computing normal forms. So, we will return to this problem in the next section 9.

It must be noticed that the requested transformation may be performed, even in the case when $\{A_{m+1}, \ldots, A_{2m}\}$, considered as a system in x_1, ..., x_m alone, does not generate a differential ideal of dimension 0, as in the following example:

Examples. — 93) Consider the explicit normal system of 4 equations in 4 variables

$$x_1''' = x_2'', x_2''' = 0, x_3' = x_1'', x_4 = x_1''.$$

If one wishes to decrease the orders of x_1 and x_2 and to increase that of x_3 and x_4, we cannot use the preceding lemma, nor any generalization of it, for the 2 last equations do not depend of x_2. However, we can achieve our goal with the following normal form:

$$x_1' = x_4, x_2'' = x_4', x_3'' = 0, x_4' = x_3'.$$

The next example shows that one can decrease the order of 2 variables, when increasing the order of a single one.

94) Consider the system:

$$x_1'' = x_2, x_2' = 0, x_3 = x_1',$$

it is possible to decrease the order of x_1 and x_2 in the following normal form:

$$x_1' = x_3, x_2 = x_3', x_3'' = 0.$$

Testing the existence of a characteristic set $\mathcal{B}$ with leading derivatives $x_j^{(\beta_j)}$, for given β_j, some characteristic set $\mathcal{A}$ being known will be the subject of the next section.

We denote by $\mathrm{orders}\mathcal{A}$ the n-uple (α_1, α_n), where $\alpha_j = \mathrm{ord}_{x_j} A_j$, assuming that the main derivative of A_j is a derivative of x_j. Let $\mathcal{P}$ be a prime ideal, we denote by $\mathrm{orders}\mathcal{P}$ the set $\{\mathrm{orders}\,\mathcal{A} | \mathcal{A}$ a char. set of $\mathcal{P}\}$. We will conclude our investigations with a description of the possible values of $\mathrm{orders}\mathcal{P}$, for a prime differential ideal $\mathcal{P}$ of diff. dim. 0 and order e. In two variables, it can take any value compatible with the order, as shows the next proposition.

Proposition 95. — *Let $I \subset [0, e]$, there exists a prime differential ideal $\mathcal{P}$ such that $\mathrm{orders}\mathcal{P} = \{(a, e-a) | a \in I\}$.*

Proof. — Let $\alpha^1 < \alpha^2 < \cdots < \alpha^r$ be the elements of I. Define recursively

$$\begin{array}{rclrcl} {}_1^1A_1 & = & x_1^{(\alpha^1)}, & {}_1^1A_2 & = & 0 \\ {}_1^{i+1}A_1 & = & {}_1^iA_2 + {}_1^iA_1^{(\alpha^{i+1}-\alpha^i)}, & {}_1^{i+1}A_2 & = & {}_1^iA_1, \\ {}_2^rA_2 & = & x_2^{(e-\alpha^r)}, & {}_2^rA_1 & = & 0 \\ {}_2^{i-1}A_2 & = & {}_2^iA_1 - {}_2^iA_2^{(\alpha^i-\alpha^{i-1})}, & {}_2^{i-1}A_1 & = & {}_2^iA_2. \end{array}$$

Let then for $j = 1, 2$ ${}^iA_j = {}_1^rA_j + {}_2^rA_j$ and $\mathcal{A}_i = \{{}^iA_2, {}^iA_2\}$. By construction, for $1 \le i \le r$, the $\mathcal{A}_i$ are characteristic sets of the same prime differential ideal $\mathcal{P}$ and by lemma 91, $\mathrm{orders}\mathcal{P} = \{(a, e-a) | a \in I\}$. ∎

For a greater number of variables, the situation is more complicated... *Tam quaestiones altioris indaginis poscuntur.*

If one try to visualize the set of possible characteristic set for a given system in 3 variables, it is convenient to use triangular coordinates, as the sum of the 3 maximal orders in the 3 variables is constant.

Examples. — 96) If one considers the system

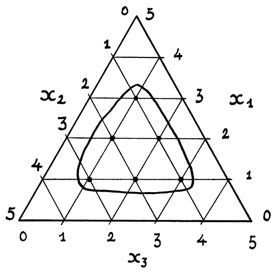
$$\begin{aligned} x_1' &= x_2'', \\ x_2''' &= x_2', \\ x_3' &= x_2'', \end{aligned}$$

easy computations show that it admits 6 normal forms.

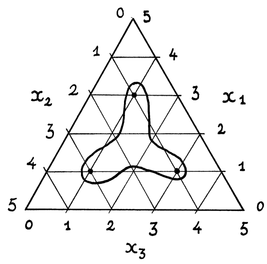
97) For the sy﬩em $x_1' - x_2' = 0$, $x_2''' = 0$, $x_3' - x_2' = 0$, only 3 normal forms exi﬩.

$$\begin{array}{llllll} x_1''' & = 0, & x_1' & = x_2', & x_1' & = x_2', \\ x_2' & = x_1', & x_2''' & = 0, & x_2' & = x_3', \\ x_3' & = x_2'; & x_3' & = x_2'; & x_3''' & = 0. \end{array}$$

The two examples may be illu﬩rated by such drawings, where the points corresponding to exi﬩ing normal forms are surrounded by a loop.

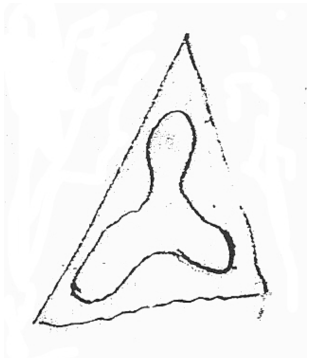


Example 96



Example 97

Those drawings look very much like these ones, that appear on the margin of [II 13 b), fo 2206a].



We conclude this subse﬩ion with a proposition, showing that for more than two variables, the set $\mathrm{orders}\mathcal{P}$ cannot be arbitrary.

Proposition 98. — *Let $\mathcal{P} \subset \mathcal{F}\{x_1, x_2, x_3\}$ be a prime differential ideal and $\mathcal{A}$ a chara﬩eri﬩ic set of $\mathcal{P}$. Assume that there exi﬩ $\alpha_1 > \beta > \gamma$ such that* $\mathrm{orders}\mathcal{P}$ *contains $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta, \alpha_2 + \alpha_1 - \beta, \alpha_3)$ and $(\gamma, \alpha_2, \alpha_3 + \alpha_1 - \gamma)$ and no element $(\delta_1, \delta_2, \alpha_1)$ with $\alpha_1 > \delta_1 > \beta$ or $(\delta_1, \alpha_2, \delta_3)$ with $\alpha_1 > \delta_1 > \gamma$), then* $\mathrm{orders}\mathcal{P}$ *contains $(\gamma, \alpha_2 + \alpha_1 - \beta, \alpha_3 + \beta - \gamma)$.*

Proof. — In such a case, we have $\mathrm{ord}_{x_1} A_1 = \alpha_1$, $\mathrm{ord}_{x_2} A_2 = \alpha_2$, $\mathrm{ord}_{x_3} A_3 = \alpha_3$ and $\alpha_1 > \beta := \mathrm{ord}_{x_1} A_2 > \gamma := \mathrm{ord}_{x_1} A_3$. Then, we get a chara﬩eri﬩ic set $\mathcal{B}$ with $\mathrm{ord}\mathcal{B} = (\beta, \alpha_2 + \alpha_1 - \beta, \alpha_3)$ by reducing A_1 with A_2, A_3 remaining unchanged. So $\mathrm{ord}_{x_1} B_3 = \gamma$, hence the result. ∎

9 Change of orderings

Change of orderings on monomials for standard bases computation (FGLM) [24], or derivatives (Pardi) [4] for characteristic set computations have been considered in the computer algebra litterature. It may be noticed that the main theoretical works of the xx$^{\text{th}}$ century often restrict to particular orderings, Janet orderings (Janet), elimination orderings (Ritt), but for many applications, one need to use particular orderings, e.g. testing identifiability or observability in control theory requires to eliminate precise sets of indeterminates. In [II/23 a) fo 2217–2220] [37, p. 36–43], Jacobi considers, in full generality, the problem of computing a normal form of an ordinary differential system, some normal form being known for a different ordering. The method he gives is quite similar to the tools of contemporary litterature and he provides moreover sharp bounds on the requested number of of derivations, that may be used to improve the efficiency of our algorithms.

Considering a system in explicit normal form $x_i^{(e_i)} = F_i(x)$, $1 \le i \le n$, the problem is to compute a new normal form of the system $x_i^{(f_i)} = G_i(x)$. In a first step Jacobi divides the indeterminates in three sets. For $i \in I_1$, $f_i > e_i$; for $i \in I_2$, $f_i < e_i$ and for $i \in I_3$, $f_i = e_i$.

Using the derivation

$$\delta := \sum_{j=1}^{n} \left(F_i(x) \frac{\partial}{\partial x_i^{(e_i-1)}} + \sum_{k=0}^{e_i-2} x_j^{(e_k+1)} \frac{\partial}{\partial x_i^{(k)}} \right),$$

Jacobi claims that it is possible to compute the new normal forms using the first ones, completed with the equations $x_i^{(e_i+k)} = \delta^k F_i(x)$, $i \in I_1$, $1 \le k \le f_i - e_i$ and that the new normal form exists iff

$$\left| \frac{\partial \delta^k F_i}{\partial x_j^{(\alpha)}} | i \in I_1,\ 0 \le k \le f_i - e_i; \quad j \in I_2,\ f_j \le \alpha < e_j \right| \ne 0.$$

The following theorem translates this result, in the framework of differential algebra, using the $\widetilde{A}_i^{(k)}$ in the place of $\delta^k \left(x_i^{(e_i)} - F_i(x) \right)$ (see def. 89). For the sake of simplicity, we restrict here to the case of prime ideals. In more general situations, splitting may occur that may be considered *à la D5* [18]...

Theorem 99. — *Let $\mathcal{P}$ be a prime ideal of differential dimension* 0 *of* $\mathcal{F}\{x_1, \ldots, x_n\}$ *and* $\mathcal{A} = \{A_1, \ldots, A_n\}$ *a characteristic set of $\mathcal{P}$ for some ordering, such that the main derivative of A_i is $x_i^{(e_i)}$.*

i) Assume that there exists a characteristic set $\mathcal{B} = \{B_1, \ldots, B_n\}$ of $\mathcal{P}$, being such that the main derivative of B_i is $x_i^{(f_i)}$ and that we have $f_i > e_i$ for $i \in I_1$, $f_i < I_2$ for $i \in I_2$ and $f_i = e_i$ for $i \in I_3 = [1, n] \setminus I_1 \cap I_2$. Then, *a)* $\mathcal{B} \subset (\widetilde{A}_i^{(k)} | i \notin I_2,\ 0 \leq k \leq f_i - e_i) : H_{\mathcal{A}}^{\infty}$, *more precisely,* $\{B_i | i \in I_2\} \subset (\widetilde{A}_i^{(k)} | i \in I_1,\ 0 \leq k \leq f_i - e_i) : H_{\mathcal{A}}^{\infty}$;

b) if for some $i_0 \in I_1$ $\ell_{i_0} < f_{i_0} - e_{i_0}$, *then* $B_{i_0} \notin (\widetilde{A}_i^{(k)} | 1 \leq i \leq s,\ 0 \leq k \leq \ell_i) : H_{\mathcal{A}}^{\infty}$.

ii) A characteristic set $\mathcal{B}$ *satisfying the hypotheses of i) does exist iff*

$$\left| \frac{\partial \widetilde{A}_i^{(k)}}{\partial x_j^{(\alpha)}} | i \in I_1,\ 0 \leq k \leq f_i - e_i; \quad j \in I_2,\ f_j \leq \alpha < e_j \right| \notin \mathcal{P}.$$

Proof. — i) The char. set $\mathcal{B}$ cannot contain polynomials involving derivatives of each x_i of order higher than the f_i^{th}. If $\mathcal{B}$ exists, it must be included in $(\widetilde{A}_i^{(\alpha(k))} | 1 \leq i \leq n,\ 0 \leq k \leq \alpha_i)$, a) is a consequence of 72. If for some $i_0 \in I_1$ $\ell_{i_0} < f_{i_0} - e_{i_0}$, then $x_{i_0}^{(f_0)}$ does not appear in the generators of $(\widetilde{A}_i^{(k)} | 1 \leq i \leq s,\ 0 \leq k \leq \ell_i)$: if B_{i_0} were in $(\widetilde{A}_i^{(k)} | 1 \leq i \leq s,\ 0 \leq k \leq \ell_i) : H_{\mathcal{A}}^{\infty}$, $\partial B_{i_0} / \partial x_{i_0}^{(f_0)}$ would be in $\mathcal{P}$, which is impossible as $\mathcal{B}$ is the char. set of a prime ideal. This proves b).

ii) Using lemma 70 iv), the problem is reduced to the existence of a standard basis for $(\mathrm{d}P)_{\mathcal{M}_{\mathcal{P}}}$ with main derivatives $\mathrm{d}x_i^{(f_i)}$. The non vanishing of the determinant implies that an autoreduced set with the requested main derivatives exists, that must be a standard basis by invariance of the order. Assume reciprocally that such a standard basis exists. Lemma 70 iii) implies that $(\mathrm{d}P)_{\mathcal{M}_{\mathcal{P}}} \cap \langle \mathrm{d}x_i^{(k)} | 0 \leq k < \max(e_i, f_i) \rangle = \langle \mathrm{d}\widetilde{A}_i^{(k)} | i \in I_1,\ 0 \leq k < f_i - e_i \rangle$. Defining the $\widetilde{B}_i^{(k)}$ as in def. 89, we see that for $i \in I_2$ and $0 \leq k < e_i - f_i$, $\widetilde{B}_i^{(k)} \in (\mathrm{d}P)_{\mathcal{M}_{\mathcal{P}}} \cap \langle \mathrm{d}x_i^{(k)} | 0 \leq k < \max(e_i, f_i) \rangle = \langle \mathrm{d}\widetilde{A}_i^{(k)} | i \in I_1,\ 0 \leq k < f_i - e_i \rangle$, so that the determinant cannot vanish in $\mathcal{G}$. ∎

Jacobi did not stop his investigations at this step. Claiming that it was sometimes more efficient to use derivatives of the A_i instead of the $\widetilde{A}_i^{(k)}$ obtained by substitutions. This strongly suggests a practical experience of computing changes of ordering, although no explicit example is found in his manuscripts. We have already noticed that derivation, introducing new derivatives, produces an exponential growth of the equations in dense representation. The situation becames worse if substitutions are done at the same time, for then the degree will increase too. The best known bounds for the required eliminations imply to use Bézout's theorem, and the degrees will be the smallest using the A_i instead of the $\widetilde{A}_i^{(k)}$... We see that Jacobi's intuition of the complexity issues meets here again contemporary research, such as D'Alfonso *et al.* [14, 15, 17, 16] in the spirit of the Kronecker algorithm [28].

This problem is considered in § 18 of [II/23 a)] [37, p. 40–43]. The end of this manuscript seems lost, as the sentence at the end of fo 2220 remains unachieved, but we can understand the general idea.

With the notations and hypotheses of th. 99, one needs to differentiate equation A_i $f_i - e_i$ times if $i \in I_1$. Let $\mathrm{ord}_{x_j} A_i := a_{i,j}$, then, generically, A_i, $1 \le i \le n$, must be differentiated ℓ_i to compute the derivatives $\widetilde{A}_i$, $i \in I_1$, with ℓ_i such that: $ell_i \ge f_i - e_i$ for $i \in I_1$ and $a_{j,j} + \ell_j \ge \max_i a_{i,j} + \ell_i$, so that the necessary reductions could be performed. The minimal solution of this problem is obtained by computing the minimal canon of the matrix $a_{i,j} + \max(f_i - e_i, 0)$, using the methods of subsection 3.6.

Remarks. — 100) Using explicit normal forms, we may assume that the leading derivatives $x_i^{(e_i)}$ do not appear in the right members $F_i(x)$. It is no longer the case with characteristic sets. All we know is that the leading derivative of A_i may only appear in A_j with a strictly smaller degree. But we may, without changing the main derivatives, assume that the A_i for $i \in I_1$ do not depend on the main derivatives of the x_j for $j \notin I_1$. With this assumption, the bound in the characteristic set setting coincides with Jacobi's one.

10 Resolvants

For a modern treatment of the question, one may refer to Cluzeau and Hubert [10].

Assume that $\mathcal{P}$ is a prime regular component of $[P] : \nabla^\infty$. One may define resolvants following Ritt [71, chap. II § 22].

Definition 101. — *We call a* resolvant *of $\mathcal{P}$ the data of two differential polynomials R and S together with a characteristic set $\mathcal{A}$ of the prime differential ideal $[\mathcal{P}, Sw - R] : S^\infty$ (in $\mathcal{F}\{x_1, \ldots, x_n, w\}$), such that $v_{A_i} = x_i$, $1 \le i \le n$.*

As $\mathcal{P}$ is regular, we know that its order is Jacobi's bound $\mathcal{O}$, so that $v_{A_{n+1}} = w^{(\mathcal{O})}$.

Jacobi [37, § 4] assume that a resolvent exists when choosing $w = x_{j_0}$ and proposes to compute the order up to which one needs to differentiate each equation P_i to be able to compute the resolvent. The word resolvent was not used by Jacobi, but he evocates the notion as something well known in the the mathematical folklore of his time: *"It is usual that this type of normal forms be considered before others by mathematicians".*

One may summarize his findings with the following theorem.

Theorem 102. — *Assume that $w = x_{j_0}$ and that $\mathcal{A}$ is a resolvent for a prime component $\mathcal{P} := [P] : T^\infty$ of differential dimension 0, then $\mathcal{A} \in [P_i^{(k)} | 0 \le k \le \mathcal{O}_{i,j_0}] : T^\infty$, where $\mathcal{O}_{i,j_0}$ is the tropical determinant of the matrix $(a_{\hat{\imath},\hat{\jmath}} | \hat{\imath} \neq i, \hat{\jmath} \neq j_0)$.*

Proof. — Replace x_{j_0} in P_i by w_i; we obtain a new system $\tilde{P}_i$. Then one may fix w_i for $i \neq i_0$ and the system $\tilde{P}_i$, $i \neq i_0$ in the variables x_j, $j \neq j_0$ has order at most $\mathcal{O}_{i,j_0}$. Eliminating the x_j, one finds an equation of order at most $\mathcal{O}_{i,j_0}$ in w_{j_0} that depends only of the w_j. Replacing then all the w_j by w, one gets a non zero polynomial that obviously belongs to the requested algebraic ideal; if it were 0, the equations P_i would not be independent, and $\mathcal{P}$ would not be of differential dimenson 0. ∎

Remark 103. —] The heuristic way in which Jacobi presents the result his interesting. He claims that one should differentiate P_{j_0} up to order $\mathcal{O}_P$, and then computes a minimal canon such that $\ell_{j_0} = \mathcal{O}_P$, in the spirit of subsection 3.6. This means that we can efficiently compute the $\mathcal{O}_{i,j_0}$, knowing a canon for the order matrix of P by using a shortest path algorithm, as explained in section 3.7[14].

Conclusion

We have seen that the corpus of results contained in Jacobi's posthumous manuscripts provides a large set of applicable methods for the resolution of ordinary differential systems. From the automatization of easy ideas, such as looking for block decompositions to more sophisticated tools, allowing to produce simpler normal form reductions or better ways to perform change of ordering, they can improve in many ways the existent computer algebra algorithms.

In all cases, Jacobi's bound by itself is able to replace advantageously Ritt's analog of the Bézout bound, in all situations where is is proved, *i.e.* at this time quasi-regular components or systems of two equations [71, Ch. VII § 6 p. 136]. Its interest to produce sharper complexity bounds, upper and lower, is obvious.

Using these tools as widely as possible is a promizing task, and generalizing them to arbitrary systems a challenging goal. Let us insist on the fact that this paper has no claim to exhaustivity and that we encourage the reading of the original works.

Primary material

Manuscripts

[II/13 a)] *Letter from* Sigismund Cohn *to* C.W. Borchardt. Hirschberg, August, 25th 1859, 3 p.

[14]This is related to Nanson's [63] and Jordan's [49] methods for proving Jacobi's bound, using the following elementary property of tropical geometry. Denote by A_{i,j_0} the order matrix where line i and column j_0 have been suppressed, we have $|A|_{\mathrm{T}} = \max_i a_{i,j_0} + |A_{i,j_0}|_{\mathrm{T}}$.

[II/13 b)] Carl Gustav Jacob Jacobi, Manuscript *De ordine systematis æquationum differentialium canonici variisque formis quas inducere potest*, folios 2186–96, 2200–2206. 35 p. Basis of Cohn's transcription. English translation in [36].

[II/13 c)] Sigismund Cohn, transcription of [II/13 b)] with corrections and notes by Carl Wilhelm Borchardt, 39 p.

[II/23 a)] *Reduction simultaner Differentialgleichungen in ihre canonische Form und Multiplicator derselben.*, manuscript by Jacobi, pages: 2214–2237. Five different fragments: 2214–2216; 2217–2220 (§ 17-18); 2221–2225 (§ 17); 2226–2229; 2230–2232, 2235, 2237, 2236, 2238 (numbered from 1 to 13). English translation in [37]

[II/23 b)] *De aequationum differentialium systemate non normali ad formam normalem revocando*, manuscript by Jacobi p. 2238, 2239–2241, 2242–2251. 25 p. Envelop by Borchardt. The basis of [Jacobi 2]. English translation in [37].

[II/24] *De aequationum differentialium systemate non normali ad formam normalem revocando*, manuscript from unknown hand, mixing [Jacobi 1] and [Jacobi 2].

[II/25] *De aequationum differentialium systemate non normali ad formam normalem revocando.* Abstract and notes by Borchardt. 8 p.

Journals and complete works

[Crelles 27] *Journal für die reine und angewandte Mathematik*, sieben und zwanzigster Band, Berlin, 1844.

[Crelles 29] *Journal für die reine und angewandte Mathematik*, neun und zwanzigster Band, Berlin, 1845.

[Crelle 65] *Journal für die reine und angewandte Mathematik*, Bd LXIV, Heft 4, p. 297-320, Berlin, Druck und Verlag von Georg Reimer, 1865.

[GW IV] *C.G.J. Jacobi's gesammelte Werke*, vierter Band, herausgegeben von K. Weierstrass, Berlin, Druck und Verlag von Georg Reimer, 1886.

[GW V] *C.G.J. Jacobi's gesammelte Werke*, fünfter Band, herausgegeben von K. Weierstrass, Berlin, Druck und Verlag von Georg Reimer, 1890.

[VD] *Vorlesungen über Dynamik von C. G. J. Jacobi nebstes fünf hinterlassenen Abhandlungen desselben*, herausegegeben von A. Clesch, Berlin, Druck und Verlag von Georg Reimer, 1866.

References

[1] Adelson-Velsky (Georgii) and Landis (Evgenii), "An Algorithm for the Organization of Information", *Doklady Akademiya Nauk SSSR* **146**, 263–266, 1962; English translation, *Soviet Math.* **3**, 1259-1263.

[2] Alt, (Helmut), Blum (Norbert), Mehlhorn (Kurt) and Paul (Manfred), "Computing a maximum cardinality matching in a bipartite graph in time $O(n^{1.5}\sqrt{m/\log n}$", *Information Processing Letters*, **37**, (4), 237–240, 1991.

[3] Bellman (Richard), "On a routing problem", *Quarterly of Applied Mathematics*, **16** 87–90, 1958.

[4] Boulier (François), Lemaire (François) and Moreno Maza (Marc), "PARDI !", *International Symposium on Symbolic and algebraic computation 2001*, ACM Press, 38–47, 2001.

[5] Boulier (François), Lazard (Daniel), Ollivier (François) and Petitot (Michel), "Computing representations for radicals of finitely generated differential ideals", Special issue Jacobi's Legacy of *AAECC*, J. Calmet and F. Ollivier eds., 20, (1), 73–121, 2009.

[6] Boole (George), *A treatise on differential equations*, Macmillan and Co, Cambridge, 1859.

[7] Burkard (Rainer), Dell'Amico (Mauro) and Martello (Silvano), *Assignment problems*, Society for Industrial and Applied Mathematics, Philadelphia, 2012.

[8] Castro-Jiménez (Francisco-Jesus), "Calculs effectifs pour les idéaux d'opérateurs différentiels", *Actas de la II Conferencia Internacional de Geometría Algebraica*, La Rábida, Travaux en Cours 24, Hermann, Paris, 1987.

[9] Chou (Shang-Ching) and Gao (Xiao-Shan), "Ritt-Wu's Decomposition Algorithm and Geometry Theorem Proving", Proceedings of CADE 10, *Lecture Notes in Artificial Intelligence* **449**, Springer-Verlag, 1990, 207–220.

[10] Cluzeau (Thomas) and Hubert (Évelyne), "Resolvent Representation for Regular Differential Ideals", *Applicable Algebra in Engineering, Communication and Computing*, **13**, (5), 395–425, 2003.

[11] Chrystal, *Transactions of the Royal Society of Edinburgh*, vol. 38, p. 163, 1895.

[12] Cohn (Richard M.), "The Greenspan bound for the order of differential systems", *Proc. Amer. Math. Soc.* **79** (1980), n° 4, 523–526.

[13] Cohn (Richard M.), "Order and dimension", *Proc. Amer. Math. Soc.* **87** (1983), n° 1, 1–6.

[14] D'Alfonso (Lisi), Jeronimo (Gabriela) and Solernó (Pablo), "On the complexity of the resolvent representation of some prime differential ideals", it Journal of Complexity, **22**, 3, (June 2006), table of contents 396–430.

[15] D'Alfonso (Lisi), Jeronimo (Gabriela) and Solernó (Pablo), "A linear algebra approach to the differentiation index of generic DAE systems", preprint, 2006, http://arxiv.org/abs/cs/0608064

[16] D'Alfonso (Lisi), Gabriela Jeronimo (Gabriela), Ollivier (François), Sedoglavic (Alexandre) and Solernó (Pablo), "A geometric index reduction method for implicit systems of differential algebraic equations", *J. Symb. Comput.* **46** (10) 1114-1138, 2011.

[17] D'Alfonso (María Elisabet), "Métodos simbólicos para sistemas de ecuaciones álgebro-diferenciales", PhD thesis, University of Buenos Aires, 2006.

[18] Della Dora (Jean), Dicrescenzo (Claire) andDuval (Dominique), "About a new method for computing in algebraic number fields", *Eurocal'85*, Springer, LNCS **204**, 289-290, 1985.

[19] Dijkstra (Edsger Wybe), "A note on two problems in connexion with graphs", *Numerische Mathematik*, **1**, 269–271, 1959.

[20] Dinic, (E. A.) and Kronrod, (M. A.), "An algorithm for the solution of the assignment problem", *Soviet Math . Dokl.* **69** (6), 1324–1326, 1969.

[21] Duan, (Ran) and Pettie, (Seth), "Approximating Maximum Weight Matching in Near-Linear Time", *Proceedings of the 51rst Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 673–682, 2010.

[22] Edmonds, (J.) and Karp, (R. M.), "Theoretical improvements in algorithmic efficiency for network flow problems", *J. ACM*, **19**, 248–264, 1972.

[23] Egerváry (Jenő), "Matrixok kombinatorius tulajdonságairól" [In Hungarian: On combinatorial properties of matrices], *Matematikai és Fizikai Lapok*, vol. 38, 1931, 16–28; translated by H. W. Kuhn as Paper 4, Issue 11 of *Logistik Papers*, Georges Washington University Research Project, 1955.

[24] Faugère (Jean-Charles), Gianni (Patricia), Lazard (Daniel) and Mora (Teo), "Efficient Computation of Zero-dimensional Gröbner Bases by Change of Ordering", *Journal of Symbolic Computation*, **16**, (4), 329–344, 1993.

[25] Frobenius (Ferdinand Georg), "Über Matrizen aus nicht negativen Elementen", *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin*, 456–477, 1912.

[26] Frobenius (Ferdinand Georg), "Über zerlegbare Determinanten", *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin*, 274–277, 1917.

[27] Gabow (H.N.) and Tarjan (R.E.), "Faster scaling algorithms for network problems", *SIAM J. Comput.*, **18** 1013–1036, 1989.

[28] Giusti (Marc), Lecerf (Gregoire) and Salvy (Bruno), "A Grobner free alternative for polynomial system solving", Proceedings FOCM'99. – 1999.

[29] Greenspan (Bernard), "A bound for the orders of components of a system of algebraic difference equations", *Pacific J. Math.*, **9**, 1959, 473–486.

[30] Grier (David Alan), "Dr. Veblen Takes a Uniform", *The American Mathematical Monthly*, vol. 108, December, 2001, p. 927.

[31] Grier (David Alan), *When computers were human*, Princeton University Press, 2005.

[32] Hopcroft (John E.) et Karp (Richard M.), "An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs", *SIAM Journal on Computing*, 2 (4), 225–231, 1973.

[33] Houtain (Louis), "Des solutions singulières des équations différentielles", *Annales des universités de Belgique*[15], années 1851–1854, 973–1323.

[34] Hubert (Évelyne), "Essential components of an algebraic differential equation", J. Symbolic Computation, 21, 1–24, 1999.

[35] Ibarra (Oscar H.) and Moran (Shlomo), "Deterministic and probabilistic algorithms for maximum bipartite matching via fast matrix multiplication", *Inform. Process. Lett.*, 12–15, 1981.

[15]Reference established from *Bull. Amer. Math. Soc.* **12** (1906), p. 212. In the copy I could consult, there is no publisher indication, and the handwritten date 1852.

[36] Jacobi (Carl Gustav Jacob), "De investigando ordine systematis aequationum differentialum vulgarium cujuscunque ", [Crelle 65] p. 297-320, [GW V] p. 193-216. English translation, *AAECC*, **20**, (1), 7–32, 2009.

[37] Jacobi (Carl Gustav Jacob), "De aequationum differentialum systemate non normali ad formam normalem revocando", [VD] p. 550–578 and [GW V] p. 485-513. English translation, *AAECC*, **20**, (1), 33–64, 2009.

[38] Jacobi (Carl Gustav Jacob), "Theoria novi multiplicatoris systemati aequationum differentialium vulgarium applicandi", first published in [Crelle 27] Heft III 199–268 (part I) and [Crelle29] Heft III 213–279 and Heft IV 333–376 (part II), reproduced in [GW IV], 317–509.

[39] Jacobi (Carl Gustav Jacob), "Lettre adressée à M. le président de l'académie des sciences", *Journal de math. pures et appliquées*, V, 1840, p. 350-351.

[40] Jacobi (Carl Gustav Jacob), "Sur un nouveau principe de la mécanique analytique", *C. R. Acad. Sci. Paris*, **XV**, 1842, 202–205; [GW IV], 289–294.

[41] Jacobi (Carl Gustav Jacob), "Sul principo dell'ultimo multiplicatore e suo uso como nuovo principio generale di meccanica", *Giornale arcadico*, **XCIX**, 129–146, 1844; [GW IV] 511-522.

[42] Jacobi (Carl Gustav Jacob), "Sur un théorème de Poisson", *C. R. Acad. Sci. Paris*, **XI**, 1840, p. 529; [GW IV], 143–146.

[43] Jacobi (Carl Gustav Jacob), *Canon arithmeticus, sive talulæ quibus exhibentur pro singulis numeris primis...*, Berolini, typis academicis, 1839.

[44] *Briefwechsel zwischen C.G.J. Jacobi und M.H. Jacobi*, herausgegeben von W. Ahrens, Abhandlungen zur Geschichte der Mathematischen Wissenschaften, XXII. Heft, Leipzig, Druck und Verlag von B.G. Teubner, 1907.

[45] Johnson (Joseph), "Differential dimension polynomials and a fundamental theorem on differential modules", *Am. J. Math.*, 91, 251-257 (1969).

[46] Johnson (Joseph), "Kähler Differentials and Differential Algebra", *The Annals of Mathematics*, 2nd Ser., Vol. 89, No. 1 (Jan., 1969), pp. 92-98.

[47] Johnson (Joseph), " A notion of regularity for differential local algebras", *Contributions to algebra*, A Collection of Papers Dedicated to Ellis Kolchin, Edited by Hyman Bass, Phyllis J. Cassidy and Jerald Kovacic, Elsevier, 211–232, 1977.

[48] JOHNSON (Joseph), "Systems of n partial differential equations in n unknown functions: the conjecture of M. Janet", *Trans. of the AMS*, vol. 242, Aug. 1978.

[49] JORDAN (Camille), "Sur l'ordre d'un système d'équations différentielles ", *Annales de la société scientifique de Bruxelles*, vol. 7, B., 127–130, 1883.

[50] KNUTH (Donald E.), THE ART OF COMPUTER PROGRAMMING III. SORTING AND SEARCHING, Addison-Wesley, Reading, 1973.

[51] KOLCHIN (Ellis Robert), *Differential algebra and algebbraic groups*, Academic Press, New York, 1973.

[52] KONDRATIEVA (Marina Vladimirovna), MIKHALEV (Aleksandr Vasil'evich), PANKRATIEV (Evgeniĭ Vasil'evich), "Jacobi's bound for systems of differential polynomials" (in Russian), Algebra. M.: MGU, 1982, s. 79-85.

[53] KONDRATIEVA (Marina Vladimirovna), MIKHALEV (Aleksandr Vasil'evich), PANKRATIEV (Evgeniĭ Vasil'evich), "Jacobi's bound for independent systems of algebraic partial differential equations", *AAECC*, , Volume **20**, (1), 65–71, 2009.

[54] KŐNIG (Dénes), "Gráfok és mátrixok", *Matematikai és Fizikai Lapok* 38: 116–119, 1931.

[55] KŐNIG (Dénes), *Theorie der endlichen und unendlichen Graphen*, (1936), Chelsey, New-York, 1950.

[56] KUHN (Harold H.), "The Hungarian method for the assignment problem", *Naval res. Logist. Quart.* **2** (1955), 83–97.

[57] KUHN (Harold H.), "A tale of three eras: The discovery and rediscovery of the Hungarian Method", *European Journal of Operational Research*, 219 (2012), 641–651.

[58] MUNKRES (James), "Algorithms for the assignment and transportation problems", *J. Soc. Industr. Appl. Math.*, **5** (1957), 32–38.

[59] LANDO (Barbara A.), "Jacobi's bound for the order of systems of first order differential equations", *Trans. Amer. Math. Soc.* **152** 1970, 119–135.

[60] MACLAGAN (Diane) and STURMFELS (Bernd), *Introduction to Tropical Geometry*, preprint, 2015.
https://homepages.warwick.ac.uk/staff/D.Maclagan/papers/TropicalBook20.4.14.pdf

[61] Morrison (Sally), "The Differential Ideal $[P] : M^{\infty}$", *Journal of Symbolic Computation*, **28**, 1999, 631–656.

[62] Monge (Gaspard), "Mémoire sur la théorie des déblais et des remblais", *Histoire de l'Académie royale des Sciences*, [année 1781. Avec les Mémoires de Mathématique & de Physique pour la même Année] (2e partie) (1784) [Histoire: 34–38, Mémoires :] 666–704.

[63] Nanson (Edward John), "On the number of arbitrary constants in the complete solution of ordinary simultaneous differential equations", *Messenger of mathematics* (2), vol. 6, 77–81, 1876.

[64] Nucci (Maria Clara) and Leach (Peter G. L.), "Jacobi's last multiplier and symmetries for the Kepler problem plus a lineal story", *J. Phys. A: Math. Gen.*, **37**, 7743–7753, 2004.

[65] Nucci (Maria Clara) and Leach (Peter G. L.), "Jacobi's last multiplier and its applications in mechanics", *Physica Scripta*, **78**, 06511, 6 p., 2008.

[66] Ollivier, (François) and Sadik (Brahim), *La borne de Jacobi pour une diffiété définie par un système quasi régulier* (Jacobi's bound for a diffiety defined by a quasi-regular system), *Comptes rendus Mathematique*, **345**, 3 (page 139) , 2007, 139–144.

[67] Ollivier, (François), " Jacobi's bound and normal forms computations. A historical survey", *Differential Algebra And Related Topics*, World Scientific Publishing Company, 2008. ISBN 978-981-283-371-6

[68] Pryce (John D.), "A simple structural analysis method for daes", *BIT: Numerical Mathematics*, vol. 41, no. 2, 364–394, Swets & Zeitlinger, 2001.

[69] Rabinowitsch, (J.L.) [Rainich (George Yuri)], "Zum Hilbertschen Nullstellensatz", *Math. Ann.*, **102**, 520, 1929.

[70] Ritt, (Joseph Fels), "Jacobi's problem on the order of a system of differential equations", *Annals of Mathematics*, vol. 36, 1935, 303–312.

[71] Ritt, (Joseph Fels), 1950. *Differential Algebra*, Amer. Math. Soc. Colloq. Publ., vol. 33, A.M.S., New-York.

[72] Shaleninov (A. A.), "Removal of topological degeneracy in systems of differential evolution equations", *Automation and Remote Control*, 51, 1599–1605, 1991 (English version). Translation from *Avtomatika i Telemekhanika*, Issue 11, 163–170, 1990.

[73] Schrijver (Alexander), *Combinatorial Optimization. Polyhedra and Efficiency*, Springer, 2003.

[74] Schrijver (Alexander), "On the history of combinatorial optimization (till 1960)", *Handbook of Discrete Optimization*, K. Aardal, G.L. Nemhauser, R. Weismantel, eds., Elsevier, Amsterdam, 2005, pp. 1–68.

[75] Schrijver (Alexander), "On the history of the shortest path problem", *Documenta Math.*, Extra Volume: Optimization Stories, 2012, 155–167.

[76] Tomasovic, Jr. (Joseph S.), *A generalized Jacobi conjecture for arbitrary systems of differential equations*, Dissertation, Columbia University, 1976. [52 feuillets, 29 cm. Columbia University Rare Books and Manuscript Library, Butler Library, New-York, LD1237.5D 1976 T552]

[77] Tomizawa, (N.), "On some techniques useful for solution of transportation network problems", *Networks*, **1**, 173–194, 1971.

[78] Volevich, (Leonid Romanovich), "On general systems of differential equations", *Soviet. Math.* **1**, 1960, 458–465.