The Ways of Paradox
and Other Essays

REVISED AND ENLARGED EDITION

W. V. Quine

Harvard University Press
Cambridge, Massachusetts
and London, England

1

The Ways of Paradox

Frederic, protagonist of *The Pirates of Penzance*, has reached the age of 21 after passing only five birthdays. Several circumstances conspire to make this possible. Age is reckoned in elapsed time, whereas a birthday has to match the date of birth; and February 29 comes less frequently than once a year.

Granted that Frederic's situation is possible, wherein is it paradoxical? Merely in its initial air of absurdity. The likelihood that a man will be more than *n* years old on his *n*th birthday is as little as 1 to 1,460, or slightly better if we allow for seasonal trends; and this likelihood is so slight that we easily forget its existence.

May we say in general, then, that a paradox is just any conclusion that at first sounds absurd but that has an argument to sustain it? In the end I think this account stands up pretty well. But it leaves much unsaid. The argument that sustains a paradox may expose the absurdity of a buried premise or of some preconception previously reckoned as central to physical theory, to mathematics, or to the thinking process. Catastrophe may lurk, therefore, in the most innocent-seeming paradox. More than once in history the discovery of paradox has been the occasion for major reconstruction at the foundations of thought. For some decades, indeed, studies of the foundation of mathematics have been confounded and greatly stimulated by confrontation with

---

I presented this as a lecture at the University of Akron, November 1961, and again at a Shop Club at Harvard. It was published under the title "Paradox" in *Scientific American* (Volume 206, 1962).

two paradoxes, one propounded by Bertrand Russell in 1901 and the other by Kurt Gödel in 1931.

As a first step onto this dangerous ground, let us consider another paradox: that of the village barber. This is not Russell's great paradox of 1901, to which we shall come, but a lesser one that Russell attributed to an unnamed source in 1918. In a certain village there is a man, so the paradox runs, who is a barber; this barber shaves all and only those men in the village who do not shave themselves. Query: Does the barber shave himself?

Any man in this village is shaved by the barber if and only if he is not shaved by himself. Therefore in particular the barber shaves himself if and only if he does not. We are in trouble if we say the barber shaves himself and we are in trouble if we say he does not.

Now compare the two paradoxes. Frederic's situation seemed absurd at first, but a simple argument sufficed to make us acquiesce in it for good. In the case of the barber, on the other hand, the conclusion is too absurd to acquiesce in at any time.

What are we to say to the argument that goes to prove this unacceptable conclusion? Happily it rests on assumptions. We are asked to swallow a story about a village and a man in it who shaves all and only those men in the village who do not shave themselves. This is the source of our trouble; grant this and we end up saying, absurdly, that the barber shaves himself if and only if he does not. The proper conclusion to draw is just that there is no such barber. We are confronted with nothing more mysterious than what logicians have been referring to for a couple of thousand years as a *reductio ad absurdum*. We disprove the barber by assuming him and deducing the absurdity that he shaves himself if and only if he does not. The paradox is simply a proof that no village can contain a man who shaves all and only those men in it who do not shave themselves. This sweeping denial at first sounds absurd; why should there not be such a man in a village? But the argument shows why not, and so we acquiesce in the sweeping denial just as we acquiesced in the possibility, absurd on first exposure, of Frederic's being so much more than five years old on his fifth birthday.

Both paradoxes are alike, after all, in sustaining prima facie absurdities by conclusive argument. What is strange but true in

the one paradox is that one can be $4n$ years old on one's nth birthday; what is strange but true in the other paradox is that no village can contain a man who shaves all and only those men in the village who do not shave themselves.

Still, I would not limit the word 'paradox' to cases where what is purportedly established is true. I shall call these, more particularly, veridical, or truth-telling, paradoxes. For the name of paradox is suited equally to falsidical ones. (This word is not so barbarous as it sounds; *falsidicus* occurs twice in Plautus and twice in earlier writers.)

The Frederic paradox is a veridical one if we take its proposition not as something about Frederic but as the abstract truth that a man can be $4n$ years old on his nth birthday. Similarly, the barber paradox is a veridical one if we take its proposition as being that no village contains such a barber. A falsidical paradox, on the other hand, is one whose proposition not only seems at first absurd but also is false, there being a fallacy in the purported proof. Typical falsidical paradoxes are the comic misproofs that $2 = 1$. Most of us have heard one or another such. Here is the version offered by Augustus De Morgan: Let $x = 1$. Then $x^2 = x$. So $x^2 - 1 = x - 1$. Dividing both sides by $x - 1$, we conclude that $x + 1 = 1$; that is, since $x = 1$, $2 = 1$. The fallacy comes in the division by $x - 1$, which is 0.

Instead of 'falsidical paradox' could I say simply 'fallacy'? Not quite. Fallacies can lead to true conclusions as well as false ones, and to unsurprising conclusions as well as surprising ones. In a falsidical paradox there is always a fallacy in the argument, but the proposition purportedly established has furthermore to seem absurd and to be indeed false.

Some of the ancient paradoxes of Zeno belong under the head of falsidical paradoxes. Take the one about Achilles and the tortoise. Generalized beyond these two fictitious characters, what the paradox purports to establish is the absurd proposition that so long as a runner keeps running, however slowly, another runner can never overtake him. The argument is that each time the pursuer reaches a spot where the pursued has been, the pursued has moved a bit beyond. When we try to make this argument more explicit, the fallacy that emerges is the mistaken notion that any infinite succession of intervals of time has to add up to all eternity. Actually when an infinite succession of intervals of time is so chosen that the succeeding intervals become shorter and shorter, the whole succession may take either a finite or an infinite time. It is a question of a convergent series.

GRELLING'S PARADOX

The realm of paradox is not clearly exhausted even by the veridical and falsidical paradoxes together. The most startling of all paradoxes are not clearly assignable to either of these domains. Consider the paradox, devised by Kurt Grelling in 1908, concerning the heterological, or non-self-descriptive, adjectives.

To explain this paradox requires first a definition of the autological, or self-descriptive, adjective. The adjective 'short' is short; the adjective 'English' is English; the adjective 'adjectival' is adjectival; the adjective 'polysyllabic' is polysyllabic. Each of these adjectives is, in Grelling's terminology, autological: each is true of itself. Other adjectives are heterological; thus 'long', which is not a long adjective; 'German', which is not a German adjective; 'monosyllabic', which is not a monosyllabic one.

Grelling's paradox arises from the query: Is the adjective 'heterological' an autological or a heterological one? We are as badly off here as we were with the barber. If we decide that 'heterological' is autological, then the adjective is true of itself. But that makes it heterological rather than autological, since whatever the adjective 'heterological' is true of is heterological. If we therefore decide that the adjective 'heterological' is heterological, then it is true of itself, and that makes it autological.

Our recourse in a comparable quandary over the village barber was to declare a *reductio ad absurdum* and conclude that there was no such barber. Here, however, there is no interim premise to disavow. We merely defined the adjective 'heterological' and asked if it was heterological. In fact, we can get the paradox just as well without the adjective and its definition. 'Heterological' was defined as meaning 'not true of self'; we can therefore ask if the adjectival phrase 'not true of self' is true of itself. We find that it is if and only if it is not, hence that it is and it is not; and so we have our paradox.

Thus viewed, Grelling's paradox seems unequivocally falsidical. Its proposition is a self-contradictory compound proposition

to the effect that our adjective is and is not true of itself. But this paradox contrasts strangely with the falsidical paradoxes of Zeno, or of '$2 = 1$', in that we are at a loss to spot the fallacy in the argument. It may for this reason be best seen as representing a third class of paradoxes, separate from the veridical and falsidical ones.

ANTINOMIES

The paradoxes of this class are called antinomies, and it is they that bring on the crises in thought. An antinomy produces a self-contradiction by accepted ways of reasoning. It establishes that some tacit and trusted pattern of reasoning must be made explicit and henceforward be avoided or revised.

Take Grelling's paradox, in the form in which it shows the adjective phrase 'not true of self' to be both true and false of itself. What tacit principles of reasoning does the argument depend on? Notably this one: the adjective 'red' is true of a thing if and only if the thing is red; the adjective 'big' is true of a thing if and only if the thing is big; the adjective 'not true of self' is true of a thing if and only if the thing is not true of itself; and so on. This last case of the principle is the case that issues directly in the paradox.

There is no denying that this principle is constantly used, tacitly, when we speak of adjectives as true of things: the adjective 'red' is true of a thing if and only if the thing is red, and correspondingly for all adjectives. This principle simply reflects what we mean in saying that adjectives are true of things. It is a hard principle to distrust, and yet it is obviously the principle that is to blame for our antinomy. The antinomy is directly a case of this principle. Take the adjective in the principle as the adjectival phrase 'not true of self' instead of the adjective 'red', and take the 'thing' in the principle, of which the adjective is to be true, as that adjective over again; thereupon the principle says outright that 'not true of self' is true of itself if and only if it is not true of itself. So the principle must be abandoned or at least somehow restricted.

Yet so faithfully does the principle reflect what we mean in calling adjectives true of things that we cannot abandon it without abjuring the very expression 'true of' as pernicious nonsense. We could still go on using the adjectives themselves that had been said to be true of things; we could go on attributing them to things as usual; what we would be cutting out in 'true of' is merely a special locution for talking about the attribution of the adjectives to the things.

This special locution, however, has its conveniences, and it would be missed. In fact, we do not have to do without it altogether. Speaking of adjectives as true or not true of things has made trouble in one special case, involving one special adjective, namely the phrase 'not true of self', in attribution to one special thing, namely that same phrase over again. If we forswear the use of the locution 'true of' in connection with this particular phrase in relation to itself as object, we thereby silence our antinomy and may go on blithely using the locution 'true of' in other cases as always, pending the discovery of further antinomies.

Actually related antinomies are still forthcoming. To inactivate the lot we have to cut a little deeper than our one case; we have to forswear the use of 'true of' not only in connection with 'not true of self' but also in connection with various other phrases relating to truth; and in such connections we have to forswear the use not only of 'true of' but also of various other truth locutions. First let us look at some of the antinomies that would otherwise threaten.

THE PARADOX OF EPIMENIDES

There is the ancient paradox of Epimenides the Cretan, who said that all Cretans were liars. If he spoke the truth, he was a liar. It seems that this paradox may have reached the ears of St. Paul and that he missed the point of it. He warned, in his epistle to Titus: "One of themselves, even a prophet of their own, said, The Cretans are always liars."

Actually the paradox of Epimenides is untidy; there are loopholes. Perhaps some Cretans were liars, notably Epimenides, and others were not; perhaps Epimenides was a liar who occasionally told the truth; either way it turns out that the contradiction vanishes. Something of paradox can be salvaged

with a little tinkering; but we do better to switch to a different and simpler rendering, also ancient, of the same idea. This is the *pseudomenon*, which runs simply: 'I am lying'. We can even drop the indirectness of a personal reference and speak directly of the sentence: 'This sentence is false'. Here we seem to have the irreducible essence of antinomy: a sentence that is true if and only if it is false.

In an effort to clear up this antinomy it has been protested that the phrase 'This sentence', so used, refers to nothing. This is claimed on the ground that you cannot get rid of the phrase by supplying a sentence that is referred to. For what sentence does the phrase refer to? The sentence 'This sentence is false'. If, accordingly, we supplant the phrase 'This sentence' by a quotation of the sentence referred to, we get: ' 'This sentence is false' is false'. But the whole outside sentence here attributes falsity no longer to itself but merely to something other than itself, thereby engendering no paradox.

If, however, in our perversity we are still bent on constructing a sentence that does attribute falsity unequivocally to itself, we can do so thus: ' 'Yields a falsehood when appended to its own quotation' yields a falsehood when appended to its own quotation'. This sentence specifies a string of nine words and says of this string that if you put it down twice, with quotation marks around the first of the two occurrences, the result is false. But that result is the very sentence that is doing the telling. The sentence is true if and only if it is false, and we have our antinomy.

This is a genuine antinomy, on a par with the one about 'heterological', or 'false of self', or 'not true of self', being true of itself. But whereas that earlier one turned on 'true of', through the construct 'not true of self', this new one turns merely on 'true', through the construct 'falsehood', or 'statement not true'. We can avoid both antinomies, and others related to them, by ceasing to use 'true of' and 'true' and their equivalents and derivatives, or at any rate ceasing to apply such truth locutions to adjectives or sentences that themselves contain such truth locutions.

This restriction can be relaxed somewhat by admitting a hierarchy of truth locutions, as suggested by the work of Bertrand Russell and Alfred Tarski. The expressions 'true', 'true of', 'false', and related ones can be used with numerical subscripts '0', '1', '2', and so on always attached or imagined; thus 'true_0', 'true_1', 'true_2', 'false_0', and so on. Then we can avoid the antimonies by taking care, when a truth locution T is applied to a sentence or other expression S, that the subscript on T is higher than any subscript inside S. Violations of this restriction would be treated as meaningless, or ungrammatical, rather than as true or false sentences. For instance, we could meaningfully ask whether the adjectives 'long' and 'short' are true_0 of themselves; the answers are respectively no and yes. But we could not meaningfully speak of the phrase 'not true_0 of self' as true_0 or false_0 of itself; we would have to ask whether it is true_1 or false_1 of itself, and this is a question that leads to no antinomy. Either way the question can be answered with a simple and unpenalized negative.

This point deserves to be restated: Whereas 'long' and 'short' are adjectives that can meaningfully be applied to themselves, falsely in the one case and truly in the other, on the other hand 'true_0 of self' and 'not true_0 of self' are adjectival phrases that cannot be applied to themselves meaningfully at all, truly or falsely. Therefore to the question 'Is 'true_0 of self' true_1 of itself?' the answer is no; the adjectival phrase 'true_0 of itself' is meaningless of itself rather than true_1 of itself.

Next let us consider, in terms of subscripts, the most perverse version of the *pseudomenon*. We have now, for meaningfulness, to insert subscripts on the two occurrences of the word 'falsehood', and in ascending order, thus: ''Yields a falsehood_0 when appended to its own quotation' yields a falsehood_1 when appended to its own quotation.' Thereupon paradox vanishes. This sentence is unequivocally false. What it tells us is that a certain described form of words is false_1, namely the form of words ''Yields a falsehood_0 when appended to its own quotation' yields a falsehood_0 when appended to its own quotation.' But in fact this form of words is not false_1; it is meaningless. So the preceding sentence, which said that this form of words was false_1, is false. It is false_2.

This may seem an extravagant way of eliminating antinomies. But it would be much more costly to drop the word 'true', and related locutions, once and for all. At an intermediate cost one could merely leave off applying such locutions to expressions

containing such locutions. Either method is less economical than this method of subscripts. The subscripts do enable us to apply truth locutions to expressions containing such locutions, although in a manner disconcertingly at variance with custom. Each resort is desperate; each is an artificial departure from natural and established usage. Such is the way of antinomies.

A veridical paradox packs a surprise, but the surprise quickly dissipates itself as we ponder the proof. A falsidical paradox packs a surprise, but it is seen as a false alarm when we solve the underlying fallacy. An antinomy, however, packs a surprise that can be accommodated by nothing less than a repudiation of part of our conceptual heritage.

Revision of a conceptual scheme is not unprecedented. It happens in a small way with each advance in science, and it happens in a big way with the big advances, such as the Copernican revolution and the shift from Newtonian mechanics to Einstein's theory of relativity. We can hope in time even to get used to the biggest such changes and to find the new schemes natural. There was a time when the doctrine that the earth revolves around the sun was called the Copernican paradox, even by the men who accepted it. And perhaps a time will come when truth locutions without implicit subscripts, or like safeguards, will really sound as nonsensical as the antinomies show them to be.

Conversely, the falsidical paradoxes of Zeno must have been, in his day, genuine antinomies. We in our latter-day smugness point to a fallacy: the notion that an infinite succession of intervals must add up to an infinite interval. But surely this was part and parcel of the conceptual scheme of Zeno's day. Our recognition of convergent series, in which an infinite number of segments add up to a finite segment, is from Zeno's vantage point an artificiality comparable to our new subscripts on truth locutions. Perhaps these subscripts will seem as natural to our descendants of A.D. 4000, granted the tenuous hypothesis of there being any, as the convergent series does to us. One man's antinomy is another man's falsidical paradox, give or take a couple of thousand years.

I have not, by the way, exhausted the store of latter-day antinomies. Another good one is attributed by Russell to a librarian named G. G. Berry. Here the theme is numbers and syllables. Ten has a one-syllable name. Seventy-seven has a five-syllable name. The seventh power of seven hundred seventy-seven has a name that, if we were to work it out, might run to 100 syllables or so; but this number can also be specified more briefly in other terms. I have just specified it in 15 syllables. We can be sure, however, that there are no end of numbers that resist all specification, by name or description, under 19 syllables. There is only a finite stock of syllables altogether, and hence only a finite number of names or phrases of less than 19 syllables, whereas there are an infinite number of positive integers. Very well, then; of those numbers not specifiable in less than 19 syllables, there must be a least. And here is our antinomy: the least number not specifiable in less than nineteen syllables is specifiable in 18 syllables. I have just so specified it.

This antinomy belongs to the same family as the antinomies that have gone before. For the key word of this antinomy, 'specifiable', is interdefinable with 'true of'. It is one more of the truth locutions that would take on subscripts under the Russell–Tarski plan. The least number not specifiable$_0$ in less than nineteen syllables is indeed specifiable$_1$ in 18 syllables, but it is not specifiable$_0$ in less than 19 syllables; for all I know it is not specifiable$_0$ in less than 23. This resolution of Berry's antinomy is the one that would come through automatically if we paraphrase 'specifiable' in terms of 'true of' and then subject 'true of' to the subscript treatment.

RUSSELL'S ANTINOMY

Not all antinomies belong to this family. The most celebrated of all antinomies, discovered by Russell in 1901, belongs outside this family. It has to do with self-membership of classes. Some classes are members of themselves; some are not. For example, the class of all classes that have more than five members clearly has more than five classes as members; therefore the class is a member of itself. On the other hand, the class of all men is not a member of itself, not being a man. What of the class of all classes that are not members of themselves? Since its members are the non-self-members, it qualifies as a member of itself if and only if it is not. It is and it is not: antinomy's by now familiar face.

Russell's antinomy bears a conspicuous analogy to Grelling's antinomy of 'not true of self', which in fact it antedates. But Russell's antinomy does not belong to the same family as the Epimenides antinomy and those of Berry and Grelling. By this I mean that Russell's antinomy cannot be blamed on any of the truth locutions, nor is it resolved by subjecting those locutions to subscripts. The crucial words in Russell's antinomy are 'class' and 'member', and neither of these is definable in terms of 'true', 'true of', or the like.

I said earlier that an antinomy establishes that some tacit and trusted pattern of reasoning must be made explicit and be henceforward avoided or revised. In the case of Russell's antinomy, the tacit and trusted pattern of reasoning that is found wanting is this: for any condition you can formulate, there is a class whose members are the things meeting the condition.

This principle is not easily given up. The almost invariable way of specifying a class is by stating a necessary and sufficient condition for belonging to it. When we have stated such a condition, we feel that we have "given" the class and can scarcely make sense of there not being such a class. The class may be empty, yes; but how could there not be such a class at all? What substance can be asked for it that the membership condition does not provide? Yet such exhortations avail us nothing in the face of the antinomy, which simply proves the principle untenable. It is a simple point of logic, once we look at it, that there is no class, empty or otherwise, that has as members precisely the classes that are not members of themselves. It would have to have itself as member if and only if it did not.

Russell's antinomy came as a shock to Gottlob Frege, founder of mathematical logic. In his *Grundgesetze der Arithmetik* Frege thought that he had secured the foundations of mathematics in the self-consistent laws of logic. He received a letter from Russell as the second volume of this work was on its way to press. "Arithmetic totters," Frege is said to have written in answer. An appendix that he added to the volume opens with the words: "A scientist can hardly encounter anything more undesirable than to have the foundation collapse just as the work is finished. I was put in this position by a letter from Bertrand Russell . . ."

In Russell's antinomy there is more than a hint of the paradox of the barber. The parallel is, in truth, exact. It was a simple point of logic that there was in no village a man who shaved all and only those men in the village who did not shave themselves; he would shave himself if and only if he did not. The barber paradox was a veridical paradox showing that there is no such barber. Why is Russell's antinomy then not a veridical paradox showing that there is no class whose members are all and only the non-self-members? Why does it count as an antinomy and the barber paradox not? The reason is that there has been in our habits of thought an overwhelming presumption of there being such a class but no presumption of there being such a barber. The barber paradox barely qualifies as paradox in that we are mildly surprised at being able to exclude the barber on purely logical grounds by reducing him to absurdity. Even this surprise ebbs as we review the argument; and anyway we had never positively believed in such a barber. Russell's paradox is a genuine antinomy because the principle of class existence that it compels us to give up is so fundamental. When in a future century the absurdity of that principle has become a commonplace, and some substitute principle has enjoyed long enough tenure to take on somewhat the air of common sense, perhaps we can begin to see Russell's paradox as no more than a veridical paradox, showing that there is no such class as that of the non-self-members. One man's antinomy can be another man's veridical paradox, and one man's veridical paradox can be another man's platitude.

Russell's antinomy made for a more serious crisis still than did Grelling's and Berry's and the one about Epimenides. For these strike at the semantics of truth and denotation, but Russell's strikes at the mathematics of classes. Classes are appealed to in an auxiliary way in most branches of mathematics, and increasingly so as passages of mathematical reasoning are made more explicit. The basic principle of classes that is tacitly used, at virtually every turn where classes are involved at all, is precisely the class-existence principle that is discredited by Russell's antinomy.

I spoke of Grelling's antinomy and Berry's and the Epimenides as all in a family, to which Russell's antinomy does not belong. For its part, Russell's antinomy has family connections of its own. In fact, it is the first of an infinite series of antinomies, as follows. Russell's antinomy shows that there is no class whose members are precisely the classes that are not members of themselves. Now there is a parallel antinomy that shows there is no class whose members are precisely the classes that are not members of members of themselves. Further, there is an antinomy that shows there is no class whose members are precisely the classes that are not members of members of members of themselves. And so on ad infinitum.

All these antinomies, and other related ones, can be inactivated by limiting the guilty principle of class existence in a very simple way. The principle is that for any membership condition you can formulate there is a class whose members are solely the things meeting the condition. We get Russell's antinomy and all the others of its series by taking the condition as non-membership in self, or non-membership in members of self, or the like. Each time the trouble comes of taking a membership condition that itself talks in turn of membership and non-membership. If we withhold our principle of class existence from cases where the membership condition mentions membership, Russell's antinomy and related ones are no longer forthcoming. This restriction on class existence is parallel to a restriction on the truth locutions that we contemplated for a while, before bringing in the subscripts; namely, not to apply the truth locutions to expressions containing any of the truth locutions.

Happily we can indeed withhold the principle of class existence from cases where the membership condition mentions membership, without unsettling those branches of mathematics that make only incidental use of classes. This is why it has been possible for most branches of mathematics to go on blithely using classes as auxiliary apparatus in spite of Russell's and related antinomies.

THE MATHEMATICS OF CLASSES

There is a particular branch of mathematics in which the central concern is with classes: general set theory. In this domain one deals expressly with classes of classes, classes of classes of classes, and so on, in ways that would be paralyzed by the restriction just now contemplated: withholding the principle of class existence from cases where the membership condition mentions membership. So one tries in general set theory to manage with milder restrictions.

General set theory is rich in paradox. Even the endless series of antinomies that I mentioned above, of which Russell's was the first, by no means exhausts this vein of paradox. General set theory is primarily occupied with infinity—infinite classes, infinite numbers—and so is involved in paradoxes of the infinite. A rather tame old paradox under this head is that you can exhaust the members of a whole class by correlating them with the members of a mere part of the class. For instance, you can correlate all the positive integers with the multiples of 10, thus: 1 with 10, 2 with 20, 3 with 30, and so on. Every positive integer gets disposed of; there are as many multiples of 10 as integers altogether. This is no antinomy but a veridical paradox. Among adepts in the field it even loses the air of paradox altogether, as is indeed the way of veridical paradox.

Georg Cantor, the nineteenth-century pioneer in general set theory and infinite arithmetic, proved that there are always more classes of things of a given kind than there are things of that kind; more classes of cows than cows. A distinct air of paradox suffuses his proof of this.

First note the definition of 'more'. What it means when one says there are more things of one kind than another is that every correlation of things of the one kind to things of the other fails to exhaust the things of the one kind. So what Cantor is proving is that no correlation of cow classes to cows accommodates all the cow classes. The proof is as follows. Suppose a correlation of cow classes to cows. It can be any arbitrary correlation; a cow may or may not belong to the class correlated with her. Now consider the cows, if any, that do not belong to the classes correlated with them. These cows themselves form a cow class, empty or not. And it is a cow class that is not correlated with any cow. If the class were so correlated, that cow would have to belong to the class if and only if she did not.

This argument is typical of the arguments in general set theory that would be sacrificed if we were to withhold the principle of class existence from cases where the membership condition mentions membership. The recalcitrant cow class that clinched the proof was specified by a membership condition that mentioned membership. The condition was non-membership in the correlated cow class.

But what I am more concerned to bring out, regarding the cow-

class argument, is its air of paradox. The argument makes its negative point in much the same way that the veridical barber paradox showed there to be no such barber, and in much the same way that Russell's antinomy showed there to be no class of all and only the non-self-members. So in Cantor's theorem—a theorem not only about cows and their classes but also about things of any sort and their classes—we see paradox, or something like it, seriously at work in the advancement of theory. His theorem establishes that for every class, even every infinite class, there is a larger class: the class of its subclasses.

So far, no antinomy. But now it is a short step to one. If for every class there is a larger class, what of the class of everything? Such is Cantor's antinomy. If you review the proof of Cantor's theorem in application directly to this disastrous example—speaking therefore not of cows but of everything—you will quickly see that Cantor's antinomy boils down, after all, to Russell's.

So the central problem in laying the foundations of general set theory is to inactivate Russell's antinomy and its suite. If such theorems as Cantor's are to be kept, the antinomies must be inactivated by milder restrictions than the total withholding of the principle of class existence from cases where the membership condition mentions membership. One tempting line is a scheme of subscripts analogous to the scheme used in avoiding the antinomies of truth and denotation. Something like this line was taken by Russell himself in 1908, under the name of the theory of logical types. A very different line was proposed in the same year by Ernst Zermelo, and further variations have been advanced in subsequent years.

All such foundations for general set theory have as their point of departure the counsel of the antinomies; namely, that a given condition, advanced as a necessary and sufficient condition of membership in some class, may or may not really have a class corresponding to it. So the various alternative foundations for general set theory differ from one another with respect to the membership conditions to which they do and do not guarantee corresponding classes. Non-self-membership is of course a condition to which none of the theories accord corresponding classes. The same holds true for the condition of not being a member of any own member; and for the conditions that give all the further antinomies of the series that began with Russell's; and for any membership condition that would give rise to any other antinomy, if we can spot it.

But we cannot simply withhold each antinomy-producing membership condition and assume classes corresponding to the rest. The trouble is that there are membership conditions corresponding to each of which, by itself, we can innocuously assume a class, and yet these classes together can yield a contradiction. We are driven to seeking optimum consistent combinations of existence assumptions, and consequently there is a great variety of proposals for the foundations of general set theory. Each proposed scheme is unnatural, because the natural scheme is the unrestricted one that the antinomies discredit; and each has advantages, in power or simplicity or in attractive consequences in special directions, that each of its rivals lacks.

I remarked earlier that the discovery of antinomy is a crisis in the evolution of thought. In general set theory the crisis began sixty years ago and is not yet over.

GÖDEL'S PROOF

Up to now the heroes or villains of this piece have been the antinomies. Other paradoxes have paled in comparison. Other paradoxes have been less startling to us, anyway, and more readily adjusted to. Other paradoxes have not precipitated sixty-year crises, at least not in our time. When any of them did in the past precipitate crises that durable (and surely the falsidical paradoxes of Zeno did), they themselves qualified as antinomies.

Let me, in closing, touch on a latter-day paradox that is by no means an antinomy but is strictly a veridical paradox, and yet is comparable to the antinomies in the pattern of its proof, in the surprisingness of the result and even in its capacity to precipitate a crisis. This is Gödel's proof of the incompletability of number theory.

What Kurt Gödel proved, in that great paper of 1931, was that no deductive system, with axioms however arbitrary, is capable of embracing among its theorems all the truths of the elementary arithmetic of positive integers unless it discredits itself by letting

slip some of the falsehoods too. Gödel showed how, for any given deductive system, he could construct a sentence of elementary number theory that would be true if and only if not provable in that system. Every such system is therefore either incomplete, in that it misses a relevant truth, or else bankrupt, in that it proves a falsehood.

Gödel's proof may conveniently be related to the Epimenides paradox or the *pseudomenon* in the 'yields a falsehood' version. For 'falsehood' read 'non-theorem', thus: ''Yields a non-theorem when appended to its own quotation' yields a non-theorem when appended to its own quotation'.

This statement no longer presents an antinomy, because it no longer says of itself that it is false. What it does say of itself is that it is not a theorem (of some deductive theory that I have not yet specified). If it is true, here is one truth that that deductive theory, whatever it is, fails to include as a theorem. If the statement is false, it is a theorem, in which event that deductive theory has a false theorem and so is discredited.

What Gödel proceeds to do, in getting his proof of the incompletability of number theory, is the following. He shows how the sort of talk that occurs in the above statement—talk of non-theoremhood and of appending things to quotations—can be mirrored systematically in arithmetical talk of integers. In this way, with much ingenuity, he gets a sentence purely in the arithmetical vocabulary of number theory that inherits that crucial property of being true if and only if not a theorem of number theory. And Gödel's trick works for any deductive system we may choose as defining 'theorem of number theory'.

Gödel's discovery is not an antinomy but a veridical paradox. That there can be no sound and complete deductive systematization of elementary number theory, much less of pure mathematics generally, is true. It is decidedly paradoxical, in the sense that it upsets crucial preconceptions. We used to think that mathematical truth consisted in provability.

Like any veridical paradox, this is one we can get used to, thereby gradually sapping its quality of paradox. But this one takes some sapping. And mathematical logicians are at it, most assiduously. Gödel's result started a trend of research that has grown in thirty years to the proportions of a big and busy branch of mathematics sometimes called proof theory, having to do with recursive functions and related matters, and embracing indeed a general abstract theory of machine computation. Of all the ways of paradoxes, perhaps the quaintest is their capacity on occasion to turn out to be so very much less frivolous than they look.