Frontiers of Time

John Archibald Wheeler

CENTER FOR THEORETICAL PHYSICS
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712
USA

International School of Physics
"Enrico Fermi"
Varenna
LXXII COURSE
25 July-6 August 1977

Frontiers of Time

John Archibald Wheeler
Center for Theoretical Physics
Department of Physics
The University of Texas
Austin, Texas 78712

[Revised manuscript of lectures presented August 1-5 and to appear in the volume, *Rendiconti della Scuola Internazionale di Fisica "Enrico Fermi," LXXII Corso*, edited by N. Toraldo di Francia and Bas van Fraassen, *Problems in the Foundations of Physics*, North Holland, Amsterdam, scheduled for 1979 publication.]

© by John Archibald Wheeler, August 1978

I. Law Without Law

"species will never vary, and have remained the same since the creation of each species."

-Charles Lyell,(1) writing almost three
decades before *The Origin of Species*

"[The astronomer Sir John Frederick William] *Herschel says my book is 'the law of higgledy-piggledy'."*

-Charles Darwin,(2) 18 days after the
November 24, 1859 publication of
The Origin of Species

Are the laws of physics eternal and immutable? or are these laws, like species, mutable(3) and of "higgledy-piggledy" origin?

The hierarchical speciation of plant and animal life, we now know, arises out of the blind accidents of genetic mutation and natural selection(5,6). Likewise the gas laws, the pressure-volume-temperature relation for water and for other substances, and the laws of thermodynamics take their origin in the chaos of molecular collisions. But as for the molecules themselves, the particles of which they are made, and the fields of force that couple them, is it conceivable that they too derive their way of action, their structure, and even their existence from multitudinous accidents?

Such questions about the "plan" of physics we would hardly raise if we had the skeleton of it in hand. But we don't.

Now and then we meet a colleague in another realm of thought who still thinks physics is in possession of this plan. He cites the words of Laplace(7) and reiterates the Laplacean vision as he understands it: the laws are definite, the initial coordinates and momenta are definite, and therefore the future is definite. The universe is a machine. No, we have to tell him; that is a cracked paradigm. Quantum mechanics allows us to know a coordinate, or a momentum, but not both. Of the initial value data that Laplace needed, the principle of complementarity(8) or indeterminacy(9) says half do not and cannot exist.

It is no use to warn our colleague of the grip of determinism, nor to compare it for him with "the hold of astrology on the Renaissance mind; neither education nor enlightenment [Jacob Burckhardt insisted(10)] could do anything against this delusion. . . because it was supported by the authority of the ancients and satisfied passionate fantasies and the fervent wish to know and determine the future"(11). He reads more physics and comes back convinced that the plan of the world is still determinism. No one can deny, he insists, that the Schrödinger equation foreordains in every detail the time development of the wave function. Yes, there is a probability element in the physics that shows up at the instant of an observation, he admits. However, that element of chance is not at all in contradiction with determinism, he tells us, but evidence that we have failed to include the observer in our bookkeeping. In support of this contention he cites the thesis of Everett(12,13,14), that the measurement postulate of quantum mechanics(15) can be derived out of the wave equation itself, rather than being added from outside as a mysterious and foreign element. On this view, our colleague reminds us, the relevant dynamical system is the system under study augmented by the observer system. The wave function for this larger system lends itself to being written as the sum of products. Each product contains one factor referring to the system under study, multiplied by a second factor referring to the observer. Measurement is described in terms of the correlation between these two factors. The first factor describes the system under study as being in a specific quantum state. The second, according to Everett's analysis, represents the observer as aware that the system under study is in that quantum state. The "coexistence" in one overall wave function of these alternative states of observer-plus-observed-system has given rise to such phrases as "branching histories" and "the many worlds interpretation of quantum mechanics." That one can get all this, our friend concludes, out of the deterministic Schrödinger equation shows more clearly than any argument ever advanced that nature is at bottom deterministic.

Imaginative Everett's thesis is, and instructive, we agree. We once subscribed to it(16). In retrospect, however, it looks like the wrong track. First, this formulation of quantum mechanics denigrates the quantum. It denies from the start that the quantum character of nature is any clue to the plan of physics. Take this Hamiltonian for the world, that Hamiltonian, or any other Hamiltonian, this formulation says. I am a principle too lordly to care which, or why there should be any Hamiltonian at all. You give me whatever world you please, and in return I give you back many worlds. Don't look to me for help in understanding this universe.

Second, its infinitely many unobservable worlds make a heavy load of metaphysical baggage. They would seem to defy Mendeléev's demand of any proper scientific theory, that it should "expose itself to destruction." Wigner(17) [see also Wigner(18, 19, 20)], Weizsäcker(21) and Wheeler(22) have made objections in more detail, but also in quite contrasting terms, to the relative-state or many-worlds interpretation of quantum mechanics. It is hard to name anyone who conceives of it as a way to uphold determinism.

You tell me what isn't the plan of physics, our friend rejoins. If you understand quantum mechanics so well, why don't you tell me what *is* the plan of physics?

No one knows, we reply. We have clues, clues most of all in the writings of Bohr(23,24,25), but no answer. That he did not propose an answer, not philosophize, not go an inch beyond the soundest fullest statement of the inescapable lessons of quantum mechanics, was his way to build a clean pier for some later day's bridge to the future.

What kind of a "plan of physics" do you think Bohr had in mind, our colleague asks. I know Einstein's words(26), "Physics is an attempt to grasp reality as it is thought independently of its being observed." I know Bohr's reply(28), "These conditions [of measurement] constitute an inherent element of any phenomenon to which the term 'physical reality' can be attached. . . .[This requires] a final renunciation of the classical ideal of causality and a radical revision of our attitude towards the problem of physical reality." But if I could have asked Bohr, how did he think the universe came into being, and what is its substance, what would he have said?

It is too late to ask. The plan is up to us to find.

The universe can't be Laplacean. It may be higgledy-piggledy. But have hope. Surely someday we will see the necessity of the quantum in its construction. Would you like a little story along this line?

Of course! About what?

About the game of twenty questions. You recall how it goes--one of the after-dinner party sent out of the living room, the others agreeing on a word, the one fated to be questioner returning and starting his questions. "Is it a living object?" "No." "Is it here on earth?" "Yes." So the questions go from respondent to respondent around the room until at length the word emerges: victory if in twenty tries or less; otherwise, defeat.

Then comes the moment when we are fourth to be sent from the room. We are locked out unbelievably long. On finally being readmitted, we find a smile on everyone's face, sign of a joke or a plot. We innocently start our questions. At first the answers come quickly. Then each question begins to take longer in the answering--strange, when the answer itself is only a simple "yes" or "no." At length, feeling hot on the trail, we ask, "Is the word 'cloud'?" "Yes," comes the reply, and everyone bursts out laughing. When we were out of the room, they explain, they had agreed not to agree in advance on any word at all. Each one around the circle could respond "yes" or "no" as he pleased to whatever question we put to him. But however he replied he had to have a word in mind compatible with his own reply--and with all the replies that went before. No wonder some of those decisions between "yes" and "no" proved so hard!

And the point of your story?

Compare the game in its two versions with physics in its two formulations, classical and quantum. First, we thought the word already existed "out there" as physics once thought that the position and momentum of the electron existed "out there," independent of any act of observation. Second, in actuality the information about the word was brought into being step by step through the questions we raised, as the information about the electron is brought into being, step by step, by the experiments that the observer chooses to make. Third, if we had chosen to ask different questions we would have ended up with a different word--as the experimenter would have ended up with a different story for the doings of the electron if he had measured different quantities or the same quantities in a different order. Fourth, whatever power we had in bringing the particular word "cloud" into being was partial only. A major part of the selection--unknowing selection--lay in the "yes" or "no" replies of the colleagues around the room. Similarly, the experimenter has some substantial influence on what will happen to the electron by the choice of experiments he will do on it; but he knows there is much impredictability about what any given one of his measurements will disclose. Fifth, there was a "rule of the game" that required of every participant that his choice of yes or no should be compatible with some word. Similarly, there is a consistency about the observations made in physics. One person must be able to tell another in plain language what he finds and the second person must be able to verify the observation.

Go on!

That is difficult! Interesting though our comparison is between the world of physics and the world of the game, there is an important point of difference. The game has few participants and terminates after a few steps. In contrast, the making of observations is a continuing process. Moreover, it is extraordinarily difficult to state sharply and clearly where the community of observer-participants begins and where it ends.

This comparison between the world of quantum observations and the game of twenty questions misses much, but it makes the vital central point. In the real world of quantum physics, *no elementary phenomenon is a phenomenon until it is an observed phenomenon.* In the surprise version of the game no word is a word until that word is promoted to reality by the choice of questions asked and answers given. "Cloud" sitting there waiting to be found as we entered the room? Pure delusion! Momentum, $p_x = 1.4\times10^{-19}$ gcm/s, or position, $x = 0.31\times10^{-8}$ cm, of the electron waiting to be found as we start to probe the atom? Pure fantasy!

Mann may be going too far when he suggests[(29)] that ". . .we are actually bringing about what seems to be happening to us." However, it is undeniable that each of us, as observer, is also *one* of the participants in bringing "reality" into being.

Until I heard your story I had never grasped what a strange and fascinating quality the universe has, and never understood how absolutely indefensible determinism is. Won't you go on? What do *you* think the quantum is trying to tell us about the structure of physics?

Nobody wants conjecture!

But how can anybody even begin to ask the right questions if he doesn't have at least some thought in his mind about how the answers look? I can see you have some suspicions about the shape of things. What are they?

Little though we know, I agree we owe it to each other to talk as frankly as we can.

Please do.

"Law without law": it is difficult to see what else than that can be the "plan" of physics. It is preposterous to think of the laws of physics as installed by a Swiss watchmaker to endure from everlasting to everlasting when we know that the universe began with a big bang. The laws must have come into being(30,3). Therefore they could not have been always a hundred percent accurate. That means that they are derivative, not primary. Also derivative, also not primary is the statistical law of distribution of the molecules of a dilute gas between the two interconnecting portions, V_1 and V_2, of a total volume $V = V_1 + V_2$,

(1) $$N_1 = V_1(N/V); \quad N_2 = V_2(N/V).$$

This law is always violated and yet also always upheld. The individual molecules laugh at it; yet as they laugh they find themselves obeying it. The statistical fluctuations about the predicted values,

(2) $$(\delta N_1)_{RMS} = (\delta N_2)_{RMS} = (\bar{N}_1\bar{N}_2/N)^{\frac{1}{2}},$$

in every normal circumstance are absolutely negligible.

Are the laws of physics of a similar statistical character? And if so, statistics of what? Of billions and billions of acts of observer-participancy which individually defy all law?

The only thing harder to understand than a law of statistical origin would be a law that is not of statistical origin, for then there would be no way for it--or its progenitor principles--to come into being. On the other hand, when we view each of the laws of physics--and no laws are more magnificent in scope or better tested--as at bottom

statistical in character, then we are at last able to forego the idea of a law that endures from everlasting to everlasting.

Individual events. Events beyond law. Events so numerous and so uncoordinated that, flaunting their freedom from formula, they yet fabricate firm form.

"Fabricate form"? Do you suggest that even the 4-dimensional spacetime manifold is only a fabrication, only a theory--irreplaceable convenience though that theory is?

Yes! Compare spacetime with cloth. Each it is useful under everyday circumstances to call a manifold. Yet each is exactly then most obviously not a manifold where it comes to an end, whether in the selvedge made by the loom, or in the geodesic terminations made by one of the "gates of time"--big bang or big crunch(31,32) or black hole.(33) Nowhere more clearly than in the ending of spacetime are we warned that time is not an ultimate category in the description of nature.(34)

Aren't you being extreme? I see the lesson of the game of twenty questions. I begin to believe with you that no elementary phenomenon is a phenomenon until it is an observed phenomenon. I accept that events of observer-participancy as you call them, occupy a special place in the scheme of things. I agree that that word "cloud" was brought into being entirely through such elementary events. But that such events, however numerous, should be the sole blocks for building the laws of physics--and space and time themselves--seems to me preposterous. You surely have been involved enough in times past with nuts-and-bolts physics to know the difference between science and poetry; yet if I appreciate the drift of what you say, you might as well be quoting Shakespeare,(35)

> . . . These our actors,
> As I foretold you, were all spirits and
> Are melted into air, into thin air:
> And like the baseless fabric of this vision,
> The cloud capp'd towers, the gorgeous palaces,
> The solemn temples, the great globe itself,
> Yes, all which it inherit, shall dissolve
> And like this insubstantial pageant faded,
> Leave not a rack behind. We are such stuff
> As dreams are made on. . .

I can't believe any such dreamlike vision of the physical world. As Samuel Johnson used to say, I have only to kick a stone to find it real enough.

Why do you say "preposterous"? Perhaps Shakespeare understood this universe of ours better than we do ourselves! You have known for years that the atom is more than 99.99 percent emptiness. If matter turns out in the end to be altogether ephermeral, what difference can that make in the

pain you feel when you kick the rock? And how can matter--and spacetime--be anything but mutable, coming into being at one gate of time and fading out of existence at the other?

No physics before the big bang, or after the big crunch?

No! The lesson of Einstein's standard closed-space cosmology is different and stronger. It denies all meaning to such terms as "before the big bang" and "after the big crunch." Particles or fields or mathematics won't do for ultimate building blocks. They can't come into being or fade out of existence.(30)

Yes, I appreciate the reasons given(36) against believing in any "magic particle" or any "magic field" or any "magic mathematics" as the foundation of physics; but isn't it even more difficult to think of acts of observer-participancy as the magic ingredient?

Difficult, yes; inconceivable, no.

Go on!

No, we have to stop here. It is beyond the power of today to fit together the pieces of the puzzle.

Don't stop! You've carried me halfway into an exciting mystery story. You can't leave me without the traditional half-way-point review of the important clues and first try at a working hypothesis.

Review? A proper review would be impossibly ambitious. And how can one advance a working hypothesis that will not be wrong tomorrow and ridiculous the day after?

I appeal to you to go on. You have told me more than once that science advances only by making all possible mistakes; that the main thing is to make the mistakes as fast as possibl and recognize them. You like to quote the motto of that engine inventor, John Kris: "Start her up and see why she don't run." You point to Einstein's definition of a scientist, "An unscrupulous opportunist." If you believe all this, and are a true colleague of mine, you must go on.

You leave no escape!

Good! Then let us agree to go on; but let us replace the comprehensive review of clues that you wanted by something more modest. How would it do, for example, to survey some of the lessons we have learned from the study of time, and how those lessons bear on "observer-participancy"?

I accept, and with many thanks. But first tell me the central point as you see it.

The absolute central point would seem to be this: The universe had to have a way to come into being out of nothingness, with no prior laws, no Swiss watchworks, no nucleus of crystallization to help it--as on a more modest level, we believe, life came into being out of lifeless matter with no prior life to guide the process.(38,5,6)

When we say "out of nothingness" we do not mean, out of

the vacuum of physics. The vacuum of physics is loaded with geometrical structure and vacuum fluctuations and virtual pairs of particles. The universe is already in existence when we have such a vacuum. No, when we speak of nothingness we mean nothingness: neither structure, nor law, nor plan.

A conception more clearly impossible I never heard!

Preposterous we have to agree is the idea that everything is produced out of nothing--as preposterous, but perhaps also as inescapable, as the view that life had its origin in lifeless matter.

But how?

"Omnibus ex nihil ducendis sufficit unum," Leibniz told us;(39) for producing everything out of nothing one principle is enough. Of all principles that might meet this requirement of Leibniz nothing stands out more strikingly in this era of the quantum than the necessity to draw a line between the observer-participant and the system under view. Without that demarcation it would make no sense to do quantum mechanics, no sense to speak of quantum theory of measurement, no sense to say that "No elementary phenomenon is a phenomenon until it is an observed phenomenon." The necessity for that line of separation is the most mysterious feature of the quantum. We take that demarcation as being, if not the central principle, the clue to the central principle in constructing out of nothing everything.

Let me ask if your reasoning couldn't be turned around. You talk of the observer-participant of quantum theory as the mechanism for the universe to come into being. If that is a proper way of speaking, would the converse not also hold: The strange necessity of the quantum as we see it everywhere in the scheme of physics comes from the requirement that--via observer-participancy--the universe should have a way to come into being?

Your point is exciting indeed. If true--and it is attractive--it should provide someday a means to <u>derive</u> quantum mechanics from the requirement that the universe must have a way to come into being.(40)

I know that in that empty courtyard many a game cannot be a game until a line has been drawn--it does not matter where--to separate one side from the other. I know that no Gaussian flux integral can be a flux integral until the 2-surface over which it runs--bumpy and rippled though we make it and deform it as we will--has been extended to closure. But how much arbitrariness is there in the this more ethereal kind of demarcation, the line between "system" and "observing device"?

Much arbitrariness! Bohr stresses(42) that the stick we hold can itself be an object of investigation, as when we run our fingers over its surface. The same stick, when grasped firmly and used to explore something else, becomes an extension of the observer or--when we depersonalize--a part of the measuring equipment. As we withdraw the stick from the one role, and recast it in the other role, we transpose the line of demarcation from one end of it to the other. The distinction between the probed and the probe, so evident at this scale of the everyday, is the without-which-nothing of every elementary phenomenon, of every "closed" quantum process.

Do we possess today any mathematical or legalistic formula for what the line is or where it is to be drawn?

No.

Then what is important about this demarcation?

Existence, yes; position, no. It is the mark of an observation to leave an "indelible" record, according to Belinfante.(43) Wigner argues that an observation is only then an observation when it becomes part of "the consciousness of the observer"(44) and points to "the impressions which the observer receives as the basic entities between which quantum mechanics postulates correlations."(45) For Bohr the central point is not "consciousness," not even an "observer," but an experimental device--grain of silver bromide, Geiger counter, retina of the eye--capable of an "irreversible act of amplification."(47) This act brings the measuring process to a "close."(48) Only then, he emphasized, is one person able "to describe the result of the measurement to another in plain language."(49) He adds that "all departures from common language and ordinary logic are entirely avoided by reserving the word 'phenomenon' solely for reference to unambiguously communicable information."(50)

I would have felt very uncomfortable if Bohr had used the term "consciousness" in defining the elemental act of observation. I would not have known what he meant. However, I am beginning to understand and accept the terms he actually adopts, "brought to a close by an irreversible act of amplification" and "communicable in plain language." What *was* his position on consciousness?

We have asked Jørgen Kalckar, who collaborated with Bohr in his last months, and he has kindly replied(51), "During work on the preparation of some lecture, to define the phenomenon of consciousness, Bohr used a phrase somewhat like this: a behaviour so complex that an adequate account would require references to the organism's 'self-awareness.' I objected jokingly that with this definition he would soon have to ascribe a consciousness to the highly developed electronic computers. This did not worry Bohr. 'I am absolutely prepared,' said he, 'to talk of the spiritual life of an electronic computer; to state that it is reflectin or that it is in a bad mood. . . .The question whether the machine *really* feels or ponders, or whether it merely looks as though it did, is of course absolutely meaningless.'"

Other outstanding thinkers have argued otherwise. For them "consciousness" makes an unclimbable difference of principle etween even the most powerful imaginable computer and he brain.(52)

o you agree with that argument? ow can we possibly accept such a difference of principle? o we not believe that brain function itself will someday e explained entirely in terms of physical chemistry and lectrochemical potentials? What escape is there from the easoning of von Neumann(53) and Bohr and many active present ay investigators? When one of the three discoverers of he mechanism of superconductivity today gives us, chapter y chapter and verse by verse, an entirely cellular account f the mechanism of memory,(54,55,56) who can dismiss it? hen a distinguished computer expert and student of the tructure of society details, one by one, the distinctions proposed in imes past between "consciousness" and the computer, and painstakingly nalyzes each down to nothingness;(57) what case can anyone ossibly maintain for *any* distinction of principle between he computer and the brain?

am happy not to have to delve today into the term "consciousess." I find it hard enough to know what to make of "irreversible act of amplification." Never have I heard f an act of amplification that was not characterized by n amplification factor, or an equivalent quantity; and ever an amplification factor that was not a finite number. etween infinity and a finite number there may be a difference f principle; but between one finite number and another there s only a difference of degree. How big does the grain of silver bromide have to be, or the avalanche of electrons in the Geiger counter, before we count the measuring process as brought to a close by an irreversible act of amplification? According as I specify one or another number as the critical level of amplification, don't I make all the difference between rating or not rating a given process as an "elementary phenomenon"?

According as the closed Gaussian surface encloses a given elementary charge or not, we find an unmistakable difference in the surface integral of the electric flux. Nevertheless we know enough about the relevant invariance principle never to question the correctness of always identifying flux with enclosed charge. About "elementary quantum phenomenon" we have not today learned, but have a deep obligation someday to learn, enough to display a similar covariance with respect to where we draw the line. That is what "complementarity" is all about.

Even if neither you nor I know how to define that line, I like the idea that the "game" in the empty courtyard is only then possible when *a* line is drawn. May I question you now about the game itself? How would you describe it if forced to commit yourself?

Let us try to squeeze an answer(36) into three sentences and a picture (Fig. 1). The universe is a self-excited circuit. As it expands, cools and develops, it gives rise to observer-participancy. Observer-participancy

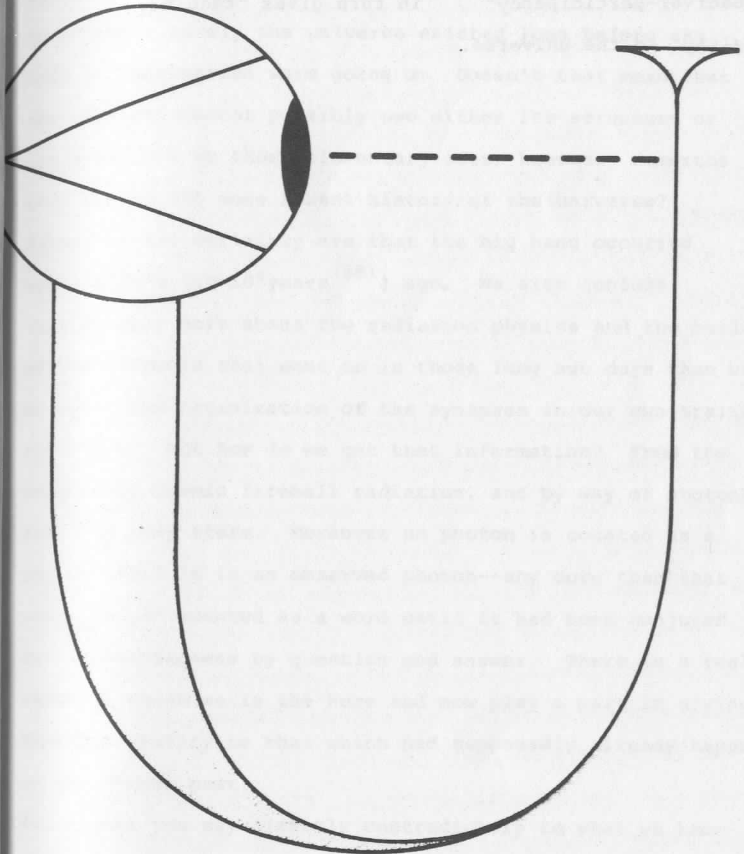


Figure 1.

Figure 1. The universe (big U) viewed as a "self-excited circuit," grows and in time gives rise to observership. "Observer-participancy" in turn gives "tangible reality" to the universe.

in turn gives what we call "tangible reality" to the universe.

Thank you for the brevity and challenge of that working hypothesis. Forgive me if I respond with an immediate objection. Surely the universe existed long before any acts of observation were going on. Doesn't that mean that the universe cannot possibly owe either its structure or its existence to those elementary acts, however numerous they are in the more recent history of the universe?

Agreed we all certainly are that the big bang occurred some 10×10^9a [10×10^9years(58)] ago. We also confess that we know more about the radiation physics and the building of the elements that went on in those long ago days than we do about the organization of the synapses in our own brains right now. But how do we get that information? From the primordial cosmic fireball radiation, and by way of photons from far away stars. Moreover no photon is counted as a photon until it is an observed photon--any more than that word "cloud" counted as a word until it had been conjured out of nothingness by question and answer. There is a real sense in which we in the here and now play a part in giving tangible reality to that which had supposedly already happened in the remote past.

Isn't what you say directly contradictory to what we know about the direction of time? How can an observation made now have any influence whatsoever on what has already happened?

Ah, but "what has already happened" is not so easy to say. Perhaps here is where we should start to outline that revie that you made me promise to give. Let us take up in our next encounter (Review II) "The 'Past' and the Delayed Choice Double-Slit Experiment." Here we shall see in what sense, after an electron or photon has *already* traversed a screen with two holes in it, we can choose whether it *shall have* gone through only one of the two holes (triggeri one or the other of two distinct counters) or through both of them (contributing to building up an interference fringe)(59). A decision in the present thus makes a striking difference in what we can rightfully say about the past.

To say, "No elementary phenomenon is a phenomenon until it is an observed phenomenon" is to make no small change in ou traditional view that something has "already happened" before we observe it. The word "cloud," we mistakenly thought, already existed in the room before we "uncovered" it. The photons of the primordial cosmic fireball radiatio that enter our telescope today, we customarily assume, alre had an existence in the very earliest days of the universe, long before life evolved. However, not until we catch a particular one of those photons in a particular state with particular parameters, not until the elementary phenomenon is an observed phenomenon, do we have the right even to call

it a phenomenon. This is the sense, the limited sense, but the inescapable sense, in which we, here, now, have a part in bringing about that which "had already happened" at a time when no observers existed.

But what about the unbelievably more numerous relict photons that escape our telescope? Surely you do not deny them "reality"?

Of course not; but their "reality" is of a paler and more theoretic hue. The vision of the universe that is so vivid in our minds is framed by a few iron posts of true observation--themselves also resting on theory for their meaning--but most of the walls and towers in the vision are of papier-maché, plastered in between those posts by an immense labor of imagination and theory. In this labor, ". . .we can never neatly separate what we see from what we know. . .what we call seeing is invariably coloured and shaped by our knowledge (or belief) of what we see."(61) "Without some initial system, without a first guess to which we can stick unless it is disproved, we could. . . make no 'sense' of the milliards of ambiguous stimuli that reach us from our environment. In order to learn, we must make mistakes. . .the simplicity hypothesis cannot be learned. It is. . .the only condition under which we could learn at all."(62) ". . .our mind will still react to the challenge of this conundrum [of what we 'see'] by throwing out a random answer, making ready to test it in terms of consistent possible worlds. It is these answers that will transform the ambiguous stimulus pattern into the image of something 'out there.'"(62)

What keeps these images of something "out there" from degenerating into separate and private universes: one observer, one universe; another observer, another universe?

That is prevented by the very solidity of those iron posts, the elementary acts of observership-participancy.

That is the importance of Bohr's point that no observation is an observation unless we can communicate the results of that observation to others in plain language.(49)

I have the impression that texts on quantum mechanics deal with the case where one observer is involved; research papers on the Einstein-Podolsky-Rosen experiment(64) address the situation where two observers are making measurements; and nobody deals with the richness of the case where many observers are at work on the system. How then is the limiting case to be analyzed that you have in mind? As I understand it you propose that the statistics of billions upon billions of elementary acts of observation gives rise all by itself, without law or plan, to all the structure and laws of physics. Ever to check that proposal would seem an impossibility.

All of us--we agree--will have to try long and hard before we learn how to do that kind of statistics; but have hope that we will! It is good fortune for this enterprise that some of the necessary ground work has already been laid in a paper by Houtappel, Van Dam and Wigner.(65) Why don't

we make it the subject of Review III, "'Development in Time' Gives Way to 'Correlation in Time'"?

Good! There will surely come a day when the concept of "law without law"--out of the statistics of acts of observation-participancy--can be tested, and either fleshed out or disproved. Until then, however, I shall be one of the many who will persist in considering Maxwell's electromagnetism, Einstein's geometric theory of gravity, and the Yang-Mills theory(66) of the quark-binding field as among the truly great achievements of science. They take an enormous range of experience, and measurements made over many years by gifted experimenters, and by way of a few simple principles bring these results into beautiful order. Those laws, in my view, mark our deepest penetration to date into the working of nature.

Beautiful, yes; marvellous in their summarizing power, yes; but depth of penetration--is that so clear? Isn't the quantum, the fact that no elementary phenomenon is a phenomenon until it is an observed phenomenon--the clear evidence that this universe of ours is in some strange sense a participatory universe--a far deeper discovery?

For me what counts is not words, but equations: we have nature boiled down into three laws, each with its own definite equation.

Then perhaps we should devote Review IV to the theme of "Many-Fingered Time, 'Imbeddability,' and the Laws of Physics."

Those equations that took the efforts of so many investigato so many years to work out can be derived today, all three of them, and in a few minutes, from the utterly simple requirement of Hojman, Kuchař and Teitelboim(67, 68,69,70) that the physics of the field should be "imbeddable" in spacetime. Never has one got so much physics from so little.

This is news to me--and exciting, too. I realize we can't get into technical details now. But can you at least give me the flavor of the idea--and appraise it?

Compare the three field theories with the theory of elastici For a homogeneous isotropic solid subject to a small strain--symmetry considerations tell us--there are only two ways to form an expression for the stored energy of second order in the strain. Either take the trace of the strain tensor and square it, or square the strain tensor and trace it. A linear combination of the two quantities with appropriately chosen coefficients gives the most general acceptab expression. By this simple line of reasoning we conclude that the elasticity of a homogeneous isotropic substance is characterized by exactly two elastic constants. If this example shows how much can be obtained from arguments of symmetry, it also illustrates that those symmetry considerat conceal from sight any view of the underlying machinery.(71,7 A hundred years of the study of elasticity would never have revealed that those elastic constants are formed by adding

the second derivatives of hundreds of complicated molecular potential energy curves, each multiplied by the appropriate direction cosines. And a hundred years of the chemistry of interatomic forces would never have revealed that these forces--and the "hundred laws" of chemistry--have their origin in something so fantastically simple as a system of positively and negatively charged masses moving in accordance with the Schrödinger wave equation. Symmetry is the quick road to the mathematics of law, but no road at all to the machinery behind law. That's why the work of Hojman, Kuchař and Teitelboim makes those three deepest laws of physics today no longer look so deep; rather, as "superficial" as elasticity is superficial. Deeper we must look if we would know in what soil those laws are rooted.

To constitute that soil, to compose that substrate, to serve as that primordial building "substance," what can we possibly propose today except the totality of elementary quantum acts of observer-participancy.

When you speak of the machinery underlying the great laws, are you suggesting that all of the important field equations have already been discovered?

Quite the opposite. Never more rapidly than today is progress being made in unravelling spinor field equations, as seen especially in that beautiful development of our times known as "supersymmetry."(74-77) Moreover, as bombarding energies go up, and distances probed in collision experiments go down, new effects will come to light which will provide--so distinguished colleagues in elementary particle physics suggest--evidence for still other fields of force, world without end.

I view that as an almost hopeless prospect.

Not at all! Cheer yourself up by remembering what it is to do a harmonic analysis of the tides. The more components we include in our Fourier analysis, the better fit we get to the past, and the better predictions of the future. No matter how many terms we include, however, they won't do a thing to forecast the splash of tomorrow morning's ship launching, or the tsunami from next week's earthquake. On the contrary, the more terms there are for dealing with th expected, the more prominently those features will show up which belong to the unexpected. Even to begin to include them in the bookkeeping requires one to go to a far more comprehensive form of analysis that altogether transcends the traditional treatment of the tides. When the number of Ptolemaic epicycles becomes too great, or the number of "elementary fields" too large, we have a compelling motive to look for a new paradigm.

In the topic you proposed for our Meeting IV you mentioned "many-fingered time," in keeping with your overall theme of "Frontiers of Time." But what about time itself, and spacetime? Are they primordial concepts? Or are they secondary and approximate?

Spacetime is a classical and approximate concept that utterly contradicts the uncertainty principle. Give up one or give up the other; you can't keep both.

I can't see how spacetime can possibly relate to the uncertainty principle, let alone violate it or be "approximate."

Then let us look at another, and simpler, classical concept that also violates the uncertainty principle: a world line. At every point along its length the world line attributes to the particle in question both a position and a velocity, or momentum. That degree of definiteness violates the uncertainty principle in the most evident way, would you not agree?

Of course. And I know what we do about it. We give up altogether the idea of a deterministic world line. In its place we speak of a wave function or probability amplitude.

You are right; but you have to go further and say more about what is right and what is wrong with the idea of a world line if later on you expect to see what is right and what is wrong with the idea of spacetime.

I don't see anything right about the idea of a world line.

Let us recall two features of a world line which you surely know and accept. First, the classical world line is a useful approximation to replace the quantum wave function you spoke about when the particle wave length is small compared to all other relevant physical dimensions. In this limit--

we agree--the predictions of "geometrical optics" closely model the predictions of "physical optics." Second, and rega less of these relative dimensions, the individual world line or history, H, of the particle's motion is the elementary building block in Feynman's prescription for the wave function.(78,79) He emphasizes what we may call "the democratic equality of all histories." The probability amplitude for one history is as great in magnitude as the probability amplitude for any other history. One of these probability amplitudes is differentiated from another only by its phase. The phase is given by the action integral for that history:

$$(3)\qquad (phase)=I_H/\hbar=(1/\hbar)\int_{x',t'}^{x'',t''} Lagrangian[x_H(t),\dot{x}_H(t),t]\,dt.$$

The total probability amplitude to transit from the original position x', at the original time t', to the final position x'', at the final time t'', is given by summing the elementary probabilities over all conceivable histories with equal weight for each; thus

$$(4)\qquad \langle x'',t''|x',t'\rangle = \int \exp(iI_H/\hbar)\, DH,$$

where DH is a suitably normalized measure over the "space" of all histories. This Feynman "sum over histories" or over world lines, and the principle of correspondence between world line and wave, are the two "rights" about a classical world line. To see what is "wrong" about a classical history, we have only to note the overwhelming preponderance of "unruly" histories(80) over smooth ones in the Feynman sum. The shorter the intervals between the times t_1, t_2, t_3, . . .t_i, t_n at which we specify the position, x_i to x_i+dx_i, of the particle, the greater are the zigzags in velocity in the histories which contribute most to the sum over histories. The more numerous and wilder the histories are that we are forced to consider, the clearer we become that the classical concept of "one history" is wrong.

What has all this to do with "spacetime"? "Spacetime" is the history of space geometry changing with time. "World line" is the history of particle position changing with time. What we have just said of the history of the particle inescapably applies to the history of space geometry. "History" is right, "spacetime" is right, for an approximate and semiclassical description of space geometry changing with time.(81) It is right too as a building block in the Feynman sum over histories to give the quantum description of the dynamics of geometry.(82-84) But a single classical history of space geometry, a single spacetime, is wrong; it is incompatible with any proper quantum description of the dynamics of geometry.

Why worry?

Because "spacetime" violates the uncertainty principle.

Take any deterministic classical spacetime, such as the Friedmann universe, the Schwarzschild geometry, or the Taub universe.(85) Make any spacelike slice whatsoever through it.(86) That slice assigns to space a definite 3-geometry. However, it also assigns to space a definite curvature with respect to the enveloping 4-geometry. That means a definite "extrinsic curvature;"(87) or, in the language of Hamiltonian field theory, a definite "field momentum."(88) Moreover, when we ascribe to space both a field coordinate--a definite intrinsic 3-geometry--and a field momentum--a definite extrinsic curvature, we collide head-on with the uncertainty principle. We can talk of "spacetime," as we talk of a "world line," but both are classical, anti-quantum concepts.

You agree that a "world line" ascribes to a particle both a coordinate and a momentum?

Yes.

And you agree that that is incompatible with the uncertainty relation?

I do.

And you concede that it is equally wrong to assign to a field both a coordinate and a momentum?

I must.

Then what escape is there from ruling out "spacetime" as

a deterministic classical concept, applicable only at the level of approximation theory?

I see no escape. But I would like to understand this matter better. Why not make it the subject of Meeting V?

Agreed. Let's give that discussion the title "Transcending Time." In giving up "spacetime" as a basic idea in the description of nature we have to give up "time," too; and with time gone, even the concepts of "before" and "after" lose all meaning.(89)

I feel completely lost. I have never heard anything in philosophy or logic that did not rest in the end, explicitly or implicitly, on the distinction between what comes first and what follows. How can I or anyone hope to make sense out of a nature in which the terms "before" and "after" have "lost all meaning"?

You have to recognize that we are discussing questions of principle. In all every day situations, and even in radiative processes and in the collisions of GeV particles, the relevant distances and times are enormous compared to the Planck distance,(90,91)

$$L^* = (\hbar G/c^3)^{\frac{1}{2}} = 1.6\times10^{-33}\text{cm}, \tag{5}$$

and the corresponding Planck time. Only at such small distances in present day geometry, or in the extreme geometry of big bang or collapse, do we expect to have to give up the idea of before and after. And why shouldn't we accept that limitation on our customary presuppositions, even welcome it? How else are we to come to terms with what Einstein's theory tells us? How else can we begin to understand that there is no such thing as a "before" before the big bang? no "after" after gravitational collapse? The only thing worse than having to give up "before" and "after" would be not having to give them up.

I can't understand how we can arrive at such limitations on our customary ideas of time starting from a theory--Einstein's geometrical theory of gravity--which accepts from the start the familiar local special-relativity theory distinction between past and future.

Be happy that we have sure and simple guides through these questions. We have not only Einstein's standard theory of the dynamics of geometry. We also have the standard principles of Hamiltonian dynamics, interpreted as we interpret them today in the light of quantum theory.(92) We shall need no more to see how and why "time" is transcended as a primary category in the description of nature.

Now I feel better prepared for Meeting V. What do you propose for Meeting VI?

"Initial Conditions and the Asymmetry in Time of Radiative Reaction." There we can come back to the idealization of flat spacetime and classical theory. There we may summarize the account of radiative reaction given by Wheeler and Feynman in 1945.(93,94) In it every charged particle is envisaged as coupled to every other charged particle by a field that is symmetric in time: half advanced, half retarded. Interconnections run forward and backward in time in such numbers as to make an unbelievable maze. That weaving together of past and future seems to contradict every normal idea of causality. However, when the number of particles is great enough to absorb completely the signal starting out from any source, then this myriad of couplings adds up to a simple result: the familiar retarded actions of everyday experience, plus the familiar force of radiative reaction with its familiar sign.

How can couplings symmetric in time add up to a result so obviously asymmetric in time?

Asymmetry in the boundary value data provide the explanation. The particles of the absorber are either at rest or in random motion before the acceleration of the source. They are correlated with it in velocity after that acceleration. Thus radiation and radiative reaction are understood in terms, not of pure electrodynamics, but of statistical mechanics.(95)

Don't I also understand why heat always flows from hot to cold and why entropy increases in terms of asymmetry in time in the boundary value data? In that reasoning don't I dispense with interactions propagated in time? Don't I idealize--and idealize with good results--to instantaneous couplings? Then why so much emphasis on half-advanced-plus-half-retarded interactions?

Our emphasis is on the directly opposite point: We need asymmetry in time of boundary value data to understand the asymmetry in time that we see in nature, regardless whether the elementary time symmetric interactions are idealized as instantaneous or are propagated in time.

I am happy with the perspective you give me; happy, that is, with all except one point, and I fear it is an absolutely central point. Why should it be initial value data that are specified in statistical mechanics? Why not final value data?

Your question couldn't be more appropriate. You put your finger on one of the great mysteries. It is even conceivable that we can't make any headway in answering your question until we finally begin doing statistical mechanics in a proper cosmological setting. Why then don't we make "Asymmetry in Time and the Expansion of the Universe" the topic of Meeting VII? It will suggest some observations and measurements.

I look forward to that topic.

And while we are on mysteries let us discuss in Meeting VIII

another: "Memory." How does it come about that we remember the past but not the future? Is this asymmetry in time a consequence of and witness to the "observer-participancy" that we would make the underpinning for all the laws of physics? It is not necessary for us to have answers to raise questions.

Cosmological issues are so central to all you have to say that before you end I would like to hear more about the big bang, the big crunch and the black hole--what you call(96) "The Gates of Time."

Then let's make that the topic for a final Meeting IX. That will bring to a natural close our survey of some of the "Frontiers of Time." Nothing indicates more clearly than those gates of time that the universe did not exist forever. No evidence gives more incentive to conceive of the laws of physics as having come into being. None suggests more forcefully that proud unbending immutability is a mistaken ideal for physics; that this science now shares, and must forever share, the more modest mutability of its sister sciences, biology and geology.

A new species of bird may appear unbelievable. The upended strata of a mountain slope may look incredible. Yet both biology and geology find their explanation in the accumulative consequences of many individual small effects. Today we do not abandon reason when we regard the kingdom of life, rich though it is, or when we look up at the Himalayas, tall though they stand. How these wonders came about we now understand in outline, and count on someday being able to describe in detail. Have equal confidence that we shall find out how the laws of physics--and the universe--came into being, incredibly remote though they today seem from being also the accumulative consequences of many individual small effects.

Small effects? Accidents? Accidents like mutations, or like the rainstorms that wear away mountains? Blind accidents?

We have to be careful with that word "blind accident." "Blind" implies blind towards future consequences. It suggests a happening that is rooted in the past and heedless of the future. Such a conception implies that an order in time is already in being. The direct opposite is lesson number one of our survey of "Frontiers of Time."

Time, we discover, is not a primordial concept in the structu of nature. It is secondary and derived. So too, it would appear, is the asymmetry between past and future that shows up so strikingly in radiative reaction, in the flow of heat from hot to cold, in biological evolution, and even in the mechanism of the memory. In contrast, how can any elementary building process be an elementary process for building existence and law unless it transcends the category of

time?

To identify an elementary building process that transcends time: is that why you put "The 'Past' and the 'Delayed Choice' Double-Slit Experiment" ahead of all other topics on our list of meetings?

Yes. The act of observer-participancy in such an experiment, right now, irretrievably alters what we have the right to say about "the past." In that sense, that carefully restricted sense, that act is an inescapable part of the actual building of "the past."

I begin to realize that not only topic II, but all the topics on our list of meetings make time their central concern. How did this come about, when the original focus of our discussion was "How did the universe come into being, and what is its substance?"

The answer is simple. We don't understand genesis and we never will until we rise to an outlook that transcends time. That is why a review of frontiers of time is precondition for any proper analysis of the ultimate issue.

Thanks! With your permission I plan to bring colleagues to our further meetings, even if that forecloses most of the questions I would like.to ask along the way. However, I worry lest they miss the bearing of "time" on "genesis." Therefore may I ask if for them--and me--you would please boil down into a few lines that I can copy the gist of what you've said today?

We would do better to have no summary at all than a summary so short we cannot analyze for each point the evidence, whether weak or strong.

I understand your concern. Nevertheless, please put everything in a dozen brief points. Leave it to me to supply later, in the light of what you have said today, the qualifications and caveats I know you would want.

Then let us try.

As surely as we now know how tangible water forms out of invisible vapor, so surely we shall someday know how the universe comes into being. We will first understand how simple the universe is when we recognize how strange it is.

The simplicity of that strangeness, Everest summit, so well directs the eye that the feet can afford to toil up and down many a wrong mountain valley, certain stage by stage to reach someday the goal.

Of all strange features of the universe, none are stranger than these: time is transcended, laws are mutable, and observer-participancy matters.

"Before" and "after" don't rule everywhere, as witness quantum fluctuations in the geometry of space at the scale of the Planck distance. Therefore "before" and "after" cannot legalistically rule anywhere. Even at the classical

level, Einstein's standard closed-space cosmology denies all meaning to "before the big bang" and "after the big crunch." Time cannot be an ultimate category in the description of nature. We cannot expect to understand genesis until we rise to an outlook that transcends time.

There never was a law of physics that did not require space and time for its statement. With collapse the framework falls down for everything one ever called a law. The laws of physics were not installed in advance by a Swiss watchmaker, nor can they endure from everlasting to everlasting. They must have come into being. They could not always have been accurate. They are derivative and superficial, not primary and revelatory.

Quantum physics teaches that no elementary phenomenon is a phenomenon until it is an observed phenomenon. The "delayed-choice experiment" shows that an act of observer-participancy in the present has an irretrievable consequence for what one can say--with the help of theory--about the past.

Conformant to these three strangenesses, how else can the universe come into being except as a "self-excited circuit?" As it expands, cools and develops, it gives rise to observer-participancy. Observer-participancy in turn gives what we call "tangible reality" to the universe.

"Omnia ex nihil ducendis sufficit unum"--one principle suffices to build everything from nothing.

From what kind of nothingness?

"Nothingness" is not the vacuum of physics, loaded with geometry and field fluctuations; it is a nothingness devoid of structure, law or plan; it is the zeroness of existence of that word "cloud" at the beginning of the surprise-version game of twenty questions.

Build how much out of nothingness?

Law; and spacetime as part of law; and out of law substance.

Build law out of the statistics of billions upon billions of acts of observer-participancy each of which by itself it utterly random. Recognize law as the accumulative consequence of many individual small effects. How else could law come into being?

No test of these views looks more like being someday doable, nor more interesting and more instructive, than a derivation of the structure of quantum theory from the requirement that everything have a way to come into being out of nothing.

If you would have an epitome of this summary, let it be this: Nothing. No time. The line. Acts. Statistics. Law. Spacetime. Substance. Observer-participant Closed circuit. Test. But all our further meetings, we have agreed, will focus on one part of this larger theme: on time, and what it means to transcend time.

II. The "Past" and the "Delayed-Choice" Double-Slit Experiment.

"Reality is theory."

-Torgny Segerstedt.

". . .the past has no existence except as it is recorded in the present.(59)*"*

[The following is abbreviated from ref. (59).]

Partway down the optic axis of the traditional double-slit experiment stands the central element, the doubly-slit screen. Can one choose whether the photon (or electron) *shall have* come through both of the slits, or only one of them, after it has *already* transversed this screen? That is the new question raised and analyzed here.

Known since the days of Young is the possibility to use the receptor at the end of the apparatus to record well defined interference fringes. How can they be formed unless the electromagnetic energy has come through both slits? In later times Einstein noted that in principle one can determine the lateral kick given to the receptor by each arriving quantum. How can this kick be understood unless the energy came through only a single slit?

Einstein's further reasoning as reported by Bohr (97) is familiar. Record both the kicks and the fringes. Conclude from the kicks that each quantum of energy comes through a single slit alone; from the fringes, that it nevertheless also comes through both slits. But this conclusion is self-contradictory. Therefore quantum theory destroys itself by internal inconsistency.

Bohr's reply(97) has become by now a central lesson of quantum physics. One can record the fringes or the kicks but not both. The arrangement for the recording of the one automatically rules out the recording of the other. The quantum has momentum p, de Broglie wave length $\lambda = h/p$, and reduced wave length $\bar{\lambda} = \hbar/p$. To record for it well defined interference fringes one must fix the location of the receptor within a latitude

$$(6) \qquad \Delta y < (\text{fringe spacing})/2\pi = (L/2S)\bar{\lambda}.$$

To tell from which slit the quantum of energy arrives one must register the transverse kick it gives to the receptor within a latitude small enough to distinguish clearly between a momentum $p = \hbar/\bar{\lambda}$ coming from below, at the inclination S/L, and a momentum coming from above at a like inclination; thus,

$$(7) \qquad \Delta p_y < (S/L)(\hbar/\bar{\lambda}).$$

However, for the receptor simultaneously to serve both functions would be incompatible with what the principle of indeterminancy has to say about receptor dynamics in the y-direction,

$$(8) \qquad \Delta y \Delta p_y > \hbar/2.$$

Not being able to observe simultaneously the two complementary features of the radiation, it is natural to focus

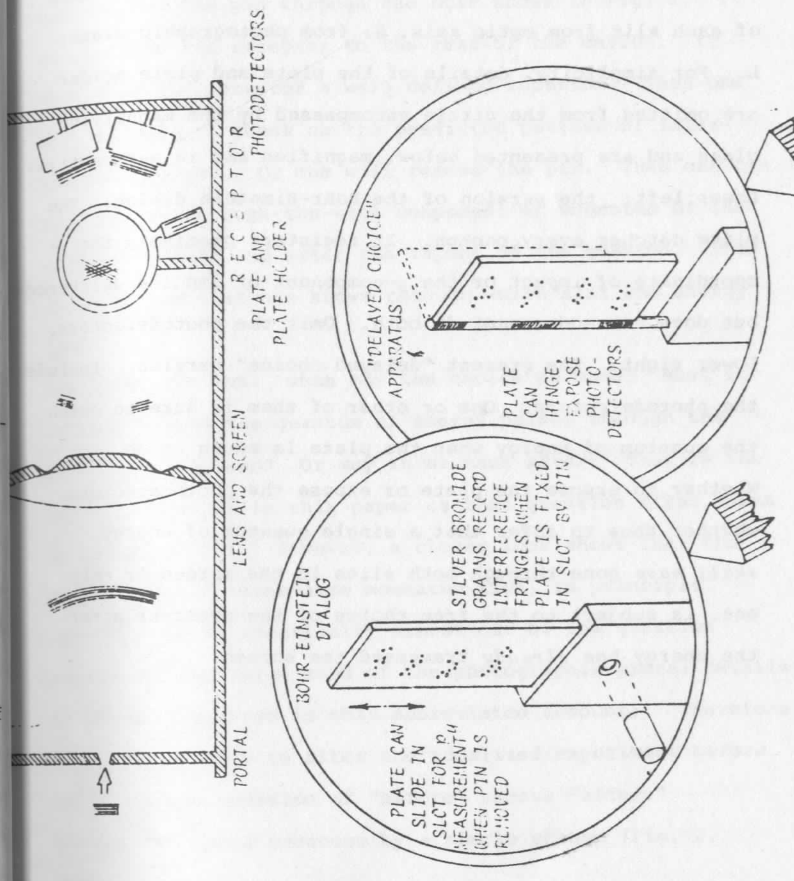


FIGURE 2.

Fig. 2. Top: Idealized double-slit experiment. Distance of each slit from optic axis, S; from photographic plate, L. For simplicity, details of the plate and plate holder are omitted from the circle encompassed by the magnifying glass and are presented below, magnified and in perspectiv Lower left: the version of the Bohr-Einstein dialog. The plate catches every photon. It registers precisely the y-coordinate of impact or the y-component of impulse delivere but does not and cannot do both. Omit the photodetectors. Lower right: The present "delayed choice" version. Inclu the photodetectors. One or other of them is sure to catch the quantum of energy when the plate is swung aside. Whether to expose the plate or expose the photodetectors, whether thus to infer that a single quantum of energy *shall have* gone through both slits in the screen or only one, is subject to the free choice of the observer after the energy has *already* traversed the screen.

on the one and forego examination of the other. Either one will insert the pin through the hole shown in Fig. 2. It will couple the receptor to the rest of the device. It will give the receptor a well defined location. Then one will be able to check on the predicted pattern of interference fringes. Or one will remove the pin. Then one can measure the through-the-slot component of momentum of the receptor before and after the impact of the quantum. Then one will say that one knows through which slit the energy came.

Pin in or pin out: when may the choice be made? Must it be made before the quantum of energy passes through the doubly slit screen? Or may it be made after? That is the central question in this paper as that question first seems to impose itself. However, a closer look shows that the measurement of transverse momentum kick, in principle conceivable, is practically almost out of the question [because of the large mass of the photographic plate; details of analysis omitted in this abbreviated account]. Therefore it is appropriate to alter the idealized experiment before taking up the question of "before" *versus* "after." The difficulty is overcome by a simple change [Fig. 2, lower right]:

(1) Give up measuring the y-component of the momentum of the photographic plate.

(2) Hold its y-coordinate fixed.

(3) By means of a hinge parallel to the y-axis arrange that this high narrow plate can be swung out of the way of the incident light--at the last minute option of the obser quicker than the flight of light from screen to plate. (Switch from "operative" to "open" position.)

(4) Sufficiently far beyond the region of the plate, the beams from upper and lower slits cease to overlap and becc well separated. There place photodetectors. Let each hav an opening such that it records with essentially 100 perce probability a quantum of energy arriving in its own beam, and with essentially zero probability a quantum arriving i the other beam.

Now the choice is clear; and the objective, too. We today cannot argue, and Einstein in his later years would not even have wanted to argue, his erstwhile case of logical inconsistency against quantum theory: the photon goes through both slits, as evidenced in interference fringes, and yet simultaneously through only one, as evidenced in lateral momentum kick. Choose we know we must between the two complementary features open to study; and choose we do by putting the plate athwart the light or turning it out o the line of fire. In the one case the quantum will transform a grain of silver bromide and contribute to the recor of a two-slit interference fringe. In the other case one

the two counters will go off and signal in which beam--and therefore from which slit--the photon has arrived.

In our arrangement the photographic plate registers only the point of impact of a photon. In the earlier idealized experiment it could additionally (Einstein) or alternatively (Bohr) record the transverse momentum delivered by the impact. We have assigned the two distinct kinds of measurement to two distinct kinds of register. We have demoted the plate from a privileged status. That demotion is irrelevant to any question now at issue. Equally irrelevant is the different distance--and time of flight--from entry portal to plate, or photodetector, according as the one or other register is exposed. But the essential new point is the timing of the *choice*--between observing a two-slit effect and a one-slit one--until after the single quantum of energy in question has *already* passed through the screen.

Let the reasoning be passed in review that leads to this at first sight strange inversion of the normal order of time. Then let the general lesson of this apparent time inversion be drawn: "No [elementary] phenomenon is a phenomenon until it is an observed phenomenon." In other words, it is not a paradox that we choose what *shall* have happened after "it has *already* happened." It has not really happened, it is not a phenomenon, until it is an observed phenomenon.

Whatever we now do to spell out the otherwise idealized experiment, we will leave idealized its most unusual feature, the "swinging door photographic plate." That term includes the arrangement, whatever it may be,

(1) for a last minute choice, to swing the plate aside or leave it athwart the beam, after the arriving energy has already traversed the doubly slit screen, and

(2) for completion of that movement before the energy arrives at the plate. In practice it will be more reasonable to swing the beam than swing the plate. Fix the plate. Halfway from screen to it, position a Kerr cell.(98) Apply to it a positive or a negative voltage according as one wishes to record fringes on the plate, or register "which beam" on a counter. Or, still better, Manfred Fink suggests, replace the experiment with the photon by an experiment with an electron. Then the last-minute deflection of the electron beam can be accomplished by a localized magnetic field centered between screen and plate. One or another of these arrangements to swing the beam will be understood hereafter to apply in practice when in principle we speak of swinging the plate.

[Other requirements and presuppositions of the idealized experiment are analyzed in the original publication but these details are omitted here. Also omitted here are six other types of "delayed choice" experiments.]

. . . .

The double slit experiment, like the other six idealized experiments (microscope, split beam, tilt-teeth, radiation pattern, one-photon polarization, and polarization of paired photons), imposes a choice between complementary modes of observation. In each experiment we have found a way to delay that choice of type of phenomenon to be looked for up to the very final stage of development of the phenomenon, *whichever* type we then fix upon. That delay makes no difference in the experimental predictions. On this score everything we find was foreshadowed in that solitary and pregnant sentence of Bohr(99), ". . .it. . . can make no difference, as regards observable effects obtainable by a definite experimental arrangement, whether our plans for constructing or handling the instruments are fixed beforehand or whether we prefer to postpone the completion of our planning until a later moment when the particle is already on its way from one instrument to another."

Not one of the seven delayed choice experiments has yet been done. There can hardly be one that the student of physics would not like to see done. In none is any justification whatsoever evident for doubting the obvious predictions.

We search here, not for new experiments or new predictions, but for new insight. Experiments dramatize and predictions spell out the quantum's consequences; but what is its central idea? A pedant of Copernican times could have calculated planetary positions from the equations of Copernicus as well as Copernicus himself; but what would we think of him if his eyes were closed to the main point, that the "Earth goes around the Sun?"

. . . .

[No analysis] in recent times moved our understanding forward more than the Einstein-Bohr dialog(97). Out of that dialog no concept emerged of greater fruitfulness tha "phenomenon"(100): ". . .[In my discussions with Einstein] I advocated the application of the word *phenomenon* exclusi to refer to the observations obtained under specified circumstances, including an account of the whole experimen arrangement."(101) No other point does the present analys of idealized delayed-choice experiments have but to investi what "phenomenon" means as applied to the "past."

After the quantum of energy has *already* gone through the doubly slit screen, a last-instant free choice on our part--we have found--gives at will a double-slit-interference record or a one-slit-beam count. Does this result mean that present choice influences past dynamics, in contravention of every formulation of causality? Or does it mea calculate pedantically and don't ask questions? Neither; the lesson presents itself rather as this, that the past has no existence except as it is recorded in the present. It has no sense to speak of what the quantum of electromag

energy was doing except as it is observed or calculable from what is observed. More generally, we would seem forced to say that no [elementary] phenomenon is a phenomenon until--by observation, or some proper combination of theory and observation--it is an observed phenomenon.

. . . .

III. "Development in Time" Gives Way to "Correlation in Time."

". . . it appears that our theory denies the existence of absolute reality--a denial which is unacceptable to many. . . I do not know how one could define operationally the reality of anything."

- E. P. Wigner(102)

Most instructive of all the idealized experiments considered in the great dialog between Bohr and Einstein was the Einstein-Podolsky-Rosen experiment,(104,105) later simplified by Bohm(106) to the version illustrated at the middle of Fig. 3. The very light isotope of hydrogen composed of one positive and one negative electron is allowed to cascade down to its ground state of 0 angular momentum. There it sits until it undergoes annihilation. Two photons come off with equal and opposite momenta, as illustrated by the two wavy lines in the diagram. An observer on the right determines whether the photon travelling to the right is circularly polarized to the right or to the left. Whatever the result, he is assured that a measurement of the circular polarization of the left-hand photon will give exactly that result, right-handed or left-handed, which is required for conservation of angular momentum.

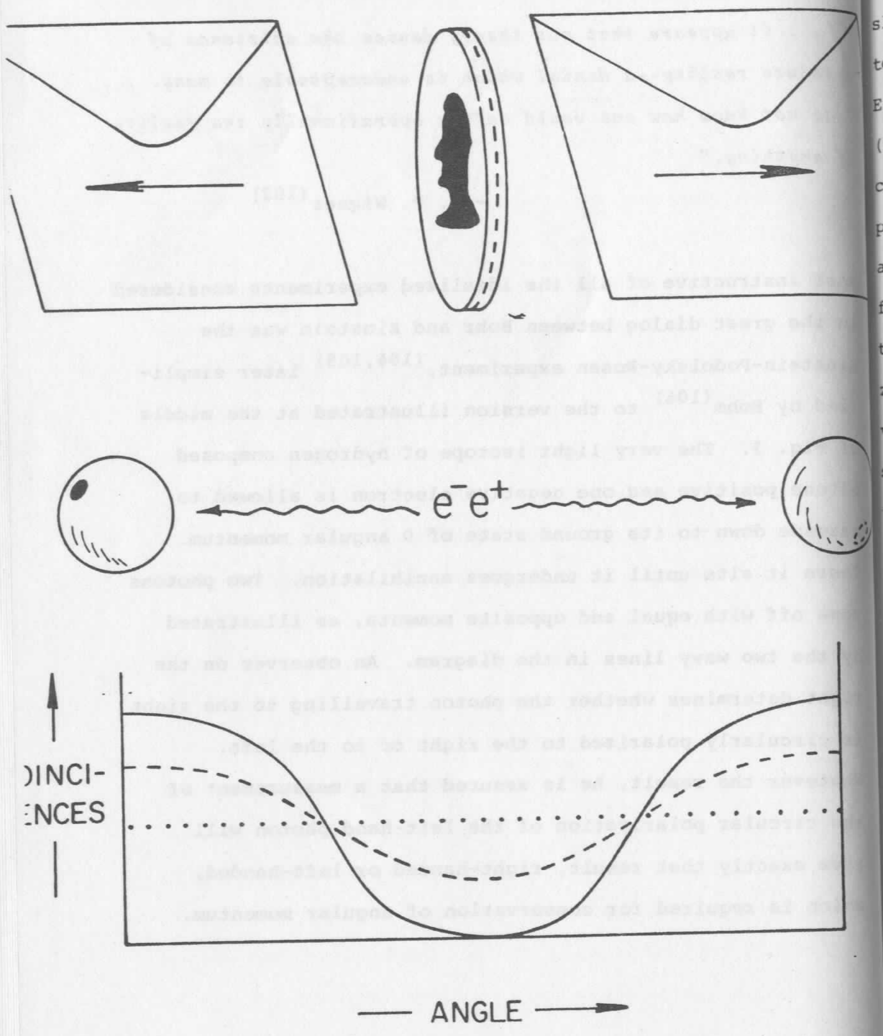


FIGURE 3.

Fig. 3. In contrast to splitting a coin, putting the two slices into envelopes, shuffling them, and sending them to two remote observers (upper part of diagram) the Einstein-Podolsky-Rosen experiment in the version of Bohm (middle part of diagram) permits a double infinity of choices (point on right-hand Stokes sphere) for the polarization to be looked for in the right-hand (e^+e^-) annihilation photon, with corresponding consequences(108) for the polarization (point on left-hand Stokes sphere) that will be found for the left-hand photon. If the polarizations were determined in the act of emission ("hidden variables") the coincidences between the two photons would show only half the dependence on relative orientation of the two polarizations (dashed curve in lower diagram) that is predicted by quantum mechanics--and observed (full curve).

Alternatively, he may choose to study the photon going to the right with the help of an analyzer of linear polarizatio Then he makes a clean measurement as to whether the polarization lies in the y direction (or the z direction). Then he is assured that a study by similar equipment of the photon travelling to the left will show it to be vibrating with 100% certainty in the z direction (or the y direction).

At first sight there is nothing very startling about the correlation in polarization between the two photons. What difference in principle is there, one might well ask, with respect to the old game in which a coin is sawed in half? The two halves are put in separate envelopes and sealed and dispatched to observers far away on the left and the right. If the observer on the right opens his envelope and finds the head in it then he knows that the other far away observer will find the tail of the coin when he opens his envelope. There is no paradox involved. There is no possibility of using the arrangement of envelope to send a signal in excess of the speed of light.

In the e^+e^- annihilation, the new feature is this: The polarization of a photon is a more sophisticated quantity than the differences between the two faces of a coin. According to Stokes' parametrization of polarization [see for example Born and Wolf(107)] each of the many alternative

ways to specify a well defined polarization can be set into one to one correspondence with a point on the surface of the unit sphere. The observer on the right can delay until the very last picosecond before the arrival of his photon, the determination of which kind of polarization he will look for, as symbolized by point B on the right-hand sphere. Whatever the choice--and he cannot, of course, know whether the photon that arrives will have the polarization B or the polarization "anti-B"--he is assured that the running of the corresponding experiment on the other photon will give for it the uniquely mated polarization [see Kagali(108) for the coincidence rate for the general case of arbitrary polarization].

What is to be said of the polarizations of the two photons in the course of all their long travel from the site of the e^+, e^- annihilation to their respective points of reception? Nothing. Nothing until the experiment is over. No elementary phenomenon is a phenomenon until it is an observed phenomenon.

Instead of accepting this lesson of the quantum one can try to quarrel with it. Why not assign a probability amplitude to the state of the two photons, as for example $[\alpha(1)\beta(2)-\alpha(2)\beta(1)]/2^{\frac{1}{2}}$? Why not go further and view the process of measurement in a Lorentz frame in which the right-hand photon arrives first at its analyzer-detector? Is the right-hand photon not suddenly recorded as having, for example, the polarization α(1)? Does it not follow that the left-hand photon, still en route from the site of annihilation to the left-hand analyzer-detector, must suddenly in mid-course be redescribed as having the mated polarization β(2)? Does not this redoing of the state function imply the existence of an effect propagated from right to left in excess of the speed of light? Then look at the whole process all over again in a Lorentz frame in which the left-hand photon arrives first at the left-hand analyzer-detector. By the same reasoning is one not led to speak of an effect propagated this time from left to right in excess of the speed of light? What a confusion! What a warning not to identify these pencil-and-paper readjustments, these pencil-and-paper supra-light-velocity effects, with anything real. What an indication that the wave function is not itself real, but a purely formalistic device, and within the present incomplete marriage of quantum theory and relativity not a very happy device, for calculating the probability of real coincidences. Only when the counters have gone off has the reality of the situation declared itself. No elementary phenomenon is a phenomenon until it is an observed phenomenon.

Continue to contest this lesson of the quantum. Argue that a reality *is* to be attributed to the polarizations of

the two photons on their way towards the two detectors, regardless of the settings of the two analyzers. Declare that the chance for the first detector to go off is

$\cos^2$(angle between the "true direction" of the polarization of photon 1 and the setting of analyzer 1),

and the chance for the second detector to go off is

$\cos^2$(angle between the "true direction" of the polarization of photon 2 and the setting of analyzer 2).

Argue that the chance for a coincidence is the product of these two expressions averaged over the random direction of the polarization of one of these two photons--the second being of necessity orthogonal to the first. In this way end up with the dashed curve of Fig. 3 for number of coincidences as a function of the relative setting of the analyzers on left and right. For the difference in rate between the least favorable and the most favorable setting one gets only half what is predicted by quantum mechanics and only half of what is observed [Polarization of e^+e^- annihilation photons: theory, Wheeler (1946)(109), Pryce and Ward (1947)(110), Snyder, Pasternack and Hornbostel (1948)(111); observations by Bleuler and Bradt (1948)(112), Hanna (1948)$^{(113)}$, Vlasov and Dzeljepov (1949)$^{(114)}$, Wu and Shakhov (1950)$^{(115)}$, Hereford (1951)$^{(116)}$, Bertolini, Bettoni and Lazzarini (1955)$^{(117)}$, Langhoff (1960)$^{(118)}$, Kasday, Ullman and Wu (1970)$^{(119)}$, Kasday (1971)$^{(120)}$, and Faraci, Gutkowski, Notarrigo and Pennisi (1974)$^{(121)}$, Wilson, Lowe and Butt (1976)$^{(122)}$; polarization of the photons given out by an atom in a 2-step transition, Kocher and Commins (1967)$^{(123)}$; Freedman and Clauser (1972)$^{(124)}$; Holt (1973)$^{(125)}$ result in contradiction with quantum predictions but not confirmed by Clauser (1976)$^{(126)}$; Freedman and Holt (1975)$^{(127)}$; and Fry and Thompson (1976)$^{(12}$ Quantum mechanics thus "exposes itself to destruction" in numerous decisive tests--and stands up to these tests. It is a central point of this quantum mechanics that it denies to photons any "real" polarizations merely in virtue of their being "on their way" and in default of any actual act of observation. In other words, an elementary phenomeno is a phenomenon only when it is an observed phenomenon.

If "development in time" of "the wave function" is not a happy way to describe the state of the two photons in the EPR-experiment, how should one describe this and more comple situations, in which observations are made at several or even many, locations in space time? Correlation of observations: this is the appropriate concept, according to Houtappel, Van Dam and Wigner$^{(65)}$ and Wigner$^{(102)}$.

In this approach one makes a conceptual reformulation of the equations of quantum mechanics, "eliminating explicit reference to the equations of motion and to state vectors. According to this [philosophy], the function of quantum mechanics is to give statistical correlations between the outcomes of successive observations."$^{(102)}$ As an example, Wigner considers the correlation between two measurements. In the first measurement the physical quantity under examination is described by some operator Q. The various possible outcomes for this measurement are labelled by an index j. Associated with the jth outcomes is a projection operator P_j, with $P_j^2=P_j$. Both Q and the P_j are envisaged to vary with time in accordance with the same law,

$$Q(t) = e^{iHt/\hbar}Q(o)e^{-iHt/\hbar}, \tag{9}$$

$$P_j(t) = e^{iHt/\hbar}P_j(o)e^{-iHt/\hbar}, \tag{10}$$

in which H is the Hamiltonian operator. As Wigner notes, we attribute "the operator $Q(t)$ to the same measurement, carried out at time t, to which we attribute the operator $Q(o)$ if carried out at time o," and $P_j(t)$ is "the projection operator which leaves the state vectors of outcome j unchanged, annihilates the state vectors of all other measurement outcomes." A measurement of quite a different physical quantity at quite a different time has associated with many quite different possible outcomes, of which the kth is associated with the projection operator P'_k. "The probability that the second measurement yields the result k if the first one's outcome was i is then given by [the ratio of traces]

$$(11)\quad (probability) = Tr(P'_k P_j)/TrP_j = Tr(P_j P'_k P_j)/TrP_j;$$

and similar expressions [see Houtappel, Van Dam and Wigner Eqs. (4.4-4.7)] can be given for the probabilities of the different outcomes of several successive measurements."

This formalism replaces the older view of dynamics as development in time by a proper quantum concept of correlation as it depends upon time. It does not go the whole way towards what is eventually envisaged under the heading of "law without law." First, the correlation treatment does not analyze, nor was it aimed at analyzing, down to the substrate of observation the ultimate make-up of the particle or field under study. It takes the existence of the dynamic entity under study as for granted. Second, it takes the concept of time for granted. It does not transcend the concept of time nor was it intended to. However, it has the great merit of providing a first framework for working towards an ultimate statistics of

---

"billions upon billions of acts of observer-participancy." At that point the idea will come closer to testability, "Is everything--including time--built from nothingness by acts of observer-participancy?"

IV. Many-Fingered Time, "Imbeddability," and the Laws of Physics.

> *"....according to the remark which I formerly made on the occasion of an optical law....final cause....even in physi serves to find and to discover hidden truths."*
>
> --G. W. Leibniz(129)

The more one learns about the laws of physics, the more one learns how little one has learned. Maxwell electrodynamics, Einstein geometrodynamics, and the chromodynamics of the Yang-Mills quark-binding field, laws won through decades of effort, summarizing an unbelievable richness of experience, and representing our deepest penetration to date into the machinery of nature, spring out, all three, full-bodied, at a single simple Alladin-like command, "let final values be independent of the choice of many-fingered time."

It seems at first sight almost unbelievable that so much hard-won experience should be deducible from a demand so simple. Therefore it may be appropriate to spell out a bit more fully this beautiful discovery of Hojman, Kuckař and Teitelboim (hereafter abbreviated as HKT).(67,68) Specify the initial value of the field in question and its conjugate momentum on an arbitrary smooth initial spacelik

hypersurface σ_1. From these initial data and from the Hamiltonian, H, of the field calculate the final value of the field and its conjugate momentum on an arbitrary smooth final spacelike hypersurface σ_2. Do the calculation over and over again moving forward from σ_1 to σ_2 by all arbitrary choices of many-fingered time (Fig. 4). Demand that the final values on σ_2 shall be the same for all these different ways of marching forward from σ_1 to σ_2. HKT, recognizers and exploiters of this requirement, call it the demand for "imbeddability." If it were violated the dynamics could not be imbedded in a single spacetime manifold. "Imbeddability" is the magic word that summarizes and delivers forth almost all we know of the laws of nature.

Following this brief overview, it is appropriate to come back for a little more detail, especially on "time." "Unfolding in time" being the essence of dynamics, it is natural that changing views of time have led to changing concepts of dynamics. The progress from Newtonian time to the time of special relativity, and from that to the many-fingered time of general relativity, was essential precondition for the discovery of H-K-T.

Time is absolute and universal in Newtonian physics. In special relativity successive times correspond to successive slices through spacetime. These slices are parallel to one another and normal to the time axis of the particular observer in question. For another observer there is another time axis and normal to it a different set of parallel flat spacelike hypersurfaces. Thus for each observer, each inertial frame, there is a globally defined time. Dynamics may be described in one of these global Lorentz frames, or another, or another; but in any one frame it is described with respect to a single time variable.

In 1932 Dirac, Fock and Podolsky(130) introduced a way of analyzing particle dynamics in which there are as many time parameters as there are particles. These time parameters are at the disposition, not of the particles, but of the analyst. He picks up business-machine printout cards telling what each particle has been doing. Nobody can keep him from placing on his desk the card that tells what particle 1 was doing at a particular time t_1, and alongside it a card that tells what particle 2 was doing at a particular time t_2, and correspondingly other cards from other times t_3, t_4, ...t_n in the lives of the other particles. Not to have any of these particles in the zone of influence of any other of these particles it is useful to impose the requirement that the chosen "events" on the several world lines should have a spacelike relation each to all the others.

This concept of what we may call a "many-fingered time" Tomonaga(131) generalized in 1946 from the dynamics of n

particles in flat spacetime to the dynamics of the electromagnetic field in flat spacetime. The field is conceived to be studied in its dependence, not on one time parameter, not on n time parameters, but on as many time parameters,

$$t = t(x,y,z), \tag{12}$$

as there are points in space--which is to say, a continuous infinity of time parameters. There is a simple way to visualize this collection of parameters. It constitutes a hypersurface, a slice through spacetime, what Landau and Lifshitz(132) call a "simultaneity." Generalizing from an arbitrarily curved, bent, or wiggly slice through flat spacetime to an arbitrary slice through the curved spacetime of general relativity, we impose the same kind of requirement that Dirac, Fock and Podolsky did. We require each point on this hypersurface to stand in a spacelike relationship to all the other points. None is to be able to send a signal to, or exert a force at, the others. In this sense the hypersurface is "spacelike." To demand the existence of such a global spacelike hypersurface is a powerful condition. Gödel's model of a rotating universe with closed timelike lines(133) does not satisfy this requirement.

On this account that spacetime is generally regarded as non-physical. It we therefore exclude from consideration along with every spacetime that does not admit global spacelike hypersurfaces.

In classical physics the electromagnetic field has a determ istic evolution in spacetime. What does this mean for the description of this field in terms of many-fingered time? Nothing startling. Pick the spacelike hypersurface. Pick one point in the 3-dimensional space thus defined. Erect at that point the unique timelike unit vector normal to the local tangent hypersurface. With respect to that vector and that local tangent 3-space the electromagnetic field falls apart into the magnetic field, $\underset{\sim}{B}$, and the electric field, $\underset{\sim}{E}$, both 3-vectors located in the local tangent 3-space. The magnetic field $\underset{\sim}{B} = \nabla\times\underset{\sim}{A}$ or, better, the vector potential $\underset{\sim}{A}$ from which $\underset{\sim}{B}$ lets itself be derived, thus specified from point to point throughout the spacelike hypersurface, may be regarded as the electromagnetic field coordinate; the electric field, divided by 4π, as the electromagnetic field momentum, $\underset{\sim}{\pi} = \underset{\sim}{E}/4\pi$, in a canonica Hamiltonian description of the electromagnetic field.

It is enough to give $\underset{\sim}{B}$ and $\underset{\sim}{E}$ as initial data on an initial spacelike hypersurface, and to know the Hamiltonian density for Maxwell's field,

$$H = (\underset{\sim}{B}^2+\underset{\sim}{E}^2)/8\pi$$

$$= (1/8\pi)(\nabla\times\underset{\sim}{A})^2+[(4\pi)^2/8\pi]\underset{\sim}{\pi}^2, \quad (13)$$

to be able to predict how the field changes with changes in many-fingered time as the hypersurface is pushed slowly forward (or backward) in time. We will not write down the necessary Hamiltonian equations; they can be imagined.

Why the particular Hamiltonian (13)? Why not some other Hamiltonian, some other law of physics? It provides only a partial answer to this question to turn back to Hilbert's famous paper of 1915(134). He derived electrodynamics and general relativity, or vacuum geometrodynamics--and the combined theory of the two field together--by postulating the simplest action principle that depends on a 4-dimensional vector field

$$A_\alpha \quad (14)$$

or on a 4-dimensional metric field

$$g_{\mu\nu} \quad (15)$$

or on the combination of the two. But why the simplest action principle? Why not some one of the thousand and one alternative action principles that contain these two fields in some other invariant combination? "Imbeddability" is the new and magic and beautiful answer that Hojman, Kuchař and Teitelboim(67,68) [see also Wheeler(135), Teitelboim(136,137,69), and Kuchař(138,139)] give to this old question. They envisage a 3-vector field

(16) $A_i(i=1,2,3)$ *(and its conjugate momentum)*

(17) $g_{jk}(j,k=1,2,3)$ *(and its conjugate momentum)*

or both (and their conjugate momenta). Whatever the dynam law that governs the evolution of these fields with time, as many-fingered time is pushed forward from the spacelike hypersurface σ_1 to the spacelike hypersurface σ_2 in Fig. 4, that law must give the same result for the dynamic variable whether this hypersurface is pushed forward first more rapidly on the "right" and then more rapidly on the "left or first more rapidly on the "left" and then more rapidly on the "right." If the conditions obtained at σ_2--by step by step forward integration of the Hamiltonian field equations on an electronic computer--depended upon the choice of history adopted in proceeding from σ_1 to σ_2, then the history of the fields could not be imbedded in any single spacetime manifold. With "independence of history" lost, imbeddability would also be lost.

No local Hamiltonian law for the development with many-fingered time of a vector field (16) will satisfy this condition of imbeddability except Maxwell's theory. No local Hamiltonian for the development with time of space

geometry (17) will give a history-independent result except Einstein's general relativity. No local Hamiltonian for a vector field with an "internal-spin" degree of freedom

(18) $A_i^{(S)}(i=1,2,3)$

is compatible with imbeddability except the Yang-Mills theory(66), today's standard and widely accepted theory of the quark-binding field. Thus simply derived from almost nothing are electromagnetism, gravitation and the current theory of the forces that hold elementary particles together, theories that summarize an unbelievable wealth of experience, years of experimentation, and the life work of some of the most gifted men of the last two centuries. No one has ever seen a simpler or more compelling theme than "imbeddability" to summarize the requirements that lead to physics as we know it.

Not only the Hamiltonian, but the gauge features of field theory follow from the argument of imbeddability. As Teitelboim showed(137), "$\underset{\sim}{A}$ is not observable but only its curl, $\underset{\sim}{B} = \nabla\times\underset{\sim}{A}$ is, because the propagation of $\underset{\sim}{A}$ itself will not be in general integrable. Thus the gauge transformation

(19) $A_i \to A_i+\partial\Lambda/\partial x^i$

must have no physical effect: Our efforts for preserving path-[history-]independence have led us to gauge invariance. He finds a similar result for the dynamics of the metric: No dynamics of the g_{jk} is imbeddable in spacetime unless the theory is left unchanged by a coordinate transformation.

The significance of gauge invariance is familiar. The physically meaningful quantity is not the vector field $A_i (i=1,2,3)$, but its curl, a quantity that rises above gauge, $B = \nabla\times A$; not the tensor field g_{jk}, but a quantity that rise above coordinates, the 3-geometry, $^{(3)}\mathcal{G}$, about which Section V has more to say; not the Yang-Mills field $A_i^{(S)}$ but a new geometrical entity that once more rises above gauge.

A scalar field, ϕ, departs in two ways from the pattern of the Yang-Mills field, gravitation, and electromagnetism. First, the requirement that its dynamics be imbeddable does not introduce gauge. If such a field existed in nature it would be directly observable. Second, imbeddability does not determine a unique local Hamiltonian. The function $f(\phi)$ in

$$H = (\nabla\phi)^2/8\pi+f(\phi) \tag{20}$$

is arbitrary.

Is this arbitrariness of the Hamiltonian for a scalar the reason why no scalar field has ever been found in nature? Is this omission a clue to how nature may build law without law? No one can rest happy with "history-independence of dynamics" as the foundation of physics, simple guide though it is to the great laws. It does not explain how it comes about that the dynamics must be imbedded in a manifold of 1 time and 3 space dimensions in the first place, nor why nature drops the scalar field.

No questions bring us closer than these to the frontiers of time. No way seems reasonable for deriving the dimensionality of 3 + 1 which does not start from a viewpoint that transcends dimensionality. No building blocks offer but elementary acts of observer-participancy. No method of construction that has been seen at work in other contexts looks more applicable than Feynman's sum over histories(78,79), applied however here to the higgledy-piggledy of yes-no-decision observations. No feature of such a sum over histories would seem more immediately susceptible to test than this, that it should kill out by destructive interference any contribution that looks like a scalar field. For the other three fields there is a uniqueness of contribution that can be imagined to lead to a constructive interference of elementary Feynman amplitudes, and therefore a non-zero representation of such fields in the physics. For the scalar field, however, does the very wealth of Hamiltonians acceptable at the classical level mean a wealth of values for the classical action I_H, and therefore wide-ranging values for the phase, $I_H/\hbar$ [see Eq. (3)] of the elementary Feynman amplitude? Does this feature of the phase in turn imply destructive interference, and therefore finally zero representation for the scalar field in the scheme of physics?

Whatever the next steps may be towards deriving "everythin out of nothing," the H-K-T result would seem to mark one of the largest leaps of recent times. Their way of analys starting from the "group" of deformations of a spacelike hypersurface, reminds us again of the power of symmetry considerations to simplify the content of physical law, and their impotence in revealing the machinery behind law. No one would dream of studying the laws of elasticity to uncover the principles of quantum mechanics. Neither would anyone investigate the work-hardening of a metal to learn about atomic physics. The order of understanding ran not

(21) *work-hardening* (1 cm) → *dislocations* (10^{-4} cm)
→ *atoms* (10^{-8} cm),

but the other way,

(22) *atoms* (10^{-8} cm) → *dislocations* (10^{-4} cm)
→ *work-hardening* (1 cm)

One had to know about atoms to conceive of dislocations, and had to know about dislocations to understand work-hardening. Is it not likewise hopeless to go from laws of physics to underlying machinery? Must the order of progress not be the direct opposite? If so, what course offers itself except to try "acts of observer-participancy" as the underlying "machinery," and see if out of them one can derive the laws of physics? Nothing does more to give a little encouragement in such an enterprise than the H-K-T achievement of deriving so much from so little, with the help of the concepts of "many-fingered time" and "imbeddability."

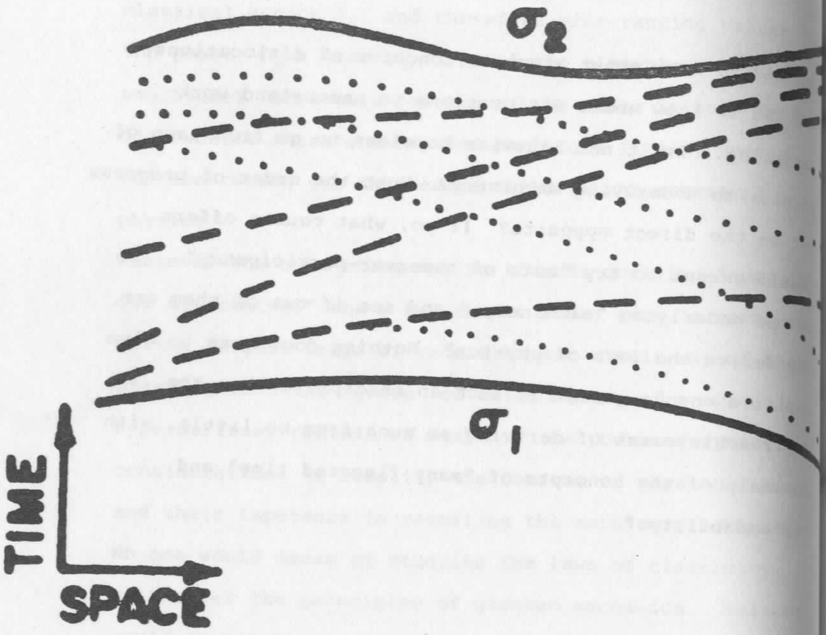


Figure 4.

Figure 4. The "history of deformation" indicated by the dashed hypersurface leads from initial-value hypersurface σ_1 to final-value hypersurface σ_2. So does the history indicated by the dotted hypersurfaces. The physics on σ_2 resulting from a complete specification of the initial value data on σ_1 must be independent of the history one chooses to integrate along in passing from σ_1 to σ_2 via the Hamiltonian equations of motion. This heavy but simple requirement suffices to fix the form of the Hamiltonian both for the dynamics of a vector field (giving Maxwell theory) and for the dynamics of the 3-geometry itself (giving Einstein's geometrodynamics) (Hojman, Kuchař, and Teitelboim, 1973).

V. TRANSCENDING TIME

> ". . .space and time are orders of things and not things"
>
> --G. W. Leibniz.(140)

> ". . .time and space are modes by which we think and not conditions in which we live"
>
> --A. Einstein.(141)

There is no such thing as spacetime, quantum mechanics tells us. Spacetime is a purely classical concept. It is a classical history of space geometry changing with the progress of time. What is meant by a "classical history of space changing with time?" How does it come about that quantum mechanics forbids this way of speaking? And what does it offer instead as acceptable way of describing the dynamics of space? But first, before any of these questions, why focus on 3-geometry at all when a casual impression might have made it seem that space<u>time</u> is the "without-which-nothing" ingredient of modern theoretical physics? How can one accept going back from four dimensions to three when one knows that going from three dimensions to four marke one of the great steps forward in the history of science? Not putting the fourth dimension into his curved space geometry accounts more reasonably than any other circumstance that one can easily name for Riemann's failure to discover general relativity. Already at the age of 27, in his habilitation lecture of June 10, 1854 on entry into the philosophical faculty of the University of Göttingen, he had set forth the mathematical tools to describe curvature in any number of dimensions; and he had declared that, "The properties which distinguish space from other conceivable triply-extended magnitudes are only to be deduced from experience. . . .At every point the three-directional measure of curvature can have an arbitrary value if only the effective curvature of every measurable region of space does not differ noticeably from zero."(142) Einstein speaks of the inspiration he derived from this lecture of Riemann in developing his own geometrical theory of gravity, "But. . . physicists were still far removed from such a way of thinking; space was still, for them, a rigid, homogeneous something, susceptible of no change or conditions. Only the genius of Riemann, solitary and uncomprehended, had already won its way by the middle of the last century to a new conception of space, in which space was deprived of its rigidity, and in which its power to take part in physical events was recognized as possible."(143) Dying of tuberculosis at Selasca on Lake Maggiore July 20, 1866, twelve years later, in his final days achieving with Betti a system for characterizing multiply connected topologies, Riemann failed in the other great enterprise to which he gave his last measure of devotion: to provide a unified explanation of gravitation and electromagnetism. It took 1905, Einstein, and special relativity to provide the missing concept: four dimensions, not three. With that recognized, it took only a decade to achieve general relativity and a fully geometrical theory of gravity in the spirit of Riemann.

It took much longer to recognize the dynamic structure of Einstein's geometrodynamics. The point that was most central, and took the longest to grasp, was also the simple The dynamic object is not spacetime. It is space. The geometric configuration of space changes with time. But it is space, three-dimensional space, that does the changing.

That 3-space is the dynamic object would have been recognize much sooner had the work and results[144,145] of Élie Cartan been more widely appreciated, whose deep insights into the theory of partial differential equations gave him a hold on many of the essential ideas. However, physics already had a standard machinery for dealing with dynamic problems, and it seemed natural to lay out general relativity in the Hamiltonian pattern without further thought. If the basic theory is 4-dimensional, should not the Hamiltonian be 4-dimensional, and was it not therefore reasonable to think of the dynamic object itself as also being 4-dimensional (spacetime)? No wonder that the resulting equations persist in yielding up zero quantities, statements that "zero equals zero," and deeper difficulties. These difficulties clouded the subject for several decades until Dirac on the

one hand[146,147] and Arnowitt, Deser and Misner on the other[148] moved from a 4-dimensional treatment to a 3-plus-1-dimensional analysis. Still further down the road one began to see the larger pattern of subject in all its basic simplicity.

The central concept lends itself to statement in a single sentence: <u>A 3-geometry describes the momentary configuration of space as it undergoes its dynamic change with time.</u>[34]

"3-geometry" is a coordinate-free concept. One does not have to use coordinates to speak of "a 2-sphere of radius *a*," nor coordinates to define "a 3-sphere of radius *a*," nor coordinates to describe the deformation of a 3-sphere of radius *a* into a 3-ellipsoid of principal dimensions *a*, *b*, *c*. But neither do coordinates hurt--nor the combination of coordinates and metric that gives the square of the element of distance,

$$ds^2 = g_{ij}dx^i dx^j . \tag{23}$$

For the 2-sphere one choice of coordinates gives

$$ds^2 = a^2(d\theta^2+\sin^2\theta d\phi^2); \tag{24}$$

another choice of coordinates on the same 2-sphere gives

$$ds^2 = \frac{dx^2+dy^2}{[1+(x^2+y^2)/4a^2]^2} ; \tag{25}$$

and there are similar options, infinite in number, for the coordinates on the 3-sphere and the 3-ellipsoid. What counts in these options is not the name given to the coordinates. The names for those coordinates one can standardize so that they always read x^1, x^2, x^3. What counts rather than name is the dependence on these coordinates of the metric coefficients. How is one to know that the metric of (24),

$$\left\| \begin{matrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{matrix} \right\| = \left\| \begin{matrix} a^2 & 0 \\ 0 & a^2\sin^2 x^1 \end{matrix} \right\| \tag{26}$$

and the metric of (25),

$$\left\| \begin{matrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{matrix} \right\| = \{1+[(x^1)^2+(x^2)^2]/4a^2\} \left\| \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right\| \tag{27}$$

describe the same ${}^{(2)}\mathcal{G}$--in this case, the same radius-a 2-sphere--whereas the metric

$$\left\| \begin{matrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{matrix} \right\| = \left\| \begin{matrix} a^2 & 0 \\ 0 & (a^2\sin^2 x^1+\varepsilon^2\sin^4 x^1) \end{matrix} \right\| \tag{28}$$

describes a figure with an equatorial bulge?

An alteration in metric coefficients that marks a real change in 3-geometry is distinguished most easily at the infinitesimal level from an alteration in the g_{ij} that arises from a mere change of coordinates.

Let a certain definite ${}^{(3)}\mathcal{G}$, with all its lumps, bumps and ripples, be expressed throughout one "local coordinate patch"(149) in terms of one set of coordinates x^i by one set of metric coefficients g_{ij}. Reexpress that 3-geometry in terms of new coordinates $\bar{x}^i$ shifted by the small amount ξ^i,

$$\bar{x}^i = x^i - \xi^i. \tag{29}$$

As a picture at the back of one's mind of what is going on, envisage the 3-geometry in question as the right hand front fender of a Ford automobile, short though it is by one dimension of measuring up to a proper mental image. It is distinct in shape from the right hand front fender of cars of a hundred other kinds. The difference is clear cut. There is no need of coordinates to see the difference. But coordinates provide a useful means to express the difference. To supply coordinates, take a sufficiently large and transparent rubber sheet. Mark on it an intersecting grid of lines. x^1= . . . ,13, 14, 15, . . .; x^2 = . . . ,77, 78, 79, 80, ... Apply it to the fender, stretching it so it fits snugly. Then every point on the fender acquires a pair of coordinates, (x^1,x^2). Therefore a measurement of the distances from a given point to several nearby points provides a straightforward way to determine the several metric coefficients g_{ij}. The distinction is clear between a mere change of coordinates and a true alteration in the shape of the fender, as for example in a collision. To ilustrate a mere change in coordinates, slip the marked rubber sheet over the surface of the fender a little way in the direction of increasing x^1. In that process a given scratch mark on the fender acquires a slightly decreased coordinate, $\bar{x}^1 = x^1 - \xi^1$; hence the minus sign in (29).

The distance from one scratch mark on the fender to the several nearby scratch marks naturally is not changed by the movement of the rubber sheet over the surface. In other words, $(ds)^2$ is coordinate independent. More concretely, we have for each pair of nearby scratch marks,

$$d\bar{s} = \bar{g}_{ij}(\bar{x})\,d\bar{x}^i d\bar{x}^j = ds^2 = g_{mn}(x)\,dx^m dx^n$$

$$= g_{mn}(\bar{x}+\xi)(d\bar{x}^m+d\xi^m)(d\bar{x}^n+d\xi^n)$$

$$(30) \quad = [g_{mn}(\bar{x})+(\partial g_{mn}/\partial\bar{x}^s)\xi^s]\times[d\bar{x}^m+(\partial\xi^m/\partial\bar{x}^p)d\bar{x}^p][d\bar{x}^n+(\partial\xi^n/\partial\bar{x}$$

an expression reminiscent of how one analyzes strain in the theory of elasticity. Comparing the coefficient of $d\bar{x}^i d\bar{x}^j$ on left and right, and taking account of the symmetry of metric coefficients in their two indices, we arrive at an expression showing how metric coefficients change as a result of the infinitesimal slippage of the sheet:

$$(31) \qquad \bar{g}_{ij} = g_{ij}+\xi_{i|j}+\xi_{j|i}\,,$$

or

$$(32) \qquad \delta g_{ij} = \xi_{i|j}+\xi_{j|i}\,.$$

Here

$$(33) \qquad \xi_{i|j} = \partial\xi_i/\partial x^j - \Gamma_{ij}^{\ \ m}\xi_m$$

is an abbreviation for the covariant derivative of the ith component of the displacement of the "rubber sheet" with respect to the jth coordinate. Moreover

$$(34)\ \Gamma_{ij}^{\ \ n} = g^{mn}\Gamma_{ij\,m} = (g^{mn}/2)(\partial g_{jm}/\partial x^i+\partial g_{im}/\partial x^j-\partial g_{ij}/\partial x^m),$$

expressed in terms of the rate of change of the metric coefficients and the elements g^{mn} of the matrix reciprocal to the metric tensor, is the typical "connection coefficient," having to do with the way the coordinate grid turns and swells or shrinks as one moves from point to point <u>on</u> the rubber sheet, regarded as fixed.

In brief, an infinitesimal change in metric coefficients, δg_{ij}, that lets itself be expressed in the form $\xi_{i|j}+\xi_{j|i}$ betokens no change in 3-geometry at all, only a change in coordinates, otherwise known as a "gauge change." In contrast, an infinitesimal change in metric coefficients that does not let itself be represented in this form is the sign of a real change in 3-geometry:

$$\delta g_{ij} = \xi_{i|j} + \xi_{j|i} + (\delta g_{ij})_{\text{real change}}. \quad (35)$$

This latter type of change, like the slow crumpling of an automobile fender, is what one means when one talks about the dynamics of geometry.

In what arena does the dynamics of geometry unroll? Superspace,(34,150) $\mathcal{S}$. Superspace, with suitable mathematical amendments(151), is the manifold made up by the totality of all 3-geometries. This manifold contains infinitely many points. Each point represents one and only one 3-geometry. A collection of these points makes up the dynamic history of space evolving with time.

How does one coordinatize superspace, and how does one describe that movement from one point to a nearby point which is the essence of dynamics?

First ever to consider superspace was Riemann himself(152), though not in the context of relativity, of course. His superspace was composed of the totality of all conformally equivalent closed Riemannian 2-geometries of the same topology. Such a superspace is known today as Teichmüller space. For more on Riemann's contribution to such superspac and the subsequent development of the relevant theory, reference may be made to the literature(150,153). For 2-geometries of genus g the superspace in question has dimens $6g-6$ for $g\geq 2$ ($g=0$, 2-sphere, dimensionality 0, $g=1$, 2-torus, dimensionality 2; $g=2$, figure eight shape, dimension 6); it i manifold of a very limited dimensionality. In contrast

the superspace built of 3-geometries requires an infinity of parameters for its representation.

For mathematical simplicity limit attention here and hereafter to closed, or in mathematical terms "compact," 3-geometries. The physics associated with such a restriction is briefly recapitulated in Section IX. Among compact 3-geometries the easiest to consider is a 3-sphere. Lifshitz and Khalatnikov have given a complete classification of the small deformations of a 3-sphere into tensorial harmonics,(154) analogous to the scalar harmonics that one finds so useful in electrostatics. The coefficients in this expansion provide countable and convenient coordinates to describe the small deformations of the geometry of the 3-sphere. * In the language of superspace, they allow one to "reach out" a little ways in every conceivable direction from one chosen point in the ∞-dimensional arena, $\mathcal{S}$. Similar ways have been discussed(150) for parametrizing, not only the small deformations of other 3-geometries, but also the general finite deformation--and thus coordinatizing superspace in its entirety.

An alternative approach to mathematizing superspace contents itself with an approximation that provides additional insight. As a smooth auditorium roof can be approximated arbitrarily closely by a geodesic dome constructed of sufficiently many sufficiently small flat triangles, so a smooth 3-geometry

* They parametrize either the tangent space at $x \in S^3$ <u>or</u> a local chart at $x \in S^3$.

can be approximated arbitrarily closely by a locked-together assembly of sufficiently many sufficiently small Euclidean tetrahedrons. This scheme of approximation, devised by Regge$^{(155)}$, has received the name of "Regge calculus" in a subsequent review$^{(156)}$.

The triangles that meet at a common vertex on the geodesic dome there ordinarily have angles that fall short by some small amount δ of adding up to $2\pi = 360°$ (Fig. 5). This "def angle" provides a measure of the curvature that is concentra at that point of the dome. Moreover, that angle and that curvature, and the analogous angles at all the other vertice of the dome--and therefore the "shape" or "2-geometry" of the dome as a whole--are all determined by a finite number of parameters, the edge lengths, $l_1, l_2, l_3, \ldots l_N$ of these triangles. Therefore it might seem reasonable directl to adopt these N lengths as coordinates to single out and specify the one 2-geometry in question in contrast to all the other 2-geometries available in the "truncated N-dimensional superspace" of the l_i. However, some changes i the l_i amount in effect to mere reexpression of essentiall the same 2-geometry in terms of triangles of slightly altered sizes and locations. Excluding such uninteresting alterations by appropriate supplementary conditions, one reduces the number of independent parameters from N to some lesser number, N', which has to be regarded as the proper dimensionality of the "truncated superspace" built

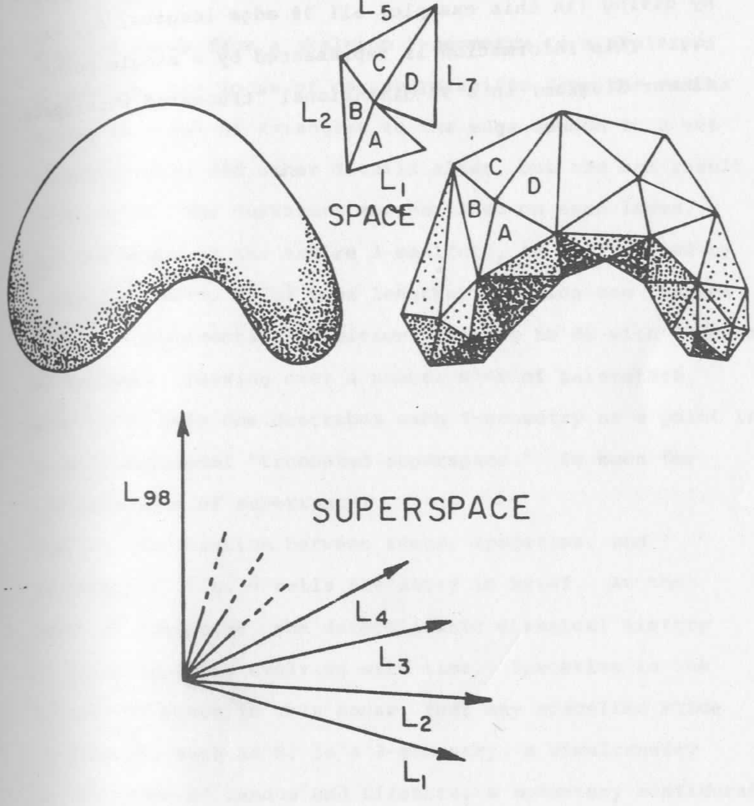


Figure 5.

Figure 5. A 2-geometry (upper left) is approximated by a skeleton 2-geometry (upper right). All the details of the shape of this skeleton 2-geometry are completely specified by giving (in this example) all 98 edge lengths, L_1, L_2,. . . L_{98}. This information is represented by a single point (lower diagram) in a 98-dimensional "truncated superspace."

of "skeleton 2-geometries" with the given number of vertices. The larger the number of vertices, the more closely one expects to be able to reproduce the results of an analysis based on the full ∞-dimensional superspace of 2-geometries. When one turns from a skeleton 2-geometry to a skeleton 3-geometry, the locus of curvature shifts from the vertex common to a set of triangles to the edge common to a set of tetrahedra, and other details alter, but the end result is similar. The curvature concentrated on each locus, and the shape of the entire 3-manifold, is fully fixed by a finite number, N, of edge lengths, on which one imposes certain supplementary conditions (having to do with "evenness of zoning"), leaving over a number $N'<N$ of parameters. With their help one describes each 3-geometry as a point in an N'-dimensional "truncated superspace." So much for illustrations of superspace!

What is the relation between space, spacetime, and superspace? Fig. 6 tells the story in brief. At the right is spacetime, the deterministic classical history of space geometry evolving with time. Spacetime is the history of space in this sense, that any spacelike slice through it, such as A, is a 3-geometry, a simultaneity in the sense of Landau and Lifshitz, a momentary configuration of space. That momentary 3-geometry, conceived here for definiteness as "closed" or "compact," and endowed with the topology of a 3-sphere, is illustrated schematically at the upper left--for want of dimensions on the paper--by a small deformed 2-sphere. In it are two bumps. They symbolize the local curvature of space produced by two large agglomerations of mass-energy at an early stage in the history of the universe when the dimensions of space were much smaller than they are today and galaxies were closer together. That entire 3-geometry A, with all its curves and bumps, is represented by a single point A in the infinite-dimensional superspace at the bottom of Fig. 6.

Another slice B through the same spacetime at the upper right provides another 3-geometry, another momentary configure for space in its dynamical evolution with time. The universe in this case is larger, but the two great clouds of mass-energy, because they happen to have started off moving towards each other, are now closer than they were in moving-picture-frame A. In the superspace description of the dynamics at the bottom of Fig. 6 this configuration of the universe is described by a single point, B.

A one-parameter family of spacelike slices through a given spacetime thus evidently "generates" a one parameter family of points running through superspace: a line or curve. However, time in general relativity has a many-fingered character. It bursts the bounds of anything so narrow as a one-parameter family of spacelike slices. The explorers

of spacetime have full liberty to push ahead their exploration faster in one place than another. They have perfect freedom to measure up the 3-geometry of the spacelike slice B'. This 3-geometry is represented by another point, B', in superspace. No one simple line in superspace can accomodate all the points A, B, B', . . . all the 3-geometries, that one gets by making spacelike slices in all conceivable ways through a given spacetime. The region of superspace occupied by all these points is not a line; it is a leaf.

A *leaf of history* (illustrated schematically by the bent leaf visible through the cut-away part of the lower diagram in Fig. 6) cuts through superspace. It describes the deterministic dynamical development of space with time.

To be more specific, consider one of the 3-geometries, say C, that is met with in the history of space, changing its shape with time. At each space point of this 3-dimensional manifold there are three independent and meaningful alterations that can be conceived in this 3-geometry (6 freely variably metric coefficients g_{ik}, diminished by the 3 types of change that arise out of mere changes in coordinates as in Eq. 30, giving a net of 3 "adjustable parameters" or "real degrees of freedom" per space point). One of these three modifications amounts to pushing the hypersurface ahead in time a small amount in the given spacetime

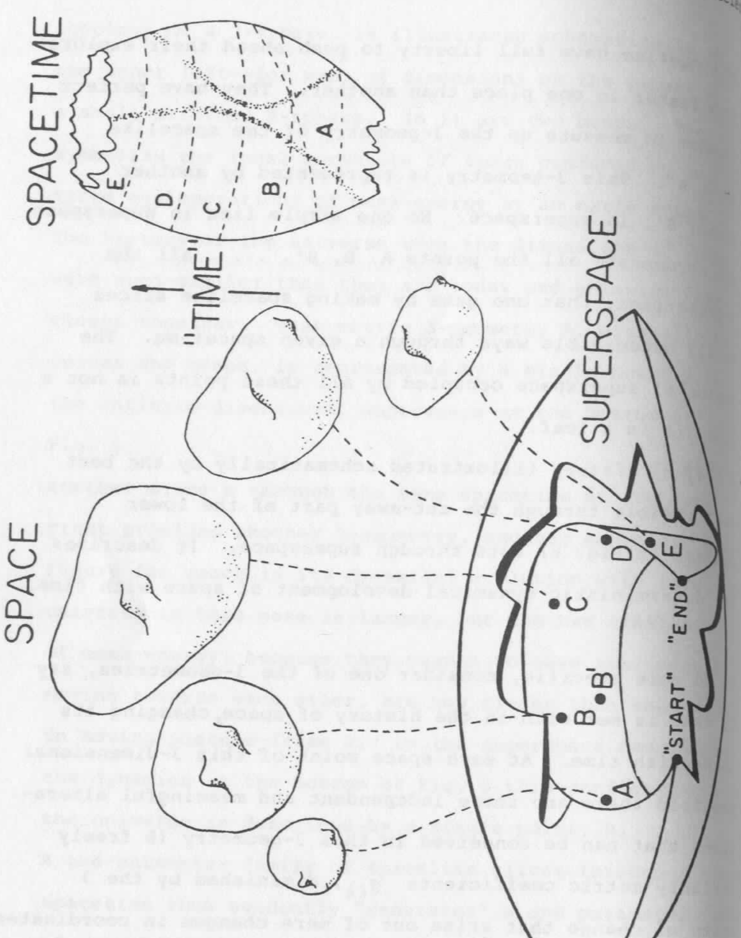


Figure 6. Space (upper left), spacetime (upper right) and superspace (below). The *leaf of history* that curves through superspace includes all the configuration (*A*, *B*, *B'*, . . .) achieved by space in its classical dynamical evolution in time; that is, all spacelike slices through the given spacetime. A different spacetime (not shown); that is, a classical history of space when the dynamics of space is started off with different initial conditions, corresponds to a different leaf of history (also not shown) cutting through superspace.

or 4-geometry: so much here, so much there, so much at each of the points in 3-space. It describes that freedom of exploration of a given spacetime which we subsume under the title of "many-fingered time." It describes a movement in superspace that leaves the representative point on the given leaf of history. It provides the tool for analyzing the given dynamics of geometry; for reaching, bit by bit, every conceivable spacelike slice that one can think of making through the given spacetime, illustrated at the upper right in Fig. 6.

Conversely, given all details of the spacetime geometry in question, and given all details of some particular 3-geometry that lies on that leaf of history in superspace, say C, then--apart from non-generic symmetries or degeneracies one can say exactly where that particular spacelike slice is located, and must necessarily be located, in that particula spacetime. In other words, in this sense the specification of a 3-geometry compatible with the given 4-geometry is entirely equivalent to the complete specification of many-fingered time. This is what one means by speaking of a 3-geometry as a "carrier of information about time."(157) "Time" conceived in these terms means nothing more or less than the location of the $^{(3)}\mathcal{G}$ in the $^{(4)}\mathcal{G}$.

Put in still other language, "time" tells how to take the $^{(3)}$ that are strung out on a given leaf of history in superspace a install them into a $^{(4)}\mathcal{G}$, an equivalent description of that same history. The child's

toy can be removed from its box only to reveal another box and--that taken away--another box, and so on, until eventually there are dozens of boxes scattered over the floor. Or conversely the boxes can be put back together, nested one inside the other, to reconstitute the original package. The packaging of $^{(3)}\mathcal{G}$'s into a $^{(4)}\mathcal{G}$ is much more sophisticated. Nature provides no monotonic ordering of the $^{(3)}\mathcal{G}$'s. Two of the dynamically allowed $^{(3)}\mathcal{G}$'s taken at random will often cross each other one or more times. When one shakes the $^{(4)}\mathcal{G}$ apart, he therefore gets enormously more $^{(3)}\mathcal{G}$'s "spread out over the floor" than he might otherwise have imagined. Conversely, when one puts back together all of the $^{(3)}\mathcal{G}$'s allowed by the condition of constructive interference, he gets a structure with a rigidity that he might not otherwise have foreseen. This rigidity arises from the infinitely rich interleaving and intercrossing of clear-cut well-defined $^{(3)}\mathcal{G}$'s one with another.

How different from the textbook concept of spacetime! There the geometry of spacetime is conceived as constructed out of elementary objects, or points, known as "events." Here, by contrast, the primary concept is 3-geometry, and the event is secondary: (1) The events lies at the "intersection" of such and such $^{(3)}\mathcal{G}$'s. (2) Its timelike relation to some other $^{(3)}\mathcal{G}$ is determined by the structure of the $^{(4)}\mathcal{G}$, which in turn derives from the intercrossings of all the other ${}^{(3)}\mathcal{G}$'s.

Whether one starts with ${}^{(3)}\mathcal{G}$'s as primary and regards the "event" as a derived concept, or vice versa, might make little difference if one were to remain in the domain of classical geometrodynamics. It makes all the difference when one turns to quantum geometrodynamics.

There is no such thing as a 4-geometry in quantum geometrodynamics, and for a simple reason. No probability amplitude function $\psi({}^{(3)}\mathcal{G})$ can propagate through superspace as an indefinitely sharp wave packet. It spreads. It has a finite probability amplitude in a domain of superspace of finite measure. This domain encompasses a set of ${}^{(3)}\mathcal{G}$'s far too numerous to be accomodated in any one ${}^{(4)}\mathcal{G}$. One can express this situation in various terms. One can say that propagation takes place in superspace, not by following any one classical history of space, not by following any one ${}^{(4)}\mathcal{G}$, but by summation of contributions from an infinite variety of such histories. In whatever way one states the matter, however, the facts are clear. The ${}^{(3)}\mathcal{G}$'s that occur with significant probability amplitude do not fit and cannot be fitted into any single ${}^{(4)}\mathcal{G}$. That "magic structure" of classical geometrodynamics simply does not exist. Without that building plan to organize the ${}^{(3)}\mathcal{G}$'s of significance into a definite relationship, one to another, even the "time ordering of events" is a notion devoid of all meaning.

These considerations reveal that the concepts of spacetime and time itself are not primary but secondary ideas in the structure of physical theory. These concepts are valid in the classical approximation. However, they have neither meaning nor application under circumstances when quantum-geometrodynamical effects become important. Then one has to forgo that view of nature in which every event, past, present, or future, occupies its preordained position in a grand catalog called "spacetime." There is no spacetime, there is no time, there is no before, there is no after. The question what happens "next" is without meaning.

How does one see these lessons of the quantum in more detail? And how close to being inescapable are they? Is geometry measurable anyway? Especially is it measurable in principle down to distances comparable to the Planck length of Eq. (5), where the concepts of "before" and "after" are predicted to lose all applicability?

Consider first geometry at the classical level. Compare the spacetime interval PQ anywhere in spacetime with a fiducial interval MN at a particular location in spacetime. (162-166)
Use any and all routes of intercomparison one pleases.

Get in every case without exception the same value for the ratio(163)

$$r = \overline{PQ}/\overline{MN} \qquad (36)$$

That is the central point and prediction of Riemannian geometry. It exposes itself to destruction on a hundred fronts. Were it not true, then for example electrons brought by different routes to the same iron atom at the center of the earth would be expected to have different properties. Then the Pauli exclusion principle would not apply. The electrons would all fall to the K-orbit. The iron atom--and the center of the earth--would collapse, contrary to observation.(163)

When one turns from the classical to the quantum dynamics of geometry, then field coordinate and field momentum have to be accepted as complementary, conjugate, and not simultaneously measurable quantities, the reciprocal uncertainty relations between which are given by the theory itself. Into these relations enters only one physical quantity, the Planck length of Eq. (5). Wigner and Saleckar,(162) looking at possible methods to measure the geometry compatible with the quantum principle, conclude that any determination of substantial precision is limited, not by the Planck distance, but by a distance many powers of ten greater. If this conclusion were to be upheld, one would have to accept that the quantum theory of the dynamics of geometry is incomplete or incorrect or both. A similar incompleteness or incorrectness in what quantum electrodynamics has to say about the possibilities for field measurements was claimed by Landau and Peierls.(167) It took the famous papers of Bohr and Rosenfeld(168,169) to show that the possibilities for making measurements had been too narrowly conceived, and that the precision predicted by theory could be attained in principle by idealized measuring equipment when one looked apart from limitations imposed by the atomic constitution of matter. In brief, devising measuring equipment that won't work is easier than devising equipment that will. In the end field theory itself would seem to be the safest guide--in the absence of other evidence--on the reciprocal uncertainties of the field quantities, and on the precision attainable in measurements of the 3-geometry intrinsic to a spacelike hypersurface or the extrinsic curvature of that geometry relative to the enveloping spacetime. We adopt this point of view pending further analysis and assessment of the conclusions of Wigner and Saleckar.

The plain straightforward conclusions of quantum geometrodynamics about uncertainties in spacetime geometry follow from an elementary line of reasoning as familiar in the physics of the simple harmonic oscillator as in the analysis of the electromagnetic field.

The essential ideas show up already in such an elementary system as a single harmonic oscillator. There we write the wave function of a typical state, for example the ground state, in the form

$$(37) \qquad \psi(x) = N\exp(-m\omega x^2/2\hbar),$$

where N is a normalization factor. We proceed similarly with a collection of harmonic oscillators; and with suitably normalized displacement coordinates $\xi_1, \xi_2, \ldots$, we have for the ground state probability amplitude function the expression

$$(38) \qquad \psi(\xi_1, \xi_2, \ldots) = N\exp(-\xi_1^2 - \xi_2^2 - \ldots).$$

More familiar in the case of the electromagnetic field than this description in terms of oscillator amplitudes is the so-called occupation number representation; but a third, *spacelike*, representation prepares the way for situations, as in general relativity, where Fourier analysis is not appropriate. The magnetic field at the point x, y, z, expressed in terms of normal modes and the amplitudes $\xi_1, \xi_2, \xi_3, \xi_4, \ldots$ of these normal modes, has the form

$$(39) \qquad \underline{\mathrm{B}}(x,y,z) = \sum_{n=1}^{\infty} \xi_n \underline{\mathrm{B}}^{(n)}(x,y,z)$$

To specify the amplitudes is to specify the magnetic field; but conversely, to specify the magnetic field everywhere is to have all the information required to

determine the ξ's and therefore the wave function (38); thus,

$$(40)\ \psi = \psi(\underline{\mathrm{B}}(x,y,z)) = N\exp-(1/16\pi^4\hbar c)\iint r_{12}^{-2}\,\underline{\mathrm{B}}(1)\cdot\underline{\mathrm{B}}(2)\,d^3x_1\,d^3x_2.$$

In expression (40) one has the probability amplitude for a given, global, configuration of the magnetic field. For example, a configuration in which $\underline{\mathrm{B}}$ is zero everywhere, except for a non-zero value $\Delta\underline{\mathrm{B}}$ in a region of extension $\sim L$, has a probability amplitude in which the exponent in (40) is of the order

$$(41) \qquad L^4(\Delta \mathrm{B})^2/\hbar c.$$

In this sense a field fluctuation $\Delta\mathrm{B}$ has a negligible probability unless its magnitude is of the order

$$(42) \qquad \Delta\mathrm{B} \sim (\hbar c)^{\frac{1}{2}}/L^2$$

or less.

In a fuller description the appropriate wave function depends on the time t as well as on the entire configuration of the magnetic field at that time. However, in a curved spacetime one generalizes from a time coordinate t to an arbitrary spacelike hypersurface σ. The probability amplitude depends as well on σ as on the configuration of the magnetic field *upon* this hypersurface:

$$\psi = \psi(\underset{\sim}{B}, \sigma). \tag{43}$$

This wave function in this spacelike representation satisfies Tomonaga's wave equation, with its "bubble time" functional differentiation,[131]

$$i h \; \delta\psi/\delta\sigma = (\underset{\sim}{B}^2/8\pi)\psi + (1/8\pi)[(4\pi h/i)\delta/\delta\underset{\sim}{A}]^2\psi \tag{44}$$

The wave function ostensibly depends on all three components of the vector potential $\underset{\sim}{A}$; thus,

$$\psi = \psi(\underset{\sim}{A}, \sigma) \tag{45}$$

However, the change in these components induced by the arbitrary infinitesimal "change in gauge" λ,

$$A_k \rightarrow A_k + \partial\lambda/\partial x^k, \tag{46}$$

produces no change in what alone counts physically, the

magnetic field B (Eq. 43). Therefore the change in ψ resulting from the transformation (46) must vanish for arbitrary choice of the gauge function λ; that is, in the last integral below,

$$\begin{aligned} \delta\psi &= \int (\delta\psi/\delta A_k)\; \delta A_k \; d^3x \\ &= -\int \lambda[(\partial/\partial x^k)(\delta\psi/\delta A_k)] \; d^3x, \end{aligned} \tag{47}$$

the expression in square brackets must vanish everywhere. This is the condition that the divergence of the electric field should vanish, expressed in operator language. It is also the condition that ψ, ostensibly dependent upon the potential $\underset{\sim}{A}$, with its three independent components per space point, should really depend only on the divergence-free field $\underset{\sim}{B}$, with its two independent components per space point.

In a similar way the superspace formulation of general relativity (here taken for simplicity to be source-free) expresses the state functional as ostensibly dependent on the six independent g_{ik} of the metric upon a spacelike hypersurface, but in reality as dependent only on the coordinate independent 3-geometry $^{(3)}\mathcal{G}$ described by this metric. This 3-geometry is not at all affected by the arbitrary infinitesimal coordinate transformation

$$\bar{x}^p = x^p - \xi^p , \tag{48}$$

$$\bar{g}_{pq} = g_{pq} + \xi_{p|q} + \xi_{q|p}$$

where the vertical slash stands for covariant differentiation. Therefore the change in ψ resulting from (48), calculated in exact analogy to (47), must vanish for arbitrary choice of the three infinitesimal coordinate shifts ξ^p, from which one concludes that the three conditions,

$$(\delta\psi/\delta g_{pq})_{|q} = 0 , \tag{49}$$

must be fulfilled everywhere. Thus ψ, instead of depending upon 6 quantities g_{ij} per space point, depends only on the three quantities per space point that are carried in ${}^{(3)}\mathcal{G}$:

$$\psi = \psi({}^{(3)}\mathcal{G}) . \tag{50}$$

Of these three "informations," two have to do with gravitational wave amplitudes, and one with time. In the case of electromagnetism these two kinds of data are cleanly separated in (43). In the case of a 3-geometry no such clean separation

---

of wave amplitude from time is possible. The 3-geometry as a whole is a "carrier of information about time"(157). Each 3-geometry requires for its specification an infinite number of paramters and can be represented as a point in an infinite dimensional manifold, superspace.(150)

The propagation of the probability amplitude, ψ, in the superspace of geometrodynamics requires a propagation law analogous to the Tomonaga equation (44) of electrodynamics; symbolically,

$$\delta^2\psi/(\delta\, {}^{(3)}\mathcal{G})^2 + {}^{(3)}R = 0 , \tag{51}$$

where ${}^{(3)}R$ is the local value of the curvature scalar of the 3-geometry. In the WKB approximation, where ψ is represented as a slowly varying amplitude factor times a rapidly varying phase factor,

$$\psi \sim A \exp(iS/\hbar) , \tag{52}$$

the "dynamical phase," or Hamilton-Jacobi function $S({}^{(3)}\mathcal{G})$ satisfies the Einstein-Hamilton-Jacobi equation of Peres,(171) the "dispersion relation,"

$$(16\pi)^2 g^{-\frac{1}{2}}\left[\tfrac{1}{2}g_{pq}g_{rs}-g_{pr}g_{qs}\right](\delta S/\delta g_{pq})(\delta S/\delta g_{rs})+g^{\frac{1}{2}}\,{}^{(3)}R = 0 \tag{53}$$

All of (source-free) classical general relativity follows from this one equation[172]. Consider a classical history H_{class} of 3-geometry developing deterministically in time in accordance with Einstein's field equation. Consider the "leaf of history" in superspace that describes this dynamics. Consider one of the $^{(3)}\mathcal{G}$'s that is met on this leaf of history. Per space point of this 3-dimensional manifold there are three independent modifications that can be conceived in this 3-geometry (6-3 arbitrary coordinates g_{ik} = 3 real degrees of freedom per space point). One of these modifications amounts to pushing the hypersurface ahead in time a small amount in the given 4-geometry. The other two modifications change gravitational wave degrees of freedom, therefore change the spacetime, and therefore carry the representative point in superspace off the given leaf of history. In other words, the infinite dimensional space of small deformations away from the given point $^{(3)}\mathcal{G}$ on the leaf of history ("local tangent space of superspace") breaks down into the product of two subspaces, each also infinite dimensional. One has one third the dimensionality of the original space. It

is the subspace of deformations that leave $^{(3)}\mathcal{G}$ on the leaf of history. The other has two thirds the dimensionality of the full tangent space. It is the subspace of deformations that move $^{(3)}\mathcal{G}$ off the leaf of history. Quantum geometrodynamics makes no such sharp distinction. It assigns a finite probability amplitude $\psi(^{(3)}\mathcal{G})$ to 3-geometries off the classical leaf. This spread of the state function in superspace is the superspace description of the quantum fluctuations in geometry. A closer analysis[173,174] tells us that in a probe region of extension L, the quantum fluctuations in the normal metric coefficients (-1, 1, 1, 1) are of the order

$$\Delta g \sim L^*/L, \tag{54}$$

where L* is the Planck length.

To summarize, the sharp division of superspace by a classical history into "Yes" and "No" $^{(3)}\mathcal{G}$'s is denied by the quantum principle, which assigns a probability amplitude $\psi(^{(3)}\mathcal{G})$ to every 3-geometry. The $^{(3)}\mathcal{G}$'s with appreciable probability amplitude are too numerous to be accommodated into any one spacetime. Thus the uncertainty principle declares that spacetime is only an approximate and classical concept. In reality there is no such thing as spacetime. "Time" itself loses its meaning, and the words "before" and "after" are without application. These long known considerations are of importance only at the Planck scale of distances. They all flow out of the straightforward analysis of the dynamics of geometry in the arena of superspace, inescapable conceptual adjunct of general relativity.

[The above ten paragraphs are from reference 170.]

VI. Causal Order Without Causal Order

Time is not primordial. It, like every concept that man works with, is secondary and derived. How time, and spacetime, force themselves upon us in our efforts to organize our observations is a question on which only a miniscule beginning has so far been made despite the impressive pioneer work of Mach(175) and Piaget(176). Not one bit of further headway into this enterprise do the present lectures intend to make. Their purpose is much less courageous. Don't try to "take time apart" into the elementary quantum acts of observer-participatorship out of which we conceive it--and everything--to be built. Instead, sticking to the solid ground of physics as we know it, identify domains where familiar concepts of time and causality come to the limit of applicability and have to be modified. We have just finished exploring one such frontier. We have seen how both time and spacetime, according to existing theory, lose all application at the Planck distance and the Planck time; but how out of a description that transcends time--out of superspace--we come back in the appropriate correspondence principle limit to familiar views of time. We now turn to another question. Can we similarly arrive at the familiar ordering of cause and effect from a description that transcends that order? We consider a system of point charges coupled with each other by elementary electromagnetic actions-at-a-distance, individually time symmetric, in the sense that the force exerted on particle b by particle a is given by half the retarded field of a, as usually calculated, plus half the advanced field. Of the motives for considering such a coupling--that it should be derivable from an action principle, that it should be compatible with a principle of action and reaction, that it should repro all the familiar physics of electrostatics and of electromagnetism--we shall say nothing here, for the subject is discussed at length in two papers written with Richard Feynman(93,94). For simplicity the interactions are treated in the context of classical theory and a pre-existing flat spacetime. The poi of interest is the field created by one of these particles wh it undergoes a sudden acceleration. Experience says that the effect produced will be confined to the future light cone of the acceleration. With this observation the model seems absolutely incompatible. It links past and future in a maze of backward and forward running light rays. Nowhere can the slightest change be made without altering motions everywhere into the indefinite past and future.

Why should we be interested in trying to derive causality out of an apparently so preposterous model? Because we want to establish in this one example a point of more general application: The apparent inability of an action taken now to influence the past by no means rules out a direct influence of the present in "bringing about that which we call the past". It is in no way suggested here that this is the actual mechani by which acts of observer-participancy in the present bri that which we call the tangible or communicable reality of the universe at an era when no observers existed. That is a deeper question with which physics is not yet prepared to deal. However, one is open to believe that the kind of consid tions that elucidate the one may clarify the other.

The concrete issue then is this: How is one to reconcile the (1/2)R + (1/2)A field that the accelerated particle produces in the model with the 1 R field that the particle in actuality produces? The answer is easily summarized. The far away absorber driven by the source, produces a field which in the neighborhood of the source, though source-free there, looks as if it were a field for which the source is directly responsible, (1/2)R - (1/2)A. Combined with the field due to the source, this field generated in the absorber gives rise in the vicinity of a to total field, R, in full agreement with experience. Thus the familiar ordering in time of cause and effect is upheld in a model which at the beginning violated that ordering as outrageously as one could well imagine.

The idea thus so briefly stated raises several questions. How can the "superposition of the advanced fields of a large number of particles...give the appearance of both retarded and advanced fields due to the source itself[?](93) The advanced field of a single charge of the absorber can be symbolized as a sphere which is converging towards the particle and which will collapse upon the source. But at the moment when the source particle itself was accelerated, the sphere in question had a substantial radius. One point on it touched, or nearly touched, the source. The shrinking sphere therefore appears to the source as a nearly plane wave which passes over it headed towards one of the particles of the absorber. When we consider the effect of all the absorbing charges, we have to visualize an array of approximately plane waves, all marching towards the source and passing over it in step. The resultant of these individual effects is an spherical wave, the envelope of the many nearly plane waves. The sphere converges, collapses on the source, and then pours out again as a divergent sphere. An observer in the neighborhood will gain the impressi that this divergent wave originated from the source."

"Why does radiation have [an] irreversible character even in a formulation of electrodynamics which is from the beginning symmetrical with respect to the interchange of past and future? ...We have to conclude with Einstein(95) that the irreversibility of the emission process is a phenomenon of statistical mechanics connected with the asymmetry of the initial conditi with respect to time. In our example the particles of the absorber were either at rest or in random motion before the time at which the impulse was given to the source.

"That it is solely the nature of the initial conditions which governs the direction of the radiation process can be seen by imagining a reversal of the direction of time... We have then a solution of the equations of motion just as consistent as the original solution. However, our interpretation of the solution is different. As the result of chaotic motion going on in the absorber, we see each one of the particles receiving at the proper moment just the right impulse to generate a disturbance which converges upon the source at the precise instant when it is accelerated. The source receives energy and the particles of the absorber are left with diminished velocity. No electrodynamic objection can be raised against this solution of the equations of motion. Small *a priori*

probability of the given initial conditions provides our only basis on which to exclude such phenomena."

What is the effect on the source particle of the (1/2)R - (1/2)A field produced by the absorber? It gives rise to the familiar and well tested force of radiative reaction(93). What for our present purpose is the central lesson of this study in electrodynamics? That an order in time, ostensibly causal, can originate from an underlying machinery that is very far from causal.

Why is this point relevant to our larger theme (Section I) of "law without law"? Because we see here a sample law, causality, emerging from a description of nature that contains no such law.

VII. Asymmetry in Time and the Expansion of the Universe

Rhenium-187 has a half-life of 40×10^9a (a=year), as measured today. In other words, of ^{187}Re atoms now present, the fracti

(55) $$-dN/N = \lambda_{\text{apparent}}\,dt$$

will disappear, on the average, in the time dt, where the familiar constant for radioactive decay has the value

(56) $$\lambda_{\text{apparent}} = 0.693/40 \text{ in units of } (10^9 a)^{-1}.$$

Therefore, it has often seemed natural to suppose that the number, N, of these atoms has been, is now, and will continue falling off as $\exp(-\lambda_{\text{app}}t)$. This assumption is a special case of the belief of older times that the universe will endure forever but that all activity in it will eventually slow down and end in a "heat death." In that final condition, it was imagined, temperature differences, net outflow of particles from radioactive nuclei, and all other measures of departure from statistical equilibrium will have sunk to zero and "the entropy of the universe" will have attained "the absolute maximum" of which N. L. Sadi Carnot was already writing in 1824, inspiration for the phrase "the heat-death of the universe" that Clausius first set down on paper in 1865,(177) and that Bertrand Russell much later took as gospel truth when he wrote,(178) "The second law of thermodynamics makes it scarcely possible to doubt that the universe is runn down, and that, ultimately, nothing of the slightest interest will be possible anywhere."

Will the amount of ^{187}Re really fall off exponentially with time? Will temperature differences really sink exponentially to zero? Is perpetual approach to equilibrium guaranteed? It is impossible to face up to such questions in our own time without encountering issues of cosmology and without having to ask, is there a connection between statistical mechanics and cosmology? An exponential can only be brought to zero in an infinite time; but a finite time is all that is available in the familiar Friedmann model of a closed universe. If the universe is to end out of equilibrium who knows enough to say that it should not end as much out of equilibrium as it started? How then can one properly predict the amount of ^{187}Re over a cosmological range of time without first coming to grips with this question of "double-ended" statistical mechanics?(179)

The first line of the first page of a recent and distinguished book(180) by two leading mathematicians declares that, "The fundamental problem of mechanics is computing, or studying qualitatively, the evolution of a dynamical system with prescribed initial data." Moreover, thus to focus on an initial time and at that time to specify all coordinates and momenta is often the most useful way to apply dynamics to a given problem and sometimes the only way. However, one states the data in a quite different and thoroughly time-symmetric, "double-ended," way when one derives dynamics in the first place from either

(1) the Euler-Lagrange variation principle of point mechanics;

(2) the Hamilton variation principle of point mechanics;

(3) the Hilbert variation principle of electrodynamics;

(4) the Hilbert variation principle of general relativi

(5) the Hojman-Kuchař-Teitelboim imbeddability argumen of Section III; or

(6) the Feynman sum-over-histories.

One deals with the coordinates of particles or fields, and coordinates only, but at two times, or on two spacelike hyper surfaces.

If one thus plumbs some of the deepest issues of dynamics in terms of "double-ended data," can one escape from asking what statistical mechanics looks like when it too is stated in terms of double-ended data? No more quickly than by this ro is one led--if one is ever led--to question(181-188) the automatic presupposition that departures from equilibri will necessarily decrease and entropy will inescapably incre in the Einstein-Freidmann-predicted phase of contraction of th universe.

A recent review(188) puts the issue in these terms: "As dynamic time marches forward, what will happen then [in the phase of contraction] to [the arrows of] statistical and biological time? Will they continue to point in the same dir or will they point in opposite directions? In the one case, to a person alive in the second phase of the universe, the universe will appear to be contracting. In the other case, it will appear to be expanding, simply because a moving pict of contraction run backwards looks like expansion. Many colleagues agree that the question is open and that the answ

is one of the great puzzles of our day; but others are strongly convinced that the one answer or the other is the only right answer and that the answer is perfectly obvious and should be accepted without question. This is the insanity of the subject [of the arrow of time]." To paraphrase, it is not a question of accepting a solution; it is a question of accepting a problem.(189)

Is there any real doubt that each revolution of the earth around the sun will see a greater statistical-mechanical disorder in the universe, down to the end of time? Doubt shows clear in the works of leading figures in statistical mechanics from Boltzmann to today. Presuppose order in the initial conditions, and randomness otherwise? That assumption, all recognize, will reproduce the evidence of experience that entropy increases. But has so cosmic an assumption any deeper foundation? Doubt begins when it is asked whether entropy will increase forever. Doubt grows when it is asked, why order in the initial conditions? Doubt takes a new turn with the advances of relativistic astrophysics of recent years. How can a cosmological requirement on initial conditions possibly be imagined to be well grounded when it presupposes the out-of-date cosmological model of a universe that endures from everlasting to everlasting? That this doubt about the right end-point conditions for statistical mechanics has a long history one can forbear from reminding oneself anew by skipping the next few pages of brief quotes, extracted for the most part from the collection of reprints and translations of reprints edited by Stephen Brush.(190) Nothing stands out more strikingly from this quick oversight of the last hundred years than the "foreverness" of the cosmology taken as forgranted in all the discussions.

Gibbs(191) (1875): "The impossibility of an uncompensated decrease in entropy seems to be reduced to an improbability."

Boltzmann(192) (1877): ". . .The fact that this integral [$\int dQ/T$] is actually $\leqq 0$ for all processes in the world in which we live (as experience shows) is not due to the nature of the forces, but rather to the initial conditions." "It is only because there are many more uniform distributions than non-uniform ones that the distribution of states will become uniform in the course of time. One therefore cannot prove that, whatever may be the positions and velocities of the spheres at the beginning, the distribution must become uniform after a long time; rather one can only prove that infinitely many more initial states will lead to a uniform one after a definite length of time than to a non-uniform one." ". . .When we follow the state of the world into the infinitely distant past [here Boltzmann is speaking without benefit of the present day evidence for big bang cosmology, and is tacitly assuming that the universe endures from everlasting to everlasting], we are actually just as correct in taking it to be very probable that we would reach a state in which all temperature differences have disappeared, as we would be in following the state of the world into the distant future. . .if we know that in a gas at a certain time there is a non-uniform distribution of states, and that the gas has been in the same container without external disturbance for a very long time, then we must conclude that much earlier the distribution of states was uniform and that the rare case occurred that it gradually became non-uniform." "If perhaps this reduction of the second law to the realm of probability makes its application to the entire universe appear dubious, yet the laws of probability theory are confirmed by all experiments carried out in the laboratory."

Poincaré(193) (1893): "A theorem, easy to prove, tells us that a bounded world, governed only by the laws of mechanics, will always pass through a state very close to its initial state. On the other hand, according to accepted experimental laws. . . the universe tends towards a certain final state [of uniform temperature], from which it will never depart. . . I do not know if it has been remarked that the English kinetic theories can extricate themselves from this contradiction. The world, according to them, tends at first toward a stage where it remains for a long time without apparent change; and this is consistent with experience; but it does not remain that way forever, if the theorem cited above is not violated; it merely stays there for an enormously long time, a time which is longer the more numerous are the molecules. This state will not be the final death of the universe, but a sort of slumber, from which it will awake after millions of millions of centuries. According to this theory, to see heat pass from a cold body to a warm one, it will not be necessary to have the acute vision, the intelligence, and the dexterity of Maxwell's demon; it will suffice to have a little patience."

Zermelo(194) (1896): "Poincaré's theorem(195) says that in a system of mass-points under the influence of forces that depend only on position in space, in general any state of motion (characterized by configurations and velocities) must recur arbitrarily often, at least to any arbitrary degree of approximation even if not exactly, provided that the coordinates and velocities cannot increase to infinity.

Hence, in such a system irreversible processes are impossible (aside from singular initial states)."

"Suppose we have a gas enclosed in a solid container with elastic sides that are impermeable to heat. In general there will indeed be an infinite manifold of states of the molecules for which the gas will undergo permanent changes of state, such as viscosity, heat conduction, or diffusion. However, there will also be a much larger number of possible initial states, which can be reached by arbitrarily small displacements from the former states, and these states, instead of undergoing irreversible changes, will come back periodically to their initial states as closely as one likes. . . ."

Boltzmann(196) (1896): "Poincaré's theorem, which Zermelo explains at the beginning of his paper, is clearly correct,

but his application of it to the theory of heat is not. ". . .according to the laws of probability a certain quantity H (which is some kind of measure of the deviation of the prevailing state from Maxwell's) can only decrease for a stationary gas in a stationary container. . .[Thereafter the H-curve] almost always runs very close to the abscissa [time] axis. Only very rarely does it rise up above this axis; we call this a peak, and indeed the probability of a peak [significant deviation from Maxwell's distribution] decreases very rapidly as the height of the peak increases. . . It is just for certain singular initial states that the Maxwell distribution is never reached, for example when all the molecules are initially moving in a line perpendicular to two sides of the container. . . Whereas Zermelo says that the number of states that finally lead to the Maxwellian state is small compared to all possible states, I assert on the contrary that by far the largest number of possible states are "Maxwellian" and that the number that deviate from the Maxwellian state is vanishingly small. . .According to the molecular-kinetic view, this [second] law [of thermodynamics] is merely a theorem of probability theory. According to this view, it cannot be proved from the equations of motion that all phenomena must evolve in a certain direction in time.

"An answer to the question--how does it happen that at present the bodies surrounding us are in a very improbable state--cannot be given, any more than one can expect science to tell us why phenomena occur at all and take place according to certain laws.

. . .One may say that according to Poincaré's theorem the entire universe must return to its initial state after a sufficiently long time, and hence there must be times when all processes take place in the opposite direction. How shall we decide, when we leave the domain of the observable, whether the age of the universe, or the number of centers of force which it contains, is infinite?"

Zermelo[197] (1896): ". . .as long as one cannot make comprehensible the <u>physical</u> <u>origin</u> of the initial state, one must merely assume what one wants to prove; instead of an explanation one has a renunciation of any explanation."

Boltzmann[198] (1897): "The second law will be explained mechanically by means of assumption A (which is of course unprovable) that the universe, considered as a mechanical system--or at least a very large part of it which surrounds us--started from a very improbable state, and is still in an improbable state.

". . .Poincaré's theorem does not contradict the applicability of probability theory, but rather supports it, since it shows that in eons of time there will occur a relatively short period during which the state probability and the entropy of the gas will significantly decrease, and that a more ordered state similar to the initial state will occur.

"One has the choice of two kinds of pictures. One can assume that the entire universe finds itself at present in a very improbable state. [Or one can assume that for] the universe as a whole the two directions of time are indistinguishable. . . [with, however] here and there relatively small regions of the size of our galaxy (which we call worlds), which during the relatively short time of eons deviate significantly from thermal equilibrium. . .a living being that finds itself in such a world at a certain period of time can define the time direction as going from less probable to more probable states (the former will be the "past" and the latter the "future") and by virtue of this deinfition he will find that this small region, isolated from the rest of the universe, is "initially" always in an improbable state."

Ehrenfest-Afanassjewa[199,200] (1959): "Although Boltzmann did not fully succeed in proving the tendency of the world to go to a final equilibrium state, there remain after all criticisms the following valuable results. First, the derivation of the Maxwell-Boltzmann distribution for equilibrium states, then the kinetic interpretation of the entropy by the H-function, and finally the <u>explanation</u> of the existence of an integrating factor for dU + dA. . . .

"The so important irreversibility of all observable processes can be fitted into the picture in the following way. The period of time in which we live happens to be a period in which the H-function of the part of the world accessible to observation decreases. This coincidence is really not an accident, since the existence and functioning of our organisms, as they are now, would not be possible in any other period. To try to explain this coincidence by any kind of probability considerations will in my opinion necessarily fail. The expectation that the irreversible behaviour will not stop suddenly is in harmony with the mechanical foundations of the kinetic theory."

Uhlenbeck(201) (1968): ". . .one then can conclude that if the system is not in thermal equilibrium it almost always will go into that state; and if the system is in thermal equilibrium, it almost always will stay in that state although fluctuations away from equilibrium will and must occur because of the quasi-periodic nature of the motion of the Γ-point. This is the Boltzmann picture; it clearly reconciles the reversibility of the mechanical motion as expressed by the Poincaré theorem with the approach to equilibrium as required by the zeroth law of thermodynamics. . . . how is it possible that a contracted description can be closed and causal [?]. In a bona-fide macroscopic theory it should of course not be necessary to go back to the microscopic, molecular picture (in this sense the theory must be closed), and it should be possible to make predictions, that is the theory must be causal. This is the macroscopic causality problem and although it is in my opinion still far from clarified, one begins to see some light thanks to the basic papers of Bogoliubov(203). Bogoliubov pointed out that in any macroscopic theory the macroscopic variables must be in some sense secular variables, that is they must vary in time much slower than all the remaining variables needed to describe the molecular system."

Cohen(205) (1973): "It is the Boltzmann Ansatz, the statistical Ansatz of molecular chaos, which introduces the arrow of time or. . .the approach to equilibrium. It is the assumption of the factorization of the s-particle distribution at time t = 0, which is a generalization of the statistical Ansatz, which introduces the irreversibility."

For more on the history and the issues, reference may be made to a review article of Prigogine,(206) Klein's biography of Paul Ehrenfest,(207) and especially the books of Gold(186), Reichenbach(208), and Davies(209).

In summary, after a century and more, half the battle has been won to understand the direction of time in heat flow and other statistical processes; but the other half looks like being a long struggle. Evidently it is generally accepted that the elementary molecular interaction is time symmetric in thermal conduction, in viscosity, and in other irreversible processes of everyday interest; and that the observed macroscopic irreversibility takes its origin in two circumstances: the enormous number of molecules involved, and the asymmetry in time of the initial conditions. In other words, conditions were ordered before the relevant observations were made and disordered afterwards. In the rare case in which conditions are guaranteed instead to be disordered before, and ordered after one measures--say every five minutes for an hour--the temperature difference between a hot block of metal and a cold one in contact with it, then the same reasoning tells us that the temperature difference, rather than falling exponentially with time, should rise exponentially. This reasoning about exponential rise has been confirmed observationally so far only at the level of small fluctuations. For the temperature difference to increase exponentially by chance fluctuations to any truly macroscopic--and macroscopically observable--level would require a time so fantastically long as to put a test at this level utterly beyond reach. All this is not only understandable, but also well understood, as the quotes indicate. Different investigations use different words to make the same by now generally agreed points: all the elementary processes normally taken into consideration are reversible in time at the microscopic level; and the macroscopic resultant of large numbers of such processes is shown to go according to the usual sense of the arrow of time only by appealing to boundary value conditions on the microscopic motions that presuppose order in the past, disorder in the future. Here consensus ends. Shall one or shall one not impose boundary conditions near the big crunch similar to those

that one imagines imposing near the big bang? Or is it even without sense, as some would suggest, to raise such issues? To flee the abstractness of these question, let us turn to a concrete model.

No model ever illustrated approach to equilibrium more instructively than the Ehrenfest double-urn model(209-213) (Fig. 7). No model allows one to see more vividly what it means to prescribe a departure from equilibrium at a "final time" $t = +T$ as great as the departure at an "initial time," $t = -T$. The idea is due to Cocke.(179) One starts as did the Ehrenfests with the 100 balls divided $75 = 50+25 = 50 + n = 50 +$ "surplus" in the left hand urn and $25 = 50-25 = 50 - n$ in the right hand one. Each spin of the roulette wheel brings up a number between 1 and 100. The ball with that number painted on it thereupon jumps from whichever urn it's in to the other urn. To start with there are three times as many balls on the left. Therefore it is three times more probable that a given number will turn up on a ball on the left. Consequently the most probable course of events is a gradual decrease in the number of balls on the left:

$$A = \begin{pmatrix}\text{chance that roulette wheel}\\ \text{causes ball on left to jump}\end{pmatrix} = \frac{50 + n}{100}$$

$$B = (\text{change of } n \text{ in such a jump}) = -1$$

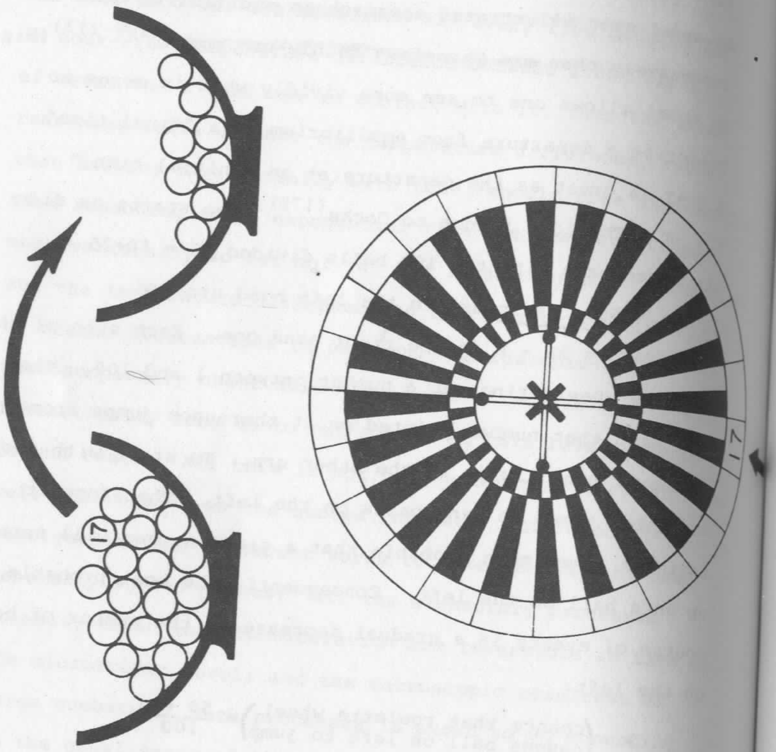


Figure 7. The Ehrenfest double urn in a 1978 rendering. When number 17 comes up on the roulette wheel, the ball carrying that number is transferred from whichever urn it happens to be in to the other urn. Thus 100 balls, 75 of them initially in the left hand urn and 25 in the right hand, gradually approach (see Fig. 8) a 50-50 distribution as "time increases" (more spins of the roulette wheel).

$$C = \begin{pmatrix}\text{chance that roulette wheel}\\ \text{causes ball on right to jump}\end{pmatrix} = \frac{50 - n}{100}$$

$$D = (\text{change of } n \text{ in such a jump}) = +1$$

$$\begin{pmatrix}\text{expectation value}\\ \text{of change in } n\end{pmatrix} = AB + CD = -n/50 \qquad (57)$$

Taking for the unit of time the interval between spins of the roulette wheel, and dealing only with averages or expectation values, one thus analyzing events at the simplest level finds a differential equation for approach to equilibrium,

$$dn/dt = -n/50. \qquad (58)$$

The solution of this equation shows the familiar feature of exponential approach to equilibrium,

$$n = 25\ \exp(-\text{time}/50), \qquad (59)$$

in agreement with the standard "law of cooling." Eq. (59) predicts that the expectation value of the "surplus number," n, in the left-hand urn will drop to 25/2.718 = 9.2 after 50 spins of the roulette wheel; to $\bar{n}$ = 3.4 after 100 spins; to $\bar{n}$ = 1.25 after 150 spins; and to $\bar{n}$ = 0.46 after 200 spins. However, superposed on this regular fall off--to be seen only by averaging over many independent runs, each starting with n = 25--will be the fluctuations about this average unique to any one individual run. These random variations quickly grow to a magnitude given to a good approximation by the familiar formula,

$$\overline{(N_{left} - \bar{N}_{left})^2} = \bar{N}_{left} \qquad (60)$$

implying a root mean square fluctuation in the "surplus," $n = N_{left} - 50$, given by

$$(\Delta n)_{RMS} = (\bar{N}_{left})^{\frac{1}{2}} = 7.07 \qquad (61)$$

These ups and downs drown out the tail of the exponential.

The fluctuations in the "surplus," n, in the left hand urn are not limited to the magnitude $(\bar{n})^{\frac{1}{2}}$. From time to time a larger variation occurs; and very rarely, a much larger one.

What is to be said about a much larger than average fluctuation as at the "Boltzmann peak" or point P in Figure 8. What about time asymmetry? There is none. The behavior prior to P is dominated by exponential rise as that later than P is dominated by exponential fall-off. In other words the one-sidedness in time of the exponential law of fall-off of the "surplus" has to be understood, not as an indicator of any asymmetry in time of the elementary process itself but as a consequence of the special initial conditions ("order" at P). We see here in an example as elementary as anyone has ever devised what is also apparent in the phenomenon of heat conduction and in the coupling of a complete absorber with an accelerated source by half-advanced, half-retarded potentials (Section VI).

All three processes--spins of the roulette wheel, molecular collisions, radiative coupling--convert ordered into disordered

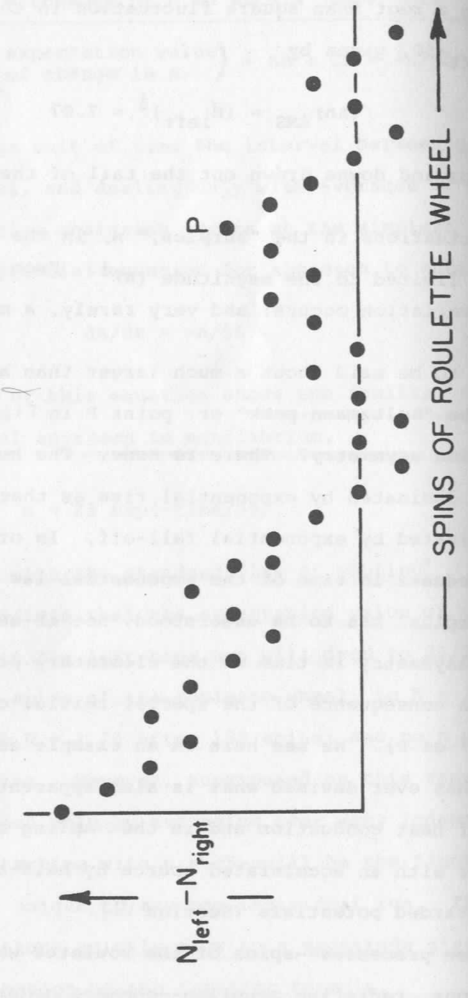


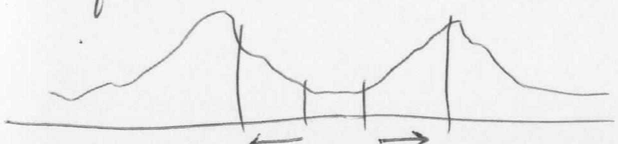
Figure 8. Approach to equilibrium, and fluctuations about equilibrium, as they show up in a typical "run" of the Ehrenfest double-urn experiment. The point P marks a larger-than-average fluctuation away from equilibrium. When one makes statistical run after statistical run, each run containing for example 300 spins of the roulette wheel in Fig. 7, one will find some runs in which statistical fluctuation brings $N_{left} - N_{right}$ at the end of the run back to its original value. When one averages over sufficiently many of the runs that satisfy these special end point conditions one washes out the statistical fluctuations and arrives at a cosh-curve (Fig. 9). If we rule out all the other--and much more numerous--runs as "might-have-been" but "never-were" runs, we have a model for what is meant by a universe ruled by "double-ended statistics."

states, although every elementary interaction is microscopically reversible. All three are epitomized by heat conduction. Wheeler and Feynman remark, (93) "A portion of matter observed at the present moment to be warmer than its surroundings will cool off in the future with a probability overwhelmingly greater than the chance for it to grow hotter. About the past of the same portion of matter Boltzmann's H-theorem however also predicts an enormously greater likelihood that the body warmed up to its present state rather than cooled down to it. In other words, we are asked to understand the present temperature of the body as the result of a simple statistical fluctuation in the distribution of energy through the entire system. This deduction is based on the premise that the system was isolated before observation. However, common experience tells us that the given portion of matter probably acquired its abnormal temperature, not via an internal statistical fluctuation, but because it had earlier not been isolated from the outside. For the radiative analogy of this example of heat conduction, conceive a charged particle bound to a position of equilibrium by a quasi-elastic force. Furthermore suppose its energy at the moment of observation is large in comparison with the agitation of the surrounding absorber particles. There is then an overwhelming probability that the oscillator will lose energy to the absorber at a rate in close accord with the law of radiative damping. What can be said of the particle prior to the moment of acceleration? In an ideal

absorbing system completely free of special disturbances, there is an equally overwhelming chance that the energy of the charge was then increasing at a rate given approximately by the inverse of the law of radiative damping. In this case as in heat conduction the abnormally high energy of the object is to be interpreted as the result of a statistical fluctuation. However, that the sun at some past age acquired its energy by such a fluctuation no one now would seriously propose. Obviously the universe is a special system with respect to the origin of which probability considerations cannot freely be applied."

How do these considerations bear on the Ehrenfest double urn? The first part of our response is immediate. We identify the point P with a statistical fluctuation. The dominant feature of n(t) before P is exponential rise, and after P an exponential fall off. However when we turn from the point P to the start of play, we do not suggest that n acquired the value n = 25 as a consequence of a prior and very large statistical fluctuation. On the contrary we understand n = 25 as an initial condition. That initial condition in the double-urn problem symbolizes the quiescence of the absorber <u>before</u> the acceleration of the source charge and--in the problem of heat flow--the initial condition of energy disequilibrium in the early universe. In other words the initial surplus, n = 25, symbolizes a cosmological boundary condition at the start of time.

Define "time" (local) as <u>entropy</u>.



Are any cosmological boundary conditions complete that do not deal with the end of time as completely as with the beginning of time? If the universe collapses to a big crunch as it begins with a big bang is it not as natural for requirements on particle motions and field configurations to be imposed at the end as at the beginning? Why not then see what are the consequences for the Ehrenfest double urn of imposing a final n = n" at a final time t = t" = T on the same footing as the initial n = n' at the initial time t = t' = -T? W. J. Cocke$^{(179)}$ was the first to ask and analyze this question for the Ehrenfest double urn. That analysis is carried further here.

What does it mean to impose on the double urn problem a final condition, n = n" = 25, symmetric to the initial condition, n = n' = 25? For definiteness let the length of play, 2T, be limited always to 2T = 200 spins of the roulette wheel. By that time the initial condition will be almost forgotten [$\bar{n} = 25 \exp(-200/50) = 0.46$] and fluctuations ($\Delta n_{RMS} = 7.07$) will dominate. Thus, let the 200 spin play be repeated over and over 10^9 times, each time starting with a surplus of balls on the left, n = n' = 25. The play will end with n = n" sometimes equal zero, sometimes +6, sometimes -10 and, very rarely, but occasionally, +25, the initial value. The probability that exactly this value is attained at the time t = t" = T = 100 is estimated most easily by neglecting altogether any "memory of the past" (surplus $\bar{n} = 0.46 \approx 0$) and treating each of the 100 balls as having a probability (1/2) to be in the left hand urn and an equal probability to be in the right hand one. All possible ways to distribute the balls between the two urns are contained--each with its characteristic probability--in the binomial expansion

$$1 = [(\tfrac{1}{2})_{\text{left}} + (\tfrac{1}{2})_{\text{right}}]^{100} \tag{62}$$

$$= \sum_{N_{\text{left}} + N_{\text{right}} = 100} \frac{(100)!}{N_{\text{left}}!\,N_{\text{right}}!} (\tfrac{1}{2})^{N_{\text{left}}} (\tfrac{1}{2})^{N_{\text{right}}}.$$

Therefore the desired probability to return at t = T = 100 to the surplus n = 25 on the left--in the stated approximation (low by not quite a factor 2, but uncorrected here)--is,

$$w(n=25) = \frac{100!}{75!\ 25!\ 2} \sim (2\pi)^{-\frac{1}{2}} \frac{100^{100.5}}{75^{75.5}\ 25^{25.5}\ 2^{100}} = 192\times10^{-9}. \tag{63}$$

In other words, out of the 10^9 repetitions of a 200-spin run, of the order of 192 will end up with $N_{\text{left}} = 50 + n = 75$. Let these ~192 histories be called "acceptable". Let existence be denied to all the others; let them be ruled out as "unacceptable", as "might-have-been" but "never born" universes. This is what we shall mean by speaking of "double-ended statistics".

What are the features of the typical history that is allowed by double-ended statistics? It is marked by an almost exponential decay of n at the beginning and an almost exponential rise of n at the end. Superposed on this general trend are the inevitable fluctuations. To iron them out we turn attention from the individual history to the average of all 192 acceptable histories. Better, increase the number of tries from 10^9 to 10^{12} and the number of acceptable histories from ~192 to ~192×10^3. Or multiply the number of trials by still further powers of ten. In this way reduce below any preassigned level the effect of the fluctuations which show so clearly in any one acceptable run and which still show a little when one takes the average of 192 acceptable runs.

The "ideal average run" in the sense just described follows a simple mathematical formula. There is a quick way to this formula: a differential equation. The appropriate differential equation is not the usual law of cooling,

$$dn/dt = -n/50 = -n . \tag{64}$$

That is asymmetric in time. The new law must treat the two directions of time symmetrically. It must make no reference to the initial time or the final time. It must make no reference to initial n' or final n". Those boundary value data must go into the final formula for n only as boundary value data. The only law that meets the physical requirements of the problem is one that treats exponentially rising and exponentially falling functions on the same footing,

$$d^2n/dt^2 = \lambda^2 n . \tag{65}$$

This is the law of change of n with time in double-ended statistics.

The general solution of (65) is a linear combination of $\exp(-\lambda t)$ and $\exp(\lambda t)$; or a linear combination of $\sinh \lambda t$ and $\cosh \lambda t$.

That solution which takes on the value n' at time t' and n" at time t" is given by the expression

$$n = [n' \sinh\lambda(t''-t) + n'' \sinh\lambda(t-t')]/\sinh\lambda(t''-t') . \tag{66}$$

This solution is characterized, for positive n' and n", by 3 regions (Fig. 9). The first is a region of nearly exponential fall off near t'. The last is a region of nearly exponential rise near t". Between is a region of transition from fall off to increase.

These considerations make a little clearer what it means to ask whether there is any correlation between statistics and cosmology. In further pursuance of this point, let "the turning of the tide" refer to the phase in the dynamics of the universe where expansion gives way to contraction, and let the term "the statistical turn of the tide" refer to the minimum in "the departure from equilibrium", as represented in Fig. 9. Even if there is any correlation between statistics and cosmology, it is not a necessary consequence of the reasoning that the statistical tide should turn at the same time as the cosmological tide, nor is it necessary that either time occur exactly half way between start and stop. There are few model universes easier to analyze in all detail than the Taub model universe(214-217). For the extreme time-asymmetric (large m') case of this model the volume varies with proper time in accordance with a relation which, written parametrically, is(217)

$$V = 32\pi^2 \ell^3 (m')^2 \sin f(1-\cos f),$$
$$\tau = \ell m'(f - \sin f) . \tag{67}$$

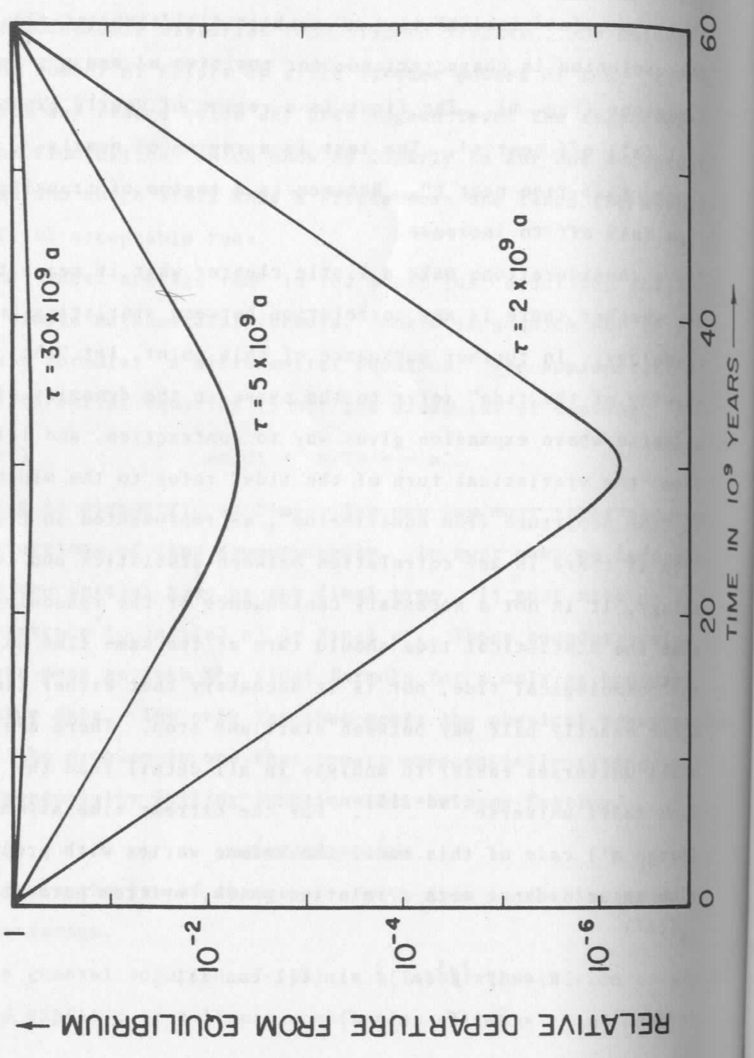


Fig. 9. Relative departure from equilibrium of temperature or number of radioactive atoms as a function of time calculated under quite schematized assumptions for one illustrative scenario [see Eq. (6c)] out of many equally conceivable alternatives; specifically: (1) initial and final departures from equilibrium identical; (2) total time available from start to end, 60×10^9 a; (3) symmetry in time; (4) no reaction chains; only one characteristic time relevant, $\lambda^{-1} = \tau = 30$, 5, or 2, in units of 10^9 a for the three cases illustrated. These are gross and highly arbitrary simplifications. The departure from standard exponential decay in the first half of time shows up strongly only in the last e-folding time before "turnaround."

Only for the special choice of the parameter, $m' = 0$, is the dynamics time-symmetric. Moreover, there is no obvious reason why the final value of $n = n''$ in the double-urn experiment should be identified with the initial value; or, to spell out the analogy, no obvious reason why the conditions at the big crunch should be in every way identical to those at the big bang. Moreover, the inescapable fluctuations that occur in any given history and that produce deviations from any idealized statistical law will normally be quite distinct in the descending and ascending phases of the curve of Fig. 9. Despite all these provisos and caveats, the simplest model makes the greatest appeal in any first sketch of the possibilities. In it the turning of the tide for the statistics is identical in its timing with the transition from expansion to contraction. Also both are mirror symmetric with respect to that common time.

The "homogeneity and isotropy" of the Friedmann model, if it applies to the universe at all, applies in the large, not in detail. Likewise "mirror symmetry in time", if it applies to the universe at all, applies in the large, not in detail. If it applied in detail the configuration of every part of the universe at the time $t_{\mathrm{mirror}} + t$ would have to be identical with its configuration at the time $t_{\mathrm{mirror}} - t$. That would mean that every motion would come to a halt at t_{mirror} itself. So detailed a requirement would plainly be incompatible with the motion of the planets around the Sun and the Moon around the Earth.

To accept double-ended statistics for investigation is to deal with no small change in familiar ideas of time and causality.

What is the observer of the roulette wheel to think as he watches the end of the play approaching? Spin after spin the wheel turns up predominantly the identifying numbers of the balls that lie in the right-hand and less occupied urn. Against all normal odds the smaller number N_{right} grows still smaller. And with the final spin of the wheel the numbers of balls in the two urns are restored as if by magic to their initial values $N_{\mathrm{left}} = N'_{\mathrm{left}} = 75$ and $N_{\mathrm{right}} = N'_{\mathrm{right}} = 25$. He would find this outcome utterly beyond understanding if he did not know that every history had been thrown out as impossible which did not end as it began with prescribed conditions.

With what words will one describe the biased probabilities spun out by the roulette wheel? "Bias" or "providence factor" are the only terms that immediately suggest themselves. A factor is at work that pushes the probabilities ever more strongly toward the predetermined end as the final time of reckoning approaches.

A providence factor defines itself naturally in the context of the smooth average number, $n(t)$, (average over many repetitions of a 200-spin run) in its dependence on time. It also shows up in quite another way in the biasing for or against certain numbers on the roulette wheel according as the balls so-numbered lie in the right-hand or left-hand urn.

Let us turn to the continuum description first as the simpler way to analyze this bias. We want to say that the surplus, $n = N_{\mathrm{left}} - 50$, decreases in time in accordance with the normal law of cooling except as modified by a bias of "unknown origin" that will see to it that the predetermined end is brought about.

Thus we write

$$\text{(68)}\qquad dn/dt = -\lambda n + \text{"bias term"}.$$

We compare this expression with what we get by taking the general solution (66), differentiating it once, and eliminating from the two expressions dn/dt and n the initial value n' (and, simultaneously, t'); thus,

$$\text{(69)}\qquad (dn/dt)\sinh\lambda(t'' - t) + \lambda n\cosh\lambda(t'' - t) = \lambda n'';$$

or

$$\text{(70)}\qquad dn/dt = -\lambda n + \underbrace{\frac{\lambda[n'' - ne^{-\lambda(t''-t)}]}{\sinh\lambda(t''-t)}}_{\text{"bias term"}}$$

In other words, we have defined the bias term in such a way that it should make reference to present value and final value alone and no reference at all to the initial value, n'. As solution to this requirement we find one and only one answer, the second term on the right hand side of (70).

The meaning of (70) is clear. Final requirements have no influence on present happenings so long as the time of reckoning lies many relaxation times in the future. However, as the time available for the final adjustments becomes of the order of a couple of relaxation times or less, the predestined end impresses itself on the game in an ever heavier bias. In the very last stage, only a few spins before the game must end, the normal decay rate is essentially without effect. The development proceeds practically deterministically to its end. In mathematical terms, the differential equation (70) reduces in this "last-moment limit" to the form

$$\text{(71)}\qquad dn/dt = \frac{n''-n}{t''-t}\ .$$

The solution is

$$\text{(72)}\qquad n = n'' - \text{constant}\ (t''-t);$$

in other words, single-minded straight-line progression towards the final goal, unmoderated by any influence of the relaxation constant, λ.

It is characteristic of "double-ended statistics" that one direction of time, t, is as good as the other, $\bar{t} = -t$, for describing it. The two very different looking equations,

$$\text{(73)}\qquad dn/dt = -\lambda n + \text{bias},$$

and

$$\text{(74)}\qquad dn/d\bar{t} = -\lambda n + \overline{\text{bias}},$$

or

$$\text{(75)}\qquad dn/dt = \lambda n - \overline{\text{bias}},$$

deal with two completely equivalent ways of describing the same time dependence, n = n(t), of the surplus in the left-hand urn. When the total length of the run, t"-t', amounts to many relaxation times, then one equation is "useful" near one limit, and the other equation is "useful" near the other limit. Here "useful" means that the term ±λn dominates, and the bias term is negligible by comparison. But either equation, and both, are valid for the entire stretch of time from t = t' to t = t". To say that the "providence factor" or "bias" is important or is negligible at such and such an epoch is therefore not a statement that is invariant with respect to the change of description,

$$(76) \qquad t \rightarrow \bar{t} = -t.$$

This existence of covariance but not of invariance under time reversal is reminiscent in some respects of the alternative descriptions of approach to equilibrium developed by Prigogine;(206, 218, 219) but the statement of boundary conditions at both ends of time is unique to "double-ended statistics."

It would be possible to go to the next step beyond the continuum description of Eqs. (65) and (70) and deal with fluctuations about the continuum description. Thus the number n(t) dealt with so far does not refer to any individual history. Rather it is the average over many acceptable histories, not 192 histories, not 192×10^3 histories, but 192×10^p histories, where the power p can be made large enough to guarantee approach to a continuum description to any preassigned degree of precision. When we turn to the characterization of individual histories in all their fluctuations about the continuum, the relevant quantity is the probability, w_n, that any given surplus of balls, n, will be found in the left hand urn. This probability will vary with time according to the equation

$$(77) \quad dw_n/dt = [50+(n+1)](w_{n+1}/100) - w_n + [50-(n-1)](w_{n-1}/100) + \text{"(bias term)"}_n$$

The first three terms follow from the elementary probabilities of Eq. (70). They will suffice to account for what goes on when many relaxation times intervene between "now" and the end. They give for the average value of the surplus on the left at any given time,

$$(78) \qquad \bar{n} = \sum_{n=-50}^{50} n w_n ,$$

the familiar cooling equation,

$$(79) \qquad d\bar{n}/dt = -\lambda\bar{n},$$

with $\lambda = (1/50)$. However, as the end comes nearer, the fourth or bias term begins to become effective. If at this stage the number of balls on the left does not measure up to the prescribed final number, this term sees to it that in the spinning of the roulette wheel (1) all those numbers show up with greater probability which belong to balls to be moved from right to left; and (2) all those numbers show up with decreased probability which represent balls to be moved from left to right. To state and derive the explicit formula for this bias term, to discuss the ostensible upper limit on the rate of change of $\bar{n}$ with time (one unit per spin of the roulette wheel), and to examine what it would mean to try to circumvent this limit by allowing negative values for jump probabilities, are all interesting questions; but they deflect attention from the main point: The double urn model of Cocke, as analyzed here, provides the simplest model that one can well imagine for what it means to speak of double-ended statistics.

Incentive though the double-urn model is for asking new questions about the universe, it is inadequate for answering them. One shortcoming is evident from the start. The double-urn model is characterized by a single transition rate, the λ of (79). In contrast, the universe is characterized by almost as many transition rates as there are physical processes, from elementary particle decay rates to the rates of thermonuclear processes in stars, and from the rate of dynamical evolution in star clusters to the rate of decay of turbulence. Nowhere does this limitation of the double urn model show more conspicuously than in the difficulties it makes for predictions about β-decay of ^{187}Re. Which is relevant, the 40 × 10^{9}a half-life for expulsion of the β-particle or the 10^{-12}s time for reducing the expelled β-particle to thermal equilibrium with its surrounding? Or a complex resultant of these two and many other characteristic times? The predictions of the double urn model, if one can call them predictions, are utterly different according as one correlates the characteristic decay constant, λ, of that model with the short time or the long time (Fig. 9), let alone some unknown third "resultant time constant."

In the one case the transition from exponential decay of ^{187}Re to exponential increase takes place within an extremely short internal of the turning of the tide. To hope to see

any evidence of that transition today, at a time when the universe is still expanding, would seem preposterous. However, if the long time of the β-decay itself is the relevant quantity, then the transition from fall to rise takes place gradually over the whole range from start to end (top curve in Fig. 9). In this case a significant difference in the effective half-life of Re^{187} might be expected as between today and 4.5 × 10^{9}a ago, when certain stony meteorites were formed.

Consider first the customary hypothesis that the decay has been exponential ever since the time, t_{form}, of the formation of the meteorite, and has continued to have a decay rate, the $\lambda_{apparent}$ of Eq. (55), equal to that found today$^{(22)}$,

(80) $$N_{form}(Re) = N_{now}(Re) \exp \lambda_{app}(t_{now} - t_{form}).$$

In this event the number of daughter ^{187}Os atoms that should have accumulated in the meteorite is

(81) $$N_{now}(Os) = N_{form}(Re) - N_{now}(Re).$$

Thus correcting for any primordial ^{187}Os present in the relevant granules of the meteorite, or verifying that the amount of primordial ^{187}Os was negligible, we have

(82) $$R = \frac{N_{now}(\text{daughter Os})}{N_{now}(\text{surviving Re})} = e^{\lambda_{app}\Delta t} - 1,$$

where we use the abbreviation

(83) $$\Delta t = t_{now} - t_{form}$$

Now ask how the situation will differ if ultimately there is to be a turnabout in statistics, a turnabout that already produces today a premonitory effect. Adopt a simple illustrative cosmology (Section IX, Table 2), with a big bang 10×10^9a in the past and a maximum in the expansion, or a turning of the tide, 20×10^9a into our future. Make further the purely illustrative assumption that the number of ^{187}Re atoms in an undisturbed meteorite is symmetric in time with respect to that same time, t_{tt}, of the turning of the tide. Then we have

$$N_{form}(Re) = N_{tt}(Re)\cosh\lambda(t_{tt} - t_{form}), \tag{84}$$

$$N_{now}(Re) = N_{tt}(Re)\cosh\lambda(t_{tt} - t_{now}). \tag{85}$$

Neither the number of ^{187}Re atoms at turnabout, N_{tt}, nor the true transformation constant, λ, is directly observable. The observable quantities are the apparent decay rate today,

$$\lambda_{app} = \lambda\tanh\lambda(t_{tt} - t_{now}), \tag{86}$$

and the ratio of accumulated ^{187}Os to surviving ^{187}Re,

$$R = \frac{N_{now}(\text{daughter Os})}{N_{now}(\text{surviving Re})} = \frac{\cosh\lambda(t_{tt}-t_{form})}{\cosh\lambda(t_{tt}-t_{now})} - 1. \tag{87}$$

In the limit where the time of turning of the tide is sufficiently far into the future ($t_{tt} \to \infty$), then statistical turnabout is destined never to arrive, and expressions (86, 87) reduce to the familiar result (82). However, for a value of $t_{tt} - t_{now} = 20 \times 10^9$a--a cosmologically reasonable order of magnitude--and a specimen that has been undisturbed for $t_{now} - t_{form} = 4 \times 10^9$a since formation, the calculated

TABLE 1. Calculated effect of future "turning of the tide of statistics" on amount of daughter ^{187}Os accumulated up to now in an ancient rock or meteorite containing ^{187}Re (present day apparent decay constant $\lambda_{app} = -dN/Ndt = 0.693/40$ billion years). There is a 7.8 percent difference between the two numbers marked in the table by arrows.

Time from now to "turning of tide" $t_{tt} - t_{now}$	λ(true) required to make $T_{\frac{1}{2}}$ (apparent, today) equal 40×10^9a	$R = \frac{N_{now}(\text{daughter Os})}{N_{now}(\text{surviving Re})}$ for age of meteorite, $t_{now} - t_{form}$,	
		2×10^9a	4×10^9a
5×10^9a	5.97×10^{-11}a^{-1}	0.0419	0.0987
10×10^9a	4.29×10^{-11}a^{-1}	0.0384	0.0843
20×10^9a	3.13×10^{-11}a^{-1}	0.0365	→ 0.0773
50×10^9a	2.18×10^{-11}a^{-1}	0.0356	0.0731
∞ (never)	1.73×10^{-11}a^{-1}	0.0353	→ 0.0717

accumulation of ^{187}Os (Table 1) is about 8 percent greater than one would have expected from the standard straightforward Rutherford-Soddy theory of radioactivity.

The calculated effect is so big in the case of the ^{187}Re-to-^{187}Os decay primarily because the relevant effective halflife, 40×10^9a, is so long. For the α-decay of ^{238}U, where the apparent halflife is 4.51×10^9a, the calculated accumulation ratio $R = [N_{now}(\text{daughter } ^{234}\text{Th})]/[N_{now}(\text{surviving } ^{238}\text{U})]$ in the same 4×10^9a-old rock or meteorite (provided that it keeps its decay ^{4}He) is increased only 0.24 percent (from 0.8490 to 0.8510) by a turning of the tide that lies ahead in the future by the same 20×10^9a. Forgetting this small correction, we can say that the ratio of daughter ^{234}Th to remaining ^{238}U tells the age of the mineral. This age once known, the past accumulation of ^{187}Os from ^{187}Re tests for a future turning of the tide.

The discussion given here for $^{187}\text{Re}(T_{\frac{1}{2}} = 40 \times 10^9\text{a})$ *versus* $^{238}\text{U}(T_{\frac{1}{2}} = 4.51 \times 10^9\text{a})$ can be extended to other familiar long-lived radioactive substances, such as $^{40}\text{K}(T_{\frac{1}{2}} = 1.3 \times 10^9\text{a})$, $^{87}\text{Rb}(T_{\frac{1}{2}} = 50 \times 10^9\text{a})$, and $^{147}\text{Sm}(T_{\frac{1}{2}} = 130 \times 10^9\text{a})$.

The apparent ages of $\sim 4 \times 10^9$a-old terrestial rocks and meteorites, as deduced from accumulations from the radioactive decay of three substances, U, ^{40}K, and ^{87}Rb, of very different apparent halflives, have been found compatible by Peebles and Dicke(221). Those ages would have been in observable discrepancy, one against the other, they conclude, if the fine structure constant were changing at a rate more than 3 parts in 10^{13} per year. On the other hand, if we assume no change in the fine structure constant, the same considerations will put an upper limit on the "turnabout effects" that we have been considering here. Otherwise stated, there is not the slightest evidence in the data cited sixteen years ago by Peebles and Dicke for anything in the way of an impending reversal of statistics coming up at a cosmologically reasonable time in the future.

The great advances that have taken place in radiochemical age determinations in the meantime give room for a reexamination of this question. Even more needed is a consistent theory of "doubled-ended statistics" in a fully cosmological context. How can such varied physical processes as heat conduction, thermonuclear reactions, electromagnetic radiation, and radioactive decay, with their very different chracteristic times, couple together to give an orchestrated turning of the tide? Until one has an answer to this question of theory, one will not really understand the first thing about what it means observationally to test for a "turning of the tide".

It is conceivable that one will someday understand the origin of initial value data so well that one can say that statistics of necessity always runs in one direction. Today we are not in that happy situation. Therefore at the moment it cannot be excluded that statistical turnaround occurs. If so, and if it can be detected, it will at one stroke, (1) give a cosmological foundation for statistical mechanics, (2) tell the scale of time from big bang to big stop and, thus (3) provide evidence that the universe is closed.

VIII. Memory and the Arrow of Time

"It's a poor memory that only works backwards."

--White Queen to Alice.(222)

"If physics is four-dimensional, and if past, present, and future are all laid out shiningly in one vast spacetime diagram, why is there any "now" in our apprehension of physics? Nothing has done more to suggest to some of us a way out of this mystery than some comments made in conversation by Hugh Everett. He compares the brain of the observer with a servomechanism, or--if I may go beyond Everett in explicitness--the computer of an aircraft gun. The radar unit mounted on the gun carriage sights on the enemy plane. [Fig. 10] Minute by minute it feeds information about the position of that plane into the computer. From this information the computer extrapolates the future position of the plane. It then fires a shell to intercept that plane an appropriate number of minutes later. The computer thus carries within it information about a few minutes of past history--and also information about a few minutes of forthcoming history.

"It would be possible for the computer to remember more, perhaps the position of the enemy plane yesterday. But that outdated information would be of no use in the present crisis. Remembering it would only impose a more complicated burden on the electronics and increase the weight to be hauled along as the gun is moved from site to site. Similarly, the computer can be forced to extrapolate the flight of the enemy plane over a much greater reach of time, even to this hour tomorrow. However, that prediction would obviously

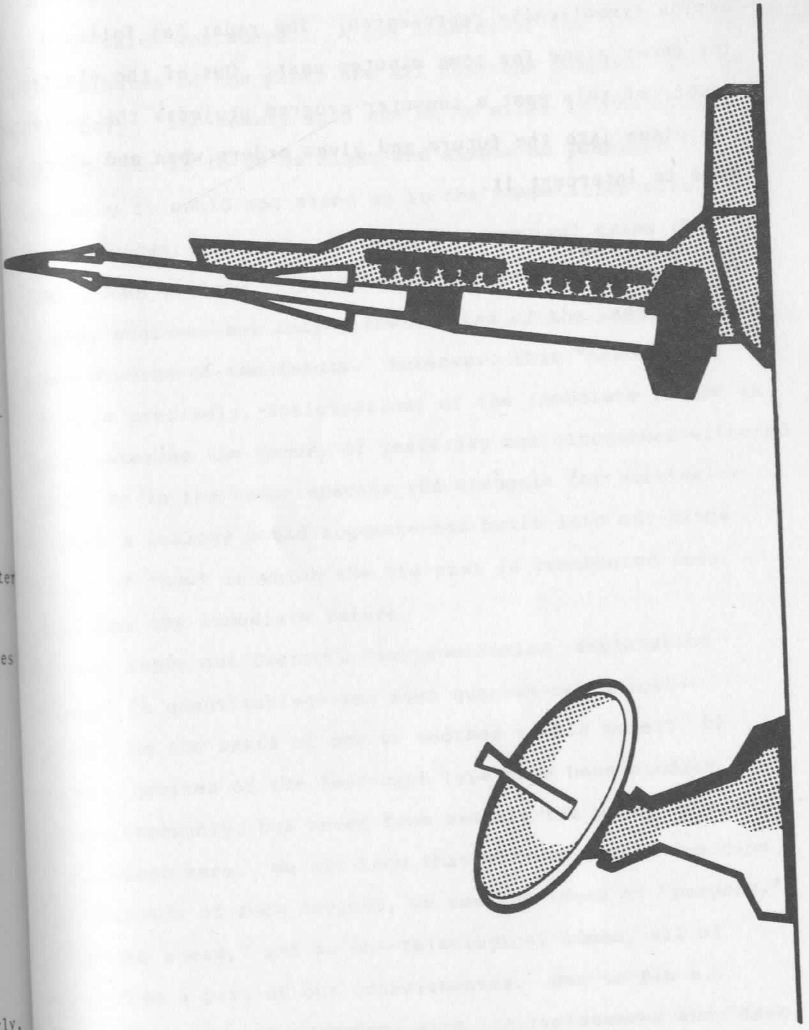
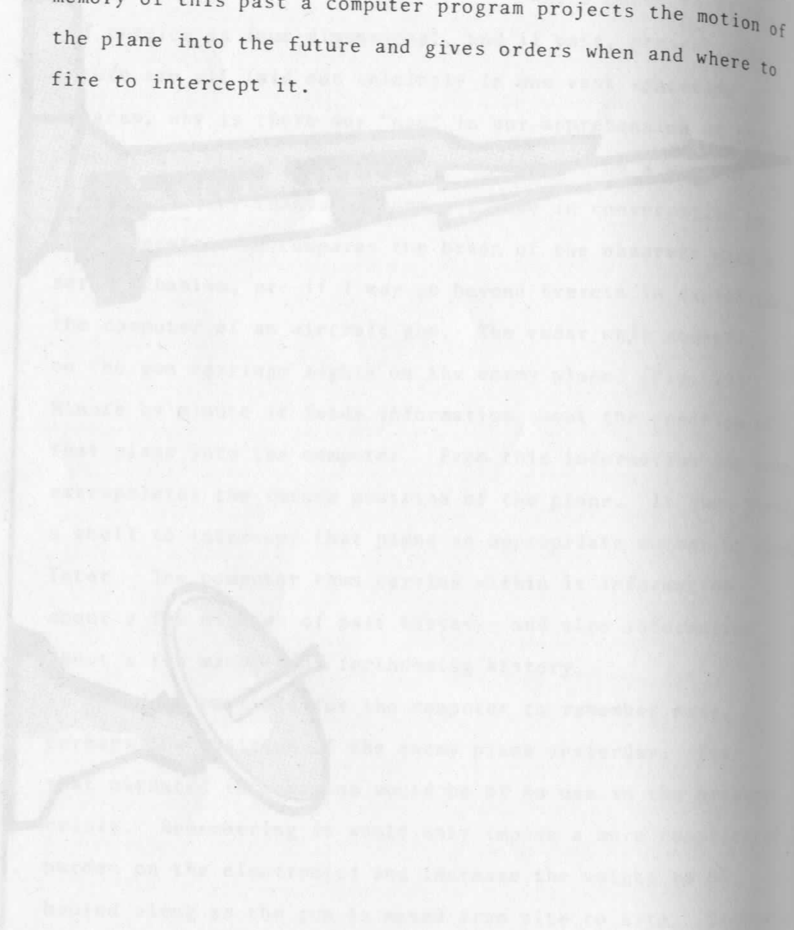


Figure 10. Role of the memory in linking observation with action symbolically represented. The radar has followed the enemy plane for some minutes past. Out of the electronic memory of this past a computer program projects the motion of the plane into the future and gives orders when and where to fire to intercept it.



have no value whatsoever. A few minutes of the future, like a few minutes of the past, are all that the computer memory will carry. The memory span can be no wider if the anti-aircraft gun is to be as light and simple as possible. Otherwise it could not stand up in the competition with rival devices. Thus the struggle for survival trims the memory down to "now." This "now" is remarkable. On it are vividly engraved not only a few minutes of the past, but also a few minutes of the future. Moreover, this "memory" (or more precisely, anticipation) of the immediate future is green, whereas the memory of yesterday has altogether withered away. So in the human species the struggle for survival--Everett's analogy would suggest--has built into our minds a type of "now" in which the old past is remembered less well than the immediate future.

Can one trace out Everett's "servomechanism explanation of now" in quantitative--and even quantum-mechanical--detail, on the basis of one or another simple model? Of course, devices of the feed-back type have been studied quite thoroughly, but never from exactly the point of view of interest here. We all know that when we try to describe the behavior of such devices, we use the ideas of "purpose," "planning ahead," and so on--teleological ideas, all of which form a part of our consciousness. But to fit a description of such a system, with its resistances and "dash pots", into the Hamiltonian formulation of the kind we now require is an "analysis that has not yet been undertaken." [The foregoing quoted from Ref. 223); for investigations on the mechanism of memory see for example Ref. (54-56).]

IX. The Gates of Time(96)

"...every substance...can only begin by creation and end only by annihilation."---G.W. Leibniz(224)

Given memory, uncover the machinery of memory: that was the challenge of VIII. In IX the concern is different. Given a machinery of memory--dynamics--explain how that machinery can ever stop remembering.

Not the slightest warrant does Einstein's equation give for thinking there can be any such thing as a "before" before the big bang or an "after" after the big crunch or after the collapse of a star to a black hole. These three processes mark three "gates of time".

For time to come to an end is to say that time is not an ultimate category in the description of nature. Therefore a deeper description of nature must transcend the category of time: this is the conclusion suggested by a review of available evidence on cosmology, theoretical and obervational; this is the theme of this final "frontier of time".

The characteristic feature of a gate of time is collapse to a singularity, not only for matter but also for the space geometry that envelops this matter. Moreover, at a singularity Einstein's field equation loses its applicability. If the mathematics fails at the singularity, how can one argue consistently about the physics at the singularity? How then can there be any foundation for believing that time ends at a gate of time? The point is simple. Time does not today stand in splendid isolation, a concept with an independent existence of its own, free of entangling alliances with the rest of physics. General relativity has subdued the concept time to membership in a larger kingdom: spacetime. There is nothing that we know about time, there is nothing we do with the concept of time, there is no meaningful attribute of time that is not subsumed, defined and given meaning through Einstein's 1915 and still standard geometrodynamics. Equations stop and time go on? That might once have seemed conceivable. Today it is not. Time has been robbed of the power to go off on a voyage of its own. There is no time today except the time of spacetime. Where the one stops, so does the other. Time before spacetime? That is a question, a proposal, a commentary to which one does not even know how to give the smallest shred of meaning.

Story though this is in brief of the gates of time, it is a story that can and must receive expansion in the rest of this section. Six topics will come into consideration: (1) the validity of general relativity, (2) evidence for the big bang, (3) do black holes exist? (4) is the universe closed? (5) will it collapse? (6) what happens to a black hole when the universe collapses?

First, how certain is one that the particular description of spacetime that is given by Einstein's general relativity is the most reasonable one? On that point the available evidence is summarized in Ref. (73): experimental tests in Chapt. 38; analysis of alternative theories and their difficulties, Chapt. 39; solar system experiments, Chapt. 40; gravitational waves and possibilities of detecting them to get new tests of relativity, Chapts. 35-37; cosmology and its relevance to general relativity, Chapts. 27-31; gravitational collapse and the theory of the black hole, Chapts. 32-34.

In brief "no theory more resembles Maxwell's electrodynamics in its simplicity, beauty, and scope than Einstein's geometrodynamics. Few principles in physics are more firmly established than those on which it rests: the local validity of special relativity, the equivalence principle, the conservation of momentum and energy, and the prevalence of second-order field equations throughout physics. Those principles and the demand for no 'extraneous fields' (e.g., Dicke's scalar field) and 'no prior geometry' lead to the conclusion that the geometry of spacetime must be Riemannian and the geometrodynamic law must be Einstein's.

"To say that the geometry is Riemannian is to say that the interval between any two nearby events C and D, anywhere in spacetime, stated in terms of the interval AB between two nearby fiducial events, at quite another point in spacetime, has a value CD/AB independent of the route of intercomparison. There are a thousand routes. By this hydraheaded prediction, Einstein's theory thus exposes itself to destruction in a thousand ways.

"Geometrodynamics lends itself to being disproven in other ways as well. The geometry has no option about the control it exerts on the dynamics of particles and fields. The theory makes predictions about the equilibrium configurations and pulsations of compact stars. It gives formulas for the deceleration of the expansion of the universe, for the density of mass-energy, and for the magnifying power of the curvature of space, the tests of which are not far off. It predicts gravitational collapse, and the existence of black holes, and a wealth of physics associated with these objects. It predicts gravitational waves. In the appropriate approximation, it encompasses all the well-tested predictions of the Newtonian theory of gravity for the dynamics of the solar system, and predicts testable post-Newtonian corrections besides, including several already verified effects.

"No inconsistency of principle has ever been found in Einstein's geometric theory of gravity. No purported observational evidence against the theory has ever stood the test of time. No other acceptable account of physics of comparable simplicity and scope has ever been put forward."(73)

Second, how certain are we of the initial gate of time, the big bang? No one has found any way to escape the big bang acceptable to even 10% of the community of physics and astrophysics. The reason is simple. There is too much evidence that is correlated by the concept of a big bang that has not been brought into line by any other reasonable proposal: the Hubble expansion(225); the primordial cosmic fireball radiation(226-229); the time scale of the astrophysical evolution of stars(230) and star clusters(231); and the physical conditions and the time required for the formation of the elements(232,233).

Third, ever so much nearer in time to today than big bang or big crunch is the black hole, the second gate of time; but how certain are we that there is even one genuine black hole anywhere in the universe? We can agree with the words of Laplace in 1795. It "would not, in consequence of its attractio

allow any of its rays to arrive at us; it is therefore possible that the largest luminous bodies in the universe may, through this cause, be invisible."(234) We can recognize that a neutron star, member of a double-star system, can receive almost unlimited amounts of matter via stellar wind from an appropriate companion in a double-star system(235,236). We can note the absolute inescapability, in theory, of collapse for such a system when it attains more than a critical mass, less than three solar masses according to the best available estimate(237,238). But has collapse to a black hole actually happened anywhere?

There is no absolutely compelling evidence for a black hole today. However, no one sees any other reasonable way to account for the unusual properties of the compact x-ray object Cyg X-1 discovered by R. Giaconni and his collaborators in 1973(239). "The optical component of the pair moves back and forth a distance of 5.2×10^6 km in the line of sight with a 5.6 day period. Its mass, from two lines of evidence (spectral character and absolute luminosity), is concluded to be of the order of 25 solar masses. The invisible component, in order to swing by its gravitational pull so big a visible mass back and forth so great a distance and so quickly has to be of the order of 10 solar masses, and certainly greater than 5 $M_\odot$, one reasons. An ordinary star of this mass would be quite visible in the optical, quite invisible in the x-ray spectrum. This is not an ordinary star. Moreover, it is too heavy to be a white dwarf or a neutron star. No one sees any natural and reasonable interpretation for it except as a black hole. [Of course the] x-radiation does not come out of a black hole. It comes out of gas on its way towards the black hole from the normal star. Gas is drawn in towards the compact component by its powerful gravitational attraction. In the ensuing 'traffic jam' it is compressed and heated. . .to temperatures so high that the gas cannot avoid emitting x-rays before it reaches the horizon of the black hole."(240)

No hope to make compelling identification of a black hole is today the focus of more numerous and more active investigations than the signature of such an object: that combination of fluctuations in time and spectral characteristics of x-ray emission which will divide putative black holes unambiguously into true and counterfeit.

In parallel with the search for black holes of few solar masses goes the search for black holes of a million solar masses or more. The characteristic distance associated by general relativity with an object of mass, m, is its Schwarzschild "radius", $2Gm/c^2$, which amounts to 29.4 km for a black hole of 10 solar masses, but for such an object with $10^6\ M_\odot$ amounts to $2.94x10^6$km or about 0.02 times the distance from the Earth to the Sun equals 0.02A.U.).

Oort(241) gives evidence that is reasonable to think of the center of the Milky Way containing a black hole of mass $\sim 4x10^6 M_\odot$. The relevant region is too obscured by intervening dust to be seeable in the visible spectrum but it can and has been investigated via radio waves and infrared. Stars are

unmistakably present at distances of 100A.U. from the center. Moreover their Doppler shifts can be measured. From the velocity and distance from the center one can deduce the amount of mass sufficient to curve such rapidly moving objects into orbits so great, $\sim 6x10^6 M_\odot$. On the other hand, from the luminosity in this region one concludes that the amount of the mass in the form of stars may be only $\sim 2x10^6 M_\odot$. Oort tentatively attributes the difference, m $\sim 4x10^6 M_\odot$, to a single black hole. His original paper has to be read for a careful statement of caveats and consequences.

Tentative evidence for a still more massive black hole, m $\sim 5x10^9 M_\odot$, at the center of the galaxy M 87 has been reported still more recently(242,243). The evidence comes from studying the distribution in luminosity with very high resolution very close to the center of M 87. The investigators looked at "slices" of the telescopic image of the galaxy M 87 at different distances from the center. They used Doppler measurements to tell how fast the stars are moving in each slice. The stars near the center are moving much faster than one would expect, and as if orbiting around a concentrated but invisible object of mass $\sim 5x10^9 M_\odot$.

It is not possible to say that the present evidence incontrovertibly establishes the existence of black holes. However, the evidence is sufficiently impressive to make one comfortable about accepting two very general considerations: the possibility of such objects is an inescapable consequence of general relativity; and there are several very plausible astrophysical scenarios, the inevitable outcome of which is the formation of a black hole.

Fourth, any evidence that the universe will some day collapse up against the third gate of time, dealing as it does with the future, necessarily contains a strong component of theory. Of this the most important ingredient is closure of the universe. "Is the universe open or closed? On no central issue of cosmology is there greater divergence of evidence today. Einstein's philosophical arguments speak for closure. [So does an appreciable body of physical evidence.] An appreciable body of astrophysical evidence speaks against it.

"To determine the so-called deceleration parameter q_0

$$q_0 = -\frac{d^2(\text{radius of universe})/dt^2}{(\text{radius})}\left[\frac{\text{radius}}{d(\text{radius})/dt}\right]^2 \tag{88}$$

from source counts is the goal of some of the greatest and most skilled observers of our times. This important measurement nevertheless requires such care in interpretation, demands so many corrections, and is afflicted with such uncertainties that the final number still today leaves the door open to either cosmology(244).

'The quickest way to see that the expansion may be slowing down is still the most elementary. One has only to compare the actual time back to the start of the expansion, a time of the order of 10×10^9 years, as judged from the rate of evolution of stars and clusters of stars, with the apparent, or extrapolate or Hubble time of $\sim 20 \times 10^9$ years. This is the time it would have taken galaxies to get to their present separations from us, moving with their present separation velocities, with no

allowance for the greater velocity in times past. Of course, considerable uncertainties attend both numbers, uncertainties of the order of 30 percent or, conceivably, even more. Even so, it is difficult to find evidence more impressive anywhere else in cosmology for the predicted slowing down of the expansion. "If to fix ideas we take the two numbers, 10×10^9 years and 20×10^9 years, as 100 percent accurate and assume a homogeneous isotropic spherical universe and neglect the pressure and energy content of radiation in comparison to the mass energy of inchoate material ('dust') then Einstein's theory straightforwardly gives all the other illustrative numbers of Table 2. The 30-fold discrepancy between the density of the universe today as called for by these calculations and the density estimated by Oort(245) gives rise to the well-known 'mystery of the missing mass'.. Of all the evidence for a low density cited by Gott, Gunn, Schramm, and Tinsley(246) and by Gunn and Oke(247), none is more impressive than the abundance of primordial deuterium. The sensitivity of the deuterium abundance to density arises. .

from the dependence of the expansion rate on density and from the fact that only a few minutes are required for primordial neutrons to decay to protons. Unhappily less satisfactory than this theoretical side of the study is the observational evidence. Determinations of deuterium abundance are made by looking at the absorption of light in interstellar space on its way from a star to the telescope. Only a few such determinations have been made. No one knows how representative are the samples of gas intervening nor how much they have been altered

TABLE 2. Major features of the universe according to Einstein's theory, as normalized by two key astrophysical data, each believed uncertain by an amount of the order of 30%: (1) the actual time, $\sim 10 \times 10^9$ Yr, back to the start of the expansion, as determined from the evolution of the stars and the elements, and (2) the "Hubble time", or time linearly extrapolated back to the start of the expansion, $\sim 20 \times 10^9$ yr, that is, the time needed for galaxies to reach their present distances if they had always been receding from us with their present velocities (adapted from Ref. 73).

Illustrative values all derived from	
Time from start to now	10×10^9yr
Hubble time now	20×10^9yr
Hubble expansion rate now	$49.0 \dfrac{\text{km/sec}}{\text{megaparsec}}$
Rate of increase of radius now	0.66 lyr/yr
Radius now	13.19×10^9 lyr
Radius at maximum	18.94×10^9 lyr
Time, start to end	59.52×10^9 lyr
Density now	14.8×10^{-30} g/cm^3
Amount of matter	5.68×10^{56} g
Equivalent number of baryons	3.39×10^{80}

between primordial times and today by cosmic ray impacts and contaminated by ejecta from stars and supernovae.

New light on missing mass comes from the recent work of Ostriker and Peebles(248) and Ostriker, Peebles and Yahil(249). They give arguments from [the gravitational theory of] galactic stability that the mass of the typical galaxy must be of the order of 3 to 20 times as great as one has previously estimated. They give reasons to believe that this matter is in the form of stars of modest mass and very low luminosity. Happily for the subject, the direct observational search for this 'halo' is now underway."(250)

Quite another way to get at the effective overall density of matter in the universe has been developed by Peebles and his associates(251). The focus of attention in this work is galaxy clustering and the correlation in space between galaxies. What comes into play here is the force of gravitation, which one understands, and the density, which one wants to understand. Negligible by comparison are other factors such as radiation pressure, degree of ionization, opacity and nuclear reactions, important though they are in the internal machinery of individual stars and galaxies. This enormous simplification in the analysis opens the door to meaningful statistical analysis of the correlation in position between galaxies and its change in time. Davis, Groth and Peebles(251) find that the logarithm of an accurately defined correlation function, plotted as a function of the logarithm of the angular separation between galaxy and galaxy, shows a sharp break in slope at a separation, $r \sim 9h^{-1}$Mpc $= 2.8 \times 10^{25}h^{-1}$cm. Here h is the ratio of the actual Hubble expansion rate today, whatever measurements of high precision may someday disclose it to be, to the working figure of 55km sec^{-1} Mpc^{-1}, adopted for convention's sake. This behavior is reproduced by gravitation theory when Peebles and his collaborators assume a simple power law spectrum of initial perturbations and density in the early universe and when they assume in addition that the actual magnitude of the density today is that called for by the condition of closure. In contrast when a substantially lower density is assumed the calculations based on gravitation theory fail to produce the observed break in the distribution function. They conclude, "the analysis presented here yielding $\Omega \gtrsim 0.3$ [a density greater than about 30% of the requisite amount], conflicts with arguments based on other lines of evidence that have been taken to show $\Omega \lesssim 0.1$ [that the density is less than about 10% of the requisite amount] (e.g. Gott *et al*(246)). Our approach will require considerable further work before it can offer a definitive constraint on the cosmology. On the other hand, we consider that the same applies to the other methods of estimating Ω, so this discrepancy is an interesting indication that something is not well understood but, at the moment, it is hardly a serious problem for the gravitational instability picture." [End of quote from ref. (251); following is completion of quote from ref. (250).] "It is difficult to name any single issue in all of astrophysics which draws together a wider variety of important investigations than those going on today concerning in one way or another the mystery of the missing mass.

"It has often been suggested that one should make a direct geometrical determination of the curvature of space in the large. In this way, the hope has been expressed, one could find out whether the universe is closed or open even prior to a reliable determination of the average mass density of the universe. More than one calculation has been made and reported(252) of the apparent angular diameter of an object of standard dimensions (if there be any such) as a function of distance (as defined by red shift). In Euclidean space, a 'standard' object has an apparent angular diameter which decreases in inverse proportion to distance. However, when the object is far enough away in an ideal spherical space, it is magnified by a kind of lens effect. Then the apparent angular diameter, rather than decreasing, increases with distance. Moreover, the double radio sources associated with quasistellar objects offer a conspicuous 'ruler'. If anything, the length of this 'ruler' will be shortened in early double radio sources as compared to more recent ones by the greater density at early times of the matter through which the 'twin exhausts'(253,254) have to plough their way. Thus if double radio sources of a sufficiently great red shift were to begin to show an increase in apparent angular diameter, one could hardly do anything but regard this effect as evidence for the predicted lens effect.

"A closer consideration shows that the situation is by no means as simple as would be indicated by these elementary considerations. It was already pointed out by Zel'dovich(255) and by Dashevsky and Zel'dovich(256) (references to this and the subsequent literature in Press and Gunn(257)) that the clustering of matter into galaxies, deviation from uniformity unimportant for the question of openness or closure, is vitally important for the focusing process. A spray of light rays that starts at a point, and spreads out as it goes, *continues* to spread out as it travels through matter-free interstellar space, even though the universe itself is contracting. Nothing like the elementary focusing effect takes place. ...R.C. Roeder(258) [stresses] the difficulties posed by this circumstance for any proposed cosmological test of closure, via measurement of apparent angular diameters as a function of red shift. However, if one hope fades another brightens. Press and Gunn(257) show that [condensed objects present to] a cosmologically significant density [have a] high probability to cause a distant point source to be gravitationally imaged into two roughly equal images--an effect with testable consequences." In spite of these difficulties hopes remain very much alive that someday an astrophysical means will be found to determine the large scale curvature of space. Among these hopes the conceivable anomalies in radioactive decay rates cited in VII may also be mentioned. Much astrophysical work, and justified astrophysical work, is underway to get a 'yes', 'no' answer to the simple question: Is the universe closed? However much as this issue belongs to science and however important general relativity is in dealing with it, one cannot forget that this science and this tool took their birth in philosophy. Therefore, it would be unbalanced not to quote Einstein's own considerations about closure, 'Thus we may present the following arguments against the conception of a space-infinite, and for the conception of a space-bounded, universe: (1) From the standpoint of the theory of relativity, the condition for a closed surface is very much simpler than the corresponding boundary condition at infinity of the quasi-Euclidean structure of the universe. (2) The idea that Mach expressed, that inertia depends upon the mutual action of bodies, is contained, to a first approximation, in the equations of the theory of relativity; ...But this idea of Mach's corresponds only to a finite universe, bounded in space, and not to a quasi-Euclidean infinite universe.'(259) In another place Einstein(260) states, 'In my opinion the general theory of relativity can only solve this problem satisfactorily if it regards the world as spatially self-enclosed.'"

In our own time a fresh consideration argues for closure: the difficulty of any alternative. [The following is quoted from ref. (250).] "[T]he 'initial value data' are essential in formulating what general relativity is all about. There are alternatives to closure as part of the formulation of the initial value data but no alternative so simple as closure. It is one alternative to postulate asymptotic flatness at infinity. It is another alternative to postulate more particularistic data on some closed 2-surface that bounds the 3-geometry embraced in the 'initial value problem'. What kind of data should be given on such a 2-surface? Mathematical tools we have on hand to try to answer such a question, but no slightest hint of any physical consideration that would make this a reasonable route to follow. And aymptotic flatness (see, for example, the 'heirarchical cosmology' of Alfven and Klein(261) and

De Vaucouleurs(262) makes double difficulties. First, it takes the geometry of faraway space out of physics and makes it part of theology, to be discovered by reading Euclid's bible. It puts us back to the days before Riemann, days when as Einstein(26 puts it, '...space was still, for them [physicists], a rigid, homogeneous something, susceptible of no change or conditions. Only the genius of Riemann, solitary and uncomprehended, had already won its way by the middle of the last century to a new conception of space, in which space was deprived of its rigidity, and in which its power to take part in physical events was recognized as possible.'

"Why accept this advance for near space and undo it for faraway space? Moreover, 'asymptotic flatness' leaves one lost. How can anyone even define the idea of asymptotic flatness? According to the most elementary considerations of quantum theory there is no such thing as *the* geometry of space. Geometry is not deterministic, it is probabilistic. There is a probability amplitude $\psi({}^{(3)}\mathcal{G})$ for this, that, and the other 3-geometry that differs from the first by an amount of the order $\Delta g \sim L^*/L$ in a region of order L. Thus, no matter how 'far away' one goes, one can never arrive at a place where the fluctuations have less than standard strength. Difficult as it is under these circumstances to define 'far away', it is even more difficult to see where else one can turn for a satisfactorily sharp boundary condition compatible with quantum fluctuations, except to *closure*."

Fifth: Will there be a big crunch? No factor bears so directly on this point as the question of closure. As simplest illustration, the distinction may be recalled between the Friedmann open model universe with the metric

$$ds^2 = -dt^2+a^2[(d\chi)^2+(\sinh\chi)^2(d\theta^2+\sin^2\theta d\phi^2)] \tag{89}$$

and the closed universe with the metric

$$ds^2 = -dt^2+a^2[(d\chi)^2+(\sin^2\chi)(d\theta^2+\sin^2\theta d\phi^2)]. \tag{90}$$

In one case the scale factor, a, and the time, t, are connected with each other by the relation,

$$\begin{aligned} a &= (a_0/2)(\cosh\eta-1) \\ t &= (a_0/2)(\sinh\eta-\eta) \end{aligned} \tag{91}$$

and both are ever growing quantities. In the case of the closed Friedmann model universe, things come to an end after a finite time and at that third gate of time the radius itself falls to zero,

$$\begin{aligned} a &= (a_0/2)(1-\cos\eta) \\ t &= (a_0/2)(\eta-\sin\eta). \end{aligned} \tag{92}$$

However, when the 3-sphere model universe is replaced by a 3-torus universe of repetition length, L(t), then the story is quite different. The dimension, L(t), following the big bang increases forever(264). This type of closed space has not been explored enough to know what its difficulties are. In default of the deeper analysis that is required we shall exclude it from attention. When we speak of a closed model universe, we shall mean a geometry that has in the large the qualitative character of a 3-sphere, however much it may be pocked in the small with multiple connectedness: wormholes or handles.

A model universe that has the topology of a 3-sphere, that obeys Einstein's geometrodynamic law, and that contains a nowhere negative density of mass-energy, almost inevitably develops a singularity according to reasoning traced out in successively greater detail by Tolman(265), Avez (266), Geroch(267), Hawking and Penrose(268), and Hawking and Ellis(269). Only in a set of cases of measure zero does the system escape big bang or big crunch or both, it is widely believed.

An illustration is provided by the Taub model universe(270). This model universe is of exceptional theoretical interest and simple in this respect, that it is curved up into closure neither by matter nor by electromagnetic radiation, but by gravitational radiation alone, and this of the longest wave length that will fit into a closed universe(271,217). Despite the fact that the volume of this system is zero at a time, t_1, and at another time, t_2, with one maximum in between, the geometry does not become singular at either t_1 or t_2. Instead it transforms itself smoothly and continuously(272) into another topology, one where there exists closed time-like lines. Such a spacetime contradicts every normal idea of causality. In it past and future are inextricably confused.

Ellis and King have given other examples of such "whimper" model universes, that just barely escape the singularity of big bang or big crunch(273). The transition from closed to open geometry had been investigated in detail only in the case of Taub universes(272). It is found that the "continuity is

achieved only at the cost of having certain classes of world lines spiral round the universe in the final stages of its collapse to tighter and tighter packing. Thus the presence of the slightest 'real matter' builds up an ever-increasing density (in this connection, see also Penrose(274)). As it goes to infinity, this density destroys the relevance of the model with which one started. One returns to something closer to a Friedmann cosmology with a Friedmann singularity."(250)

Near both gates of time, it is reasoned in several interesting papers by members of the Moscow group(275-282,154) the singularity in the generic case is characterized by a general "mixmaster oscillation" of the local geometry(283) with the phase, amplitude and orientation of the principal axes of this deformation of the geometry varying from point to point (see also Eardley, Liang and Sachs(284)). If this is the characteristic behavior of the generic solution then, it is suggested by Belinsky, Khalatnikov and Collins(285), the "whimper" solutions form a set of measure zero among these generic solutions. (See also Wheeler(286).) If these tentative conclusions are sustained by more detailed mathematical analysis then one would seem justified to say that the big crunch is "almost inevitable".

Sixth, the singularity of the black hole is not separate and distinct from the singularity of the big crunch in a model universe that collapses to a singularity. The icicle that hangs from the ceiling of a cave of ice is not separate and distinct from the ceiling. The one is part and parcel of the other. How best to bring into mathematical evidence this point, first suggested by Penrose(287) [see also ref. (288)], is a point under active investigation. One proposal has it that spacetime is best foliated by a 1-parameter family of spacelike hypersurfaces distinguished one from another by the value of the trace of the extrinsic curvature, constant on each hypersurface but differing from hypersurface to hypersurface. As successive members of this family are examined, each higher within the ice cave than the one before, none will touch the hanging icicle. Instead each will envelop it more closely than the one before, after the manner of a glove. The value of the trace, apart from a numerical constant is identical with the so-called York time(289). With respect to growing values of this time parameter one expects to see the mixmaster oscillations of Belinsky, Khalatnikov and Lifshitz played out. The black hole shows itself up, not as something new and strange, but as a special case of the mixmaster oscillations. Ahead though this description is of what the mathematics of the moment allows one to say, and afflicted though it is with some uncertainty as to the appropriate scheme of foliation, it nevertheless puts together the major features of the best thinking of today as to how the generic singularity is approached. When there are several black holes they coalesce: but the singularity of the individual black holes and of their coalescence, are still described in terms of deformation oscillations of the geometry as the foliation parameter rises without limit.

This review of the three gates of time shows in what sense and with what caveats and with what degree of certainty one can

say that the universe begins with a singularity and ends with a singularity. Little escape is evident from these words: there is no "before" before the big bang and no "after" after the big crunch. Time ends with spacetime. The universe does not endure from everlasting to everlasting. Everything came from "nothing". Of all the frontiers of time examined here, that one would seem to be most pregnant with the future.

that the universe begins with a singularity and ends with a singularity. Little wonder is evident from these words there is no "before" before the big bang and no "after" after the big crunch. Time ends with spacetime. The universe does not endure from everlasting to everlasting. Everything came from nothing. Of all the frontiers of time examined here this one would seem to be most pregnant with the future.

References

[Note to editor and publisher: Please *leave titles of journal articles in these references* because otherwise there is very great difficulty for the reader to know *which* references he wants to look up. Also please leave last page number as well as first page number of references so reader can tell whether article is a minor note or a major contribution. I shall be most appreciative if you will not change either of these features of the references without first consulting me about any proposed changes. Many thanks! --Author]

(1) C. Lyell: *Principles of Geology, being an attempt to explain the former changes of the earth's surface, by reference to causes now in operation,* (London, 1830-1833), Vol. II.

(2) C. Darwin: as quoted in H. Ward, *Darwin: The Man and His Warfare* (Indianapolis, Indiana, 1927), p. 297.

(3) J. A. Wheeler: "From relativity to mutability," in *The Physicists' Conception of Nature*, J. Mehra, ed. (Dordrecht, Holland, 1973), pp. 202-247.

(4) J. Mehra, ed.: *The Physicists' Conception of Nature* (Dordrecht, Holland, 1973).

(5) M. Eigen: "The origin of biological information," in *The Physicists' Conception of Nature*, J. Mehra, ed. (Dordrecht, Holland, 1973), pp.594-632.

(6) M. Eigen and R. Winkler: *Das Spiel: Naturgesetze steuern den Zufall* (München, 1975).

(7) P. S. Laplace: *Essai philosophique sur les probabilités* (Paris, 1814), 2nd ed., pp.3-4.

(8) N. Bohr: "Das Quantenpostulat und die neuere Entwicklung der Atomistik," *Naturwiss. 16*, 245-257 (1928).

(9) W. Heisenberg: "Uber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," *Zeits. für Physik, 43*, 172-198 (1927).

(10) J. Burckhardt: *Die Kultur der Renaissance in Italien* (Leipzig, 1860); authorized English translation from the 15th edition, S. G. C. Middlemore: *The Civilization of the Renaissance in Italy* (New York, 1929).

(11) P. Gay: *Style in History* (New York, 1974).

(12) H. Everett III: "The theory of the universal wave function," doctoral dissertation, Princeton University (1957); published in abbreviated form in ref. 13; published in full for the first time in ref. 14.

(13) H. Everett, III: "'Relative state' formulation of quantum mechanics," *Rev. Mod. Phys. 29*, 454-462 (1957); reprinted in ref. 14.

(14) B. S. DeWitt and N. Graham, eds.: *The Many-Worlds Interpretation of Quantum Mechanics* (Princeton, New Jersey, 1973).

(15) J. von Neumann: *Mathematischen Grundlagen der Quantenmechanik* (Berlin, 1932); translated as *Mathematical Foundations of Quantum Mechanics* (Princeton, N. J., 1955).

(16) J. A. Wheeler: "Assessment of Everett's 'relative state' formulation of quantum mechanics," *Rev. Mod. Phys. 29*, 463-465 (1974); reprinted in ref. 14.

(17) E. P. Wigner: "Epistemological perspective on quantum theory," in *Contemporary Research in the Foundations and Philosophy of Quantum Theory*, C. A. Hooker, ed. (Dordrecht, 1973), pp. 369-385.

(18) E. P. Wigner: "Remarks on the mind-body question," in *The Scientist Speculates*, I. J. Good, ed. (London, 1962), pp. 284-302.

(19) E. P. Wigner: "Are we machines?," *Proc. American Philos. Soc. 113*, 95-101 (1969).

(20) E. P. Wigner: "Physics and the explanation of life," *Foundations of Phys. 1*, 35-45 (1970).

(21) C. F. von Weizsäcker: "Classical and quantum descriptions," pp. 635-667 in *The Physicists' Conception of Nature*, J. Mehra, ed. (Dordrecht, Holland, 1973), pp. 635-667.

(22) J. A. Wheeler: "Include the observer in the wave function?," in *Quantum Mechanics, a Half Century Later*, J. Leite Lopes and M. Paty, eds. (Dordrecht, Holland, 1977).

(23) N. Bohr: *Atomic Theory and the Description of Nature* (Cambridge, 1934).

(24) N. Bohr: *Atomic Physics and Human Knowledge* (New York, 1958).

(25) N. Bohr: *Essays 1958-1962 on Atomic Physics and Human Knowledge* (New York, 1963).

(26) A. Einstein: "Autobiographical Notes," in *Albert Einstein: Philosopher-Scientist*, P. A. Schilpp, ed. (Evanston, Illinois,1949), quotation from p. 81.

(27) P. A. Schilpp, ed.: *Albert Einstein: Philosopher-Scientist* (Evanston, Illinois, 1949), and subsequent paperback editions elsewhere.

(28). N. Bohr: "Can quantum-mechanical description of physical reality be considered complete?," *Phys. Rev. 48*, 696-702 (1935); the quotation comes from p. 697.

(29) T. Mann: *Freud, Goethe, Wagner* (New York, 1937), p. 20; translated by H. T. Lowe-Porter from *Freud und die Zukunft* (Vienna, 1936); the cited words were included in the lecture given at the 80th birthday celebration for Sigmund Freud, 8 May 1936.

(30) C. S. Peirce: *The Philosophy of Peirce: Selected Writings*, ed. by J. Buchler (London, 1940); paperback reprint, *Philosophical Writings of Peirce* (New York, 1955); "fallibilist," p. 358.

(31) B. K. Harrison, M. Wakano, and J. A. Wheeler: "Matter-energy at high density; end point of thermonuclear evoluti in *La Structure et l'évolution de l'univers, Onzième conseil de physique Solvay*, (Bruxelles, 1958) pp. 124-146 terminology "crushing points," pp. 134-136.

(32) J. R. Gott III, J. E. Gunn, D. N. Schramm and B. M. Tinsley: "Will the universe expand forever," *Sci. American 234*, 62-79 (March, 1976); terminology "big crunch," p. 69.

(33) J. A. Wheeler: "Our universe: the known and the unknown," address before the American Association for the Advancement of Science," New York, Dec. 29, 1967; *American Scholar 37*, 248-274 (1968); terminology "black hole," pp. 258-262.

(34) J. A. Wheeler: "Superspace and the nature of quantum geometrodynamics," pp. 242-307 in *Battelle Rencontres: 1967 Lectures in Mathematics and Physics*, C. DeWitt and J. A. Wheeler, eds. (New York, 1968); "no before, no after," p. 253.

(35) W. Shakespeare: "The Tempest," London, about 1610; Prospero in Act IV, Scene I, lines 148-158.

(36) J. A. Wheeler: ref. 3; no magic particle, no magic field, p. 235.

(37) C. M. Patton and J. A. Wheeler: "Is physics legislated by cosmogony?," in *Quantum Gravity*, C. J. Isham, R. Penrose and D. W. Sciama, (Oxford, 1975); no magic mathematics, pp. 589-591.

(38) A. I. Oparin, ed.: *Evolutionary Biochemistry, Proceedings of the Vth International Congress on Biochemistry* (Moscow, 1961, London, 1963), pp. 12-51.

(39) G. W. Leibniz: source of quotation not traced.

(40) J. A. Wheeler: "Genesis and Observership," pp. 3-33 in Butts and Hintikka, eds., ref. 41; see p. 29.

(41) R. E. Butts and K. J. Hintikka: *Foundational Problems in the Special Sciences*, Reidel, (Dordrecht, 1977).

(42) N. Bohr: ref. 23; Bohr's stick, p. 99.

(43) F. J. Belinfante: *Measurements and Time Reversal in Objective Quantum Theory* (Oxford, 1975); terminology "indelible", p. 39.

(44) E. P. Wigner: "Are we machines?", *Proc. American Phil. Soc. 113*, 95-101 (1969); quotation from p. 97.

(45) E. P. Wigner: "The philosophical problem," pp. 1-3 in ref. 46; quotation from p. 3.

(46) B. d'Espagnat, ed.: *Foundations of Quantum Mechanics*, (New York, 1971).

(47) N. Bohr: ref. 24; irreversible amplification, p. 88.

(48) N. Bohr: ref. 24; closed by irreversible amplification, p. 73.

(49) N. Bohr: ref. 25; plain language, p. 3.

(50) N. Bohr: ref. 25; unambiguously communicable, pp. 5, 6.

(51) J. Kalckar: Letter to J. A. Wheeler dated June 10, 1977. Appreciation is expressed here to Dr. Kalckar both for the letter and for subsequent permission to quote from it. He adds that he cannot guarantee to have correct wording, and that Bohr's reply, as so often, was a joke with a definite "point"--hence to be taken seriously, but not <u>quite</u> seriously."

(52) K. R. Popper and J. C. Eccles: *The Self and Its Brain*, (Berlin, New York and London, 1977); see especially pp. 2 208; 438-440, and 515.

3) J. von Neumann: *The Computer and the Brain*, (New Haven, Connecticut, 1958); see especially pp. 60-61.

4) L. N. Cooper: "A possible organization of animal memory and learning," pp. 252-264 in *Nobel Symposium on the Collective Properties of Physical Systems*, B. Lundqvist and S. Lundqvist, eds., (New York, 1973).

5) M. M. Nass and L. N. Cooper: "A theory for the development of feature-detecting cells in visual cortex," *Biological Cybernetics 19*, 1-18 (1975).

6) L. N. Cooper: "A theory for the acquisition of animal memory," *Lepton and Hadron Structure (1974 Int. School of Subnuclear Physics, Erice, Trapani, Sicily: 14-31 July 1974)*, A. Zichichi (ed), (New York, 1975), 808-839.

57) G. E. Pugh: *On the Origin of Human Values*, (New York, 1976), chapter "Human values, free will, and the conscious mind; preprinted in *Zygon, 11*, 2-24 (1976).

58) International Union of Pure and Applied Physics: *Report of the Commission on Symbols, Units and Nomenclature*, Amsterdam, 1948.

59) J. A. Wheeler, "The 'past' and the 'delayed-choice' double slit experiment," in Marlow, ref. 60.

60) A. R. Marlow, ed.: *Mathematical Foundations of Quantum Theory*, (New York, 1978).

61) E. H. Gombrich: *Art and Illusion: A Study in the Psychology of Pictorial Representation*, (Princeton, New Jersey, 1961) 2nd ed., revised, p. 394.

(62) E. H. Gombrich: ref. (61), p. 273.

(63) E. H. Gombrich: ref. (61), p. 329.

(64) A. Einstein, B. Podolsky and N. Rosen: "Can quantum-mechanical description of physical reality be considered complete?", *Phys. Rev. 47*, 777-780 (1935).

(65) R. M. F. Houtappel, H. Van Dam and E. P. Wigner: "The conceptual basis and use of the geometric invariance principles," *Rev. Mod. Phys. 37*, 595-632 (1965); see especially §§ 4.1-4.5 on pp. 610-616.

(66) C. N. Yang and R. L. Mills, "Conservation of isotopic spin and isotopic gauge invariance," *Phys. Rev. 96*, 191-195 (1954).

(67) S. Hojman, K. Kuchař and C. Teitelboim: "New approach to general relativity," *Nature Phys. Sci. 245*, 97-98 (1973).

(68) S. A. Hojman, K. Kuchař and C. Teitelboim: "Geometrodynamics regained," *Ann. of Physics 76*, 88-135 (1976).

(69) C. Teitelboim: "Surface deformations, spacetime structure and gauge invariance," in C. Aragone, ed., *Relativity, Fields, Strings and Gravity: Proceedings of the Second Latin American Symposium on Relativity and Gravitation SILARG II* held in Caracas December 1975, Universidad Simon (Bolivar, Caracas, 1976).

(70) J. E. Nelson and C. Teitelboim: "Hamiltonian formulation of the theory of interacting gravitational and electron fields," Princeton University preprint, Nov., 1977.

71) J. A. Wheeler: ref. (3); machinery hidden, pp. 236-240.

72) J. A. Wheeler: "Beyond the end of time," Marchon lecture, University of Newcastle Upon Tyne, May 18, 1971 and Nuffield lecture, Cambridge University, July 19, 1971; unpublished lectures, updated in chapter of same title in ref. 72; see especially pp. 1206-1208.

73) C. W. Misner, K. S. Thorne and J. A. Wheeler: *Gravitation*, (San Francisco, 1973).

74) D. Z. Freedman, P. van Nieuwenhuizen and S. Ferrara: "Progress toward a theory of supergravity," *Phys. Rev. D 13*, 3214-3218 (1976).

75) S. Deser and B. Zumino: "Consistent supergravity," *Phys. Lett. 62B*, 335-337 (1976).

76) C. Teitelboim: "Supergravity and square roots of constraints," *Phys. Rev. Lett. 38*, 1106-1110 (16 May 1977).

(77) R. Tabensky and C. Teitelboim: "The square root of general relativity," *Phys. Letters 63B*, 453-456 (29 Aug., 1977).

(78) R. P. Feynman: *The Principle of Least Action in Quantum Mechanics*, doctoral dissertation, (Princeton University, 1942); unpublished; available from University Microfilms, Inc., (Ann Arbor, Michigan, 48106).

(79) R. P. Feynman and A. R. Hibbs: *Quantum Mechanics and Path Integrals*, (New York, 1965).

(80) J. R. Klauder: "The action option and a Feynman quantization of spinor fields in terms of ordinary c-numbers," *Ann. of Physics 11*, 123-168 (1960).

(81) J. A. Wheeler: ref. (34); spacetime an approximation, pp. 252-253.

(82) C. W. Misner: "Feynman quantization of general relativity," *Rev. Mod. Phys. 29*, 497-509 (1957).

(83) L. D. Faddeev: "Symplectic structure and quantization of the Einstein gravitation theory," *Actes du Congres International des Mathématiciens, Nice 1-10 Septembre 1970*, (Paris III, 1971), pp. 35-40.

(84) G. W. Gibbons and S. W. Hawking: "Action integrals and partition functions in quantum gravity," *Phys. Rev. D 15*, 2752-2756 (1977).

(85) Ref. (72): pp. 733-742 (Friedmann); pp. 1274-1275 (Schwarzschild); p. 940 (Taub).

(86) Ref. (72): pp. 419-423; 713ff; and 1181.

(87) Ref. (72): pp. 511-516 and 520.

(88) Ref. (72): pp. 493-498 and Chap. 21.

(89) J. A. Wheeler: ref. (34), p. 253

(90) M. Planck: "Über irreversible Strahlungsvorgänge," *Sitzungsber. Deut. Akad. Wiss. Berlin, Kl. Math.-Phy. Tech.*, 440-480 (1899).

(91) J. A. Wheeler: "On the nature of quantum geometrodynamics," *Ann. of Phys. 2*, 604-614 (1957).

(92) Ref. (72): Sections 1 and 2 of Chap. 21.

(93) J. A. Wheeler and R. P. Feynman: "Interaction with the absorber as the mechanism of radiation," *Rev. Mod. Phys. 17*, 157-181 (1945).

(94) J. A. Wheeler and R. P. Feynman: "Classical electrodynamics in terms of direct interparticle action," *Rev. Mod. Phys. 21*, 425-433 (1949).

(95) W. Ritz and A. Einstein: Zum gegenwartigen Stand des Strahlungsproblems, *Physik. Zeits. 10*, 323-324 (1909).

(96) J. A. Wheeler: ref. (40), p. 11, terminology "gates of time."

(97) N. Bohr: "Discussion with Einstein on epistemological problems in atomic physics," p. 199-241, in P. A. Schilpp, ed.: *Albert Einstein: Philosopher-Scientist*, (Evanston, Illinois, 1949).

(98) M. A. Duguay, "The ultrafast optical Kerr shutter," pp. 161-193 in E. Wolf, ed.: *Progress in Optics*, vol. 14 (North Holland, Amsterdam, 1976), compares the new ultrafast *optical* Kerr shutter to the older and slower conventional and *electrically* operated Kerr cell.

(99) N. Bohr: ref. (97), p. 230.

(100) A. Petersen: *Quantum Mechanics and the Philosophical Tradition*, (Cambridge, Massachusetts, 1968), gives a preliminary account of the stages in the development of the concept of "phenomenon."

(101) N. Bohr: *Atomic Physics and Human Knowledge*, (New York, 1958), p. 64.

(102) E. P. Wigner: "Epistemological perspective on quantum theory," pp. 369-385 in Hooker, ref. (103).

(103) C. A. Hooker, ed.: *Contemporary Research in the Foundations and Philosophy of Quantum Theory*, (Dordrecht, Holland, 1973).

(104) A. Einstein, B. Podolsky and N. Rosen: "Can quantum-mechanical description of physical reality be considered complete?", *Phys. Rev. 47*, 777-780 (1935). See (64)

(105) N. Bohr: ref. (28).

(106) D. Bohm: *Quantum Theory*, (Englewood Cliffs, New Jersey, 1951), pp. 614-619.

(107) M. Born and E. Wolf: *Principles of Optics*, 4th ed., (Pergamon, New York, 1970).

(108) B. A. Kagali: "A note on the correlation of photon polarizations in positronium annihilation," accepted for publication by *Phys. Letters*.

(109) J. A. Wheeler: "Polyelectrons," *Ann. New York Acad. Sci. 48*, 219-238 (1946). An error in one of the numbers in this paper is corrected in ref. (110).

(110) M. H. I. Pryce and J. C. Ward: "Angular correlation effects with annihilation radiation," *Nature 160*, 435 (1947).

(111) H. S. Snyder, S. Pasternack and J. Hornbostel: "Angular correlation of scattered annihilation radiation," *Phys. Rev. 73*, 440-448 (1948).

(112) E. Bleuler and H. L. Bradt: "Correlation between the states of polarization of the two quanta of annihilation radiation," *Phys. Rev. 73*, 1398 (1948).

(113) R. C. Hanna: "Polarization of annihilation radiation," *Nature 162*, 332 (1948).

(114) N. A. Vlasov and B. S. Dzeljepov: "Poljarizatzija annigiljatzionnikh gamma-kvantov," *Dokl. Akad. Nauk. SSSR 69*, 777-779 (1949).

(115) C. S. Wu and I. Shakhov: "The angular correlation of scattered annihilation radiation," *Phys. Rev. 77*, 136 (1950).

(116) F. Hereford: "The angular correlation of photo-electrons ejected by annihilation quanta," *Phys. Rev. 81*, 482 (1951), see also pp. 627-628.

(117) G. Bertolini, M. Bettoni and E. Lazzarini: "Angular correlation of scattered annihilation radiation," *Nuovo Cim. 2*, 661-662 (1955).

(118) H. Langhoff: "Die Linearpolarisation der Vernichtungsstrahlung von Positronen," *Zeits. f. Physik 160*, 186-193 (1960).

(119) L. Kasday, J. Ullman and C. S. Wu: "The Einstein-Podolsky-Rosen argument: positron annihilation experiment," *Bull. Amer. Phys. Soc. 15*, 586 (1970).

(120) L. Kasday: "Experimental test of quantum predictions for widely separated photons," in B. d'Espagnat, ed., *Foundations of Quantum Mechanics, Proceedings of the International School of Physics "Enrico Fermi," Course 49*, (New York, 1971), pp. 195-210.

(121) G. Faraci, D. Gutkowski, S. Notarrigo and A. R. Pennisi: "An experimental test of the EPR paradox," *Lettere al Nuovo Cimento 9*, 607-611 (1974).

(122) A. R. Wilson, J. Lowe and D. K. Butt: "Measurement of the relative planes of polarization of annihilation quanta as a function of separation distance," *J. Phys. G: Nucl. Phys. 2*, 613-624 (1976).

(123) C. A. Kocher and E. D. Commins: "Polarization correlation of photons emitted in an atomic cascade," *Phys. Rev. Lett. 18*, 575-577 (1967).

(124) S. J. Freedman and J. F. Clauser: "Experimental test of local hidden-variable theories," *Phys. Rev. Lett. 28*, 938-941 (1972).

(125) R. A. Holt: "Atomic cascade experiments," doctoral dissertation, Harvard University, Cambridge, Massachusetts, unpublished; available from University Microfilms, Inc., Ann Arbor, MI 48106, (1973).

(126) J. F. Clauser: "Experimental investigation of a polarization correlation anomaly," *Phys. Rev. Lett. 36*, 1223-1226 (1976).

(127) S. J. Freedman and R. A. Holt: "Tests of local hidden-variable theories in atomic physics," *Comments on Atomic and Molecular Physics 5*, 55-62 (1975).

(128) E. S. Fry and R. C. Thompson: "Experimental test of local hidden-variable theories," *Phys. Rev. Lett. 37*, 465-468 (1976).

(129) G. W. Leibniz: "On nature in itself; or on the force residing in created things, and their actions," pp. 119-134 in G. M. Duncan: *The Philosophical Works of Leibniz*, 2nd (revised) edition,(New Haven, Connecticut, 1908); translated from the original "De ipsa natura sive de vi insita creaturarum" of 1698.

(130) P. A. M. Dirac, V. A. Fock and B. Podolsky: "On quantum electrodynamics," *Phys. Zeits. Sowjetunion 2*, 468-479 (1932).

(131) S. Tomonaga: "On a relativistically invariant formulation of the quantum theory of wave fields," *Prog. Theor. Phys. 1*, 27-42 (1946).

(132) L. D. Landau and E. Lifshitz: *The Classical Theory of Fields*, 3rd revised English edition (Reading, Massachusetts and London, 1971).

(133) K. Gödel: "An example of a new type of cosmological solution of Einstein's field equations of gravitation," *Rev. Mod. Phys. 21*, 447-450 (1949).

(134) D. Hilbert: "Die Grundlagen der Physik," *Konigl. Gesell. der Wiss. Göttingen, Nachr., Math.-Phys. Kl.* 395-407 (1915).

(135) J. A. Wheeler: ref. (34), p. 273.

(136) C. Teitelboim: "How commutators of constraints reflect spacetime structure," *Ann. of Phys. 79*, 542-557 (1973).

(137) C. Teitelboim: "The Hamiltonian structure of spacetime," doctoral thesis (Princeton, 1973); available from University Microfilms, Inc., Ann Arbor, MI 48106.

(138) K. Kuchař: "Canonical quantization of gravity," in W. Israel, ed., *Relativity, Astrophysics and Cosmology* (Dordrecht, Holland, 1973), pp. 238-288.

(139) K. Kuchař: "Geometrodynamics regained: a Lagrangian approach," *J. Math. Phys. 15*, 708-715 (1974).

(140) G. W. Leibniz: "Refutation of Spinoza," written about 1708 as "Animadversiones ad Joh. George Wachteri librum de recondita Hebraeorum philosophia;" English translation, pp. 264-273 in G. M. Duncan, *The Philosophical Works of Leibniz*, 2nd (revised) ed., (New Haven, 1908).

(141) A. Einstein: as quoted by Alyesa Forsee in *Albert Einstein, theoretical physicist*, Macmillan (New York, 1963), p. 81.

(142) B. Riemann, edited by H. Weber: *Gesammelte Mathematische Werke*, 2nd edition, paperback reprint, (New York, 1953). The *Habilitationsvorlesung* of 1854 was translated into English by W. K. Clifford, *Nature 8*, 14 (1873).

143) A. Einstein: *Essays in Science*, Philosophical Library (New York, 1934), p. 68. Translated from *Mein Weltbild* Querido Verlag, (Amsterdam, 1933).

144) E. Cartan: "Sur les équations de la gravitation de Einstein," *J. Math. Pures Appl. 1*, 141-203 (1922).

145) E. Cartan: "La théorie des groupes et les recherches récentes de géométrie différentielle," *Conference Proceedings International Congress of Mathematicians Toronto, 1924) L'Enseign. math., t. 24*, 1-18 (1925).

146) P. A. M. Dirac: "Fixation of coordinates in the Hamiltonian theory of gravitation," *Phys. Rev. 114*, 924-930 (1959).

147) P. A. M. Dirac: *Lectures on Quantum Mechanics*, Belfer Graduate School of Science Monograph Series Number Two (New York, 1964).

148) R. Arnowitt, S. Deser and C. W. Misner: "The dynamics of general relativity," in L. Witten, ed., *Gravitation: An Introduction to Current Research*, (New York, 1962), pp. 227-265.

149) C. W. Misner, K. S. Thorne and J. A. Wheeler: ref. (73), pp. 10-12.

150) J. A. Wheeler: "Superspace," in R. P. Gilbert and R. G. Newton, eds., *Analytic Methods in Mathematical Physics*, (New York, 1970), pp. 335-378.

151) A. E. Fischer: "The theory of superspace," M. Carmeli, S. I. Fickler and L. Witten, eds., *Relativity*, (New York, 1970), pp. 303-357.

(152) B. Riemann: *Gesammelte Mathematische Werke*, 2nd ed., ed. by H. Weber, paperback reprint (New York, 1953).

(153) L. Bers: "Universal Teichmüller Space," in R. P. Gilbert and R. Newton, eds., *Analytic Methods in Mathematical Physics*, (New York, 1970), pp. 65-83.

(154) E. M. Lifshitz and I. M. Khalatnikov: "Investigations in relativistic cosmology," *Advances in Physics 12*, 185-249 (1963); translated from the earlier Russian version by J. L. Beeby; results available in boiled down form in L. D. Landau and E. M. Lifshitz: ref. (132); see especially pp. 350-355.

(155) T. Regge: "Relativity without coordinates," *Nuovo Cimento 19*, 558-571 (1961).

(156) C. W. Misner, K. S. Thorne and J. A. Wheeler: ref. (73), Chapter 42.

(157) R. F. Baierlein, D. H. Sharp and J. A. Wheeler: "Three-dimensional geometry as carrier of information about time," *Phys. Rev. 126*, 1864-1865 (1962).

(158) H. Leutwyler: "Gravitational field: Equivalence of Feynman quantization and canonical quantization," *Phys. Rev. 134*, B1155-B1182 (1964).

(159) B. DeWitt: "Quantum theory of gravity. I," *Phys. Rev. 160*, 1113-1148 (1967).

(160) B. DeWitt: "Quantum theory of gravity II; the manifestly covariant theory," *Phys. Rev. 162*, 1195-1239 (1967).

(161) B. DeWitt: "Quantum theory of gravity. III; Applications of the covariant theory," *Phys. Rev. 162*, 1239-1256 (1967).

(162) E. P. Wigner and H. Saleckar: "Quantum limitations of the measurement of space-time distances," *Phys. Rev. 109*, 571-577 (1958).

(163) R. F. Marzke and J. A. Wheeler: "Gravitation as geometry. I. The geometry of spacetime and the geometrodynamical standard meter," in H. -Y. Chiu and W. F. Hoffman, ed., *Gravitation and Relativity* (New York, 1964), pp. 40-64.

(164) B. DeWitt: "Dynamical theory of groups and fields," in C. DeWitt and B. DeWitt, eds., *Relativity, Groups and Topology* (New York, 1964), pp. 587-822; see especially pp. 598-614.

(165) J. Ehlers, F. A. E. Pirani and A. Schild: "The geometry of free fall and light propagation," in L. O'Raifeartaigh, ed., *General Relativity, Papers in Honor of J. L. Synge*, (London, 1972), pp. 63-84.

(166) C. W. Misner, K. S. Thorne and J. A. Wheeler: ref. (73), p. 72.

(167) L. Landau and R. Peierls: "Erweiterung des Unbestimmtheitsprinzips für die relativistische Quantentheorie," *Zeits. f. Physik 69*, 56-69 (1931).

(168) N. Bohr and L. Rosenfeld: "Zur Frage der Messbarkeit der electromagnetischen Feldgrössen," *Kgl. Danske Videnskab. Selskab, Mat-fys. Medd. 12*, No. 8 (1933).

(169) N. Bohr and L. Rosenfeld: "Field and charge measurements in quantum electrodynamics," *Phys. Rev. 78*, 794-798 (1950).

(170) C. M. Patton and J. A. Wheeler: ref. (37), pp. 577 ff.

(171) A. Peres: "On the Cauchy problem in general relativity, II," *Nuovo Cimento 26*, 53-62 (1962).

(172) U. Gerlach: "Derivation of the ten Einstein equations from the semiclassical approximation to classical geometrodynamics," *Phys. Rev. 177*, 1929-1941 (1969).

(173) J. A. Wheeler: *Geometrodynamics* (New York, 1962), pp. 71-83.

(174) K. Kuchař: "Ground state functional of the linearized gravitational field," *J. Math. Phys. 11*, 3322-3334 (1970).

(175) E. Mach: *Beiträge zur Analyse d. Empindungen* (Jena, 1886; 5th ed., 1906, entitled *Die Analyse d. Empfindungen)*.

(176) J. Piaget: *The Construction of Reality in the Child*, translated from the French by M. Cook, Basic Books, New York, 1954; also the summary of the lesson of this and his books, *The Child's Conception of Movement and Speed*, and *The Child's Conception of Time*, in J. Piaget, *Le Structuralisme*, Presses Universitaires de France, Paris, 1968; translated by C. Maschler as *Structuralism*, Basic Books, New York, 1971.

(177) R.J.E. Clausius: *Abhandlungen über die mechanische Wärmetheorie*, Vieweg, Braunschweig 1864-1867.

(178) B. Russell, as quoted by G. Steiner: "Has truth a future?-The first annual Bronowski Memorial Lecture" *BBC Listener 99, No. 2542*, 42-46 (Jan. 12, 1978).

(179) W.J. Cocke: "Statistical time symmetry and two-time boundary conditions in physics and cosmology", *Phys. Rev. 160*, 1165-1170 (1967).

(180) V.I Arnold and A. Avez: *Ergodic Problems of Classical Mechanics*, Benjamin, New York, 1968.

(181) T. Gold, "The arrow of time", pp 81-95 in *Onzième Conseil de Physique Solvay: La structure et l'évolution de l'univers*, Editions Stoops, Brussels, 1959.

(182) T. Gold: "The arrow of time", *Amer. J. Phys. 30*, 403-410 (1962).

(183) T. Gold: "The arrow of time", pp. 225-234 in *Recent Developments in General Relativity*, Pergamon-Macmillan, New York and Warsaw, 1962.

(184) H. Bondi: "Physics and Cosmology" *The Observatory 82*, 133-143 (1962).

(185) P.G. Bergmann, H. Bondi, T. Gold, A. Grünbaum, F. Hoyle, D. Layzer, P. Morrison, I. Robinson, L. Rosenfeld, L. Schiff, J.A. Wheeler, and "X": "The instability of the future", pp. 138-148, and "General discussion", pp. 223-243, in T. Gold, ed., Ref. (186).

(186) T. Gold, ed: *The Nature of Time*, Cornell University Press, Ithaca, New York, 1967.

(187) H. Schmidt: "Model of an oscillating cosmos which rejuvenates during contraction", *J. Math. Phys. 7*, 495-509 (1966).

(188) J.A. Wheeler: "At last--a sane look at the 'arrow of time'", a review of P.C.W. Davies: *The Physics of Time Asymmetry*, *Physics Today 28*, 49-50 (June 1975).

(189) G.K. Chesterton: quoted in A.L. Mackay, collector: *The Harvest of the Quiet Eye: A Selection of Scientific Quotations*, Institute of Physics, Bristol, 1977, as adapted for use here.

(190) S.G. Brush, ed.: *Kinetic Theory. Volume 2: Irreversible Processes*, Pergamon, Oxford, 1966.

(191) J.W. Gibbs: "On the equilibrium of heterogeneous substances", *Trans. Connecticut Acad. 3*, 108-248 (1875).

(192) L. Boltzmann: "Über die Beziehung eines allgemeine mechanischen Satzes zum zweiten Hauptsatze der Warmetheorie", *Sitzungsberichte Akad. Wiss.*, part II, *75*, 67-73 (1877), translated into English as "On the relation of a general mechanical theorem to the second law of thermodynamics", pp. 188-193 in S.G. Brush (190); quotes from pp. 190, 192 and 193.

(193) H. Poincaré: "Le méchanisme et l'expérience", *Revue de Metaphysique et de Morale 1*, 534-537 (1893), translated into English as *Mechanism and Experience*, pp. 203-207 in S.G. Brush (190); quote from p. 206.

(194) E. Zermelo: "Über einen Satz der Dynamik und die mechanische Warmetheorie", *Ann. der Physik 57*, 485-494 (1896), translated into English as "On a theorem of dynamics and the mechanical theory of heat", pp. 208-217 in S.G. Brush (190); quotes from pp. 208, 209 and 214, 215.

(195) H. Poincaré: "Sur le problème des trois corps et les équations de dynamique", *Acta Math. 13*, 1-270 (1890).

(196) L. Boltzmann: "Entgegnung auf die wärmetheoretischen Betrachtungen der Hrn. E. Zermelo", *Ann. der Physik 57*, 773-784 (1896); translation, "Reply to Zermelo's remarks on the theory of heat", pp. 218-228 in S.G. Brush (190); quoted from pp. 219-221 and 223-226.

(197) E. Zermelo: "Ueber mechanische Erklärungen irreversibler Vorgänge", *Ann. der Physik 59*, 793-801 (1896); translation "On the mechanical explanation of irreversible processes", pp. 229-237 in S.G. Brush (190); quote from p. 235.

(198) L. Boltzmann: "Zu Hrn. Zermelo's Abhandlung Über die mechanische Erklärung irreversibler Vorgänge", *Ann. der Physik 60*, 392-398 (1897); translation, "On Zermelo's paper 'On the mechanical explanation of irreversible processes'", pp. 238-245 in S.G. Brush (190); quotes from pp. 238, 239, 241, 242.

(199) T. Ehrenfest-Afanassjewa: "Note", appended to the 1959 reprinting of P. and T. Ehrenfest: "Begriffliche Grundlagen der statistischen Auffassung in der Mechanik", pp. 1-90 in *Encyklopädie der mathematischen Wissenschaften IV 2, II*, Heft 6, Teubner, Leipzig, 1912, and printed on

pp. 301-302 of M.J. Klein, ed.(200); quote from p. 302.

(200) M.J. Klein, ed.: *Paul Ehrenfest: Collected Scientific Papers*, North Holland, Amsterdam, 1959.

(201) G.E. Uhlenbeck: "An outline of statistical mechanics", pp. 1-29 in E.G.D. Cohen, ed.(202); quotes from p. 4 and p. 19.

(202) E.G.D. Cohen, ed.: *Fundamental Problems in Statistical Mechanics II*, North-Holland, Amsterdam, 1968.

(203) N.N. Boguliubov: "Problem dynamischeskoi teorii v statisticheskoi fizike," Mowcow, 1946; translation by E.K. Gora: "Problems of a dynamical theory in statistical physics", pp. 1-118 in de Boer and Uhlenbeck(204).

(204) J. de Boer and G.E. Uhlenbeck: *Studies in Statistical Mechanics, Vol. I.*, North-Holland, Amsterdam, 1962.

(205) E.G.D. Cohen: "Kinetic approach to non-equilibrium phenomena", p. 548-560 in J. Mehra, ed.(4); the quote comes from p. 560.

(206) I. Prigogine: "Time, irreversibility and structure", pp. 561-591 in J. Mehra, ed.(4).

(207) M.J. Klein: *Paul Ehrenfest. Vol. I. The Making of a Theoretical Physicist*, North-Holland, Amsterdam, 1970.

(208) H. Reichenbach: *The Direction of Time*, University of California Press, Berkeley, 1956.

(209) P.C.W. Davies: *The Physics of Time Asymmetry*, The University of California Press, Berkeley and Los Angeles, 1974.

(210) P. and T. Ehrenfest: "Über eine Aufgabe aus der Wahrscheinlichkeitsrechnung, die mit der kinetischen Deutung der Entropievermehrung zusammenhängt", *Math.-Naturw. Blätter 3*, Nos. 11 and 12, 1906; discussed and key passages translated in M.J. Klein (207), pp. 115-119.

(211) P. and T. Ehrenfest: "Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem", *Physik. Zeits. 8*, 311-314 (1907); reprinted in M.J. Klein, ed.(200)

(212) E. Schrödinger and K.W.F. Kohlrausch: "Die Ehrenfestsche Modell der H-Kurve", *Physikal.Zeits. 27*, 306-313 (1926).

(213) M. Kac: *Probability and Related Topics in Physical Science*, Interscience, New York, 1959.

(214) A.H. Taub: "Empty space times admitting a three parameter group of motions", *Ann. of Math. 53*, 472-490, 1951.

(215) C.W. Misner: "Taub-NUT space as a counterexample to almost anything", pp. 160-169 in J. Ehlers, ed.: *Relativity Theory and Astrophysics I: Relativity and Cosmology* (Lectures in Applied Mathematics, Vol. 8), American Mathematical Society, Providence, Rhode Island, 1967.

(216) C.W. Misner and A.H. Taub: "A singularity-free empty universe", *Soviet Phys. JETP 28*, 122-133 (1969). See (272).

(217) J.A. Wheeler: "The Beam and the Stay of the Taub Universe", preprint, University of Texas, Austin (July 1978).

(218) I. Prigogine: "The dynamical theory of irreversible processes", lectures of 25 July-6 August 1977 at International School of Physics "Enrico Fermi", Varenna, Italy.

(219) I. Prigogine, "Time, structure and fluctuations", Nobel Prize Lecture in Chemistry, December 1977, University of Texas at Austin preprint.

(220) K.L. Hainebach and D.N. Schramm, "Comments on galactic evolution and nucleocosmochronology", *Astrophys. J. 212*, 347-359 (1977) give references to the literature on the ^{187}Re decay constant.

(221) R.H. Dicke and P.J.E. Peebles, "Cosmology and the Radioactive Decay Ages of Terrestrial Rocks and Meteorites", *Phys. Rev. 128*: 2006-2011 (1962).

(222) C.L. Dodgson, *alias* Lewis Carroll: *Through the Looking-Glass*, London, 1871; on p. 170 in this as reprinted in *Alice's Adventures in Wonderland and Through the Looking-Glass*, Dent, London, 1954 (reprinted 1966).

(223) J.A. Wheeler: "General discussion", pp 234-235 in T. Gold, ed.(186)

(224) G.W. Leibniz: "A new system of nature...", *Journal des Savans*, June 27, 1695; reprinted in P.P. Wiener, ed., *Leibniz Selections*, Scribners, New York, 1951, p. 108.

(225) E.P. Hubble: "A relation between distance and radial velocity among extragalactic nebulae", *Proc. Nat. Acad. Sci. U.S. 15*, 169-173 (1929).

(226) R.A. Alpher, H.A. Bethe and G. Gamow: as cited on p. 760 of Misner, Thorne and Wheeler(73).

(227) R.H. Dicke: as cited on p. 760 of Misner, Thorne and Wheeler(73).

(228) A.A. Penzias and R.W. Wilson: as cited on p. 760 of

Misner, Thorne and Wheeler(73).

(229) R.H. Dicke, P.J.E. Peebles, P.G. Roll and D.T. Wilkinson: as cited on p. 760 of Misner, Thorne and Wheeler(73).

(230) M. Schwarzschild: *Structure and Evolution of the Stars*, Princeton University Press, Princeton, N.J., 1958.

(231) A.P. Lightman and S.L. Shapiro: "The dynamical evolution of globular clusters", *Rev. Mod. Phys. 50*, 437-481 (1978); see especially Table I on p. 442 for ages $\sim 10^{10}$a for nine representative clusters in the Milky Way.

(232) W. Fowler and F. Hoyle: *Nucleosyntheses in Massive Stars and Supernovae*, University of Chicago Press, Chicago, 1964.

(233) R.V. Wagoner: "Big-Bang Nucleosynthesis Revisted", *Astrophys. J. 179*, 343-360 (1973).

(234) P.S. Laplace: *Exposition du Système du monde, vol. 1*, Paris, 1796; the relevant page is reproduced photographically in Misner, Thorne and Wheeler(73), p. 623.

(235) B. Paczynski: "A model of accretion disks in close binaries", *Astrophys. J. 216*, 822-826 (1977).

(236) E. van den Heuvel: "Discussion of Binary Stars", in *Astrophysics and Gravitation: Proceedings of the Sixteenth Conference in Physics of the University of Brussels, September 1973*, Editions de l'Universite de Bruxelles, Bruxelles, Belgium, 1974.

(237) B. K. Harrison, K. S. Thorne, M. Wakano and J.A. Wheeler: *Gravitation Theory and Gravitational Collapse*, University of Chicago Press, Chicago, 1965.

(238) J. Wilson: "Rapidly rotating neutron stars", *Phys. Rev. Lett. 30*, 1082-1084 (1973).

(239) R. Giacconi: "Progress in X-ray astronomy", *Physics Today, Vol. 26, No. 5*, 38-47 (1973).

(240) J.A. Wheeler: "The Universe as Home for Man", in *The Nature of Scientific Discovery*, O. Gingerich, ed., Smithsonian Institution Press, Washington, D.C., 1975, pp. 276-278.

(241) J.H. Oort: "The galactic center", *Ann. Rev. Astron. and Astrophys. 15*, 295-362 (1977).

(242) P. Young, J.A. Westphal, J. Kristian, C.P. Wilson and F.P. Landauer: "Evidence for a supermassive object in the nucleus of the galaxy M 87 from SIT and CCD area photometry", *Ap. J. 221*, 721-730 (1978).

(243) W.L.W. Sargent, P.J. Young, A. Boksenberg, K. Shortridge, C.R. Lynds and F.D.A. Hartwick: "Dynamical evidence for a central mass concentration in the galaxy M 87", *Ap. J. 221*, 731-744 (1978).

(244) C.W. Misner, K.S. Thorne and J.A. Wheeler, ref. (73), pp. 797-798.

(245) J.H. Oort: "Distribution of galaxies and the density of the universe", in *Onzième Conseil de Physique Solvay: La Structure et l'Évolution de l'univers*, Editions Stoops, Brussels, Belgium, 1958.

(246) J.F. Gott, J.E. Gunn, D.N. Schramm, and B.M. Tinsley: "An Unbound Universe?" *Astrophys. J. 194*, 543-553 (1974).

(247) J.E. Gunn and J.B. Oke: "Spectrophometry of faint cluster galaxies and the Hubble diagram: An Approach to Cosmology", *Astrophys. J. 195*, 255-268 (1975).

(248) J.P. Ostriker and P.J.E. Peebles: "A numerical study of the stability of flattened galaxies; or, can cold galaxies survive?", *Astrophys. J. 186*, 467-480 (1973).

(249) J.P. Ostriker, P.J.E. Peebles and A. Yahil: "The size and mass of galaxies and the mass of the universe", *Astrophys. J. Lett. 193*, L1-4 (1974).

(250) J.A. Wheeler: "Conference summary: more results than ever in gravitation physics and relativity", pp. 299-344 in *General Relativity and Gravitation*, Proceedings of the Seventh International Conference (GR7), Tel-Aviv, June 1974, G. Shaviv and J. Rosen, ed., John Wiley and Sons, New York, 1975.

(251) M. Davis, E.J. Groth and P.J.E. Peebles: "Study of galaxy correlations: Evidence for the gravitational instability picture in a dense universe", *Astrophys. J. 212*, L107-111 (1977). Also see earlier references cited therein.

(252) C.W. Misner, K.S. Thorne, and J.A. Wheeler, ref. (73), pp. 795-796.

(253) R.D. Blandford and M.J. Rees: "Extragalactic double radio sources-the current observational and theoretical position", *Contem. Phys. 16*, 1-16 (1975).

(254) R.D. Blandford and M.J. Rees: "A 'twin-exhaust' model

for double radio sources", *Mon. Not. R. Astr. Soc. 169*, 395-415 (1974).

(255) Ya. B. Zel'dovich: "Observations in a universe homogeneous in the mean", *Astronom. Zh. 41*, 19-24 (1964); English translation in *Sov. Astronom.--AJ8*, 13-16 (1964).

(256) V.M. Dashevsky and Ya. B. Zel'dovich: "The propagation of light in a nonhomogeneous nonflat universe II", *Astronom. Zh. 41*, 1071-1074 (1964); English translation in *Sov. Astr.--AJ8*, 854-856 (1965).

(257) W.H. Press and J.E. Gunn: "Method for etecting a cosmological density of condensed objects", *Astrophys. J. 185*, 397-412 (1973).

(258) R.C. Roeder: Report presented at the Seventh International Conference (GR7) at Tel-Aviv University, June 23-28, 1974.

(259) A. Einstein: *The Meaning of Relativity*, 3rd ed., Princeton University Press, Princeton, N.J., 1950, pp. 107-108.

(260) A. Einstein: *Essays in Science*, Philosophical Library, New York, 1934. Translated from *Mein Weltbild*, Querido Verlag, Amsterdam, 1933. The quote appears on p. 52 of the 1934 translation.

(261) H. Alfvén and O. Klein: "Matter-antimatter annihilation and cosmology", *Ark. Fys. 23*, 187-194 (1962).

(262) G. De Vaucouleurs: "The large-scale distribution of galaxies and clusters of galaxies", *Astron. Soc. Pacific Publ. 83*, 113-143 (1971).

(263) A. Einstein: p. 68 of Ref. (260).

(264) D.R. Brill and J.A. Isenberg: "Dynamics of an open universe", *Bull. American Phys. Soc. 21*, 766 (1976).

(265) R.C. Tolman: *Relativity, Thermodynamics, and Cosmology*, Clarendon Press, Oxford, 1934.

(266) A. Avez: "Propriétés globales des espace-temps périodiques clos", *Acad. des Sci. Paris, Comptes Rend. 250*, 3583-3587 (1960).

(267) R.P. Geroch: *Singularities in the Spacetime of General Relativity: Their Definition, Existence, and Local Characterization*, doctoral dissertation, Princeton University, 1967.

(268) S.W. Hawking and R. Penrose: "The singularities of gravitational collapse and cosmology", *Proc. R. Soc. London A314*, 529-548, 1969.

(269) S.W. Hawking and G.F.R. Ellis: *The Large Scale Structure of Space-Time*, Cambridge University Press, Cambridge, U.K., 1973.

(270) A.H. Taub: "Empty space-times admitting a 3-parameter group of motions", *Ann. of Math. 53*, 472-490 (1951).

(271) J.A. Wheeler: "Geometrodynamics and the issue of the final state", pp. 315-520 in C. DeWitt and B.S. DeWitt, eds., *Relativity, Groups and Topology*, Gordon and Breach, New York, 1964, p. 501 and p. 519.

(272) C.W. Misner and A.H. Taub: "A singularity-free empty universe", *Zh. Eksp. and Teor. Fiz. 55*, 233-255 (1968); English original in *Sov. Phys.-JETP 28*, 122-133 (1969).

(273) G.F.R. Ellis and A.R. King: "Was the big bang a whimper?", *Commun. Math. Phys. 38*, 119-156 (1974).

(274) R. Penrose: "Internal instability in a Reissner-Nordstrøm black hole", *Internat. J. Theor. Phys. 7*, 183-197 (1973).

(275) E.M. Lifshitz and I.M. Khalatnikov: "Problems of relativistic cosmology", *Usp. Fiz. Nauk 80*, 391-438 (1963); English translation in *Sov. Phys.--Uspekhi 6*, 495-522 (1964).

(276) E.M. Lifshitz and I.M. Khalatnikov: "Oscillatory approach to singular point in the open cosmological model", *Zh. Eksp. & Teor. Fiz. Pis'ma 11*, 200-203 (1970); English translation in *Sov. Phys.--JETP Lett. 11*, 123-125 (1971).

(277) V.A. Belinsky and I.M. Khalatnikov: "On the nature of the singularities in the general solution of gravitational equations", *Zh. Eksp. & Teor. Fiz. 56*, 1700-1712 (1969); English translation in *Sov. Phys.--JETP 29*, 911-917.

(278) V.A. Belinsky and I.M. Khalatnikov: "General solution of the gravitational equations with a physical singularity", *Zh. Iksp. & Fiz. 57*, 2163-2175 (1969); English translation in *Sov. Phys.--JETP 30*, 1174-1180 (1970).

(279) V.A. Belinsky and I.M. Khalatnikov: "General solution of the gravitational equations with a physical oscillatory singularity", *Zh. Eksp. & Teor. Fiz. 59*, 314-321 (1970); English translation in *Sov. Phys.--JETP 32*, 169-172.

(280) I.M. Khalatnikov and E.M. Lifshitz: "General cosmological solutions of the gravitational equations with a singularity in time", *Phys. Lett. 24*, 76-79 (1970).

(281) V.A. Belinsky, I.M. Khalatnikov and E.M. Lifshitz: "Oscillatory approach to a singular point in the relativistic cosmology", *Usp. Nauk 102*, 463-500 (1970); English translation in *Advances in Physics 19*, 525-573.

(282) V.A. Belinsky, E.M. Lifshitz and I.M. Khalatnikov: "Oscillatory mode of approach to a singularity in homogeneous cosmological models with rotating axes", *Zh. Eksp. & Teor. Fiz. 60*, 1969-1979 (1971); English translation in *Sov. Phys.--JETP 33*, 1061-1066 (1971).

(283) C.W. Misner, K.S. Thorne and J.A. Wheeler, Ref. (73), pp. 805-814.

(284) D. Eardley, E. Liang and R. Sachs: "Velocity-dominated singularities in irrotational dust cosmologies", *J. Math. Phys. 13*, 99-107 (1972).

(285) V.A. Belinsky, I.M. Khalatnikov and B. Collins: 24 June 1976 draft preprint entitled "Dealing with the instability of special solutions of the field equations". I thank Dr. Belinsky for the opportunity to see this unpublished paper.

(286) J.A. Wheeler: "Singularity and unanimity", *Gen. Rel. & Grav., vol. 8, no. 8*, pp. 713-715 (1977).

(287) R. Penrose: "Singularities in cosmology", in M.S. Longair, ed., *Confrontation of Cosmological Theories with Observational Data*, I.A.U., Reidel, Dordrecht, 1974.

(288) J.A. Wheeler: "The black hole", in *Astrophysics and Gravitation: Proceedings of the 16th Conference on Physics at the University of Brussels, September 1973*, Editions de l'Université de Bruxelles, 1040 Bruxelles, Belgium, 1974.

(289) C.W. Misner, K.S. Thorne and J.A. Wheeler, Ref. (73), pp. 539-543.

ALPHABETICAL LISTING OF REFERENCES