DEFINITION OF AFFINE GEOMETRY BY A GROUP OF TRANSFORMATIONS

Joe Lipman

(received July 10, 1961)

Introduction. In his Geometric Algebra (New York, 1957) E. Artin poses the problem of co-ordinatizing an affine plane in the following terms.

How little do we have to assume, from a geometric point of view, about an affine plane, in order to be able to describe its points by pairs of elements of a field, and its lines by linear equations?

This suggests a more general question.

(Q) What axioms for affine geometry will give a structure that is equivalent to that of a vector space (of arbitrary dimension) in such a way that the lines of the vector space (i. e. the one-dimensional linear varieties) correspond to the lines of the geometry?

Artin gives five axioms concerning points, lines, and a certain group of transformations called dilatations; i. e. (intuitively) the group consisting of magnifications from points and of translations (magnifications from infinity). These axioms enable him to introduce coordinates. What we are going to do here is to concentrate on the transformations alone. We start out with a set and a group of transformations on the set. Straight lines and parallelism are defined in terms of the action of the group on the set. The affine structure is then determined by two geometric axioms, (resembling Artin's axioms 4b and 3), which are really assumptions about the transitivity properties of the group.

It turns out that these assumptions provide an answer to (Q).

Canad. Math. Bull. vol. 4, no. 3, September 1961

https://doi.org/10.4153/CMB-1961-030-1 Published online by Cambridge University Press

Although the basic approach and final results here are different, the development is taken from Artin; proofs of several of the theorems are included only so that the exposition may be self-contained.

The end of a proof is indicated, where necessary, by ///.

NOTE: It is well known that given any group of transformations, transitive on a set, one can identify the points of the set with the left cosets of a certain subgroup. [cf. N. Bourbaki: Structures Algébriques, p. 106] Therefore, one could dispense with the set, and, taking the point of view of F. Bachmann [Aufbau der Geometrie aus dem Spiegelungsbegriff, Springer, 1959], interpret the present paper as a development within group theory of affine geometry over a skew field.

1. Terminology.

DEFINITION 1. A geometry is a set G together with a group of transformations of G into itself.

For convenience, we call the transformations dilatations.

A dilatation has an inverse, and hence is 1-1 onto.

DEFINITION 2. $A, B \equiv C, D$ (for points A, B, C, D of G) means that there is a dilatation g such that $gA = C$, $gB = D$.

$\equiv$ is obviously an equivalence relation.

DEFINITION 3. We write "[ABC]" and say "A, B, C are collinear" if $A = B$ or $A = C$, or if there is a dilatation which leaves A fixed and sends B into C (i.e. $A, B \equiv A, C$).

THEOREM 1. "[ACB]" is implied by any one of:

a) $B = C$

b) [ABC]

c) [ACE], [AEB], $A \neq E$.

Proof. The dilatations form a group.

https://doi.org/10.4153/CMB-1961-030-1 Published online by Cambridge University Press

DEFINITION 4. If $A \neq B$, then $A + B$ denotes the set consisting of A and of all images of B under dilatations leaving A fixed. Thus

$$A + B = \{C \mid [ABC]\}.$$

Any set $A + B$ is called a line.

Note that this definition is not symmetric in A and B; and that $A + A$ is left undefined.

COROLLARY. A dilatation g takes the line $A + B$ onto the line $gA + gB$.

Proof. Clearly $[ABC]$ iff $[gA, gB, gC]$.

2. Straight lines.

THEOREM 2. (i) The following statements are equivalent:

a) $A + B = B + A$ (where A, B are any two distinct points).

b) The truth or falsity of "$[ABC]$" does not depend on the order of the arguments (where A, B, C are any three points).

c) If $C \neq D$ and $C, D \in A + B$, then $A + B \subset C + D$.

d) If $A \neq B$, there is one and only one line containing A, B, namely $A + B$.

(ii) Statements a), b), c), d) are all implied by:

e) The group of dilatations is doubly transitive on lines; that is to say: if A, B, C, D all belong to a line, and $A \neq B$, $C \neq D$, then $A, B \equiv C, D$.

Proofs. a) $\rightarrow$ b): a) can be restated as "if $[ABC]$ then $[BAC]$". In conjunction with theorem 1(b) this clearly gives statement b).

b) $\rightarrow$ c): Let $E \in A + B$ and assume $E \neq A$ (if $E = A$, then $E \neq B$ and the proof is similar).

In view of statement b), theorem 1(c) may be applied three times as follows:

$$[ABC], [ABE] \rightarrow [ACE]$$
$$[ABD], [ABE] \rightarrow [ADE]$$
$$[ACE], [ADE] \rightarrow [CDE]$$

i.e. $E \in C + D$.

https://doi.org/10.4153/CMB-1961-030-1 Published online by Cambridge University Press

c) $\to$ d): c) says that if $A, B \in C + D$ then $C + D \subset A + B$. But then $C, D \in A + B$, so that $A + B \subset C + D$.

d) $\to$ a): Obvious.

e) $\to$ a): Let $C \in A + B$. If $C = B$, then [BAC]. Otherwise, by e) $B, A \equiv B, C$, i.e. [BAC]. In either case, $C \in B + A$. Statement a) now follows. ///

The geometry is provided with a structure by the adoption of statement e) as our first assumption. Statement e) will be referred to henceforth as AXIOM 1.

a), b), c), d) can, then, be considered to be established theorems. Also, we deduce easily

COROLLARY A1. For any A, B, there exists a dilatation interchanging A and B.

A propos, we prove

THEOREM 3. If some dilatation other than the identity has a fixed point, then a), b), c), d), and axiom 1 are all equivalent.

Proof. It is enough to show that d) implies axiom 1.

Since some dilatation other than the identity has a fixed point, some line contains at least three points. It follows that <u>every line contains at least three points</u>.

Otherwise, there is a line $A + B$ containing precisely two points, viz. A, B. This means that every dilatation leaving A fixed also leaves B fixed. If some line through A contains three distinct points, A, C, D, then, by d), [ACD]; and by the preceding remark, this implies that [BCD]. Thus, by d), both A and B belong to $C + D$ and so again by d), $C, D \in A + B$ which is absurd. In other words, every line through A contains exactly two points. But any point E in G lies on a line through A, viz. $E + A$; and so, repeating the argument, we conclude that any line through any point contains exactly two points, which is a contradiction.

https://doi.org/10.4153/CMB-1961-030-1 Published online by Cambridge University Press

Axiom 1 is now deduced as follows:

If $A \neq D$ then $hg(A) = C$, $hg(B) = D$ where g is such that $gA = A$, $gB = D$; and h is such that $hD = D$, $hA = C$.

If $B \neq C$ a similar procedure applies.

Finally, if $A = D$, $B = C$, choose $E \neq A$ or B, $E \in A + B$. Then, by the preceding two paragraphs, $A, B \equiv C, E \equiv C, D$. ///

Note that the assumption in theorem 3 is necessary, as can be seen by considering a group of dilatations consisting of the identity alone.

We now return to the main development.

3. Parallel lines.

DEFINITION 5. Two lines m_1, m_2 are parallel (written $m_1 \| m_2$) if there exists a dilatation mapping m_1 onto m_2.

COROLLARY P1. Parallelism is an equivalence relation.

COROLLARY P2. A dilatation maps any line onto a parallel line.

Proof. Refer to the corollary to definition 4.

COROLLARY P3. For any three points A, C, D, with $A \neq C$, $D \neq C$, there is a point B such that $C + D \| B + A$, $C + A \| B + D$.

Proof. Let h interchange A and D (corollary A1) and let $B = hC$. The required properties of B follow from corollary P2. ///

From axiom 1, we get:

COROLLARY P4. If $m_1 \| m_2$ and A_i, $B_i \in m_i$ $(i = 1, 2)$

https://doi.org/10.4153/CMB-1961-030-1 Published online by Cambridge University Press

then $A_1, B_1 \equiv A_2, B_2$.

With theorem 2(d), this implies:

COROLLARY P5. Two parallel lines are either coincident or disjoint.

Proof. If $A + B \parallel B + C$, then $A, B \equiv B, C$ i.e. $[ABC]$.

THEOREM 4. For any point A and any line $C + D$ there is a unique line which contains A and is parallel to $C + D$.

Proof. (Existence). If $A = C$ this is trivial. Otherwise use corollary P3.

(Uniqueness). Corollaries P1, P5.

4. Traces.

DEFINITION 6. A g-trace of A is any line through A and gA.

COROLLARY 4.1. If m is a g-trace of A, then $g(m) = m$.

Proof. $g(m)$ and m are parallel (corollary P2) and both contain gA.

COROLLARY 4.2. Two distinct g-traces which intersect do so in a fixed point of g.

Proof. Corollary 4.1, theorem 2(d).

COROLLARY 4.3. If B is not a fixed point of g, then there is one and only one g-trace containing B; viz. $B + gB$.

Proof. Corollary 4.2. ///

If it happens that any three points are collinear, then any line $C + D$ contains all points (because $[CDP]$ for all P) which means that $C + D$ is the only line. There does not seem to be much point in studying incidence if every point is incident

https://doi.org/10.4153/CMB-1961-030-1 Published online by Cambridge University Press

with every line. So we restrict our attention to those geometries satisfying

AXIOM 2. There exist three non-collinear points.

NOTE. In view of axiom 1, an equivalent statement is: The dilatation group is not doubly transitive on G.

For any two distinct points C, D there will now be a point A such that A, C, D are not collinear. Otherwise, as in the preceding remark, C + D is the only line, and axiom 2 is violated.

THEOREM 5. A dilatation g with two distinct fixed points is the identity.

Proof. Let C and D be fixed points ($C \neq D$) and choose $A \notin C + D$. Then C + A, D + A are distinct g-traces, and so their intersection A is a fixed point (corollary 4.2). Since C and A are fixed points, the same argument shows that any $B \in C + D$ $(B \neq C)$ is a fixed point. Thus, all points are fixed points of g, i.e. g is the identity.

COROLLARY 5.1. A dilatation is uniquely determined by the images of two distinct points C, D.

Proof. If $gC = hC$, $gD = hD$, then C and D are fixed points of $g^{-1}h$, so that $g^{-1}h$ is the identity, i.e. $g = h$.

THEOREM 6. If g is different from the identity, and has a fixed point A, then the g-traces are precisely the lines through A.

Proof. Any line through A is a g-trace. Conversely, if $B \neq A$, then B is not a fixed point; so that $B + A = B + gB$ (corollary 4.3) i.e. every g-trace passes through A.

THEOREM 7. If g is not the identity, then any two non-parallel g-traces m_1 and m_2 have a point in common.

[This point is a fixed point of g (corollary 4.2)]

https://doi.org/10.4153/CMB-1961-030-1 Published online by Cambridge University Press

Proof. Let $A \in m_1$, $C \in m_2$ where neither A nor C is a fixed point. Then $m_1 = A + gA$, $m_2 = C + gC$ (corollary 4.3). We may assume $A \neq C$.

Find B such that $C + gC \parallel B + A$, $C + A \parallel B + gC$ (corollary P3).

Since $C + A \parallel gC + gA$, $B \in gC + gA$. Moreover $B \neq gA$, since $C + gC \nparallel gA + A$. Hence there is an h such that

$$h(gA) = gA \qquad hB = gC.$$

Then $hA \in A + gA$ and

$$C + gC \parallel A + B \parallel hA + hB \parallel hA + gC$$

so that $hA \in C + gC$. Thus $A + gA$ meets $C + gC$ in hA.

COROLLARY 7.1 . For any dilatation g, there are three mutually exclusive possibilities:

1) g is the identity.

2) All g-traces are parallel. There is no fixed point.

3) g is not the identity, and some two g-traces are not parallel. Then there is a fixed point and all traces pass through it.

5. Translations.

DEFINITION 7. In the cases 1) and 2) of corollary 7.1 g is called a translation. Thus a non-identity translation is characterised either by the absence of a fixed point or by the fact that some two of (and hence any two of) its traces are parallel. The equivalence class of parallels determined by the traces is referred to as the direction of g. The identity will be said to have every direction. Two translations having a common direction will be said to be parallel. Thus the identity is parallel to every translation.

THEOREM 8. Given any two points C, D there is a unique translation taking C to D.

Proof. If $C = D$ we use the identity. Otherwise choose

$A \notin C + D$ and find B as in corollary P3. By corollary P4, there is a dilatation h with $hA = B$, $hC = D$. Since h has two distinct parallel traces, (viz. $A + B$, $C + D$) h is a translation.

Unicity is proved in

THEOREM 9. A translation t is uniquely determined by the image of one point A.

Proof. If $A = tA$, t must be the identity. Otherwise the image of any $B \notin A + tA$ is determined as the intersection of the parallel to $A + tA$ through B and the parallel to $A + B$ through tA. Then the image of any $C \in A + tA$ is determined by B and tB.

COROLLARY. The translations form a subgroup of the group of dilatations.

Proof. Clearly if t_1 is a translation, then so is t_1^{-1}. Moreover if $t_1^{-1} t_2 A = A$, then $t_2 A = t_1 A$ whence $t_2 = t_1$ and $t_1^{-1} t_2$ is the identity.

THEOREM 10. The translations having a given direction form a normal subgroup of the group of dilatations.

Proof. a) Let A be any point, and let t_1 and t_2 have a certain direction. Obviously t_2^{-1} has the same direction.

Now, either one of the translations t_1, t_2, $t_2^{-1} t_1$ is the identity, or

$$A + t_2^{-1} A \parallel A + t_1 A \parallel t_2^{-1} A + t_2^{-1} t_1 A \qquad \text{(corollary P2)}$$

so that $[A, t_2^{-1} A, t_2^{-1} t_1 A]$; i.e. the $t_2^{-1} t_1$-trace of A is the same as the t_2^{-1}-trace. q.e.d.

b) From $g^{-1}A + tg^{-1}A \parallel g^{-1}B + tg^{-1}B$ it follows that

$$A + gtg^{-1}A \parallel B + gtg^{-1}B \qquad \text{(corollaries P2, P1).}$$

Thus gtg^{-1} is a translation, and since

$$g^{-1}A + tg^{-1}A \parallel A + gtg^{-1}A \qquad \text{(corollary P2),}$$

gtg^{-1} is parallel to t.

THEOREM 11. The group of translations is Abelian.

Proof. Let t_1, t_2 have different directions. By theorem 10, $t_1(t_2t_1^{-1}t_2^{-1}) = (t_1t_2t_1^{-1})t_2^{-1}$ is parallel to both t_1 and t_2 so that $t_1t_2t_1^{-1}t_2^{-1}$ is the identity i.e. $t_1t_2 = t_2t_1$.

If t_1 and t_2 are parallel, and t_3 is not parallel to t_2, then t_2t_3 is not parallel to t_2 (otherwise t_2 and $t_2^{-1}(t_2t_3)$ are parallel). By the first paragraph

$$(t_1t_2)t_3 = t_1(t_2t_3) = (t_2t_3)t_1 = (t_3t_2)t_1 = t_3(t_2t_1) = (t_2t_1)t_3$$

whence $t_1t_2 = t_2t_1$.

6. <u>The vector space.</u> Since the group of translations is commutative, it may be considered as a unitary left module over its ring of endomorphisms.

(In other words, if α, β are endomorphisms, we write the group of translations additively and define $\alpha + \beta$, $\alpha\beta$ by

$$(\alpha + \beta)t = \alpha t + \beta t \qquad (\alpha\beta)t = \alpha(\beta t)$$

respectively. One sees quickly that $\alpha + \beta$, $\alpha\beta$ are endomorphisms; that the endomorphisms form an additive Abelian group, (with $-\alpha$ defined by $(-\alpha)t = -\alpha t$.); that the distributive laws $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$, $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$ hold; and that $1.\beta = \beta.1 = \beta$, where 1 is the identity mapping of the translations onto themselves; in short, the endomorphisms form a ring with a unit. Finally, by definition,

$$\alpha(t_1 + t_2) = \alpha t_1 + \alpha t_2 \qquad 1(t) = t$$

so that all the conditions for a left unitary module are fulfilled.)

We will be particularly interested in those endomorphisms which are direction preserving, i.e. those which take any translation into a parallel translation. Using theorem 10, one verifies easily that these endomorphisms form a subring F of the ring of endomorphisms. The object of the next few theorems is to establish the fact that F is a skew field. The translations then become a left vector space over the skew field, and, as will be seen, this structure can be transferred over to the geometry.

THEOREM 12. Let $\alpha \in F$ and let t_1 be a translation different from the identity. If αt_1 is the identity, then $\alpha = 0$ (the 0-endomorphism takes all translations into the identity).

Proof. Let t_2 have a different direction than that of t_1. Then t_2 and $t_2 t_1^{-1}$ have different directions [otherwise, by theorem 10, t_2 and t_1 are parallel, since $t_1 = (t_2 t_1^{-1})^{-1} t_2$].

But $t_2 t_1^{-1}$ is parallel to $\alpha(t_2 t_1^{-1}) = (\alpha t_2).(\alpha t_1)^{-1} = \alpha t_2$, and αt_2 is parallel to t_2. Thus, there is a contradiction unless αt_2 is the identity.

If, now, t_3 has the direction of t_1, then, as above, since αt_2 is the identity, αt_3 is also the identity.

COROLLARY 12.1. Any member of F is uniquely determined by its effect on one non-identity translation t.

Proof. If $\alpha t = \beta t$, then $(\alpha - \beta)t$ is the identity, so $\alpha - \beta = 0$, i.e., $\alpha = \beta$.

DEFINITION 8. "$\gamma(g)$" denotes the transformation $t \to g t g^{-1}$.

https://doi.org/10.4153/CMB-1961-030-1 Published online by Cambridge University Press

COROLLARY. $\gamma(g) \in F$.

Proof. Theorem 10.

THEOREM 13. Let $\alpha \in F$, $\alpha \neq 0$, and let A be a given point. There is a unique dilatation g having A as fixed point such that $\alpha = \gamma(g)$.

Proof. (Existence) Let t be a non-identity translation. Then $\alpha tA \neq A$ (theorems 9, 12). But $A + tA \parallel A + \alpha tA$, i.e. $[A, tA, \alpha tA]$, so there is a g such that

$$gA = A, \qquad gtA = [\alpha t]A.$$

Then

$$[\gamma(g)t]A = gtg^{-1}A = gtA = [\alpha t]A.$$

Hence
$\gamma(g)t = \alpha t$ (theorem 9) and $\gamma(g) = \alpha$ (corollary 12.1).

(Uniqueness) If $hA = A$ and $\alpha = \gamma(h)$, then

$$[\alpha t]A = [\gamma(h)t]A = hth^{-1}A = htA$$

so that g and h have the same effect on A and on tA. By corollary 5.1, $g = h$.

COROLLARY 13.1. The direction preserving endomorphisms are 0 and $\gamma(g)$ where g runs through the dilatations leaving A fixed.

THEOREM 14. F is a skew field.

Proof. It has already been pointed out that F is a ring. In view, then, of corollary 13.1, it is enough to observe that $g^{-1}(gtg^{-1})g = t$. ///

If A is a given point, then with any point B one can associate the translation taking A to B. This correspondence is 1-1 onto (theorem 8), and so it enables us to carry the left vector space structure from the translations over to the geometry G. To endow G with such a structure has been our main objective, and this is now achieved.

7. Equivalences. It is now possible to add ($\oplus$) points

and to multiply them by scalars (i.e. direction preserving endomorphisms $\alpha, \beta, \gamma, \ldots$).

In fact, it is easily checked that $C \oplus B = t_B C$, where t_B is the translation taking A to B; that for $\alpha = \gamma(g)$ [where $gA = A$] $\alpha C = gC$; and that $0C = C - C = A$ (i.e. A is the null-vector of our vector space).

THEOREM 15(a). For any $\alpha \neq 0$, B, the transformation $C \to \alpha C \oplus B$ is a dilatation.

Proof. By theorem 13, $\alpha = \gamma(g)$ where $gA = A$. As in the above paragraph, $\alpha C \oplus B = t_B gC$.

THEOREM 15(b). The resulting correspondence is bijective; i.e. for any dilatation h there is a unique pair (α, B), $\alpha \neq 0$, such that $hC = \alpha C \oplus B$ for all C.

Proof. (Existence) Let t be a translation such that $thA = A$, and let $g = th$. As above, $hC = t^{-1} gC = \gamma(g)C \oplus hA$.

[Note that $\gamma(g) = \gamma(t)\gamma(h) = \gamma(h)$.]

(Uniqueness) If $\alpha C \oplus B = \bar{\alpha} C \oplus \bar{B}$ for all C, then

$$(\alpha - \bar{\alpha})C \oplus (B - \bar{B}) = A \qquad \text{for all } C.$$

By theorem 15(a), this cannot happen unless $\alpha - \bar{\alpha} = 0$, in which case $B - \bar{B} = A$; i.e. $\alpha = \bar{\alpha}$, $B = \bar{B}$.

COROLLARY. The straight line $D + C$ is the set of points $\{\alpha C \oplus (1 - \alpha)D \mid \alpha \in F\}$.

Proof. Recall the definition of $D + C$, and note that $D = \alpha D \oplus B$ iff $B = (1 - \alpha)D$. ///

Theorem 15 suggests a procedure inverse to the one developed in the first six sections; namely, given any left vector space L of more than one dimension over a skew field (with vectors $V, W, \ldots$ and scalars $x, y, \ldots$) consider the group of transformations $V \to xV + W$, $x \neq 0$. For this group, axioms 1 and 2 are readily checked, which means that L becomes a geometry.

As before, the translations can be turned into a vector space $\bar{L}$. The members of the group $V \to xV + W$ which have no fixed point are clearly those for which $x = 1$. There is then a natural map from L onto $\bar{L}$; namely to each $W \in L$ associate the translation $V \to V + W$.

In theorem 13, take A to be the null-vector of L. The dilatations leaving A fixed are then those of the form $V \to xV$, $x \neq 0$, and so there is a natural map from the field of L to the field of $\bar{L}$.

These two mappings set up an isomorphism between L and $\bar{L}$; that is, from the geometry on L, we recover the original vector space structure of L. The details are left to the reader.

Conversely, if we start with a geometry G, give it a vector space structure, and then turn the vector space into a new geometry, then theorem 15 says that the new geometry is in fact the same as G.

To sum up

THEOREM 16. The geometric structure given by axioms 1 and 2 is <u>equivalent</u> to the structure of a left vector space of more than one dimension over a skew field.

The author is indebted to Professor P. Scherk for his encouragement and valuable suggestions.

Summer Research Institute,
Queen's University.

https://doi.org/10.4153/CMB-1961-030-1 Published online by Cambridge University Press