Nexus Netw J
DOI 10.1007/s00004-015-0269-1

NEXUS NETWORK JOURNAL
Architecture and Mathematics

RESEARCH

Mulling Over Shapes, Rules and Numbers

Ramesh Krishnamurti[1]

Abstract This paper explores the relationships between geometric constructability, numbers, shapes and shape grammars. Shapes are based on compositional constructs in geometry, which rely upon drawing instruments. Implementing shape grammars relies upon numeric encodings, properties of which specify whether shape algorithms are decidable and/or tractable.

Introduction

Nothing fundamentally new, just a different look at geometry and numbers, shapes and rules, which I hope appeals to Lionel.

We are often surprised by some of the intricate geometric forms produced in software. In a cognitive and aesthetic sense these surprises are real. I am interested in a different kind of surprise, a surprise of the mechanistic kind, where one questions whether forms produced by digital means are really radically distinct from those drawn by hand. What is surprising to me is how difficult it is to get a computer to achieve exactly what we intend it to do. It is in this context that I explore geometric constructability and shape grammars by comparing the construction of shapes through analogue and digital devices; in short, I connect Euclid (1956), Descartes (1954) and Stiny-Gips (1972). Euclid gave us mechanical drawing instruments, namely, the straightedge and compass; Descartes, an algebraic mechanism, namely, the coordinate system; and Stiny-Gips, an automaton, a

Dedicated to Lionel March.

Ramesh Krishnamurti
ramesh@cmu.edu

1 School of Architecture, Carnegie Mellon University, College of Fine Arts 201, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA

Published online: 01 September 2015

Birkhäuser

system of rules, namely, the shape grammar. Of particular interest is the classification of shapes and shape grammars: that is, the relationship between tractable shape grammar computation and geometric constructability, in other words, between shapes based on a number system and shapes based on specific drawing instruments. In this paper, shapes are restricted to planar geometrical figures made up of finite lines.

The gist of this article is as follows. In principle, a computer system for geometry specifies a universe of constructible shapes. Compositional constructs in such systems are based on encodings of Euclidean constructs, which rely on drawing instruments to aid manual construction. Thus, we can classify shapes according to their constructive device.[1] Encodings are based on numbers, which can be classified according to specific properties. There is an association between number systems and drawing instruments through the Cartesian coordinate system. Thus, we can classify shapes according to their numeric properties. Shape grammars are compositional systems that can be used to specify particular languages of shape. Implementing shapes *precisely* in a computer system depends upon determinate processes; practical implementation relies upon the tractability of such algorithms.

In writing this article, to make for a compelling narrative, I include succinct summaries of known work instead of being content to just cite references. Sections "Drawing Instruments" through "Geometric Construction" draw heavily upon established material and information.

Drawing Instruments

In the process, classes of shapes that can be produced using 'mechanical' drawing instruments are examined, translating the mechanics of geometric construction into algorithms and procedures. When doing so, one arrives at a classification of shapes (and hence, shape grammars) in terms of their constructive devices.

Euclidean geometry relates to construction restricted to the straightedge (ruler) and compass. With a ruler one can construct a (infinite) line, which passes through two given points; with a compass one can construct a circle (or circular arc) centred on one point and passing through another; lines and circles may intersect and in doing so result in new points from which new lines and new circles can likewise be constructed. Shapes are defined as made up of finite lines, that is, segments, with end-points specified by such constructed points.

In Euclidean geometry, geometric figures are either constructible or otherwise. Of the latter category, three celebrated problems are squaring the circle, doubling the cube and trisecting an angle. The first construction is impossible; the latter two are solvable by verging or the insertion principle.[2] Viète's construction for trisecting an angle centred at O is given in Fig. 1.

[1] Geometric entities such as NURBS, splines and other constructive curves do exactly this.

[2] Given two curves and a point, the line touching the two curves at points corresponding to marked locations on a straightedge *verging* through the given point, is said to be drawn by the *insertion principle*, which was discovered by Fançois Viète in 1593.

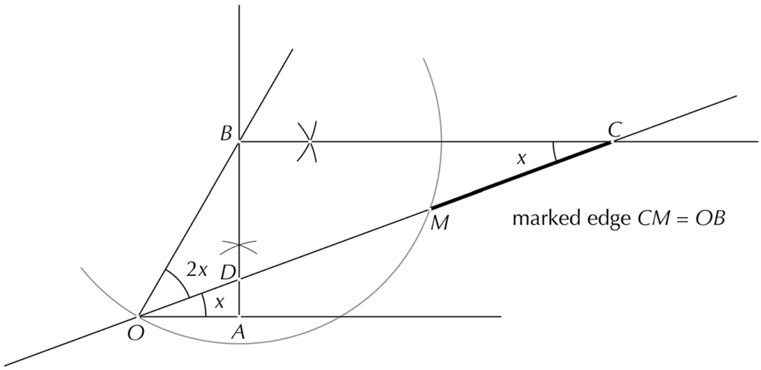


Fig. 1 Trisecting an angle using a 'marked' straightedge

Notation The notation –*AB*…–, *AB*…–, *AB*, *A*(*B*), *A*(*r*), $\angle AOB$, $\angle O$ and ΔABC represent a line, a ray through points *A*, *B*, …, a line segment between points *A* and *B*, a circle centred at *A* passing through *B*, a circle of radius *r* centred at *A*, angles centred at *O* and a triangle with points *A*, *B* and *C*. When context is understood, *AB* and $\angle$ also represent distance and angle measure.

The proof is straightforward ($OM = OB$, from triangles ΔBMC and ΔOBM angle calculations show $\angle AOD = x$, that is, an angle trisection). As seen in the figure, the construction requires a 'marked' straightedge, where distance *CM* (=*OB*) is marked on the straightedge. This is distinct from normal Euclidean constructions, which rely solely on a compass and an *unmarked* straightedge as the constructions for the perpendicular lines –*AB*– and –*BC*– clearly show. In this construction, the straightedge has to be aligned so that the marked points *C* and *M* are coincident with respectively the horizontal line and the circular arc centred at *B*. Note that the alignment of points is non-determinate although the fact that the points can be aligned is certain.

For purposes of this paper, Euclid's original devices will be augmented with two other drawing instruments: divider and marked ruler. The former adapts a compass for transferring distance whilst the latter extends the straightedge with two designated marked points. The four devices—ruler, compass, divider and marked ruler—are referred to as the *standard* drawing instruments.

The insertion principle, of course, suggests that one might consider a range of drawing implements as part of the standard toolkit. However, analogue drawing tools such as the rusty compass (fixed circle), double or parallel ruler, right angle, cannon, stick (fixed divider), tomahawk (trisector), cissoid (a cubic curve used to duplicate a cube), paper-folds (or origami folds), etc., constructively used either singly or in combination are equivalent to the standard instruments.

Lastly, congruence can be established by three or fewer reflections and similarity is congruence with scaling (see Sect. "Scale Transformation and Reflections"). Reflection and scale can be accomplished by ruler-compass constructions. The constructions are illustrated in Fig. 2. Consequently, for any pair of corresponding points, it is always possible to construct two possible pairings of similar triangles

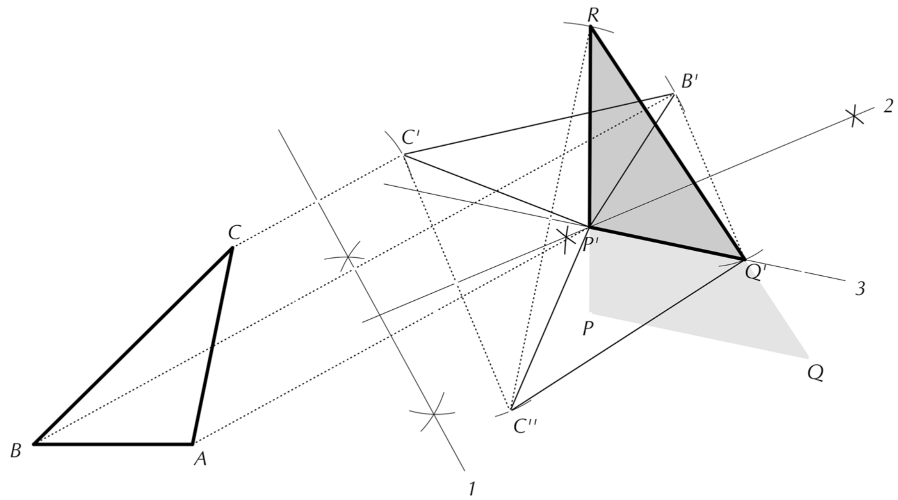


Fig. 2 A similarity is an isometry composed with a scale. An isometry is specified by no more than three reflections (triangles ΔABC and $\Delta P'Q'R$ are congruent). *1, 2* and *3* represent the axes of reflection. Triangles ΔPQR and $\Delta P'Q'R$ are similar by a scale about *R*. Some ruler-compass compass constructions are indicated.

formed by three potentially corresponding points. It then follows that testing for similarity between shapes given two corresponding points is Euclidean constructible. Note that—apropos our motivation—such constructions are independent of any numeric specification for points.

Constructible Numbers

George Martin, in his book, *Geometric Constructions* (1998) examines the connection between (combinations of) standard drawing instruments and spaces of constructible numbers, each satisfying the properties of an algebraic field, and each based on the union of extensions of the space of rational numbers. Shapes are based on points, which are specified by coordinates with values from these constructible numbers. Geometric transformations between shapes defined over any specific field of constructible numbers are shown to have coefficients in the same field. Decidability of shape grammars then hinges upon whether these constructible numbers can be represented so as to ensure determinate equality checking. The constructible numbers form a sequence from the space of rational numbers to the space of real numbers, each a subspace of the next space in the sequence. In this way, shapes can be classified in terms of their standard drawing instruments, and shape grammar computation is closed within the number fields represented by these instruments. The material below is chiefly taken from Martin.

Number Fields

The coordinates of a point satisfy the properties of an algebraic *field*, which is a subset of the real numbers closed under addition and multiplication with an additive and multiplicative inverse, 0 and 1. Whenever a, b, $c \neq 0$, are in field F, $a + b$, $a - b$, ab and a/c are also in F.[3] A *subfield* is a subset of a field that satisfies the properties of a field. The real numbers, **R**, form a field, as do its subfield of rational numbers, **Q**. The set of integers, **Z** does not, although it is a subset of every number field. That is, if F is a number field, $\mathbf{Z} \subset F$.

Quadratic Extension

Suppose c is a positive number in F but $\sqrt{c}$ is not in F. The set $F(\sqrt{c}) = \{a + b\sqrt{c} | a$ and b in $F\}$ is termed a *quadratic extension* of F. $F(\sqrt{c})$ is a field.[4] Quadratic extensions have the following sum of squares properties:

1. For all a, b and c are in F with $\sqrt{c}$ not in F, $a + b\sqrt{c}$ is a sum of squares in $F(\sqrt{c})$ if and only if $a - b\sqrt{c}$ is a sum of squares in $F(\sqrt{c})$.
2. For all a, b and $c \neq 0$ that are sums of squares in F, then so too are $a \pm b$, ab, and a/c.

If $F_1 = F(\sqrt{c_1})$, $F_2 = F_1(\sqrt{c_2})$, …, $F_n = F_{n-1}(\sqrt{c_n})$, we write $F_n = F(\sqrt{c_1}, \sqrt{c_2}, \ldots, \sqrt{c_n})$. Each of F_1, F_2, …, F_n is an *iterated quadratic extension* of F. Each F_i can be shown to be a field: that is, F_i is closed under +, –, × and ÷. Moreover, F_i is closed under $\sqrt{(.)}$, namely, quadratic extension.

Consider a *tower* of fields over fields F_i, $i = 0, 1, \ldots, n$ built by an iterated quadratic extension of the rational numbers, **Q**:

$$F_0 = \mathbf{Q}, F_1 = \mathbf{Q}(\sqrt{c_1}), \ldots, F_i = \mathbf{Q}(\sqrt{c_1}, \sqrt{c_2}, \ldots, \sqrt{c_i}).,\ \ldots, F_n = \mathbf{Q}(\sqrt{c_1}, \sqrt{c_2}, \ldots, \sqrt{c_n}).$$

That is, one can write such numbers using just integers and the symbols +, –, × , ÷ and $\sqrt{\ }$. An example of such radical expressions is:

$$123\sqrt{1 + 4567\sqrt{89}} - \sqrt{4 + \sqrt{5 - \sqrt{6}} + 7\sqrt{10 - 3\sqrt{11}}}.$$

Euclidean Field

A field F is *Euclidean* whenever a in F and $a > 0$ implies that $\sqrt{a}$ is in F. That is, every positive number in a Euclidean field is a square in the field.

Let **E** denote the union of all iterated quadratic extensions of **Q**. **E** is a field. By definition, **E** is Euclidean. As **Q** is the smallest rational field, **E** is closed under

[3] Fields are also associative, commutative and distributive.

[4] $F(\sqrt{c})$ contains 0 and 1. If $p = a + b\sqrt{c}$ and $q = d + e\sqrt{c}$, then $p \pm q = (a \pm d) + (b \pm d)\sqrt{c}$ are in $F(\sqrt{c})$. Likewise, $pq = (ad + bec) + (bd + ae)\sqrt{c}$ is in $F(\sqrt{c})$. Consider $q^{-1} = (d - e\sqrt{c})/(d^2 - ec) = f + g\sqrt{c}$, where f and g are again in F. q^{-1} is in $F(\sqrt{c})$ and therefore, $p/q = pq^{-1}$ is in $F(\sqrt{c})$.

quadratic extension, and therefore, **E** is the smallest Euclidean field. A number is *Euclidean* if it is in **E**.

Pythagorean Field

A field F is *Pythagorean* whenever a, b in F implies that $\sqrt{a^2+b^2}$ is in F. That is, the sum of squares in the field is a square in the field. Equivalently, if $a \neq 0$, $\pm a\sqrt{1+(b/a)^2}$ is in F.

Suppose a, b are in F, but $\sqrt{c}$ is not in F where c is the sum of two squares. Then, $F(\sqrt{c}) = \{a + b\sqrt{c}|a$ and b in $F\}$ is a *Pythagorean extension* of F. $F(\sqrt{c})$ can be shown to be a field. If F_{i+1} denotes a Pythagorean extension of F_i for $i = 0, 1, 2, \ldots, n-1$, then $F_1, F_2, \ldots, F_n$ are each termed an *iterated Pythagorean extension* of F $(= F_0)$. Note that a Pythagorean extension is a quadratic extension with the following additional sum of squares properties.

3. Whenever d in F is a sum of squares in its Pythagorean extension $F(\sqrt{c})$, d is a sum of squares in F
4. Whenever d in F is a sum of squares in the iterated Pythagorean extension $F(\sqrt{c_1}, \sqrt{c_2}, \ldots, \sqrt{c_n})$, d is a sum of squares in F

Let **P** denote the union of all iterated Pythagorean extensions of **Q**. **P** is a field. Note that is in **P** for every a and b in **P**. For every $c_1, c_2, \ldots, c_{n-1}, c_n$, we have:

$$\sqrt{c_1^2+c_2^2+\cdots+c_{n-1}^2+c_n^2} = \sqrt{\sqrt{(c_1^2+c_2^2+\ldots+c_{n-1}^2)}^2+c_n^2}.$$

That is, the sum of every square in **P** is a square in **P**. Whence, **P** is a Pythagorean field. As **Q** is the smallest rational field, **P** has no Pythagorean extension, and therefore, **P** is the smallest Pythagorean field. Note that **P** is closed under quadratic extension. A number is *Pythagorean* if it is in **P**.

A Euclidean field is necessarily Pythagorean as the sum of squares is nonnegative; however, the converse is not the case.[5] That is, $\mathbf{P} \neq \mathbf{E}$ and $\mathbf{Z} \subset \mathbf{Q} \subset \mathbf{P} \subset \mathbf{E} \subset \mathbf{R}$.

Dioclesian, Glotin and Vietean Fields

There are other number fields that sit between **E** and **R**. Consider the roots of cubic equations. If a cubic equation $ax^3 + bx^2 + cx + d = 0$ has no rational root, then none of its roots are in **E**. The equation $x^3 - 3x^2 - 2\,cos\,3t = 0$ has roots: $2\,cos\,t$, $2\,cos\,(t + 2\pi/3)$, $2\,cos\,(t + 4\pi/3)$. When $t = \pi/3$, $cos\,t = ½$, its roots are in **E**. When t is a third this, that is, $t = \pi/9$, none of its roots are rational, thus, not Euclidean, although all its roots are real On the other hand, $x^3 - 2 = 0$ has only one real root from which it follows that $\sqrt[3]{2}$ is not Euclidean A Dioclesian[6] field F is

[5] $1 + \sqrt{5}$ is in both **P** and **E** whereas $\sqrt{1+\sqrt{5}}$ is in **E** but not in **P**.

[6] Named for Diocles (circa 180 BC) who invented the cissoid.

closed under cube root if x in F implies $\sqrt[3]{x}$ is in F. A Glotin[7] field F is *closed under trisection* if *cos* $3t$ in F implies *cos* t is in F. A Euclidean field is *Vietean* if it is closed under cube root and trisection. Let **V** be the smallest Vietean field closed under cube root and trisection. Likewise, let **G** and **D** denote the smallest Euclidean fields that are closed under, respectively, trisection and cube root. $\mathbf{E} \subset \mathbf{V} \subset \mathbf{R}$. Moreover, $\mathbf{V} = \mathbf{G} \cup \mathbf{D}$, $\mathbf{D} \cap \mathbf{G} = \emptyset$. **V** includes the real roots of all quartic and cubic equations. If an equation has only one real root it is in **D**; otherwise, it is in **G**.

Geometry

The standard drawing instruments produce lines and circles. We now examine them in relation to algebraic fields. The constructions described in this paper are registered in the Cartesian plane.

Terminology Suppose coordinates of points are defined in field F. A point is an *F-point* if it has all its coordinates in field F. An *F-line* is specified by two F-points. An F-circle has an F-point for its centre and an F-point on its circumference. We extend the terminology to include shapes and shape grammars, namely, an *F-shape* is specified by a set of maximal F-lines and an *F-shape grammar* is defined on F-shapes. In this way, appropriately, we may classify points, lines circles, shapes and shape grammars as rational, Pythagorean, Euclidean, Glotin, Dioclesian, Vietean, real, etc.

Lines and Circles

The equation of a line passing through two F-points (x_1, y_1) and (x_2, y_2) is given by:

$$X(y_2 - y_1) - Y(x_2 - x_1) + (x_2 y_1 - x_1 y_2) = 0.$$

As the equation depends only on the definition of a field, it follows then that the coefficients are also in F. That is, F-lines have equations with coefficients defined in F.

A similar argument is given for circles. The general equation of a circle with centre (x_1, y_1) and passing through (x_2, y_2) is given by:

$$(X - x_1)^2 + (Y - y_1)^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2,$$

or,

$$X^2 + Y^2 + (-2x_1)X + (-2y_1)Y + (x_2(2x_1 - x_2) + y_2(2y_1 - y_2)) = 0.$$

As the above expression depends only on the definition of a field, it follows then that the coefficients are also in F. That is, F-circles have equations with coefficients defined in F.

[7] Named for Pierre Glotin who in 1863 made the first serious study of the trisector.

Intersections

Next, consider the intersection of lines and circles.

Suppose two lines have equations $a_1X + b_1Y + c_1 = 0$ and $a_2X + b_2Y + c_2 = 0$.

The point of intersection of the two lines is given by (x_{int}, y_{int}) where

$$x_{int} = (b_1c_2 - b_2c_1)/(a_1b_2 - a_2b_1) \text{ and } y_{int} = (a_2c_1 - a_1c_2)/(a_1b_2 - a_2b_1),$$

provided that $(a_1b_2 - a_2b_1) \neq 0$.

However, when $(a_1b_2 - a_2b_1) = 0$, the lines are parallel and therefore, do not intersect. The coordinates of the points of intersection thus depend only on the definition of a field. That is, F-lines intersect at F-points.

Line with equation $aX + bY + c = 0$ and circle with equation $X^2 + Y^2 + fX + gY + h = 0$ intersect at the points (x_{int}, y_{int}) where

$$d = (fb - ag)^2 + 4c(af + gb - c) - 4h(a^2 + b^2)$$

$$x_{int} = (abg - 2ac - b^2f \pm b\sqrt{d})/\left(2\left(a^2 + b^2\right)\right)$$

$$y_{int} = (abf - 2bc - a^2g \pm (-a\sqrt{d}))/\left(2\left(a^2 + b^2\right)\right)$$

Without loss in generality, suppose that the coefficients a, b, c, f, g and h are in F. For the line and circle to intersect, d must be nonnegative. There are two possibilities: either d is square in which the points of intersection are in F, or else, if d is not square in F, then at least one of x_{int} or y_{int} is not in F, since both a and b cannot simultaneously be zero. That is, x_{int} and y_{int} are both in $F(\sqrt{d})$, a quadratic extension of F.

When a number field F is closed under quadratic extension, it follows that F-lines and F-circles intersect F-points. Since $\mathbf{Q}$ is not closed under quadratic extension, rational circles do not necessarily intersect rational lines at rational points.

Lastly, the pair of equations:

$$X^2 + Y^2 + aX + bY + c = 0$$

$$X^2 + Y^2 + fX + gY + h = 0$$

is equivalent to the pair of equations:

$$(a-f)X + (b - g)Y + (c - h) = 0$$

$$X^2 + Y^2 + fX + gY + h = 0$$

This system of equations has solutions either in F or in a quadratic extension of F.

Points on a Perpendicular

Consider the line through two F-points (x_1, y_1) and (x_2, y_2). Consider the line perpendicular at (x_2, y_2). Consider the expressions:

$$x_3, x_4 = x_2 \pm (y_1 - y_2) \text{ and } y_3, y_4 = y_2 \pm (x_2 - x_1).$$

(x_3, y_3) and (x_4, y_4) represent two points on the perpendicular, which are equidistant from (x_2, y_2) as (x_2, y_2) is from (x_1, y_1). Clearly, (x_3, y_3) and (x_4, y_4) are F-points.

Again, consider the line through two F-points (x_1, y_1) and (x_2, y_2). The foot of the perpendicular from an F-point (x_3, y_3) not coincident with the line satisfies:

$$x_{foot} = \left(b^2x_3 + aby_3 - ac\right)/\left(a^2 + b^2\right) \text{ and } y_{foot} = \left(abx_3 + a^2y_3 + bc\right)/\left(a^2 + b^2\right),$$

where $a = (y_2 - y_1)$, $b = (x_2 - x_1)$, and $c = (x_2y_1 - x_1y_2)$. As the equations depend only on the definition of a field, it follows that (x_{foot}, x_{foot}) is an F-point.

Scale Transformation and Reflections

When computing with a shape grammar, there may be multiple possible sub-shapes to which a rule applies. Each sub-shape corresponds to a similarity transformation that maps a shape (the left side of a rule) to the sub-shape in the given (current) shape. A similarity transformation can always be specified by a scale transformation and an isometry (Martin 1982: 141).

Let $aX + bY + c = 0$ be the equation of a line with coefficients in field F. The reflection of any point (x, y) in this line is given by (x', y'):

$$x' = x - 2a(ax + by + c)/\left(a^2 + b^2\right)$$

$$y' = y - 2b(ax + by + c)/\left(a^2 + b^2\right)$$

Suppose (x, y) is an F-point. Then, the reflected point (x', y') is also an F-point.

The midpoint of two points (x_1, y_1) and (x_2, y_2) defined over F has coordinates in F. Let $aX + bY + c = 0$ be the equation of a line passing through these two points, where a, b and c are in F. The equation of a line perpendicular to this line has equation $bX - aY + d = 0$. If this line passes through the midpoint, then d is in F. That is, the line perpendicular to the line joining two F-points at their midpoint is an F-point. This line is the axis of reflection, which maps (x_1, y_1) onto (x_2, y_2) and vice versa.

By the definition of a field, compositions of reflections have coefficients in F. Two congruent figures are related by an isometry, which can always be defined by composing three or fewer reflections (Martin and George 1982: 35). That is, if the coefficients of reflection are in defined in F, then so too are the coefficients of the isometry that specifies the congruence. A similarity transformation corresponds to a scale composed with an isometry. If similar triangles are defined by F-points, then the scale factor is in F. That is, coefficients of a similarity transformation are defined in F.

Geometric Construction

Each specific set of drawing instruments is associated with a number field. Lines and circles are basic geometric constructions and each requires two points for their definition. Accordingly, we have a starter set S with at least two points in the Cartesian plane.

Let M denote a drawing machine, that is, a subset of the standard drawing instruments.

We define a *M-point* to be the last in a finite sequence $P_1, P_2, \ldots, P_n$ of points such that each point P_i is in S or is obtained in the following possible ways, termed an *M-intersection*:

(i) Intersection of two lines each passing through two points that appear earlier in the sequence;
(ii) Point of intersection of a line that pass through two points that appear earlier in the sequence and a circle centred at an earlier point and passing through an earlier point;
(iii) Point of intersection of two circles, each centred at an earlier point and each passing through an earlier point;
(iv) Point P_i is on a line –AB–, $A \neq B$, such that $AP_i = CD$ where A, B, C and D are points that appear earlier in the sequence;
(v) Either one of two points that are one unit apart, that are collinear with a point that appears earlier in the sequence, and that are such that each lies on a different line through two earlier points in the sequence.

In each case, the construction depends on specific drawing instruments. For instance, (i) through (iii) require both ruler and compass with (i) requiring just a ruler and (iii) just a compass. Rule (iv) accounts for the divider, and (v) for the marked ruler. Similar M-intersection rules can be enunciated for other drawing instruments, for instance, the stick, parallel-ruler, right angle, tomahawk, cissoid, paper-fold, etc.

By convention, points in the starter set are considered M-points.

An *M-line* passes through two M-points. An *M-circle* is a circle centred at an M-point passing through an M-point. A number (or measure) x is a *M-number* if and only if $(x, 0)$ is a M-point, that is, if $(x, 0)$ is M-constructible. Note the terminology can be naturally extended to include *M-shapes* and *M-shape grammars* as shapes and shape grammars that are M-constructible.

M-numbers form a field; this can be demonstrated by constructions for the arithmetic operations. If M includes the compass, an M-number can be extended by a quadratic by constructing the square root of the M-number. The following inductive argument illustrates how M-numbers can be constructed.

Consider the coordinates of any M-point, P. By definition, P is the last point in a sequence $P_1, P_2, \ldots, P_n$. Each P_i is in S, or is obtained as the M-intersection of two points that appear earlier in the sequence. We can associate P_1 with a rational number. By the arguments presented in Sect. "Quadratic Extension", each point P_i,

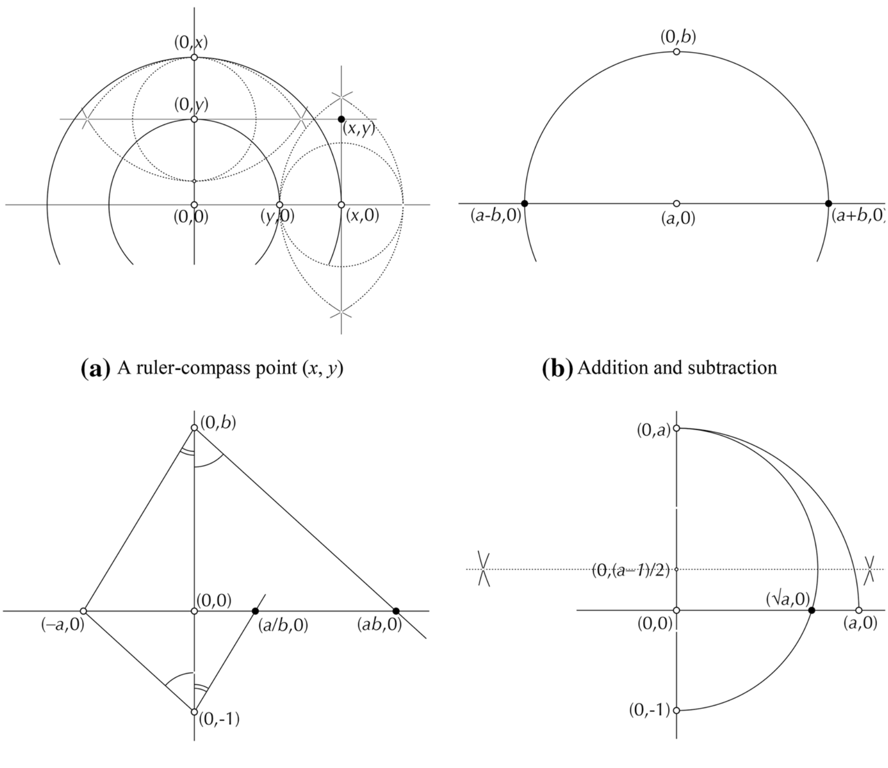
$i > 1$, can be associated with a field F_i, which either equals $F_{i\text{-}1}$ or else, if possible, is a quadratic extension of $F_{i\text{-}1}$.

In the case of ruler-compass constructions we choose a starter set with two points a unit apart, say $S = \{(0, 0), (1, 0)\}$. Figure 3 gives constructions for a ruler-compass point, and for arithmetic and square root operations. The ruler-compass numbers form the Euclidean field **E**.

Consider using just a ruler to produce shapes. As a single line cannot intersect itself, we need at least two lines. Alternatively, we start with a quadrilateral from which new points can be constructed. Given two parallel lines we can bisect a line using just a ruler. Likewise, given a line and its midpoint we can construct a line parallel to it through a given point again just using a ruler. These simple constructions are known as the *trapezoid* theorem.

By definition, the intersection of two ruler-lines is a ruler point. Figure 4 shows construction of ruler points from the starter set, say, $S = \{(0, 1), (1, 0), (2, 0), (0, 2)\}$ using the trapezoid theorem.

The trapezoid theorem can be adapted to show that the midpoint of two ruler-points is a ruler-point, and the line through a ruler-point parallel to a ruler-line is a



(a) A ruler-compass point (x, y)

(b) Addition and subtraction

(c) Multiplication and division

(d) Square root

Fig. 3 The ruler-compass numbers form a field and its quadratic extensions

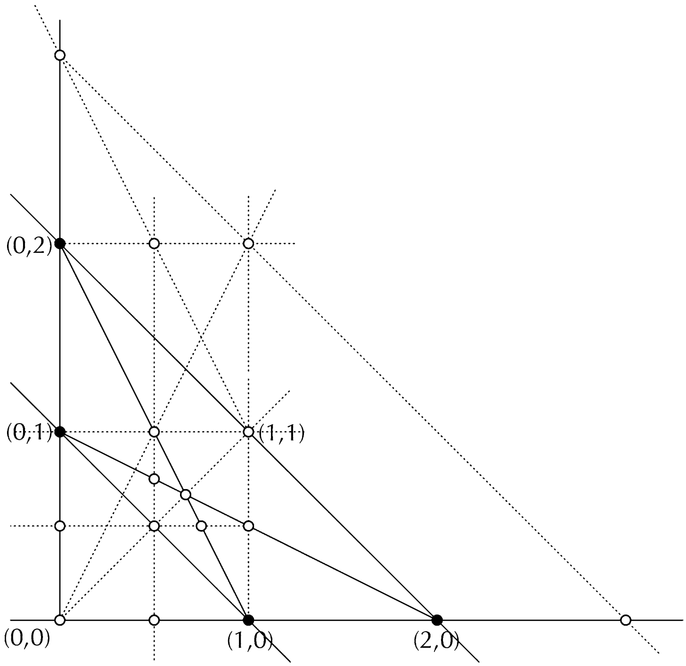


Fig. 4 Construction of ruler points from a starter set {(1,0), (0,1), (2,0), (0,2)}

ruler-line. A ruler-line contains at least three ruler-points. This construction is particularly useful in the arsenal of ruler construction techniques as it justifies the existence of ruler points on or off particular constructed lines. An immediate consequence is that all integer points in the set $\{(\pm\ n, 0), (0, \pm\ n) \mid n$ is an integer$\}$ are ruler points.

Two variations of a 'push-up pull-down' construction show that ruler-numbers form a field. Suppose a, b and c are ruler-numbers. That is, $OA = a$, $OB = b$, $OC = c$. D is a ruler-point off the X-axis. The line through D parallel to –OA– is a ruler-line.

In the first variation B is pushed onto the ruler line through D off the X-axis and pulled back onto the X-axis to produce ruler points X such that $BX = OA$. That is, if a and b are ruler numbers, then so too are $a + b$ and $a - b$. In the second variation, there is a point E on the ruler-line through D such that neither –BE– nor –CE– are parallel to –OD–. The construction lines –OD– and –BE– intersect at F, –DE– and –AF– at G, –CE– and –OD– at H, and –HG– and –OB– at X. X is a ruler-point and $OB/OA = OC/OX$. By choosing $c = 1$, we have $OX = a/b$. By choosing $OC = b$ and $OB = 1$, we have $OX = ab$. Since integers are ruler-numbers, by the above construction it follows that rational numbers are also ruler-numbers. That is, the field of the ruler numbers is **Q**.

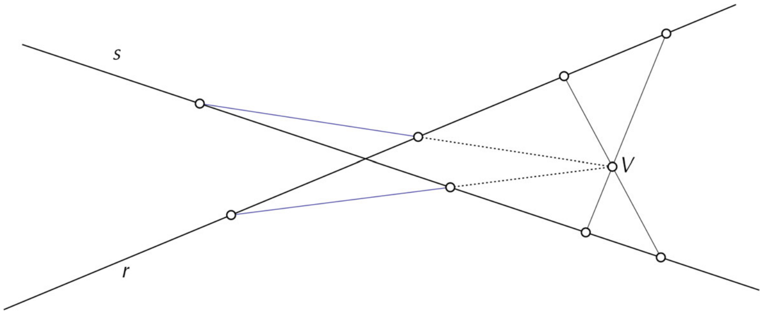


Fig. 5 Four possible ways of verging with a point

Ruler and divider constructions subsume ruler constructions. $O = (0, 0)$ and $(0, 1)$ are ruler-divider points. Suppose $P = (a, b)$ is a ruler divider point. By (iv) there is a ruler-divider point $X = (0, x)$ such $OX = OP$. That is, $x = \sqrt{a^2 + b^2}$. The field of ruler-divider numbers is thus Pythagorean.

Marked ruler constructions subsume ruler constructions with an added verging construction. Figure 5 shows four possible ways of verging with point V between lines r and s, Verging produces points the coordinates of which are solutions to polynomial equations of degree at most four with coefficients in $\mathbf{V}$.

In a similar manner one can establish relationships between specific drawing tools and specific number fields for example, the tomahawk and cissoid constructions are equivalent to the Glotin and Dicoclesian fields respectively.[8] Both are subfields of the Vietean field and as previously stated such constructions are equivalent to marked-ruler constructions.

M-constructible Shape Grammars

Computing with shapes is central to shape grammar research (Stiny 2006); I have previously written on the subject (Krishnamurti 1980, 1981; with Giraud 1986; with Stouffs 1993). One way of looking at shapes is through the manner by which they are created—typically, through a process of manipulation and change of basic geometric forms. A shape grammar computation is expressed by the equation:

$$v = u - f(a) + f(b) \text{ with } f(a) \le u.$$

[8] There are a number of celebrated results relating to construction by specific drawing tools. For example, Georg Mohr in 1672 (and independently discovered in 1797 by Lorenzo Mascheroni) showed that every ruler-compass point is a compass point and conversely. Jean-Victor Poncelet suggested in 1822 and Jakob Steiner proved in 1833 that every ruler-compass point is equivalent to a ruler-fixed circle point. A ruler construction is also known as Steiner construction. August Adler in 1890 showed the equivalence of ruler-compass and parallel-ruler constructions. Sundara Row (1893) introduced paper-fold constructions; subsequently, in 1945 Yates proved their equivalence to marked-ruler constructions. In 1939, Dawson introduced matchstick geometry, which turns out to be equivalent to ruler-compass constructions. See Eves (1972) and Martin (1998).

Variables v, u, a, and b represent shapes, in particular, a, and b specify the shape rule $a \rightarrow b$, and f is a geometric transformation which maps a to some part of u. The symbol $\leq$ denotes the part or sub-shape relationship.

There are two issues to consider: shape arithmetic and determining f. Stiny, in his thesis (1975) lays out the complete theoretical foundation for doing shape arithmetic and the recognition of sub-shapes, upon which all subsequent worthwhile[9] implementations are based.

Doing arithmetic on shapes is a process of producing maximal lines using 'reduction' rules (Stiny 1975, 2006; Krishnamurti and Stouffs 2004). Under addition, collinear point-wise non-disjoint lines are combined to form longer lines; likewise, under subtraction, such lines are combined to produce shorter, even empty, lines.

Determinacy of f depends upon a (Krishnamurti 1981). For line shapes in the plane, when there are at least two distinguished points in a, it is always possible to find all possible geometric transformations f, for which $f(a) \leq u$. A point in a shape is *distinguished* if it is possible to find a corresponding point in the other shape that preserves the same spatial relationship under the transformation. For example, the point of intersection of two lines in a shape is distinguished if it is mapped to the point of intersection of two corresponding lines in the other shape. Likewise, the foot of the perpendicular from a distinguished point to a line in the shape is distinguished. In the absence of such distinguished points, f is defined by a set of representative transformations from which all possible indeterminate f's can be generated. Distinguished points are closed within a number field (see Sect. "Points on a Perpendicular").

Decidability of shape computation depends on whether two shapes can be checked for equality. This requires encoding of shapes; that is, by embedding shapes within a coordinate system (Stiny 1975). When shapes are specified in terms of rational coordinates, then such computations are always decidable (Krishnamurti 1981). The question explored here is whether for other classes of shapes similar computations are decidable and practically possible.

Implementing arbitrary shape grammars have proven difficult. In many ways this has to with our ability to exactly compute shape rule application, which requires the encoding of points as definite numerical expressions. Exact computation requires the facility to performing exact arithmetic and comparing numbers for equality.

Shape arithmetic on shapes takes time linear in the number of its maximal lines (Stouffs and Krishnamurti 1993). That is, the time complexity of an algorithm depends on the complexity of doing numerical calculations.

Arithmetic on rational numbers is straightforward. So too is comparing pairs of rational numbers a/b, c/d for equality, which is easily determined by testing whether $ad - bc$ equals zero. These require unit time.

In general, constructible numbers involve radical expressions, which can be parenthesized and thus data-structured as trees. As quartic equations are solvable algebraically, it should be possible to represent Euclidean and Vietean numbers in this manner. Comparing trees for equality is straightforward—a simple in-order tree traversal achieves this in linear time. On the other hand, ensuring that a radical

[9] Computer implementations that intrinsically deal with ambiguity in shape rule application.

expression is irreducible is not.[10] Reducing radical expressions is neither intuitive nor straightforward. There are illustrative examples in the literature: for instance, determining something as simple as

$$\sqrt{5+2\sqrt{6}}=\sqrt{2}+\sqrt{3},$$

or

$$\sqrt{3+2\sqrt{3}}=\frac{1}{4}\sqrt[4]{12}(2+\sqrt{12})$$

can be hard.

Radical expressions can be nested to arbitrary depth, for example,

$$\sqrt{16-2\sqrt{29}+2\sqrt{55-10\sqrt{29}}}=\sqrt{5}+\sqrt{11-2\sqrt{29}}.$$

Expressions can be more deeply nested such as the doubly nested expression below:

$$\sqrt{(112+70\sqrt{2})+(46+34\sqrt{2})\sqrt{5}}=(5+4\sqrt{2})-(3+\sqrt{2})\sqrt{5}.$$

Irreducible radical expressions in combination can be reduced further:

$$\sqrt{1+\sqrt{3}}+\sqrt{3+3\sqrt{3}}-\sqrt{10+6\sqrt{3}}=0.$$

Reducing a radical expression to its irreducible form is termed *denesting* (Landau 1994). The best-known algorithms to denest certain kinds of radical expressions take exponential time (Landau 1992; Borodin et al. 1985). The Borodin algorithm is specifically directed at denesting expressions involving just square roots. Euclidean numbers fall into this classification. Consequently, while it may be possible to deal with exact shape grammar computation for certain Euclidean or Pythagorean numbers, it is not essentially tractable and therefore, unlikely to yield practical implementations for Euclidean or Pythagorean shape grammar interpreters. Constructing Vietean numbers by mechanical means is a 'trial and error' process – essentially a procedure. Whether one can do so precisely by analytic means remains an open question.

Sticks

In 1980 Lionel (mainly) and I (in a minor way) wrote a research proposal to The Leverhulme Trust entitled, *A New Grammar of Ornament*, "the purpose of which is to carry out preparatory design work which will be constructively defined through 'the principles beyond appearance,' which are now known to us." In the spirit of that proposal I would like to conclude this paper by exploring an aspect of shapes made up of 'matchstick' lines specified by 'matchstick' points.

[10] Tongue in cheek, unless one can impersonate the famous Indian mathematician Srinivasa Ramanujan who made the subject popular by producing with some astonishing results on continued fractions and radical expressions. See (Berndt and Bruce 1998).

Imagine having an unlimited supply of sticks of unit length. Endpoints of sticks are stick points; others arise as a result of intersecting sticks. We can place sticks on known stick points to create new stick points. There are four possible ways—say rules—as illustrated in Fig. 6 where the open node represents the new stick point.

From the starting stick points, two are assumed, say: {(0, 0), (½, 0)}. By just these first two rules we can produce overlaying triangular tilings. By rule 4 we create stick-points where sticks overlap (see Fig. 7). The stick points have coordinates as indicated, namely, of the form $(2m\ {}^1/_4,\ 2n\ {}^{\sqrt{3}}/_4)$ or $((2m + 1)\ {}^1/_4,\ (2n + 1)\ {}^{\sqrt{3}}/_4)$. That is, $(2m, 2n)$, $(2m + 1, 2n + 1)$ where the x- and y- axes are scaled by ${}^1/_4$ and ${}^{\sqrt{3}}/_4$, respectively. Computing with such radicals is simple and computationally tractable. Term the set of stick points generated as $S_1 = \{(2m, 2n), (2m + 1, 2n + 1) \mid m, n \text{ are integers}\}$.

As a result of further overlaying two rhombic tilings as shown in Fig. 8 we can generate stick-points of intersection by applying rules 2 and 4 repeatedly. Here

$$\begin{aligned} S_2 = S_1 \cup \Big\{ &\Big((2m \pm 1 - \sqrt{5})^1/_4, (2n+1)^{\sqrt{3}}/_4\Big), \\ &\Big((\sqrt{5} + 4m - 1)^1/_8, (4n+1)(\sqrt{5} \pm 1)^{\sqrt{3}}/_8\Big), \\ &\Big((4m + 5 - \sqrt{5})^1/_8, (4n+1)(\sqrt{5} \pm 1)^{\sqrt{3}}/_8\Big) \,|\, m, n \quad \text{are integers}\Big\}. \end{aligned}$$

Again, computation with such radical numbers, although tedious, will still be tractable.

The set grows exponentially. The rapidity with which the number of stick points is produced at each generation depends on the start set. Any start set with two points that are less than a stick's length apart will suffice.

We could repeat this process ad infinitum. In the limit, $\lim_{n \to \infty} S_n = \mathbf{E}$, the Euclidean numbers and the best known denesting algorithms are computationally intractable. In fact anything that is ruler-compass constructible is stick constructible, which is an interesting result in its own right. See (Dawson 1939; Martin 1998: Ch. 8) for a proof as well as for interesting examples of stick constructions. The point at which the process breaks down from a tractable set of stick numbers to an intractable set of stick numbers is an open question. However, with enough stick points we can make interesting stick shapes and consequently interesting stick shape grammars, which are computationally tractable.

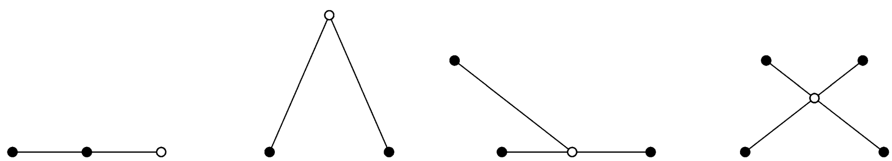


Fig. 6 Rules for constructing a new stick point from known

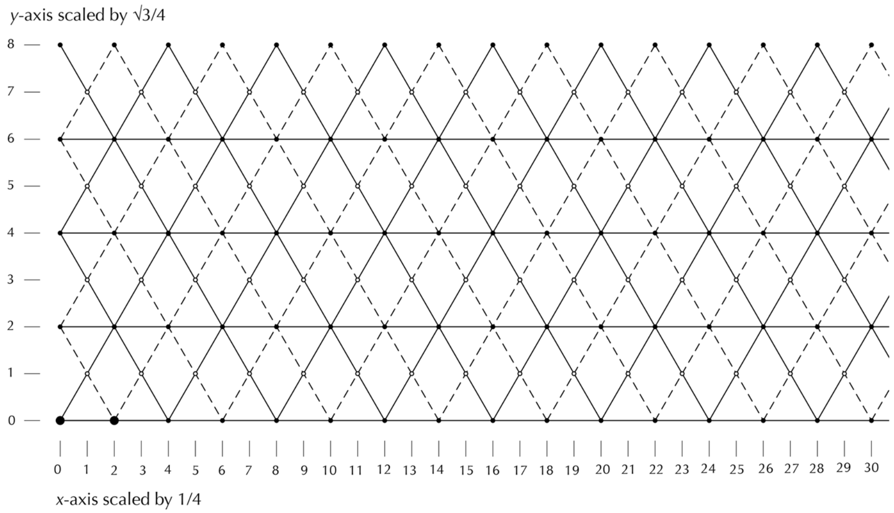


Fig. 7 Constructing a set of stick points by overlaying two triangular tilings created by using rules 1, 2 and 4. The x- and y-coordinates are scaled by $^1/_4$ and $^{\sqrt{3}}/_4$ respectively

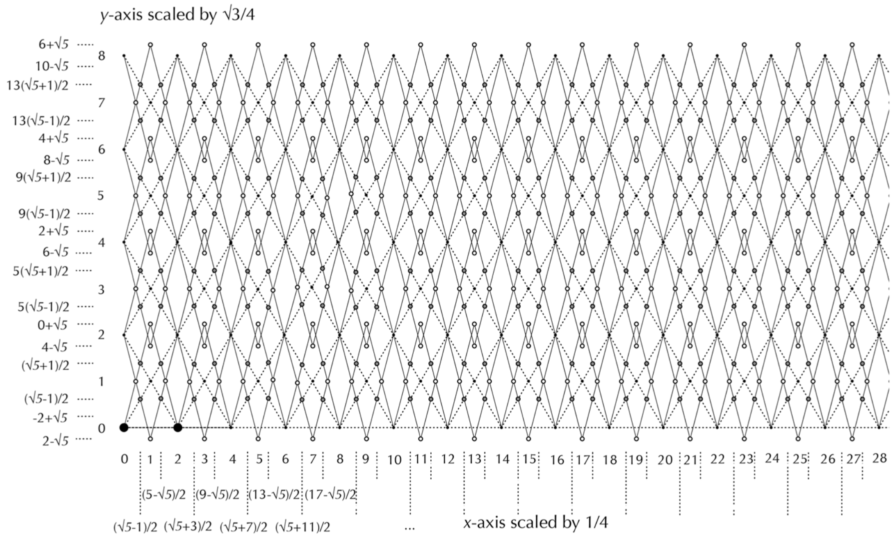


Fig. 8 Constructing more stick-points by further overlaying two rhombic tilings produced by applying rules 2 and 4

An Ending

A lack of time renders this article necessarily incomplete. I began in hopes of solving a problem that has long vexed me, namely, of being able to explain how to do on a machine what one can do naturally and with ease by hand; more generally to specify the universe of shapes constructed by various classes of machine. For me it has been a journey of discovery … about numbers and geometry. I implemented the prototypical shape grammar interpreter that catered for ambiguity (Krishnamurti 1982); at the time I gravitated towards working with rational shapes not from a consideration of number fields, but because such shapes supported precise calculation, and also from a general dislike of the floating point. Numbers expressed as π and e have an aesthetic in the way that 3.14159… or 2.71828… do not.

Over the years I have looked at this problem and many a time have put it aside. Memories of a youthful dexterity with the aptly named surds ineluctably induced long hours in attempts to denest arbitrary radical expressions, a feat that has bested better, more agile minds. I had also hoped to illustrate this paper with examples of an exotic kind of shape grammar, say, for origami forms using shape rules that embody paper-folds. Settling upon matchsticks, I can claim to commune with Euclid. Lionel would approve.

Forty years ago Lionel took me under his wing and introduced me to the world of combinatorial configurations—my first attempt was an algorithm to generate 'poly-animals' on regular tilings; this led to exploring designs on Archimedean tessellations (Krishnamurti and Roe 1979). Parametric modellers have provided the modern designer with an easy means of working with Vornonoi and other popular tilings, which are normally defined over a set of simple fixed-point coordinates. Configurations on such tilings are readily rational and shape grammars defined thereupon decidedly tractable. Emulating Owen Jones still remains a goal: *A Grammar of Ornament* (Jones 1856) presents an imagery of patterns, symmetries and colours drawn upon from various cultures. There are contained these challenges: What is an image? What is a pixel? What is its shape? Conventionally, an image no matter its resolution is always a rational shape made up of rectangular pixels located at rational points. Imagine instead a more interesting world of Pythagorean images made up of Pythagorean pixels located at Pythagorean points.[11] In many ways Lionel's own fascination with numbers, shapes, rules, symmetry, colour, and in general, geometry has captured my wonderment and for that I thank him. Hopefully, this paper encourages others to explore this fascinating world of geometry, numbers, rules and shapes.

References

Berndt, Bruce C. 1998. An overview of Ramanujan's notebooks. In *Charlemagne and His Heritage: 1200 Years of Civilization and Science in Europe, vol. 2: Mathematical Arts*, P. L. Butzer, H.Th. Jongen and W. Oberschelp, eds., pp. 119–146. Turnhout: Brepols. (http://www.math.uiuc.edu/~berndt/articles/aachen.pdf).

Borodin, Allan, Fagin, R., Hopcroft, J. E., Tompa, M. 1985. Decreasing the Nesting Depth of Expressions Involving Square Roots. *Journal of Symbolic Computation* **1**: 169–188.

[11] It is always possible to tessellate any collection of Pythagorean points in the plane as a Voronoi tiling where each tile and its barycentre are respectively a Pythagorean shape and point.

Eves, Howard. 1972. *A Survey of Geometry*. Revised Edition. Boston: Allyn and Bacon Inc.
Dawson, T.R. 1939. "Match-stick" Geometry. *Mathematical Gazette* **23**:161–168.
Euclid. 1956. *The Thirteen Books of Euclid's Elements*, 3 vols. Thomas L. Heath, ed. and trans. New York: Dover.
Jones, Owen. 1856. *A Grammar of Ornament*. Rpt. London: Omega Books.
Krishnamurti, Ramesh. 1980. The arithmetic of shapes. *Environment and Planning B: Planning and Design* **7**: 463–484.
Krishnamurti, Ramesh. 1981. The construction of shapes. *Environment and Planning B: Planning and Design* **8**: 5–40.
Krishnamurti, Ramesh. 1982. SGI: an interpreter for shape grammars. Technical report, Centre for Configurational Studies, Design Discipline, The Open University. (http://www.andrew.cmu.edu/~ramesh/pub/distribution/technical/SGI.pdf).
Krishnamurti, Ramesh, Giraud, C. 1986. Towards a shape editor: an implementation of a shape generation system, *Environment and Planning B: Planning and Design* **13**: 391–404.
Krishnamurti, Ramesh, Roe, P. H. O'N. 1979. On the generation and enumeration of tessellation designs, *Environment and Planning B: Planning and Design* **6**: 191–260.
Krishnamurti, Ramesh, Stouffs, R. 1993. Spatial Grammars: Motivation, Comparison and New Results. In *CAAD Futures 93* (Edited by U. Flemming and S. Van Wyk) Amsterdam: North-Holland. 57-74. (http://www.andrew.cmu.edu/~ramesh/pub/distribution/conference/cf1993-pittsburgh.pdf).
Krishnamurti, Ramesh, Stouffs, R. 2004. The boundary of a shape and its classification. *Journal of Design Research* **4**(1): 28. (http://www.inderscience.com/jdr/backfiles/articles/issue2004.01/stouffs.pdf).
Landau, Susan. 1992. Simplification of Nested Radicals. *SIAM Journal of Computing* **21**(1): 85–110.
Landau, Susan. 1994. How to Tangle with a Nested Radical. *The Mathematical Intelligencer* **16**(2): 49–55.
Martin, George E. 1982. *Transformation Geometry: An Introduction to Symmetry*. New York: Springer-Verlag.
Martin, George E. 1998. *Geometric Constructions*. New York: Springer-Verlag.
Descartes, René. 1954. *The Geometry of René Descartes*. Edited and translated by David E. Smith and Marcia L. Latham. New York: Dover Publications.
Stiny, George, Gips, J. 1972. Shape grammars and the generative specification of painting and sculpture. In *Information Processing 71* (Edited by C.V. Freiman) Amsterdam: North-Holland. 1460–1465.
Stiny, George. 1975. *Pictorial and Formal Aspects of Shape and Shape Grammars*. Basel: Birkhäuser.
Stiny, George. 2006. *Shape: Talking About Seeing and Doing*. Cambridge: MIT Press.
Stouffs, Rudi, Krishnamurti, R. 1993. The complexity of the maximal representation of shapes. In *Proceedings of the IFIP WG 5.2 Workshop on Formal Design Methods for CAD* (Edited by J.S. Gero) Dordrecht: Kluwer. 53–66. (http://www.andrew.cmu.edu/~ramesh/pub/distribution/workshop/complexity.pdf).
Sundara Row, T. 1893. *Geometric Exercises in Paper Folding*. Chicago: The Open Court Publishing Company.

Ramesh Krishnamurti has a BE (Honours) in electrical engineering from the University of Madras, a BA in computer science from the University of Canberra, and MASc and PhD in systems design from the University of Waterloo. He has taught and worked in Canada, United Kingdom, United States and Taiwan. He is currently a Professor at Carnegie Mellon University, where he directs the Graduate Programs in the School of Architecture. His research focuses on the formal, semantic, generative, and algorithmic issues in computational design. His research activities have had a multidisciplinary flavor and include shape grammars, generative designs, spatial topologies, spatial algorithms, geometrical and parametric modeling, sensor-based modeling and recognition, analyses of design styles, knowledge-based design systems, integration of graphical and natural language, interactivity and user interfaces, graphic environments, computer simulation, 'green' CAD and war games.