PHILOSOPHICAL TRANSACTIONS A

rsta.royalsocietypublishing.org

Research

Cite this article: Malheiro A, Reis JF. 2019 Identification of proofs via syzygies. *Phil. Trans. R. Soc. A* **377**: 20180275.
http://dx.doi.org/10.1098/rsta.2018.0275

Accepted: 9 November 2018

One contribution of 11 to a theme issue 'The notion of 'simple proof' - Hilbert's 24th problem'.

Subject Areas:
mathematical logic, complexity

Keywords:
Hilbert's 24th problem, identification of proofs, syzygies

Author for correspondence:
José Francisco Reis
e-mail: jfd.reis@campus.fct.unl.pt

Identification of proofs via syzygies

António Malheiro[1,2] and José Francisco Reis[2]

[1]Departamento de Matemática, and [2]Centro de Matemática e Aplicações (CMA), Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829–516 Caparica, Portugal

AM, 0000-0003-1186-6216; JFR, 0000-0001-8022-7499

In 1900, Hilbert gave a lecture at the International Congress of Mathematicians in Paris, for which he prepared 23 problems that mathematicians should solve during the twentieth century. It was found that there was a note on a 24th problem focusing on the problem of simplicity of proofs. One of the lines of research that was generated from this problem was the identification of proofs. In this article, we present a possible method for exploring the identification of proofs based on the membership problem original from the theory of polynomial rings. To show this, we start by giving a complete worked-out example of a membership problem, that is the problem of checking if a given polynomial belongs to an ideal generated by finitely many polynomials. This problem can be solved by considering Gröbner bases and the corresponding reductions. Each reduction is a simplification of the polynomial and it corresponds to a rewriting step. In proving that a polynomial is a member of an ideal, a rewriting process is used, and many different such processes can be considered. To better illustrate this, we consider a graph where each rewriting step corresponds to an edge, and thus a path corresponds to a rewriting process. In this paper, we consider the identification of paths, within the context of the membership problem, to propose a criterion of identification of proofs.

This article is part of the theme issue 'The notion of 'simple proof' - Hilbert's 24th problem'.

1. Introduction

This article is about the identification of proofs, a line of research that emerged from a note by Hilbert on his 24th

THE ROYAL SOCIETY PUBLISHING

problem. We quote the note from [1], and we keep the comments by R. Thiele:

> 'The 24th problem in my Paris lecture was to be: Criteria of simplicity, or proof of the greatest simplicity of certain proofs. Develop a theory of the method of proof in mathematics in general. Under a given set of conditions there can be but one simplest proof. Quite generally, if there are two proofs for a theorem, you must keep going until you have derived each from the other, or until it becomes quite evident what variant conditions (and aids) have been used in the two proofs. Given two routes, it is not right to take either of these two or to look for a third; it is necessary to investigate the area lying between the two routes. Attempts at judging the simplicity of a proof are in my examination of syzygies and syzygies [Hilbert misspelled the word syzygies] between syzygies [see Hilbert [2, lectures XXXII-XXXIX]]. The use or the knowledge of a syzygy simplifies in an essential way a proof that a certain identity is true. Because any process of addition [is] an application of the commutative law of addition etc. [and because] this always corresponds to geometric theorems or logical conclusions, one can count these [processes], and, for instance, in proving certain theorems of elementary geometry (the Pythagoras theorem, [theorems] on remarkable points of triangles), one can very well decide which of the proofs is the simplest. [Author's note: Part of the last sentence is not only barely legible in Hilbert's notebook but also grammatically incorrect. Corrections and insertions that Hilbert made in this entry show that he wrote down the problem in haste.]'

From this note, it is clear that Hilbert's 24th problem is about the simplicity of proofs. Therefore, in this problem, one has to compare proofs and consequently it is important to know when two proofs can be said to be the same. For instance, one may formalize a proof in two different ways and if we are given two different formalizations, it is important to know whether these formalizations correspond to the same proof. In fact, if they do correspond to the same proof and we want some kind of criteria measuring the simplicity of proofs, that criteria must assign the same simplicity for those two formalizations.

In this paper, we hope to give a possible reason for Hilbert mentioning syzygies in his 24th problem. In order to do such, we consider the problem of identification of proofs in the context of the membership problem of a polynomial f in an ideal. The main idea throughout this article is the following: given a Gröbner basis, we consider the corresponding reductions and apply those to f in all possible ways. With this information, we can draw a graph where a rewriting step corresponds to an edge. We then can easily identify some paths with each other because they essentially come from applying the rewriting rules in a different order. After these identifications, how can we identify two paths that are not a simple reordering of the application of the rewriting rules? To each path, there is a corresponding expression for f in terms of the Gröbner bases. So, our aim is to identify those expressions of f, and we can do it modulo a certain module. In fact, we do it modulo a syzygy module and we hope this explains why Hilbert mentioned syzygies in his 24th problem.

In §2, we apply this reasoning to a polynomial f and an ideal I. In §2a, we introduce the context of this membership problem. For instance, we introduce the concept of a rewriting rule, which is well known from the theory of polynomial rings, and through a rewriting process, we are able to solve this problem. The rewriting rules that we used in order to solve the membership problem could have been used in a different way, giving rise to a different rewriting process. In §2b, we consider all rewriting processes that solve our problem and translate this information into a graph. In this graph, we have paths from the vertex corresponding to f to the vertex corresponding to zero. We identify some paths and in §2c, we build an analogy between the paths in the graph and proofs of the membership of f to I. For instance, each way of reducing f to zero can be seen as a proof that f is a member of the ideal. Hence, a path from f to zero, that corresponds to a way of reducing f to zero, can be seen as a proof of the membership of f to I. With this reasoning, we build the analogy between paths and proofs. Therefore, we translate the problem of identification of proofs into a problem of identification of paths. In this paper, we propose a criterion of identification of proofs. We start by translating proofs into paths but this is not enough,

given that when we identify some paths with one another, we were not able to identify all those paths. Hence, in §2d, considering the identified paths, we obtain two expressions for f in terms of g_1 and g_2, which are going to be elements of a Gröbner basis. Finally, in §2e we are able, by 'subtracting' a path from another path, to identify those expressions and consequently all paths. Therefore, we identify all proofs. For a general case, it is not obvious which paths to consider in the subtraction. Thus, in §2f instead of subtracting paths, we consider a resolution in the context of our example. Considering resolutions has the advantage that we can generalize our reasoning to any membership problem of a polynomial to an ideal. This is done in §2g. Also, we obtain a criterion to identify proofs in the context of a membership problem: consider all coordinates (in a Gröbner basis) of a polynomial f and note that they are the same modulo a syzygy module. Or, more informally, translate proofs into paths, consider the corresponding expression of each path and note that they are the same modulo a module that happens to be a syzygy module.

In the last section of this article, we ask if this kind of reasoning, translate proofs into paths of a graph, can be formalized. In addition, we mention that it might be interesting to propose a distance between proofs, using the established analogy, by taking into account the minimal number of second syzygy modules that are needed to identify the corresponding paths. For instance, if we needed three second syzygy modules to identify two proofs, the distance between those proofs would be three. While if they only needed one second syzygy module, the distance would be one and hence being the distance shorter it would mean that the proofs are more similar.

2. From reduction of polynomials to identification of proofs

(a) Reduction of polynomials—an example

We are going to present an example of a membership problem and relate it to the identity of proofs. We are considering polynomials in two variables with complex coefficients. That is, we are working in $\mathbb{C}[x,y]$ and we want to know if a polynomial f, in our example, $-xy^4+xy^3$, is in the ideal I generated by the polynomials $g_1=y^4$ and $g_2=xy+y^2$. The ideal generated by the polynomials g_1 and g_2 is denoted by $I=(g_1,g_2)$ and it corresponds to the set of 'linear combinations' of g_1 and g_2: that is, the polynomials of the form $f_1g_1+f_2g_2$, where $f_1,f_2\in\mathbb{C}[x,y]$. So, the membership problem that we consider is the problem of determining whether polynomials $f_1,f_2\in\mathbb{C}[x,y]$ exist in such a way that we can have the equality

$$f=f_1g_1+f_2g_2.$$

To make this problem easier to solve, one usually considers Gröbner bases. The concept of a Gröbner bases only makes sense when we have an order among monomials. In this example, we will consider the lexicographic order among monomials, here denoted $>$. In this order, given two monomials $x^{a_1}y^{a_2}$ and $x^{b_1}y^{b_2}$ in $\mathbb{C}[x,y]$, with a_1,a_2,b_1,b_2 non-negative integers, we say that $x^{a_1}y^{a_2}>x^{b_1}y^{b_2}$ if $a_1>b_1$ or, in case $a_1=b_1$, we have $a_2>b_2$. To give examples in the context we are working under, we have $xy>y^2$ and $x>y^2$, even though x has degree 1 and y^2 has degree 2.

Having fixed an ordering on monomials, we can check whether a set of polynomials ($\{g_1,g_2\}$ in our case) is a Gröbner basis (with respect to the fixed ordering). We will not define Gröbner bases but, in a few paragraphs, we will mention a nice property they have that will be important in this example. For more details on Gröbner bases, we refer the reader to [3] or [4]. In our running example, the set $\{g_1,g_2\}$, is already a Gröbner basis. An important tool to obtain a Gröbner basis is the so-called Buchberger's Algorithm [4, Buchberger's Algorithm 15.9]. This algorithm provides us with a method to obtain, from a given set of polynomials, a set of polynomials which is a Gröbner basis and which generates the same ideal as the original set of polynomials. It is worth noting, that throughout this article, we assume commutative. Nevertheless, the Grobner bases theory can also be adapted for non-commuting variables, being in that case known as Gröbner–Shirshov bases (c.f. [5]).

We now informally introduce the concept of reduction, which is well known from the theory of polynomial rings. Let us assume that we have $h=f_1g_1+h'$, where h' and f_1 are polynomials

in $\mathbb{C}[x, y]$ (in the univariate polynomial case we might get such a decomposition by applying the usual Euclidean division algorithm) and we want to know whether h is in I. Given that g_1 is in I, it is enough to show that h' is in I. Actually, $h \in I$ if, and only if, $h' \in I$. So, we might look at g_1 as a 'zero' in the sense that it can be ignored when checking the membership in I. More formally, we can write $g_1 \longrightarrow 0$ to mean that g_1 rewrites as zero. Hence, we would write $h = f_1 g_1 + h' \longrightarrow h'$ and say that h rewrites to h' modulo g_1. We can repeat this process, that is, if we could isolate one of the g_1 or g_2 in h' we could apply these rewriting rules: $g_1 \longrightarrow 0$ and $g_2 \longrightarrow 0$. Something that is clear is that if by applying these rewriting rules we obtain zero, then the polynomial h is indeed in I. Not so clear, but a consequence of having a Gröbner bases, is that we can deduce that if we reach a non-zero polynomial for which we cannot apply more reductions (modulo g_1 or g_2), then h is not in I.

From the above reasoning, we can check whether the polynomial h is in I by checking if it is possible to write h as a linear combination of the g_1 and g_2. This is not an easy process, and this is where Gröbner bases simplify the process. Let us give a bit more detail on the reduction step, which can be simplified if we look at each reduction modulo g_i (interpreted by the equation $g_i = 0$) as rewriting the greatest monomial in g_i to the remaining monomials of g_i. More precisely, consider the greatest monomial (with respect to the order on monomials considered above) of each g_i and call this monomial the initial term of g_i, denoted by $in(g_i)$. Then a reduction modulo g_i corresponds to apply the rewriting rule

$$in(g_i) \longrightarrow in(g_i) - g_i.$$

In our running example, we have the rewriting rules

$$\begin{aligned} (1) \quad y^4 &\longrightarrow 0 \\ (2) \quad xy &\longrightarrow -y^2. \end{aligned}$$

To be able to track each application of a rewriting rule, we shall write the corresponding number above the arrow. This will become more clear in the following reduction of f:

$$\begin{aligned} f = -x(y^4) + xy^3 &\xrightarrow{1} -x(0) + xy^3 = xy^3 = (xy)y^2 \\ &\xrightarrow{2} (-y^2)y^2 = -y^4 \\ &\xrightarrow{1} -0 = 0. \end{aligned}$$

Since f reduces to 0, we conclude that f is in I.

(b) Identifying paths

Now, one may wonder if this is the only way of using these rewriting rules to prove that f is in I. It is not

$$\begin{aligned} f = -(xy)y^3 + xy^3 &\xrightarrow{2} y^5 + xy^3 = y(y^4) + xy^3 \\ &\xrightarrow{1} y(0) + xy^3 = (xy)y^2 \\ &\xrightarrow{2} -y^4 \\ &\xrightarrow{1} 0. \end{aligned}$$

Thus, we have found a different way of applying these rewriting rules to show that f is I. Let us then apply the rewriting rules in all possible ways and, with that, we draw the graph in figure 1 that shows all possible paths from f to zero by applying these rewriting rules. There, an edge corresponds to a rewriting step and a rewriting process corresponds to a path from f to zero. Note that by a rewriting process we mean a certain application of the rewriting rules that shows, in this context, that f is in I.

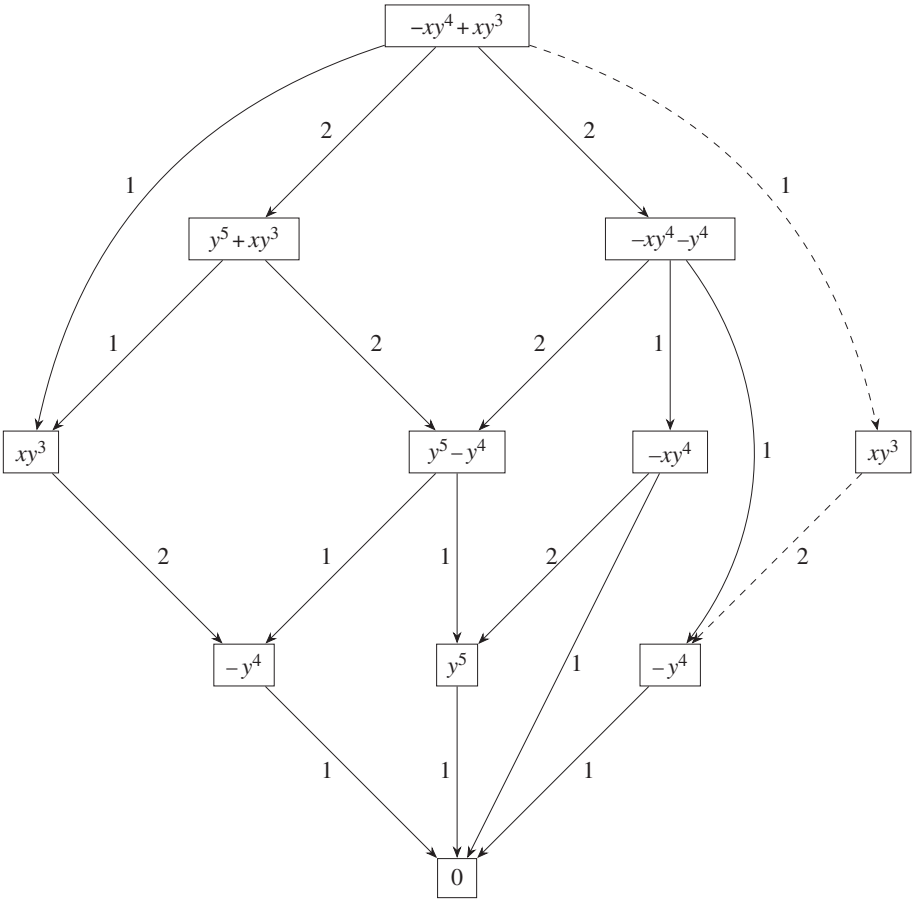

Figure 1. This graph is drawn considering all possible ways of applying the rewriting rules to f. The vertices correspond to polynomials obtained from f and the application of the rewriting rules. The edges are oriented and labelled by the number of the rewriting rule that was applied. That is, each edge corresponds to a rewriting step. The dashed path is a duplication, and later on we will explain why we duplicated a path.

Now, considering this graph, we see nine different paths from f to zero. Nevertheless, we realize quite easily that some paths are essentially the same.

In fact, whenever we have a polynomial to which we can apply the rewriting rules in different ways if there is no overlapping in monomials to which we apply the rules, the paths are going to be essentially the same. An example of such 'no overlapping' is: $f = -xy^4 + xy^3 \xrightarrow{2} y^5 + xy^3 \xrightarrow{2} y^5 - y^4$ and $f = -xy^4 + xy^3 \xrightarrow{2} -xy^4 - y^4 \xrightarrow{2} y^5 - y^4$.

So, whenever there is no overlapping in the application of the rewriting rules, let us colour the 'area' between those applications of the rewriting rule. An 'area' betweeen rewriting rules is called a cell.

An example where there is overlapping (in y) is: $f = -x(y^4) + xy^3 \xrightarrow{1} xy^3$ and $f = -(xy)y^3 + xy^3 \xrightarrow{2} y^5 + xy^3$. In this case, we do not colour the area between the application of these rewriting rules. Applying this reasoning to the entire graph, we obtain figure 2. Huet already studied overlap situations and gave some criteria to know when rewriting rules commute, see [6].

As we can see in the graph, there are some blank cells and following our reasoning means that paths separated by those blank cells cannot be identified. Therefore, we end up having only two different paths. To make this more clear, suppose we have two different paths and we are allowed to drag edges of those paths as long as we do it on coloured cells. If we can drag one path to

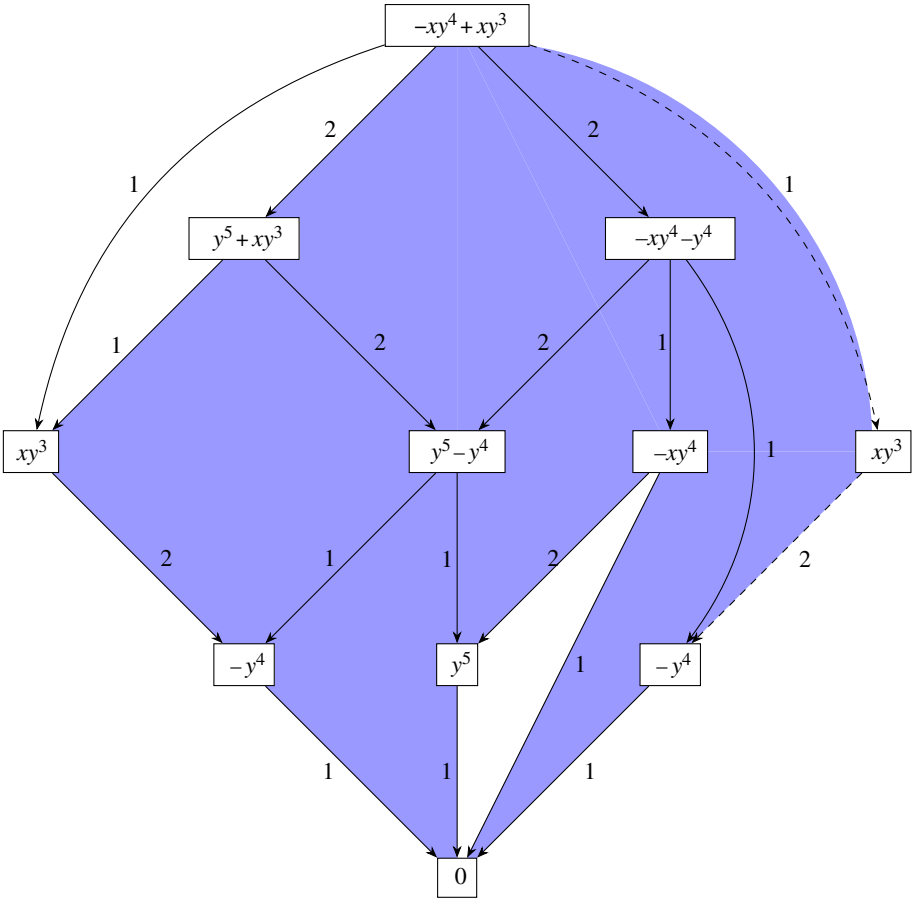

Figure 2. Possible reductions of f, where no overlapping in different applications of rewriting rules corresponds to colouring cells between them. Some cells are left blank because there is overlap in the application of different rewriting rules. (Online version in colour.)

the other one following this rule, we have that those paths are, in some sense, the same, and we identify those paths with one another. In figure 3, we show what we mean by dragging edges within coloured cells (we omit some information from the graphs). If we kept dragging paths, as shown in the figure, we would conclude that those two paths were the same. Now, as there are two blank cells, we see that we cannot identify all paths. In fact, we conclude that if before we had nine different paths, after these identifications we only have two paths.

(c) The analogy between paths and proofs

We now start to relate paths and proofs. Let us consider figure 1. There we had nine different paths and these paths came from different applications of the rewriting rules, and these different applications of the rewriting rules mean that we have different proofs of the statement f is in I. Given that in figure 1 we have nine paths from f to zero, we can say that we have nine different proofs. So a path in the graph of figure 1 corresponds to a proof. Hence, we are rephrasing our problem on the identification of proofs into a problem on the identification of paths.

We noted that some paths could be identified with each other and with those identifications we end up with two paths. Hence, we could say, as a path corresponds to a proof, that we have two proofs up to this identification. It is the blank triangle that makes it impossible to identify all paths, and hence all proofs. Note that it happens because there is overlapping in the application of the

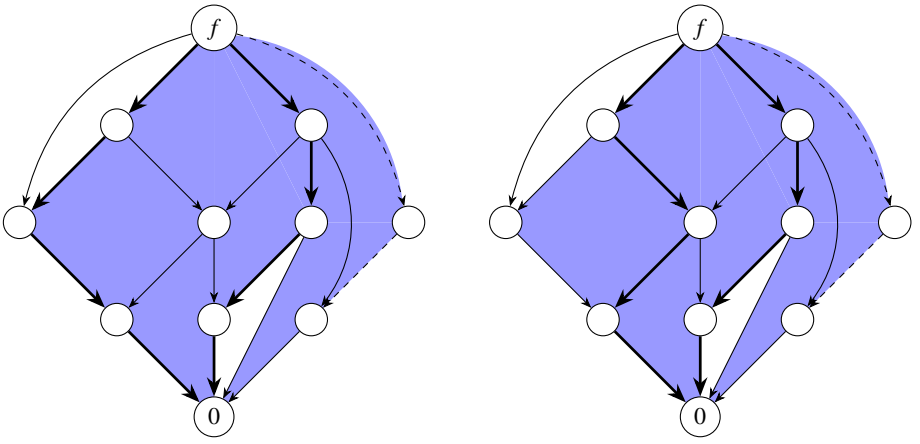

Figure 3. We are dragging two edges of a path in order to move it into another path without passing through blank cells. We just show one step, if we kept moving edges we would have moved one path to the other. (Online version in colour.)

rewriting rules and one path is longer than the other. So, considering proof ordering, the proof that corresponds to the longer path is larger than the other, see [7, Chapter 2]. Note, however, that the general conflict case is not given by a triangle, as in our example, but is given by a parallelogram.

(d) Paths as expressions

It might not be obvious, but when we apply a rewriting rule we are subtracting, from the polynomial to which we apply that rewriting rule, a polynomial divisible by a g_i, depending on which rule we apply. For instance, in the reduction $f = -xy^4 + xy^3 \xrightarrow{1} xy^3$ we are subtracting $-xg_1$ from f. Thus, since f is in I it means that to a path in the graph it corresponds to write f as a linear combination of g_1 and g_2.

More interesting is that when two paths are identifiable, they give rise to the same expression of f in terms of the g_i's, for $i = 1, 2$. This was expected since there is no overlapping in the application of the rewriting rules for identifiable paths.

Thus, in our running example, we have two paths up to this identification and to these two paths correspond expressions of f in terms of the Gröbner bases we have been considering. We have

$$\text{E1: } f = (y-1)g_1 + (y^2 - y^3)g_2$$
$$\text{E2: } f = -(x+1)g_1 + y^2 g_2.$$

(The first expression E1 corresponds to the paths on the left of the triangular blank cell and the second expression E2 to the paths on the right of that cell).

Consider the path around the blank cell, as in figure 4. We get

$$0 = (x+y)\,g_1 - y^3\,g_2.$$

In fact, if we consider the two paths mentioned before and we subtract the corresponding expressions, we get exactly the same result.

Thus, we have that the pair $(x+y, -y^3)$ are coordinates of zero with respect to g_1 and g_2. Now, from E1 we have that $(y-1, y^2 - y^3)$ are coordinates of f, with respect to g_1 and g_2, and from E2 we obtain $(-(x+1), y^2)$. Thus, in terms of g_1 and g_2, the coordinates of E1 and of E2 are the same modulo $(x+y, -y^3)$. Hence, we say that E1 and E2 are the same modulo $(x+y, -y^3)$. Therefore,

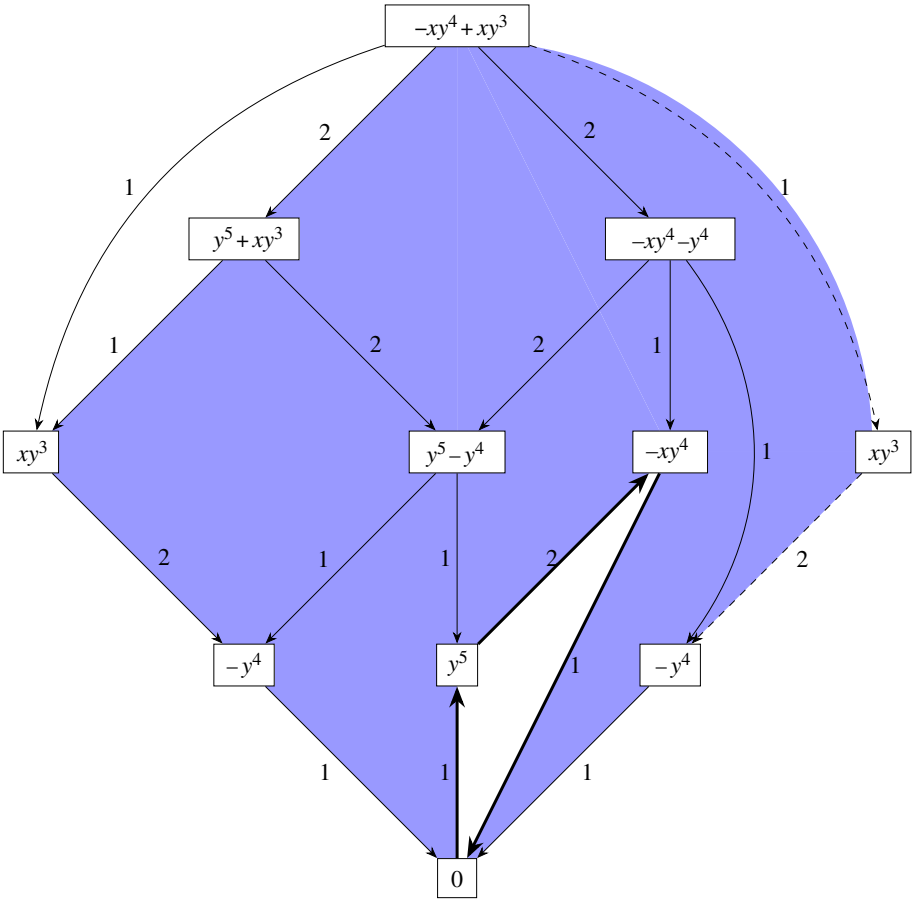

Figure 4. The path around the blank cell. (Online version in colour.)

as these expressions corresponded to some paths, we have that these paths are the same modulo $(x+y, -y^3)$.

It is important to note that if we had considered the other blank cell we would obtain the same result.

(e) Identifying proofs

We began by identifying paths which correspond to a reordering of the non-overlapping rewriting rules. In the example, we got two expressions that corresponded to two paths, and in the analogy established previously, these two paths correspond to two proofs. Thus, we have that these two expressions correspond to two proofs. We then noticed that these expressions gave different coordinates of f with respect to g_1 and g_2, but these coordinates are the same modulo $(x+y, -y^3)$. Now, the identification is not simply a reordering of the rewriting rules.

As the coordinates were the same modulo $(x+y, -y^3)$, the same reasoning can be extended to claim that we have a unique proof modulo $(x+y, -y^3)$ by some 'known reasoning'. Now, something interesting is that the module generated by $(x+y, -y^3)$ ends up being the second syzygy module of a resolution of $\mathbb{C}[x,y]/I$. And here, we may say that Hilbert had a good reason to consider syzygies in this context.

(f) Framed with resolutions

In the previous subsections, we identified all possible ways of proving the membership of f in I arguing that the coordinates, in terms of our Gröbner bases, of all the expressions (of the paths in the graph) are the same modulo $(x+y, -y^3)$. We mentioned that the module that made this possible was the second syzygy module of a certain resolution. In this subsection, we will briefly explain these ideas and concepts. For more details, we refer the reader to [4, Section 1.10] and to [8]. Indeed, if we consider the first paragraph of [8], let M be a module, over a commutative ring R, generated by $f_1, \ldots, f_n$. A syzygy is an element $\{a_1, \ldots, a_n\}$ such that $a_1 f_1 + \cdots a_n f_n = 0$. Now, the set of all syzygies is a submodule of R^n, called a syzygy module. In fact, it is easy to see that this submodule is the kernel of the map that sends the standard bases element e_i of R^n to f_i. In addition, one may repeat this process to obtain syzygy of syzygies. We now give a more formal treatment of what is a syzygy module.

Let R be an abelian group and let M be an R-module. A resolution of M is an exact sequence of R-modules:

$$\cdots \to F_n \xrightarrow{\phi_n} F_{n-1} \to \cdots \xrightarrow{\phi_2} F_1 \xrightarrow{\phi_1} F_0,$$

where the image of the homomorphism ϕ_{i+1} is equal to the kernel of the homomorphism ϕ_i, i.e., $\operatorname{Im}\phi_{i+1} = \ker \phi_i$, and we must also have $\operatorname{coker}\phi_1 := F_0/\operatorname{Im}\phi_1 = M$. If this happens, we say we have a resolution of M and we define the ith syzygy module to be the image of ϕ_i. Moreover, if there is an n such that $F_i = \{0\}$ for all $i > n$ and $F_n \neq \{0\}$ we say that we have a finite resolution of M of length n.

In the context of our example, let us find a resolution for $\mathbb{C}[x,y]/I$. We start by defining F_0, F_1 and ϕ_1. So, let F_0 be $\mathbb{C}[x,y]$, F_2 be $\mathbb{C}[x,y]^{\oplus 2}$ and ϕ_1 is

$$\begin{aligned} \phi_1 : F_1 = \mathbb{C}[x,y]^{\oplus 2} &\longrightarrow \mathbb{C}[x,y] = F_0 \\ (1,0) &\longmapsto y^4 = g_1 \\ (0,1) &\longmapsto xy + y^2 = g_2. \end{aligned}$$

We note that $\operatorname{Im}\phi_1 = I$ and that $\ker \phi_1 = \langle (x+y, -y^3) \rangle$. It is clear that $\operatorname{coker} := F_0/\operatorname{Im}\phi_1 = \mathbb{C}[x,y]/I$.

We now want to define F_2 and ϕ_2 such that $\operatorname{Im}\phi_2 = \ker \phi_1$. We set $F_2 = \mathbb{C}[x,y]$ and define

$$\begin{aligned} \phi_2 : F_2 = \mathbb{C}[x,y] &\longrightarrow F_1 = \mathbb{C}[x,y]^{\oplus 2} \\ 1 &\longmapsto (x+y, -y^3). \end{aligned}$$

Clearly, $\operatorname{Im}\phi_2 = \ker \phi_1$. Moreover, this homomorphism is injective, consequently, $\ker \phi_2 = \{0\}$. Therefore, we can set $F_3 = \{0\}$ and define $\phi_3 : F_3 = \{0\} \longrightarrow F_2 = \mathbb{C}[x,y]$ sending zero to zero. It is clear that its image is equal to the kernel of ϕ_2 and we have obtained a finite resolution of $\mathbb{C}[x,y]/I$ of length 2.

Again, $(y-1, y^2 - y^3)$ and $(-x-1, y^2)$ are coordinates of f, with respect to g_1 and g_2, and they represent the same element modulo $\ker \phi_1 = \operatorname{Im}\phi_2$, the second syzygy module. With this, we can say that there is a unique expression for f as a linear combination of g_1 and g_2 modulo the second syzygy module.

(g) A general analysis

Let us just summarize the reasoning we have been following. We first said that a path corresponded to a proof. But then we saw that some paths were essentially the same. So, we identified those paths and obtain only two different paths. We then said that these two different paths originated two different expressions for f as a linear combination of g_1 and g_2. Thus, we wanted to find a way of claiming that those expressions were the same so that we could identify the proofs that gave rise to these two expressions. What we now did states that the two expressions are the same modulo the second syzygy module of a resolution. Interestingly enough,

we can find this second syzygy module from the graph, as we did before. Now, it is in this sense that we say that there is a unique proof of the membership of f to I.

In addition, one may wonder if this is always the case in the context of a membership problem, that is if we can always say that there is a unique proof modulo the second syzygy module of a resolution. Let us try to work in a more general way.

Suppose we are working over the ring $\mathbb{K}[x_1, \ldots, x_r]$, with $\mathbb{K}$ a field, and we have I a finitely generated ideal of $\mathbb{K}[x_1, \ldots, x_r]$ with Gröbner bases formed by $g_1, \ldots, g_t$. (It is important to note that we can always obtain a Gröbner basis by Buchberger's Algorithm.) Bearing in mind our example, we want to know if the coordinates of f with respect to the $g_1, \ldots, g_t$ are unique modulo a second syzygy modulo of a resolution. Hence, let us consider $\mathbb{K}[x_1, \ldots, x_r]/I$, a $\mathbb{K}[x_1, \ldots, x_r]-$module. This is a natural choice, giving that we are just doing the same we did in our example. The next step is to consider a resolution and to obtain it we just need a sharpened version of Hilbert's Syzygy Theorem, see [4, Corollary 15.11], that states that every finitely generated module over a polynomial ring in f variables has a finite resolution of length less or equal than r. Note that Hilbert's Syzygy Theorem states something similar for graded modules, see [4, Theorem 1.13]. Thus, we obtain a finite resolution for $\mathbb{K}[x_1, \ldots, x_r]/I$ of length $n \leq r$:

$$0 \to F_n \xrightarrow{\phi_n} F_{n-1} \longrightarrow \cdots \longrightarrow F_1 \xrightarrow{\phi_1} F_0 = \mathbb{K}[x_1, \ldots, x_r].$$

Note that we must have $n \geq 1$. Indeed, if $n = 0$, we would have $F_1 = \{0\}$ and $\operatorname{Im}\phi_1 = \{0\}$ Given this is a resolution, we would have $\operatorname{coker}\phi_1 = \mathbb{K}[x_1, \ldots, x_r]/\operatorname{Im}\phi_1 = \mathbb{K}[x_1, \ldots, x_r]/I$. Thus, $I = \operatorname{Im}\phi_1 = \{0\}$. But I is not the zero ideal, as we are assuming that f is in I. Moreover, the membership problem is already solved if I was the zero ideal.

Thus, as $\operatorname{coker}\phi_1 = \mathbb{K}[x_1, \ldots, x_r]/\operatorname{Im}\phi_1 = \mathbb{K}[x_1, \ldots, x_r]/I$ and we are assuming that f is in I, we have f in $\operatorname{Im}\phi_1$. Now, suppose that there are $w, z \in F_1$, with $w \neq z$ such that $\phi_1(w) = \phi_1(z) = f$. In the analogy we have been considering, we would say that there were two 'proofs' for the membership of f in I, given there were two coordinates for f, namely w and z. Now, note that w and z are in the same coset of $\ker\phi_1$, that is, there is a $v \in F_1$ such that $w, z \in v + \ker\phi_1$. So, w and z are the same modulo $\ker\phi_1 = \operatorname{Im}\phi_2$, the second syzygy module. So, we have a 'unique proof' module the second syzygy module.

We would like to note that we do not need a finite resolution. In fact, we only need, in this context, a resolution of length $n \geq 1$.

3. Conclusion and future work

In this article, we established an analogy between the reduction of polynomials and proofs. Since the reductions can be seen as paths in a graph, we also have an analogy between paths in the graph and proofs. Doing this, we translated the problem of identifying proofs into a problem of identifying paths. We were able to easily identify (by reordering) some paths. To identify all of them we needed to consider the second syzygy module of a certain module. This syzygy module could be obtained from the graph as well. Thus, in this context, we can always identify all proofs and therefore the problem of identifying proofs is done. We may think that Hilbert may have thought something similar, at least informally, since he referred to 'routes' while we refer to 'paths' and we actually 'investigate the area lying between the two routes' (from [1]). In addition, we considered syzygies as a tool to identify these proofs.

Now, one may wonder if we can formalize this idea: translate the problem of identification of proofs into a problem of identification of paths. One idea could be to generalize this reasoning to first-order logic. Thus, instead of considering polynomials we could consider formulae and instead of rewriting rules we could consider the rules of inference. Many questions arise here, can we draw a similar graph to the ones we have been drawing? If so, can we easily identify some paths up to reordering? If not, can some paths be identified modulo some 'well known' reasoning of first-order logic?

In another direction, it would be interesting to measure 'how similar' are two proofs, ignoring the possible similar proofs up to reordering. Bearing in mind our example, we would only need to consider the case of 'proofs' that cannot be dragged to one another without crossing blank cells. We propose to measure the distance between two proofs (paths in the graph) to be the minimum number of blank cells that are needed to cross to drag one path into the other. In our example, that means that the distance between the two paths corresponding to expressions E1 and E2 would be one.

Data accessibility. No supporting data here.

Authors' contributions. A.M. conceptualized the main idea of the article, critically revised it and gave his final approval of the version to be published. J.F.R. gave the worked-out example and developed the main idea of the article, wrote the drafts of the article and gave his final approval of the version to be published.

Competing interests. We declare we have no competing interests.

Funding. For both authors, this work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2013 (Centro de Matemática e Aplicações) and the project PTDC/MHC-FIL/2583/2014. The first author was also funded by the FCT project PTDC/MAT-PUR/31174/2017.

References

1. Thiele. R. 2003 Hilbert's twenty-fourth problem. *Am. Math. Mon.* **110**, 1–24. (doi:10.1080/00029890.2003.11919933)
2. Hilbert D. 1993 *Theory of algebraic invariants* (ed. B Sturmfels) (Transl. by RC Lauenbacher). Cambridge Mathematical Library. Cambridge, UK: Cambridge University Press.
3. Cox DA, Little J, O'Shea D. 2007 *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*, 3rd edn. Undergraduate Texts in Mathematics. Secaucus, NJ: Springer-Verlag, New York Inc.
4. Eisenbud D. 1995 *Commutative algebra: with a view toward algebraic geometry*. Graduate Texts in Mathematics, 150. London, UK: Springer-Verlag New York.
5. Bokut LA, Chen Y. 1980 Gröbner-Shirshov bases and their calculation. *Bull. Math. Sci.* **4**, 325–395. (doi:10.1007/s13373-014-0054-6)
6. Huet G. 1980 Confluent reductions: abstract properties and applications to term rewriting systems. *J. Assoc. Comput. Mach.* **27**, 797–821. (doi:10.1145/322217.322230)
7. Bachmair L. 1991 *Canonical equational proofs (Progress in theoretical computer science)*. Basel, Switzerland: Birkhäuser.
8. Wiegand R. 2006 What is . . . a syzygy? *Notices Am. Math. Soc.* **53**, 456–457.