Escher's Art, Smith Chart, and Hyperbolic Geometry

The purpose of this article is to point out a conceptual link between an artistic creation by the well-known Dutch artist M.C. Escher (1898–1972), called *Circle Limit IV* (woodcut, 1960), and the most commonly used graphical aid in microwave engineering work, called the Smith chart (1939–1944), created by the U.S. engineer P.H. Smith (1905–1987). The basis of Escher's art and of the Smith chart can both be traced back to the invariance of the cross ratio of four complex numbers under a Möbius transformation on the domain of complex numbers. When measured using a hyperbolic distance metric that is induced by the invariant cross ratio in Poincare's open-disk model of hyperbolic space, the visually different geometrical figures in Escher's work are found to have a mosaic-like fixed size and periodicity. The Smith chart can be used as an aid in constructing other Escher-like drawings that display periodic mosaic patterns and at the same time convey the perception of infinite progression within a unit circle.

The public image of an engineer in the media notwithstanding, engineers do enjoy, engage in, contribute to, enable, and inspire the fine arts, both personally and professionally. This is to be expected, because engineering, like art, requires creativity, discipline, attention to detail, and sensitivity to human perception. The vibrant interface between technology and art is exemplified by numerous aesthetic and creative works, artifacts, and exhibits from a variety of art forms, including music, dramatics, photography, sculptors, and paintings [1]–[3]. This article points out yet another little-known interface between art and engineering tools, specifically that connecting certain woodcuts by Escher and the Smith chart pervasive in microwave engineering. The readers can use it to enhance their appreciation of that art, to stimulate an interest in microwaves among nonprofessionals, or to create Escher-like art themselves.

Madhu S. Gupta

Madhu S. Gupta is with San Diego State University, San Diego, CA 92182 USA, +1 619 594 7015, mgupta@mail.sdsu.edu.

This article is an enlarged transcript of the invited talk given by the author at MIKON-2006, Krakow, Poland, in May 2006.

More information about MIKON-2006 can be found in "Transnational News" on p. 94.

1527-3342/06/$20.00©2006 IEEE

Escher and His Art

Artist with Appeal Among Technologists

There have been a number of artists, such as Leonardo da Vinci, whose work holds a special appeal and affords a deeper level of appreciation to the technologically oriented viewers. One of the most popular among the 20th century artists of this kind is Escher, whose works have fascinated scientists and engineers for about one-half century. They have been reproduced in scientific journals, magazines, and monographs and have appeared on the covers of textbooks, posters, calendars, and in popular media. One reason for their appeal is Escher's use of such figures as regular polyhedra, periodic designs, mirror images, and objects like the Möbius strip, which are familiar to the technologists and that convey the sense of harmony and order in his drawings. A second reason for their popularity has been the unusual ways in which space is treated in the works, through reflections, stretching, deformation, projection, and other such transformations, which have struck a chord with viewers having a technological inclination. This is particularly remarkable given that Escher himself was not a professional mathematician or even trained in mathematics. His intuitive manipulation of spatial regions, arrived at from aesthetic grounds, nevertheless represent sophisticated mathematical tools, such as conformation mapping and hyperbolic geometry, thus suggesting that these mathematical operations are not entirely arbitrary in the abstract sense but have a relationship to human perception.

Escher's Life

Maurits Cornelis Escher was born in Leeuwarden, Holland, on 17 June 1898 and, along with four brothers, grew up in Arnhem. While his father was a civil engineer and three of his brothers pursued science or engineering fields, he was not mathematically inclined and pursued his interest in graphic arts. He attended the School of Architecture and Decorative Arts in Haarlem, where he learned woodcut technique from Samuel de Mesquita and thereafter migrated to Italy in 1922 to settle in Rome, where he lived until 1934. During this period, he rendered a large number of sketches of scenery and buildings from southern Italy. In 1935, he left Italy for Switzerland, where he lived for two years, followed by three years in Belgium, and finally in 1941, settled in Baarn, Holland, for the next three decades until his death in 1972. A self-portrait of Escher from 1935 is shown in Figure 1.

Escher's Works

Escher's creations, produced over four decades, include some 450 works of art, including woodcuts, wood engravings, lithographs, and drawings [4]. Although there is a large variety in Escher's lifetime of output, several enduring themes can be identified in his creations, of which the following six are the principal ones. An illustrative example of each is shown in Figure 2.

Landscapes

During his early years in Italy, Escher produced sketches of landscapes, both real and imaginary, with a striking visual effect due to his clever choice of vantage points or use of light, shadow, and color.

Unusual Perspectives

Escher produced a number of sketches showing fine details in things ranging from the mundane objects or scenes of daily life to the architectural details of monumental buildings but with unusual vantage points and with captivating results.

Symmetries and Periodicities

Tessellations are forms or arrangements of periodic, mosaic-like patterns used as adornments. A large number of Escher's works drew inspiration from the art of the Moors who occupied Spain from 711 to 1492 and decorated walls and floors with congruent

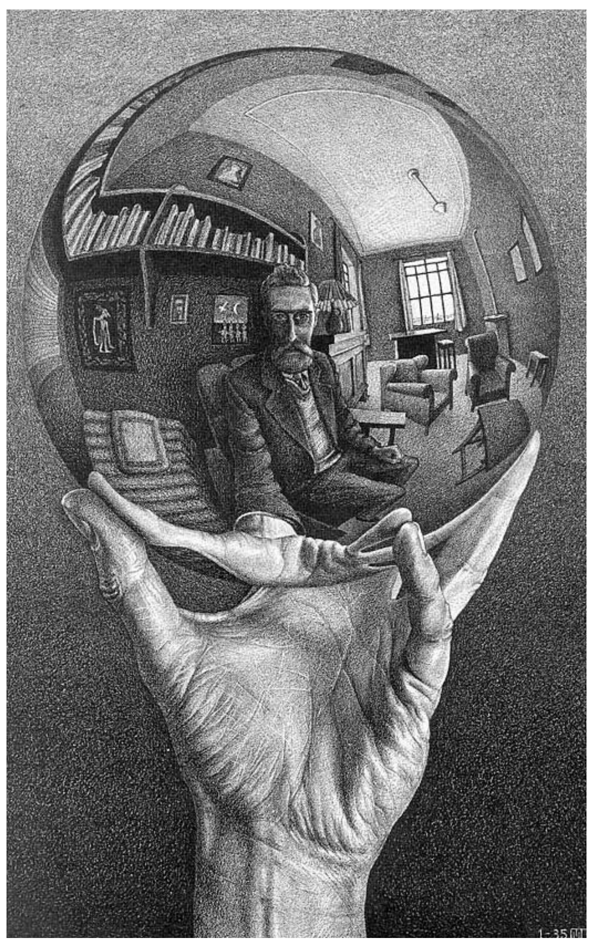

Figure 1. *A self-portrait of M.C. Escher (1898–1972) in spherical mirror, dating from 1935 titled* Hand with Reflecting Globe.

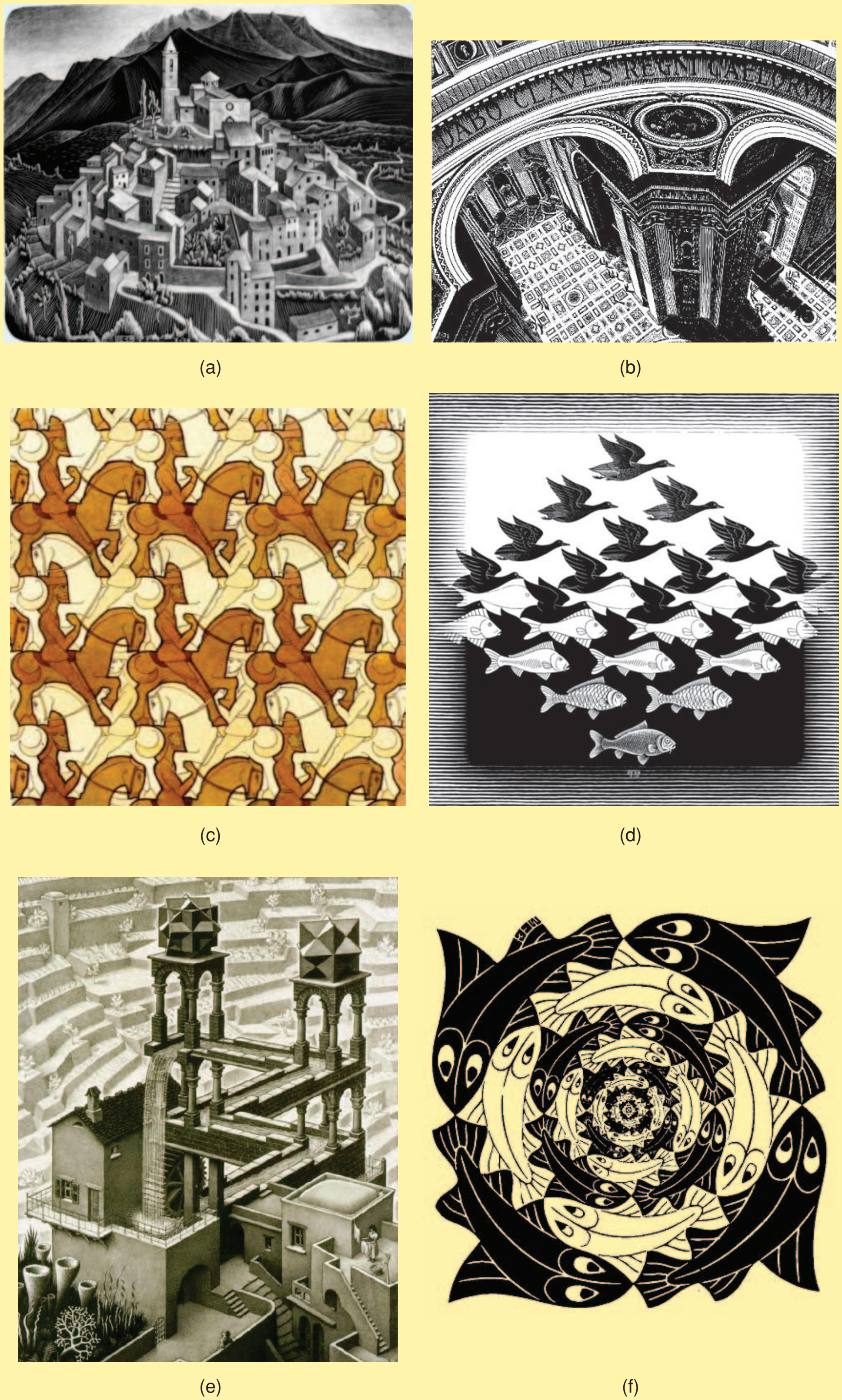

Figure 2. *Representative examples of Escher's art, illustrating the major themes in his work. (a) Unusual landscapes.* Goriano Sicoli, Abruzzi *(1929). (b) Unusual perspectives.* Inside St. Peter's *(1935). (c) Tessellations.* Regular Division of Plane III *(1957). (d) Gradual transformations.* Sky and Water I *(1938). (e) 2-D–3-D illusions.* Waterfall *(1961). (f) Representations of infinity.* Fish Vignette *(1956).*

multi-colored tiles that covered the surface completely. However, while the Moors were forbidden from depicting animate objects due to religious reasons, Escher made it his hallmark. His drawings illustrate many types of symmetries, which have delighted crystallographers and group theorists [5], [6].

Gradual Transmutations

A distinctive motif in many of Escher's creations and in some of his most recognized works is the gradual transformation of one figure or tessellation into another in successive tiles of a mosaic-like periodic drawing. Combined with the unusual perspectives, they create an arresting effect.

Two- to Three-Dimensional Illusions

The representation of a three-dimensional (3-D) object in a two-dimensional (2-D) picture results in some ambiguities in human observation and thus provides many opportunities for optical illusions. Escher was a master of this art form of tricking the human eye in delightful ways.

Representations of Infinity

A number of Escher's works convey a graphical impression of the infinite within a finite boundary by suggesting an indefinite continuation of a tessellation in space. Works of this type will be further examined here.

Representation of Infinity with Tessellations

The concept of infinity has long had the aura of mystery and intrigue among mathematicians and laymen alike [7], [8]. Many of Escher's mosaic-like works suggest a sense of infinity, due to their periodicity and unlimited extendibility [8], but in practice, they must come to an abrupt stop at the boundary of the artwork. Clearly, more is required to convey the sense of truly infinite extent, and Escher was fascinated with such representation of infinity in a finite space. Late in his career, starting in the 1950s, Escher made a number of attempts to represent an infinite mosaic within a circular or square boundary, as illustrated in Figure 3.

Escher attempted to capture the perception of an infinite continuation within a finite area through a progressive reduction in tile size, either toward the center or toward the edge of the drawing. In one of his early attempts [Figure 3(a)], the approach to the infinite occurs toward the center of the drawing with the shrinking of the tile size. This was reversed in subsequent works [Figure 3(b) and (c)], where the size reduction occurs toward the periphery of the drawing as the infinite is approached. Escher himself found these earlier works unsatisfying and continued to look for ways to improve upon them. His later works include two woodcuts, titled *Circle Limit III* [Figure 3(d)] and *Circle Limit IV* [Figure 3(e)], which are the best-known works of this genre. The procedure for the construction of tessellations with a fixed figure of any desired outline has been described in the literature [9], [10], as has the analysis of the geometrical structure of *Circle Limit III* [11], [12]. In this article, we analyze the structure of *Circle Limit IV* in detail.

Engineers enjoy, engage in, contribute to, enable, and inspire the fine arts, both personally and professionally.

The Structure of *Circle Limit IV*

Tessellations Constructed by Transformations

Considered one of Escher's masterpieces, *Circle Limit IV*, also titled *Heaven and Hell*, was completed in July 1960 (Figure 4). The original is a woodcut, printed in black and ochre, measures 416 mm in diameter, and shows angels and demons in a tessellation that completely fills the plane. The size of these figures gradually diminishes from the center toward the edge, with over two dozen different sizes identifiable, until they merge into the visual limit achievable in a woodcut. Although the figures are different in size in a Euclidean sense, we will show that they are indeed congruent when measured with a hyperbolic metric of distance; viewed in this way, the work is simply a periodic mosaic with a constant tile size and periodicity in the hyperbolic space.

The first requirement and the principal defining attribute of a mosaic is the repeated appearance of a unit element after periodic spatial displacements. Escher's specialty was in introducing a gradual transformation in the unit element with each successive repetition. For a transformed figure to still be recognized as a transformation of the previous figure, it is necessary that certain basic traits of the figure remain unchanged. The transformation to be applied within the unit element should therefore have some invariants, and those invariants should have a visually appealing geometrical manifestation. This is the basis of many of Escher's tessellations that display gradual change.

A second requirement imposed here is the need to fit an infinite lattice on a finite canvas, which in turn implies that the sizes and the spatial displacements of the unit element in the mosaic cannot be constant. A second transformation is therefore needed to be applied to the size and displacement of the unit element upon each successive replication. As a result of this transformation, the tile dimensions should approach zero, thereby producing the desired perception of infinity.

Figure 3. *Escher's attempts to represent the infinite in various ways. (a)* Smaller and Smaller *(1956). (b)* Circle Limit I *(1958). (c)* Circle Limit II *(1959). (d)* Circle Limit III *(1959). (e)* Circle Limit IV *(1960). (f)* Square Limit *(1964).*

The woodcut *Circle Limit IV* (Figure 4) and other works in this category possess both of the above two features—a repeated pattern subjected to a continuing alteration and a periodicity subject to shrinking with each repetition of the tile. Therefore, such works must employ two simultaneous transformations. In the case of *Circle Limit IV*, the only transformation applied to the unit element, such as the figure of an angel, is a rotation. The transformation applied to the size and displacement of the unit tile is considerably more involved and is the subject of our scrutiny below.

Escher's Circle Limit IV *and Smith Chart*

The clue that the Smith chart might be relevant to an understanding of Escher's graphical art lies in the radial scales typically provided with a Smith chart. Among the several scales for reading reflection coefficients, return loss, and standing wave ratio in various units, there is one scale that expresses the voltage standing-wave ratio (VSWR) in decibels, and extends from zero at the center of the Smith chart to ∞ at the periphery of the chart. This range corresponds to that required by Escher in an attempt to represent infinity graphically.

For a preliminary test of this idea, the following simple measurement can be performed on *Circle Limit IV*. While most of the angels depicted in the drawing are visually asymmetric, those with bilateral symmetry have an axis of symmetry that passes through the center of the circle. The height of a bilaterally symmetric angel (defined in the radial direction from head to foot) can therefore be obtained by measuring, for each of the two ends, the radial distance from the origin, and taking the magnitude of the difference between their radial coordinates. If we scale *Circle Limit IV* to fit exactly within the unit circle of a Smith chart and measure the radial distance of a point with the help of the scale provided for VSWR in decibels, the results shown in Table 1 are obtained. It is found that all bilaterally symmetric angels measure about 8 dB in height. This result encourages the conjecture that the figures in *Circle Limit IV* are congruent with each other and holds out the hope that an engineering tool such as the Smith chart might be relevant to understanding an aesthetic artwork. Demonstrating that the other angels without bilateral symmetry are also congruent requires a more detailed analysis.

The VSWR scale allows us to measure the distance only along the radial direction in the unit circle. To measure the distance between arbitrary pairs of points Γ_1 and Γ_2 within the unit circle, we must examine the geometrical origin of the VSWR scale used with the Smith chart. Such generalized measures of distance have already been in use in microwave work in a number of contexts, as will be pointed out later.

The Smith chart was originally intended to be a graphical aid for eliminating the drudgery of computation with complex numbers.

Smith Chart and Möbius Transformation

The Smith Chart as a Graphical Aid

The Smith chart (Figure 5) has become an icon of microwave engineering. For example, it is often used in the design of logos and is the most remembered part of a microwave engineering curriculum decades later when everything else learned in the classroom has been forgotten. The Smith chart was originally

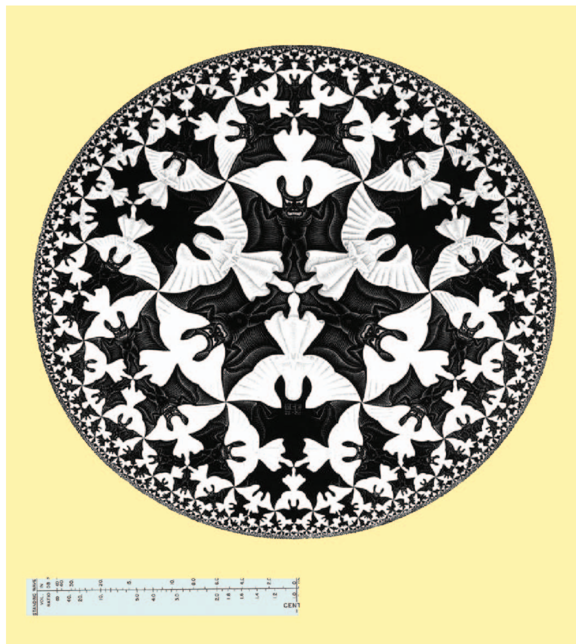

Figure 4. *Escher's woodcut titled* Heaven and Hell, *also known as* Circle Limit IV *(July 1960), placed in the unit circle along with the VSWR scale from the Smith chart.*

TABLE 1. Measurements of heights of figures in *Circle Limit IV*.

Figure	Radial Coordinate of the Head, H	Radial Coordinate of the Head, F	Height of the Figure, $\|H - F\|$
Angel #1	8.3 dB	0 dB	8.3 dB
Angel #2	7.2 dB	15.5 dB	8.3 dB
Angel #3	23.5 dB	15.3 dB	8.2 dB
Angel #4	22 dB	30.2 dB	8.2 dB
Angel #5	39 dB	31 dB	8.0 dB

Measurements limited to bilaterally symmetric Angels. *Circle Limit IV* is superimposed over the Smith chart, and radial distances from center are measured using the VSWR (in decibels) scale provided with the Smith Chart.

intended to be a graphical aid for eliminating the drudgery of computation with complex numbers [13], [14]. While that has no longer been a necessity since the appearance of electronic calculators in the 1960s, the Smith chart remains highly useful for visual representation and comprehension of information.

The Möbius transformation in the complex plane has several useful properties both in art and in microwave engineering.

Of the numerous extensions and applications of the chart that have been proposed over the years [15], those that aid thinking (rather than merely computation) continue to be useful for professionals who have developed an intuitive feel for it. So entrenched is the Smith chart in microwave engineers' conceptualization that despite their powerful and highly sophisticated computational capability, both the modern computer-aided design software and the computer-controlled microwave measurement equipment continue to present results on Smith chart overlays.

The Smith chart is constructed to perform essentially two basic tasks:

1) the transformation between a reflection coefficient Γ defined with respect to a reference impedance Z_0 and the corresponding normalized impedance Z/Z_0:

$$\Gamma = \frac{Z - Z_0}{Z + Z_0} \quad \text{and} \quad Z = \frac{Z_0 + Z_0\Gamma}{1 - \Gamma} \tag{1}$$

2) the transformation of either Γ or Z upon shifting the reference plane at which they are defined by a distance l_{12} along a uniform transmission line with characteristic impedance Z_0 and propagation constant $\gamma = j\beta$:

$$Z_2 = Z_0 \frac{Z_1 + Z_0 \tanh\gamma l_{12}}{Z_1 \tanh\gamma l_{12} + Z_0} \quad \text{and} \quad \Gamma_2 = \Gamma_1 e^{-2\gamma l_{12}}. \tag{2}$$

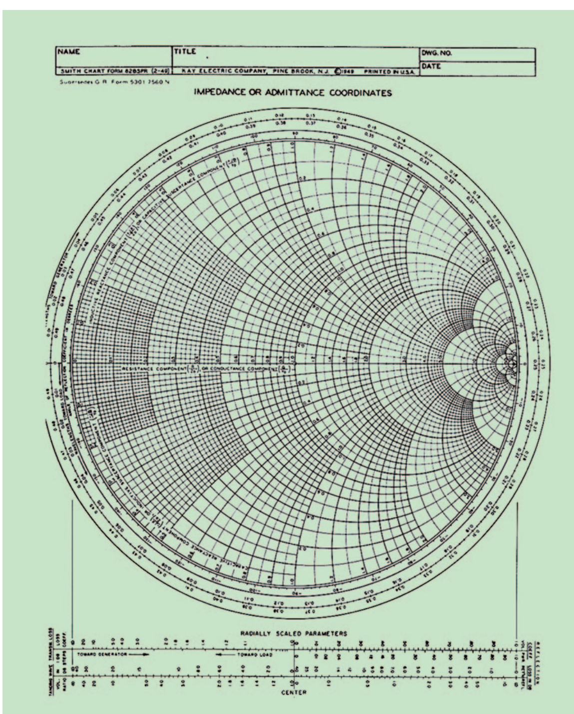

Figure 5. *The Smith chart, along with radial scales. (Courtesy of Analog Instrument Co.)*

The Smith chart accomplishes the first task by 1) drawing the polar coordinate scales on a plane for plotting Γ, restricted within a unit circle (i.e., to Γ values for a passive one-port); 2) drawing the rectangular Cartesian coordinate lines on another, distortable plane for plotting Z/Z_0, limited to the right half plane (also for a passive one-port); and then 3) distorting the Cartesian coordinate system such that when the two coordinate systems are superimposed on top of each other, each point of one coordinate system coincides exactly with its mapping in the other coordinate system. The second task is easily accomplished by a rotation of Γ by the angle 2 $\mathrm{Im}[\gamma]l_{12}$ and requires only a relabeling of the angular scale in the polar coordinate system in the units of l/λ. Since each of these two tasks is simply an example of a Möbius bilinear transformation of a complex number, we might be permitted to claim that the Smith chart is a graphical aid for carrying out Möbius transformation.

Möbius Bilinear Transformations

Möbius transformations are ubiquitous in microwave work for practical reasons. The desired or required reference plane for defining a Z or a Γ is very frequently different from the one at which measurement or computation can be carried out accurately and conveniently; reasons include inaccessibility of the plane for measurement, the unavailability of reference standards usable at the plane, and structural complexity that creates multimodal fields or coupling across the plane. Hence there is a frequent need to transform, or de-embed, impedances or reflection coefficients between two reference planes. The electromagnetic structure intervening between these two reference planes can often be adequately represented by a linear two-port network, characterized in the frequency domain by an impedance matrix [Z] or a scattering matrix [S]. The transformation of a response function (such as a reflection coefficient or an impedance) by an arbitrary linear two-port network is given by

$$\Gamma_{\text{in}} = \frac{(S_{11}S_{22} - S_{12}S_{21})\Gamma_L + S_{11}}{S_{22}\Gamma_L + 1}$$

$$\text{or} \quad Z_{\text{in}} = \frac{Z_{11}Z_L - (Z_{11}Z_{22} - Z_{12}Z_{21})}{Z_L + Z_{22}}, \tag{3}$$

and each of these is an example of a Möbius transformation of a complex number.

The Möbius bilinear (fractional linear) transformation of a complex number Z is defined by

$$M(Z) \equiv W \equiv \frac{AZ+B}{CZ+D}, \quad (4)$$

where A, B, C, and D are complex constants (meaning independent of Z). Properties of this transformation can be stated more compactly by supplementing this basic definition with some additional requirements, such as the following:

1) $M(\infty) \equiv A/C$, and $M(-D/C) \equiv \infty$, which allows us to extend the complex plane by including the point $Z = \infty$ in specifying the domain and the range of the transformation, thus making the Möbius transformation homeomorphic
2) $AD - BC \neq 0$, which allows us to normalize the transformation by dividing each of the constants A, B, C, and D by the quantity $AD - BC$ without altering the transformation in any way.

The Möbius transformation in the complex plane has several useful properties [16], useful both in art and in microwave engineering [17], [18]. These properties can be described in several alternative languages, e.g., algebraically, in matrix form, geometrically, and topologically. The geometrical approach is most useful here in view of the need to relate them to Escher's graphical work.

Geometrical Properties of Möbius Transformation

A Möbius transformation maps a complex number Z into another complex number W, each of which requires a two-dimensional plane for its geometrical representation, in which it is represented by a point. A curve C_Z in the Z-plane is thus a set of points, each of which is mapped by the Möbius transformation into a point in the W-plane, and the set of transformed points in the W-plane together defines another curve C_W; this process can be called a transformation of the curve. Such a transformation preserves certain of the properties of the curve, and it is these invariants shared by C_Z and C_W that are of interest to us here.

To visualize the effect of this transformation in the complex plane, it is helpful to consider some special cases of the Möbius transformation in (4).

- translation: $W = Z + B$ (where $A = D = 1$; $C = 0$)
- scaling: $W = |A|Z$ (where $B = C = 0$; $D = 1$; A real, positive)
- rotation: $W = Z\exp(j\angle A)$ (where $B = C = 0$; $D = 1$; $|A| = 1$)
- inversion: $W = 1/Z$ (where $A = 0$; $B = C = 1$; $D = 0$).

The geometrical effect of the first three on a curve can be visualized in a straightforward manner; the last operation can be viewed as a reflection in the unit circle. The utility of these special cases lies in the fact that they serve as building blocks for any Möbius transformation, which can be expressed as a concatenation of these elementary operations; thus W in (4) can be arrived at by the following sequence of transformations:

The Smith chart is a graphical aid for carrying out Möbius transformation.

$$\begin{aligned} &Z \to CZ \to CZ + D \to 1/(CZ+D) \\ &\quad\to [(BC-AD)/C]/(CZ+D) \\ &\quad\to [(BC-AD)/C]/(CZ+D) + (A/C) \\ &\quad= (AZ+B)/(CZ+D) = W. \end{aligned}$$

The Möbius transformation of a curve in the complex plane has several useful properties of which the following three are particularly relevant here.

- The transformation maps circles (and hence straight lines, which are a special case of the circles) in the Z-plane into circles in the W-plane. [Hence the appearance of the lines in a Smith chart, representing the transformation of constant Re[Z] and Im[Z] curves by the Möbius transformation in (1)].
- The transformation is conformal, meaning that the angle between two curves C_{Z1} and C_{Z2} (defined as the angle between tangents to those curves at their point of intersection) in the Z-plane remains unaltered, both in magnitude and in sign, upon transformation to the W-plane.
- With appropriate choice of distance metric, the length of a geodesic curve between two points Z_1 and Z_2 in the Z-plane is the same as that between their images W_1 and W_2; Euclidean metric is not such a metric.

Proofs, examples, and applications of these properties can be found in cited works [16] and [19] and elsewhere.

It is clear that the first two properties can assist in keeping the shape of some figure recognizable following a transformation, while the third can provide the scaling in the Euclidean plane where the distance is not invariant. Those are the two essential requirements mentioned in "Tessellations Constructed by Transformations" for the construction of tessellations representing infinity. We focus on the third property and the distance metric required for invariance to understand the distance and size scaling in Escher's work.

Hyperbolic Distance Metric

Definition of a Distance Metric

The distance $d(Z_1, Z_2)$ between two points Z_1 and Z_2 is defined as the length of the shortest path or curve joining the two points. The curve having the shortest length between two points is called a geodesic, which is a generalization of the concept of a straight line from Euclidean geometry. The length of a curve, in turn, is defined as the integral of (i.e., the summation over) the elementary lengths between successive points along the curve, each separated from the previous one by an

The hyperbolic distance between two points is invariant under a Möbius transformation.

infinitesimal distance. Finally, the infinitesimal distance between two points can be defined in the usual Euclidean manner, because in the infinitesimal limit, all spaces become essentially Euclidean [20]. The resulting distance metric has the properties intuitively expected of a distance, namely

- nonnegativity: $d(Z_1, Z_2) \geq 0$ for all Z_1, Z_2
- identity: $d(Z_1, Z_2) = 0$ if and only if $Z_1 = Z_2$
- bilateral symmetry: $d(Z_1, Z_2) = d(Z_2, Z_1)$ for all Z_1, Z_2
- triangle inequality: $d(Z_1, Z_3) \leq d(Z_1, Z_2), + d(Z_2, Z_3)$ for all Z_1, Z_2, Z_3
- continuity of $d(Z_1, Z_2)$, ensured by its one-to-one correspondence with real numbers.

It is easy to see that the Euclidean distance metric remains invariant under the operations of displacement and rotation; it also remains invariant under the operation of complex conjugation ($W = Z^*$), which can be interpreted as reflection in the real axis. By contrast, as shown above, a Möbius transformation is composed of scaling and inversion in addition to translation and rotation, and as a result, in general, it does not leave the Euclidean distance between two points invariant. Instead, one of the invariants of the Möbius transformation is a cross-ratio defined as:

$$\frac{(W_1 - W_3)(W_2 - W_4)}{(W_1 - W_4)(W_2 - W_3)} = \frac{(Z_1 - Z_3)(Z_2 - Z_4)}{(Z_1 - Z_4)(Z_2 - Z_3)}, \quad (5)$$

where W_1, W_2, W_3, and W_4 are the images of Z_1, Z_2, Z_3, and Z_4, respectively, under the transformation. Therefore, a distance metric based on the cross-ratio, or any monotonic function thereof, would be invariant to the Möbius transformation. Moreover, because the distance metric is dependent upon the geodesic, which in turn is governed by the rules of geometry laid down in the space, it is necessary to introduce a different space wherein the geometrical rules are different from the Euclidean. Hence the need for a non-Euclidean space.

Non-Euclidean Geometry

Elementary school level geometry is called Euclidean geometry, in honor of Euclid of Alexandria (circa 300 B.C.), who authored *Elements of Geometry*, a treatise on geometry, that is apparently the most widely read book in all of science, mathematics, and technology in the history of mankind. Little can be said definitively about Euclid, as the earliest surviving copy of this work, written in Latin by Plonus, dates from around 900 A.D., 12 centuries after Euclid (indeed, of the 15 books or chapters contained in the *Elements*, the last two seem to be later additions and not by Euclid). Euclid consolidated the geometrical knowledge of his times by giving it an axiomatic structure, wherein all results are deduced from the smallest possible set of a priori postulates. Such a minimal set consists of five Euclidean axioms, which basically postulate the existence of a straight line, continuity, a metric of distance, the congruence of angles, and parallel lines. The fifth postulate, the postulates of parallels, can be stated in a number of alternate forms that are equivalent (i.e., deducible from each other), in one of which it specifies that there exists a unique line that passes through a given point P and is parallel to a given line L.

Following Euclid, over a period of over 2,000 years, there is a long history of efforts by mathematicians, from Archimedes to Legendre, to further reduce the set of Euclidean postulates, with Euclid's fifth postulate receiving the most scrutiny in an attempt to deduce it from the other four. Only in the 19th century did it become clear that 1) the fifth postulate is not a consequence of the other four, 2) it is also not essential to the internal consistency of geometry, and 3) its replacement by alternative postulates can result in a perfectly self-consistent geometry. Such geometries are called non-Euclidean, and there are two distinct varieties of them, depending on the replacement selected for the fifth postulate. If the number of lines passing through P and parallel to L is zero, the geometry is elliptical (or Riemannian), in which the sum of the internal angles of a triangle is greater than 180°. If the number is two or more, the geometry is hyperbolic, wherein the sum of the angles of a triangle is less than 180°.

Perhaps the easiest way to illustrate and comprehend the non-Euclidean geometry and in that process, demonstrate the existence and consistency of its axioms, is by constructing a model of a non-Euclidean geometry that borrows from the already familiar Euclidean geometry. In the following, we present only one from among the many possible models and for only the hyperbolic variety of non-Euclidean geometry, which will be required for later discussion.

Poincaré's Open Disc Model of Hyperbolic Space

To construct the new geometry, we need objects that can serve as embodiments of the concepts of points, lines, and planes, and to which a set of self-consistent postulates can be applied; a minimal set consists, for example, of the Euclidean axioms, which basically postulate the existence of a straight line, continuity, a metric of distance, the congruence of angles, and parallel lines. The last postulate, the postulate of parallels, is however, not essential for self-consistency and can therefore be violated to obtain a new geometry. These objects can be selected from the Euclidean world itself, and many alternative choices are available. One of the choices is that of points in the interior of a unit circle, leading to the so-called Poincaré open disc model of hyperbolic geometry.

In a Euclidean plane, wherein each point can be specified by a complex number Γ, consider a circle of unit radius with its center at the origin. We designate the interior of this circle as hyperbolic plane, and the set of all points with $|\Gamma| < 1$ in the open unit disc as hyperbolic points (h-points). Next, consider a Euclidean circle (having a center C and a radius R) that is normal to the unit circle, i.e., it intersects the unit circle at a right angle at P and Q, their points of intersection (by symmetry—if the circles are mutually orthogonal at P, they are also orthogonal at Q). We designate the set of all h-points on the interior segment of any such orthogonal circle as a hyperbolic line (h-line), as shown in Figure 6; each choice of the location of C and magnitude of R gives rise to a different h-line. Given two h-points Γ_1 and Γ_2, there is a unique Euclidean circle that passes through them and is at the same time normal to the unit circle; hence there is a unique h-line through two given h-points. The Euclidean postulate on angle congruence is ensured by retaining the Euclidean measure of an angle between two lines. These definitions meet the first four Euclidean postulates, including the existence, continuity, and unlimited extendibility, of a line (toward P and Q, which are not part of the line, because they are not h-points).

Finally, we consider the fifth postulate concerning parallel lines. Two h-lines are parallel to each other if they do not intersect, i.e., have no h-point in common. Since more than one h-line can be drawn through a given h-point Γ_3, each of them nonintersecting with a given h-line passing through Γ_1 and Γ_2, it is clear that in this model, the number of parallel lines passing through a given point and parallel to a given line is greater than one, which is the distinguishing characteristic of hyperbolic geometries.

Testing line segments for congruence requires first having a distance metric. In the hyperbolic space, requiring that the distance metric satisfy the conditions listed in the previous section leads to a natural metric, to be defined in the following.

The Hyperbolic Distance Metric

The hyperbolic distance $d_H(\Gamma_1, \Gamma_2)$ between two h-points Γ_1 and Γ_2 (i.e., the length of the h-line joining them) is different from the Euclidean distance d_E between two Euclidean points Γ_1 and Γ_2. Indeed, d_H can be defined in terms of the Euclidean distances of each of the points Γ_1 and Γ_2 from the intersection points P and Q of the orthogonal Euclidean circle mentioned earlier (caution: P is the limiting point approached when the h-line $\Gamma_1\Gamma_2$ is extended toward Γ_1, while Q is the limiting point when it is extended toward Γ_2); then the hyperbolic distance metric, obtained by the procedure outlined previously, is as follows:

The number of parallel lines passing through a given point and parallel to a given line is greater than one, which is the distinguishing characteristic of hyperbolic geometries.

$$d_H(\Gamma_1, \Gamma_2) = \log_e \left[\left(\frac{d_E(\Gamma_1, Q)}{d_E(\Gamma_2, Q)} \right) \left(\frac{d_E(\Gamma_2, P)}{d_E(\Gamma_1, P)} \right) \right], \quad (6)$$

where d_E denotes the familiar Euclidean distance between two complex numbers, defined as $d_E(\Gamma_1, Q) = |\Gamma_1 - Q|$. These Euclidean distances can be explicitly evaluated for the open disc model, and after their substitution, the hyperbolic distance between two points Γ_1 and Γ_2 can be expressed solely in terms of the two points as

$$d_H(\Gamma_1, \Gamma_2) = 2 \tanh^{-1} \left| \frac{\Gamma_1 - \Gamma_2}{1 - \Gamma_1 \Gamma_2^*} \right|. \quad (7)$$

This definition of the hyperbolic distance between two points satisfies all of the properties expected of a distance metric listed previously. Moreover, it is invariant under a Möbius transformation, because the distance in (6) is a monotonic function of a cross-ratio, as defined in (5).

In the Smith chart, the same cross-ratio invariance and distance metric are manifested in many ways, such as in the radial VSWR scale employed, and are used in certain geometrical constructions and for defining the measure of variance between two different impedances. In the special case where one of the two points is at the origin, the radial distance of the other point from the origin is given by

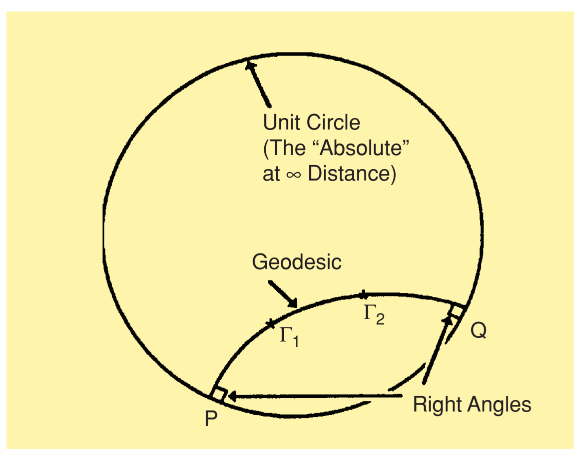

Figure 6. *A hyperbolic line in the unit circle.*

TABLE 2. Measurements on features of figures in *Circle Limit IV*.

Figure	Coordinates of Wing Tips: Left Wing	Coordinates of Wing Tips: Right Wings	Wing Span in decibels
Angel #1	$0.515 \angle 0°$	$0.515 \angle 60°$	10.0 dB
Angel # 2	$0.515 \angle 0°$	$0.81 \angle 13°$	10.5 dB
Angel # 3	$0.815 \angle 0°$	$0.87 \angle 11°$	9.5 dB
Angel # 4	$0.887 \angle 0°$	$0.887 \angle 9°$	9.5 dB
Angel # 5	$0.815 \angle 0°$	$0.91 \angle 8°$	10.5 dB

Dimensions in hyperbolic distance metric deduced from polar coordinates in the unit circle and rounded to the nearest half decibel.

$$d_H(0, \Gamma) = 2 \tanh^{-1}\Gamma = \log_e\left(\frac{1+|\Gamma|}{1-|\Gamma|}\right). \quad (8)$$

This is identical, to within a scale factor, with the definition of the VSWR (in decibels) on a uniform lossless transmission line terminated in a load with reflection coefficient Γ. As a result, the VSWR scale in decibels, provided with the Smith chart, is essentially a scale for the measurement of radial distance from the origin in the hyperbolic metric. Moreover, (7) can be viewed as a generalization of the VSWR definition and quantifies the distance between any two reflection coefficients. Such a measure of distance has been employed in microwave engineering in a number of situations, such as for measuring the distance between two impedance values of a switching diode [21] or defining the unilateral power gain of an active device in a lossless embedding [18].

Circle Limit IV Analyzed with Hyperbolic Distance Metric

Returning to *Circle Limit IV*, we can now measure all the figures appearing in the drawing, including those that lack bilateral symmetry. When the individual figures are rotated, the parts farther from the origin suffer greater shrinkage due to the crowding of the hyperbolic scale toward the periphery, and this causes the apparent asymmetry in the figures.

For the purpose of illustration, we will select just one feature, the wing span of the angels, for comparison. Table 2 shows the polar coordinates of the wing tips for a number of angel figures, measured with reference to each other. The wingspan is then calculated from the hyperbolic metric in (7) and is also listed in the table. The results show that the dimensions are identical to within the accuracy attainable in a woodcut, and allow us to conclude that the figures are indeed congruent. *Circle Limit IV* is therefore just a regular tessellation when measured with the hyperbolic metric, in which individual tiles are rotated.

Acknowledgment

I wish to record my indebtedness to late Prof. Robert L. Kyhl (1918–2003) of Massachusetts Institute of Technology (MIT), Cambridge, who introduced me to this subject in a short lecture, delivered during the "Independent Activities Period" at MIT. Thanks are also due to The M.C. Escher Company (Baarn, The Netherlands), the copyright holder, for permission to reproduce the works of Escher in this article.

References

[1] E. Strosberg, *Art and Science.* Paris: Unesco, 1999 (also New York: Abbeville Press, 2001).

[2] S. Ede, *Art and Science*. London, U.K.: L.B. Tauris, 2005.

[3] S. Wilson, *Information Arts: Intersection of Arts, Science, and Technology.* Cambridge, MA: MIT Press, 2002.

[4] M.C. Escher, *M.C. Escher, His Life and Complete Graphic Works (Transl.:* from Dutch, F.H. Bool). New York: Abradale Press/Harry N. Abrams, 1992.

[5] D. Schattschneider, *Visions of Symmetry: Notebooks, Periodic Drawings, and Related Works of M.C. Escher.* San Francisco, CA: Freeman, 1990.

[6] C.H. MacGillavry, *Fantasy and Symmetry: The Periodic Drawings of M.C. Escher.* New York: Harry N. Abrams, 1976.

[7] E. Maor, *To Infinity and Beyond.* Cambridge, MA: Birkhauser, 1987.

[8] M.C. Escher, *Escher on Escher: Exploring the Infinite* (Transl. from Dutch, K. Ford). New York: Abrams, 1989.

[9] E.R. Ranucci, "Master of tessellations: M.C. Escher, 1898–1972," *Math. Teacher,* vol. 67, pp. 299–306, Apr. 1974.

[10] J.L. Teeters, "How to draw tessellations of the Escher type," *Math. Teacher*, vol. 67, pp. 307–310, Apr. 1974.

[11] H.S.M. Coxeter, "The non-Euclidean symmetry of Escher's picture *Circle Limit III*," *Leonardo*, vol. 12, no. 1, pp. 19–25, winter 1979.

[12] H.S.M. Coxeter, "The trigonometry of Escher's woodcut *Circle Limit III*," *Math. Intell.,* vol. 18, no. 4, pp. 42–46, 1996.

[13] P.H. Smith, "Transmission line calculator," *Electronics,* vol. 12, no. 1, p. 29, Jan. 1939.

[14] P.H. Smith, "An improved transmission line calculator," *Electronics,* vol. 17, no. 1, pp. 130–133, Jan. 1944.

[15] P.H. Smith, *Electronic Applications of the Smith Chart.* New York: McGraw-Hill, 1969.

[16] F.J. Flanigan, *Complex Variables. Harmonic and Analytical Functions.* Boston: Allyn and Bacon, 1972, pp. 304–324.

[17] G.F. Engen, *Microwave Circuit Theory and Foundations of Microwave Metrology.* Stevenage, U.K.: Peregrinus, 1992.

[18] M.S. Gupta, "Power gain in feedback amplifiers, a classic revisited," *IEEE Trans. Microwave Theory Tech.,* vol. 40, no. 5, pp. 864–879, May 1992.

[19] A. Ramsay and R.D. Richtmyer, *Introduction to Hyperbolic Geometry.* New York: Springer-Verlag, 1995.

[20] H.S.M. Coxeter, *Non-Euclidean Geometry.* Toronto, Canada: Univ. Toronto Press, 1942.

[21] K. Kurokawa and W.O. Schlosser, "Quality factor of switching diodes for digital modulation," *Proc. IEEE,* vol. 58, no. 1, pp. 180–181, Jan. 1970.