Math.Comput.Sci. (2016) 10:5–25
DOI 10.1007/s11786-016-0246-4

Mathematics in Computer Science

CrossMark

On the Unavoidable Uncertainty of Truth in Dynamic Geometry Proving

Francisco Botana · Tomas Recio

Received: 10 July 2015 / Revised: 11 January 2016 / Accepted: 21 January 2016 / Published online: 5 March 2016

Abstract The aim of this note is to discuss some issues posed by the emergency of universal interfaces able to decide on the truth of geometric statements. More specifically, we consider a recent GeoGebra module allowing general users to verify standard geometric theorems. Working with this module in the context of Varignon's theorem, we were driven—by the characteristics of the GeoGebra interface—to perform a quite detailed study of the very diverse fate of attempting to automatically prove this statement, when using two different construction procedures. We highlight the relevance—for the theorem proving output—of expression power of the dynamic geometry interface, and we show that the algorithm deciding about the truth of some—even quite simple—statements can fall into a not true and not false situation, providing a source of confusion for a standard user and an interesting benchmark for geometers interested in discovering new geometric facts.

Keywords Dynamic geometry · Automated theorem proving · GeoGebra · Varignon theorem

Mathematics Subject Classification Primary 68T15; Secondary 68W30

1 Introduction

In September 2014, GeoGebra was the first well spread dynamic geometry program to include an automatic theorem proving feature. Thus, it is also the first time that some achievements from a half century history of automated deduction in geometry (ADG) research, are actually exposed to a global customer through an open use, moving away from university labs and controlled learning experiments.

The GeoGebra Theorem Proving feature can be roughly described as follows. Each of the steps of a geometric construction is, first, performed by using different GeoGebra graphic commands and, then, internally translated

F. Botana (✉)
Depto. de Matemática Aplicada I, University of Vigo, EE Forestal, 36005 Pontevedra, Spain
e-mail: fbotana@uvigo.es

T. Recio
Depto. de Matemáticas, Estadística y Computación, University of Cantabria, Avenida Los Castros, 39071 Santander, Spain
e-mail: tomas.recio@unican.es

into algebraic terms, by following an a priori program-established geometry/algebra dictionary. Next, on this construction, the user can formulate different queries (are collinear points X, Y, Z?, are parallel lines r, s?, …and the like) from a given list of *relation questions*, that are also automatically translated by the program into algebraic terms. Then, following different criteria, constraints and heuristics, a collection of algebraic methods in ADG, some of them using the own GeoGebra symbolic computation features, some others connecting with an external server (see, for instance [1]) are activated and sequentially attempt dealing with the proposed question, until one of them eventually succeeds or the program yields a failure warning. In the successful case, the output is the grant/denial of the truth of the proposed statement, optionally including a list of non-degeneracy conditions for the validity of the proposition. Further details are provided in [2].

Moreover, it can be remarked that most of the algebraic methods implemented in the GeoGebra Prover portfolio follow an approach and terminology that has become standard after Chou [3]. In Sect. 2 we present a summary of the main definitions and results in this framework, as described in Recio and Vélez [5].

An interesting issue in this context, and one of the motivations for this paper, has to do with the existence of some statements that are neither generally true nor generally false (following the terminology of [5]). Roughly speaking, statements dealing with constructions that have multiple instances for a single value of the free parameters of the corresponding construction and which are neither true for all such instances nor false for all of them. Sect. 3 introduces a simple illustrative example,[1] where a geometric fact, neither generally true nor generally false if naively formulated, can be easily turned into a generally true one, by adding an "intuitive" and natural complementary hypothesis. Furthermore, it advances some proving related consequences when defining, in dynamic geometry, the midpoint of a segment.

Finally, Sect. 4 deals in detail with a surprisingly hard example: Varignon's theorem (Sect. 4.1) and its converse (Sect. 4.2). Actually, each of these statements is not hard in itself, but its proof becomes surprisingly hard when using a non standard formulation which seems simple. More precisely: both statements deal with properties of midpoints of some segments and it turns that, depending on some precise formulation of the concept of midpoint, the statements can greatly vary in many different aspects: truth, computing time, etc. (see Table 2). In fact, one of the formulations (Sect. 4.2.2) leads to a Converse Varignon theorem that is neither true nor false. This example also shows that—without performing a primary decomposition, something out of the scope of most dynamic geometry programs with automatic proving features—it is quite non trivial to guess complementary conditions to break off such undesirable status of confusion (not true, not false). It is, also, an opportunity for graduate students to use effective algebraic geometry tools to attempt discovering such complementary conditions, as achieved in Sect. 4.2.3.

As remarked above, all these reflections—on the hidden subtleties of the topic—can be specially interesting nowadays, in view of the current trend concerning the inclusion of proving features in dynamic geometry programs.

2 Automatic Theorem Proving Through Elimination and Refutation: Short Survey of Key Concepts

Before getting into details, let us provide the reader with a rough description of the theoretical framework we are working with. This is particularly important here because our claim on the unavoidable uncertainty of theorem proving pretends to be intrinsic, i.e., we say that uncertainty is unavoidable by whatever methods. Intrinsic, yes, but intrinsic to a precise theoretical formulation of the theorem proving problem.

Given a geometric statement ($H \implies T$), we are assuming that, by some means—not to be considered here—, the geometric elements involved in the hypotheses H and theses T have been automatically converted into an algebraic system of equations, as it is shown, for instance, in the algebraic window of GeoGebra, if one builds the elements of a given geometric statement by means of this dynamic geometry program. Say, a point in a circle is to be replaced by the equation showing that the coordinates of the point verify the equation of the circle; if we claim that three points are aligned, then we are actually dealing with the vanishing of some determinant, etc.

[1] Kept specially simple in this section because the purpose here is merely introductory, but one can show it hides some unexpected complications.

This paper is, among other things, about the importance of the geometric-algebraic translation; yet we consider that the translation is a previous step and our theoretical framework starts with the collections H, T of polynomial equations[2] representing hypotheses and theses.

Moreover, we also assume, as "a priori" provided, a maximum size collection of free coordinates involved in the algebraic description of the hypotheses. For example, if we are dealing with the three heights of an arbitrary triangle, it is clear there will be six free coordinates, corresponding to the three vertices of the given triangle; the heights will be represented by linear equations on some new variables (say, x, y), with coefficients involving some expression on the coordinates of the vertices. If we now claim that the three heights meet at one point, we can consider as the hypotheses the set of equations of two of the heights. Then we will state as the thesis the fact that the solution of these two equations in two variables also satisfies the equation of the third height. In total there will be six plus two variables in the hypotheses, but only six of them will be free, since x, y will be determined as the intersection of the two given heights. The thesis will be that this precise point (x, y) verifies the equation of the other height, i.e., an equation in x, y with coefficients polynomial in the coordinates of the vertices.

Of course, this is just one possible way of thinking about a figure consisting on a triangle and two of its heights. We could also consider that this figure means, instead, that we are given two free points (the two vertices of the triangle where the two given heights are built from) and an extra free point, with coordinates (x, y), and then, two lines passing from one of the vertices and point (x, y). Since these two lines must be the heights of the given triangle, we know the third vertex is totally determined. What we want to emphasize is the fact that the assignment of free variables in a construction is something "user-driven", although in most cases there is a quite obvious assignment, related, for instance, to the sequence of steps in the performed geometric construction. Anyway, we just want to recall the reader that what we are presenting here is a very rough description of the whole framework. See [3–5] for a much more precise formulation and a quite sophisticated discussion on the choice and requirements for the free variables.

In summary, the precise input for our automatic proving problem is a collection $H(x_1, \ldots, x_n)$ of equations describing the hypotheses and an equation $T(x_1, \ldots, x_n)$ describing the (single) thesis; the required modifications for the case of several simultaneous theses are quite straightforward. A selected subset of variables $\{x_1, \ldots, x_r\}$ is also provided, as the set of free coordinates ruling our hypotheses, see Definition 2.1 below. And the truth of the given statement follows if every solution—over some given field K, say $\mathbb{R}$ or $\mathbb{C}$—of the system H is also a solution for T. Yet, it is often the case that most statements that we consider as valid fail because of the unexpected behaviour of our geometric construction (or of our algebraic translation) in some limit cases, e.g., if a triangle collapses to a line, what will happen to its heights? Thus, for automatic theorem proving, we usually consider a weaker concept of truth of a given statement, namely, a statement is said to be generally true if for "almost all" placements of the free variables $\{x_1, \ldots, x_r\}$ it happens that the theses $T(x_1, \ldots, x_n)$ hold over all the corresponding points in the hypotheses variety $H(x_1, \ldots, x_n)$. The question is, first, to propose an acceptable measure for estimating how big "almost all" should be considered. And, then, when an statement turns to be generally true, to automatically provide some extra hypotheses so that the statement would turn absolutely true under these new conditions.

Since we are working with algebraic equations, a natural answer to the first issue is to consider "small" those elements of the set of K^r that are constrained to verify an algebraic equation, say, those lying within a hyperplane. We think it is a natural proposal because (a) it is measure or probability zero, for instance, if we think of $K = \mathbb{R}$ or $\mathbb{C}$, (b) it makes easy to algebraically describe its complement, i.e., the points out of the exceptional set where the statement fails are those that verify some algebraic in equation, stating that such points are not in some exceptional hyperplane.

Once we agree in this setting, its development is straightforward algebraic geometry. Let us assume, for simplicity, that we are working on an algebraically closed field K, such as $\mathbb{C}$. Then, the condition on the "small size" of the failing set of free variables is equivalent to the fact that the elimination of the non-free variables on the set of equations $\{H = 0 \text{ and } T \neq 0\}$ (the set of geometric instances that do not verify the theses) is non zero; it means that the failing cases can be "wrapped" around by some non-zero equation in the free variables. Any element of

[2] Even if we are dealing with equations H, T, we do not represent them as $H = 0$, $T = 0$, following Maple's notation.

this non-zero elimination set provides—by considering its complement—moreover, an extra hypothesis for the statement to be absolutely true. See Definition 2.2 below for a more precise statement.

The introduction of the concept of generally true statements can be an adequate reaction to the problems arising with the algebraic translation of a given geometric statement because of the existence of degenerate instances. But there is another problem associated to the algebro-geometric dictionary. It is often the case that for a concrete value of the free coordinates ruling our construction there is more than one set of depending variables verifying the hypotheses. See, for instance, the example in Sect. 3 of this paper, where t_1, t_2, t_3, t_4 are free variables and each couple $E(e_1, e_2)$, $F(f_1, f_2)$ is defined as solution of a system of two, degree two, bivariate equations with coefficients polynomials in $\{t_1, t_2, t_3, t_4\}$. Both solutions can not be distinguished by means of equations, but only one pair (E, F, with $E \neq F$) corresponds to the construction we are thinking of. Thus the statement in Sect. 3 is not generally true, since it does not hold that H is included in T for all solutions of H over any values of the free variables, such as the points of H verifying the equality $E = F$.

Moreover, it happens that this statement can not be labeled as generally false, either. By a "generally false" statement we understand one that follows the definition of generally true in a converse way. That is, we say the implication $H \Rightarrow T$ is generally false if, for almost all values of the free variables, all of the corresponding solutions of H are not included in T. Again, checking if a statement is generally false involves just deciding whether the elimination of the non-free variables on the set of equations $\{H = 0 \text{ and } T = 0\}$ (the set of geometric instances that do verify the theses) is non zero, "wrapping around into a small set, the values of the free variables where all hypotheses verify the thesis. In our example of Sect. 3, it happens that it is not generally false, since for arbitrary values of the free variables, the solutions of the hypotheses verifying $E \neq F$, do hold the thesis.

In summary, what we want to emphasize here is the fact that this framework is not depending on the algebraic method that we might use for elimination. It has nothing to do with using tools such as Groebner bases, triangular or characteristic sets. Once this framework is assumed, it yields unavoidably—and not method dependent—, as explained in the paragraphs following Definition 2.2 below, to the consideration of the irreducible components of the hypotheses variety. In fact, the introduction of this ellaborated algebro-geometric concept (irreducible components) seems alien to the very down-to-earth reflections we were taking into consideration till this point. We can, in fact, avoid introducing the idea of component if the statement turns to be generally true. This is precisely what we have done till now. Or if it is generally false (a symmetric notion, considering statements that are false except for a small set of values of the free variables). But it does not help if a statement turns to be, simultaneously, not generally true and not generally false as in the very trivial example of Sect. 3 and in many much more relevant others, as shown in the following sections.

In general, being simultaneously not generally true and not generally false, as in the above mentioned example, is something that refers to a behavior that corresponds to all values of the free variables, so it can not be dilucidated via a finer analysis of such variables, i.e., by adding new conditions of degeneracy. But then, it means we can not go any further by considering "all" solutions of H, "all" possible constructions for a given position of the free variables. We need to make a finer analysis and consider the truth or falsity of each isolated solution, i.e., the different behavior of the components of H. And this what we mean, in our paper, as unavoidable, because of its computational cost. For instance, after more than one hour with Maple 17, on a Mac Book Pro 2.5 GHz Intel Core i7, we have failed obtaining the primary decomposition of the hypotheses ideal of the example in Sect. 3, in order to distinguish the true and false components and in order to avoid "guessing" the need to add the intuitively obvious condition $E \neq F$, for the statement to hold generally true.

Once we have described our basic idea in a narrative style, let us introduce precisely the main notation and concepts. Let $H(x_1, \ldots, x_n)$ denote the collection of equations describing the hypotheses of a statement and let $T(x_1, \ldots, x_n)$ be the single thesis. The required modifications for the case of several simultaneous theses are quite straightforward. Let $I = (H, T * z - 1)$ be the ideal of hypotheses and negated thesis in $K[x_1, \ldots, x_n, z]$, where K is an algebraically closed field.

Definition 2.1 Let $\{x_1, \ldots, x_n\}$ be the collection of coordinates involved in the algebraic description of the hypotheses, with $\{x_1, \ldots, x_r\}$ taken as a maximum-size set of free variables for the hypotheses.

This means:

(a) the dimension (HilbertDimension) of the ideal of hypotheses is r, and
(b) the elimination in the hypotheses ideal of all the variables minus $\{x_1, \ldots, x_r\}$ yields 0, i.e., these variables are free modulo H.

Both conditions imply that:

(i) there is no polynomial relation in the ideal of hypotheses holding for the variables $\{x_1, \ldots, x_r\}$ alone, and
(ii) because of the concept of Ideal Dimension, r is the largest number of variables having that property with respect to the ideal of hypotheses. Therefore, for any extra variable x_m, with $m > r$, the elimination with respect to $\{x_1, \ldots, x_r, x_m\}$ is not 0, i.e. there is at least one non zero polynomial in the variables $\{x_1, \ldots, x_r, x_m\}$ belonging to the ideal H. See several detailed comments on this issue in Recio et al. [6].

Now, let us project the variety $V = V(H, T * z - 1) \subseteq K^{n+1}$ over the affine space of free variables K^r. Then, the Zariski closure of the projection $\pi(V)$ is the zero set $V_r \subseteq K^r$ of the elimination ideal $I_r = (H, T * z - 1) \cap K[x_1, \ldots, x_r]$. Let J be the ideal (H, T) in $K[x_1, \ldots, x_n]$ and let W be its zero set. Let J_r be the elimination ideal $J_r = (H, T) \cap K[x_1, \ldots, x_r]$ and let W_r be its zero set in K^r.

Definition 2.2 The statement $H \implies T$ is said to be *generally true* if $I_r \neq 0$, and *generally false* if $J_r \neq 0$.

It should be remarked that in the generally true case, the lifting of points in V_r to the zero set of H provides instances of the hypotheses where the statement fails. That is, values of $(x_1, \ldots, x_r)$ in I_r such that there is a value $(x_{r+1}, \ldots, x_n)$ verifying H and not T. But it could be also true that, for the same value of $(x_1, \ldots, x_r)$, there is a different value $(x_{r+1}, \ldots, x_n)$ verifying H and T. Obviously, the irreducible components of H yielding values where it holds that there are values of $(x_1, \ldots, x_r)$ in I_r such that there is a value $(x_{r+1}, \ldots, x_n)$ verifying H and not T, are irreducible components of H where the variables $(x_1, \ldots, x_r)$ do not remain independent, since the elements of I_r belong to this component. Thus, they are labeled as *degenerate*, since it is implicit some kind of intuition that geometrically sound constructions should be those where $(x_1, \ldots, x_r)$ are free... Same kind of reflection can be considered for the generally false case.

Finally we must recall that a statement is generally true if and only if the thesis holds over all irreducible components of the hypotheses variety that are non-degenerate: i.e. such that $(x_1, \ldots, x_r)$ remain independent modulo this irreducible component. It is generally false if and only if the thesis does not hold over any of the non-degenerate components [4, Propositions 1 and 2]. Both generally true and generally false cannot simultaneously happen, as it can be derived from the following proposition.

Proposition 2.3 *If a statement is generally true (resp. false), then it is not generally false (resp. true). In symbols,* $I_r \neq 0 \implies J_r = 0$, *and* $J_r \neq 0 \implies I_r = 0$.

Proof It is enough to show that it can not simultaneously happen generally true and generally false.

Let us prove that J_r and I_r can not be simultaneously not zero. In fact, assume they are both zero and let $g \in (H, T)$, $q \in (H, T * z - 1)$ be both non zero elements of $K[x_1, \ldots, x_r]$.Thus,

$$g = \text{ combination of } H + \text{multiple of } T,$$
$$q = \text{ combination of } H + \text{multiple of } (T * z - 1),$$

where combination of H is a way of expressing a sum of polynomials in n variables times elements of H; multiple of T [resp. $(T * z - 1)$] is a way of expressing a polynomial in n variables times T [resp. a polynomial in $n + 1$ variables times $(T * z - 1)$].

Replacing z by $1/T$ and multiplying by a suitable power of T, say T^m, the last equality turns to be

$$q * T^m = \text{ combination of } H.$$

Table 1 Summary of possibilities of being generally true or false

	J_r	I_r
Not generally true and not generally false	(0)	(0)
Generally true (and, thus, not generally false)	(0)	Not (0)
Generally false (and, thus, not generally true)	Not (0)	(0)

Analogously, the expression of g above can be rewritten as

multiple of $T =$ combination of $H - g$.

Thus,

(multiple of $T)^m =$ (combination of $H - g)^m =$ combination of $H + g^m$,

and hence

$q *$ (multiple of $T)^m = q *$ combination of $H + q * g^m$.

Since $q *$ (multiple of $T)^m =$ multiple of $q * T^m =$ combination of H, finally we arrive to

$$\text{combination of } H - q * \text{combination of } H = \text{combination of } H = q * g^m$$
$$= \text{ non zero element of } K[x_1, \ldots, x_r],$$

which is impossible, since we assume $x_1, \ldots, x_r$ to be free variables for H.

Thus, we see that both elimination ideals can not be simultaneously different from zero. Table 1 summarizes the possibilities.

Moreover, neither not generally true does imply being generally false, nor not generally false does imply being generally true, since there are examples of statements that are simultaneously not generally true and not generally false. A nice example is, precisely, the converse of Varignon, in the case of a particular algebraic interpretation of the concept of midpoint, see Sect. 4.2.

Thus, in order to decide if a statement is generally true or not generally true (beware, this is not the same as being generally false), all we have to do is to establish a procedure for deciding, given a polynomial ideal of hypotheses and negated theses, whether the result of eliminating in the ideal some variables, yields the zero ideal or not.

In conclusion, in our framework we identify the concept of *proving a statement* with *proving that it is generally true*. If it is not, then we would like to learn, first, if it is generally false, i.e., false everywhere it makes sense. If not, we will conclude that it is true over some relevant components and false over some other relevant components of the hypotheses variety. Then we will learn that there is some hidden important fact, holding just in some special cases, to be discovered with further computations and insight! See the second formulation of the converse of Varignon in Sect. 4.2.

3 A Simple Example

Let us consider the following construction: Given two points $A(t_1, t_2)$, $B(t_3, t_4)$, construct circle c_1 with center A and going through B, and circle c_2 with center B and going through A. Then, consider the intersection of c_1 and c_2, points $E(e_1, e_2)$ and $F(f_1, f_2)$.

Clearly the construction depends only on t_1, t_2, t_3, t_4, but E and F are not uniquely determined, since they are described as the solution of the system

$(x - t_1)^2 + (y - t_2)^2 - (t_3 - t_1)^2 - (t_4 - t_2)^2, (x - t_3)^2 + (y - t_4)^2 - (t_1 - t_3)^2 - (t_2 - t_4)^2.$

The expression of the coordinates for E, F, through the output of the Maple Solve command expressing the solutions of a second degree univariate polynomial equation, shows that there are two different, yet indistinguishable solutions when using polynomials to define or manipulate them:

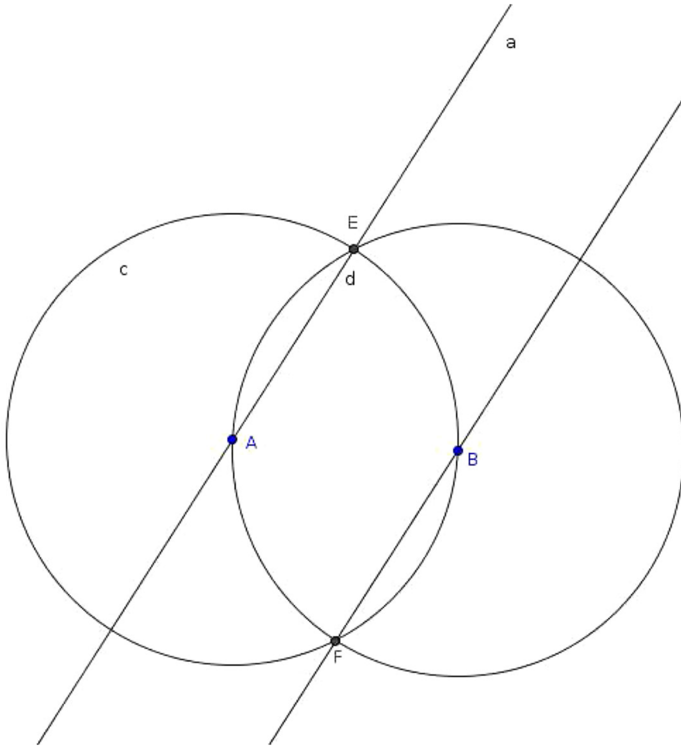

Fig. 1 A simple example

```
> solve({(x-t1)^2+(y-t2)^2-((t3-t1)^2+(t4-t2)^2),(x-t3)^2+(y-t4)^2
-((t1-t3)^2+(t2-t4)^2)},{x,y});

{y = RootOf(4_Z^2+(-4t2-4t4)_Z-3t1^2+6t3t1+t2^2+2t4t2-3t3^2+t4^2),
x = -(1/2)(2RootOf(4_Z^2+(-4t2-4t4)_Z-3t1^2+6t3t1+t2^2+2t4t2
-3t3^2+t4^2)t2-2RootOf(4_Z^2+(-4t2-4t4)_Z-3t1^2+6t3t1+t2^2+
2t4t2-3t3^2+t4^2)t4-t1^2-t2^2+t3^2+t4^2)/(t1-t3)}
```

Anyway, let us suppose that a standard user wants to prove the following thesis:

Line AE (a in Fig. 1) is parallel to line BF (resp. b), i.e.,

$$(e_1 - t_1)(f_2 - t_4) - (e_2 - t_2)(f_1 - t_3) = 0.$$

Then, this user will probably consider—within the framework described in Sect. 2—that one has to eliminate all but the free variables t_1, t_2, t_3, t_4 in the ideal generated by the following hypotheses:

$E(e_1, e_2)$ is in the intersection of circle c_1 and circle c_2,

$$(e_1 - t_1)^2 + (e_2 - t_2)^2 - ((t_3 - t_1)^2 + (t_4 - t_2)^2),$$
$$(e_1 - t_3)^2 + (e_2 - t_4)^2 - ((t_1 - t_3)^2 + (t_2 - t_4)^2),$$

$F(f_1, f_2)$ is in the intersection of circle c_1 and circle c_2,

$$(f_1 - t_1)^2 + (f_2 - t_2)^2 - ((t_3 - t_1)^2 + (t_4 - t_2)^2),$$
$$(f_1 - t_3)^2 + (f_2 - t_4)^2 - ((t_1 - t_3)^2 + (t_2 - t_4)^2),$$

and by the negation of the thesis

$$((e_1 - t_1)(f_2 - t_4) - (e_2 - t_2)(f_1 - t_3))z - 1.$$

Surely, this naive user is not going to reflect on the particular lack of determination for E and F in the construction and, thus, he/she might be perplexed to find the following Maple output

```
> with(PolynomialIdeals): EliminationIdeal(<(e1-t1)^2+(e2-t2)^2-(
(t3-t1)^2+(t4-t2)^2),(e1-t3)^2+(e2-t4)^2-((t1-t3)^2+(t2-t4)^2),
(f1-t1)^2+(f2-t2)^2-((t3-t1)^2+(t4-t2)^2),(f1-t3)^2+(f2-t4)^2-(
(t1-t3)^2+(t2-t4)^2), ((e1-t1)*(f2-t4)-(e2-t2)*(f1-t3))*z-1>,
{t1,t2,t3,t4});

                              <0>
```

showing that this statement is not generally true. Then, to check if the statement is generally false, one computes the elimination of the hypotheses + thesis ideal, yielding

```
> EliminationIdeal(<(e1-t1)^2+(e2-t2)^2-((t3-t1)^2+(t4-t2)^2),(e1-
t3)^2+(e2-t4)^2-((t1-t3)^2+(t2-t4)^2),(f1-t1)^2+(f2-t2)^2-((t3-
t1)^2+(t4-t2)^2),(f1-t3)^2+(f2-t4)^2-((t1-t3)^2+(t2-t4)^2), ((e1
-t1)*(f2-t4)-(e2-t2)*(f1-t3))>,{t1,t2,t3,t4});

                              <0>
```

This shows that the statement is also not generally false. The reason behind this seemingly strange behavior is that the statement is true or false depending on the particular choices of E, F as solutions of the defining system of equations

$$(x-t_1)^2+(y-t_2)^2-((t_3-t_1)^2+(t_4-t_2)^2)=0,$$
$$(x-t_3)^2+(y-t_4)^2-((t_1-t_3)^2+(t_2-t_4)^2)=0.$$

In fact, by just adding the condition

$$((e_1-f_1)t-1)((e_2-f_2)s-1)=0,$$

that means

$$((e_1-f_1)t-1)=0 \text{ or } ((e_2-f_2)s-1)=0,$$

and this is equivalent to

$$e_1 \neq f_1 \quad \text{or} \quad e_2 \neq f_2,$$

i.e., to

$$E \neq F,$$

it happens that the statement is generally true, as checked by

```
> EliminationIdeal(<(e1-t1)^2+(e2-t2)^2-((t3-t1)^2+(t4-t2)^2),(e1-
t3)^2+(e2-t4)^2-((t1-t3)^2+(t2-t4)^2),(f1-t1)^2+(f2-t2)^2-((t3-
t1)^2+(t4-t2)^2),(f1-t3)^2+(f2-t4)^2-((t1-t3)^2+(t2-t4)^2), ((e1
-f1)*t-1)*((e2-f2)*s-1), ((e1-t1)*(f2-t4)-(e2-t2)*(f1-t3))*z-1>,
{t1,t2,t3,t4});
                              <t3-t1,t4-t2>
```

under the non-degeneracy condition

$$t_1 \neq t3 \quad \text{or} \quad t_2 \neq t_4,$$

or, equivalently,

$$\text{not } (t_1 = t_3 \;\text{ and }\; t_2 = t_4),$$

that is,

not $(A = B)$.

As shown in Table 1, then the statement is not generally false in this case:

```
> EliminationIdeal(<(e1-t1)^2+(e2-t2)^2-((t3-t1)^2+(t4-t2)^2),(e1-
t3)^2+(e2-t4)^2-((t1-t3)^2+(t2-t4)^2),(f1-t1)^2+(f2-t2)^2-((t3-
t1)^2+(t4-t2)^2),(f1-t3)^2+(f2-t4)^2-((t1-t3)^2+(t2-t4)^2),((e1-
f1)*h-1)*((e2-f2)*s-1),((e1-t1)*(f2-t4)-(e2-t2)*(f1-t3))>,{t1,
t2,t3,t4});
                              <0>
```

4 The Varignon Theorem and Its Converse

The preceding section highlights—in a deceivingly simple instance—the need to search for complementary and reasonable hypotheses for a given statement to hold generally true or generally false. In that example, it seems quite “intuitive”, to require that not $(E = F)$, and, then, to add as a non-degeneracy condition, that not $(A = B)$ for the statement to be true. But it is not obvious at all, in many statements, what could be some of the implicit requirements that should be added to yield a clear conclusion in the given context, i.e., to turn the given statement to be generally true or generally false. In fact, the straightforward approach to solve this issue involves performing a very costly and usually hard to interpret, geometrically, primary decomposition, and then determining which are the components where the thesis hold and which are those where the thesis fails.

This section introduces one particular illustrative example of this problem: the Varignon theorem and its converse. Varignon theorem states that the midpoints of the sides of an arbitrary quadrilateral form a parallelogram. Since different quadrilaterals can be associated, considering the midpoints of their sides, with the same parallelogram, Varignon’s converse statement declares that the vertices of an arbitrary parallelogram are the midpoints of the sides of a quadrilateral with an arbitrarily chosen vertex. We remark that the direct and converse statements involve, in different ways, the definition of midpoint of a segment. In the direct statement we need to describe the midpoint M of some given points A, B. In the converse statement, given M and A, we need to describe B so that M is the midpoint of segment AB.

The standard treatment of both cases is to consider

(a) the coordinates of M as $(A+B)/2$ in the direct Varignon theorem, and, thus, in the converse case, the coordinates of B as $2M - A$.

A different option is to consider

(b) the midpoint M as a point equidistant of A and B and in the line AB; accordingly, point B as the intersection of the line MA with the circle with center M and radius MA.

We might think that it should be evident that the first option is the correct one and the second is by far too artificial. But we have remarked, at the beginning of this note, that we want to consider the case of automated theorem proving (ATP) in the context of the popularization of dynamic geometry software, such as GeoGebra. And, in this particular software, it happens that there is a built-in tool for constructing the midpoint M of a segment AB, with algebraic translation $M = (A + B)/2$, but the determination of B such that M is the midpoint of AB is not yet provided by the program and requires a specific construction, such as building line MA and its intersection with the circle of center M and radius MA through, for instance, GeoGebra’s compass tool.

In the following subsections we explore the consequences, both in the direct and converse Varignon statement, of choosing each of the above mentioned options for describing—in the hypotheses of the direct and the converse Varignon—the midpoint M of some given segment AB, and the point B such that a given point M is the midpoint of AB, for some other given point A.

Table 2 Output of our exploration

	Option (a) for describing hypotheses	Option (b) for describing hypotheses
Varignon direct	Generally true (in fact, no non-degeneracy condition)	Generally true (but with several non-degeneracy conditions)
	Easy to compute	Several hours computation of elimination
		Use of alternative, faster but less canonical algorithm (e.g. Maple's *solve* command) requires understanding the algebraic geometry associated to the statement (lifting construction free points to the hypotheses variety, deciding how many liftings can be done, deciding if some/all/none of them verify the theses)
Varignon converse	Generally true (but with one non-degeneracy condition)	Not generally true, not generally false
	Easy to compute	Several hours computation of elimination
	The generally true conclusion can be, alternatively, deduced from the non-generally false conclusion, through the use of alternative, faster but less canonical algorithm (e.g. Maple Solve command) requires understanding the algebraic geometry associated to the statement (lifting construction free points to the hypotheses variety, deciding how many liftings can be done, deciding if some/all/none of them verify the theses)	Use of alternative, faster but less canonical algorithm (e.g. Maple's *solve* command) requires understanding the algebraic geometry associated to the statement (lifting construction free points to the hypotheses variety, deciding how many liftings can be done, deciding if some/all/none of them verify the theses)
		This alternate process, requiring exploring the algebraic geometry of the involved statement, provides an excellent benchmark for graduate students, requiring several challenging algebra-geometry interpretations and yielding to, perhaps, new geometric facts

Let us advance that the output of our exploration is rather diverse, as detailed in Table 2. Because of this diversity, we do not consider necessary to explore new options (such as defining the point B by considering the symmetric of A with respect to M, etc.).

Note that the case Converse Varignon | option (b) is the currently natural one for users of ATP tools embedded in dynamic geometry systems such as GeoGebra, but its output (not generally true and not generally false) could be rather disappointing for a standard user. One should not claim that this is a rather artificial example: we have all learned from the past that the algebraic approach to ATP in geometry involves uncontrolled (by the user) problems with the algebraic translations that could rise in the most unexpected contexts. In fact, the origin of this paper has been the search for an answer to what we, not "naive" users at all, obtained when proceeding, in the most direct way, to address the Converse Varignon statement with the current GeoGebra Proving tool.

4.1 The Direct Varignon Theorem

Using option (a) for the concept of midpoint M between P and Q as the point $M = (P + Q)/2$, that is, the coordinates of Q are those of $M + (M - P)$, consider a quadrilateral $A(0, 0)$, $B(1, 0)$, $C(t_1, t_2)$, $D(d_1, d_2)$, and the side midpoints $E(t_3, t_4)$, $F(f_1, f_2)$, $I(i_1, i_2)$, $K(k_1, k_2)$. The construction is a 4-dimensional one, with free variables the coordinates of C, D, and the hypotheses ideal is

$$\langle 2t_3 - 1, t_4, 2f_1 - t_1 - 1, 2f_2 - t2, 2i_1 - t_1 - d_1), 2i_2 - t_2 - d_2, 2k_1 - d_1, 2k2 - d_2\rangle.$$

The theses state that $EFIK$ is a parallelogram (Fig. 2), that is,

$EF||IK : (f_1 - t_3)(k_2 - i_2) - (f_2 - t_4)(k_1 - i_1) = 0$, and
$EK||FI : (k_1 - t_3)(f_2 - i_2) - (k_2 - t_4)(f_1 - i_1) = 0$.

The elimination of the free variables (t_1, t_2, d_1, d_2) in the ideal of hypotheses and negation of the theses

```
> EliminationIdeal(<2*t3-1,t4,2*f1-(t1+1),2*f2-(t2),2*i1-(t1+d1),
2*i2-(t2+d2),2*k1-(0+d1),2*k2-(0+d2), (((f1-t3)*(k2-i2)-(f2-t4)*
(k1-i1))*z-1)*(((k1-t3)*(f2-i2)-(k2-t4)*(f1-i1))*t-1)>,{t1,t2,
d1,d2});
                                <1>
```

shows that the only conditions for the statement to be false is $1 = 0$, that is, never. It is generally true, each of the theses is a combination of the hypotheses:

```
> (f1-t3)*(k2-i2)-(f2-t4)*(k1-i1) in <2*t3-1,t4,2*f1-(t1+1),2*f2-
(t2),2*i1-(t1+d1), 2*i2-(t2+d2),2*k1-(0+d1),2*k2-(0+d2)>;
                                true
> (k1-t3)*(f2-i2)-(k2-t4)*(f1-i1) in <2*t3-1,t4,2*f1-(t1+1),2*f2-
(t2),2*i1-(t1+d1), 2*i2-(t2+d2),2*k1-(0+d1),2*k2-(0+d2)>;
                                true
```

Now, using the concept of midpoint M between P and Q as the center M of a circle of radius MP that intersects the line MP in the other point Q, the ideal of hypotheses is

$< t_3^2 - (1-t_3)^2, (1-f_1)^2 + f_2^2 - (t_1 - f_1)^2 - (t_2 - f_2)^2, (t_1 - i_1)^2 + (t_2 - i_2)^2 - (d_1 - i_1)^2 - (d_2 - i_2)^2,$
$(d_1 - k_1)^2 + (d_2 - k_2)^2 - k_1^2 - k_2^2, t_4, (1 - f_1)(t_2 - f_2) + f_2(t_1 - f_1), (d_1 - i_1)(t_2 - i_2) - (d_2 - i_2)(t_1 - i_1),$
$-(d_1 - k_1)k_2 + (d_2 - k_2)k_1 >$.

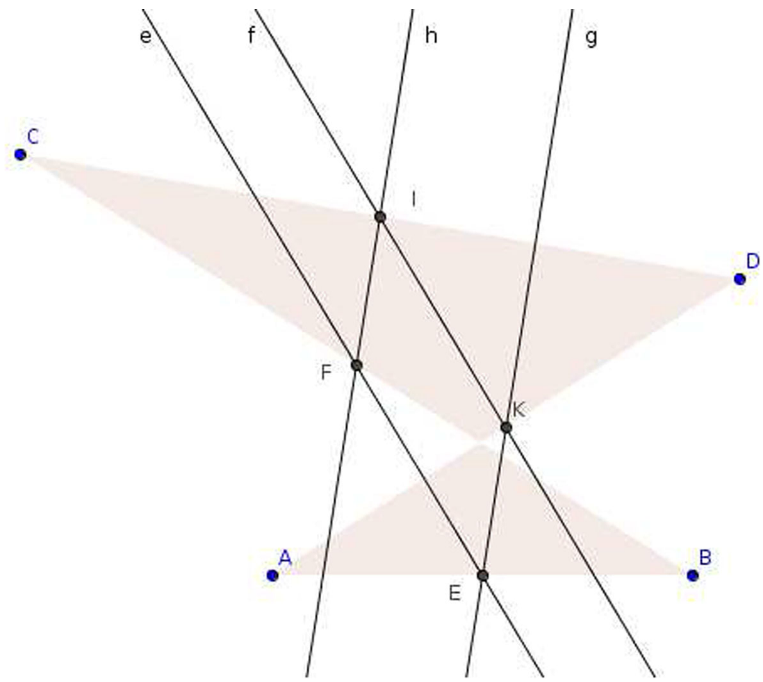

Fig. 2 Varignon theorem for a non convex quadrilateral

It is again a 4-dimensional construction, with free variables the coordinates of C, D. Solving with respect to the free variables shows that for every position of C, D, there is a unique value of the depending variables:

$$t_4 = 0,\ \ f_2 = \frac{1}{2}t_2,\ \ i_1 = \frac{1}{2}d_1 + \frac{1}{2}t_1,\ \ k_2 = \frac{1}{2}d_2,\ \ i_2 = \frac{1}{2}d_2 + \frac{1}{2}t_2,\ \ t_3 = \frac{1}{2},\ \ f_1 = \frac{1}{2} + \frac{1}{2}t_1,\ \ k_1 = \frac{1}{2}d_1.$$

With the second option to describe midpoints the theses remain unchanged, and the elimination ideal of hypotheses and negation of theses is

$$\langle (d_1^2 + d_2^2)(-2t_1 + 1 + t_1^2 + t_2^2)(t_1^2 - 2t_1d_1 + d_2^2 + d_1^2 - 2t_2d_2 + t_2^2)\rangle,$$

meaning that the statement is generally true. That is, the projection, over the free variables, of the variety where the theses do not hold is reduced to (the geometric interpretation is done over the real plane)

- $d_1^2 + d_2^2 = 0$, i.e. $d_1 = d_2 = 0$, a degenerate case $D = A$, or
- $-2t_1 + 1 + t_1^2 + t_2^2 = 0$, i.e. $t_1 = 1$, $t_2 = 0$, so $C = B$, or
- $t_1^2 - 2t_1d_1 + d_2^2 + d_1^2 - 2t_2d_2 + t_2^2$, i.e. $d_1 = t_1$ and $d_2 = t_2$, a degenerate case $D = C$.

Moreover, each of the theses belongs to the ideal of the hypotheses, extended with the negation of

$$(d_1^2 + d_2^2)(-2t_1 + 1 + t_1^2 + t_2^2)(t_1^2 - 2t_1d_1 + d_2^2 + d_1^2 - 2t_2d_2 + t_2^2) = 0,$$

as it can be checked by

```
> (f1-t3)*(k2-i2)-(f2-t4)*(k1-i1) in <(0-t3)^2+(0-t4)^2-((1-t3)^2+
(0-t4)^2),t4,(1-f1)^2+(0-f2)^2-((t1-f1)^2+(t2-f2)^2),(1-f1)*(t2-
f2)-(0-f2)*(t1-f1), (t1-i1)^2+(t2-i2)^2-((d1-i1)^2+(d2-i2)^2),
(d1-i1)*(t2-i2)-(d2-i2)*(t1-i1),(d1-k1)^2+(d2-k2)^2-((0-k1)^2+(0
-k2)^2),(d1-k1)*(0-k2)-(d2-k2)*(0-k1),(d1^2+d2^2)*(-2*t1+1+t1^2+
t2^2)*(t1^2-2*t1*d1+d2^2+d1^2-2*t2*d2+t2^2)*z-1>;
                              true
```

and

```
> (k1-t3)*(f2-i2)-(k2-t4)*(f1-i1) in <(0-t3)^2+(0-t4)^2-((1-t3)^2+
(0-t4)^2),t4,(1-f1)^2+(0-f2)^2-((t1-f1)^2+(t2-f2)^2),(1-f1)*(t2-
f2)-(0-f2)*(t1-f1), (t1-i1)^2+(t2-i2)^2-((d1-i1)^2+(d2-i2)^2),
(d1-i1)*(t2-i2)-(d2-i2)*(t1-i1),(d1-k1)^2+(d2-k2)^2-((0-k1)^2+(0
-k2)^2),(d1-k1)*(0-k2)-(d2-k2)*(0-k1),(d1^2+d2^2)*(-2*t1+1+t1^2+
t2^2)*(t1^2-2*t1*d1+d2^2+d1^2-2*t2*d2+t2^2)*z-1>;
                              true
```

thus confirming that this extended statement is generally true. Even more, the computation

```
> 1 in <(0-t3)^2+(0-t4)^2-((1-t3)^2+(0-t4)^2),t4,(1-f1)^2+(0-f2)^2
-((t1-f1)^2+(t2-f2)^2),(1-f1)*(t2-f2)-(0-f2)*(t1-f1), (t1-i1)^2+
(t2-i2)^2-((d1-i1)^2+(d2-i2)^2),(d1-i1)*(t2-i2)-(d2-i2)*(t1-i1),
(d1-k1)^2+(d2-k2)^2-((0-k1)^2+(0-k2)^2),(d1-k1)*(0-k2)-(d2-k2)*
(0-k1),(((f1-t3)*(k2-i2)-(f2-t4)*(k1-i1))*t-1)*(((k1-t3)*(f2-i2)
-(k2-t4)*(f1-i1))*s-1),(d1^2+d2^2)*(-2*t1+1+t1^2+t2^2)*(t1^2-2*
t1*d1+d2^2+d1^2-2*t2*d2+t2^2)*z-1>;
                              true
```

shows that 1 is a combination of the ideal of extended hypotheses and the negation of the theses, remarking the absolute truth of this extended statement.

As an alternative, but less rigorous proof, since computing the above elimination ideal took several hours, it can be checked that the values of the depending coordinates do verify the theses:

```
> subs(t3 = 1/2, t4 = 0, i2 = (1/2)*d2+(1/2)*t2, i1 = (1/2)*d1+
(1/2)*t1, f2 = (1/2)*t2, f1 = 1/2+(1/2)*t1, k2 = (1/2)*d2, k1 =
(1/2)*d1,{(f1-t3)*(k2-i2)-(f2-t4)*(k1-i1),(k1-t3)*(f2-i2)-(k2-
t4)*(f1-i1)}):expand(%);
                             0
```

4.2 The Converse Varignon Theorem

4.2.1 Converse Varignon: Option (a)

Using option (a) for the concept of midpoint, let us consider a parallelogram $A(0,0)$, $B(1,0)$, $C(t_1,t_2)$, $D(d_1,d_2)$ such that $DC||AB$ and $AD||BC$ (Fig. 3), that is, $d_1t_2 = d_2(t_1-1)$, $d_2 = t_2$.

Fix a free point $E(t_3,t_4)$, build the line EA, and, on this line, the point $F(f_1, f_2)$ such that A is the midpoint of segment EF, so $t_3 = -f_1$, $t_4 = -f_2$. Analogously, point $I(i_1,i_2)$ such that B is FI midpoint ($f_1+i_1 = 2$, $f_2 = -i_2$), and point $K(k_1,k_2)$ such that C is IK midpoint ($k_1+i_1 = 2t_1$, $k_2+i_2 = 2t_2$).

Thus, the eight hypotheses are

$$d_1t_2-d_2(t_1-1),\ d_2-t_2,\ t_3+f_1,\ t_4+f_2,\ f_1+i_1-2,\ f_2+i_2,\ k_1+i_1-2t_1,\ k_2+i_2-2t_2.$$

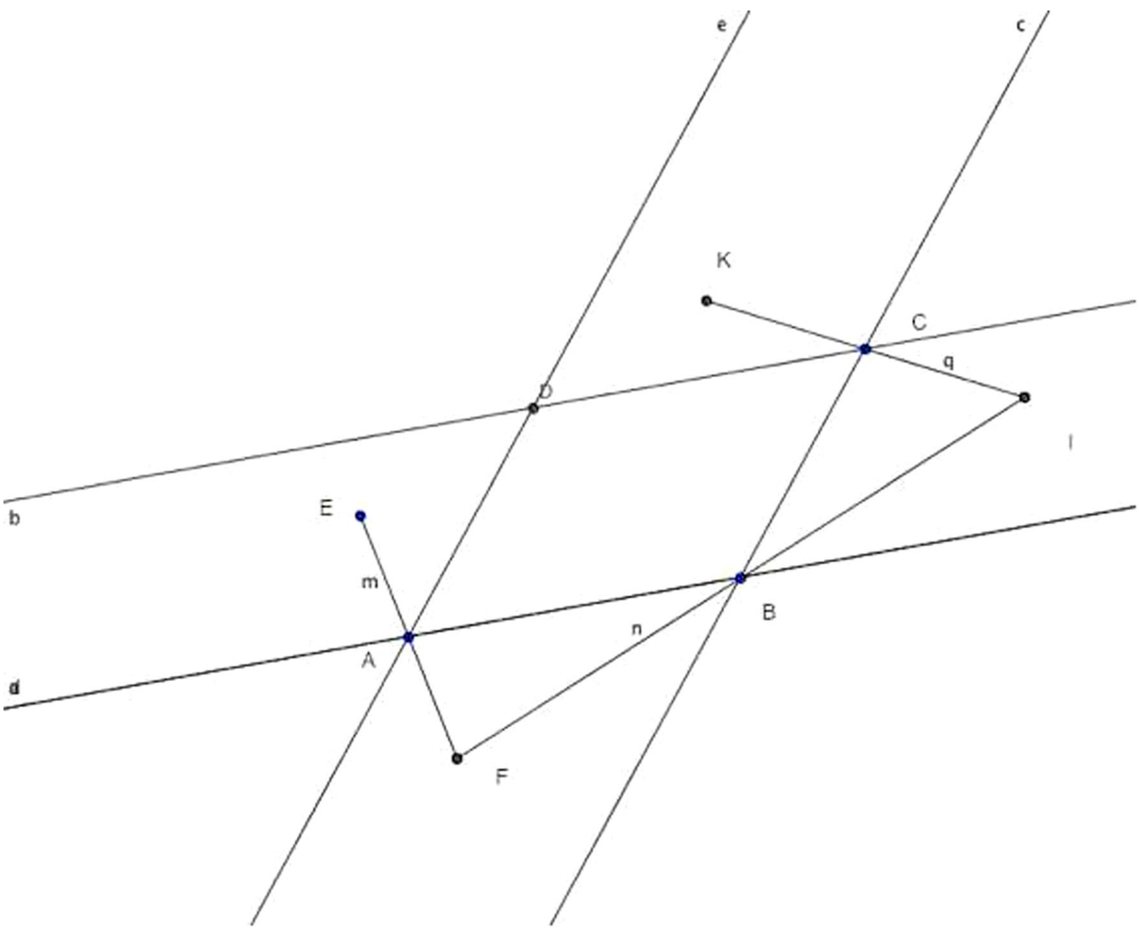

Fig. 3 Using option (a) for converse Varignon theorem

The construction involves the 12 variables $\{t_1, t_2, t_3, t_4, d_1, d_2, f_1, f_2, i_1, i_2, k_1, k_2\}$; the dimension of the construction is 4

```
> HilbertDimension(<(d1)*(t2)-d2*(t1-1),(d2-t2),t3+f1, t4+f2,f1+i1
-2, f2+i2, k1+i1-2*t1, k2+i2-2*t2 >,{t1,t2,t3,t4, d1,d2, f1,f2,
i1,i2,k1,k2});
                              4
```

The first four variables are free

```
> EliminationIdeal(<(d1)*(t2)-d2*(t1-1),(d2-t2),t3+f1, t4+f2,f1+i1
-2, f2+i2, k1+i1-2*t1, k2+i2-2*t2>,{t1,t2,t3,t4});
                              <0>
```

while the remaining ones are uniquely determined if values are assigned to the free variables $\{t_1, t_2, t_3, t_4\}$

```
> solve({(d1)*(t2)-d2*(t1-1),(d2-t2),t3+f1, t4+f2,f1+i1-2, f2+i2,
k1+i1-2*t1, k2+i2-2*t2},{d1,d2, f1,f2, i1,i2,k1,k2});
{f1 = -t3, k1 = -2 + 2t1 - t3 , f2 = -t4, i2 = t4, d2 = t2, d1 =
t1 - 1, k2 = 2t2 - t4 , i1 = 2 + t3}
```

We claim that D is the midpoint of segment EK, i.e. the simultaneous vanishing of $k_1+t_3-2d_1$ and $k_2+t_4-2d_2$. We add these two theses to the hypotheses and see what are the consequences of this claim:

```
> EliminationIdeal(<(d1)*(t2)-d2*(t1-1),(d2-t2),t3+f1, t4+f2,f1+i1
-2, f2+i2, k1+i1-2*t1, k2+i2-2*t2, k1+t3-2d1, k2+t4-2d2>, {t1,t2,
t3,t4});
                              <0>
```

So, the consequence (on the space of free points) of our claim is the whole space, i.e. all values of t_1, t_2, t_3, t_4 are on the closure of the projection of the set of points verifying the hypotheses and the theses. In other words, for almost every value of t_1, t_2, t_3, t_4 there is a value of $d_1, d_2, f_1, f_2, i_1, i_2, k_1, k_2$ verifying that the midpoints of the quadrilateral $EFIK$ form the parallelogram $ABCD$.

In principle, from the last elimination result, one can just conclude that the statement is not generally false, because for almost each value of the free parameters there is a value of the depending variables $d_1, d_2, i_1, i_2, k_1, k_2$, so that the statement is true. But it could happen, *in principle*, that there is also a different value of $d_1, d_2, f_1, f_2, i_1, i_2, k_1, k_2$ where the statement is false; we know in this particular case that this can not happen, since there is only one value of the depending variables for each value of the free ones; but this kind of argument depends on the `Solve` command, which is not easy to handle or canonical, in general, and thus it is not used as a standard in the dynamic geometry proving routines.

So, let us see by a different, more general method, that our statement is generally true by considering the collection of hypotheses and the negation of the theses:

$$((k_1 + t_3 - 2d_1)z - 1)((k_2 + t_4 - 2d_2)t - 1) = 0.$$

Then, we project, over the free parameter space, the variety given by the hypotheses and the negation of the theses

```
> EliminationIdeal(<(d1)*(t2)-d2*(t1-1),(d2-t2),t3+f1, t4+f2,f1+i1
-2, f2+i2, k1+i1-2*t1, k2+i2-2*t2, ((k1+t3-2*d1)*z-1)*((k2+t4-2*
d2)*t-1)>, {t1,t2,t3,t4});
                              <t2>
```

The result means that if $t_2 \neq 0$, then the statement is true, i.e. there is no solution to the set of hypotheses and negation of theses

```
> 1 in <(d1)*(t2)-d2*(t1-1),(d2-t2),t3+f1, t4+f2,f1+i1-2, f2+i2,
k1+i1-2*t1, k2+i2-2*t2, ((k1+t3-2*d1)*z-1)*((k2+t4-2*d2)*t-1),
t2*s-1>;
                          true
```

In conclusion, with this formulation we can prove in a straightforward way that the statement is not generally false and that it is generally true. Moreover, we have also shown that the latter conclusion can also be obtained, in this particular case and in a non-automatic way, by analyzing the number of points in the variety described by (H, T) over each value of the free ones.

4.2.2 Converse Varignon: Option (b). Proving It is Not Generally False

Using midpoint concept (b), the conditions on the parallelogram $A(0, 0)$, $B(1, 0)$, $C(t_1, t_2)$, $D(d_1, d_2)$ such that $DC||AB$ and $AD||BC$ remain as above $d_1t_2 = d_2(t_1 - 1)$, $d_2 = t_2$. Next, see Fig. 4, we fix a free point $E(t_3, t_4)$, build the line EA, and, on this line, the point $F(f_1, f_2)$ such that A is the midpoint of segment EF. Here A is now the center of a circle passing through E and F, and F is in the line EA: $f_1^2 + f_2^2 = t_3^2 + t_4^2$, $t_3 f_2 = t_4 f_1$. Idem, point $I(i_1, i_2)$ such that B is the midpoint of segment FI: $(i_1 - 1)^2 + i_2^2 = (f_1 - 1)^2 + f_2^2$, $(f_1 - 1)i_2 = f_2(i_1 - 1)$. Idem, point $K(k_1, k_2)$ such that C is the midpoint of IK: $(k_1 - t_1)^2 + (k_2 - t_2)^2 = (i_1 - t1)^2 + i_2 - t_2)^2$, $(k_1 - t_1)(i_2 - t_2) = (k_2 - t_2)((i_1 - t_1)$.

The construction involves 12 variables $\{t_1, t_2, t_3, t_4, d_1, d_2, f_1, f_2, i_1, i_2, k_1, k_2\}$. It is of dimension 4

```
> HilbertDimension(<(d1)*(t2)-d2*(t1-1),d2-t2,((f1-0)^2+(f2-0)^2)-
((t3-0)^2+(t4-0)^2), (t3-0)*(f2-0)-(t4-0)*(f1-0),((i1-1)^2+(i2)
^2)-((f1-1)^2+(f2)^2), (f1-1)*(i2-0)-(f2-0)*(i1-1), ((k1-t1)^2+
(k2-t2)^2)-((i1-t1)^2+(i2-t2)^2), (k1-t1)*(i2-t2)-(k2-t2)*(i1-
t1)>,{t1,t2,t3,t4, d1,d2, f1,f2, i1,i2,k1,k2});
                          4
```

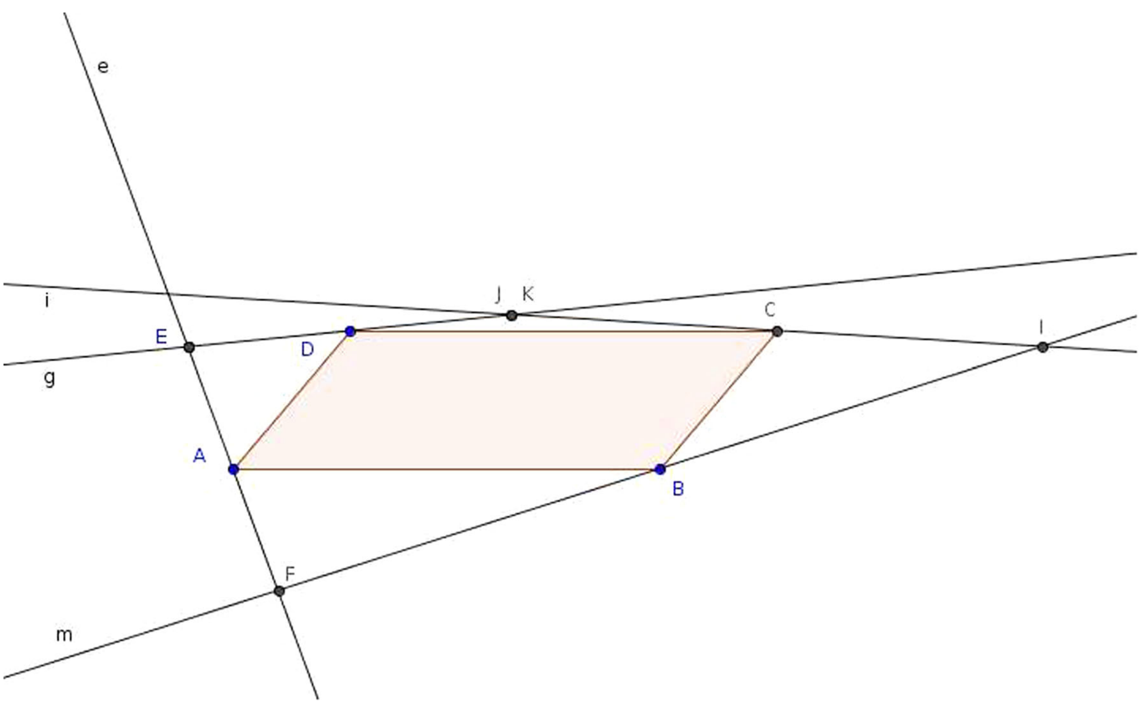

Fig. 4 Using option (b) with converse Varignon theorem

where the first four variables are free

```
> EliminationIdeal(<(d1)*(t2)-d2*(t1-1),d2-t2,((f1-0)^2+(f2-0)^2)-
((t3-0)^2+(t4-0)^2), (t3-0)*(f2-0)-(t4-0)*(f1-0),((i1-1)^2+(i2)
^2)-((f1-1)^2+(f2)^2), (f1-1)*(i2-0)-(f2-0)*(i1-1), ((k1-t1)^2+
(k2-t2)^2)-((i1-t1)^2+(i2-t2)^2), (k1-t1)*(i2-t2)-(k2-t2)*(i1-
t1)>,{t1,t2,t3,t4});
                              <0>
```

and the remaining variables are finitely determined. There are eight possible options for every $\{t_1, t_2, t_3, t_4\}$: two for point E; for each of them, two for point I, and, for every value of I, two for point K

```
> S:=solve({(d1)*(t2)-d2*(t1-1),d2-t2,((f1-0)^2+(f2-0)^2)-((t3-0)
^2+(t4-0)^2), (t3-0)*(f2-0)-(t4-0)*(f1-0),((i1-1)^2+(i2)^2)-((f1
-1)^2+(f2)^2), (f1-1)*(i2-0)-(f2-0)*(i1-1), ((k1-t1)^2+(k2-t2)
^2)-((i1-t1)^2+(i2-t2)^2), (k1-t1)*(i2-t2)-(k2-t2)*(i1-t1)},{d1,
d2, f1,f2, i1,i2,k1,k2});nops([S]);
{d1 = t1-1, d2 = t2, f1 = t3, f2 = t4, i1 = t3, i2 = t4, k1 = t3,
k2 = t4}, {d1 = t1-1, d2 = t2, f1 = t3, f2 = t4, i1 = t3, i2 = t4,
k1 = 2*t1-t3, k2 = 2*t2-t4}, {d1 = t1-1, d2 = t2, f1 = t3, f2
= t4,i1 = -t3+2, i2 = -t4, k1 = -t3+2, k2 = -t4}, {d1 = t1-1, d2
= t2, f1 = t3, f2 = t4, i1 = -t3+2, i2 = -t4, k1 = 2*t1+t3-2, k2 =
2*t2+t4}, {d1 = t1-1, d2 = t2, f1 = -t3, f2 = -t4, i1 = -t3,
i2 = -t4, k1 = -t3, k2 = -t4}, {d1 = t1-1, d2 = t2, f1 = -t3,
f2 = -t4, i1 = -t3, i2 = -t4, k1 = 2*t1+t3, k2 = 2*t2+t4},
{d1 = t1-1, d2 = t2, f1 = -t3, f2 = -t4, i1 = t3+2, i2 = t4,
k1 = t3+2, k2 = t4}, {d1 = t1-1, d2 = t2, f1 = -t3, f2 = -t4,
i1 = t3+2, i2 = t4, k1 = 2*t1-t3-2, k2 = 2*t2-t4}
                              8
```

Recall that the thesis is that D is the midpoint of segment KE, that is, $k_1 + t_3 = 2d_1$ and $k_2 + t_4 = 2d_2$. We will add these polynomials to the hypotheses and will see what are the consequence of this claim. Note that there is a hypothesis (on parallelism) stating that $d_2 = t_2$. Thus the ideal of hypotheses and theses is isomorphic to the ideal obtained replacing d_2 by t_2. In this way we define an equivalent, simpler, ideal with less variables and easier to handle:

```
> subs(d2=t2, {(d1)*(t2)-d2*(t1-1),d2-t2,((f1-0)^2+(f2-0)^2)-((t3
-0)^2+(t4-0)^2), (t3-0)*(f2-0)-(t4-0)*(f1-0),((i1-1)^2+(i2)^2)-(
(f1-1)^2+(f2)^2), (f1-1)*(i2-0)-(f2-0)*(i1-1), ((k1-t1)^2+(k2-
t2)^2)-((i1-t1)^2+(i2-t2)^2), (k1-t1)*(i2-t2)-(k2-t2)*(i1-t1),
k1+t3-2*d1,k2+t4-2*d2});
{0, d1t2-t2*(t1-1), (f1-1)i2-f2(i1-1), (k1-t1)(i2-t2)-(k2-t2)
(i1-t1), k1+t3-2d1, k2-2t2+t4, (i1-1)^2+i2^2-(f1-1)^2-f2^2,
(k1-t1)^2+(k2-t2)^2-(i1-t1)^2-(i2-t2)^2, -f1t4+f2t3, f1^2+
f2^2-t3^2-t4^2}

> IIdeal:=<(k1-t1)^2+(k2-t2)^2-(i1-t1)^2-(i2-t2)^2, t3*f2-t4*f1,
(k1-t1)*(i2-t2)-(k2-t2)*(i1-t1), (i1-1)^2+i2^2-(f1-1)^2-f2^2,
f1^2+f2^2-t3^2-t4^2, (f1-1)*i2-f2*(i1-1), -2*t2+k2+t4, d1*t2-t2*
(t1-1), k1+t3-2*d1>:
```

In order to simplify the costly elimination process, we realize that `IIdeal` contains a polynomial $d_1t_2-t_2(t_1-1)$ that factors as t_2 times $d_1 - t_1 + 1$. So, we replace in `IIdeal` the polynomial $d_1t_2 - t_2(t_1 - 1)$ by $d_1 - t_1 + 1$, defining IdealA; likewise, replacing in IIdeal $d_1t_2 - t_2(t_1 - 1)$ by t_2 we define IdealB. It can be checked that the intersection of IdealA and IdealB is equal to IIdeal

```
> IdealC:=Intersect(IdealA,IdealB): IdealContainment(IdealC,
IIdeal, IdealC);
                        true
```

So, the elimination of IIdeal over the variables $\{t_1, t_2, t_3, t_4\}$ is $\langle 0\rangle$ because the elimination of IdealA is $\langle 0\rangle$ and the elimination of IdealB is $\langle t_2\rangle$.

```
> EliminationIdeal(IdealA,{t1,t2,t3,t4});EliminationIdeal(IdealB,
{t1,t2,t3,t4});
                        <0>
                        <t2>
> EliminationIdeal(IIdeal, {t1,t2,t3,t4});
                        <0>
```

This means

- the statement is not generally false, because for almost each value of the free parameters there is a value of the depending variables $d_1, f_1, f_2, i_1, i_2, k_1, k_2$ where the theses and the hypotheses hold, so that the statement is true; but it could happen that there is also a different value of $d_1, f_1, f_2, i_1, i_2, k_1, k_2$ where the statement is false. We know in this particular case, by solving the system given by IIdeal that there is only one value of the depending variables verifying the hypotheses and theses for each value of the free ones.

```
> solve({(k1-t1)^2+(k2-t2)^2-(i1-t1)^2-(i2-t2)^2, t3*f2-t4*f1,
(k1-t1)*(i2-t2)-(k2-t2)*(i1-t1), (i1-1)^2+i2^2-(f1-1)^2-f2^2,
f1^2+f2^2-t3^2-t4^2, (f1-1)*i2-f2*(i1-1), -2*t2+k2+t4, d1*t2-
t2*(t1-1), k1+t3-2*d1},{d1,f1,f2,i1,i2,k1,k2});
{f1 = -t3, i1 = 2 + t3, k1 = -t3 - 2 + 2t1, d1 = t1 - 1, k2 =
-t4  2t2, f2 = -t4 , i2 = t4}
```

- this one value is, precisely, the one that corresponds to the intuitive idea about $F, I, K \ldots$
- remark that the elimination of IIdeal coincides with that of IdealA; again, the above Solve output for IIdeal coincides with that for IdealA, but there is not general solution for IdealB (a degenerate case, with $t_2 = 0$, the parallelogram degenerates to a line), getting an empty output to the following command:

```
> solve({(k1-t1)^2+(k2-t2)^2-(i1-t1)^2-(i2-t2)^2, t3*f2-t4*f1,
(k1-t1)*(i2-t2)-(k2-t2)*(i1-t1), (i1-1)^2+i2^2-(f1-1)^2-f2^2,
f1^2+f2^2-t3^2-t4^2, (f1-1)*i2-f2*(i1-1), -2*t2+k2+t4, t2,
k1+t3-2*d1},{d1,f1,f2,i1,i2,k1,k2});
```

4.2.3 Converse Varignon: Option (b). Proving It is Not Generally True

Unfortunately, we can not prove that this formulation of the Converse Varignon statement is generally true. In fact, let us try to see that it is not generally true, by considering the collection of hypotheses and the negation of theses $(((k_1+t_3-2d_1)t-1)((k_2+t_4-2d_2)s-1)=0)$. Then, we project over the free parameter space the variety given by the hypotheses and the negation of the theses, by eliminating all variables except t_1, t_2, t_3, t_4. If this elimination is 0, then it is not generally true, because it means that for almost all values of the free points C, E, there are values of the remaining variables so that the thesis is not true. Taking in consideration that we have already proved the statement is not generally false, proving it is not generally true, it would mean that for almost all positions of the free points C, E, there are values of the remaining points such that the thesis holds (because the statement is not generally false), but there are also values of such points where the thesis does not hold (because it is not generally true). Let us see if we can achieve proving it is not generally true, by

Elimination (Ideal of Hypotheses + Negation of Theses, $\{t_1, t_2, t_3, t_4\}$).

But this elimination, directly, is too involved concerning time and memory. Let us attempt to simplify it, as above, by, first, substituting $d_2 = t_2$ in all polynomials of the ideal of hypotheses and negation of theses, yielding `IdealN`. Then we split in two factors the generator $d_1 t_2 - t_2(t_1 - 1)$ obtaining factor $d_1 - (t_1 - 1)$ and factor t_2. Likewise, we split in two factors the generator $((k_1 + t_3 - 2d_1)t - 1)((k_2 + t_4 - 2d_2)s - 1)$. Thus, combining the four resulting factors, we build four ideals. `IdealP1` and `IdealP2`, both with factor $d_1 - (t_1 - 1)$ and with factors $((k_1 + t_3 - 2d_1)t - 1)$ or $(k_2 + t_4 - 2d_2)s - 1)$, respectively. Same, `IdealQ1` and `IdealQ2`, both with factor t_2 and with factors $((k_1 + t_3 - 2d_1)t - 1)$ or $(k_2 + t_4 - 2d_2)s - 1)$. `IdealR` is the intersection of the four ideals. Obviously, `IdealN` is contained in each of the four ideals, and, thus, `IdealN` is contained in `IdealR`, and the converse is also true (by brute force computation with Maple, via IdealContainment command). Thus, `IdealR` is equal to `IdealN`, and the elimination in `IdealN` can be carried out by eliminating in `IdealP1`, `IdealP2`, `IdealQ1`, `IdealQ2`, and then finding the intersection of elimination.

Note that the elimination in `IdealQ1` and `IdealQ2`

```
> EliminationIdeal(IdealQ1,{t1,t2,t3,t4});
                          <t2>
> EliminationIdeal(IdealQ2,{t1,t2,t3,t4});
                          <t2>
```

shows that the elimination applied to `IdealN` is also contained in $\langle t_2 \rangle$. We are not able to eliminate `IdealP1`, `IdealP2` with Maple, but we do achieve it with Sage:

```
sage: K=QQ['t1, t2, t3, t4, d1, f1, f2, k1, k2, i1, i2, t,z'] sage:
K.inject_variables() Defining t1, t2, t3, t4, d1, f1, f2, k1, k2,
i1, i2, t, z sage: P1=Ideal((k1-t1)^2+(k2-t2)^2-(i1-t1)^2-(i2-t2)^2,
t3*f2-t4*f1, (k1-t1)*(i2-t2)-(k2-t2)*(i1-t1), (i1-1)^2+i2^2
-(f1-1)^2-f2^2, f1^2+f2^2-t3^2-t4^2, (f1-1)*i2-f2*(i1-1), d1-(t1-1),
((k1+t3-2*d1)*z-1)) sage:
P2=Ideal((k1-t1)^2+(k2-t2)^2-(i1-t1)^2-(i2-t2)^2, t3*f2-t4*f1,
(k1-t1)*(i2-t2)-(k2-t2)*(i1-t1), (i1-1)^2+i2^2 -(f1-1)^2-f2^2,
f1^2+f2^2-t3^2-t4^2, (f1-1)*i2-f2*(i1-1), d1-(t1-1),
((-2*t2+k2+t4)*t-1)) sage: time
P1.elimination_ideal([k1,k2,f1,f2,i1,i2,d1,t]) CPU times: user 46
min 11 s, sys: 752 ms, total: 46 min 12 s Wall time: 46 min 10 s
Ideal (0) of Multivariate Polynomial Ring in t1, t2, t3, t4, d1, f1,
f2, k1, k2, i1, i2, t, z over Rational Field
```

Thus, the statement is not generally true.

4.2.4 Converse Varignon: Option (b). Learning from Failure

Yet, there is an indirect way of proving—within Maple—that the statement is not generally true. Namely, we ask Maple to solve the system of equations given by the hypotheses and the negation of theses, so that the constrained variables $d_1, f_1, f_2, i_1, i_2, k_1, k_2, t, z$ (recall we have applied the equality $d_2 = t_2$ and, thus, there are no terms in the variable d_2 in the system of equations) are solved in terms of the free variables t_1, t_2, t_3, t_4. If for almost all (i.e. all except for a closed set in the space of t_1, t_2, t_3, t_4) values of the free variables there is a solution to the system of equations given by the hypotheses and the negation of theses, it is clear that the projection of its solution set over the t_1, t_2, t_3, t_4-space will be almost all that space (i.e. the elimination ideal will be $\langle 0 \rangle$, since 0 is the only equation verified by the closure of the projection, that is, by the whole t_1, t_2, t_3, t_4-space).

We can see that, in fact, for every value of t_1, t_2, t_3, t_4 (except for some values that would vanish some denominators), there are 13 different values of $d_1, d_2, i_1, i_2, k_1, k_2, t, z$ in the system of equations given by the hypotheses and the negation of theses:

```
> SS:=solve({d1*t2-t1*t2+t2, (k1-t1)^2+(k2-t2)^2-(i1-t1)^2-(i2-t2)
^2, t3*f2-t4*f1, (k1-t1)*(i2-t2)-(k2-t2)*(i1-t1), (i1-1)^2+i2^2-
(f1-1)^2-f2^2, f1^2+f2^2-t3^2-t4^2, (f1-1)*i2-f2*(i1-1), ((k1+t3
-2*d1)*z-1)*((-2*t2+k2+t4)*t-1)}, {d1,f1,f2,i1,i2,k1,k2,t,z});

SS :=
 1) {z=-1/(2(-t3+t1-1)), d1=t1-1, i2=t4, f1=t3, k1=t3, k2=t4,
i1=t3, f2=t4, t=t},
 2) {d1=t1-1, i2=t4, f1=t3, i1=t3, f2=t4, t=t, k2=-t4+2t2,
k1=-t3+2*t1, z=1/2},
 3) {d1=t1-1, f1=t3, f2=t4, t=t, k1=-t3+2, z=-1/(2(-2+t1)),
i=-t4, k=-t4, i=-t3+2},
 4) {d1=t1-1, f1=t3, f2=t4, t=t, i2=-t4, i1=-t3+2, k2=2t2+t4,
z=1/(2t3), k1=-2+t3+2t1},
 5) {d1=t1-1, t=t, i2=-t4, k2=-t4, f1=-t3, f2=-t4, k1=-t3,
z=-1/(2(t1-1)), i1=-t3},
 6) {d1=t1-1, t=t, i2=-t4, k2=2t2+t4, f1=-t3, f2=-t4, i1=-t3,
k1=t3+2t1, z=1/(2(t3+1))},
 7) {d1=t1-1, i2=t4, k2=t4, t=t, f1=-t3, f2=-t4, i1=2+t3,
z=-1/(2(-2-t3+t1)), k1=2+t3},
 8) {d1=t1-1, i2=t4, f1=t3, k1=t3, k2=t4, i1=t3, f2=t4,
t=-1/(2*(-t4+t2)), z=z},
 9) {d1=t1-1, f1=t3, f2=t4, k1=-t3+2, i2=-t4, k2=-t4, i1=-t3+2,
z=z, t=-1/(2*t2)},
 10) {d1=t1-1, f1=t3, f2=t4, i2=-t4, i1=-t3+2, k2=2t2+t4,
k1=-2+t3+2t1, z=z, t= 1/(2t4)},
 11) {d1=t1-1, i2=-t4, k2=-t4, f1=-t3, f2=-t4, k1=-t3, i1=-t3,
z=z, t=-1/(2t2)},
 12) {d1= 1-1, i2=-t4, k2=2t2+t4, f1=-t3, f2=-t4, i1=-t3,
k1=t3+2t1, z=z, t=1/(2t4)},
 13) {d1=t1-1, i2=t4, k2=t4, f1=-t3, f2=-t4, i1=2+t3, k1=2+t3,
t=-1/(2(-t4+t2)), z=z}
```

Why 13 solutions? We have seen that for every t_1, t_2, t_3, t_4 there are eight different values of the remaining coordinates (i.e., of the cartesian product of the coordinates of the points $F \times K \times I$). Let us repeat that output S here below, identifying each of the blocks with a number from 1 to 8:

```
 S :=
1) {d2=t2, k1=t3, i1=t3, f1=t3, d1=t1-1, f2=t4, k2=t4, i2=t4}
2) {d2=t2, i1=t3, f1=t3, d1=t1-1, k1= 2t1-t3, f2=t4 , k2=-t4+
2t2, i2=t4}
3) {d2=t2, k2=-t4, f1=t3, d1=t1-1, k1=-t3+2, i1=-t3+2, f2=t4,
i2 =t4}
4) {d2=t2, k1=-2+t3+2t1, k2=t4+2t2, f1=t3, d1=t1-1, i1=-t3+2,
f2=t4,i2=-t4}
5) {k1=-t3, d2=t2, i1=-t3, f1=-t3, k2=-t4, d1=t1-1, i2=-t4,
```

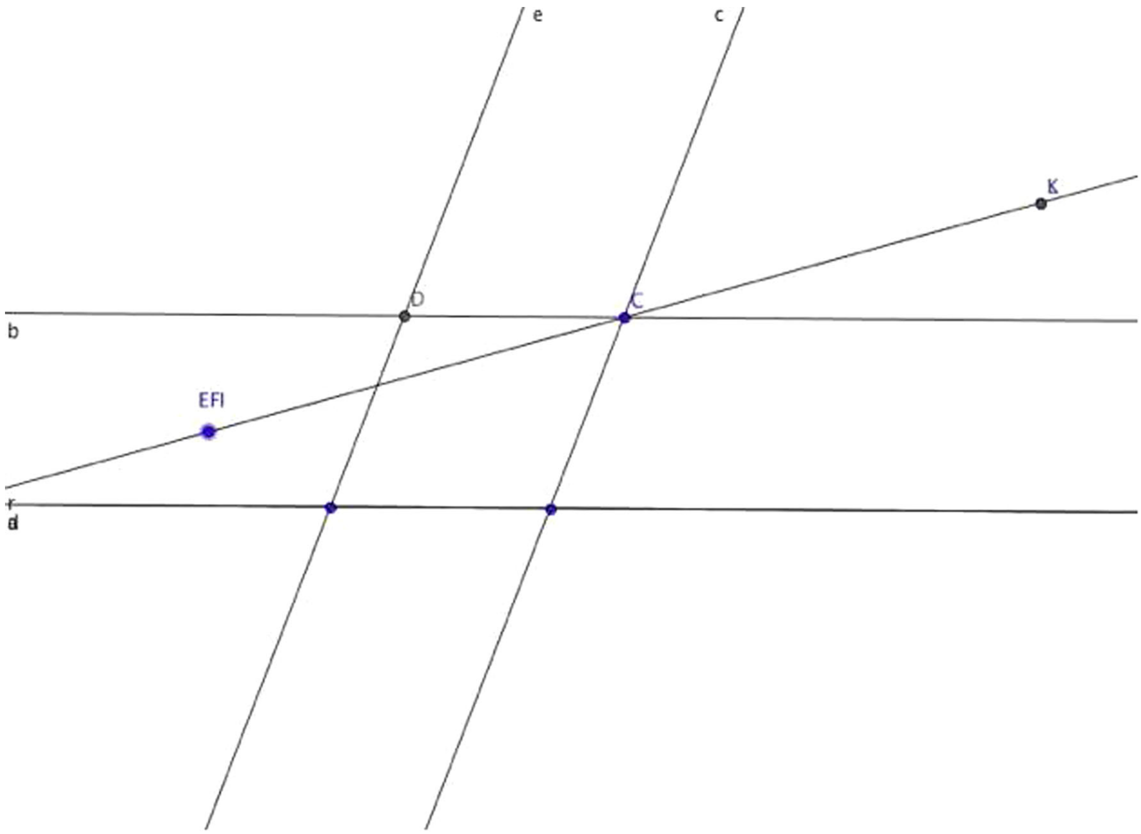

Fig. 5 Learning from failure

```
f2=-t4}
6) {d2=t2, i1=-t3, f1=-t3, k1=2t1+t3, k2=t4+2t2, d1=t1-1, i2=
-t4, f2=-t4}
7) {d2=t2, f1=-t3, i1=2+t3, k1=2+t3, d1=t1-1, k2=t4, f2=-t4, i2
=t4}
8) {d2=t2, f1=-t3, i1=2+t3, k1=-t3-2+2t1, d1=t1-1, k2=-t4+2t2,
f2=-t4, i2=t4}
```

Recall that number (8) corresponds to the values of F, K, I verifying the theses. So, the values of F, K, I in SS must come from one of the other blocks of S, say, (1)–(7).

In fact we can easily verify that the 13 blocks of SS can be described as follows:

- twelve blocks of SS correspond to the blocks (1), (3), (4), (5), (6), (7) of S regarding the values of $d_1, f_1, f_2, i_1, i_2, k_1, k_2$. In all cases $d_2 = t_2$ and $d_1 = t_1 - 1$. Given these values of $d_1, f_1, f_2, i_1, i_2, k_1, k_2$, then the value of z is, automatically, $z = 1/((k_1 + t_3 - 2d1)$—because the denominator is not identically zero—and t can take any value; this description includes the first six blocks of SS above. Then, there are other six blocks for the same values of $d_1, f_1, f_2, i_1, i_2, k_1, k_2$, where z takes any value and t is automatically the value of $t = 1/(k_2 + t_4 - 2d_2)$.
- a 13th block where $d_1, f_1, f_2, i_1, i_2, k_1, k_2$ correspond to block number (2) of S, but where $-2t_2 + k_2 + t_4$ is then identically zero and $k_1 + t_3 - 2d_1$ is equal to $2t_1 - 2d_1$, so equal to 2 ($z = 1/2$), because $d_1 = t_1 - 1$;—so one the two theses is true—but where the other thesis does not hold, so t is any value.

A remarkable asymmetry can be noted here in the fact that there is only one instance of the construction verifying precisely one of the thesis (the one about the y coordinate) and violating the other (and there is not one instance verifying the thesis about the x coordinate and violating the other).

In fact, the meaning of the block of values for the depending variables

```
k2 = -t4+2*t2, i2 = t4, f1 = t3, k1 = -t3+2*t1, i1 = t3, f2 = t4
```

is, essentially that $E = F = I$, so that K is actually the symmetrical of E respect to C (see Fig. 5), so that t_1 is $(k_1+t_3)/2$ and t_2 is $(k_2+t_4)/2$. So, now the theses are that D is the midpoint of K and E, so that $d_2 = (k_2+t_4)/2$, and $d_1 = (k_1+t_3)/2$. Obviously, the first equality, bearing in mind that $d_2 = t_2$ in all the hypotheses, it holds because is part of the block description; the second equality does not, because, bearing in mind the block hypotheses, $d_1 = (k_1+t_3)/2$ is equivalent to $d_1 = t_1, \ldots$, and this is not true, since $d_1 = t_1 - 1$ in the construction, if t_2 is not zero. In other words, the thesis here is that $C = D$, and it is true that the y coordinate of C is equal to the y coordinate of D, but not the x coordinate.

In conclusion, the statement is not generally false, because for one interpretation of the construction both theses hold, but it is not generally true because for seven other interpretations of the construction none of the theses hold (in six cases times two, i.e., 12) or just one thesis does not hold (in the only one remaining case). This could be considered as the discovery of a (subtle) new geometric fact …

5 Conclusion

A detailed study of the theorem of Varignon has been performed. In a graphic environment, as the one provided by GeoGebra, the Varignon parallelogram requires for its specification a thorough understanding of the midpoint definition inside the system. We show that depending on the used midpoint definition, the statement can be declared generally true if the native GeoGebra midpoint command is used. Nevertheless, if the midpoint is given and the user must construct one of the endpoints (as in the converse Varignon statement), the computation can become very involved (for lack of a standard GeoGebra protocol) and the Varignon conclusion is generally true for the direct case, while it is neither generally true nor generally false for the converse case. *Guessing* complementary conditions to avoid falling in this confusing circumstance (without performing a primary decomposition, something not realistic in terms of required computing time and memory) is, sometimes intuitive, sometimes very complicated, yielding, if achieved, to the discovery of new geometric facts (such as the ones expressed by the different blocks in the above discussion). We think that these reflections should be carefully considered when designing user interfaces for massive use of theorem proving features in popular dynamic geometry programs, such as GeoGebra.

Acknowledgments Authors partially supported by the Spanish “Ministerio de Economía y Competitividad” and by the European Regional Development Fund (ERDF), under the Project MTM2014-54141-P.

References

1. Botana, F., Kovács, Z.: A Singular web service for geometric computations. Ann. Math. Artif. Intell. **74**, 359–370 (2015)
2. Botana, F., Hohenwarter, M., Janičić, J., Kovács, Z., Petrović, I., Recio, T., Weitzhofer, S.: Automated theorem proving in GeoGebra: current achievements. J. Autom. Reason. **55**, 39–59 (2015)
3. Chou, S.C.: Mechanical Geometry Theorem Proving. Reidel, Dordrecht (1988)
4. Dalzotto, G., Recio, T.: On protocols for the automated discovery of theorems in elementary geometry. J. Autom. Reason. **43**, 203–236 (2009)
5. Recio, T., Vélez, M.P.: Automatic discovery of theorems in elementary geometry. J. Autom. Reason. **23**, 63–82 (1999)
6. Recio, T., Sterk, H., Vélez, M.P.: Project: automatic geometry theorem proving. In: Cohen, A., Cuipers, H., Sterk, H. Some Tapas of Computer Algebra. Algorithms and Computations in Mathematics, vol. 4, pp. 276–296. Springer, Heidelberg (1998)