**FEATURE COLUMN** Monthly essays on mathematical topics

# How Did Escher Do It?

*How did M. C. Escher draw his **Circle Limit** figures...*

Bill Casselman
University of British Columbia, Vancouver, Canada
cass at math.ubc.ca

Mail to a friend        Print this article

## Introduction

*How did M. C. Escher draw his **Circle Limit** figures?* (You can find a good exhibition of these at the web site Math and the Art of Escher.) The interesting part of this question is that the mathematics involved is not at all trivial. and the execution of the figures must have been extremely difficult. There is much literature on the topic, but most of it seems to be mostly concerned with linking the matter to other topics of higher mathematics. I am more interested in figuring out the details of how Escher might have proceeded. There is apparently only one place - a recent article in the January 2001 issue of the *American Mathematical Monthly* by Chaim Goodman-Strauss - that comes close to answering this question at the low level I am interested in, but it seems not to provide a completely satisfactory analysis. Escher worked in the days long before computers made this sort of work relatively simple, so what's involved here is an elementary mathematical question:

> *How can one construct tesselations of the non-Euclidean plane, conceived of as a certain geometry on the unit disc in the model attributed to Henri Poincaré, with ruler and compass?*

A tesselation is a covering by repeated patterns all in some sense the same. Familiar examples of a different, simpler type are tilings of a floor, invariant under certain translations, rather than of the unit disk. Much of Escher's early work involved these Euclidean tilings. I begin with an account of how he came to those related to non-Euclidean geometry. For this, I follow the two articles [Schattschneider: 2006] and [Schattschneider: 2010].

In writing this, I have been helped much by correspondence with Doris Schattschneider.

### Coxeter and Escher

In 1954 the International Congress of Mathematicians (an event held in different locations every 4 years, with hundreds or even thousands of mathematicians attending) was located in Amsterdam, and associated with the Congress was a display of Escher's work. Much of this involved artistic rendering of familiar symmetries of the Euclidean plane. It was with this exhibition that the well known geometer H. S. M. Coxeter first became acquainted with Escher's art. Three years later Coxeter gave an address on symmetry to the Royal Society of Canada, and he asked Escher if he could include some of his pictures as somewhat unconventional illustrations of symmetry. The published account of this in the _Transactions_ of the Society was apparently the first mathematical publication to contain an Escher tesselation.

Coxeter sent Escher a reprint of the lecture. It happened that Figure 7 of the article (a redrawn version is just below) resonated strongly with one of Escher's own interests, the problem of producing an infinitely repeating pattern in a finite figure. To a mathematician, Coxeter's figure represents a non-Euclidean analogue of a periodic tiling of the Euclidean plane, but I am going to largely ignore this, and just ask, how did Escher see it? I cannot answer this, but only offer a few suggestions.
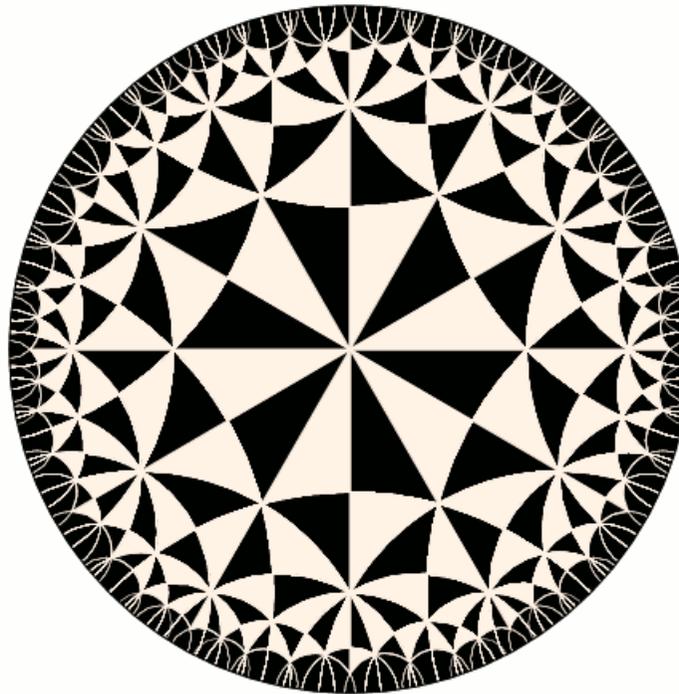
Figure 7 from Coxeter's address to the Royal Society of Canada

We do have some record of Escher's reaction. First of all, in the Gemeente Museum in the Hague is Escher's copy of Coxeter's reprint (this is MAP N82 in the Museum). What is interesting about it is that Escher has annotated it with some of his own drawing, presumably because he was trying to figure out the geometric rules by which it was constructed. It does give some insight into what he was thinking.

At any rate, Escher wrote back to Coxeter, saying that he had been "shocked" by Figure 7. He had tried to reproduce the figure by hand, and included a hand-drawn sketch illustrating his attempt. He had only partly succeeded, and the sketch showed how far he had gotten. He also asked Coxeter for advice on how to proceed. For reasons that we shall see later, Coxeter added some marks to Escher's sketch and returned it. (This sketch is now to be found as MAP N80 in the Escher Archives in the Haags Gemeentemuseum. Doris Schattschneider has reproduced her redrawn versions of both these diagrams in a June/July 2010 *NOTICES* article, one based on a slide she made many years ago and the other on a photocopy of Escher's sketch.) I am going to attempt to follow him in doing this, through the following sequence of pictures.

I should point out at the beginning that the problem is completely specified by the following conditions:

- (2) the arcs bounding all of the curved `triangles' intersect the unit circle orthogonally;
- (3) at each triangular vertex there are either *n=4*, *8*, or *12* uniformly distributed triangles;
- (4) at the origin *n = 12*.

The problem at hand is to draw the bounding arcs for arbitrarily small triangles in the tesselation. This means finding their centers and radii. In practice, once the centers are known finding the radii is immediate. I shall first simply describe how one finds the centers, for the most part without mathematical justification. This is probably what Escher did, since he seems to have thought of mathematics as an empirical science. It is not a view of the subject to be despised, especially in light of Escher's own substantial if informal mathematical talent.
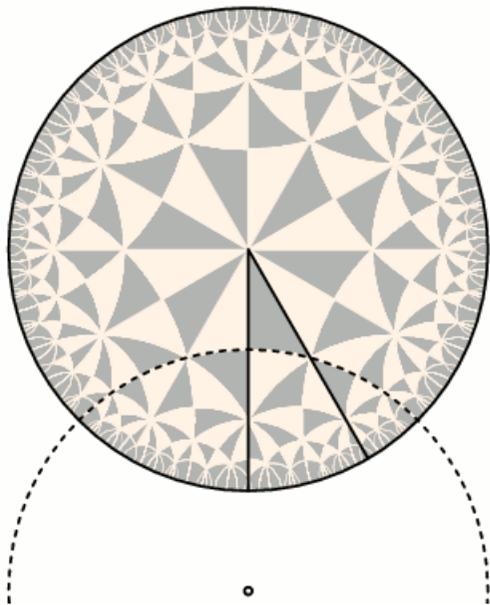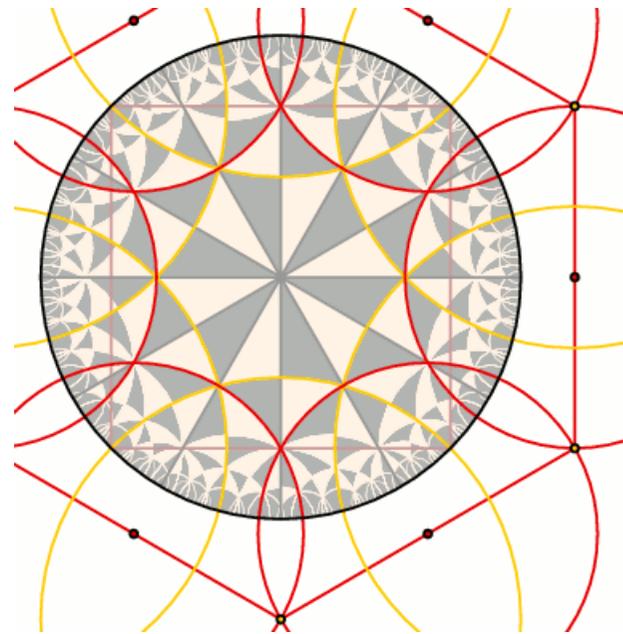


First determine the central triangles. The only hard part is to determine the circular arc bounding one side. It *seems* to be of radius *1*.
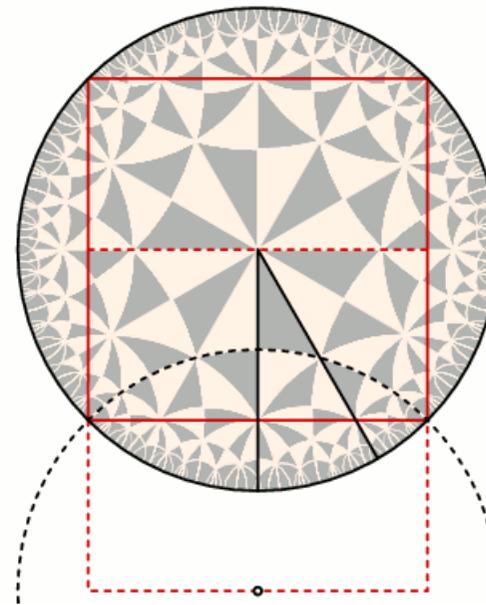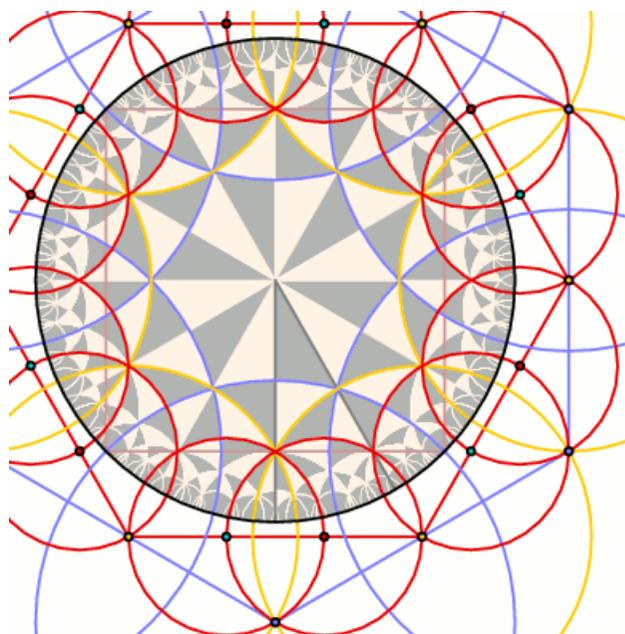
This is confirmed, and the distance of its center from the origin also found, by constructing an inscribed square and taking symmetry into account.
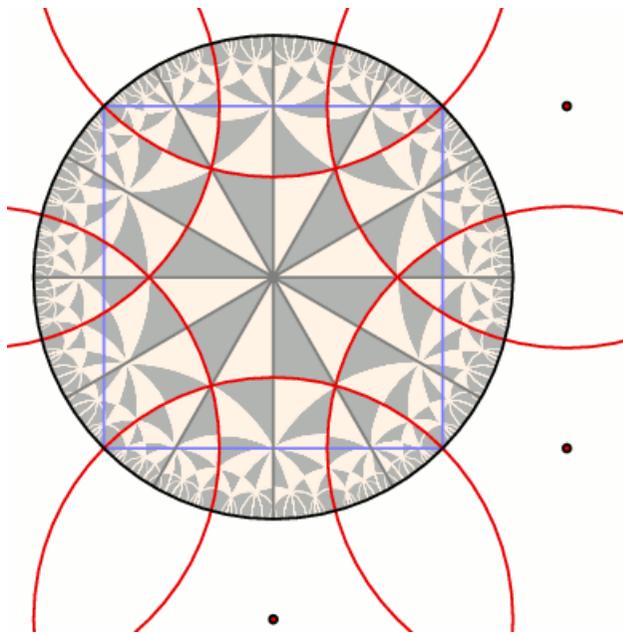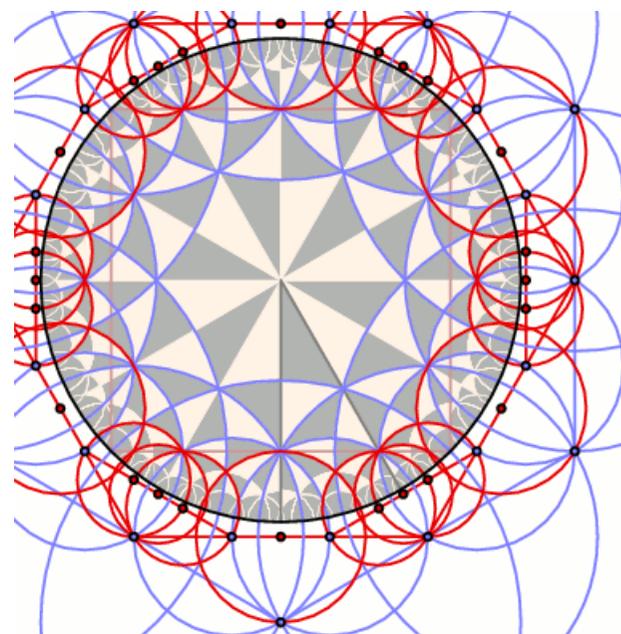
Then apply rotational symmetry. This gives us all the triangles in the first layer.



The centers of the next group of arcs lie half-way between those of the previous group.

The next group of centers split segments between the previous group in thirds.



Similar observations give another three groups of arcs by interpolating halves and thirds.

This is exactly as far as Escher got, and this is essentially what he wrote to Coxeter. *But he didn't see what to do next.* In fact, the simple process of interpolating halves and thirds comes to an end at this point. In his letter, Escher asked Coxeter what he should do in order to continue indefinitely, and I'll come back to this question in a moment. It is not a trivial problem. But first I want to summarize what one might deduce from what has occurred so far. In the figures above, groups of centers occur often on line segments exterior to the unit circle. What characterizes these groups? *A group of arc centers lie on a single line if the arcs all pass through a common point.* For example:
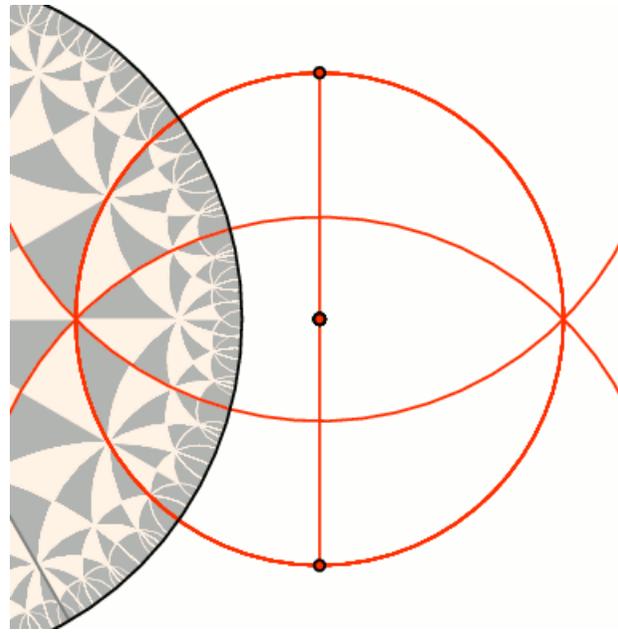
For us this is an empirical observation, but in fact it is a general principle:

> **Theorem.** *Given a point in the interior of the unit disk, the centers of all the circles that pass through it and are also orthogonal to the unit circle lie on a straight line outside the disk.*

This principle is all one needs to justify every bit of the construction so far. It is simple to state, but it is not so simple to explain why it is valid - it is already not immediately apparent why the figure above is not deceptive. I'll say something later about why this principle is true, but for the moment I'll just take it for granted. It is in fact one of two fundamental, if somewhat subtle, mathematical principles involved in constructing figures like Coxeter's.

Whether or not Escher had formulated this principle explicitly is not clear. As we shall see, Coxeter didn't think he had. I have said that Escher sent Coxeter a diagram he had drawn with compass and straight-edge more or less equivalent to what we have seen so far. He asked Coxeter what he should do to continue. Coxeter's response to this (which I quote from Doris Schattschneider's article) was to make some marks on the sketch Escher had sent him, and to send it back with the comment

> *I am ... interested that you succeeded in reconstructing so much of the surrounding "skeleton"*

*which serves to locate the centers of the circles. This can be continued in the same manner. For instance, the point that I have marked on your drawing (with a red o on the back of the page) lies on three of your circles ... These centers therefore lie on a straight line (which I have drawn faintly in red) and the fourth circle through the red point must have its center on this same red line.*

Coxeter had added the point and the line in the following figure; I have also shown in red the circle he was referring to.
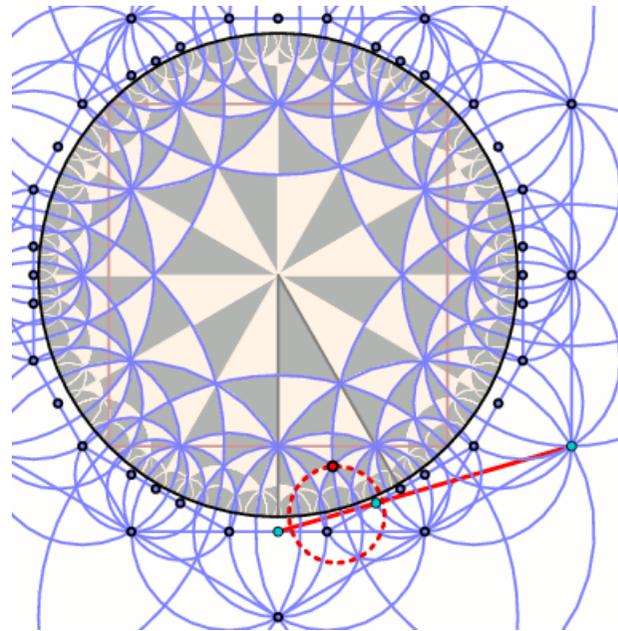


All that Coxeter had done was to make explicit to Escher the principle formulated above. *It was not sufficient to determine the center of that fourth circle on the line he had drawn.* It is not easy to understand what was going on in Coxeter's mind. He must have known he had given only an incomplete answer. In [Coxeter: 1997] he later wrote about the process of constructing tesselations that whereas usually each new center arises as the common point of lines associated to interior points by the principle enunciated above, occasionally one had to do something a bit trickier. Even later on, when writing for mathematicians, he did not specify exactly how to carry out the trickier step. What it is easy to do is to imagine the frustration Escher felt with this reply. In a letter to his son he wrote "I am so often at cross-purposes with those theoretical mathematicians ... it seems very difficult for Coxeter to write intelligibly to a layman." ([Schattschneider: 2006])

So we ask, what is missing from Coxeter's response? What does it take to locate the center of the red circle in the last figure? There are several possible answers to this. One is given by Chaim Goodman-Strauss' article, where in effect he suggests intersecting a certain rotated line with Coxeter's red line. This is equivalent to observing that the missing center is on a tangent to the large circle passing through Coxeter's red point. Maybe this option explains why Coxeter thought he had explained enough. But drawing tangent lines with ruler and compass is delicate, and a method of doing this implicit in Goodman-Strauss's article is more complicated than one would wish to carry out over and over again many times.

Of course the one evident fact is that someone knew how to make these figures. The history begins apparently sometime in the 19th century. Many of Felix Klein's works contain such figures, most famously the book with Robert Fricke. Here is one of those early ones with with Klein's own annotation (discovered by accident not so long ago in a drawer in the Mathematics Department of Göttingen University):
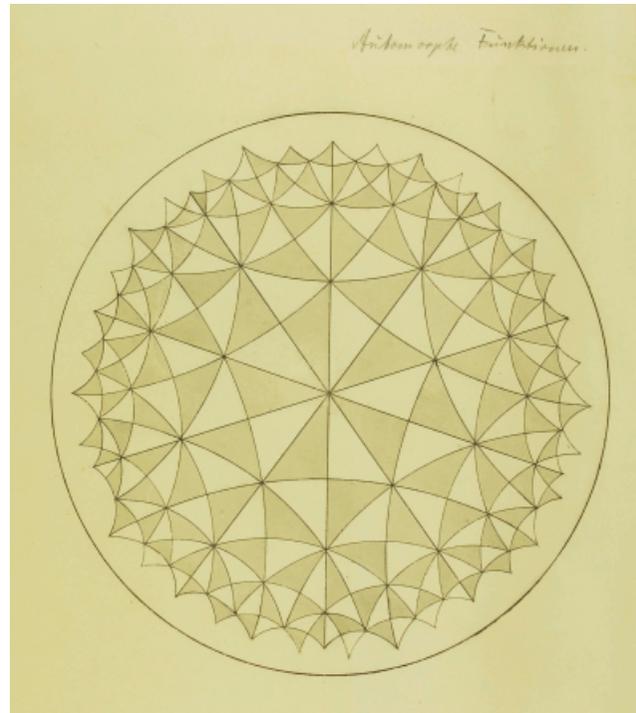


The natural question to ask is, how were all of these produced? For that matter, how was Coxeter's Figure 7 drawn? I do not know of any record of the methods used, nor even do I know who produced Coxeter's figure. It is very unlikely to have been Coxeter, who usually got others to make figures for him. A hint about at least a few can be found in a slightly different version of Figure 7 published in [Coxeter: 1979]. In this version, one can see not only the exterior lines we already know about, but also some other exterior lines proceeding from some of the exterior arc centers to the endpoints of some of the arcs. (These are called secants to the unit circle, as opposed to tangents.) Also in [Coxeter: 1997] Coxeter says of such lines, "... those lines which are secants should not have been drawn."

that the arc centers we are looking for will be intersections of the lines we already know about with certain secants. Here is the second basic principle for constructing non-Euclidean tesselations:

**Theorem.** *Given two intersecting circles, the centers of all circles intersecting both of these orthogonally lie on the straight line passing through their intersection.*
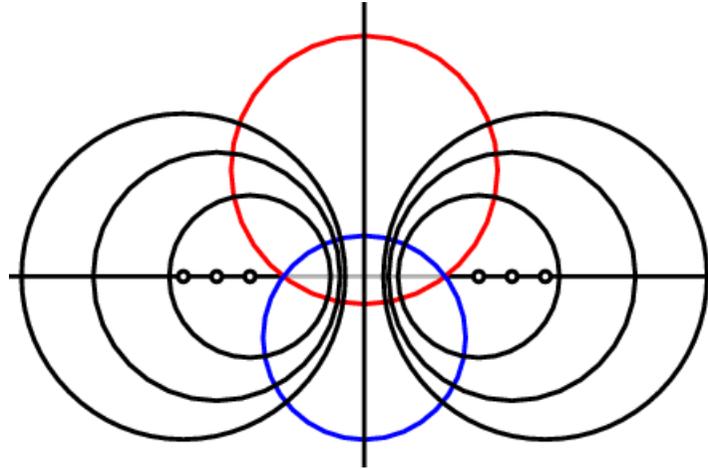
At any rate, this tells us how to find the missing center:
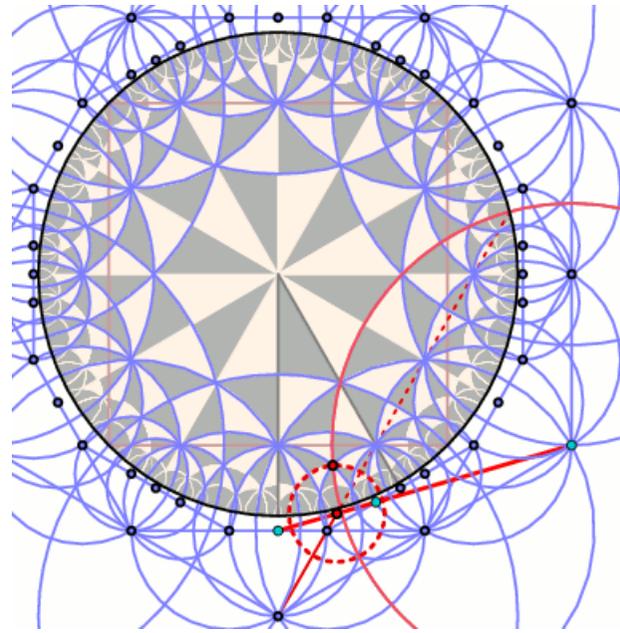
### Turning things inside out

I now want to give some idea of why the two principles stated above are valid. The mathematical tool that explains the two principles I have stated is *inversion*. An inversion associated to any circle in the plane is a map from the plane to itself that turns the circle inside out, swapping points on the inside with points on the outside. If the circle has center $O$ and radius $R$, it takes a point at distance $r$ from $O$ to a point on the same ray out from $O$ at distance $R^2/r$. There is a problem if $r=0$ (when the point is $O$ itself), in which case inversion moves it out to infinity (and moves infinity in to $O$). Points on the circle are mapped to themselves.
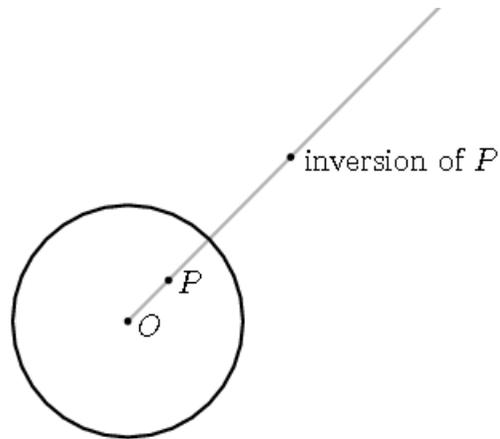
There are a small number of important facts about inversions that we'll need to know:

*Any circle that does not go through **O** is mapped to some other circle.*

*A circle that goes through **O** is mapped to a line not through **O** whose direction is that of the tangent to the circle at **O**.*

*A line through **O** is mapped to itself.*

*An inversion preserves angles. That is to say, if two curves intersect at an angle **A**, their images under inversion intersect at the same angle.*

These are proved in just about any college textbook on elementary geometry, for example Coxeter's **Introduction to Geometry**. Let's see now how they help us. First comes a simple corollary of these facts.

*If **D** is a circle orthogonal to **C** then inversion in **C** maps **D** to itself.*

The points of intersection are mapped to themselves. Because inversion preserves angles, it maps **D** into another circle orthogonal to **C**. But there is only one circle with given intersection with **C** orthogonal to it.

**Theorem.** *Given a point in the interior of the unit disk, the centers of all the circles that pass through it and are also orthogonal to the unit circle lie on a straight line.*

Suppose $P$ to be a point in the interior of the unit disk, $Q$ its image with respect to inversion in the unit circle. Any circle through $P$ orthogonal to the unit circle is taken to itself by inversion, and must therefore pass through $Q$. The center of this circle must lie on the straight line perpendicular to $PQ$ and halfway in between.
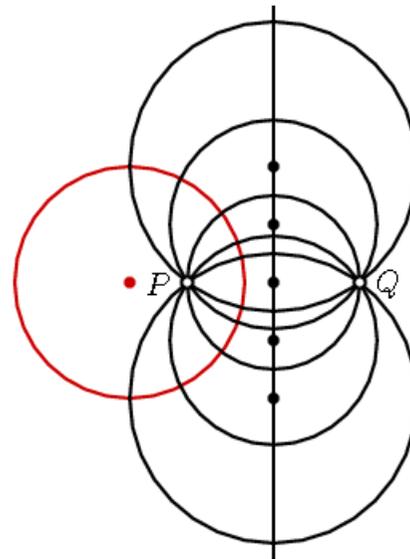


**Theorem.** *Given two intersecting circles, the centers of all circles intersecting both of these orthogonally lie on the straight line passing through their intersection.*

I use the following elementary observation: *a straight line intersects a circle orthogonally if and only if it is a diameter, which is to say passes through its center.*

Suppose given circles $C$ and $D$ intersecting each other at $P$ and $Q$. Suppose $E$ is a third circle orthogonal to both $C$ and $D$. According to the observation, it must be shown that any circle intersecting both circles orthogonally intersects the straight line through $P$ and $Q$ at an angle of $90^o$.

Let $\Gamma$ be the circle with center $O$ passing through $P$. Inversion in $\Gamma$ maps $O$ to infinity. Inversion in $\Gamma$ takes $C$ and $D$ into a pair of lines

Hence the set of all circles intersecting both *C* and *D* orthogonally is mapped to the set of all circles with center *P*. The line through *P* and *Q* is mapped to a line through *P*, All the circles with center *P* cut it orthogonally, which implies that the same is true of the original circles. But this means they all have centers on that line.

## PostScript

I have not in fact answered the original question, how did Escher produce his **Circle limit** series? To tell the truth, probably by trial and error, although they exhibit extraordinary skill. All of them are woodcuts, which do not encourage mathematical precision.

## To find out more ...

- H. S. M. Coxeter, *Crystal symmetry and its generalizations*, volume 51 of the *Transactions of the Royal Society of Canada*.
- H. S. Coxeter, *The non-Euclidean symmetry of Escher's picture "Circle Limit III"*, *Leonardo* 12 (1979).
- H. S. Coxeter, *The trigonometry of hyperbolic tessellations*, *Canadian Mathematics Bulletin* 40 (1997).
- Douglas Dunham, Hyperbolic art and the poster pattern.
- Douglas Dunham, *Creating repeating hyperbolic patterns - old and new*, NOTICES of the American Mathematical Society, April 2003.
- M. C. Escher, MAP N80 in the Escher Archives at the Haags Gemeentemuseum. This is the sketch he sent to Coxeter.
- M. C. Escher, MAP N82 in the Escher Archives at the Haags Gemeentemuseum. This is his annotated copy of Coxeter's reprint.
- Martin von Gagern and Jürgen Richter-Gebert, *Hyperbolization of Euclidean Ornaments*, in a special issue of the *Electronic Journal of Combinatorics*, 2009. Most computer-generated non-Euclidean tilings have trouble around the boundary - convergence is in some sense very slow, and it takes an unrealistically large number of triangles in order to hide the gap. These two German mathematicians explain an elegant technique, one applicable only to digital devices, that solves this problem attractively.
- Chaim Goodman-Strauss, *Compass and straightedge in the Poincaré disk*, *American Mathematical Monthly*, January 2001.
- Doris Schattschneider, *Coxeter and the artists: two-way inspiration*, in **The Coxeter Legacy: Reflections and Projections**, American Mathematical Society, 2006.
- Doris Schattschneider, *The mathematical side of M. C. Escher*, NOTICES of the American Mathematical Society, June/July 2010.

Bill Casselman
University of British Columbia, Vancouver, Canada
cass at math.ubc.ca