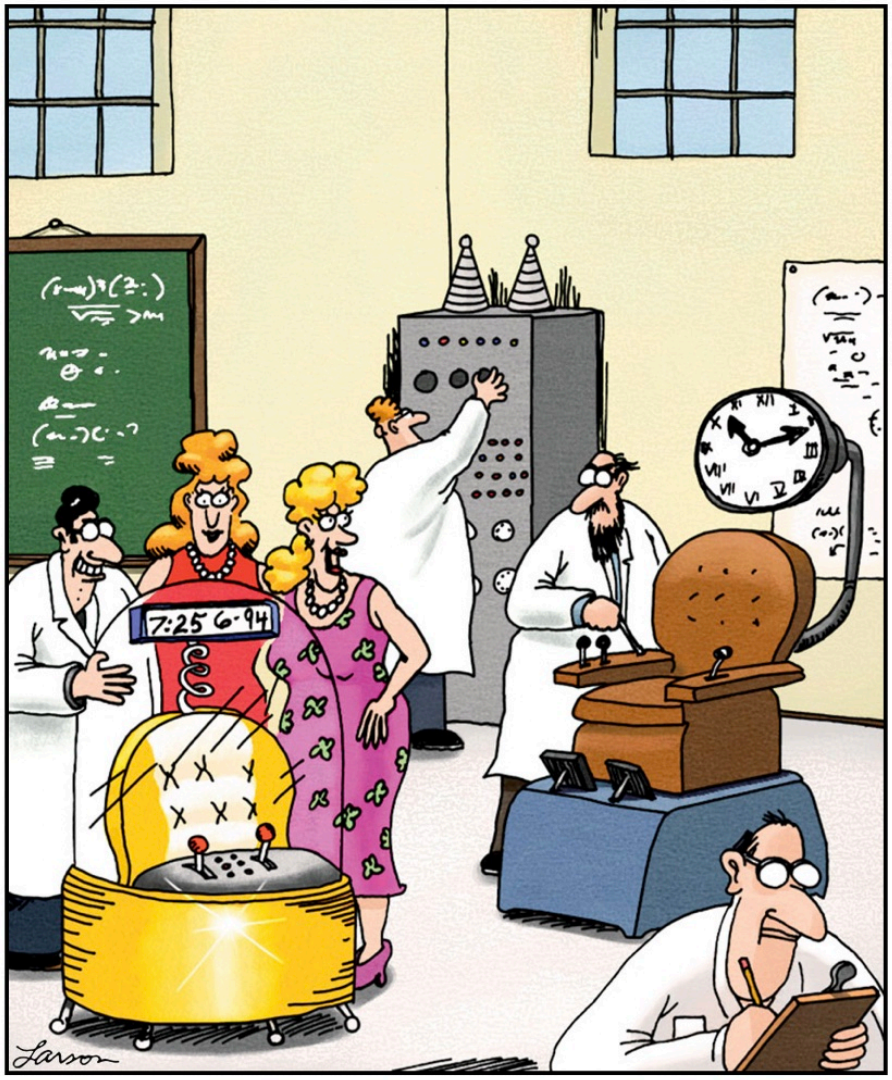
III.

The nature of time (1973)



"Oh, Professor DeWitt! Have you seen Professor *Weinberg's* time machine? ... It's digital!"

> Quid est ergo tempus? si nemo ex me quaerat, scio; si quaerenti explicare velim, nescio.[1]
>
> — Augustine: *Confessions*, XI.14
>
> Well I don't fuck much with the past but I fuck plenty with the future.
>
> — Patti Smith: "Babelogue".

— 111 —

Everything here derives from an old notebook entry which went, roughly: "'Information can only propagate forward in time'. — Why? — Because that is the *definition* of time."

Though admittedly it took me a while to understand why that isn't circular.

But the question about direction isn't why time goes one way and not the other; that is conventional. The question is why it doesn't go both ways at once. — Or sideways. — Or anyway you like.

Indeed, why does it *go*?

{…}

At first glance it all seemed tautological: the past is what you know about, the future is what you don't know about. — This is how you

---

[1] In the Loeb Classical Library edition — based (as the editor W.H.D. Rouse explains in a preface) on the rather verbose 1631 translation of William Watts — this becomes "What is time then? If nobody asks me, I know, but if I were desirous to explain it to one who should ask me, plainly I know not." Not exactly concise.

dispose of Aristotle's problem of the sea-battle, for instance: you think in terms of a set of propositions and the set of *partial* valuations upon them; these have a natural partial order defined by inclusion/extension/consistency; the sea-battle on the morrow is *after* the question today because the domain of the valuation that decides it (properly) *includes* the domain of the valuation of propositions determined to date, and the truth-values they assign are the same on the intersection of the domains. — All considerations about causality, etc., are interpreted within this essential picture.

— Or, well, something like that. — Once I'd thought it through this far I found out Kripke had worked it all out in detail in his semantics for intuitionistic logic — the logic of Becoming and not Being: stages of knowledge correspond to domains of valuations, hereditary sets form a Heyting algebra, etc.[2] — Very simple, very elegant; almost completely done.

{…}

As for why it *goes*, that is because the experience of the conscious observer is ordered by the consecutive acquisition of states of knowledge.

It seemed obvious that the essential perspective was four-dimensional,[3] and that the ideas of progression, the passing of events, the flow of time, etc., only made sense when referred to some world-line[4] along which memory accumulates, and that the past was

---

[2] See (for instance) Section 8.4 of Robert Goldblatt, *Topoi*. [Amsterdam: North-Holland, 1979.] This comes at it from the opposite direction, stages of knowledge are indexed by time rather than the reverse, but it is the equivalence that is significant.

[3] The number four has no particular significance. It usually suffices to consider 1 + 1 dimensions. The only interesting question is whether in n = p + q dimensions, the number of temporal dimensions q can be greater than 1.

[4] One refers to a "spacelike hypersurface" in theoretical physics, but this is overkill.

separated from the future by the assignment of truth-values to propositions.

But more than that, the relation of present to past was like that of metalanguage to language: one could not mix present with future because this would mean mixing propositions with valuations. — That each moment was *about* what preceded it. — That was why the paradoxes of time travel, getting into a time machine and going back to hoot your grandfather, or whatever, all sounded like the Cretan liar.

So temporality is logical. — Well. Complications ensue.

Being There (1973)

VIZZINI
Now, a clever man would put the poison into his own goblet, because he would know that only a great fool would reach for what he was given. I'm not a great fool, so I can clearly not choose the wine in front of you. But you must have known I was not a great fool; you would have counted on it, so I can clearly not choose the wine in front of me.

MAN IN BLACK
You've made your decision then?

VIZZINI
Not remotely. Because iocane comes from Australia, as everyone knows. And Australia is entirely peopled with criminals. And criminals are used to having people not trust them, as you are not trusted by me. So I can clearly not choose the wine in front of you.

MAN IN BLACK
Truly, you have a dizzying intellect.

VIZZINI
Wait till I get going! Where was I?

MAN IN BLACK
Australia.

— William Goldman: *The Princess Bride*.

Annals of futility, continued:

The *Scientific American* used to have a regular monthly column written by Martin Gardner, a famous guy back in the day, which posed mathematical puzzles for its readers. I picked it up shortly after I got back to Colorado and found a description of what later gained notoriety as Newcomb's Paradox. Gardner had found this in a paper written by Robert Nozick (then a Famous Professor of Philosophy at Harvard, which caught my eye), and passed it on to his own wider public for discussion. I solved it on inspection[5] and forget the matter for a week or two, when I realized that though there was no point in writing Gardner a letter, since as always he would get hundreds, I might try writing Nozick himself. So I typed up a few pages — honestly, less cryptic than usual — stuffed them into an envelope, committed them to the mails, and settled in to await the telegram announcing my accession to the Harvard faculty.

Of course what transpired was nothing of the kind. Gardner, who ironically[6] seemed to have read my mind, announced that he had received so many responses that he turned around and sent *all* of them to Nozick, meaning that my clever attempt to receive individual

---

[5] A rarity. Usually I stare at a problem without comprehension, forget about it, and the solution pops into my head two years later when I am taking a shower.

[6] See below.

attention would now be buried under an avalanche of bullshit. Thus naturally I got no reply, and when months later Nozick's summary of the proposals he had received appeared in Gardner's column, again as always opinions divided neatly between the two antithetical positions I had carefully explained were both wrong. I doubt he ever read my letter, and if he did, obviously he didn't understand it. — Otherwise, I realized later, there was no guarantee he wouldn't have published it as his own work.

I suppose a sensible person would have written a real paper about this and submitted it to a journal. But then a sensible person would have had access to a university library that got the volume containing the original paper (the essential reference) sooner than ten years later, would have had enough money to promote his manuscript from the slush pile, and would have been able to delude himself this was more than a silly puzzle only worth attending to on the off chance a Famous Professor would take notice of him.

I have, at any rate, the vague impression that there is now a literature on this subject, and that it is completely worthless. — Though really, who gives a shit. — But (modulo a couple of afterthoughts) what I said in the letter was this:

{…}

Suppose a game involving two players, yourself and a mysterious Being, and a pair of boxes, one of which you can see into, one which you cannot. The Being moves first, and puts a thousand dollars into the transparent box and either a million dollars or nothing into the second box. You then have the choice of taking both boxes, or the second box alone.

The twist here is that we suppose the Being can predict what you are going to do, and will punish greed. So if he[7] knows that you are going to take the second box alone, then and only then will it contain the million dollars; if on the other hand he's sure you are going to take both boxes, the second is empty.

Thus there is a payoff matrix which looks something like this:

Being predicts you will take both boxes	$0	$1000
Being predicts you will take the second box only	$1,000,000	$1,001,000
	You take the second box only	You take both boxes

So what should you do?

On the face of it, the arguments are these:

The Being is infallible, and knows what you will do. If you choose to take both boxes, this will have been foreseen; the second box will be empty, and your payoff will be a thousand dollars. On the other hand if you aren't greedy and choose the second box alone, it will contain the million. Obviously the sensible thing is to take just the second box.

---

[7] I take it for granted that an asshole who thinks he knows everything and is trying to hose you out of a million bucks would have to be male.

On the other hand the Being moved first, and the money is in the box, or it is not. Whatever he did you will make more by taking both boxes than by taking one. To suppose otherwise is to believe that you can change something that has already happened by occult influence, which is absurd. In effect you are saying that the contents of the box are not determined until you make your decision — that they do not yet lie, as it were, in your back light cone. But that too is ridiculous, because you can imagine that some friend of yours has already looked[8] for you. You can picture him staring at the million bucks and sending you urgent telepathic signals: "Take both.....Take both.... ."[9]

But neither addresses the real question, which is: who is the second player?

You are supposed to picture, i.e., someone like the mysterious stranger in *Last Year at Marienbad* who baffles everyone by always winning at the game of Nim[10] — you imagine an enigmatic smile, a mocking glance which says, I'm looking through you — this is some entity[11] who has read your source code, knows your Gödel sentence, has looked up the serial number on the back of your eyeballs, possesses

---

[8] Actually though this argument is superficially plausible it's also bullshit; the situation is like Schrödinger's Cat, not the Wheel of Fortune (a not-quite-paradoxical conundrum so universally known and discussed that it is explained, e.g., by Kevin Spacey in the movie *21* [Robert Luketic, 2008]). — The cat may know whether it's alive or dead, but I don't acquire the information until I open the box and look; in effect the determination still lies in my future. Same here.

[9] Nozick made various attempts to sharpen these arguments, none convincing and at least one based on an elementary blunder involving Bayes' theorem, but of course I didn't see them for another decade. — In any case everything he said is irrelevant or simply annoying. I'm just telling you a story here about my own folly.

[10] A solvable game with a known winning strategy; as was explained, of course, by Gardner, in another column.

[11] Wolfe [*The Right Stuff*]: "the anonymous and uncanny Chief Designer, D-503, Builder of the Integral….. — He computes the future! the mighty Integral!"

some sort of X-ray vision that allows him to see into the black box housing the freedom of the will.

But this is a trick and a con, the diversion that misdirects your eye from the shell that hides the little pea. If this were what he was doing, there wouldn't be a problem. If you were choosing on a whim, or an irrational hunch, or flipping a coin, or throwing the I Ching to decide what to do, there wouldn't be a puzzle. Maybe he could guess your choice in advance, but this would be no more problematic than one of those computers that fits in your shoe and predicts where the roulette wheel is going to stop. That wouldn't be a paradox.

No. There's only a paradox when you try to make *the rational choice*. You are trying to decide what you *should* do.

So what is the Being's problem then? What does he have to be able to predict?

Exactly the same thing: what is the *rational* choice?

So both you and the "mysterious Being" are trying to solve the same problem. The black box is transparent.

And the paradox is essentially the same as with the Cretan liar, i.e., self-reference: what the Being is going to have done (invent tenses as necessary) depends on what you are going to do. So what you are going to do depends on what you are going to do.

You know what the Being is going to do: the rules have explained it. The Being knows what you are going to do: you are going to try to maximize your payoff. Nothing is hidden.

{…}

Why was that so obvious? I had the following example[12] in the back of my mind:

Suppose that the physical world is classical and deterministic and you have a computer (for obvious reasons this could be called a Laplacian machine) that can predict the evolution of any system from its initial conditions by solving the differential equations — or whatever — in a fixed amount of time which can be estimated beforehand.

[— Well. — That isn't it exactly. — There are two distinct ideas here: first, that the system can be *modeled* by the computer; second, that for the given evolution, the computation modeling the system gets done before the system does; as it were, that the one computation is *faster* than the other. That makes the computation a *prediction*.]

There would be many questions about how precisely the initial state would have to be measured, whether or not you might have to employ a machine that could compute with real numbers and not floating-point approximations to them (Smale later worked out such a theory), etc., but ignore those for the moment. — Suffice it that it makes perfect predictions *about* the world, *within* the world.

Then suppose you ask the machine to tell you whether a light bulb is going to be on at the end of the computation. And then plug the output of the machine into the power switch for the bulb, so that if the machine outputs "on", it turns it off, and if it says "off", it turns it on.

---

[12] I think this is due to John Kemeny. See (perhaps) *A Philosopher Looks At Science*, though I can't find a copy of the book with which to verify the reference.

What this demonstrates, obviously, is that even if complete predictions were possible, they would be self-defeating if allowed to feed back into the system; for essentially the same reasons that language must be segregated from metalanguage. The machine whose output negates its own prediction presents the same problem as the attempt to assign valuations to the statements on a card which read "The statement on the other side of this card is true" and "The statement on the other side of this card is false."

{…}

It should be obvious, incidentally, that prediction is essentially computation.

This follows, really, from Church's thesis: an algorithm must be employed to make a prediction; any algorithm can be realized as a computation by a Turing machine.

Successive approximations can be realized by providing the answers to a series of binary questions. — There is nothing deep here, in practice it is straightforward.[13]

---

[13] Here elided is a lengthy digression in the original manuscript on the question of successive approximation, i.e. whether improvements in precision must generally be efficacious. The natural way to formulate that was in the familiar style, for every epsilon to which you wish a numerical prediction to be accurate there must exist a delta within which the initial conditions should be specified, etc., and that raised the embarrassing possibility that in the general case dynamical systems could amplify small errors in precision and erase the possibility of prediction entirely. — This was already obvious for systems with even a modest number of degrees of freedom, see the ergodic theorems of statistical mechanics, but it seemed a novel idea that it might hold as well for relatively simple systems. — Later, of course, this became known as the butterfly effect, and it would have annoyed me not to have worked it out in greater detail had it not become apparent that Poincaré had beaten everyone to publication before the turn of the century. — I did, however, include this analysis in a lengthy précis of the difficulties in the concept of prediction for a friend who was a graduate student in philosophy; he didn't understand it, but incorporated it in his paper nonetheless, and his instructor parroted all of it in a public lecture a few weeks later, without attribution either to him or (of course) to the ghostwriter, me. — I briefly considered beating the shit out of the guy, but then realized, as usual: why bother. What's the use.

But the ease with which the unrestricted extension of the idea leads to paradox makes it seem very strange that we can build computers within the physical world. Something about that doesn't smell right. How is it possible?

Which raises the complementary question, how complex must a mechanical system be to allow the construction of a universal Turing machine? What is the simplest system that can realize one? Because the evolution of such a system would be recursively indeterminate. It would be impossible to predict.

So this would mean that, even within an apparently deterministic physics, there would be elementary dynamical questions that would be effectively undecidable. There would be mechanical systems instantiating the halting problem.

And then: what is the relation to the question of "exact solvability"? "Integrability"? Can something as simple as the classical problem of three bodies be undecidable in this sense?

{…}

You also see that, in general, time travel paradoxes are essentially the same as paradoxes of self-reference and paradoxes of prediction. The ability to see the future is equivalent to the ability to send messages into the past. You don't need to imagine anything as grisly as physically traveling back in time to shoot your grandfather; information transfer is sufficient to generate paradox. The Being may be able to predict what you will do, or the Being may have a tachyonic

telephone[14] with which he can call himself in the past the moment after you have made the choice, it makes no difference.[15]

So that's the story at first glance: self-reference should be forbidden; the game and the Being are, therefore, impossible.

{…}

Indeed it is a mistake to assume the proposition "There is money in the first box" *has* a truth-value; that its contents have determinate value; that it *has* contents.

{…}

At second glance it's a trifle more interesting.

The Being predicts you'll take one box [1] or both [2].

If [1], then the second box contains a million, thus taking both boxes yields a million plus a thousand, thus that is the optimal choice, thus the correct prediction is [2].

If on the other hand he predicts [2], then there's nothing in the second box, but you still gain more by taking both. Thus he should still

---

[14] Tachyons are hypothetical particles which travel faster than light, invented by bored theoreticians to entertain themselves by bullshitting their way out of paradoxes. Absent *ad hoc* baroque complication, anything that travels faster than light can, in the special theory of relativity, be turned by Lorentz transformation into something travelling backward in time; thus permitting the communication with the past of useful information like where the markets will close and which way to swerve to avoid an oncoming bus. A tachyonic telephone is, accordingly, a useful shorthand for precognition on demand.

[15] Once again: (nonrelativistically) the past is what is known; the future is what isn't. If the Being *knows* what you will do, your future lies in his past. (Relativistically the past light cone is what you know about, the future light cone is what will know about you, and the rest is elsewhere, causally disconnected from here and now.)

predict [2] — though: what happened to the million dollars? — and the system has, as it were, an attractor.

Well. — We might believe that for a moment. But consider this: you and the Being have essentially the same problem. This means that the Being, too, is trying to make the choice that optimizes your payoff. Therefore when the Being analyzes the payoff matrix, by the same argument that leads you to select the second column, he must select the second row; thus to maximize your payoff, he's compelled to make the wrong prediction, and say that you will take the first box only. So the inconsistency, or metainconsistency, seems intrinsic after all.

So from this point of view the problem is that the payoff matrix should be symmetric with respect to transposition; the fact that it is not is, then, the root of the confusion. — This isn't consistent with the story we have been telling about taking one box versus taking both, but we can always make up another story. At any rate the matrix should look like this:

Being's choice 2	y	z
Being's choice 1	x	y
	Your choice 1	Your choice 2

where if $x < y < z$ or $x > y > z$ there's a self-consistent strategy, whereas if $x, z < y$ or $y < x, z$ there is not.

I don't know that I take this argument seriously, but it's no dumber than what we started with.

{…}

Superficially it might seem that the paradox might be tamed by fuzzier logic, but introducing probabilities makes no difference: if e.g. to evaluate your optimal choice you assign p[1], p[2] as the probabilities the Being will make those choices, you then observe that

$$p_1 = prob(\eta_{11}p_1 + \eta_{12}p_2 > \eta_{21}p_1 + \eta_{22}p_2)$$

which means $p_1 = 0$ or $p_1 = 1$ (since the inequality is true or false)

but if $p_1 = 1$ then $p_2 = 0$ and

$$p_1 = prob(1000000 > 1001000) = 0$$

so $p_1 = 1$ implies $p_1 = 0$.

Similarly $p_1 = 0$ implies $p_1 = 0$, etc., so the argument is identical, and we are driven to the fixed point $p_1 = 0$, $p_2 = 1$.

{…}

It isn't difficult to translate the logic of the situation into a computer program. Any language that permits recursive definition will do; in Lisp, e.g., taking the expected payoff as a function of choice, and taking another function with no arguments to represent the prediction, the relevant definitions are:

```
(defun payoff (choice)
   (cond
      ((eq choice 'one)  (if (eq (prediction) 'one) 1000000 0))
      ((eq choice 'both) (if (eq (prediction) 'one) 1001000 1000))
   (T nil)))

(defun prediction ()
   (cond
      ((> (payoff 'one)  (payoff 'both)) 'one)
      ((> (payoff 'both) (payoff 'one))  'both)
   (T nil)))
```

(You may read the first as "if the choice is one box, if the prediction was one box then the payoff is one million, else it is zero; if the choice was both boxes, if the prediction was one box the payoff is one million plus one thousand, else it is one thousand; these are all the possibilities," and the second as "if the payoff for choosing one box is greater than the payoff for choosing both boxes, predict 'one'," etc.)

Naturally though these compile into working code (their mutual dependence does not in itself entail that they are ill-defined), if you try evaluating either function the result is a stack overflow, i.e. the computational equivalent of smoke pouring out from under the hood[16] or a loud feedback squawk.

But there's an ambiguity here as well, related to the distinction in programming semantics between call-by-name and call-by-value. — Which is actually much more complicated, there are a bewildering

---

[16] One of the earliest electromechanical logic machines was built by two students of Quine, William Burkhart and Theodore Kalin, in 1947, and solved problems in the propositional calculus by evaluating truth tables. "It is interesting to note," says Gardner [*Logic Machines and Diagrams*, New York: McGraw-Hill, 1958, p. 130] "that when certain types of paradoxes are fed to the Kalin-Burkhart machine it goes into an oscillating phase, switching rapidly back and forth from true to false. In a letter to Burkhart in 1947 Kalin described one such example and concluded, 'This may be a version of Russell's paradox. Anyway, it makes one hell of a racket.'"

variety of possible strategies for evaluation[17] — but: in evaluating an expression one may have the option of performing a syntactic transformation upon it first; for instance, some algebraic manipulation that may simplify it.

(The traditional [Leibnizian] interpretation of the derivative, as a quotient of infinitesimals, involves a kind of call-by-name strategy: you compute the ratio *before* allowing the values to go to zero. — Unsurprisingly, this procedure becomes difficult to analyze in cases involving nested series of limiting processes — the order does of course affect the result — and in this sense the subtleties first encountered in the formulation of the calculus prove symptomatic of deeper difficulties in the theory of computation.)

In general if a program terminates on all inputs the results are the same, but if it does not the order of execution can make a difference.[18]

Lisp functions usually employ call-by-value,[19] so that every subexpression is evaluated and the result is passed to the routine that calls it, but conditional expressions are an exception, and whether a function terminates or does not can depend on the order in which the tests are performed.

---

[17] See Harold Abelson and Gerald Jay Sussman, *Structure and Interpretation of Computer Programs*. [Cambridge: The MIT Press, 1996.] — The discussion that follows is drastically oversimplified; there may be no subject more complex than the semantics of programing languages.

[18] This is already true in the lambda calculus, see for instance section 5.8 of Joseph Stoy, *Denotational Semantics*. [Cambridge: MIT Press, 1977.] Whether a computation terminates depends, in general, on the strategy employed in the reduction of an expression.

[19] Evaluation can be turned off with the (metalinguistic) quote function (and turned back on within the scope of a quote with a backquote). These devices are useful, e.g., when writing programs that can rewrite their own text while they are running. — I suppose I should illustrate this by exhibiting a Lisp program which on execution erases its own text and thus can only produce an output if and only if it does not, but let's leave that as an exercise for the reader.

This is because an expression like (reverting to a more Pascal-like syntax)

    if ([Boolean] test) then A else B

is evaluated by first evaluating the test, and then evaluating A or B according to whether it returns true or false. Thus if one knew for some other reason that the test always returns true, the expression can be replaced with A; while otherwise B might be some function that fails to terminate.

E.g. one might define

    f(x) = if (x = 1) then 6 else f(2)

which depending on how x is defined elsewhere in the program will return 6 or a stack overflow.

What this means in the case at hand is that the functions given above might be defined in some other way to avoid the ungrounded recursion. One might, e.g., try something like

```
(defparameter *payoff-matrix* '((1000000 1001000) (0 1000)))

(defun payoff (i j) (nth i (nth j *payoff-matrix*)))

(defun best-choice ()
   (cond
        ((and
           (> (payoff 0 0) (payoff 1 0))
           (> (payoff 0 1) (payoff 1 1)))
           'one)
        ((and
           (> (payoff 1 0) (payoff 0 0))
           (> (payoff 1 1) (payoff 0 1)))
           'both)
        (T nil)))

(defun players-move () (best-choice))

(defun beings-move () (best-choice))
```

which produces the (unconvincing) result

```
7 > (players-move)
BOTH
7 > (beings-move)
BOTH
```

{…}

Let's amend the rules slightly to construct a more consistent paradox: say that a clause in the contract specifies that if you take the second box only and the Being predicts you will take both, you can sue for breach of infallibility. Then Mister Mastermind will have to settle out of court for, say, $100,000, and the revised matrix reads:

Being's choice 2	$100,000	$1000
Being's choice 1	$1,000,000	$1,001,000
	Your choice 1	Your choice 2

This eliminates the bogus attractor, and reduces the paradox to the pure Cretan form: if you *should* take one box, you *should* take two boxes; if you *should* take two boxes, you *should* take one box. — Moreover if you start at (1,1) and alternate moves with the Being, you proceed through every node in the matrix: (1,1) entails (2,1) entails (2,2) entails (2,1) entails (1,1). — Surely that's more like it.

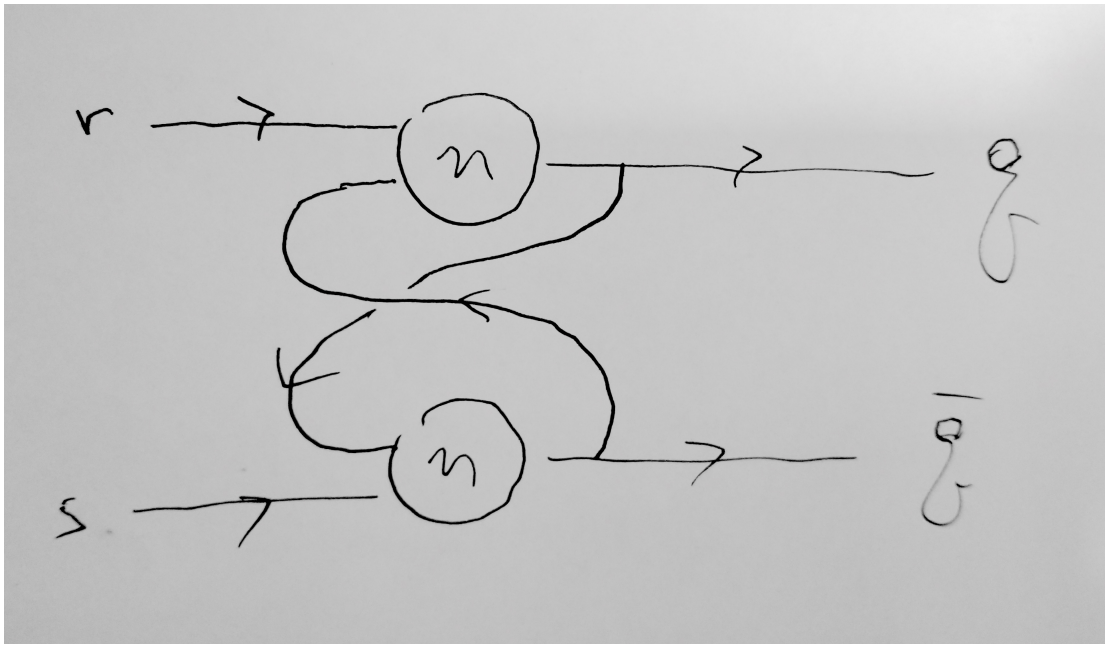
Taking p = "*should* take one box" and q = "*will* take one box" — thus not p = "*should* take both boxes", not q = "*will* take both boxes" — the sequential logic of the situation can be diagrammed as follows:

p	q	p'	q'
F	F	T	F
F	T	F	F
T	F	T	T
T	T	F	T

{…}

Apparently, then, for simple problems anyway, a program of elementary complexity suffices, the logic of the situation can be modeled by a Boolean circuit, and the self-referential part of it, the part that seems to require the tachyonic telephone, is expressed by feeding back the outputs into the inputs.

In other words another simple kind of temporal paradox might be expressed by a circuit such as the following:



where $\eta(r, s)$ is the NAND function):

r	s	NAND(r,s)

F	F	T
F	T	T
T	F	T
T	T	F

which in Lisp is:

```
(defun nand (p q) (not (and p q)))
```

This circuit is the famous flip-flop.[20] Far from being paradoxical, it is an extremely useful electronic component,[21] because its stable states can be used to store information.

If this is coded recursively as

```
(defun top (r s) (nand r (bottom r s)))

(defun bottom (r s) (nand s (top r s)))
```

the result, unsurprisingly, is a stack overflow, but if you observe that a false input to a NAND gate always produces a true output, then the (syntactically)[22] equivalent definitions

---

[20] One of them, anyway. There are many variations on the theme.

[21] It was used as a storage device as early as the codebreaking Colossus of 1943.

[22] I.e., insinuating a call-by-name strategy.

```
(defun flip-flop-top (r s)
   (if (not r) t (nand r (flip-flop-bottom r s))))

(defun flip-flop-bottom (r s) (flip-flop-top s r))

(defun flip-flop (r s)
   (list (flip-flop-top r s)
         (flip-flop-bottom r s)))
```

terminate on inputs (F,F), (T,F), and (F,T).

The behavior of the circuit is summarized by the truth table:

r	s	q	q'
F	F	T	T
F	T	T	F
T	F	F	T
T	T	q	q'

which can be interpreted as follows: the values (q, q') are assumed given (grounding the recursion) and are to be maintained as complements; thus the input (F, F) is forbidden. The inputs (F, T) and (T, F) flip the values of (q, q') — thus the name. The input (T, T) produces a well-defined result if (q, q') are given, and leaves them unchanged.

In other words what seems an impenetrable conundrum to the philosopher is a trivial commonplace for the electrical engineer. I suppose this should be embarrassing.

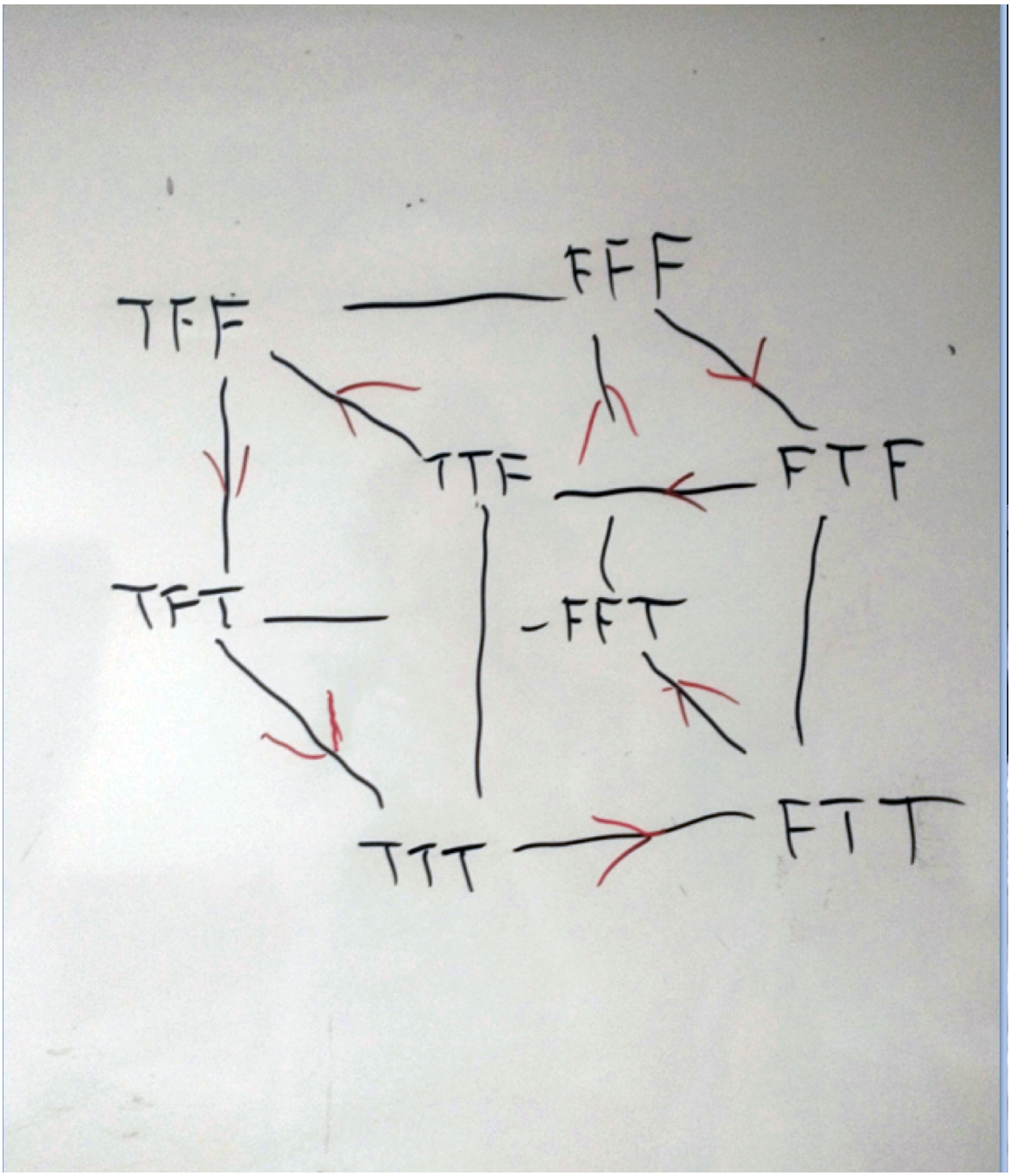
{…}

Why stop with two players? why not a game with three? Again we suppose the Player, the infallible Being, and as a third party introduce — not a Cartesian Demon, exactly — a Prankster, let us say, who may as well be female, who can intervene in the game as follows: she has the power (say by hacking into their computers)[23] to reverse the Being's perception of what move the Player has made/will make, and when she does so she also reverses the Player's judgment as to which payoff is greater than the other (switches "<" to ">" and vice-versa); since it pisses her off when the Being tries to weasel out of forking over the million bucks, she only flips this switch when he predicts the Player will take both boxes and intends to pocket the money himself — but then doesn't bother flipping the switch back until there is another change from Being-predicts-one to Being predicts-two.

Interpreting this sequentially, and supposing the Player, the Being, and the Prankster takes turns in that order, the following state transition diagram results:

---

[23] Since the point of this entire discussion is that the players can all be replaced by computers, there is no loss of generality.



i.e.

> TTT —> FTT —> FFT —> FFF —> FTF —> TTF —> TFF —> TFT —> TTT

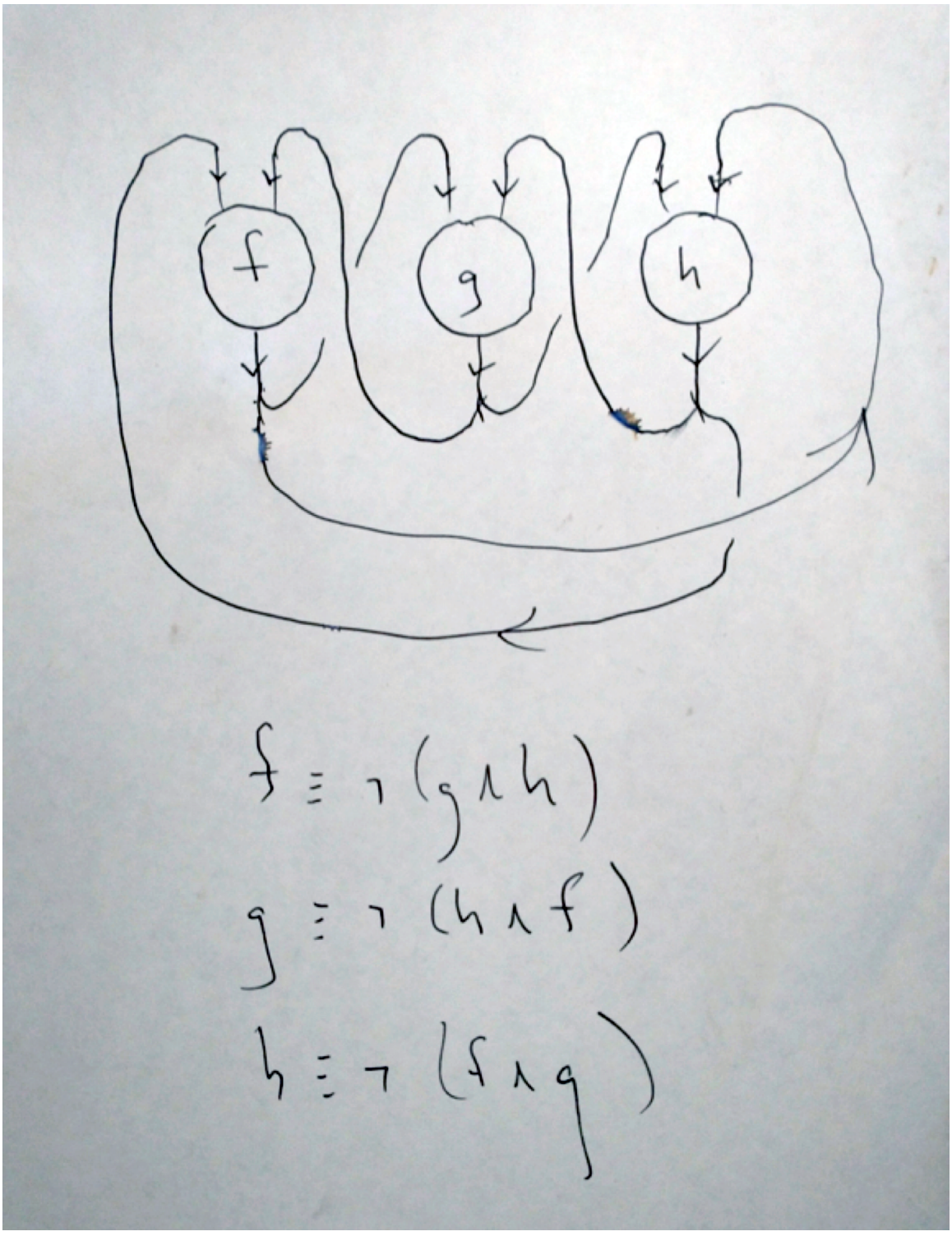
with the truth table:

x	y	z	x´	y´	z´
F	F	F	F	T	F
F	F	T	F	F	F
F	T	F	T	T	F
F	T	T	F	F	T
T	F	F	T	F	T
T	F	T	T	T	T
T	T	F	T	F	F
T	T	T	F	T	T

where x = “Player does/does not take one box”, y = “Being does/does not predict Player takes one box”, and z = “Prankster does/does not confuse the perceptions of the other two”. So here there are no fixed points and only one cycle.

{...}

A better illustration of the general case (make up your own story) is provided by the Boolean circuit:



where f, g, h are all NAND gates.

The mapping of inputs to outputs this defines is

x	y	z	x´	y´	z´
F	F	F	T	T	T
F	F	T	T	T	T
F	T	F	T	T	T
F	T	T	F	T	T
T	F	F	T	T	T
T	F	T	T	F	T
T	T	F	T	T	F
T	T	T	F	F	F

which has the stable states

FTT —> FTT
TFT —> TFT
TTF —> TTF

while

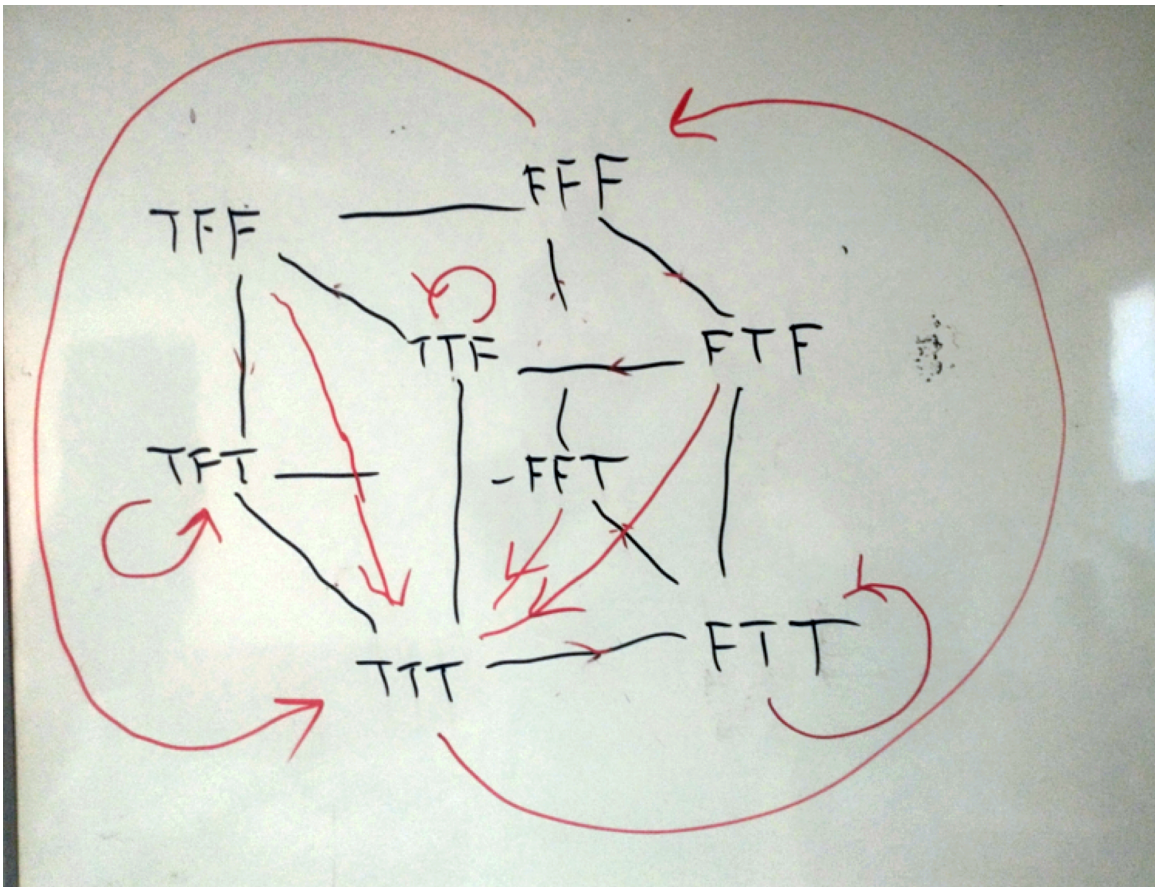
{FFT, FTF, TFF} —> TTT

and

TTT —> FFF —> TTT —> …

is a cycle.

This may be summarized by the state-transition diagram:



There are, in other words, three fixed points — representing, if you will, the self-consistent world histories in which nobody shoots his grandfather — and three states driven to an attractor, a fundamental cycle that flips back and forth between all on and all off.

(Note incidentally that oscillating behavior may be exactly what you're trying to produce with the feedback loop; in fact the original use of the flip-flop was as a circuit to produce square waves, thus the alternative designator "multivibrator".)

The generalizations to any number of players and arbitrary directed graphs with Boolean functions at the nodes[24] is straightforward. The

---

[24] Waving my hands here. A bit of care in the definitions is required.

arrangement of fixed points and cycles is largely arbitrary, but it is obvious that, for any finite network, any evolution must terminate in either a fixed point or a cycle.[25] So in this simplified model of the history of the world, at least, some form of eternal recurrence is guaranteed.

{…}

A related application of the idea of a Boolean circuit whose outputs feed back into its inputs is that of the genetic regulatory network: groups of related genes are governed by sets of rules of the form "A is expressed if B is expressed and C is not expressed or …"; Stuart Kaufmann conjectures that the stable states of such networks (self-consistent sets of choices) correspond to cell types, has shown empirically that the number of such fixed points in random networks is proportional to the square root of the number of nodes, and presents evidence that the number of cell types (which would correspond to given sets of genes being turned on and off) is correlated to the size of the genome in roughly this fashion for a variety of species.[26]

{…}

If we think of these networks as dynamical systems — we might extend the Boolean model by making the functions probabilistic, for instance — questions about the stability of equilibria become significant: how do they behave under perturbations? What is the expected lifetime of a (quasi)stable state? — etc., etc.

---

[25] Similar theorems hold for continuous dynamical systems, given the appropriate topological preconditions (some form of compactness). — The situation for infinite discrete Boolean networks is more complicated, containing as it does the case of the two-state cellular automata studied by Wolfram among others, and can entail difficulties like the halting problem for Turing machines.

[26] Kauffman, Stuart. *The Origins of Order*. Oxford: Oxford University Press, 1993.

But in any case the problem of the temporal loop has been reduced to the problem of feedback; meaning the idea isn't as absurd as it first sounds, though (as always) complications appear on a closer analysis.

{…}

Irwin translates all this into postmodernism:

> As Johnson sees it, taking a position on the numerical structure of the tale means, for Lacan and Derrida, taking a numerical position, choosing a number, but that means playing the game of even and odd, the game of trying to be one up on a specular, antithetical double. And playing that game means endlessly repeating the structure of "The Purloined Letter" in which being one up inevitably leads to being one down. For if the structure created by the repeated scenes in the tale involves doubling the thought processes of one's opponent in order to use his own methods against him—as Dupin does with the Minister, as Derrida does with Lacan, and as Johnson does with Derrida—then the very method by which one outwits one's opponent, by which one comes out one up on him, is the same method that will be employed against oneself by the next player in the game, the next interpreter in the series, in order to leave the preceding interpreter one down.[27]

Admittedly cute. But I think Vizzini said it better.

---

[27] John T. Irwin, "Mysteries We Read, Mysteries of Rereading: Poe, Borges, and the Analytic Detective Story." *MLN* Vol. 101, No. 5, Comparative Literature (Dec. 1986), pp 1168-1215.

Direction

There is an apparent paradox in the seeming contradiction between microscopic reversibility and macroscopic irreversibility: the laws of physics, at an elementary level, are time-symmetric; they look the same one way or the other. This means any instantaneous snapshot[28] of the system could belong to a movie that runs forward or backward.

So what makes entropy increase?

The usual way of putting it is that for almost all configurations of positions and velocities,[29] in one direction or the other the snapshot will be more disordered; and that determines the direction of time.

But that doesn't quite work either, because it's obvious that for any snapshot in which the velocities are aligned in such a way as to make time go in one direction, from symmetry there's another exactly like it with the positions the same and the velocities reversed, in which time goes the other way.[30]

So we have to state the principle slightly differently: suppose we have two snapshots, at times a, b; then for systems with more than a few degrees of freedom, for almost all dynamical evolutions, the measures of disorder in the two snapshots will be different, and if b is more disordered than a, then if we take another snapshot at time c, the temporal order of the three snapshots will almost certainly be the same as the order of their measures of disorder.

---

[28] Of positions *and* velocities. The picture captures infinitesimal motion blur.

[29] It is easiest to picture a collection of atoms governed by classical mechanics, because the quantum-mechanical argument isn't any different.

[30] Penrose does make this argument. In fact I think at some point I invented it and fell for it myself. I think it was in the notes I gave you in 1972.

Precise mathematical statements of "almost all", "almost certainly", etc., are notoriously difficult, see the so-called ergodic theorems of statistical mechanics, which are perniciously subtle and perennially in the process of reformulation. But that's the gist of it.

Note also that since the definitions of these concepts involve taking limits as the number of degrees of freedom[31] goes to infinity, directionality is essentially an emergent property; in very simple systems it need not exist.

---

[31] Half the dimension of the phase space. For a system of mass points moving in three dimensions, this would be 3 times the number of particles — each would have three pairs of spatial coordinates and momentum coordinates (proportional to velocities). In nonrelativistic quantum mechanics this is the dimension of the configuration space on which the Schrödinger equation is defined.

127

A few preliminary notes

First, the lengthy preamble explaining the rudiments of the theory of Riemann surfaces is more than slightly ridiculous, since any mathematician would know all this already and anyone who is not probably wouldn't be able to understand it; I only included it to be able to try the argument out on a few friends who lie in the uncanny valley that separates these two classes. Results were not great, but better than expected.

Second, the case of analytic continuation is not perfectly analogous, since the boundary value problems for the Cauchy-Riemann equations and for field equations of hyperbolic type are different: given knowledge of the gravitational field, e.g., on a patch on the spacetime manifold, its values can be inferred only on a sort of top-shaped set, conical forward and backward in time. (I.e. for any point whose back or forward light cone lies within the given set.) When you're lucky, values everywhere can be inferred from information on a spacelike hypersurface. These technicalities aside, the moral of the story remains the same.

And third, though a later entry may make it look that way, I am ashamed to admit I didn't steal this idea from Thomas Pynchon.

This is just another case of a sketch that lay at the bottom of the trunk for decades until I pulled it out and polished it up for this occasion.

128

Time machines

> There in the flickering light of the lamp was the Machine, sure enough, squat, ugly, and askew, a thing of brass, ebony, ivory, and translucent, glimmering quartz. Solid to the touch — for I put out my hand and felt the rail of it — and with brown spots and smears upon the ivory, and bits of grass and moss upon the lower parts, and one rail bent awry.
>
> H.G. Wells: *The Time Machine*.

Of course, you *can* travel through time, but only in one direction. Wells' protagonist goes several hundred thousand years into the future,[32] while within the bubble his Machine creates around him only a few hours pass; relativity say this is perfectly possible, if you just run in circles fast enough. (Say, within one part in a quintillion of the speed of light.) The problem would be that once he got there, he couldn't come back; he couldn't return to tell the tale around the table to his dinner companions. It is curious how much the mystique of the Intrepid Explorer depends upon this possibility, but there it is: without the

[32] To 802,701 A.D., to be precise. One wonders how he pulled that number out of his ass.

gripping eye-witness account, the Future is a tree falling in the forest with no one here to watch.

{…}

Logical necessity prohibits travelling into the past, but doesn't preclude the possibility of observing it. It might not be possible to go back to the Late Mesozoic to hunt dinosaurs, as in the story of L. Sprague De Camp,[33] but one could in principle watch them on television. The difficulty there would be entropic: information sufficient to run the clock backwards would require the exertion of godlike powers,[34] something like keeping track of every molecule on the planet[35] but multiplied by many orders of magnitude; and though one can imagine a kind of simulacrum of the world of the past, some sort of computer simulation, almost all of it would be guesswork.

That said, there's always a loophole.

{…}

Gödel famously found a solution to Einstein's equations which describes a universe in which time travel is possible if you

---

[33] See *A Gun For Dinosaur, and Other Imaginative Tales*. [New York: Curtis Books, 1968.]

[34] And, given the Boltzmann relation, cosmological energies.

[35] And every photon any one of them had emitted.

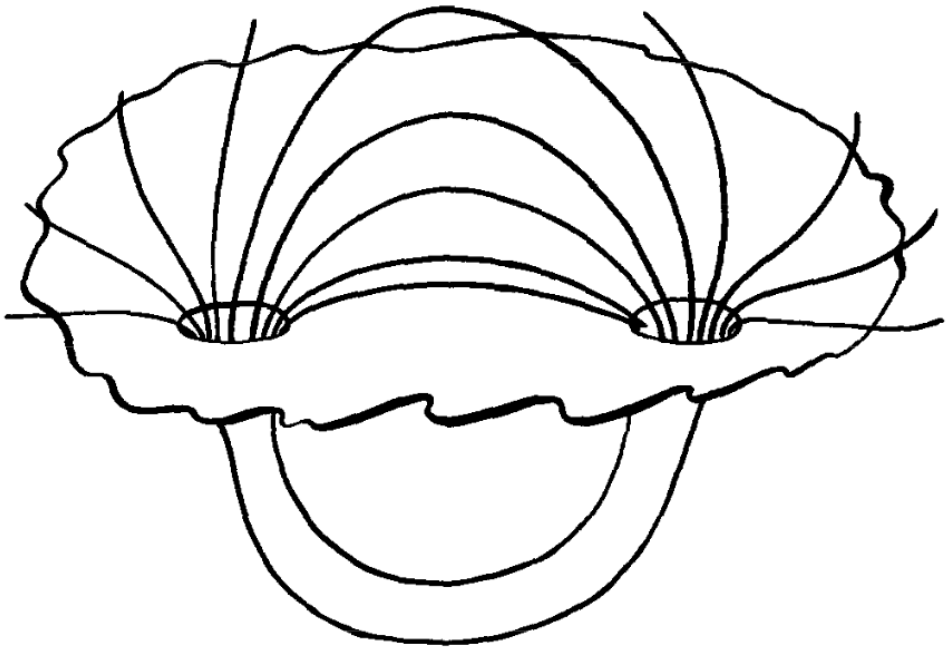
circumnavigate it you can end up in your own past.[36] His model isn't particularly "physical" it is extremely symmetric, in fact completely homogeneous in space and time but once the possibility has been demonstrated one must wonder whether other such Einsteinian spacetimes with more flexible architectures might exist; allowing solutions of Maxwell's equations, for instance, and thus the possibility of sending signals backward in time.

These are essentially mathematical questions, there are theorems about them, and it appears the existence of closed timelike paths causal loops requires the existence of negative energy densities. Classically these would not make sense, but quantum field theory doesn't make sense either, and there are circumstances in which they can arise in particular, in the explanation of the Casimir effect, which involves a negative energy density in the vacuum.

---

[36] "…by making a round trip on a rocket ship in a sufficiently wide curve, it is possible in these worlds to travel into any region of the past, present, and future, and back again, exactly as it is possible in other worlds to travel to distant parts of space." Kurt Gödel, "Relativity and Idealistic Philosophy", in Paul Arthur Schilpp, ed., *Albert Einstein, Philosopher-Scientist*. [New York: MJF Books, 1970.] See also section 5.7 of S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space-Time*. [Cambridge: Cambridge University Press, 1973.] Gödel does not, in fact, explain how you could "build a rocket ship" in such a universe, which is basically just a receptacle for a homogeneous fluid with zero pressure, but in some related cosmological model it may be possible.

These are also required for the construction of traversable wormholes, and unsurprisingly those are the basis of Thorne's time machine.[37] The idea that the spacetime manifold might be multiply connected, and that one might attach a kind of handle to it that would provide a short cut between widely separated points (thus in effect allowing you to exceed the speed of light) has been suggested on many occasions, not only by authors of science fiction; Misner and Wheeler,[38] e.g., proposed it as an explanation for the origin of charge:



[37] Michael S. Morris, Kip S. Thorne, and Ulvi Yurtsever, Wormholes, Time Machines, and the Weak Energy Condition", *Physical Review Letters*, **61** (13), 1446-1449, (1988).

[38] Charles W. Misner and John A. Wheeler, "Classical physics as geometry", *Annals of Physics* **2**, 525-603 (1957). The illustration is Figure 3.

So this isn't preposterous a priori. The potential for paradox only arises when you take into consideration the fact that the endpoints, or portals, would be if not objects at least elements of physical reality, and could accordingly move relative to one another or lie at different gravitational potentials; in either case the clocks at each end would run at different rates, and thus if you brought the ends into proximity, one end would lie in the back light cone of the other, and a temporal loop could be formed.

Actually it's a trifle more complicated than that, but a digression is necessary to explain why.

{…}

Note, first, that the usual (set-theoretic) definition of a function, as a set of ordered pairs, is incorrect, or at least misleading.[39] Generally a space is given beforehand, and a function then defined upon it. But it can work the other way around: the function can define the space. This is what one might call the sheaf-theoretic point of view; and it originated, as did so much else, with Riemann.

Pro forma this necessitates a brief summary of the theory of functions of a complex variable, though whether that will clarify matters is an open question.[40]

An analytic function w = f(z) is one which is differentiable under the usual definition, but with the additional requirement that the limit dw/dz is the same no matter how you approach z in the complex plane.

---

[39] Taking it for granted that the most basic ideas in mathematics are those of *set* and *function*, the set-theoretic viewpoint derives the latter from the former, the categorial viewpoint, which is intuitively more appealing, does the reverse. See R. Goldblatt, *Topoi, the Categorial Analysis of Logic*. [Amsterdam: North-Holland, 1979.]

[40] Here I give up (mostly) on trying to draw the pictures myself, and simply borrow the relevant illustrations from three reliable sources: Lars Ahlfors, *Complex Analysis* [New York: McGraw-Hill, 1979]; Serge Lang, *Complex Analysis* [New York: Springer, 1999]; and George Springer, *An Introduction to Riemann Surfaces* [New York: Chelsea, 1957].

This has startling consequences: if a function is once-differentiable, it is infinitely differentiable; its values at a point are determined by its values on any loop around that point according to the Cauchy integral formula [Ahlfors]

Theorem 6. *Suppose that $f(z)$ is analytic in an open disk Δ, and let γ be a closed curve in Δ. For any point a not on γ*

$$n(\gamma,a) \cdot f(a) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)\,dz}{z-a}, \tag{20}$$

where $n(\gamma,a)$ is the index of a with respect to γ.

[index is the number of times the curve wraps around the point]; an analytic function is essentially a conformal mapping, i.e. one which preserves angles; in consequence you can treat the complex plane as equivalent (in this conformal sense, by stereographic projection) to the Riemann sphere, which can be pictures sitting on the plane at the origin and onto which zero is mapped to the south pole and infinity to the north pole:

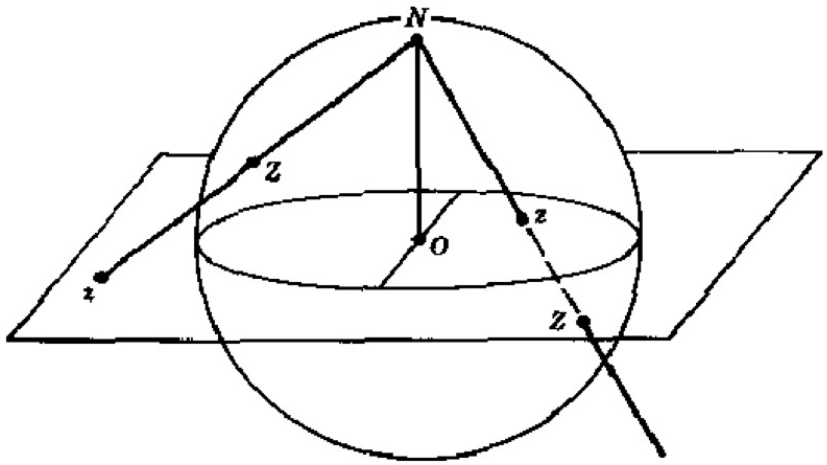


FIG. 1-3. Stereographic projection.

[Ahlfors]

I.e. in this geometry “congruence” = “conformal equivalence” and the plane and the sphere are essentially the same.

Moreover we have the Euler formula

$$e^{i\theta} = cos\theta + isin\theta$$

(angles in radians), with the famous corollary

$$e^{i\pi} = -1$$

which follows easily from the power series, and the simple representation of a complex number in polar coordinates

$$z = x + iy = r(cos\theta + isin\theta) = re^{i\theta}$$

so that, for instance, the complex nth roots of unity are given by

$$\cos\left(\frac{2\pi}{n}\right) + isin(\frac{2\pi}{n})$$

and lie on the unit circle.

Fast forward to the Weierstrass idea of analytic continuation: suppose you have a function known to be analytic on some domain it may be for some range of real numbers, e.g., as with the gamma function

$$\Gamma(n) = (n-1)!$$

which can be extended to real and complex values, as Euler did with the integral representation

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$$

Though the canonical form of the problem is to suppose that you know the values of the function in a small region around a

given point, say z = a, so that its derivatives can be computed, allowing an expansion in a power series

$$f(z) = f(a) + f'(a)(z-a) + \frac{1}{2!}f''(a)(z-a)^2 + \dots$$

It then turns out that this series converges in a disk whose radius is determined by the distance of the nearest singularity of the function (this is easy to picture, e.g., for z = 0 and f(z) = 1/(1 - z)]) and inside that disk you can compute the values of f . But! then you can pull the same trick for any *other* base point within this disk, and proceed along a curve by overlapping disks

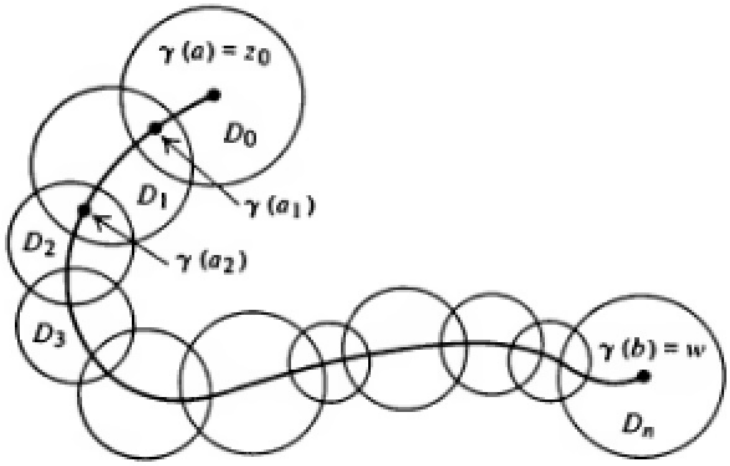


Figure 1

[Lang]

anywhere you please. So unless continuation is obstructed by a wall of singularities not impossible, consider e.g. trying to escape the unit disk with

$$f(z) = \prod_{n=1}^{\infty} \frac{1}{(1 - z^n)}$$

we can turn a local representation of a function into a global one unambiguously.[41]

Well. The procedure is *almost* unambiguous. There is a problem, which can be illustrated with algebraic functions y = f(z), defined by

$$p_0(x) + p_1(x)y + \ldots + p_n(x)y^n = 0$$

where the $p_i(x)$ are polynomials, which we are accustomed to thinking of as multiply-valued.

The simplest example is the square root,

$$-x + y^2 = 0$$

Here we don't even need to compute the power series, because there is a simple equation that determines the value of the function, i.e. if

---

[41] Put another way, an analytic function is completely causal, in a sense that would have satisfied Leibniz: if its values are known in an arbitrarily small region about some given point, then they can be inferred everywhere.

$$z = re^{i\theta}$$

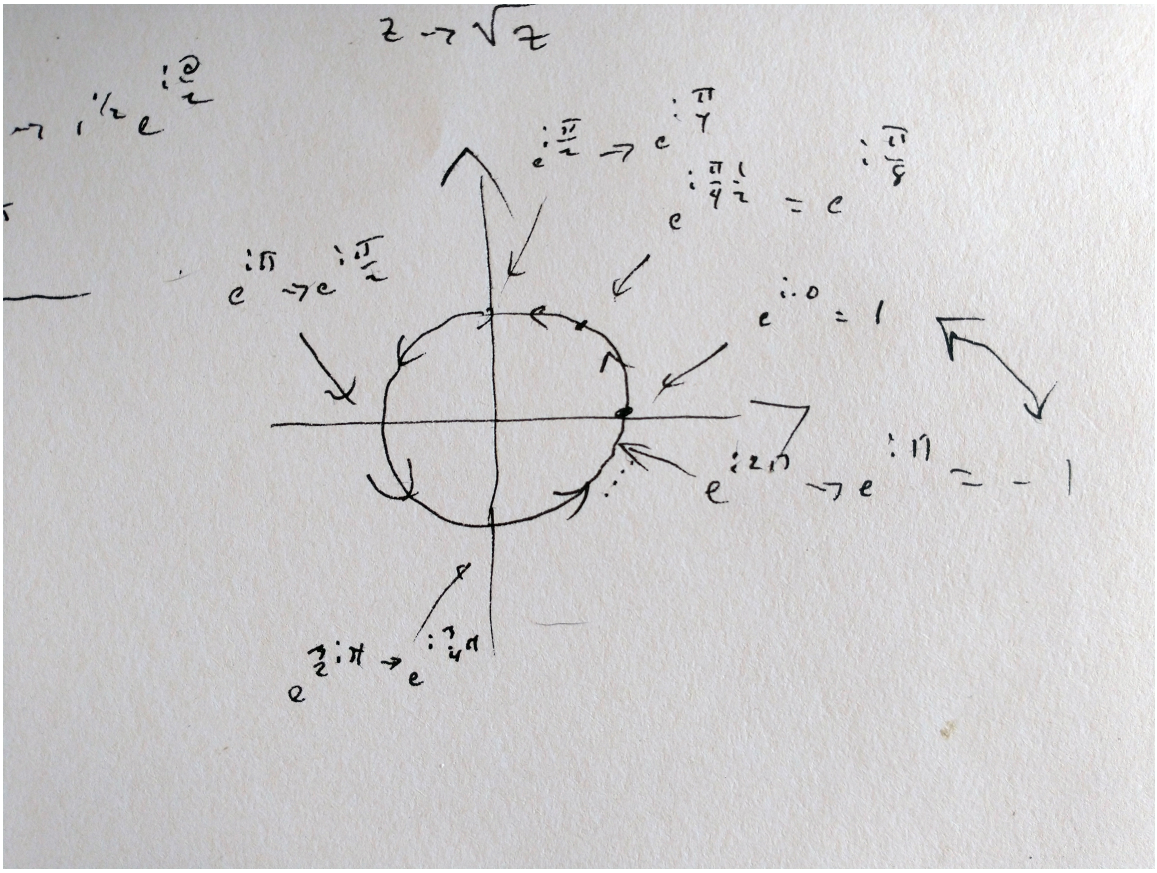
then (taking the positive real square root of r)

$$z^{\frac{1}{2}} = r^{\frac{1}{2}} e^{i\frac{\theta}{2}}$$

However if we start at z = 1 in the complex plane, let

$$\sqrt{1} = 1$$

and work our way around the unit circle counterclockwise, we compute



but when we get back to z = 1, approaching it from below, what we arrive at is the other value of the square root,

$$\sqrt{1} = -1$$

Of course if we repeat the circuit of the unit circle we get back to the original value, but even though our continuation procedure is *locally* unambiguous — if we know the value of the function at a point, we know it at any adjacent point — the square root is not.

There are two traditional solutions to this problem. One is to restrict the domain of the function: you draw a line (a branch cut) in the complex plane from 0 to infinity, and say that you aren't allowed to cross it. On the north side you have one limiting value, and on the south side you have another. So you pretend that the smooth function you have constructed on the smooth plane is actually a discontinuous function on a cut plane. This is awkward and ugly..

The other involves introducing a so-called Riemann surface, in this case a double-sheeted contraption stitched together along the branch cut that allows you to say that the square root *is* single-valued, it's just that you were defining it on the wrong space. So everything still looks like the complex plane in a small

patch, but when you sew the patches together you can end up with something much more complicated topologically.[42]

In fact it turns out that you can treat the general algebraic function in terms of attaching handles to the sphere though nth roots don't require them: you simply draw n branch cuts from 0 to infinity, and partition the surface into n regions. For n = 2 [the square root]:

[42] This is, of course, an instance of the manifold construction.

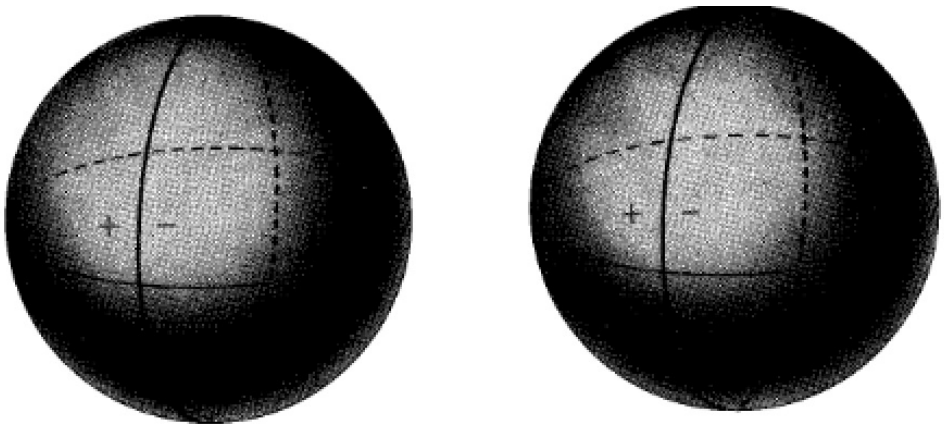


FIGURE 1–1.

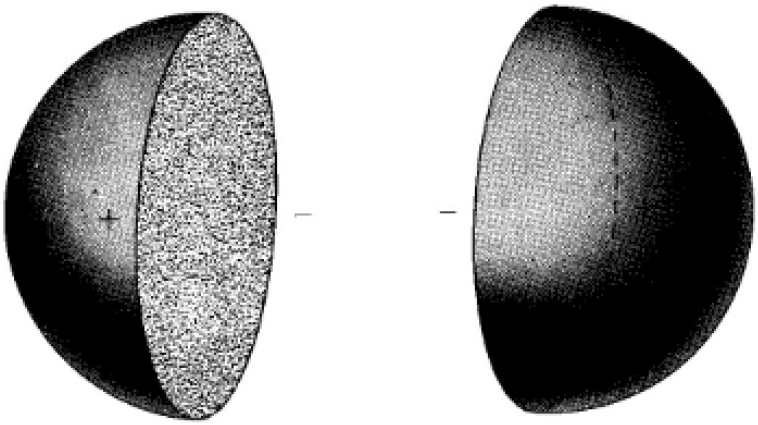


FIGURE 1–2.

[Springer]

For more complicated algebraic expressions, e.g. square roots of cubic polynomials (aka elliptic curves), there are multiple branch points [i.e. where the function takes only one value, 0 and infinity in the case of nth roots], you connect them in pairs, and paste the surface together along edges:

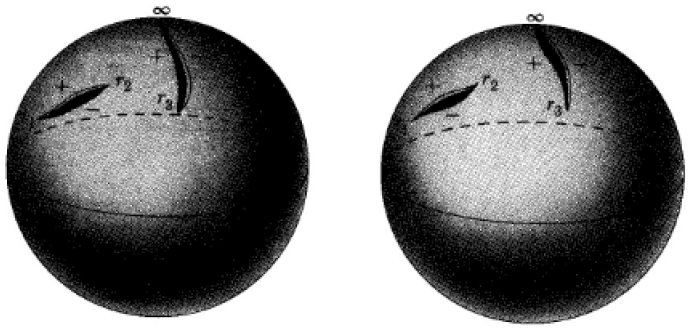


FIGURE 1–5.

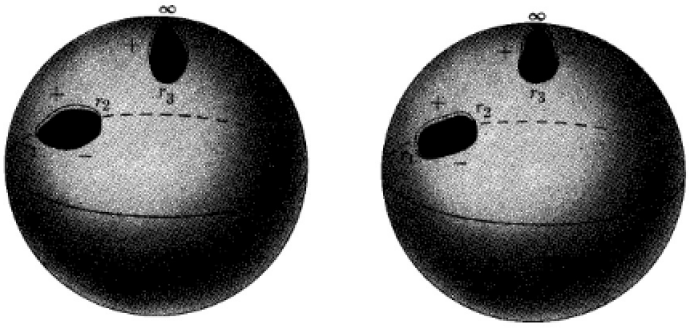


FIGURE 1–6.

[Springer]

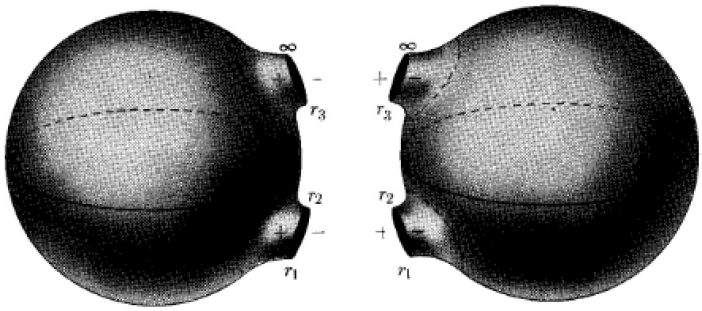


FIGURE 1–7.

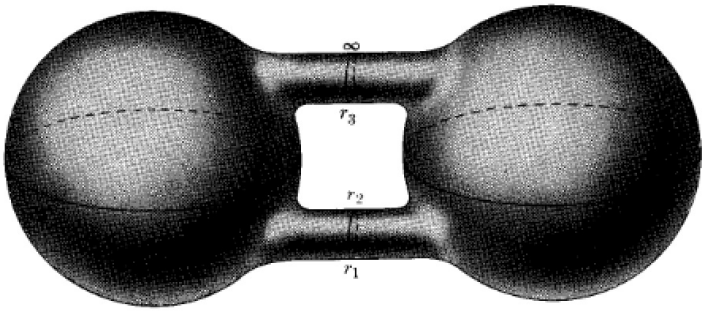


FIGURE 1–8.

[Springer]

Well. There is, naturally, a vast literature on this subject, in which, should you be curious, you can find corrections of all the mistakes I have made in this abbreviated exposition. The point toward which I am meandering is that the equations obeyed by the real and imaginary parts of analytic functions (the Cauchy-Riemann equations) look a lot like simplified field equations, Maxwell's e.g., and the whole process of sniffing out what the function does in a neighborhood around the point, and then iterating the procedure, looks a lot like the way reality gets pieced together by local causality, where "continuation" = "following events along a world line."

So hold that thought, and let's talk about time machines.

{…}

To reiterate what I am more or less taking for granted: the paradoxes these entail are well known. You climb into Bill and Ted's phone booth, dial yourself into the past, shoot your grandfather, and negate your own existence. But then who shot your grandfather? So you exist after all, and repeat the cycle. These are just the familiar logical paradoxes, and in fact physical transport is wholly unnecessary and "free will" is irrelevant: it suffices to be able to transmit information into the past; a single bit suffices, as in Kemeny's thought-experiment; we can analyze all this in terms of Boolean circuits with feedback and ask whether they have stable states or not. And the answer, of course, is "it depends."

If we're more ambitious, and try to construct models that look more like the physical world we actually inhabit, say the world of mechanics, at first glance there's no difficulty in supposing an isolated system to be periodic: a mass on a spring is the simplest example, and the dreaded time loop is no more than the equivalence of a periodic function on the line with a function on the circle.

The same could hold true for the entire universe, albeit with a caveat: if the world consisted of *two* oscillators, periodicity would

require that their periods be commensurable; otherwise we have the familiar billiard-table example of chaos theory, though you can invoke the Poincaré recurrence theorems to show that eventually the joint system returns arbitrarily close to the original configuration. Or you could say that only rational numbers or integers are really involved, as is implicitly the case with cellular automata; certainly most of the configurations you can make up in the Game of Life are recurrent.

There is also, however, the problem which manifests itself every time you say "But what if I change something?" These situations always generate paradox, because implicitly you are imagining this isolated periodic system and then coupling it to a larger system ("the rest of the world") which you are assuming to be irreversible and directional; again, it isn't necessary to invoke some act of will on the part of a free agent, a computer will suffice, or anything that sends a signal.

But the whole world could, in principle, be in on the game: you step into the phone booth and re-enter a past *that already happened*, and so it works out somehow that you didn't shoot your grandfather after all. Heinlein summarized this point of view in a couple of beautifully convoluted stories ("By His Bootstraps" and "All You Zombies", the latter the basis of the 2014 movie *Predestination)*, and though it looks weird one must admit that it appears to be logically possible: events might simply loop.

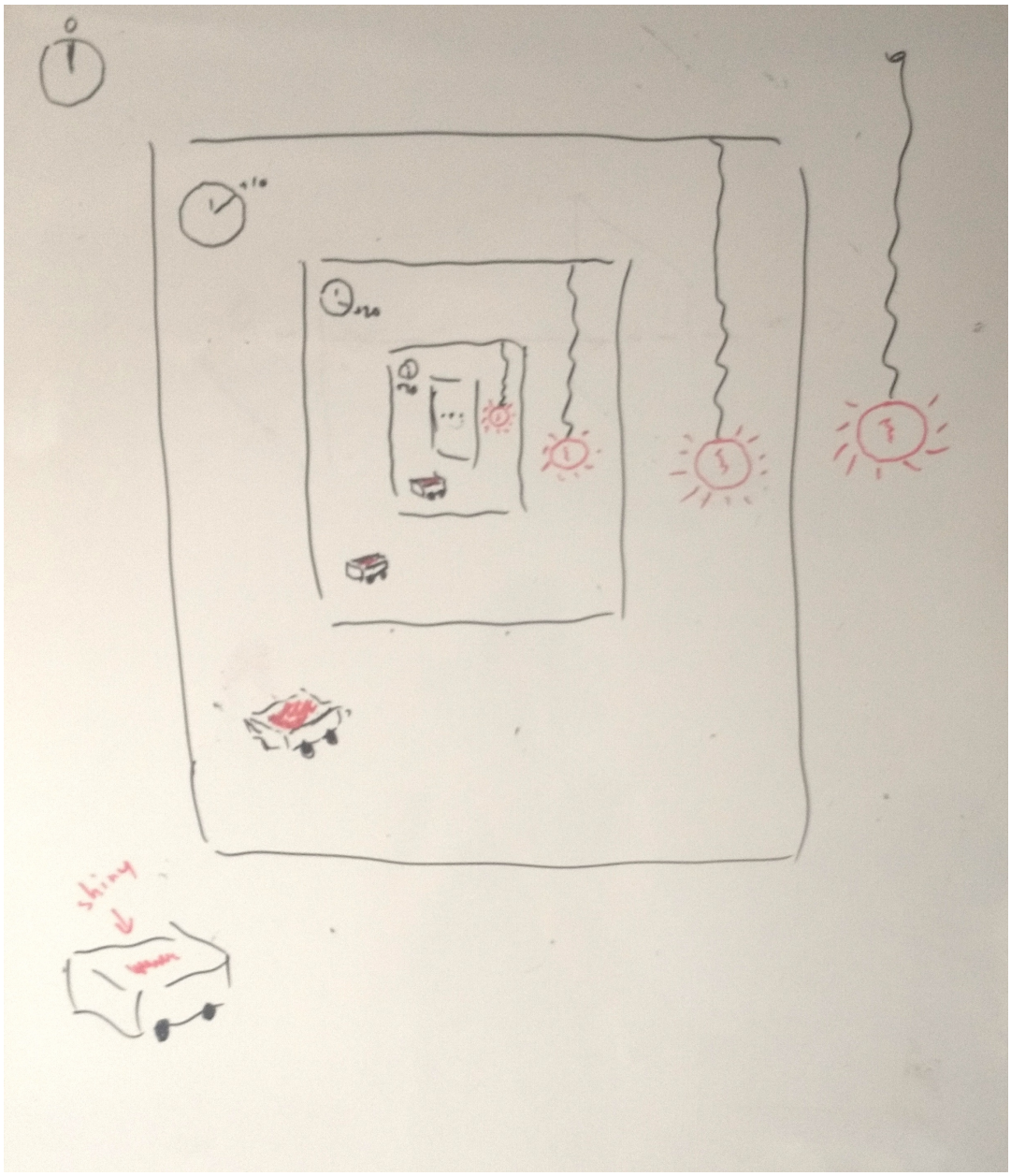
The other point of view is the one popularized by science fiction and exploited by Everett in his many-worlds interpretation of quantum mechanics: there is a manifold of possible worlds which branches at every instant, and if you return to the past to shoot your grandfather you've simply shunted yourself onto a different branch. William Gibson's novel *The Peripheral* explores this theme, with the added twist that he imagines the possibility of communication between the branches.

But there is a third alternative that is more interesting than the other two. So let's return to Thorne's wormhole, and perform a few experiments.

{…}

Suppose we have a wormhole connecting a couple of spatially adjacent points separated by a finite temporal interval; without loss of generality we can think of the portals as a couple of doorways on opposite sides of a room, and the time difference as ten minutes. Suppose as well that there's a light bulb hanging by a chain from the ceiling in the middle of the room, and there's a wheeled cart, a Radio Flyer, say, with a gold bar in it. I'll suppose I'm standing there, facing into the future, so that if I turn around and look behind me I see a nested set of doorways, one within the other, each ten minutes further into the past, extending back to when the wormhole was dragged into this configuration, and

if I look ahead I see a potentially infinite nested set of doorways, each ten minutes further into the future, as far as the eye can see.



So basically this is (again) just like Welles' infinite mirror shot in *Citizen Kane*; with lots and lots of copies of me, of course, though I won't worry about that particularly. Note however that these are less like Kane's reflections than the mirror scene in *Duck Soup*: we can wave at one another, and even reach through the portal and shake hands if we want. My presence is not, however, essential to the argument, and I'm planning on staying put, so that part of it isn't especially interesting. The bulbs and the carts are, however; they are *physically present*, available for interaction, in a way that mere reflections are not.

What is interesting? well, for instance, consider the illumination in the room. If the doors are open I am looking at an infinite number of light bulbs, and assuming intensity to fall off as the inverse square the contribution of all the doppelgängers in either direction is roughly[43] the intensity of the bulb right here multiplied by

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots = \zeta(2) \approx 1.6449\ldots$$

So it's more than three times as bright in here as one would expect from a single bulb. Photons are leaking into the room from elsewhere/elsewhen.

---

[43] Dependent on the offset.

Another thing is that local causality extends through the portals. By this I mean that I can reach through the doorway and pick up the gold bar from ten minutes ago — it's *right there* — or hand a rope to my doppelgänger ten minutes previous or future and have him pass it on through the doorways in either direction, so that I am *physically connected* to the gold bars, or the handles of the Radio Flyers, or the light switches in the rooms that I can see, backward or forward. For instance I can turn the light off in my room while holding on to the switch ten minutes forward to make sure it stays on, and then wait fifteen minutes before I turn the light back on in my room. — So these aren't like reflections in a mirror. If I flip the switch in my room, it *doesn't affect* the state of the next room. You can make up some story about magical action at a distance that turns the other lights off, or introduce ever-more-ridiculous interventions of *deus ex machina* — just then a pelican flies in the window! — but those simply wouldn't make sense. So once you've introduced this physical connection between now and ten minutes before and afterward, now is still now, but then isn't then. Each room that I can see has its own distinct future.

Or something like that. Let's put it this way: suppose I take the gold bar in my cart and hand it through the forward portal to be deposited in *that* cart. — First, do two gold bars magically appear in the other carts I see through the nested series of doorways framed by the doorway on the other side? Of course not, that

would be ridiculous, action at a distance. Second, if I wait ten minutes do two gold bars magically reappear in my cart? No.

Try it the other way: suppose I hand my gold bar back through the other door, so that now "ten minutes ago" there are two gold bars in the cart. Do two bars miraculously appear in the cart I just emptied? Where would they *come from*?

Even better, suppose the two gold bars in the cart in the adjacent room, forward or back, are handed through the doorway to be added to the next cart in that direction. Obviously this can be repeated indefinitely. So I can be sure that, even if I personally haven't any gold, some one of my doppelgängers in some distant room that looks just like this one, with a clock that reads that many steps times ten minutes plus or minus, is sitting on Fort Knox.

{…}

So what does that mean?

I'll put it this way: introducing a wormhole into the spatiotemporal continuum introduces a topological complication, like a branch point in the complex plane; the rooms visible through the portals to front and rear lie on different sheets; if a time machine is constructed, it renders physical reality multiply-valued. The wormhole connects, not two

locations on a fixed manifold, but two copies of the manifold. The analogue of analytic continuation, which connects the values of a function on overlapping patches, is local causality, which pieces together events in the neighborhood of a world-line. [44] So physical reality is still unambiguous in its construction. It's just a lot more complicated than you thought it was.

{…}

So I doubt this says anything about the capacities of advanced civilizations to exceed the speed of light, as Thorne originally speculated. It would, however, explain a lot of things about cosmology; particularly when you consider that passing one wormhole through another provides a mechanism for connecting the original sheet to one on which their numbers have been multiplied to arbitrary extent, and that if you can imagine tunnels popping up out of the vacuum in the era of the very early universe dominated by quantum gravity, then they could [a] conjure up a lot of matter/energy out of nothing, at least on some connected sheet, and [b] explain other weird phenomena like inflation, in which the size of the universe is supposed to have increased by a zillion orders of magnitude in a scintillionth of a second. (Time need not have gone forward so

---

[44] More precisely, causal topology for hyperbolic partial differential equations says that you know the values of the fields at a point in space/time if you know their values on a spacelike section of its back light cone.

much as slipped sideways.) The questions then arise as to how many of these things remain on the sheet we inhabit, and what they might look like. [Galactic nuclei?]

Moreover there's the question of what it would mean if the field-theoretic vacuum were creating virtual wormholes all the time. Somehow I am picturing something like this:

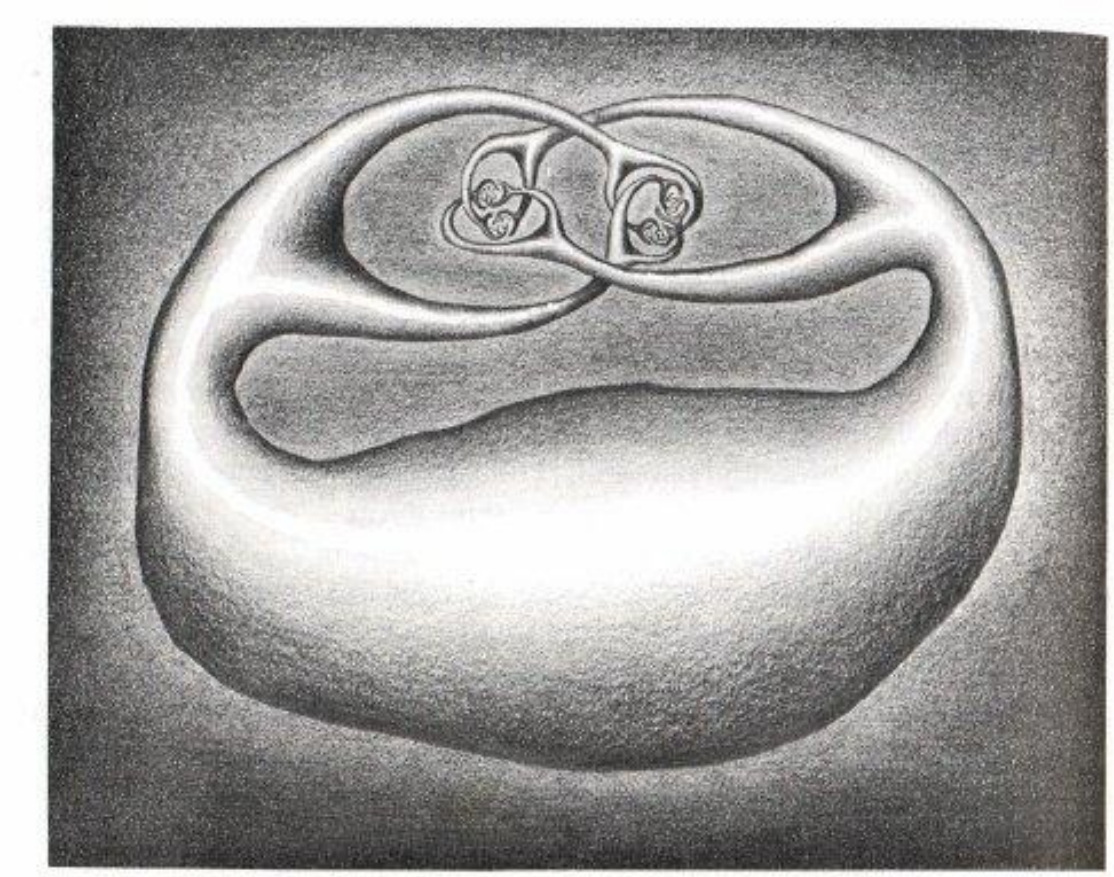


FIG. 4–11. The Alexander horned sphere.

Maybe Wheeler wasn't crazy after all.