

FIGURE 1. Johann Sebastian Bach, in 1748. From a painting by Elias Gottlieb Haussmann.

Introduction: A Musico-Logical Offering

Author:

FREDERICK THE GREAT, King of Prussia, came to power in 1740. Although he is remembered in history books mostly for his military astuteness, he was also devoted to the life of the mind and the spirit. His court in Potsdam was one of the great centers of intellectual activity in Europe in the eighteenth century. The celebrated mathematician Leonhard Euler spent twenty-five years there. Many other mathematicians and scientists came, as well as philosophers—including Voltaire and La Mettrie, who wrote some of their most influential works while there.

But music was Frederick's real love. He was an avid flutist and composer. Some of his compositions are occasionally performed even to this day. Frederick was one of the first patrons of the arts to recognize the virtues of the newly developed "piano-forte" ("soft-loud"). The piano had been developed in the first half of the eighteenth century as a modification of the harpsichord. The problem with the harpsichord was that pieces could only be played at a rather uniform loudness—there was no way to strike one note more loudly than its neighbors. The "soft-loud", as its name implies, provided a remedy to this problem. From Italy, where Bartolomeo Cristofori had made the first one, the soft-loud idea had spread widely. Gottfried Silbermann, the foremost German organ builder of the day, was endeavoring to make a "perfect" piano-forte. Undoubtedly King Frederick was the greatest supporter of his efforts—it is said that the King owned as many as fifteen Silbermann pianos!

Bach

Frederick was an admirer not only of pianos, but also of an organist and composer by the name of J. S. Bach. This Bach's compositions were somewhat notorious. Some called them "turgid and confused", while others claimed they were incomparable masterpieces. But no one disputed Bach's ability to improvise on the organ. In those days, being an organist not only meant being able to play, but also to extemporize, and Bach was known far and wide for his remarkable extemporizations. (For some delightful anecdotes about Bach's extemporization, see *The Bach Reader*, by H. T. David and A. Mendel.)

In 1747, Bach was sixty-two, and his fame, as well as one of his sons, had reached Potsdam; in fact, Carl Philipp Emanuel Bach was the Capellmeister (choirmaster) at the court of King Frederick. For years the King had let it be known, through gentle hints to Philipp Emanuel, how

pleased he would be to have the elder Bach come and pay him a visit; but this wish had never been realized. Frederick was particularly eager for Bach to try out his new Silbermann pianos, which he (Frederick) correctly foresaw as the great new wave in music.

It was Frederick's custom to have evening concerts of chamber music in his court. Often he himself would be the soloist in a concerto for flute. Here we have reproduced a painting of such an evening by the German painter Adolph von Menzel, who, in the 1800's, made a series of paintings illustrating the life of Frederick the Great. At the cembalo is C. P. E. Bach, and the figure furthest to the right is Joachim Quantz, the King's flute master—and the only person allowed to find fault with the King's flute playing. One May evening in 1747, an unexpected guest showed up. Johann Nikolaus Forkel, one of Bach's earliest biographers, tells the story as follows:

One evening, just as he was getting his flute ready, and his musicians were assembled, an officer brought him a list of the strangers who had arrived. With his flute in his hand he ran over the list, but immediately turned to the assembled musicians, and said, with a kind of agitation, "Gentlemen, old Bach is come." The flute was now laid aside, and old Bach, who had alighted at his son's lodgings, was immediately summoned to the Palace. Wilhelm Friedemann, who accompanied his father, told me this story, and I must say that I still think with pleasure on the manner in which he related it. At that time it was the fashion to make rather prolix compliments. The first appearance of J. S. Bach before so great a King, who did not even give him time to change his traveling dress for a black chanter's gown, must necessarily be attended with many apologies. I will not here dwell on these apologies, but merely observe, that in Wilhelm Friedemann's mouth they made a formal Dialogue between the King and the Apologist.

But what is more important than this is that the King gave up his Concert for this evening, and invited Bach, then already called the Old Bach, to try his fortepianos, made by Silbermann, which stood in several rooms of the palace. [Forkel here inserts this footnote: "The pianofortes manufactured by Silbermann, of Freyberg, pleased the King so much, that he resolved to buy them all up. He collected fifteen. I hear that they all now stand unfit for use in various corners of the Royal Palace."] The musicians went with him from room to room, and Bach was invited everywhere to try them and to play unpremeditated compositions. After he had gone on for some time, he asked the King to give him a subject for a Fugue, in order to execute it immediately without any preparation. The King admired the learned manner in which his subject was thus executed extempore; and, probably to see how far such art could be carried, expressed a wish to hear a Fugue with six Obligato parts. But as it is not every subject that is fit for such full harmony, Bach chose one himself, and immediately executed it to the astonishment of all present in the same magnificent and learned manner as he had done that of the King. His Majesty desired also to hear his performance on the organ. The next day therefore Bach was taken to all the organs in Potsdam, as he had before been to Silbermann's fortepianos. After his return to Leipzig, he composed the subject, which he had received from the King, in three and six parts, added several artificial passages in strict canon to it, and had it engraved, under the title of "Musikalisches Opfer" [*Musical Offering*], and dedicated it to the inventor.¹

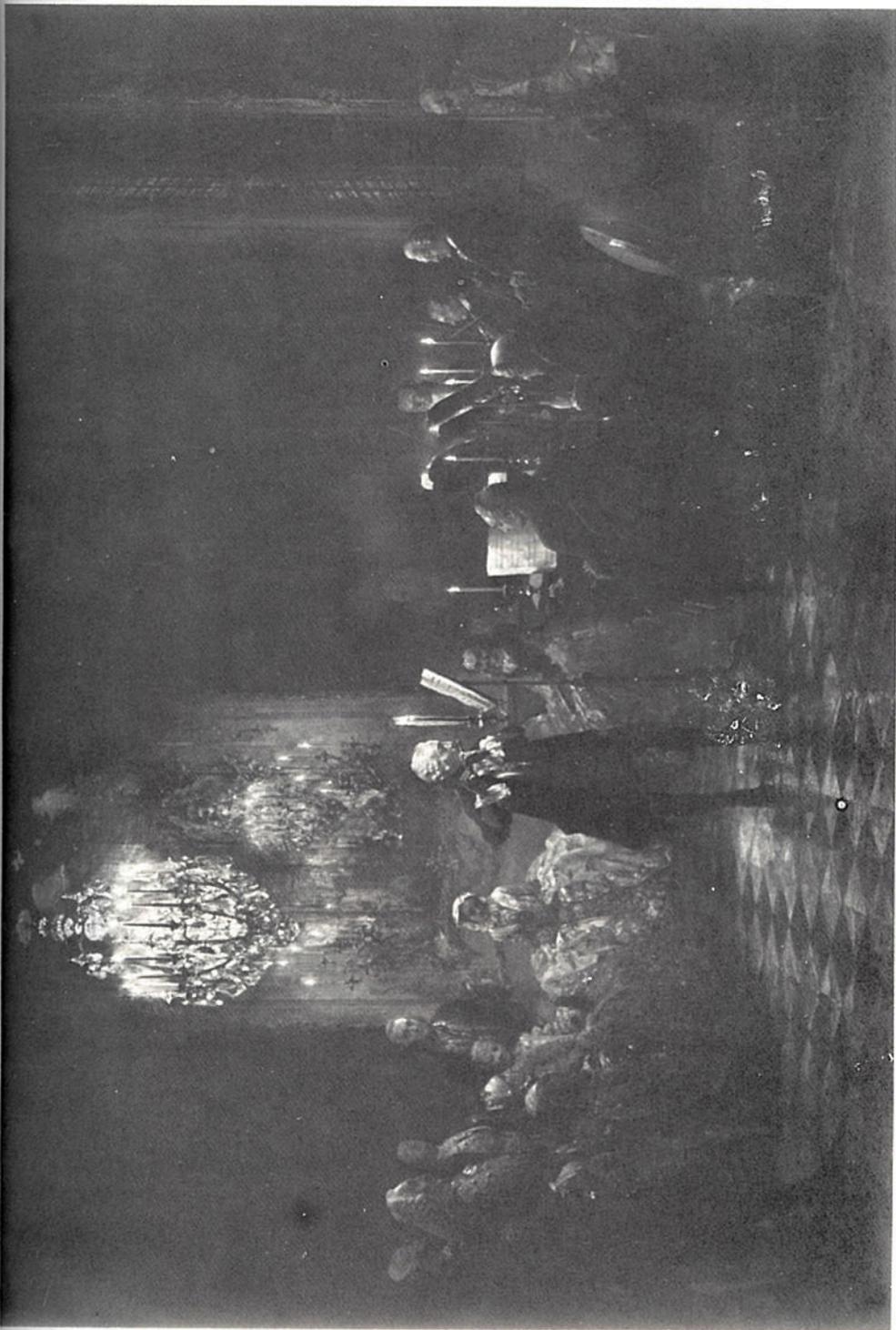


FIGURE 2. Flute Concert in Sanssouci, by Adolph von Menzel (1852).



FIGURE 3. *The Royal Theme.*

When Bach sent a copy of his *Musical Offering* to the King, he included a dedicatory letter, which is of interest for its prose style if nothing else—rather submissive and flattersome. From a modern perspective it seems comical. Also, it probably gives something of the flavor of Bach's apology for his appearance.²

MOST GRACIOUS KING!

In deepest humility I dedicate herewith to Your Majesty a musical offering, the noblest part of which derives from Your Majesty's own august hand. With awesome pleasure I still remember the very special Royal grace when, some time ago, during my visit in Potsdam, Your Majesty's Self deigned to play to me a theme for a fugue upon the clavier, and at the same time charged me most graciously to carry it out in Your Majesty's most august presence. To obey Your Majesty's command was my most humble duty. I noticed very soon, however, that, for lack of necessary preparation, the execution of the task did not fare as well as such an excellent theme demanded. I resolved therefore and promptly pledged myself to work out this right Royal theme more fully, and then make it known to the world. This resolve has now been carried out as well as possible, and it has none other than this irreproachable intent, to glorify, if only in a small point, the fame of a monarch whose greatness and power, as in all the sciences of war and peace, so especially in music, everyone must admire and revere. I make bold to add this most humble request: may Your Majesty deign to dignify the present modest labor with a gracious acceptance, and continue to grant Your Majesty's most august Royal grace to

Your Majesty's
most humble and obedient servant,
THE AUTHOR

Leipzig, July 7
1747

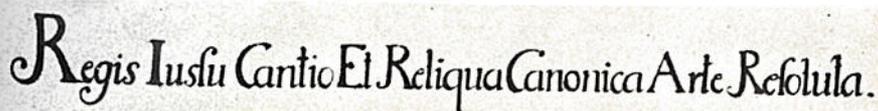
Some twenty-seven years later, when Bach had been dead for twenty-four years, a Baron named Gottfried van Swieten—to whom, incidentally, Forkel dedicated his biography of Bach, and Beethoven dedicated his First Symphony—had a conversation with King Frederick, which he reported as follows:

He [Frederick] spoke to me, among other things, of music, and of a great organist named Bach, who has been for a while in Berlin. This artist [Wilhelm Friedemann Bach] is endowed with a talent superior, in depth of harmonic knowledge and power of execution, to any I have heard or can imagine, while those who knew his father claim that he, in turn, was even greater. The King

is of this opinion, and to prove it to me he sang aloud a chromatic fugue subject which he had given this old Bach, who on the spot had made of it a fugue in four parts, then in five parts, and finally in eight parts.³

Of course there is no way of knowing whether it was King Frederick or Baron van Swieten who magnified the story into larger-than-life proportions. But it shows how powerful Bach's legend had become by that time. To give an idea of how extraordinary a six-part fugue is, in the entire *Well-Tempered Clavier* by Bach, containing forty-eight Preludes and Fugues, only two have as many as five parts, and nowhere is there a six-part fugue! One could probably liken the task of improvising a six-part fugue to the playing of sixty simultaneous blindfold games of chess, and winning them all. To improvise an eight-part fugue is really beyond human capability.

In the copy which Bach sent to King Frederick, on the page preceding the first sheet of music, was the following inscription:



Regis Iussu Cantio Et Reliqua Canonica Arte Resoluta.

FIGURE 4.

("At the King's Command, the Song and the Remainder Resolved with Canonic Art.") Here Bach is punning on the word "canonic", since it means not only "with canons" but also "in the best possible way". The initials of this inscription are

R I C E R C A R

—an Italian word, meaning "to seek". And certainly there is a great deal to seek in the *Musical Offering*. It consists of one three-part fugue, one six-part fugue, ten canons, and a trio sonata. Musical scholars have concluded that the three-part fugue must be, in essence, identical with the one which Bach improvised for King Frederick. The six-part fugue is one of Bach's most complex creations, and its theme is, of course, the Royal Theme. That theme, shown in Figure 3, is a very complex one, rhythmically irregular and highly chromatic (that is, filled with tones which do not belong to the key it is in). To write a decent fugue of even two voices based on it would not be easy for the average musician!

Both of the fugues are inscribed "Ricercar", rather than "Fuga". This is another meaning of the word; "ricercar" was, in fact, the original name for the musical form now known as "fugue". By Bach's time, the word "fugue" (or *fuga*, in Latin and Italian) had become standard, but the term "ricercar" had survived, and now designated an erudite kind of fugue, perhaps too austere intellectually for the common ear. A similar usage survives in English today: the word "recherché" means, literally, "sought out", but carries the same kind of implication, namely of esoteric or high-brow cleverness.

The trio sonata forms a delightful relief from the austerity of the fugues and canons, because it is very melodious and sweet, almost dance-

able. Nevertheless, it too is based largely on the King's theme, chromatic and austere as it is. It is rather miraculous that Bach could use such a theme to make so pleasing an interlude.

The ten canons in the *Musical Offering* are among the most sophisticated canons Bach ever wrote. However, curiously enough, Bach himself never wrote them out in full. This was deliberate. They were posed as puzzles to King Frederick. It was a familiar musical game of the day to give a single theme, together with some more or less tricky hints, and to let the canon based on that theme be "discovered" by someone else. In order to know how this is possible, you must understand a few facts about canons.

Canons and Fugues

The idea of a canon is that one single theme is played against itself. This is done by having "copies" of the theme played by the various participating voices. But there are many ways to do this. The most straightforward of all canons is the round, such as "Three Blind Mice", "Row, Row, Row Your Boat", or "Frère Jacques". Here, the theme enters in the first voice and, after a fixed time-delay, a "copy" of it enters, in precisely the same key. After the same fixed time-delay in the second voice, the third voice enters carrying the theme, and so on. Most themes will not harmonize with themselves in this way. In order for a theme to work as a canon theme, each of its notes must be able to serve in a dual (or triple, or quadruple) role: it must firstly be part of a melody, and secondly it must be part of a harmonization of the same melody. When there are three canonical voices, for instance, each note of the theme must act in two distinct harmonic ways, as well as melodically. Thus, each note in a canon has more than one musical meaning; the listener's ear and brain automatically figure out the appropriate meaning, by referring to context.

There are more complicated sorts of canons, of course. The first escalation in complexity comes when the "copies" of the theme are staggered not only in *time*, but also in *pitch*; thus, the first voice might sing the theme starting on C, and the second voice, overlapping with the first voice, might sing the identical theme starting five notes higher, on G. A third voice, starting on the D yet five notes higher, might overlap with the first two, and so on. The next escalation in complexity comes when the *speeds* of the different voices are not equal; thus, the second voice might sing twice as quickly, or twice as slowly, as the first voice. The former is called *diminution*, the latter *augmentation* (since the theme seems to shrink or to expand).

We are not yet done! The next stage of complexity in canon construction is to *invert* the theme, which means to make a melody which jumps *down* wherever the original theme jumps *up*, and by exactly the same number of semitones. This is a rather weird melodic transformation, but when one has heard many themes inverted, it begins to seem quite natural. Bach was especially fond of inversions, and they show up often in his work—and the *Musical Offering* is no exception. (For a simple example of

inversion, try the tune "Good King Wenceslas". When the original and its inversion are sung together, starting an octave apart and staggered with a time-delay of two beats, a pleasing canon results.) Finally, the most esoteric of "copies" is the retrograde copy—where the theme is played backwards in time. A canon which uses this trick is affectionately known as a *crab canon*, because of the peculiarities of crab locomotion. Bach included a crab canon in the *Musical Offering*, needless to say. Notice that every type of "copy" preserves all the information in the original theme, in the sense that the theme is fully recoverable from any of the copies. Such an information-preserving transformation is often called an *isomorphism*, and we will have much traffic with isomorphisms in this book.

Sometimes it is desirable to relax the tightness of the canon form. One way is to allow slight departures from perfect copying, for the sake of more fluid harmony. Also, some canons have "free" voices—voices which do not employ the canon's theme, but which simply harmonize agreeably with the voices that are in canon with each other.

Each of the canons in the *Musical Offering* has for its theme a different variant of the King's Theme, and all the devices described above for making canons intricate are exploited to the hilt; in fact, they are occasionally combined. Thus, one three-voice canon is labeled "Canon per Augmentationem, contrario Motu"; its middle voice is free (in fact, it sings the Royal Theme), while the other two dance canonically above and below it, using the devices of augmentation and inversion. Another bears simply the cryptic label "Quaerendo invenietis" ("By seeking, you will discover"). All of the canon puzzles have been solved. The canonical solutions were given by one of Bach's pupils, Johann Philipp Kirnberger. But one might still wonder whether there are more solutions to seek!

I should also explain briefly what a fugue is. A fugue is like a canon, in that it is usually based on one theme which gets played in different voices and different keys, and occasionally at different speeds or upside down or backwards. However, the notion of fugue is much less rigid than that of canon, and consequently it allows for more emotional and artistic expression. The telltale sign of a fugue is the way it begins: with a single voice singing its theme. When it is done, then a second voice enters, either five scale-notes up, or four down. Meanwhile the first voice goes on, singing the "countersubject": a secondary theme, chosen to provide rhythmic, harmonic, and melodic contrasts to the subject. Each of the voices enters in turn, singing the theme, often to the accompaniment of the countersubject in some other voice, with the remaining voices doing whatever fanciful things entered the composer's mind. When all the voices have "arrived", then there are no rules. There are, to be sure, standard kinds of things to do—but not so standard that one can merely compose a fugue by formula. The two fugues in the *Musical Offering* are outstanding examples of fugues that could never have been "composed by formula". There is something much deeper in them than mere fugalitv.

All in all, the *Musical Offering* represents one of Bach's supreme accomplishments in counterpoint. It is itself one large intellectual fugue, in

which many ideas and forms have been woven together, and in which playful double meanings and subtle allusions are commonplace. And it is a very beautiful creation of the human intellect which we can appreciate forever. (The entire work is wonderfully described in the book *J. S. Bach's Musical Offering*, by H. T. David.)

An Endlessly Rising Canon

There is one canon in the *Musical Offering* which is particularly unusual. Labeled simply "Canon per Tonos", it has three voices. The uppermost voice sings a variant of the Royal Theme, while underneath it, two voices provide a canonic harmonization based on a second theme. The lower of this pair sings its theme in C minor (which is the key of the canon as a whole), and the upper of the pair sings the same theme displaced upwards in pitch by an interval of a fifth. What makes this canon different from any other, however, is that when it concludes—or, rather, *seems* to conclude—it is no longer in the key of C minor, but now is in D minor. Somehow Bach has contrived to *modulate* (change keys) right under the listener's nose. And it is so constructed that this "ending" ties smoothly onto the beginning again; thus one can repeat the process and return in the key of E, only to join again to the beginning. These successive modulations lead the ear to increasingly remote provinces of tonality, so that after several of them, one would expect to be hopelessly far away from the starting key. And yet magically, after exactly six such modulations, the original key of C minor has been restored! All the voices are exactly one octave higher than they were at the beginning, and here the piece may be broken off in a musically agreeable way. Such, one imagines, was Bach's intention; but Bach indubitably also relished the implication that this process could go on ad infinitum, which is perhaps why he wrote in the margin "As the modulation rises, so may the King's Glory." To emphasize its potentially infinite aspect, I like to call this the "Endlessly Rising Canon".

In this canon, Bach has given us our first example of the notion of *Strange Loops*. The "Strange Loop" phenomenon occurs whenever, by moving upwards (or downwards) through the levels of some hierarchical system, we unexpectedly find ourselves right back where we started. (Here, the system is that of musical keys.) Sometimes I use the term *Tangled Hierarchy* to describe a system in which a Strange Loop occurs. As we go on, the theme of Strange Loops will recur again and again. Sometimes it will be hidden, other times it will be out in the open; sometimes it will be right side up, other times it will be upside down, or backwards. "Quaerendo inveniatis" is my advice to the reader.

Escher

To my mind, the most beautiful and powerful visual realizations of this notion of Strange Loops exist in the work of the Dutch graphic artist M. C. Escher, who lived from 1902 to 1972. Escher was the creator of some of the

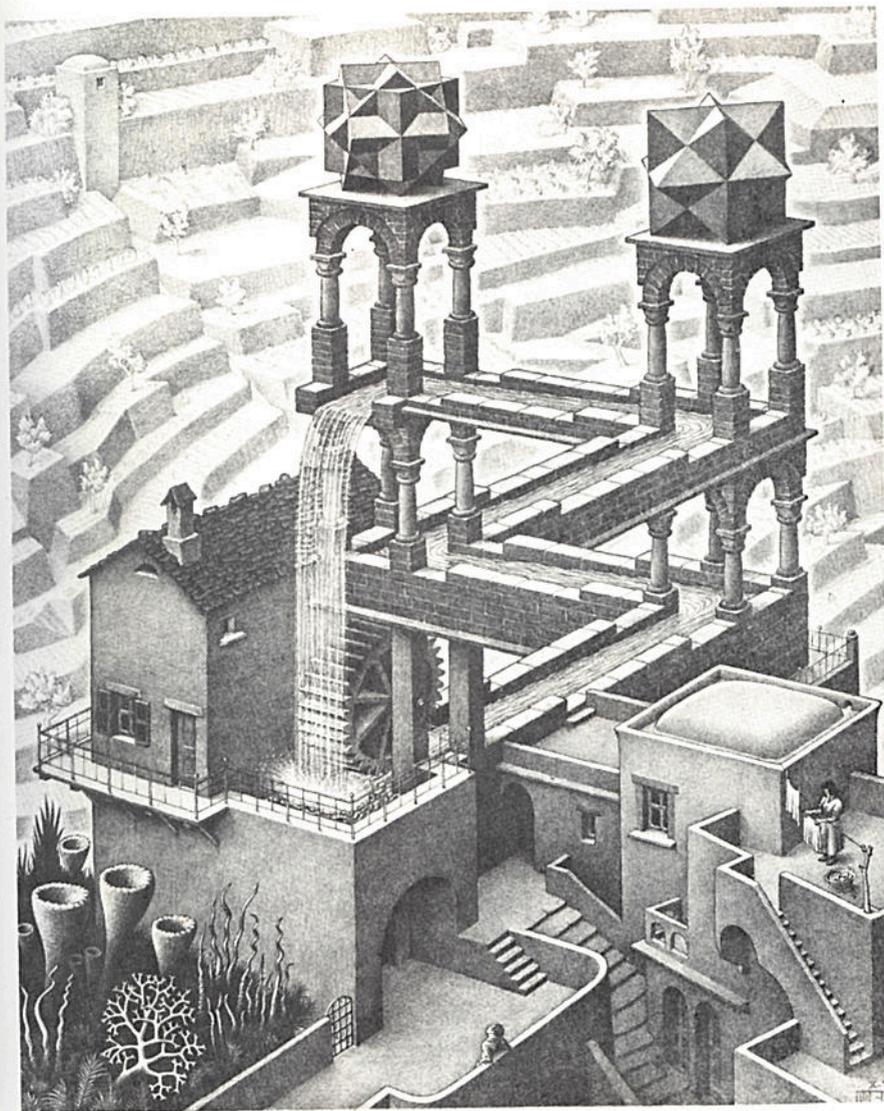


FIGURE 5. Waterfall, by M. C. Escher (lithograph, 1961).

most intellectually stimulating drawings of all time. Many of them have their origin in paradox, illusion, or double-meaning. Mathematicians were among the first admirers of Escher's drawings, and this is understandable because they often are based on mathematical principles of symmetry or pattern . . . But there is much more to a typical Escher drawing than just symmetry or pattern; there is often an underlying idea, realized in artistic form. And in particular, the Strange Loop is one of the most recurrent themes in Escher's work. Look, for example, at the lithograph *Waterfall* (Fig. 5), and compare its six-step endlessly falling loop with the six-step endlessly rising loop of the "Canon per Tonos". The similarity of vision is

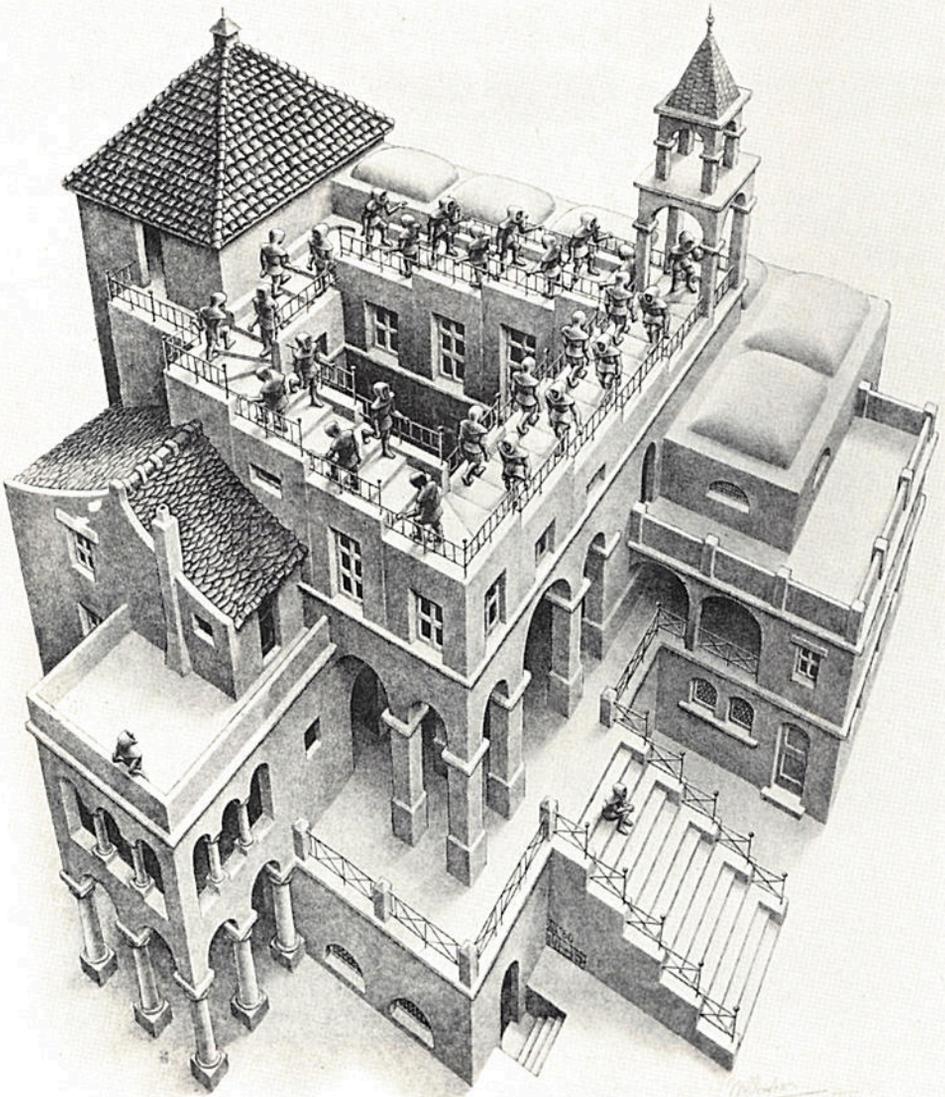
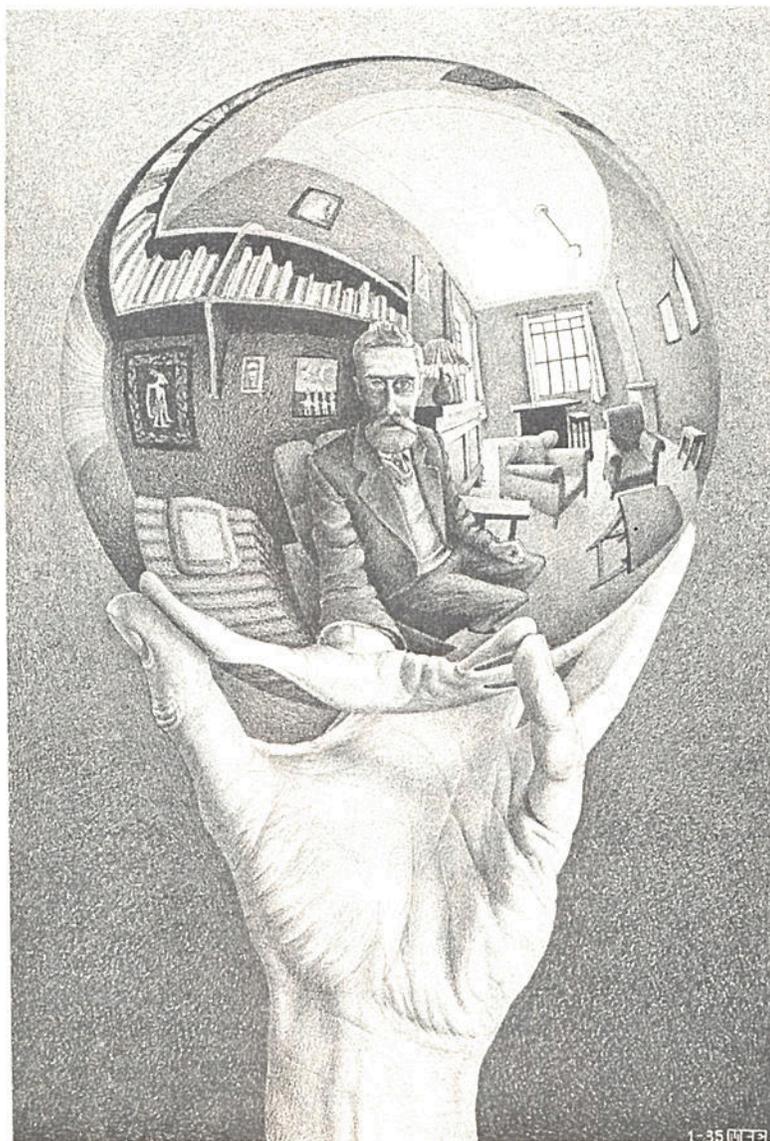


FIGURE 6. Ascending and Descending, by M. C. Escher (lithograph, 1960).

remarkable. Bach and Escher are playing one single theme in two different “keys”: music and art.

Escher realized Strange Loops in several different ways, and they can be arranged according to the tightness of the loop. The lithograph *Ascending and Descending* (Fig. 6), in which monks trudge forever in loops, is the loosest version, since it involves so many steps before the starting point is regained. A tighter loop is contained in *Waterfall*, which, as we already observed, involves only six discrete steps. You may be thinking that there is some ambiguity in the notion of a single “step”—for instance, couldn’t *Ascending and Descending* be seen just as easily as having four levels (staircases) as forty-five levels (stairs)? It is indeed true that there is an inherent

FIGURE 7. Hand with Reflecting Globe. Self-portrait by M. C. Escher (lithograph, 1935).



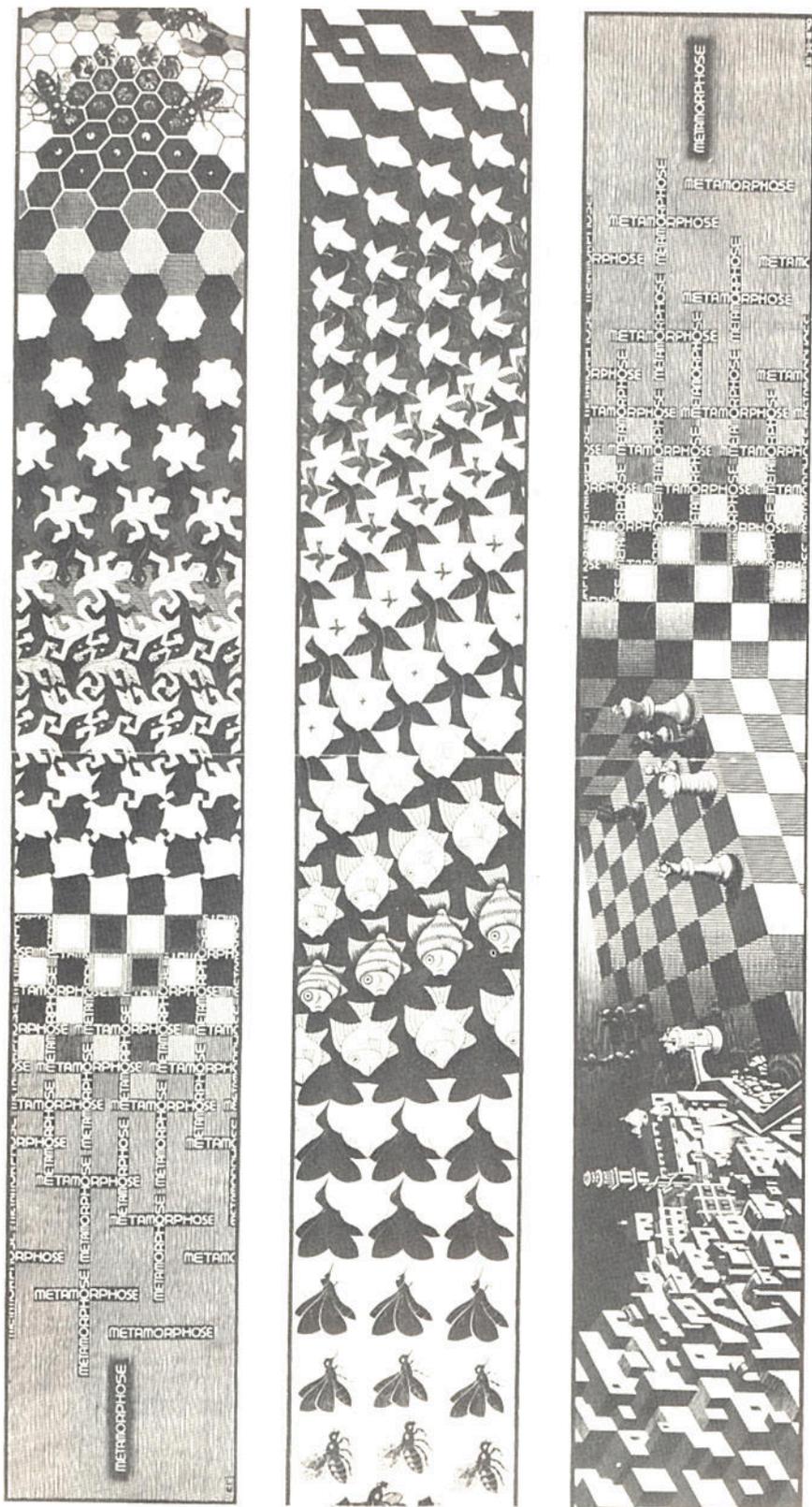


FIGURE 8. Metamorphosis II, by M. C. Escher (woodcut, 19.5 cm. \times 400 cm., 1939-40).

haziness in level-counting, not only in Escher pictures, but in hierarchical, many-level systems. We will sharpen our understanding of this haziness later on. But let us not get too distracted now! As we tighten our loop, we come to the remarkable *Drawing Hands* (Fig. 135), in which each of two hands draws the other: a two-step Strange Loop. And finally, the tightest of all Strange Loops is realized in *Print Gallery* (Fig. 142): a picture of a picture which contains itself. Or is it a picture of a gallery which contains itself? Or of a town which contains itself? Or a young man who contains himself? (Incidentally, the illusion underlying *Ascending and Descending* and *Waterfall* was not invented by Escher, but by Roger Penrose, a British mathematician, in 1958. However, the theme of the Strange Loop was already present in Escher's work in 1948, the year he drew *Drawing Hands*. *Print Gallery* dates from 1956.)

Implicit in the concept of Strange Loops is the concept of infinity, since what else is a loop but a way of representing an endless process in a finite way? And infinity plays a large role in many of Escher's drawings. Copies of one single theme often fit into each other, forming visual analogues to the canons of Bach. Several such patterns can be seen in Escher's famous print *Metamorphosis* (Fig. 8). It is a little like the "Endlessly Rising Canon": wandering further and further from its starting point, it suddenly is back. In the tiled planes of *Metamorphosis* and other pictures, there are already suggestions of infinity. But wilder visions of infinity appear in other drawings by Escher. In some of his drawings, one single theme can appear on different levels of reality. For instance, one level in a drawing might clearly be recognizable as representing fantasy or imagination; another level would be recognizable as reality. These two levels might be the only explicitly portrayed levels. But the mere presence of these two levels invites the viewer to look upon himself as part of yet another level; and by taking that step, the viewer cannot help getting caught up in Escher's implied chain of levels, in which, for any one level, there is always another level above it of greater "reality", and likewise, there is always a level below, "more imaginary" than it is. This can be mind-boggling in itself. However, what happens if the chain of levels is not linear, but forms a loop? What is real, then, and what is fantasy? The genius of Escher was that he could not only concoct, but actually portray, dozens of half-real, half-mythical worlds, worlds filled with Strange Loops, which he seems to be inviting his viewers to enter.

Gödel

In the examples we have seen of Strange Loops by Bach and Escher, there is a conflict between the finite and the infinite, and hence a strong sense of paradox. Intuition senses that there is something mathematical involved here. And indeed in our own century a mathematical counterpart was discovered, with the most enormous repercussions. And, just as the Bach and Escher loops appeal to very simple and ancient intuitions—a musical scale, a staircase—so this discovery, by K. Gödel, of a Strange Loop in

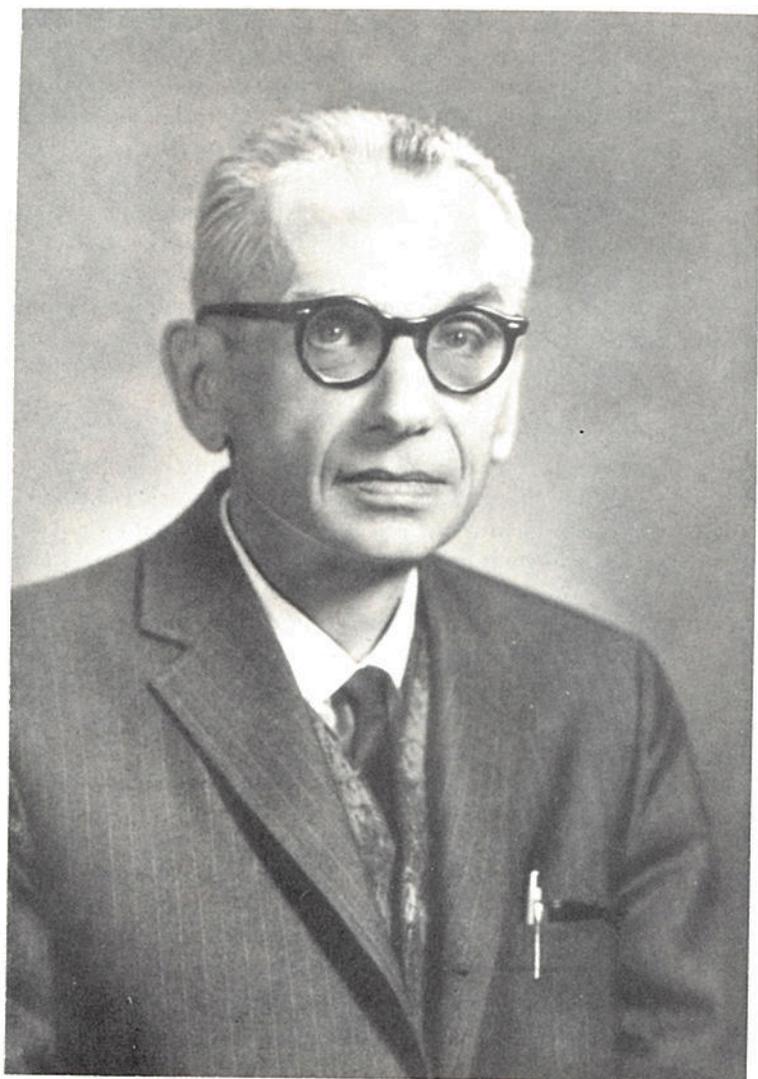


FIGURE 9. Kurt Gödel.

mathematical systems has its origins in simple and ancient intuitions. In its absolutely barest form, Gödel's discovery involves the translation of an ancient paradox in philosophy into mathematical terms. That paradox is the so-called *Epimenides paradox*, or *liar paradox*. Epimenides was a Cretan who made one immortal statement: "All Cretans are liars." A sharper version of the statement is simply "I am lying"; or, "This statement is false". It is that last version which I will usually mean when I speak of the Epimenides paradox. It is a statement which rudely violates the usually assumed dichotomy of statements into true and false, because if you tentatively think it is true, then it immediately backfires on you and makes you think it is false. But once you've decided it is false, a similar backfiring returns you to the idea that it must be true. Try it!

The Epimenides paradox is a one-step Strange Loop, like Escher's *Print Gallery*. But how does it have to do with mathematics? That is what Gödel discovered. His idea was to use mathematical reasoning in exploring mathematical reasoning itself. This notion of making mathematics "introspective" proved to be enormously powerful, and perhaps its richest implication was the one Gödel found: Gödel's Incompleteness Theorem. What the Theorem states and how it is proved are two different things. We shall discuss both in quite some detail in this book. The Theorem can be likened to a pearl, and the method of proof to an oyster. The pearl is prized for its luster and simplicity; the oyster is a complex living beast whose innards give rise to this mysteriously simple gem.

Gödel's Theorem appears as Proposition VI in his 1931 paper "On Formally Undecidable Propositions in *Principia Mathematica* and Related Systems I." It states:

To every ω -consistent recursive class κ of *formulae* there correspond recursive *class-signs* r , such that neither $v \text{ Gen } r$ nor $\text{Neg } (v \text{ Gen } r)$ belongs to $\text{Flg } (\kappa)$ (where v is the *free variable* of r).

Actually, it was in German, and perhaps you feel that it might as well be in German anyway. So here is a paraphrase in more normal English:

All consistent axiomatic formulations of number theory include undecidable propositions.

This is the pearl.

In this pearl it is hard to see a Strange Loop. That is because the Strange Loop is buried in the oyster—the proof. The proof of Gödel's Incompleteness Theorem hinges upon the writing of a self-referential mathematical statement, in the same way as the Epimenides paradox is a self-referential statement of language. But whereas it is very simple to talk about language in language, it is not at all easy to see how a statement about numbers can talk about itself. In fact, it took genius merely to connect the idea of self-referential statements with number theory. Once Gödel had the intuition that such a statement could be created, he was over the major hurdle. The actual creation of the statement was the working out of this one beautiful spark of intuition.

We shall examine the Gödel construction quite carefully in Chapters to come, but so that you are not left completely in the dark, I will sketch here, in a few strokes, the core of the idea, hoping that what you see will trigger ideas in your mind. First of all, the difficulty should be made absolutely clear. Mathematical statements—let us concentrate on number-theoretical ones—are about properties of whole numbers. Whole numbers are not statements, nor are their properties. A statement of number theory is not *about* a statement of number theory; it just *is* a statement of number theory. This is the problem; but Gödel realized that there was more here than meets the eye.

Gödel had the insight that a statement of number theory could be *about* a statement of number theory (possibly even itself), if only numbers could somehow stand for statements. The idea of a *code*, in other words, is at the heart of his construction. In the Gödel Code, usually called “Gödel-numbering”, numbers are made to stand for symbols and sequences of symbols. That way, each statement of number theory, being a sequence of specialized symbols, acquires a Gödel number, something like a telephone number or a license plate, by which it can be referred to. And this coding trick enables statements of number theory to be understood on two different levels: as statements of number theory, and also as *statements about statements* of number theory.

Once Gödel had invented this coding scheme, he had to work out in detail a way of transporting the Epimenides paradox into a number-theoretical formalism. His final transplant of Epimenides did not say, “This statement of number theory is false”, but rather, “This statement of number theory does not have any proof”. A great deal of confusion can be caused by this, because people generally understand the notion of “proof” rather vaguely. In fact, Gödel’s work was just part of a long attempt by mathematicians to explicate for themselves what proofs are. The important thing to keep in mind is that proofs are demonstrations *within fixed systems* of propositions. In the case of Gödel’s work, the fixed system of number-theoretical reasoning to which the word “proof” refers is that of *Principia Mathematica* (*P.M.*), a giant opus by Bertrand Russell and Alfred North Whitehead, published between 1910 and 1913. Therefore, the Gödel sentence *G* should more properly be written in English as:

This statement of number theory does not have any proof
in the system of *Principia Mathematica*.

Incidentally, this Gödel sentence *G* is not Gödel’s Theorem—no more than the Epimenides sentence is the observation that “The Epimenides sentence is a paradox.” We can now state what the effect of discovering *G* is. Whereas the Epimenides statement creates a paradox since it is neither true nor false, the Gödel sentence *G* is unprovable (inside *P.M.*) but true. The grand conclusion? That the system of *Principia Mathematica* is “incomplete”—there are true statements of number theory which its methods of proof are too weak to demonstrate.

But if *Principia Mathematica* was the first victim of this stroke, it was certainly not the last! The phrase “and Related Systems” in the title of Gödel’s article is a telling one; for if Gödel’s result had merely pointed out a defect in the work of Russell and Whitehead, then others could have been inspired to improve upon *P.M.* and to outwit Gödel’s Theorem. But this was not possible: Gödel’s proof pertained to *any* axiomatic system which purported to achieve the aims which Whitehead and Russell had set for themselves. And for each different system, one basic method did the trick. In short, Gödel showed that provability is a weaker notion than truth, no matter what axiomatic system is involved.

Therefore Gödel’s Theorem had an electrifying effect upon logicians, mathematicians, and philosophers interested in the foundations of mathematics, for it showed that no fixed system, no matter how complicated, could represent the complexity of the whole numbers: 0, 1, 2, 3, . . . Modern readers may not be as nonplussed by this as readers of 1931 were, since in the interim our culture has absorbed Gödel’s Theorem, along with the conceptual revolutions of relativity and quantum mechanics, and their philosophically disorienting messages have reached the public, even if cushioned by several layers of translation (and usually obfuscation). There is a general mood of expectation, these days, of “limitative” results—but back in 1931, this came as a bolt from the blue.

Mathematical Logic: A Synopsis

A proper appreciation of Gödel’s Theorem requires a setting of context. Therefore, I will now attempt to summarize in a short space the history of mathematical logic prior to 1931—an impossible task. (See DeLong, Kneebone, or Nagel and Newman, for good presentations of history.) It all began with the attempts to mechanize the thought processes of reasoning. Now our ability to reason has often been claimed to be what distinguishes us from other species; so it seems somewhat paradoxical, on first thought, to mechanize that which is most human. Yet even the ancient Greeks knew that reasoning is a patterned process, and is at least partially governed by storable laws. Aristotle codified syllogisms, and Euclid codified geometry; but thereafter, many centuries had to pass before progress in the study of axiomatic reasoning would take place again.

One of the significant discoveries of nineteenth-century mathematics was that there are different, and equally valid, geometries—where by “a geometry” is meant a theory of properties of abstract points and lines. It had long been assumed that geometry was what Euclid had codified, and that, although there might be small flaws in Euclid’s presentation, they were unimportant and any real progress in geometry would be achieved by extending Euclid. This idea was shattered by the roughly simultaneous discovery of non-Euclidean geometry by several people—a discovery that shocked the mathematics community, because it deeply challenged the idea that mathematics studies the real world. How could there be many differ-

ent kinds of “points” and “lines” in one single reality? Today, the solution to the dilemma may be apparent, even to some nonmathematicians—but at the time, the dilemma created havoc in mathematical circles.

Later in the nineteenth century, the English logicians George Boole and Augustus De Morgan went considerably further than Aristotle in codifying strictly deductive reasoning patterns. Boole even called his book *“The Laws of Thought”*—surely an exaggeration, but it was an important contribution. Lewis Carroll was fascinated by these mechanized reasoning methods, and invented many puzzles which could be solved with them. Gottlob Frege in Jena and Giuseppe Peano in Turin worked on combining formal reasoning with the study of sets and numbers. David Hilbert in Göttingen worked on stricter formalizations of geometry than Euclid’s. All of these efforts were directed towards clarifying what one means by “proof”.

In the meantime, interesting developments were taking place in classical mathematics. A theory of different types of infinities, known as the *theory of sets*, was developed by Georg Cantor in the 1880’s. The theory was powerful and beautiful, but intuition-defying. Before long, a variety of set-theoretical paradoxes had been unearthed. The situation was very disturbing, because just as mathematics seemed to be recovering from one set of paradoxes—those related to the theory of limits, in the calculus—along came a whole new set, which looked worse!

The most famous is Russell’s paradox. Most sets, it would seem, are not members of themselves—for example, the set of walruses is not a walrus, the set containing only Joan of Arc is not Joan of Arc (a set is not a person)—and so on. In this respect, most sets are rather “run-of-the-mill”. However, some “self-swallowing” sets *do* contain themselves as members, such as the set of all sets, or the set of all things except Joan of Arc, and so on. Clearly, every set is either run-of-the-mill or self-swallowing, and no set can be both. Now nothing prevents us from inventing R: *the set of all run-of-the-mill sets*. At first, R might seem a rather run-of-the-mill invention—but that opinion must be revised when you ask yourself, “Is R itself a run-of-the-mill set or a self-swallowing set?” You will find that the answer is: “R is neither run-of-the-mill nor self-swallowing, for either choice leads to paradox.” Try it!

But if R is neither run-of-the-mill nor self-swallowing, then what is it? At the very least, pathological. But no one was satisfied with evasive answers of that sort. And so people began to dig more deeply into the foundations of set theory. The crucial questions seemed to be: “What is wrong with our intuitive concept of ‘set’? Can we make a rigorous theory of sets which corresponds closely with our intuitions, but which skirts the paradoxes?” Here, as in number theory and geometry, the problem is in trying to line up intuition with formalized, or axiomatized, reasoning systems.

A startling variant of Russell’s paradox, called “Grelling’s paradox”, can be made using adjectives instead of sets. Divide the adjectives in English into two categories: those which are self-descriptive, such as “pentasyllabic”, “awkwardnessful”, and “recherché”, and those which are not, such

as “edible”, “incomplete”, and “bisyllabic”. Now if we admit “non-self-descriptive” as an adjective, to which class does it belong? If it seems questionable to include hyphenated words, we can use two terms invented specially for this paradox: *autological* (= “self-descriptive”), and *heterological* (= “non-self-descriptive”). The question then becomes: “Is ‘heterological’ heterological?” Try it!

There seems to be one common culprit in these paradoxes, namely self-reference, or “Strange Loopiness”. So if the goal is to ban all paradoxes, why not try banning self-reference and anything that allows it to arise? This is not so easy as it might seem, because it can be hard to figure out just where self-reference is occurring. It may be spread out over a whole Strange Loop with several steps, as in this “expanded” version of Epimenides, reminiscent of *Drawing Hands*:

The following sentence is false.

The preceding sentence is true.

Taken together, these sentences have the same effect as the original Epimenides paradox; yet separately, they are harmless and even potentially useful sentences. The “blame” for this Strange Loop can’t be pinned on either sentence—only on the way they “point” at each other. In the same way, each local region of *Ascending and Descending* is quite legitimate; it is only the way they are globally put together that creates an impossibility. Since there are indirect as well as direct ways of achieving self-reference, one must figure out how to ban both types at once—if one sees self-reference as the root of all evil.

Banishing Strange Loops

Russell and Whitehead did subscribe to this view, and accordingly, *Principia Mathematica* was a mammoth exercise in exorcising Strange Loops from logic, set theory, and number theory. The idea of their system was basically this. A set of the lowest “type” could contain only “objects” as members—not sets. A set of the next type up could only contain objects, or sets of the lowest type. In general, a set of a given type could only contain sets of lower type, or objects. Every set would belong to a specific type. Clearly, no set could contain itself because it would have to belong to a type higher than its own type. Only “run-of-the-mill” sets exist in such a system; furthermore, old R —the set of all run-of-the-mill sets—no longer is considered a set at all, because it does not belong to any finite type. To all appearances, then, this *theory of types*, which we might also call the “theory of the abolition of Strange Loops”, successfully rids set theory of its paradoxes, but only at the cost of introducing an artificial-seeming hierarchy, and of disallowing the formation of certain kinds of sets—such as the set of all run-of-the-mill sets. Intuitively, this is not the way we imagine sets.

The theory of types handled Russell’s paradox, but it did nothing about the Epimenides paradox or Grelling’s paradox. For people whose

interest went no further than set theory, this was quite adequate—but for people interested in the elimination of paradoxes generally, some similar “hierarchization” seemed necessary, to forbid looping back inside language. At the bottom of such a hierarchy would be an *object language*. Here, reference could be made only to a specific domain—not to aspects of the object language itself (such as its grammatical rules, or specific sentences in it). For that purpose there would be a *metalanguage*. This experience of two linguistic levels is familiar to all learners of foreign languages. Then there would be a *metametalanguage* for discussing the metalanguage, and so on. It would be required that every sentence should belong to some precise level of the hierarchy. Therefore, if one could find no level in which a given utterance fit, then the utterance would be deemed meaningless, and forgotten.

An analysis can be attempted on the two-step Epimenides loop given above. The first sentence, since it speaks of the second, must be on a higher level than the second. But by the same token, the second sentence must be on a higher level than the first. Since this is impossible, the two sentences are “meaningless”. More precisely, such sentences simply cannot be formulated at all in a system based on a strict hierarchy of languages. This prevents all versions of the Epimenides paradox as well as Grelling’s paradox. (To what language level could “heterological” belong?)

Now in set theory, which deals with abstractions that we don’t use all the time, a stratification like the theory of types seems acceptable, even if a little strange—but when it comes to language, an all-pervading part of life, such stratification appears absurd. We don’t think of ourselves as jumping up and down a hierarchy of languages when we speak about various things. A rather matter-of-fact sentence such as, “In this book, I criticize the theory of types” would be doubly forbidden in the system we are discussing. Firstly, it mentions “this book”, which should only be mentionable in a “metabook”—and secondly, it mentions *me*—a person whom I should not be allowed to speak of at all! This example points out how silly the theory of types seems, when you import it into a familiar context. The remedy it adopts for paradoxes—total banishment of self-reference in any form—is a real case of overkill, branding many perfectly good constructions as meaningless. The adjective “meaningless”, by the way, would have to apply to all discussions of the theory of linguistic types (such as that of this very paragraph) for they clearly could not occur on any of the levels—neither object language, nor metalanguage, nor metametalanguage, etc. So the very act of discussing the theory would be the most blatant possible violation of it!

Now one could defend such theories by saying that they were only intended to deal with formal languages—not with ordinary, informal language. This may be so, but then it shows that such theories are extremely academic and have little to say about paradoxes except when they crop up in special tailor-made systems. Besides, the drive to eliminate paradoxes at any cost, especially when it requires the creation of highly artificial formalisms, puts too much stress on bland consistency, and too little on the

quirky and bizarre, which make life and mathematics interesting. It is of course important to try to maintain consistency, but when this effort forces you into a stupendously ugly theory, you know something is wrong.

These types of issues in the foundations of mathematics were responsible for the high interest in codifying human reasoning methods which was present in the early part of this century. Mathematicians and philosophers had begun to have serious doubts about whether even the most concrete of theories, such as the study of whole numbers (number theory), were built on solid foundations. If paradoxes could pop up so easily in set theory—a theory whose basic concept, that of a set, is surely very intuitively appealing—then might they not also exist in other branches of mathematics? Another related worry was that the paradoxes of logic, such as the Epimenides paradox, might turn out to be internal to mathematics, and thereby cast in doubt all of mathematics. This was especially worrisome to those—and there were a good number—who firmly believed that mathematics is simply a branch of logic (or conversely, that logic is simply a branch of mathematics). In fact, this very question—“Are mathematics and logic distinct, or separate?”—was the source of much controversy.

This study of mathematics itself became known as *metamathematics*—or occasionally, *metalogue*, since mathematics and logic are so intertwined. The most urgent priority of metamathematicians was to determine the true nature of mathematical reasoning. What is a legal method of procedure, and what is an illegal one? Since mathematical reasoning had always been done in “natural language” (e.g., French or Latin or some language for normal communication), there was always a lot of possible ambiguity. Words had different meanings to different people, conjured up different images, and so forth. It seemed reasonable and even important to establish a single uniform notation in which all mathematical work could be done, and with the aid of which any two mathematicians could resolve disputes over whether a suggested proof was valid or not. This would require a complete codification of the universally acceptable modes of human reasoning, at least as far as they applied to mathematics.

Consistency, Completeness, Hilbert's Program

This was the goal of *Principia Mathematica*, which purported to derive all of mathematics from logic, and, to be sure, without contradictions! It was widely admired, but no one was sure if (1) all of mathematics really was contained in the methods delineated by Russell and Whitehead, or (2) the methods given were even self-consistent. Was it absolutely clear that contradictory results could *never* be derived, by any mathematicians whatsoever, following the methods of Russell and Whitehead?

This question particularly bothered the distinguished German mathematician (and metamathematician) David Hilbert, who set before the world community of mathematicians (and metamathematicians) this chal-

lenge: to demonstrate rigorously—perhaps following the very methods outlined by Russell and Whitehead—that the system defined in *Principia Mathematica* was both *consistent* (contradiction-free), and *complete* (i.e., that every true statement of number theory could be derived within the framework drawn up in *P.M.*). This was a tall order, and one could criticize it on the grounds that it was somewhat circular: how can you justify your methods of reasoning on the basis of those same methods of reasoning? It is like lifting yourself up by your own bootstraps. (We just don't seem to be able to get away from these Strange Loops!)

Hilbert was fully aware of this dilemma, of course, and therefore expressed the hope that a demonstration of consistency or completeness could be found which depended only on “finitistic” modes of reasoning. These were a small set of reasoning methods usually accepted by mathematicians. In this way, Hilbert hoped that mathematicians could partially lift themselves by their own bootstraps: the sum total of mathematical methods might be proved sound, by invoking only a smaller set of methods. This goal may sound rather esoteric, but it occupied the minds of many of the greatest mathematicians in the world during the first thirty years of this century.

In the thirty-first year, however, Gödel published his paper, which in some ways utterly demolished Hilbert's program. This paper revealed not only that there were irreparable “holes” in the axiomatic system proposed by Russell and Whitehead, but more generally, that no axiomatic system whatsoever could produce all number-theoretical truths, unless it were an inconsistent system! And finally, the hope of proving the consistency of a system such as that presented in *P.M.* was shown to be vain: if such a proof could be found using only methods inside *P.M.*, then—and this is one of the most mystifying consequences of Gödel's work—*P.M.* itself would be inconsistent!

The final irony of it all is that the proof of Gödel's Incompleteness Theorem involved importing the Epimenides paradox right into the heart of *Principia Mathematica*, a bastion supposedly invulnerable to the attacks of Strange Loops! Although Gödel's Strange Loop did not destroy *Principia Mathematica*, it made it far less interesting to mathematicians, for it showed that Russell and Whitehead's original aims were illusory.

Babbage, Computers, Artificial Intelligence . . .

When Gödel's paper came out, the world was on the brink of developing electronic digital computers. Now the idea of mechanical calculating engines had been around for a while. In the seventeenth century, Pascal and Leibniz designed machines to perform fixed operations (addition and multiplication). These machines had no memory, however, and were not, in modern parlance, programmable.

The first human to conceive of the immense computing potential of machinery was the Londoner Charles Babbage (1792-1871). A character who could almost have stepped out of the pages of the *Pickwick Papers*,

Babbage was most famous during his lifetime for his vigorous campaign to rid London of “street nuisances”—organ grinders above all. These pests, loving to get his goat, would come and serenade him at any time of day or night, and he would furiously chase them down the street. Today, we recognize in Babbage a man a hundred years ahead of his time: not only inventor of the basic principles of modern computers, he was also one of the first to battle noise pollution.

His first machine, the “Difference Engine”, could generate mathematical tables of many kinds by the “method of differences”. But before any model of the “D.E.” had been built, Babbage became obsessed with a much more revolutionary idea: his “Analytical Engine”. Rather immodestly, he wrote, “The course through which I arrived at it was the most entangled and perplexed which probably ever occupied the human mind.”⁴ Unlike any previously designed machine, the A.E. was to possess both a “store” (memory) and a “mill” (calculating and decision-making unit). These units were to be built of thousands of intricate geared cylinders interlocked in incredibly complex ways. Babbage had a vision of numbers swirling in and out of the mill under control of a *program* contained in punched cards—an idea inspired by the Jacquard loom, a card-controlled loom that wove amazingly complex patterns. Babbage’s brilliant but ill-fated Countess friend, Lady Ada Lovelace (daughter of Lord Byron), poetically commented that “the Analytical Engine *weaves algebraic patterns* just as the Jacquard-loom weaves flowers and leaves.” Unfortunately, her use of the present tense was misleading, for no A.E. was ever built, and Babbage died a bitterly disappointed man.

Lady Lovelace, no less than Babbage, was profoundly aware that with the invention of the Analytical Engine, mankind was flirting with mechanized intelligence—particularly if the Engine were capable of “eating its own tail” (the way Babbage described the Strange Loop created when a machine reaches in and alters its own stored program). In an 1842 memoir,⁵ she wrote that the A.E. “might act upon other things besides *number*”. While Babbage dreamt of creating a chess or tic-tac-toe automaton, she suggested that his Engine, with pitches and harmonies coded into its spinning cylinders, “might compose elaborate and scientific pieces of music of any degree of complexity or extent.” In nearly the same breath, however, she cautions that “The Analytical Engine has no pretensions whatever to *originate* anything. It can do whatever we *know how to order it to perform*.” Though she well understood the power of artificial computation, Lady Lovelace was skeptical about the artificial creation of intelligence. However, could her keen insight allow her to dream of the potential that would be opened up with the taming of electricity?

In our century the time was ripe for computers—computers beyond the wildest dreams of Pascal, Leibniz, Babbage, or Lady Lovelace. In the 1930’s and 1940’s, the first “giant electronic brains” were designed and built. They catalyzed the convergence of three previously disparate areas: the theory of axiomatic reasoning, the study of mechanical computation, and the psychology of intelligence.

These same years saw the theory of computers develop by leaps and

Ada Lovelace
on MUSIC

!

bounds. This theory was tightly linked to metamathematics. In fact, Gödel's Theorem has a counterpart in the theory of computation, discovered by Alan Turing, which reveals the existence of ineluctable "holes" in even the most powerful computer imaginable. Ironically, just as these somewhat eerie limits were being mapped out, real computers were being built whose powers seemed to grow and grow beyond their makers' power of prophecy. Babbage, who once declared he would gladly give up the rest of his life if he could come back in five hundred years and have a three-day guided scientific tour of the new age, would probably have been thrilled speechless a mere century after his death—both by the new machines, and by their unexpected limitations.

By the early 1950's, mechanized intelligence seemed a mere stone's throw away; and yet, for each barrier crossed, there always cropped up some new barrier to the actual creation of a genuine thinking machine. Was there some deep reason for this goal's mysterious recession?

No one knows where the borderline between non-intelligent behavior and intelligent behavior lies; in fact, to suggest that a sharp borderline exists is probably silly. But essential abilities for intelligence are certainly:

- to respond to situations very flexibly;
- to take advantage of fortuitous circumstances;
- to make sense out of ambiguous or contradictory messages;
- to recognize the relative importance of different elements of a situation;
- to find similarities between situations despite differences which may separate them;
- to draw distinctions between situations despite similarities which may link them;
- to synthesize new concepts by taking old concepts and putting them together in new ways;
- to come up with ideas which are novel.

Here one runs up against a seeming paradox. Computers by their very nature are the most inflexible, desireless, rule-following of beasts. Fast though they may be, they are nonetheless the epitome of unconsciousness. How, then, can intelligent behavior be programmed? Isn't this the most blatant of contradictions in terms? One of the major theses of this book is that it is not a contradiction at all. One of the major purposes of this book is to urge each reader to confront the apparent contradiction head on, to savor it, to turn it over, to take it apart, to wallow in it, so that in the end the reader might emerge with new insights into the seemingly unbreachable gulf between the formal and the informal, the animate and the inanimate, the flexible and the inflexible.

This is what Artificial Intelligence (AI) research is all about. And the strange flavor of AI work is that people try to put together long sets of rules in strict formalisms which tell inflexible machines how to be flexible.

What sorts of "rules" could possibly capture all of what we think of as intelligent behavior, however? Certainly there must be rules on all sorts of

different levels. There must be many “just plain” rules. There must be “metarules” to modify the “just plain” rules; then “metametarules” to modify the metarules, and so on. The flexibility of intelligence comes from the enormous number of different rules, and levels of rules. The reason that so many rules on so many different levels must exist is that in life, a creature is faced with millions of situations of completely different types. In some situations, there are stereotyped responses which require “just plain” rules. Some situations are mixtures of stereotyped situations—thus they require rules for deciding which of the “just plain” rules to apply. Some situations cannot be classified—thus there must exist rules for inventing new rules . . . and on and on. Without doubt, Strange Loops involving rules that change themselves, directly or indirectly, are at the core of intelligence. Sometimes the complexity of our minds seems so overwhelming that one feels that there can be no solution to the problem of understanding intelligence—that it is wrong to think that rules of any sort govern a creature’s behavior, even if one takes “rule” in the multilevel sense described above.

. . . and Bach

In the year 1754, four years after the death of J. S. Bach, the Leipzig theologian Johann Michael Schmidt wrote, in a treatise on music and the soul, the following noteworthy passage:

Not many years ago it was reported from France that a man had made a statue that could play various pieces on the *Fleuttraversiere*, placed the flute to its lips and took it down again, rolled its eyes, etc. But no one has yet invented an image that thinks, or wills, or composes, or even does anything at all similar. Let anyone who wishes to be convinced look carefully at the last fugal work of the above-praised Bach, which has appeared in copper engraving, but which was left unfinished because his blindness intervened, and let him observe the art that is contained therein; or what must strike him as even more wonderful, the Chorale which he dictated in his blindness to the pen of another: *Wenn wir in höchsten Nöthen seyn*. I am sure that he will soon need his soul if he wishes to observe all the beauties contained therein, let alone wishes to play it to himself or to form a judgment of the author. Everything that the champions of Materialism put forward must fall to the ground in view of this single example.⁶

Quite likely, the foremost of the “champions of Materialism” here alluded to was none other than Julien Offroy de la Mettrie—philosopher at the court of Frederick the Great, author of *L’homme machine* (“Man, the Machine”), and Materialist Par Excellence. It is now more than 200 years later, and the battle is still raging between those who agree with Johann Michael Schmidt, and those who agree with Julien Offroy de la Mettrie. I hope in this book to give some perspective on the battle.

“Gödel, Escher, Bach”

The book is structured in an unusual way: as a counterpoint between Dialogues and Chapters. The purpose of this structure is to allow me to

present new concepts twice: almost every new concept is first presented metaphorically in a Dialogue, yielding a set of concrete, visual images; then these serve, during the reading of the following Chapter, as an intuitive background for a more serious and abstract presentation of the same concept. In many of the Dialogues I appear to be talking about one idea on the surface, but in reality I am talking about some other idea, in a thinly disguised way.

Originally, the only characters in my Dialogues were Achilles and the Tortoise, who came to me from Zeno of Elea, by way of Lewis Carroll. Zeno of Elea, inventor of paradoxes, lived in the fifth century B.C. One of his paradoxes was an allegory, with Achilles and the Tortoise as protagonists. Zeno's invention of the happy pair is told in my first Dialogue, *Three-Part Invention*. In 1895, Lewis Carroll reincarnated Achilles and the Tortoise for the purpose of illustrating his own new paradox of infinity. Carroll's paradox, which deserves to be far better known than it is, plays a significant role in this book. Originally titled "What the Tortoise Said to Achilles", it is reprinted here as *Two-Part Invention*.

When I began writing Dialogues, somehow I connected them up with musical forms. I don't remember the moment it happened; I just remember one day writing "Fugue" above an early Dialogue, and from then on the idea stuck. Eventually I decided to pattern each Dialogue in one way or another on a different piece by Bach. This was not so inappropriate. Old Bach himself used to remind his pupils that the separate parts in their compositions should behave like "persons who conversed together as if in a select company". I have taken that suggestion perhaps rather more literally than Bach intended it; nevertheless I hope the result is faithful to the meaning. I have been particularly inspired by aspects of Bach's compositions which have struck me over and over, and which are so well described by David and Mendel in *The Bach Reader*:

His form in general was based on relations between separate sections. These relations ranged from complete identity of passages on the one hand to the return of a single principle of elaboration or a mere thematic allusion on the other. The resulting patterns were often symmetrical, but by no means necessarily so. Sometimes the relations between the various sections make up a maze of interwoven threads that only detailed analysis can unravel. Usually, however, a few dominant features afford proper orientation at first sight or hearing, and while in the course of study one may discover unending subtleties, one is never at a loss to grasp the unity that holds together every single creation by Bach.⁶

I have sought to weave an Eternal Golden Braid out of these three strands: Gödel, Escher, Bach. I began, intending to write an essay at the core of which would be Gödel's Theorem. I imagined it would be a mere pamphlet. But my ideas expanded like a sphere, and soon touched Bach and Escher. It took some time for me to think of making this connection explicit, instead of just letting it be a private motivating force. But finally I realized that to me, Gödel and Escher and Bach were only shadows cast in different directions by some central solid essence. I tried to reconstruct the central object, and came up with this book.

more poetic
form of
thesis

Three-Part Invention

Achilles (a Greek warrior, the fleetest of foot of all mortals) and a Tortoise are standing together on a dusty runway in the hot sun. Far down the runway, on a tall flagpole, there hangs a large rectangular flag. The flag is solid red, except where a thin ring-shaped hole has been cut out of it, through which one can see the sky.

Achilles: What is that strange flag down at the other end of the track? It reminds me somehow of a print by my favorite artist, M. C. Escher.

Tortoise: That is Zeno's flag.

Achilles: Could it be that the hole in it resembles the holes in a Möbius strip Escher once drew? Something is wrong about that flag, I can tell.

Tortoise: The ring which has been cut from it has the shape of the numeral for zero, which is Zeno's favorite number.

Achilles: But zero hasn't been invented yet! It will only be invented by a Hindu mathematician some millennia hence. And thus, Mr. T, my argument proves that such a flag is impossible.

Tortoise: Your argument is persuasive, Achilles, and I must agree that such a flag is indeed impossible. But it is beautiful anyway, is it not?

Achilles: Oh, yes, there is no doubt of its beauty.

Tortoise: I wonder if its beauty is related to its impossibility. I don't know; I've never had the time to analyze Beauty. It's a Capitalized Essence; and I never seem to have the time for Capitalized Essences.

Achilles: Speaking of Capitalized Essences, Mr. T, have you ever wondered about the Purpose of Life?

Tortoise: Oh, heavens, no.

Achilles: Haven't you ever wondered why we are here, or who invented us?

Tortoise: Oh, that is quite another matter. We are inventions of Zeno (as you will shortly see); and the reason we are here is to have a footrace.

Achilles: A footrace? How outrageous! Me, the fleetest of foot of all mortals, versus you, the ploddingest of all plodders! There can be no point to such a race.

Tortoise: You might give me a head start.

Achilles: It would have to be a huge one.

Tortoise: I don't object.

Achilles: But I will catch you, sooner or later—most likely sooner.

Tortoise: Not if things go according to Zeno's paradox, you won't. Zeno is hoping to use our footrace to show that motion is impossible, you see. It is only in the mind that motion seems possible, according to Zeno. In truth, Motion Is Inherently Impossible. He proves it quite elegantly.

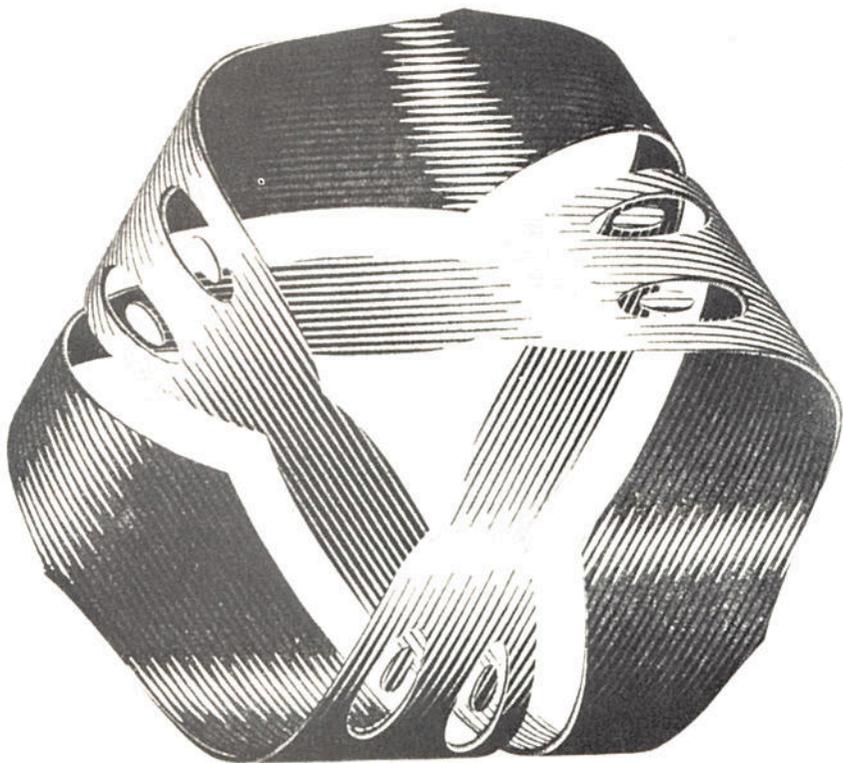


FIGURE 10. Möbius Strip I, by M. C. Escher (wood-engraving printed from four blocks, 1961).

- Achilles:* Oh, yes, it comes back to me now: the famous Zen kōan about Zen Master Zeno. As you say, it is very simple indeed.
- Tortoise:* Zen kōan? Zen Master? What do you mean?
- Achilles:* It goes like this: Two monks were arguing about a flag. One said, "The flag is moving." The other said, "The wind is moving." The sixth patriarch, Zeno, happened to be passing by. He told them, "Not the wind, not the flag; mind is moving."
- Tortoise:* I am afraid you are a little befuddled, Achilles. Zeno is no Zen master; far from it. He is, in fact, a Greek philosopher from the town of Elea (which lies halfway between points A and B). Centuries hence, he will be celebrated for his paradoxes of motion. In one of those paradoxes, this very footrace between you and me will play a central role.
- Achilles:* I'm all confused. I remember vividly how I used to repeat over and over the names of the six patriarchs of Zen, and I always said, "The sixth patriarch is Zeno, the sixth patriarch is Zeno . . ." (*Suddenly a soft warm breeze picks up.*) Oh, look, Mr. Tortoise—look at the flag waving! How I love to watch the ripples shimmer through its soft fabric. And the ring cut out of it is waving, too!

Tortoise: Don't be silly. The flag is impossible, hence it can't be waving.
The wind is waving.

(At this moment, Zeno happens by.)

Zeno: Hallo! Hulloo! What's up? What's new?

Achilles: The flag is moving.

Tortoise: The wind is moving.

Zeno: O Friends, Friends! Cease your argumentation! Arrest your vitriolics! Abandon your discord! For I shall resolve the issue for you forthwith. Ho! And on such a fine day!

Achilles: This fellow must be playing the fool.

Tortoise: No, wait, Achilles. Let us hear what he has to say. Oh, Unknown Sir, do impart to us your thoughts on this matter.

Zeno: Most willingly. Not the wind, not the flag—neither one is moving, nor is anything moving at all. For I have discovered a great Theorem, which states: "Motion Is Inherently Impossible." And from this Theorem follows an even greater Theorem—Zeno's Theorem: "Motion Unexists."

Achilles: "Zeno's Theorem"? Are you, sir, by any chance, the philosopher Zeno of Elea?

Zeno: I am indeed, Achilles.

Achilles (scratching his head in puzzlement): Now how did he know my name?

Zeno: Could I possibly persuade you two to hear me out as to why this is the case? I've come all the way to Elea from point A this afternoon, just trying to find someone who'll pay some attention to my closely honed argument. But they're all hurrying hither and thither, and they don't have time. You've no idea how disheartening it is to meet with refusal after refusal. Oh, but I'm sorry to burden you with my troubles. I'd just like to ask one thing: Would the two of you humor a silly old philosopher for a few moments—only a few, I promise you—in his eccentric theories?

Achilles: Oh, by all means! Please do illuminate us! I know I speak for both of us, since my companion, Mr. Tortoise, was only moments ago speaking of you with great veneration—and he mentioned especially your paradoxes.

Zeno: Thank you. You see, my Master, the fifth patriarch, taught me that reality is one, immutable, and unchanging; all plurality, change, and motion are mere illusions of the senses. Some have mocked his views; but I will show the absurdity of their mockery. My argument is quite simple. I will illustrate it with two characters of my own Invention: Achilles (a Greek warrior, the fleetest of foot of all mortals), and a Tortoise. In my tale, they are persuaded by a passerby to run a footrace down a runway towards a distant flag waving in the breeze. Let us assume that, since the Tortoise is a much slower runner, he gets a head start of, say, ten rods. Now the race begins. In a few bounds, Achilles has reached the spot where the Tortoise started.

Achilles: Hah!

Zeno: And now the Tortoise is but a single rod ahead of Achilles. Within only a moment, Achilles has attained that spot.

Achilles: Ho ho!

Zeno: Yet, in that short moment, the Tortoise has managed to advance a slight amount. In a flash, Achilles covers that distance, too.

Achilles: Hee hee hee!

Zeno: But in that very short flash, the Tortoise has managed to inch ahead by ever so little, and so Achilles is still behind. Now you see that in order for Achilles to catch the Tortoise, this game of "try-to-catch-me" will have to be played an INFINITE number of times—and therefore Achilles can NEVER catch up with the Tortoise!

Tortoise: Heh heh heh heh!

Achilles: Hmm . . . That argument sounds wrong to me. And yet, I can't quite make out what's wrong with it.

Zeno: Isn't it a teaser? It's my favorite paradox.

Tortoise: Excuse me, Zeno, but I believe your tale illustrates the wrong principle, does it not? You have just told us what will come to be known, centuries hence, as Zeno's "Achilles paradox", which shows (ahem!) that Achilles will never catch the Tortoise; but the proof that Motion Is Inherently Impossible (and thence that Motion Unexists) is your "dichotomy paradox", isn't that so?

Zeno: Oh, shame on me. Of course, you're right. That's the one about how, in getting from A to B, one has to go halfway first—and of that stretch one also has to go halfway, and so on and so forth. But you see, both those paradoxes really have the same flavor. Frankly, I've only had one Great Idea—I just exploit it in different ways.

Achilles: I swear, these arguments contain a flaw. I don't quite see where, but they cannot be correct.

Zeno: You doubt the validity of my paradox? Why not just try it out? You see that red flag down there, at the far end of the runway?

Achilles: The impossible one, based on an Escher print?

Zeno: Exactly. What do you say to you and Mr. Tortoise racing for it, allowing Mr. T a fair head start of, well, I don't know—

Tortoise: How about ten rods?

Zeno: Very good—ten rods.

Achilles: Any time.

Zeno: Excellent! How exciting! An empirical test of my rigorously proven Theorem! Mr. Tortoise, will you position yourself ten rods upwind?

(The Tortoise moves ten rods closer to the flag.)

Are you both ready?

Tortoise and Achilles: Ready!

Zeno: On your mark! Get set! Go!