

## UNIFORM FINITE GENERATION OF THE AFFINE GROUP

FRANKLIN LOWENTHAL

**A connected Lie group  $H$  is said to be uniformly finitely generated by a given pair of one-parameter subgroups if there exists a positive integer  $n$  such that every element of  $H$  can be written as a finite product of length at most  $n$  of elements chosen alternately from the two one-parameter subgroups. Define the order of generation of  $H$  as the least such  $n$ . It is shown that the order of generation of the affine group is either 4 or 5 while its connected Lie subgroups (with two exceptions) have order of generation equal to their dimension.**

A connected Lie group  $H$  is generated by a pair of one-parameter subgroups if every element of  $H$  can be written as a finite product of elements chosen alternately from the two one-parameter subgroups. If, moreover, there exists a positive integer  $n$  such that every element of  $H$  possesses such a representation of length at most  $n$ , then  $H$  is said to be uniformly finitely generated by the pair of one-parameter subgroups. In this case define the order of generation of  $H$  as the least such  $n$ ; otherwise define it as infinity. Since the order of generation of  $H$  will, in general, depend upon the pair of one-parameter subgroups,  $H$  may have many different orders of generation. However, it is a simple consequence of Sard's theorem [4] that the order of generation of  $H$  must always be greater than or equal to the dimension of  $H$ .

The orders of generation of the isometry groups of the Euclidean and Non-Euclidean geometry are known. The order of generation of the isometry group of the spherical geometry may be any integer  $\geq 3$ ; it is determined by the cross-ratio of the fixed points of the pair of elliptic one-parameter subgroups [1]. The order of generation of the isometry group of the Euclidean geometry is infinite if both one-parameter subgroups are elliptic and it is 3 if one is elliptic and the other parabolic [2]. The order of generation of the isometry group of the hyperbolic geometry is finite if both one-parameter subgroups are elliptic, 3 if exactly one is elliptic and 4 in all other cases except that it is 6 if both are hyperbolic with interlacing fixed points [2].

Here all possible orders of generation of the affine group, i.e., the group of all transformations  $w = az + \beta$  ( $\alpha, \beta$  complex,  $\alpha \neq 0$ ) and of all its connected Lie subgroups are determined. It is shown that for the affine group the possible orders of generation are 4 and 5 while its connected Lie subgroups (excluding the isometry group of the Euclidean geometry and the group  $w = az + \beta, a > 0$ ) have order of

generation equal to their dimension.

II. Preliminaries. The affine group  $A$ , as defined above, may also be viewed as the group of  $2 \times 2$  complex invertible matrices of the form  $\begin{pmatrix} \alpha & \beta \\ 0 & 1 \end{pmatrix}$ . It is easily shown that the Lie algebra  $\mathfrak{A}$  of the affine group consists of all  $2 \times 2$  complex matrices  $\begin{pmatrix} \gamma & \delta \\ 0 & 0 \end{pmatrix}$ . These are in 1-1 correspondence with the transformations  $\varepsilon = \gamma w + \delta$  of the complex plane, so that  $\mathfrak{A}$  may be viewed as the set of these (infinitesimal) transformations; this viewpoint will be adopted throughout this paper. By identifying the (proper) subalgebras of  $\mathfrak{A}$  and using exp, it can be shown that the proper connected Lie subgroups of  $A$ , excluding the isometry group of the Euclidean geometry and the one-parameter subgroups, are (to within an inner automorphism) (cf. [3])

- (1) (a)  $w = e^{(1+bi)t}z + \beta$  ( $t$  real,  $\beta$  complex); for each fixed real number  $b$  this is a 3-dimensional subgroup of the affine group.  
 (b)  $w = z + \beta$ ; the 2-dimensional translation group.  
 (c)  $w = \alpha z$ ,  $\alpha \neq 0$ ; the 2-dimensional subgroup that leaves the origin fixed.  
 (d)  $w = az + b$  ( $a, b$  real,  $a > 0$ ); the 2-dimensional component of the identity of the real affine group.

The one-parameter subgroups of  $A$ , identified by the procedure outlined above, may be viewed as the solutions of the differential system

$$(2) \quad \frac{dw}{dt} = \gamma w + \delta \quad w(0, z) = z$$

i.e., for each  $t$ ,  $-\infty < t < +\infty$ ,  $w(t, z)$  is an element of  $A$  and the set of all solutions (2) forms a one-parameter subgroup. Under the transformations of the Lie algebra induced by an inner automorphism of  $A$ , the discriminant  $\gamma^2$  of the infinitesimal transformation  $\varepsilon$  is an absolute invariant. An infinitesimal transformation  $\varepsilon$  together with the one-parameter subgroup that it generates is classified as loxodromic if its discriminant has nonzero imaginary part, and as elliptic, parabolic or hyperbolic if its discriminant is real and respectively negative, zero or positive. The infinitesimal transformations of the subgroups of  $A$  are respectively:

- (3) (a)  $\varepsilon = (1 + bi)w + \delta$  for  $w = e^{(1+bi)t}z + \beta$ .  
 (b)  $\varepsilon = \delta$  for the translation group.  
 (c)  $\varepsilon = \gamma w$  for the group that leaves the origin fixed.  
 (d)  $\varepsilon = qw + r$  ( $q, r$  real) for the real affine group.

The affine group  $A$  as well as all its subgroups listed in (1) except for the group  $w = e^t z + \beta$  can be generated by an appropriate pair of one-parameter subgroups [3]. A pair of infinitesimal transformations

whose one-parameter subgroups generate  $A$  can be simultaneously transformed into the normal form

$$(4) \quad \varepsilon = \gamma w, \eta = \lambda(w - 1) \quad \text{Im}\left(\frac{\lambda}{\gamma}\right) \neq 0$$

by means of a suitable inner automorphism of  $A$ . To establish this one need merely determine necessary and sufficient conditions for  $\varepsilon, \eta, \sigma = [\varepsilon, \eta] = \eta(d\varepsilon/dw) - \varepsilon(d\eta/dw)$ ,  $[\varepsilon, \sigma]$ , and  $[\eta, \sigma]$  to span a 4 dimensional vector space over the reals. A similar analysis shows that a pair of infinitesimal transformations whose one-parameter subgroups generate the subgroups of  $A$  listed in (1) (except for 1(a) with  $b = 0$ ) can be simultaneously transformed into the respective normal form

$$(5) \quad \begin{aligned} (a) \quad & \varepsilon = (1 + bi)w, b \neq 0, \eta = 1; \varepsilon \text{ loxodromic, } \eta \text{ parabolic or} \\ & \varepsilon = (1 + bi)w, \eta = (1 + bi)(w - 1), b \neq 0; \text{ both loxodromic} \\ (b) \quad & \varepsilon = 1, \eta = \delta \quad \text{Im}(\delta) \neq 0 \\ (c) \quad & \varepsilon = \gamma w, \eta = \lambda w \quad \text{Im}(\lambda/\gamma) \neq 0 \\ (d) \quad & \varepsilon = w, \eta = 1; \varepsilon \text{ hyperbolic, } \eta \text{ parabolic or} \\ & \varepsilon = w, \eta = w - 1; \text{ both hyperbolic} \end{aligned}$$

by an inner automorphism.

Denote by  $T_\varepsilon(z)$  and  $S_\eta(z)$  the one-parameter subgroups generated by  $\varepsilon$  and  $\eta$  respectively. The orbit under  $\varepsilon$  of the point  $z_0$  is defined as  $\{T_t(z_0), -\infty < t < +\infty\}$ ; in an analogous manner one defines the orbit of  $z_0$  under  $\eta$ . The orbit of  $z_0$  is a line parallel to the translation vector, a circle centered at the fixed point, a ray emanating from the fixed point or a spiral centered at the fixed point as  $\varepsilon$  is respectively parabolic, elliptic, hyperbolic or loxodromic. The lemma below gives a simple sufficient condition for a curve to intersect a spiral orbit.

**LEMMA 1.** *Let  $C$  denote the range of the continuous map  $f: [0, 1] \rightarrow$  complex plane and suppose  $f(0) = 0, f(u) \neq 0$  on any subinterval  $[0, b]$ . Assume there exists a ray  $L$  emanating from the origin ( $0 \notin L$ ) and an  $a > 0$  such that  $f([0, a]) \cap L = \emptyset$ . Then in any neighborhood of the origin,  $C$  and the spiral orbit  $\{z_0 e^{it}, -\infty < t < +\infty\}$  of  $z_0 \neq 0$  under  $\varepsilon = \gamma w$  ( $\text{Im}(\gamma^2) \neq 0$ ) must intersect infinitely often. In particular, the conclusion holds if  $f'(0)$  exists and is not zero.*

*Proof.* The proof is based on the fact that  $C$  is connected; the details are omitted.

**III. THEOREM 1.** *The order of generation of all the proper connected Lie subgroups of the affine group listed in (1) (except 1(a) with  $b = 0$ ) is always equal to the dimension of the respective subgroup.*

*Proof.* (a)  $T_t(z) = e^{(1+bi)t}tz$  and  $S_s(z) = z + s$  if  $\eta = 1$  or  $S_s(z) =$

$e^{(1+bi)s}z + 1 - e^{(1+bi)s}$  if  $\eta = (1 + bi)(w - 1)$ . Let  $V(z) = e^{(1+bi)t}z + \beta = e^{(1+bi)t}(z - \sigma)$ , where  $\sigma = -\beta/e^{(1+bi)t}$ , be an arbitrary element of the subgroup. Note that  $V(\sigma) = 0$  and that  $V(0)$  lies on the spiral orbit of  $-\sigma$  under  $\varepsilon$ . The orbit of 0 under  $\eta$  is either the real axis if  $\eta = 1$  or a spiral centered at 1 if  $\eta = (1 + bi)(w - 1)$ . By Lemma 1 both curves must intersect the spiral orbit of  $\sigma$  under  $\varepsilon$ . Choose  $t_1$  and  $s_1$  such that  $S_{s_1}T_{t_1}(z)$  takes  $\sigma$  into 0. Since  $S_{s_1}T_{t_1}(z)$  is an element of the subgroup,  $S_{s_1}T_{t_1}(0)$  must lie on the spiral orbit of  $-\sigma$  under  $\varepsilon$ . Choose  $t_2$  such that  $T_{t_2}S_{s_1}T_{t_1}(0) = V(0)$ . Since  $V$  and  $T_{t_2}S_{s_1}T_{t_1}$  agree on  $\sigma$ , 0 (and  $\infty$ ), it follows that  $T_{t_2}S_{s_1}T_{t_1} = V$ .

(b)  $T_t(z) = z + t, S_s(z) = z + s\delta$ . Since  $S_sT_t(z) = T_tS_s(z) = z + t + s\delta$  to represent an arbitrary translation  $V(z) = z + \beta$  as a product of length 2 it suffices to choose  $s$  and  $t$  such that  $t + s\delta = \beta$ —this is possible as  $\text{Im } \delta \neq 0$ .

(c)  $T_t(z) = e^{\gamma t}z, S_s(z) = e^{\lambda s}z$ . Since  $S_sT_t(z) = T_tS_s(z) = e^{\gamma t + \lambda s}z$ , to represent  $V(z) = \alpha z = e^{\log \alpha}z$  as a product of length 2 it suffices to choose  $s$  and  $t$  such that  $\gamma t + \lambda s = \log \alpha$ —this is possible as  $\text{Im}(\lambda/\gamma) \neq 0$ .

(d)  $T_t(z) = e^t z$  and  $S_s(z) = z + s$  if  $\eta = 1$  or  $S_s(z) = e^s z + 1 - e^s$  if  $\eta = w - 1$ . In the first case, since  $S_sT_t(z) = e^t z + s$ , to represent an arbitrary real affine transformation  $V(z) = az + b, a > 0$  as a product of length 2 choose  $s = b, t = \ln a$ . In the second case note that  $S_sT_t(z) = e^{s+t}z + 1 - e^s$  and  $T_tS_s(z) = e^{s+t}z + e^t(1 - e^s)$ . If  $b < 1, V(z)$  can be represented as a product  $S_sT_t(z)$  by choosing  $s = \ln(1 - b)$  and  $t = \ln(a/(1 - b))$ . If  $b \geq 1$ , then  $a + b > 0$  and  $V(z)$  can be represented as a product  $T_tS_s(z)$  by choosing  $t = \ln(a + b), s = \ln(a/(a + b))$ .

IV. THEOREM 2. *The order of generation of the affine group is 4 except that if one infinitesimal transformation is elliptic and the other hyperbolic then it is 5.*

*Proof.*  $T_t(z) = e^{\gamma t}z, S_s(z) = e^{\lambda s}z + 1 - e^{\lambda s}, \text{Im}(\lambda/\gamma) \neq 0$ . Assume first that at least one infinitesimal transformation is loxodromic—without loss of generality let it be  $\varepsilon$ . To represent an arbitrary affine transformation  $V(z) = \alpha z + \beta = \alpha z + \sigma - \alpha$  where  $\sigma = \alpha + \beta = V(1)$  as a product  $T_{t_2}S_{s_2}T_{t_1}S_{s_1}(z)$  of length 4 it is necessary and sufficient that the two equations

$$(6) \quad T_{t_2}S_{s_2}T_{t_1}S_{s_1}(1) = \sigma$$

$$(7) \quad e^{\gamma(t_1+t_2)+\lambda(s_1+s_2)} = \alpha = e^{\log \alpha}$$

be satisfied. If  $t_0 = t_1 + t_2, s_0 = s_1 + s_2$  equation (7) is satisfied if and only if

$$(8) \quad \gamma t_0 + \lambda s_0 = \log \alpha \pmod{2\pi i}.$$

Since  $\text{Im}(\lambda/\gamma) \neq 0$  it is always possible to satisfy equation (8) with real numbers  $t_0$  and  $s_0$ . Thus it suffices to prove that equation (6) can be solved with real numbers  $t_1, t_2, s_1$  and  $s_2$  such that  $t_1 + t_2 = t_0, s_1 + s_2 = s_0$ . Note that as  $S_{s_1}(1) \equiv 1$ , equation (6) is independent of  $s_1$ ; hence it suffices to prove that there are real numbers  $t_1, t_2$  and  $s_2$  with  $t_1 + t_2 = t_0$  such that

$$(9) \quad e^{\gamma t_2}(e^{\lambda s_2} e^{\gamma t_1} + (1 - e^{\lambda s_2})) = \sigma$$

or equivalently that there are real numbers  $t_1$  and  $s_2$  such that

$$(10) \quad e^{\gamma(t_0-t_1)}(e^{\lambda s_2} e^{\gamma t_1} + (1 - e^{\lambda s_2})) = \sigma.$$

If  $\sigma = e^{\gamma t_0}$ , equation (10) will be satisfied if  $s_2 = 0$ . If  $\sigma \neq e^{\gamma t_0}$ , then equation (10) is equivalent to

$$(11) \quad \frac{1 - e^{\lambda s_2}}{\sigma e^{-\gamma t_0} - e^{\lambda s_2}} = e^{\gamma t_1}.$$

If  $F(s_2) = (1 - e^{\lambda s_2}) / (\sigma e^{-\gamma t_0} - e^{\lambda s_2})$ , then  $F(0) = 0, F'(0) = \lambda(1 - \sigma e^{-\gamma t_0}) \neq 0$ . Hence by Lemma 1 there exist real  $t_1$  and  $s_2$  that satisfy equation (11) and the proof in case at least one infinitesimal transformation is loxodromic is complete.

If neither infinitesimal transformation is loxodromic, it follows from (4) that one is elliptic and the other hyperbolic; assume without loss of generality that  $\varepsilon = iw, \eta = w - 1$ . First it will be shown that the affine transformation  $V_0(z) = -z + 2$  ( $\alpha = -1, \sigma = 1$ ) cannot be expressed as a product of length 4. If  $V_0(z)$  were expressible as a product  $T_{t_2} S_{s_2} T_{t_1} S_{s_1}(z)$ , equations (8) and (11) with  $\gamma = i, \lambda = 1$  require that

$$(12) \quad it_0 + s_0 = \pi i \pmod{2\pi i} \quad \text{and}$$

$$(13) \quad \frac{1 - e^{s_2}}{e^{-it_0} - e^{s_2}} = e^{it_1}$$

both be satisfied. Now (12) implies that  $e^{-it_0} = -1$ . Since

$$(14) \quad \left| \frac{e^{s_2} - 1}{e^{s_2} + 1} \right| < 1 \quad \text{for all real } s_2$$

equations (12) and (13) cannot be simultaneously satisfied.

If  $V_0(z)$  were expressible as a product  $S_{s_2} T_{t_2} S_{s_1} T_{t_1}(z)$ , then equation (12) must be satisfied together with

$$(15) \quad S_{s_2} T_{t_2} S_{s_1} T_{t_1}(0) = V_0(0) = 2.$$

Now (12) implies that  $s_0 = 0$ , i.e.,  $s_1 = -s_2$  and with this substitution (15) becomes

$$(16) \quad \frac{e^{s_2} + 1}{e^{s_2} - 1} = e^{it_2}$$

which has no solution.

To prove that in case  $\varepsilon$  is elliptic, and  $\eta$  hyperbolic the order of generation of  $A$  is 5 it suffices to represent  $V(z) = \alpha z + \beta$  as a product  $T_{t_3} S_{s_2} T_{t_2} S_{s_1} T_{t_1}(z)$  of length 5. This is possible if and only if there exist real numbers  $t_1, t_2, t_3, s_1$  and  $s_2$  such that the equations

$$(17) \quad it_0 + s_0 = \log \alpha, t_0 = t_1 + t_2 + t_3, s_0 = s_1 + s_2$$

$$(18) \quad e^{it_3}(e^{s_2}e^{it_2}(1 - e^{s_1}) + 1 - e^{s_2}) = \beta$$

can be simultaneously satisfied. Clearly it suffices to prove the existence of real numbers  $t_2, s_1$  and  $s_2$  such that

$$(19) \quad s_1 + s_2 = \operatorname{Re} \log \alpha = \ln |\alpha|$$

$$(20) \quad e^{s_2}e^{it_2}(1 - e^{s_1}) + 1 - e^{s_2} = -|\beta|$$

both hold, since if equation (20) holds  $t_3$  can be chosen so that equation (18) holds and  $t_1$  can then be chosen so that equation (17) holds. In view of equation (19) equation (20) becomes

$$(21) \quad e^{it_2}(e^{s_2} - |\alpha|) = e^{s_2} - 1 - |\beta|.$$

If  $|\alpha| = 1 + |\beta|$ , choose  $t_2 = 0$ ; otherwise a simple application of the intermediate value theorem shows that for some real  $s$  the function

$$(22) \quad F(s) = \frac{e^s - 1 - |\beta|}{e^s - |\alpha|}$$

assumes the value  $-1$ .

V. Let  $n \neq \infty$  be the order of generation of a connected Lie group  $H$  by  $T_i$  and  $S_s$ . It is of interest to determine whether every element of  $H$  can, in fact, be represented as a product of length  $n$  whose last element is a  $T_i$ ; a dual question may be asked of  $S_s$  (both questions are trivial in the commutative case). Note that any element that can be expressed as a product of length  $< n$  can be expressed both as a product of length  $n$  whose last element is a  $T_i$  and one whose last element is an  $S_s$  by inserting the identity  $I = T_0 = S_0$  an appropriate number of times.

If there is an inner automorphism of the group (or even if there is any automorphism of the group) that interchanges  $T_i$  and  $S_s$ , then both questions must have the same answer. The same conclusion holds under the quite different assumption that  $n$  is even; if an element is

not representable as a product of length  $n$  ending in a  $T_i$ , then its inverse is not representable as a product of length  $n$  ending in an  $S_s$ .

**THEOREM 3.** *If both infinitesimal transformations are loxodromic, then every element of the group  $w = e^{(1+b\epsilon)t}z + \beta$  ( $b \neq 0$ ) has both a representation as a product of length 3 ending in a  $T_i$  and one of length 3 ending in an  $S_s$ . If  $\epsilon$  is loxodromic,  $\eta$  parabolic, then every element of the group has a representation as a product of length 3 ending in a  $T_i$  but there are elements of the group that cannot be represented as a product of length 3 ending in an  $S_s$ .*

*Proof.* In both cases the existence of a product of length 3 ending in a  $T_i$  was shown in the proof of Theorem 1. If both infinitesimal transformations are loxodromic, then there is an automorphism of the group that interchanges  $T_i$  and  $S_s$ . If  $\epsilon$  is loxodromic,  $\eta$  parabolic (assume  $\eta = 1$ ) then the translations  $w = z + \beta$ ,  $\text{Im } \beta \neq 0$ , cannot be represented as a product  $S_s T_i S_{s_1}(z)$  since this would imply that  $T_i, t_1 \neq 0$ , were also a translation.

**THEOREM 4.** *If one infinitesimal transformation is hyperbolic and the other parabolic, then every real affine transformation has both a representation as a product of length 2 ending in a  $T_i$  and one of length 2 ending in an  $S_s$ . If both  $\epsilon = w$  and  $\eta = w - 1$  are hyperbolic, then all real affine transformations  $w = az + b$  with  $a + b \leq 0$  cannot be expressed as a product of length 2 ending in a  $T_i$  and those with  $b \geq 1$  cannot be expressed as a product of length 2 ending in an  $S_s$ .*

*Proof.* See the proof of Theorem 1.

**THEOREM 5.** *If at least one infinitesimal transformation is loxodromic, then every affine transformation has both a representation as a product of length 4 ending in a  $T_i$  and one of length 4 ending in an  $S_s$ . If  $\epsilon = iw$  is elliptic,  $\eta = w - 1$  is hyperbolic, then every element of the affine group can be represented as a product of length 5 ending in a  $T_i$  but none of the transformations  $w = -az + b$ ,  $a > 0$ ,  $b - a \geq 1$  can be represented as a product of length 5 ending in an  $S_s$ .*

*Proof.* Assume at least one infinitesimal transformation is loxodromic. That every affine transformation can be represented as a product of length 4 ending in a  $T_i$  was, in fact, established in the proof of Theorem 2; since the order of generation is even, the dual result for  $S_s$  follows.

Assume  $\epsilon = iw$ ,  $\eta = w - 1$ . The assertion concerning products ending in a  $T_i$  again was established in the proof of Theorem 2. Next observe

that a simple modification of the argument used in Theorem 2 to show that  $V_0(z) = -z + 2$  was not representable as a product  $T_{t_2}S_{s_2}T_{t_1}S_{s_1}(z)$  shows that all transformations  $V(z) = -az + b$ ,  $a > 0$ ,  $b - a \geq 1$  (and hence  $V(1) \geq 1$ ) are not so representable. If  $V(z) = -az + b$  were equal to  $S_{s_3}T_{t_2}S_{s_2}T_{t_1}S_{s_1}(z)$ , then it would follow that

$$(23) \quad S_{s_3}^{-1}V(z) = T_{t_2}S_{s_2}T_{t_1}S_{s_1}(z).$$

But one may directly verify that  $S_{s_3}^{-1}V(z) = -cz + d$ ,  $c > 0$ ,  $d - c \geq 1$  (note that the ray  $x > 1$  as well as  $\{1\}$  are both orbits under  $\eta$ ).

#### REFERENCES

1. F. Lowenthal, *Uniform finite generation of the rotation group*, Rocky Mountain J. Math., **1** (1971), 575-586.
2. ———, *Uniform Finite Generation of the Isometry Groups of Euclidean and Non-Euclidean Geometry*, Canadian J. Math., **23** (1971), 364-373.
3. ———, *On Generating Subgroups of the Moebius Groups by Pairs of Infinitesimal Transformations*, Pacific J. Math., **26**, (1968), 141-147.
4. S. Sternberg, *Lectures on Differential Geometry*, Prentice Hall, Englewood Cliffs, N. J., (1964), 45-55.

Received September 30, 1970.

UNIVERSITY OF OREGON