Revista Mexicana de Física 23 (1974) 1-57

FROM RELATIVITY TO MUTABILITY*

John Archibald Wheeler

Department of Physics, Princeton University

> "... the end of all our exploring
> Will be to arrive where we started
> And know the place for the first time."
>
> T. S. Eliot[1]

1. INTRODUCTION

From Relativity to Gravitational Collapse; and from the Consequences of Collapse to the Principle that Nature Conserves Nothing

Relativity and the quantum principle constitute the two overarching concepts of 20th century physics. To review relativity here is to have opportunity for something new. Casting an eye over what we have learned in this domain, can we discover out of it all some consideration that might guide us into tomorrow? The lesson need not be positive. It could be negative. In

---

* Reprinted, with permission, from J. A. Wheeler, "From Relativity to Mutability" in J. Mehra, ed., *The Physicist's Conception of Nature*, Dirac 70th anniversary volume, published by Reidel Publishing Co., Dordrecht and Boston.

physics there are many negative principles. None is better known, and none has ever proved itself more powerful, than the original principle of relativity itself, saying that it is impossible to discover any difference in the laws of physics between two inertial reference frames that would distinguish the one frame from the other. If both special and general relativity were founded on negative principles, it is also true that some of the most remarkable consequences of relativity have a negative flavour, as for example these: (1) The universe cannot be static. (2) The volume of a closed model universe is not constant. (3) Total energy and total angular momentum cannot be defined for a closed universe. They are meaningless concepts. For such a system, no global law of conservation of energy and angular momentum has meaning or relevance. (4) Baryon number and lepton number are well defined quantities for a normal star; but when this star collapses to a black hole, the well established laws of conservation of particle number lose all applicability.

These and other interesting negative conclusions out of relativity have been recognized for some time. They are still startling enough to call for some review. However, after such a review, it is even more important for us to try to pull all these individual negatives together into a larger formulation of the way that nature acts. I have not been able to find any more reasonable way to state the situation than this: nature conserves nothing; there is no constant of physics that is not transcended; or, in one word, mutability is a law of nature.

"...it would have been difficult to establish any laws of nature," Wigner reminds us,[2] "if these were not invariant with respect to displacements in space and time." However, displacements in flat spacetime, or even in a curved spacetime that is asymptotically flat, make no sense in the closed universe of Einstein's general relativity. In that universe there is no global law of conservation of momentum and energy. More startling, such a universe undergoes gravitational collapse. In that collapse, classical space and time themselves come to an end. With their end, the framework falls down for everything that one has ever called a law of physics.

Nothing that relativity has ever predicted is more revolutionary than collapse, and nothing that collapse puts in question is more central than the very possibility of any enduring laws of physics.

The golden trail of science is surely not to end in nothingness. There may be no such thing as "the glittering central mechanism of the universe" to be seen behind a glass wall at the end of the trail. Not machinery but magic may be the better description of the treasure that is waiting. Rather than Newtonian law it may resemble more the logic of relationships that Leibniz envisaged. But that is an issue for tomorrow. Today we look at the breakage that relativity has made among the laws of physics.

Not the slightest question is implied here about the everyday laws of physics undereveryday conditions. No one will turn to relativity to learn something new about the physics of liquids or solids, unless he is concerned with a neutron star, or with conditions still more extreme.

Looking in on General Relativity through Six Windows

What relativity takes away, and what it gives, show nowhere better than on a tour about the structure that Einstein built, looking into in first through one window, then another. Framing each window is a different derivation of Einstein's law for the dynamic change of geometry with time. We scrutinize physics in turn through these six windows: (1) Einstein's original derivation, based on the principles of equivalence and correspondence; (2) Élie Cartan's derivation, resting on the fact that the boundary of a boundary is zero; (3) the most compact derivation one knows, based on the idea that density of mass-energy governs curvature; (4) the derivation of Hilbert and Palatini, founded upon the principle of least action; (5) the derivation of Hojman, Kuchař and Teitelboim, that introduces the group-theoretic concept of "'group' of deformations of a spacelike hypersurface in spacetime"; and (6) the schematic derivation of Andrei Sakharov, founded upon the concept of "the metric elasticity of space". Another derivation starts from the theory of a spin-2 field in flat space, but it and other interesting derivations will not be touched upon here.

Several comments to be made here about these derivations come from the book of Charles W. Misner, Kip S. Thorne and myself, *Gravitation*,[3] now in the course of publication. Warm appreciation is expressed to these and other colleagues and not least to Paul Dirac himself, for insights into the structure and consequences of Einstein's standard 1915 geometrodynamics.

"Tide-Producing Acceleration" or "Riemann Curvature" as Local Measure of the Effect of Geometry on Motion

There is no derivation of the effect of a moving mass upon geometry that is not best prefaced by the effect of geometry upon the movement of a mass. That there appears at first to be no such effect comes as a shock to the beginning student of relativity. He expects to see the analogue of electromagnetism faithfully pursued, where the invariantly measured acceleration of a test charge is a direct measure of the electromagnetic field strength,

$$\frac{D^2 x^\mu}{D\tau^2} = \frac{e}{m} F^\mu_{\ \nu} \frac{dx^\nu}{d\tau} \quad . \tag{1}$$

Instead, the neutral test particle of general relativity moves in a straight line with uniform velocity in the local Lorentz frame, a statement that expresses itself in an arbitrary curvilinear coordinate system in the form

$$\frac{D^2 x^{\mu}}{D\tau^2} = 0 \tag{2}$$

Gravitation seems to have disappeared. Has not Einstein gone too far, the beginning student may ask, in emphasizing that the only right description of a force is a local description? However, gravitation, at first apparently extinguished in this local description, springs into evidence again as a tide-producing force; that is, as a measure of the relative acceleration of two nearby test particles endowed with an initial separation η^{α}; thus,

$$\frac{D^2 \eta^{\alpha}}{D\tau^2} + R^{\alpha}_{\beta\gamma\delta} \frac{dx^{\beta}}{d\tau} \eta^{\gamma} \frac{dx^{\delta}}{d\tau} = 0 \ . \tag{3}$$

The tide producing force or Riemann curvature $R^{\alpha}_{\beta\gamma\delta}$, as seen in its effect on a fleet of nearby test particles, is the central descriptor in Einstein's geometrical account of gravitation. At issue from this point onward is not the effect of curvature on mass but the back action of mass on curvature.

2. EINSTEIN'S ROUTE TO GENERAL RELATIVITY: ANTECEDENTS AND CONSEQUENCES

Riemann's "Physical Geometry", Mach's Concept of Inertia, and Einstein's Equivalence Principle as Elements in Einstein's General Relativity

Einstein credits Riemann (Fig. 1) with one central idea of general relativity. Matter gets its moving order from geometry. In other words, geometry acts on matter. By the principle of action and reaction matter must therefore act on geometry. Thereupon geometry ceases to be a God-given Euclidean participant standing high above the battles of matter and energy. Geometry steps forward as a new participant in the world of physics.

A second idea that led him to relativity Einstein attributes to Mach. Acceleration can have no meaning unless there are objects with respect to which the acceleration takes place. Einstein could see consequences from this Mach principle. Thus inertia here must take its origin in mass-energy there. But gravitation here also arises from mass-energy there. Therefore

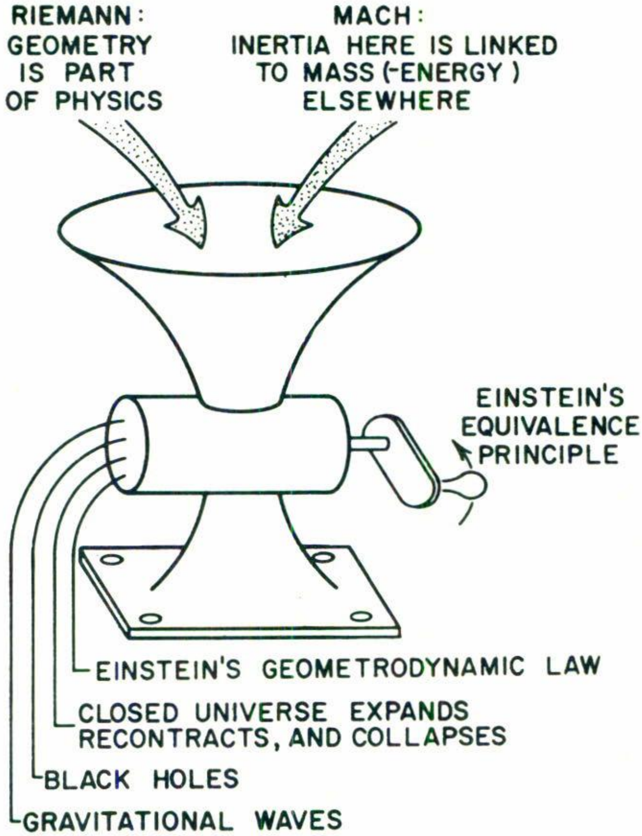


Fig. 1. Einstein's great achievement, to use his new 1907 principle of the local equivalence of "gravitational" and "propulsive" accelerations to bring together two currents of thought, going back to Riemann and Mach, and formulate (1907–1915) "general relativity" or "geometrodynamics", with all its consequences. In an unpublished essay of 1919[4] Einstein describes the equivalence principle (that came only two years after special relativity) as "the happiest thought of my life": "Thus, for an observer in free fall from the roof of a house there exists, during his fall, no gravitational field." To Mach Einstein wrote enthusiastically from Zurich on 25 June 1913, more than two years before he had arrived at the final formulation of general relativity[5], "If so [i. e., if the eclipse observations confirm the new theory], then your helpful investigations on the foundations of mechanics – Planck's unjustified criticisms notwithstanding – will receive a brilliant confirmation. For it necessarily turns out that *inertia* has its origin in a kind of interaction, entirely in accord with your considerations on the Newton pail experiment." Einstein also gives warm testimony to the contribution of Riemann[6], "... 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 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."

tentatively conclude that gravitation and inertia are transmitted by the same machinery.

That the machinery required to carry both gravitation and inertia is geometry was Einstein's great synthesis of the two currents of thought going back to Mach and Riemann. No consideration impelled him more directly to this synthesis than his 1909 principle of the local equivalence of gravitational and propulsive accelerations. Gravitation stops producing a curvilinear track in a flat spacetime. Motion becomes straight in every local Lorentz frame. Gravitation becomes the curvature encountered in passing from one local Lorentz tangent space to the next.

The principle of correspondence with the Newtonian theory of gravitation requires Einstein's conserved tensorial measure of curvature, $G_{\mu\nu}$, to agree (Einstein's papers[7] before the Berlin Academy, on 4, 18 & 25 November 1915) with 8π times the conserved measure of the density of mass-energy*, that is, the standard tensor of stress and density of momentum and energy, T; thus,

$$G_{\mu\nu} = 8\pi T_{\mu\nu} \quad . \qquad (4)$$

Schwarzschild Geometry as Source of Four Predictions

Fig. 2 illustrates the geometry calculated by Schwarzschild from Einstein's general relativity for the region within and around a centre of attraction such as the Sun. This Schwarzschild geometry leads directly to four well known predictions:

(1) the bending of light by the Sun.

$$\theta = 4M_{\odot}/R_{\odot} = 4 \times 1.47 \text{ km}/6.96 \times 10^{5} \text{ km} = 1.75'' \ ; \qquad (5)$$

(2) the redshift of light from the Sun,

$$z = \Delta\lambda/\lambda = M_{\odot}/R_{\odot} = 2.12 \times 10^{-6} \ , \qquad (6)$$

* The units here are geometrical. The factor of conversion from the conventional unit of time to the geometrical unit of time is $c = 3.00 \times 10^{10}$ cm/sec ; from the conventional unit of mass to the geometrical unit of mass is $G/c^{2} = 0.742 \times 10^{-28}$ cm/g (Earth mass, 0.44 cm; Sun mass, 1.47×10^{5} cm).

and on earth,

$$z = g_{\text{conv}} h/c^2 = gh = h/0.92 \times 10^{18} \text{ cm.}; \qquad (7)$$

(3) the relativistic precession of the perihelion of Mercury about the Sun,

$6\pi M_{\odot}/a(1 - e^2)$ radians per revolution or $43.''15$ per century; (8)

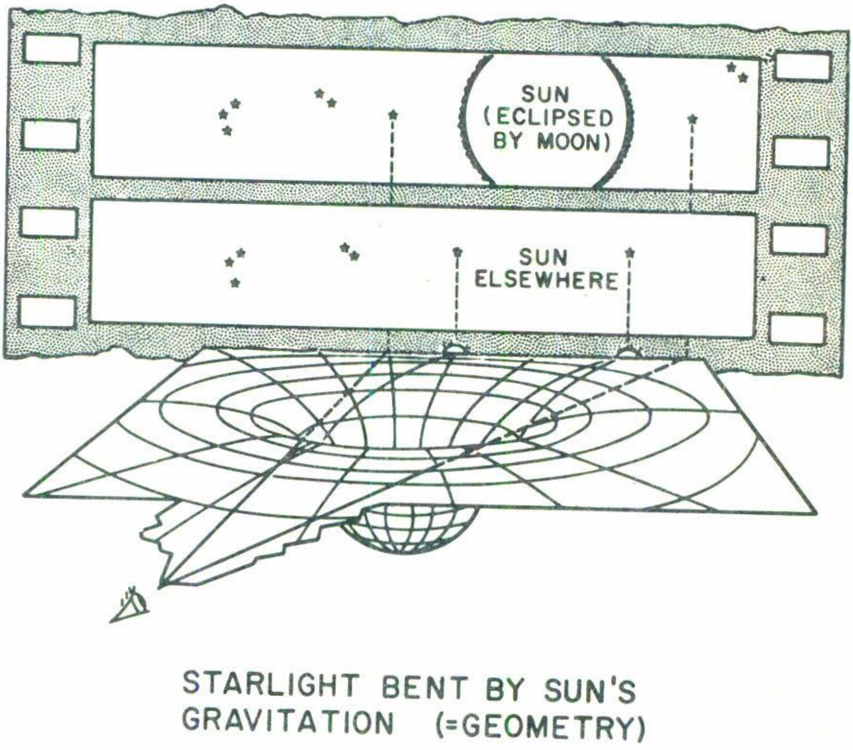


Fig. 2. Geometry within and around the Sun. Both inside and outside, the geometry departs from flatness (non-zero components of the "tide producing acceleration" or Riemannian curvature $R^{\alpha}_{\beta\gamma\delta}$); but outside, "Einstein's conserved tensorial measure of curvature", $G_{\mu\nu} = R^{\alpha}_{\mu\alpha\nu} - \frac{1}{2} g_{\mu\nu} g^{\beta\delta} R^{\alpha}_{\beta\alpha\delta}$, is zero. The analogy is close with electrostatics, where (1) the individual second derivatives $\partial^2\phi/\partial x^2$. $\partial^2\phi/\partial y^2$, $\partial^2\phi/\partial z^2$ of the electric potential have non-zero values both outside and inside a spherically symmetric cloud of electric charge, but (2) the combination of second derivatives, $\partial^2\phi/\partial x^2 + \partial^2\phi/\partial y^2 + \partial^2\phi/\partial z^2 = 4\pi\rho_{\epsilon}$, vanishes wherever there is no electric charge ($\rho_{\epsilon} = 0$).

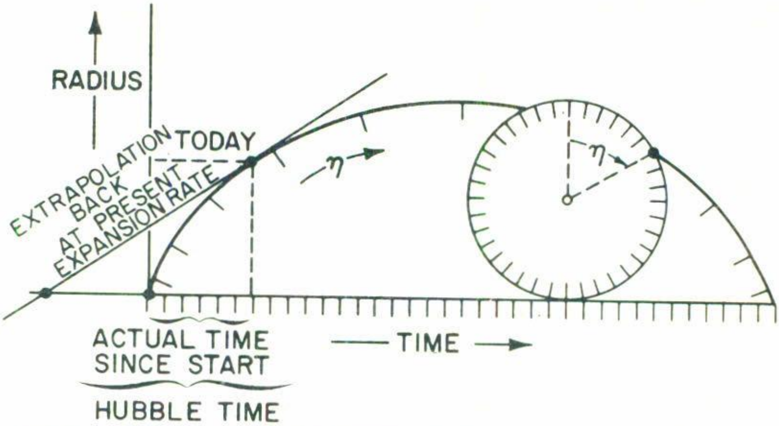
(4) the $\sim 200\,\mu$sec delay experienced by a radar pulse on its trip from Earth to Venus and back, when it passes close to the Sun. But by far greater than any of these consequences of general relativity is the revolutionary prediction that the universe itself is dynamic.

Geometry as Dynamic New Participant in Physics

In the first days of general relativity geometry had been only the slave of matter. Matter here curved space here. Curvature here meant curvature there. Curvature there meant grav.:ation there. Thus Einstein's Riemannian geometry seemed to do nothing more than carry Newtonian pull from one mass to another. Then, before Einstein's eyes, geometry cast off its chains. It stepped onto the stage of physics as a participant in its own right. It asserted dynamic degrees of freedom of its own. Earlier, under Maxwell, the electromagnetic field had also won liberation and a position as an independent dynamic entity. However, geometry became all this and more; not only new dynamic entity, but also background and home for all other fields.

Cosmology and the Closure of the Universe

Nowhere did the new dynamics of geometry display itself more dramatically or more simply than in the predicted expansion and recontraction of a closed model universe, filled to effectively uniform density with a "dust" of stars. The uniformity and the "dust" (i.e., negligible pressure) were conveniences in the analysis; but the closure was to Einstein a matter of principle. This closure, moreover, owing to the advance of astrophysics, looks like someday being a testable prediction. For example, the apparent angular diameter of objects of standard size, that goes down forever with increase in distance in Euclidean geometry, is predicted in the Friedmann universe to go up again with distance at sufficiently great distances, owing to the lens-like action of the great curve of space itself.[8] Test or no test, I would be omitting an important point if I did not suggest that *Einstein's general relativity means today not only the set of differential equations that bear his name, but also the boundary condition of closure that marks solutions of these equations as interesting*. Closure was demanded in Einstein's eyes by Mach's principle[9]: "... this idea of Mach's corresponds only to a finite universe, bounded in space, and not to a quasi-Euclidean, infinite universe. From the standpoint of epistemology it is more satisfying to have the mechanical properties of space completely determined by matter, and this is the case only in a space-bounded universe."



ILLUSTRATIVE VALUES ALL DERIVED FROM	
• TIME FROM START TO NOW	10×10^9 yr
• HUBBLE TIME NOW	20×10^9 yr
• HUBBLE EXPANSION RATE NOW	$49.0 \frac{\text{km/sec}}{\text{Megaparsec}}$
• 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 yr
• 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}

Fig. 3. The dynamics of the Friedmann matter-dominated universe is spelled out by tying a paint brush to the rim of a wagon wheel and rolling the wheel along beside the side of the barn. Vertical coordinate gives radius of curvature of the 3-sphere (cm); horizontal coordinate gives time from the start of the expansion (in cm of light travel time). Illustrative numbers are adapted from ref. 3, Box 27.4. At very early times and very late time radiation dominates over matter in any model universe at all compatible with what one knows of the actual universe; but the resulting corrections to the cycloid curve and the listed numbers are small and, for simplicity, are not shown.

Elsewhere[10] he remarked, "In my opinion the general theory of relativity can only solve this problem [of the origin of inertia] satisfactorily if it regards the world as spatially self enclosed." Some able physicists disagree: but this is *Einstein's* relativity.

No one can forget that it was the Russian meteorologist and physicist A. Friedmann who first worked out the dynamics of Einstein's simple closed model universe (Fig. 3). All follows from the decisive "00" or "tt" component of the standard geometrodynamic law,

$$G_{00} = 8\pi T_{00} \ . \tag{9}$$

Specialize to a 3-sphere universe, of time dependent radius $a(t)$, filled with "dust" of density ρ given by some constant divided by the volume of the 3-sphere. Then this equation reads

$$\frac{3}{a^2}\left(\frac{da}{dt}\right)^2 + \frac{3}{a^2} = 8\pi\rho = \frac{\text{constant}}{a^3} \ . \tag{10}$$

Multiply by $a^2/3$, rearrange and give the so-far-unspecified constant of proportionality (a measure of the total of the masses of the individual "dust" grains or stars) the name a_0, to stand for the radius of the universe at the phase of maximum expansion; thus,

$$\left(\frac{da}{dt}\right)^2 - \frac{a_0}{a} = -1 \ . \tag{11}$$

In what way does this Friedmann result differ from what one would expect for a compact cluster of rocks sitting out in space, suddenly driven apart by a blast of dynamite at the centre of the cluster? There the corresponding formula, with a now identified as the radius of the cluster, reads

$$\left(\frac{da}{dt}\right)^2 - \frac{a_0}{a} = \text{constant} \ . \tag{12}$$

The term $-a_0/a$ is a measure of the gravitational potential energy of binding of the cluster, and is always negative. The constant on the right, on the other hand, measures the excess of the energy of the dynamite over the original binding of the cluster. If the explosion is strong enough, the "constant of energy" is positive, and the cluster flies apart for ever. If the explosion is weaker than a certain critical amount, the "constant of energy" is negative, and eventually pulls the Newtonian cluster back together again.

Einstein's closed universe has no such option. There is no adjustable "constant of energy" on the right hand side of the equation. The system is gravitation-dominated at all times. The radius rises to the maximum amount $a = a_0$ (proportional to the amount of matter present) and then recontracts and collapses to zero. That the Einstein geometrodynamics of a closed universe always ends in collapse has been proved in recent times without any appeal to spherical symmetry and under remarkably general conditions.[11, 12] Einstein's condition of closure is essential to this reasoning.

Old textbooks deal with closed model universes that expand, pause or nearly pause at a certain radius, and then start again to expand, at first slowly, then more and more rapidly. A spacial case of these now disfavoured models is a universe that stands forever in unstable equilibrium at a certain radius, like a pencil balanced for all time on its tip. Such models lie outside Einstein's 1915 (and today standard) general relativity. The radius as a function of time does not fulfil the normal geometrodynamic law. It satisfies another law, obtained by adding to Einstein's equation a so-called cosmological term. Without that ill-starred term relativity would have shouted out the greatest of predictions, the prediction that the universe itself is dynamic. Today, letting that term fade into the oblivion of the past, one can say that the "would-have-been prediction" *is* the greatest prediction of all. It is a prediction almost too fantastic to be believed, and a prediction that is nevertheless dramatically confirmed by observation.

In 1915 one thought of the universe as enduring from everlasting to everlasting. Einstein could not believe the prediction that the universe is dynamic. He tried to escape it. But the considerations that lead to general relativity are compelling. He could find no natural way out. Therefore he took the least unnatural way out that he could find. He introduced the "cosmological term". Its whole purpose was to make possible a static universe. Then came 1927 and Hubble and the discovery that the universe is dynamic. Thereafter Einstein spoke of the cosmological term as [13] "the biggest blunder of my life".

Two Other Cycles of Doubt and Test

That was the first of the three great cycles of doubt and test of general relativity. The second came when the Hubble time, the "extrapolated time" for galaxies to arrive at their present distances expanding at their present recession velocities, turned out to be shorter (of the order of 2 to 3×10^9 years) then the best estimates one could make of the actual age of the universe (of the order of 10×10^9 years). This meant that the expansion had been speeding up. In contrast, the Einstein–Friedmann predictions say it must slow down ("pull of gravitation"). So "Give up general relativity," more than one group said. Thus came the era of theories outside the framework of relativity and at variance with principle that the laws of physics are local in character: theories of the "steady state expansion of the universe" and theories of "the continuous creation of matter". Then, thanks not least to Walter Baade [14], a revolution took place in one's understanding of the scale of astrophysical distances. Previously accepted distances to the galaxies in the Hubble catalogue had to be revised upward by a factor between 6 and 10. The linearly extrapolated time, H^{-1}, back to the start of the expansion rose by the same factor to a value now estimated to be not far from 20×10^9 years[15]. But galaxies actually got where they are in about half this time, according to more than one way of evaluating the time back to the beginning of all astrophysical processes ($\sim 10 \times 10^9$ years ago.) Therefore one believes today that one has clear evidence for the predicted slowing down of the expansion.

The third cycle of doubt and test began in 1958, when Jan Oort[16] gave 0.31×10^{-30} g/cm^3 as the best available figure for the averaged out density of matter present in the form of galaxies. In contrast, an amount of mass-energy of the order of 15×10^{-30} g/cm^3 is required to curve space up into closure (Fig. 3). Einstein's geometrodynamics thus predicts that there is of the order of 10 to 100 times as much mass-energy in space as one sees in the form of galaxies. But where? And in what form? This "mistery of the missing mass" is the central point of much present-day work.[17] Hydrogen gas unassembled into galaxies[18] will some day be detected by its ultraviolet absorption[19] or by its X-ray emission[20] if it is present in the required amount.

If cosmology once seemed a subject fit only for dreamers, today it is the heartland of observational astrophysics. For example, nothing did more to destroy the concept of a "steady state expansion" of the universe than the observation of the 3°K primordial cosmic fireball microwave radiation.[21, 22]

The Firedmann-Einstein prediction that the universe itself is dynamic, in the beginning too incredible for even Einstein himself to believe, has now become a central fact of modern physics. With the universe proved dynamic, one is the readier to accept three other ideas from Einstein's general relativity: (1) that other incredible prediction, that collapse is inevitable; and two prior ideas, (2) that the universe is closed, and (3) that geometry is a new dynamic participant on the stage of physics.

"Total Energy" and "total Momentum" as Concepts with No Meaning for a Closed Universe

The closure of Einstein's universe has a special consequence for energy. The law of conservation of energy connects the amount of mass-energy inside a closed surface with the value of a certain integral extended over that surface. Deform this surface of integration bit by bit in imagination at any one time so as to engulf more and more volume. At first the surface swells. Then it reaches a maximum extent. When it includes the entire volume, it has collapsed to nothingness. Thus the law of conservation of energy, applied to the complete closed system, degenerates to the trivial identity, "zero equals zero". The concept of "total mass-energy" makes no sense for a closed universe. How could it? (1) There is no natural Lorentz frame in which to do the pointing and measuring off of a 4-vector of energy and momentum even if one had such a 4-vector. (2) There is no platform' on which to stand to measure the gravitational attraction of the closed system. (3) There is no place outside the system to put a planet into Keplerian orbit around it. It is satisfying that the mathematics kills at the start a concept that is bad physics. There is no such thing as the energy (or the angular momentum) of a closed universe.[23] *The dynamics of a closed geometry transcends the laws of conservation of angular momentum and energy.*

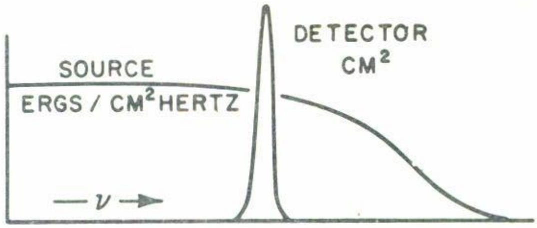
Gravitational Radiation from Gravitational Collapse

The dynamics of geometry, so central to these cosmological considerations, must also reveal itself in a testable way at a smaller scale in gravitational radiation, according to Einstein's standard general relativity. Moreover, the basic factor in the formula for gravitational radiation[24] is the very large number

$$P_{0,\mathrm{grav}} = c^5/G = 3.6 \times 10^{59}\ \mathrm{erg/sec}. \quad (13)$$

In other words, any system, big or small, that is highly asymmetric, and that changes its configuration in a time comparable to the time required for light to cross it, will give off gravitational radiation at a rate of the order of magnitude of $P_{0,\mathrm{grav}}$.

Few events are more spectacular than a supernova, nor more relevant as a source of gravitational radiation. A normal star with slowly rotating white dwarf core develops gravitational instability in the course of its standard astrophysical evolution. The core collapses to a rapidly rotating neutron star. As the core implodes, it generates a powerful shock, in consequence of which the envelope explodes. This is now accepted picture goes back for its beginnings to 1934 and Baade and Zwicky.[25] It received dramatic support in 1968 when Hewish and his collaborators[26] discovered the first few pulsars, among them one pulsing 30 time a second. It lies at that point in the Crab Nebula where Baade and Zwicky, 34 years before, had said the neutron star (from the July 1054 supernova) should be.



SOURCE: 1 TO 10 $M_\odot$ → SUPERNOVA PLUS {n-STAR OR BLACK HOLE

- IN MILKY WAY, ONCE /100yr. UP TO ONCE /yr.
- IN VIRGO CLUSTER, ~1000-FOLD MORE
- SCENARIO OF COLLAPSE, BREAKUP, CIRCULAR PURSUIT, DAMPING BY GRAVITATIONAL RADIATION, AND RECOMBINATION STRONGLY DEPENDENT UPON INITIAL MASS AND ANGULAR MOMENTUM
- PULSE (OR SEVERAL PULSES, MINUTES TO HOURS APART, DEPENDENT UPON SCENARIO), LENGTH $\sim 10^{-3}$ SEC; FREQUENCIES UP TO $\sim 10^3$ HZ
- INTENSITY $\sim 3 \times 10^5 \frac{\text{ERG}}{\text{CM}^2\,\text{HZ}} \left(\frac{\text{DISTANCE, DETECTOR TO CENTER OF MILKY WAY}}{\text{DISTANCE, DETECTOR TO SOURCE}} \right)^2$

BAR DETECTOR: $M = 10^6$ G: $\nu_0 = 1660$ HZ

- INTEGRATED CROSS SECTION $\int \sigma \, d\nu = 10^{-21} \text{cm}^2 \text{HZ}$
- CALCULATED TO BE ABLE TO DETECT SOURCE ANYWHERE IN MILKY WAY OPERATING AT 3 °K WITH DAMPING TIME OF 20 SEC IF ATTACHED SENSOR CAN MEASURE CHANGE IN VIBRATION AMPLITUDE $\lesssim 10^{-16}$ CM ON SCALE $\lesssim 0.1$ SEC (1972 TECHNOLOGY)
- ANYWHERE IN VIRGO CLUSTER WHEN OPERATING AT 3×10^{-3} °K WITH DAMPING TIME $\sim 10^6$ SEC IF SENSOR DETECT $\sim 10^{-19}$ ON TIME SCALE $\lesssim 0.1$ SEC. (LATE 1970's, EARLY 1980's)

Fig. 4. Gravitational radiation, from source to detector.

To follow the internal dynamics of future supernovae, an optical telescope does not suffice, nor does X-ray or infrared astronomy, remarkable though the advances are today in all three methods of observation. Oceans of star stuff block all view of the Niagara Falls that pours its tumult inward at the centre. One signal nevertheless makes its way out, a pulse of gravitational radiation, with characteristic shape, yet to be calculated, dependent on the mass and angular momentum of the collapsing core. What has been calculated is the order of magnitude of the pulse. Press and Thorne conclude[27] that a Weber-bar detector,[28] built of a 100 kg monocrystal of quartz, cooled to $T = 3 \times 10^{-3}$K, if it has a $(1/T)$-proportional damping time of $\sim 10^6$ sec (technology of late 1970's or early 1980's) should suffice to detect gravitational waves from a supernova in the Virgo cluster of galaxies (distance 3×10^7 light-years), if one can construct a sensor to measure changes in vibration amplitude $\lesssim 10^{-19}$ cm on a time scale of $\lesssim 0.1$ sec.(Fig. 4). Supernovae flash out within that distance once a month or more.

Thanks to the initiative of Joseph Wéber and the subsequent work of many other able investigators, at least twenty detectors of gravitational radiation have been constructed and exploited to give upper limits to the flux of energy streaming past the earth at selected frequencies. The time nears for a decisive test of one of the greatest of Einstein's predictions, energy-bearing waves in the geometry of spacetime itself.

Collapse of a Too Massive Neutron Star to a Black Hole

Complete gravitational collapse of an overcritical mass, $M > M_{crit}$, is another great prediction. The precise value of the critical mass of a neutron star is uncertain but is believed to lie in the range $0.5M_{\odot} < M_{crit} < 3M_{\odot}$ (unless, as would seem possible at most for the first few days of its life, it is endowed with large amounts of differential rotation, in which case James Wilson gives a figure about 50% larger). When a neutron star of greater mass is formed by the gravitational collapse of the core of a star with white dwarf core, the collapse may slow down temporarily as neutron-star densities are reached (10^{14} to 10^{15} g/cm^3); but the collapse is then predicted to continue and to speed up, with the matter becoming more and more compact, until a horizon forms and a black hole comes into being. The proper circumference of the horizon divided by 2π, otherwise known as the Schwarzschild radius of the black hole, is $2M$ (cm) $= 2(G/c^2)\,M_{conv}$ (g); that is, 3 km (roughly a tenth the size of a neutron star) for an object of solar mass, and 10 to 10^4 light seconds for a black hole of 10^6 to $10^9 M_{\odot}$, such as one may expect to find in a compact and highly evolved galactic nucleus.

In contrast to the "dead" or Schwarzschild black hole of the traditional text, the object formed in the collapse of matter with any net spin angular momentum at all, $S \neq 0$, is a "live" black hole, as first emphasized by Bardeen[29]; and it can give up energy to an external particle of field, as first pointed out by Penrose.[30] Hawking[31] showed that neither in the Penrose process, nor in any other process, can be surface area of the horizon of a black hole ever increase. Independently Christodoulou[32] showed that a black hole is characterized by an "irreducible mass", M_{ir} (later shown to be connected with the area of the horizon by the formula $A = 16\pi M_{ir}^2$). The "irreducible mass" is constant in any process that reversibly exchanges energy with a black hole, but it always rises in any irreversible process. Christodoulou, and Christodoulou and Ruffini[33], derived the wonderfully simple formula

$$M^2 = \left[M_{ir} + \frac{Q^2}{4M_{ir}}\right]^2 + \frac{S^2}{4M_{ir}^2} \tag{14}$$

for the mass-energy of a "live" black hole in terms of its charge and spin.

Three processes offer themselves for the detection of a black hole: (1) the pulse of gravitational radiation given out at the time of formation: (2) the X-rays given out in the traffic-jam of matter accreting onto a black hole after formation, as analyzed by Zel'dovich and Novikov[34]; and (3) "activity": activity arising from energy imparted to outside matter, or fields, or both, out of the stockpile of energy in a live black hole (see for example ref. 35 and 36). All three processes are being actively investigated, and have many interesting astrophysical consequences, most of which are reviewed in some detail in the 1972 Les Houches lecture series[37] and in ref. 3.

Roughly 50% of all stars are "married"; and of such double star systems, roughly 40% are near-binaries, with periods of the order of a few days. When one component of a revolving double star system is a neutron star or black hole, it has a good chance to feed on the envelope of its companion, and in consequence become a powerful source of X-rays. When the compact component is a neutron star, its rotating off-axis magnetic field produces the normal pulsar phenomenon, but in a denser than normal plasma. Whether the compact component is a neutron star or a black hole, the impouring gas, adiabatically compressed to $10^{10} - 10^{11}$ K, emits far more radiation in the X-ray region than in the visible. Only in this way has one been able to understand some of the spectacular eclipsing X-ray sources observed in recent months by Giacconi and his collaborators. Leach and Ruffini emphasize[38] the sharp division of these double-star X-ray sources into two classes. In one class the X-ray source flashes regularly like an optical pulsar. In this case, it is generally agreed, the compact (and optically invisible) component is to be identified with a neutron star. In the other class (two cases so far, Cygnus X - 1, and the X-ray sources 2U1700-37) the X-ray intensity fluctuates, with the fluctuations amounting to as much as a factor of a hundred in a time as short as 50 msec. Ruffini reasons that this effect indicates (1) small size and (2) hydrodynamic instability of the flow of plasma into the black hole. In conformity with this reasoning, the mass of the compact component (as deduced from the period and range of Doppler velocities of the visible component) appears in the one case to be more than $8M_\odot$, and in the other case more than $4M_\odot$. If this object were a normal star, it would be far too bright (luminosity $\sim 8^3 L_\odot$ and $\sim 4^3 L_\odot$, respectively) to escape observation. It cannot be a white dwarf, because for these objects the critical mass limit is $M_{crit} \simeq 1.2M_{\odot}$; and likewise it cannot be a neutron star if for such objects the critical mass is indeed $M_{crit} < 3.2M_{\odot}$. Few see any alternative for these two X-ray sources except to conclude that the compact object is a black hole. Moreover, it is difficult to imagine how a neutron star continuously fed from a sufficiently massive companion can ever end up as anything except a black hole. Therefore it seems reasonable to conclude that science has now been launched, quietly but momentously, into the age of black hole astrophysics.

Black Hole as "Experimental Model" for the Collapse of the Universe Itself

With black holes one has come full circle around the application of Einstein's geometrodynamics, past the traditional tests of general relativity, through the world of gravitational radiation, and into the world of gravitational collapse. The black hole of today is more than a black hole. It is symbol, "experimental model", and provider of lessons for the collapse Einstein predicted in far later days for the universe itself.

If collapse is the most startling prediction that physics has ever made, it is also true that general relativity (except for the quantum principle) is the strangest edifice that physics has ever reared. Therefore it is appropriate to look into this structure from windows other than Eisntein's original point of entry, aiming especially in the later derivations to enlarge one's view of what collapse is and what it means.

3. CARTAN'S DERIVATION: CONSERVATION OF THE SOURCE COMES ABOUT VIA THE PRINCIPLE THAT 'THE BOUNDARY OF A BOUNDARY IS ZERO'

Riemann Rotation or "Tide-Producing Effect" Associated with Each Face of a Cube

The central point of electrodynamics is conservation of charge. The central point of geometrodynamics is conservation of mass-energy. Take the relevant field – the electromagnetic field in the one case, the Riemannian curvature or "tide-producing acceleration" in the other – and *"wire the field up" to the source in such a way that* this *conservation comes about automatically, through the principle that "the boundary of a boundary is zero"*. These ideas go back for their origin to Cartan[39] (see note 39 & ref. 6 for a more complete exposition) and are illustrated in Fig. 5. Rather than look at all of spacetime,

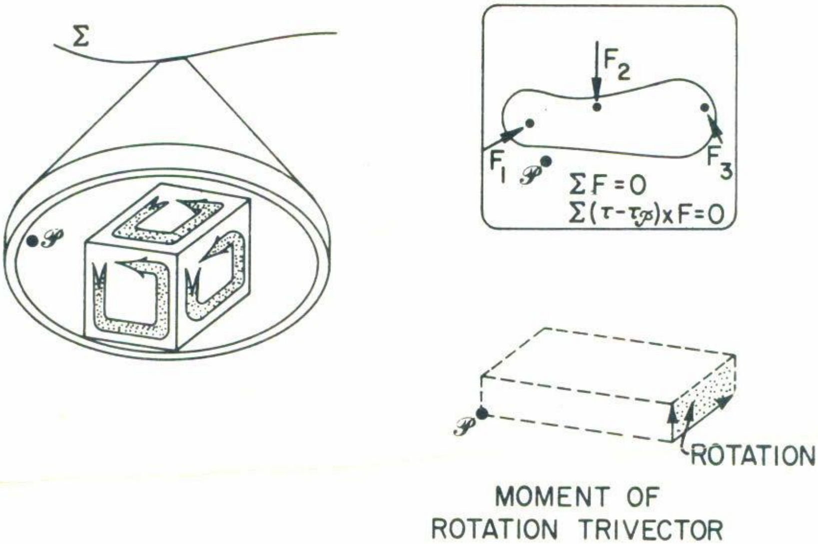


Fig. 5. The analogy with mechanics as background for Cartan's "moment of rotation". The value of the moment of rotation, totalled over all six faces of the elementary cube, is independent of the location of the point ℘. Likewise in mechanics the location of the point ℘ makes no difference in the statement of the conditions for mechanical equilibrium.

direct attention to an arbitrary "simultaneity" or spacelike hypersurface Σ slicing through spacetime. Rather than examine all of Σ, focus (see enlarged view through magnifying lens) on a small cubical 3-dimensional element of volume located anywhere on Σ, and narrow attention to the "front" face of this cube. Place a vector at the upper left hand corner (ULHC) of this face. Transport the vector parallel to itself around the periphery of this route, in the sense indicated by the arrow, ending up back at the original starting point. The vector undergoes a rotation. This rotation is proportional to (1) the size of the face and (2) the relevant component of the Riemann curvature of the 4-dimensional geometry. Repeat, taking the same vector on a tour from the same starting point and ending up at the same end point but this time around the top face of the cube. Repeat for all 6 faces of the elementary cube. Then the combined effect of all six rotation totals to zero. The cancellation of rotations occurs because each edge of the cube has been traversed as often in one direction as in the opposite direction. In other words, the 3-cube has a boundary that is made of six 2-dimensional surfaces; and each surface has a boundary that is made of four 1-dimensional edges. However, each edge occurs twice. Thus, when due account is taken of sign, the contributions of all edges cancel. In brief, the 1-dimensional boundary of the 2-dimensional boundary of an elementary 3-dimensional volume, V, is automatically zero; or, in the symbolism of algebraic geometry,

$$\partial \partial V \equiv 0 \,, \tag{15}$$

where ∂ stands for "boundary of". The resulting statement about the Riemannian curvature of spacetime, the so-called Bianchi identity, takes the form

$$\sum_{\substack{\text{all six}\\ \text{faces}}} \begin{pmatrix} \text{rotation associated} \\ \text{with each face} \end{pmatrix} \equiv 0 \,. \tag{16}$$

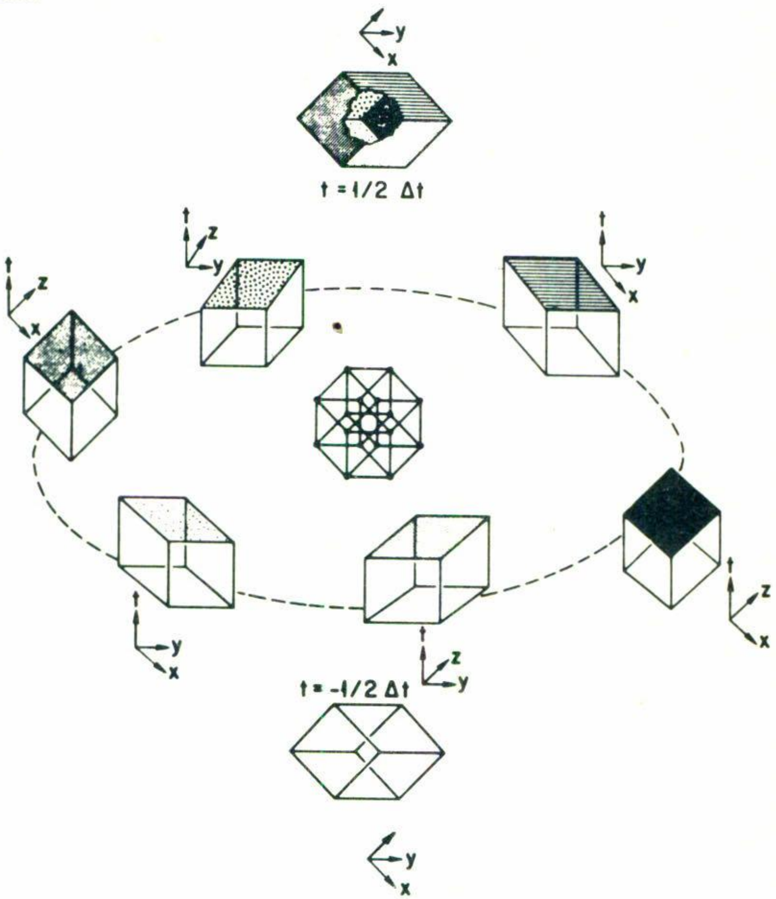


Fig. 6. The principle that "the boundary is zero" in its 4-3-2 dimensional form. Exploded off the 4-cube at the centre of the figure are its eight 3-dimensional faces, every one a cube. Each of these cubes has six 2-dimensional faces. However, these 2-dimensional faces counterbalance each other in pairs; or, otherwise stated, and with due account of sign, the $8 \times 6 = 48$ faces "add up to zero". As example, the black face of the top cube nests against the black face of the right hand cube. Thus the 4-dimensional cube exposes no 2-dimensional face to the outside world; it is "faceless". The boundary of the boundary of the 4-dimensional cube is zero.

Moments in Mechanics and in Geometrodynamics

Compare geometry to mechanics. The body in the inset in Fig. 5 cannot be in equilibrium unless the forces all add to zero:

$$\sum_{\text{all forces}} \mathbf{F} = 0 \,. \tag{17}$$

However, for equilibrium, another requirement must also be satisfied. The moments must add to zero:

$$\sum (\boldsymbol{r} - \boldsymbol{r}_{\wp}) \times \mathbf{F} = 0 \tag{18}$$

About what point the moments are taken does not matter, by reason of the requirement $\sum F = 0$ (cancellation of the multiplier or $r_{\wp}$).

Turn from the idea of "the moment of a force" in mechanics to the idea of "the moment of a rotation" in geometrodynamics. It will not matter about what point one evaluates these moments. Therefore select the arbitrary point $\wp$ shown in Fig. 5, both in "the view through the lens", and repeated, for better seeing, at the lower right. Also shown at the lower right, depicted as a bivector, is the rotation (measure of Riemann curvature) associated with one of the faces of the cube. This bivector, together with the vector from $\wp$ to the center of the relevant face of the cube, defines a trivector. The value of this trivector depends upon the location of the point $\wp$. However, the location of $\wp$ drops out from, and has no influence on the value of, the sum of these trivectors taken over all six faces of the cube:

$$\sum_{\substack{\text{all six}\\ \text{faces}}} (\boldsymbol{r} - \boldsymbol{r}_{\wp}) \wedge \begin{pmatrix}\text{rotation associated}\\ \text{with each face}\end{pmatrix} =$$

$$= \begin{bmatrix}\text{moment of rotation}\\ \text{trivector associated}\\ \text{with elementary cube}\end{bmatrix} \rightarrow \left\{\begin{array}{l}\text{identified by general relativity with } 8\pi \text{ times}\\ \text{the trivector representation (dual to an ordi-}\\ \text{nary vector) of the amount of energy and mo-}\\ \text{mentum contain in this cube (= `content of}\\ \text{source' in the cube)}\end{array}\right. \tag{19}$$

Identify this sum with 8π times the amount of energy-momentum contained in this elementary volume. Repeat this statement for all spacelike slices through the given region of spacetime, and for all regions of spacetime. Then one has stated the entire content of Einstein's 10-component field equation. This is relativity in brief!

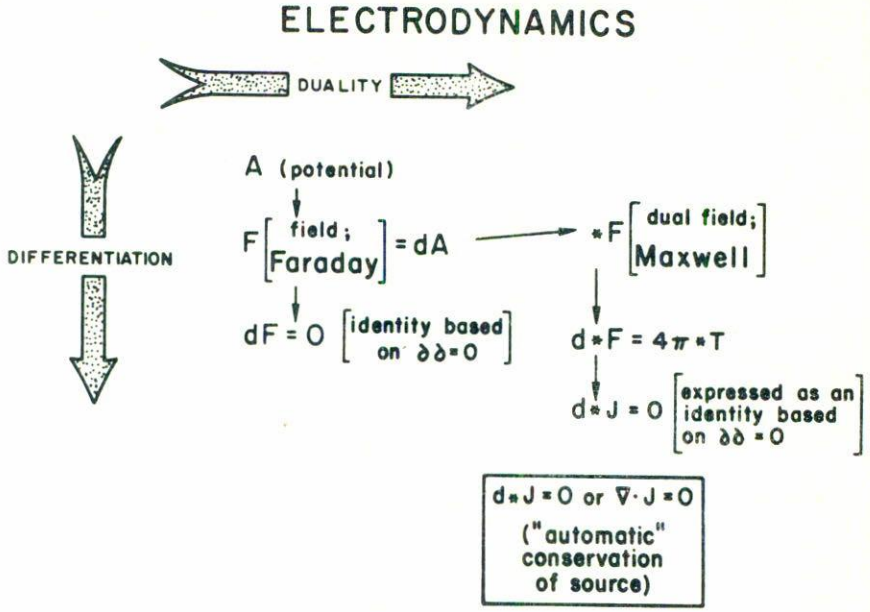


Fig. 7a.

Conservation of Energy-Momentum via the "2-Facelessness" of the 4-Cube

How does this "wiring up" of the "field" (geometry) to the "source" (energy-momentum) guarantee the desired conservation of the source? How does it guarantee that, as time goes on, say from $t = -\frac{1}{2}\Delta t$ to $t = +\frac{1}{2}\Delta t$, no source is created in the element of 4-volume, $\Omega = V\,\Delta t$ (Fig. 6)? To have conservation means that the amount of source in the top cube (V at $t = \frac{1}{2}\Delta t$) must turn out to be equal to the amount of source in the bottom cube (V at $t = -\frac{1}{2}\Delta t$) plus the inflow of source during the time Δt (as described by the "inflow" or "content of source" in the six remaining cubes of Fig. 6); or means that the "content of source" in all cubes together, with due account of sign, must add up to zero. But equation (19) wires up the source to the field in such a way that the content of source in any one cube is given by the sum of (moments of rotation) associated with the faces of that cube; and the contributions of all 8×6 faces together cancel out identically; thus

$$\begin{pmatrix}\text{creation of}\\ \text{source in } \Omega\end{pmatrix} = \sum_{\substack{\text{all eight}\\ \text{3-cubes}\\ \text{bounding } \Omega}} \begin{pmatrix}\text{content of source in}\\ \text{3-cube}\end{pmatrix} =$$

$$= \sum_{\substack{\text{all eight}\\ \text{3-cubes}}} \; \sum_{\substack{\text{all six faces}\\ \text{of given 3-cube}}} \frac{1}{8\pi} \begin{bmatrix}P\text{-dependent moment of}\\ \text{rotation associated with}\\ \text{that face}\end{bmatrix} \equiv 0 \qquad (20)$$

$\uparrow$ (because the 48 faces cancel out identically, in pairs)

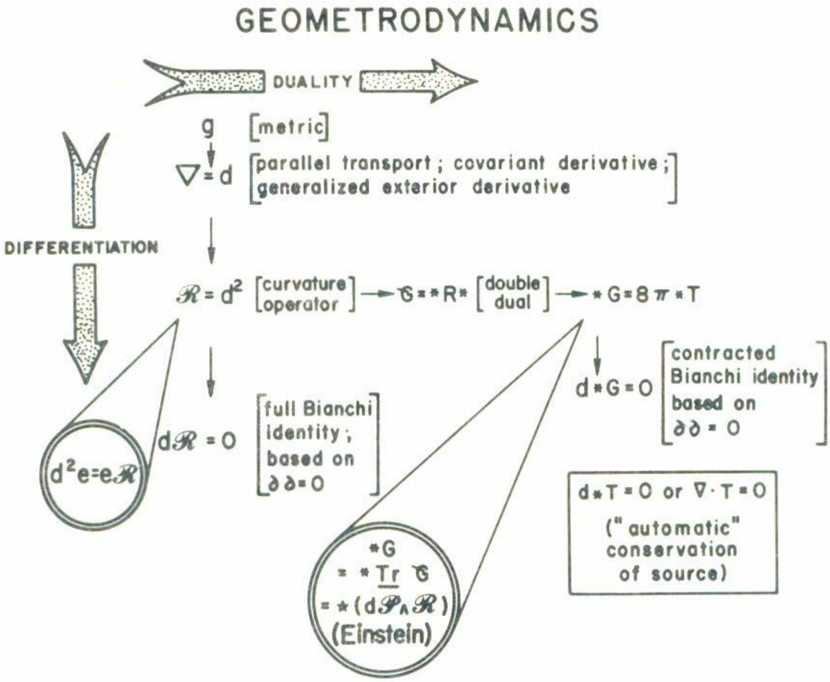


Fig. 7. (a) The structure of electrodynamics compared to (b) the structure of geometrodynamics. In both diagrams the principle that "the boundary of a boundary is zero" appears twice, once in the left hand column in its 3-2-1 dimensional form, and again in the right hand column, in its 4-3-2-dimensional form (diagrams adapted from Misner, Thorne and Wheeler[3]).

The same "conservation via the principle $\partial\partial \equiv 0$" applies in electromagnetism, as one sees by comparing Figs. 7a and 7b.

Algebraic Geometry Rises above Dimensionality

One used to believe, and often still finds it useful to postulate, that the source comes first in the scheme of things, and the field second. However, one sees that today the possibility is open to think of the field as coming first. On this view the conservation of the source, and therefore in some sense even the existence of the source, is a consequence from and mere aspect of the existence of the field. Moreover, the principle of algebraic geometry ($\partial\partial \equiv 0$) that legislates and enforces "conservation of the source" is a principle that rises above any particular dimensionality in its most general mathematical version. But the concepts of "manifold" and "dimensionality" are presupposed in the laws of physics as they look today. Can one look beyond and above existing statements of physics to a formulation that does not presuppose dimensionality? If so, the principle "$\partial\partial \equiv 0$" would seem an essential part of such a formulation. No principle reaches closer to the heart of general relativity.

4. THE MOST COMPACT FORMULATION OF GENERAL RELATIVITY

Intrinsic Curvature Plus Extrinsic Curvature Equals Energy Density

One knows no more compact statement of general relativity than this:

$$(\text{curvature}) = 8\pi\,(\text{density of mass-energy}) \tag{21}$$

More specifically, take any event P in spacetime, and any spacelike hypersurface Σ through P, and that local Lorentz frame at P in which Σ is a "simultaneity". Take the density ρ (in cm^{-2}; cm of mass-energy per cm^3 of volume) in this frame, multiply it by 8π, and equate the product to the linear scalar measure of the 4-dimensional curvature projected on Σ; thus (after doubling)

$$\underbrace{\underbrace{{}^{(3)}R}_{\substack{\text{intrinsic}\\ \text{curvature}}} + \underbrace{(\mathrm{Tr}\,K)^2 \qquad\qquad \mathrm{Tr}\,K^2}_{\substack{\text{"second invariant" of the extrinsic curvature; or, more}\\ \text{briefly, "extrinsic curvature"}}}}_{\substack{\text{twice the linear scalar measure of the}\\ \text{4-dimensional curvature projected on } \Sigma}} = 16\pi\rho \tag{22}$$

Make this demand for every inclination of the hypersurface through P, and for every choice of P, and have in this one demand the whole content of all ten components of Einstein's field equation.

In electrodynamics one similarly requires

$$\mathrm{div}\ \mathbf{E} = 4\pi\rho_e \tag{23}$$

and imposes (covariance plus) this demand for every inclination of the hypersurface Σ through P and in this way recovers the other three Maxwell equations,

$$\mathrm{curl}\ \mathbf{B} = \dot{\mathbf{E}} + 4\pi\mathbf{j}_e \tag{24}$$

In geometrodynamics, additional to the inclination of Σ, the curvature of Σ seems to matter, as evidenced not least in the appearance of the 3-dimensional scalar curvature invariant, ${}^{(3)}R$, in (22). However, the remaining two terms in (22) not only compensate for this curvature, but even follow uniquely[3] from the requirement that they should compensate for this curvature of Σ. Thus the left hand side of (22) is a measure of the 4-dimensional curvature. In this equation K (units cm^{-1}) is the so-called tensor of extrinsic curvature of the hypersurface Σ. It measures the fractional contraction of any local geometric object in Σ when all points of this object are projected forward a unit distance in time (cm) normal to Σ.

For another window into the content of general relativity we now turn from geometry to dynamics as the guiding idea.

5. FROM HILBERT'S DERIVATION TO SUPERSPACE

Hilbert's Principle of Least Action

In no branch of dynamics does a variational principle give a more comprehensive grip on the whole subject than in general relativity. David Hilbert recognized this point and presented the new variational principle to the Göttingen Academy[40] on 20 November, 1915. His step forward derived its guidance and inspiration from Einstein's earlier work. However, it based itself upon a principle of least action from the start. The resulting geometrodynamic law, independent of Einstein in its derivation, was nevertheless identical in form with what Einstein was to lay before the Berlin Academy only five days later.

The idea is simple. Give one spacelike hypersurface σ_0 and a second spacelike hypersurface σ and fill in between them a 4-geometry, ${}^{(4)}\mathcal{G}$. Try different 4-geometries. For each calculate the action integral*,

$$I = (1/16\pi)\int_{\sigma_0}^{\sigma} {}^{(4)}R \; \mathrm{d}\,(4\text{-volume})\,. \tag{25}$$

*Here the element of 4-volume, generalizing an expression like $r^2 \sin\theta\, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\phi$, is

$$\mathrm{d}(4\text{-volume}) = (-g)^{\frac{1}{2}}\, \mathrm{d}^4 x\,.$$

The integrand is the 4-dimensional scalar curvature invariant, ${}^{(4)}R$, in a problem of pure geometrodynamics; or this supplemented by the Lagrangian of the other fields when other fields are present.

That 4-geometry is allowed by classical physics that maximizes or minimizes or, more generally, extremizes this integral.

What is a 4-geometry? An automobile fender is a 2-geometry. Stretch a ruled transparent rubber sheet over the fender. In this way assign x and y coordinates to every tiny bump and pit in the metal surface. Now pull the rubber harder here and there and thus change the coordinates everywhere Yet the fender continues to keep its 2-geometry. The difference between a Ford and a Fiat fender is invariant with respect to all changes in coordinatization. Hilbert understood well that the 4-geometry resulting from his variational principle is also invariant with respect to all changes in coordinates.

What is Fixed at the Boundaries Defines "the Initial Value Problem"

To understand in addition and in coordinate-free geometrical terms what it is that one fixes on the two hypersurface σ_0 and σ is an achievement of recent times. It is also an important achievement. It permits one to state (1) what are appropriate initial value data for the classical dynamics and (2) on what the state function or probability amplitude function depends in quantum dynamics (as illustrated for the physics of a single particle in Fig. 8).

Arnowitt, Deser and Misner[41] turned away from any direct attempt to discover what was fixed at the boundaries, σ and σ_0, in Hilbert's action principle. They added a complete divergence to the Hilbert integrand. Such an addition affects in no way the resulting Einstein field equation, but does alter the quantities fixed at limits. The new quantities, expressed in coordinate-free geometrical form, turned out to be the 3-geometries, ${}^{(3)}\mathcal{G}_0$ and ${}^{(3)}\mathcal{G}$, of the bounding hypersurface, σ_0 and σ. Among other consequences of this result it follows[42] that there is a representation of quantum geometrodynamics in which the state function depends upon and is fixed by the 3-geometry:

$$\psi = \psi({}^{(3)}\mathcal{G}). \tag{26}$$

The totality of all closed 3-geometries with positive definite signature is called superspace, and what has just been discussed is often known as the superspace representation of general relativity

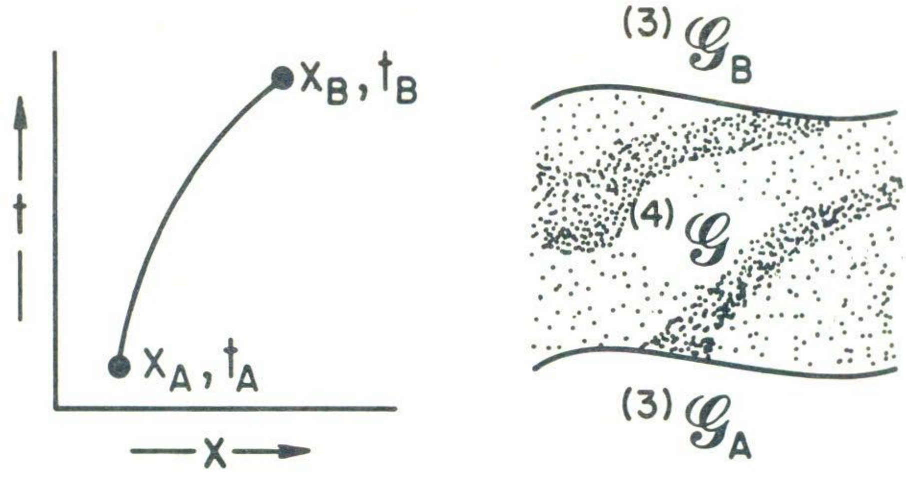


Fig. 8. The classical history of particle *in* spacetime (world line; left) compared and contrasted with the classical history *of* space (4-geometry; right). In both cases the classical history is selected out from the other conceivable histories that connect initial configuration, A, and final configuration, B, by the circumstance that it extremizes the "action integral" or "dynamical path length", $I = I(A, B)$, from A to B. In quantum physics the "wave function" or "state function" or "probability amplitude" depends upon the same variables that define the final state configuration, B; thus, $\psi = \psi(x, t)$ in particle dynamics. In geometrodynamics, one has $\psi = \psi({}^{(3)}\mathcal{G})$ in the superspace representation; or, in the York representation (conformal part of the 3-geometry and local Hubble contraction rate specified), $\psi = \psi({}^{(3)}<\Upsilon\Upsilon)$

York's Formulation of the Initial Value Data

In recent months James W. York, Jr., returning to the Hilbert principle in its original form, has discovered[43] that it demands that one should specify at each point on σ_0 and σ (1) the conformal part, ${}^{(3)}<$, of the 3-geometry and (2) the local extrinsic or Hubble time, τ, a concept first introduced by Karel Kuchař. To give the conformal part of a 3-geometry is to give for each point, not the absolute distance, but the relative distance, to every nearby point. In other words, angles are fixed, but not distances. Missing from the information that would be contained in a full 3-geometry at each space point is a scale factor; but in its place one has to specify at each space point something like the dynamical conjugate of this cale factor; namely, the rate at which this scale is decreasing with time, the local Hubble time τ, symbolized by the angular spread between two timelike vectors that stand perpendicular to the given spacelike hypersurface; thus, symbolically,

$$\Upsilon\Upsilon \text{ represents } \tau . \tag{27}$$

In this mathematical representation, York, following earlier work of André Lichnerowicz[44] and Yvonne Choquet-Bruhat[45], has been able to show that one can determine the future from the given information by simple and elegant methods. The solution of an elliptic differential equation yields the unknown scale factor. Moreover, the solution always exists and is unique.

Wave function, Wave Equation, and Hamilton-Jacobi Equation for Phase of the Wave

Quantum geometrodynamics in the York representation leads to a state function

$$\psi = \psi({}^{(3)}\!<,\ \Upsilon\ )\ ; \tag{28}$$

in the superspace representation, a state function

$$\psi = \psi({}^{(3)}\mathcal{G})\ . \tag{29}$$

In neither case is the proper order of factors in the relevant wave equation quite free of all ambiguity, despite a most valuable analysis of this problem by Bryce DeWitt.[46] However, in the semiclassical approximation, one writes

$$\psi \simeq \begin{pmatrix}\text{slowly varying}\\ \text{amplitude factor}\end{pmatrix} e^{iS/\hbar} \tag{30}$$

with the important physics showing up in the rapidly varying phase factor, $S/\hbar$. There is no ambiguity in the order of factors in the equation satisfied by the Hamilton-Jacobi function S. This definiteness follows not least because a value for S is directly given by the extremal value I of the action integral:

$$S(\sigma) = I_{\text{extremal}}\,(\sigma, \sigma_0) \tag{31}$$

Moreover, in the superspace representation, the equation for the dynamical evolution of this Hamilton-Jacobi function is a local equation. This equation was first written down by Peres.[47] It reads

$$(16\pi)^2\,(\tfrac{1}{2}g^{-\frac{1}{2}})(g_{ik}g_{jl} + g_{il}g_{jk} - g_{ij}g_{kl})(\delta S/\delta g_{ij})(\delta S/\delta g_{kl}) + g^{\frac{1}{2}}\,{}^{(3)}R = 0\ . \tag{32}$$

Here $S = S({}^{(3)}\mathcal{G})$ is, up to a factor, the phase of the wave function in superspace. Wave crests in superspace are described by surfaces of constant S. Three features of the geometry ${}^{(3)}\mathcal{G}$ put in an appearance in the Peres or 'Einstein-Hamilton-Jacobi' equation: its metric, $g_{ij}(x, y, z)$; the square root of the determinant of the metric tensor, $g^{\frac{1}{2}}(x, y, z)$; and the local value of the 3-dimensional scalar curvature invariant of the 3-geometry, ${}^{(3)}R(x, y, z)$.

Out of the law of propagation of wave crests in superspace one can deduce the law of propagation of a wave packet. In other words, one can discover how a 3-geometry evolves with time in the semi-classical approximation. In this way Ulrich Gerlach[48] has succeeded in deriving from the one Einstein-Hamilton-Jacobi equation all ten components of Einstein's standard geometrodynamic law.

Superspace as Arena for the Dynamics of Geometry

In no formulation of dynamics is the leap from the classical to the quantum outlook shorter than in Hamilton-Jacobi theory. Sharply intersecting wave crests reproduce the determinism of classical dynamics; waves of finite wavelength reproduce the finite wave packets and indeterminism of quantum dynamics. All this is familiar. What is new is superspace. It imposes itself on our attention exactly because we insist on analyzing the dynamics of geometry from the wave point of view. Demand Einstein geometrodynamics, demand the quantum principle, and end up with superspace.

What kind of an arena for dynamics is superspace? And what lessons does it teach? Fig. 9 illustrates at the left a smooth closed 2-geometry. One can approximate this 2-geometry arbitrarily closely by a polyhedron or "skeleton 2-geometry" (illustration at right) built of a sufficiently great number of faces. Euclidean geometry rules in each face. In this illustration the 98 edge lengths determine all the details of the shape of the polyhedron. Represent this information by a single point in a space of 98 dimensions. The projections of this point onto 98 coordinate axes give back all the original information about the 98 edge lengths. Move this "representative point" slightly in the 98-dimensional space. Then all 98 coordinates of this point – and therefore all 98 edge lengths of the triangles in the polyhedron – also change slightly. The skeleton 2-geometry bends, twists, swells and otherwise changes in shape in obedience to the motion of the representative point. Take the analysis given here for skeleton 2-geometries built out of triangles and redo it[49] for skeleton 3-geometries built out of tetrahedrons. Also go from finite-dimensional or "truncated" superspace to the limit where (1) the skeletonization is infinitely finegrained, (2) the edge lengths are infinitely numerous, and (3) superspace rises in dimensionality from the purely illustrative number of 98 to the actual number of infinity.

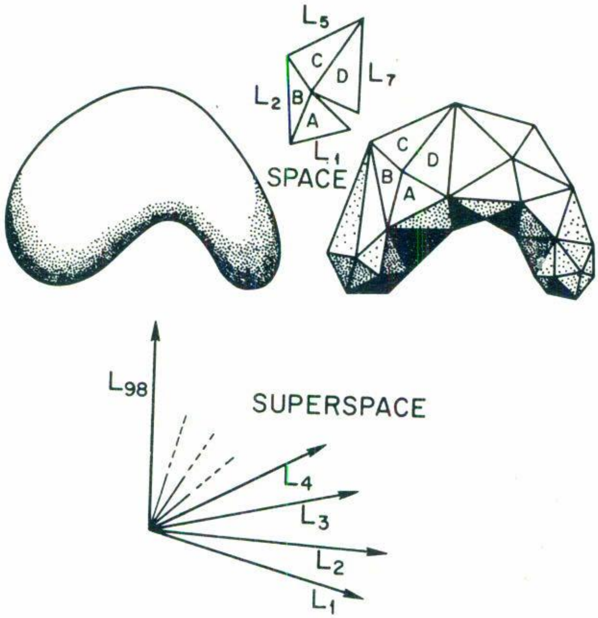


Fig. 9. A 2-geometry (upper left) is approximated by a skeleton 2-geometry (upper right). All the details of the shape of this skeleton 2-geometry are completely specified by giving (in this example) all 98 edge lengths, $L_1, L_2, \ldots, L_{98}$. This information is represented by a single point (lower diagram) in a 98-dimensional "truncated superspace".

Dynamics of the Universe as a Leaf of History in Superspace

A *leaf of history* cuts through superspace. It describes the deterministic dynamic development of the geometry of space with time. Fig. 10 illustrates how. At the right is spacetime, the usual deterministic classical picture of space evolving with time. Any spacelike slice through this spacetime, such as A, is a 3-geometry, a momentary configuration of space. It is represented in superspace by a single point, also denoted by A. Another slice B through the same spacetime provides another 3-geometry, and thus another point B in superspace. A one-parameter family of spacelike slices through spacetime thus "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. It is a perfectly

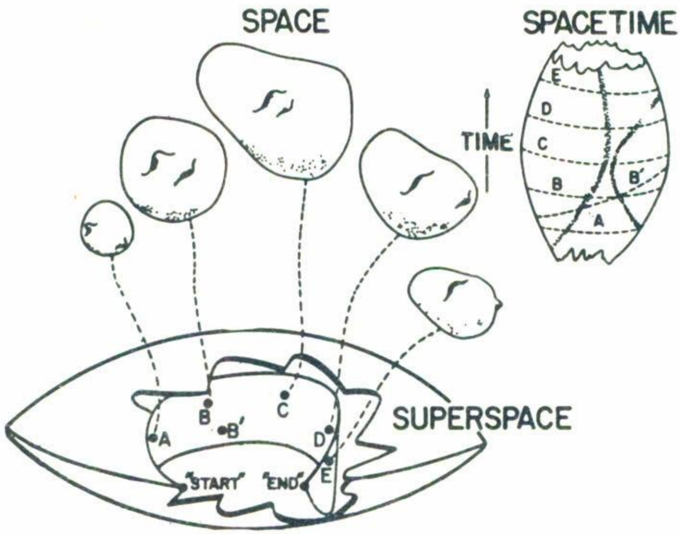


Fig. 10. Space (upper left), spacetime (upper right) and superspace (below). The "leaf of history" that curves through superspace includes all the configuration $(A, B, B' \ldots)$ achieved by space in its classical dynamical evolution in time; that is, all spacelike slices through the given spacetime. 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.

legitimate action for them to measure up the 3-geometry of the spacelike slice B'. This 3-geometry is a new point in superspace. No line in superspace can accommodate all the points, 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.

Given the spacetime, we have seen how we construct the leaf of history in superspace. Conversely, given the leaf of history in superspace, we obtain all the 3-geometries we need to reconstruct the spacetime. The procedure required, and used by Gerlach, but not spelled out here, reminds us in some ways of how we interlock together the disassembled wooden pieces of a Chinese-puzzle elephant to reconstitute the elephant.

New Features of Quantum Geometrodynamics

Quantum geometrodynamics differs drastically in principle from classical geometrodynamics. No longer is there the sharp yes-no difference between 3-geometries. The classical analysis clearly marked off the YES 3-geometries, that lie on a given leaf of history, from the NO 3-geometries, that do not. In the quantum analysis there is instead a probability amplitude $\psi({}^{(3)}\mathcal{G})$

for this, that and the other 3-geometry. The 3-geometries with appreciable probability amplitude are far more numerous than can be accommodated in any one spacetime. There are too many wooden pieces to be fitted into one elephant. The concept of a deterministic classical spacetime has to be abandoned.

The idea has been discussed for many years that quantum effects smear out the local light cone.[50] A much more drastic conclusion emerges out of quantum geometrodynamics and displays itself before our eyes in the machinery of superspace: ***there is no such thing as spacetime in the real world of quantum physics***. Spacetime is a classical concept. It is incompatible with the quantum principle. It has to be discarded in any deep-going analysis of the foundations of physics. It is an approximation idea, an extremely good approximation under most circumstances, but always only an approximation.

If we had a deterministic spacetime, we could take spacelike slices through it at two immediately succeeding instants, and thus find both a 3-geometry and a time rate of change of this 3-geometry. But complementarity forbids. It does not forbid our determining the 3-geometry alone an on initial spacelike hypersurface within arbitrarily narrow limits. However, the reciprocal uncertainty in the time rate of change of this 3-geometry is then arbitrarily great. This uncertainty deprives us of any possibility whatsoever to give any sharply defined meaning either to "spacetime" or to "the dynamical history of space".

In summary, superspace leaves us space but not spacetime and therefore not time. With time gone the very ideas of "before" and "after" also lose their meaning.

Quantum Fluctuations in the Geometry of Space

These quantum effects show up in significant measure only at small distances. There is a convenient name for them – "quantum fluctuations in the geometry". They have nothing directly to do with particle physics. They are a property of all space.[51]

Analogous quantum fluctuations in the electromagnetic field are also a property of all space. To analyze these fluctuations, to calculate their effect upon the motion of the electron in the hydrogen atom, and to observe the resulting shifts in the spectral lines of hydrogen, together constitute one of the greatest triumphs of physics since World War II.[52] Thus today it is fully confirmed that the quantum fluctuations of the electric field in a region of extension L are of the order of magnitude

$$\Delta\mathcal{E} \sim (\hbar c)^{\frac{1}{2}}/L^2 \, . \tag{33}$$

Apply the same kind of analysis to the gravitational field and equally directly conclude[53, 54] that the inescapable fluctuations in the metric (-1, 1, 1, 1) are of the order

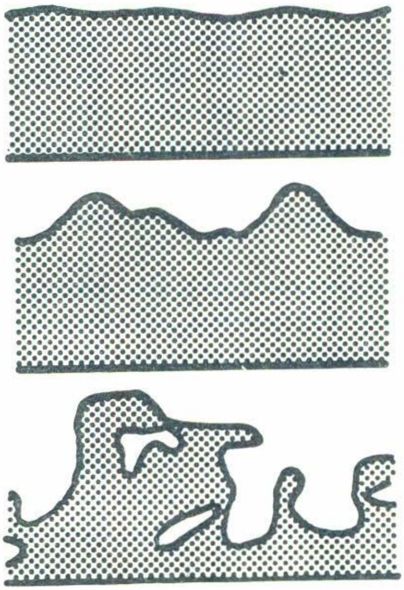


Fig. 11. Symbolic representation of the quantum fluctuations that take place everywhere and all the time in the geometry of space: above, as evidenced at a scale of observation, L, far larger than the Planck length, L^*; middle, L only a little larger than L^*; below, L comparable to L^*. Compare with the view of a stormy ocean as seen by an aviator flying miles above it, flying a hundred metres above it, and tossing in a lifeboat on the surface.

$$\Delta g \sim L^*/L \, . \tag{34}$$

Here

$$L^* = (\hbar G/c^3)^{\frac{1}{2}} = 1.6 \times 10^{-33}\,\text{cm} \tag{35}$$

is the Planck length.

These quantum fluctuations in the geometry of space are completely negligible at the scale of atoms and nuclei and elementary particles (L from 10^{-8} cm to 10^{-15} cm; Δg from 10^{-25} to 10^{-18}). In the domain of everyday physics space can be considered to be flat. Therefore, it is not surprising that no immediately measurable effect of the fluctuations, like the Lamb-Rutherford shift in hydrogen, has yet come to light. However, at smaller and smaller distances of observation the predicted fluctuations in the geometry become larger and larger, until at dimensions of the order of the Planck length one is open to believe that fluctuations take place even in the topology or connectivity itself (Fig. 11).

Quantum Fluctuations and Multiple Connectivity

Without any thought of quantum fluctuations, William Clifford[55], a century ago, had considered local changes in the connectivity of space as connected with the physics of particles. Again half a century ago, Hermann Weyl[56] pointed out that space, here and there, may be multiply connected in the small and, consequently, "The argument that the charge of an electron must be spread over a finite region, because otherwise it would possess infinite inertial mass, has thus lost its force. One cannot at all say, here is charge, but only, this closed surface encloses charge". The writer gave reasons[51] for the first time in 1957 out of fluctuation theory to consider "wormholes" a property, not of particles, but of all space, and all electric charge as "lines of electric force trapped in the topology of space". In the same year Charles Misner[57] showed the beautiful ties that connect Maxwell's theory in a multiply connected space with the mathematics of differential forms and homology groups.

Today, reconsidering electric charge, we can turn around the order of history in our imagination. Deny the existence in nature of any such thing as a mystic magic electric jelly. Rule out also any point singularity in any solution of Maxwell's equations. Agree with Einstein that once one admits the possibility of a singularity here, he has to admit it there, and therefore everywhere, and then he has destroyed the force of his field equation. Insist then that Maxwell's source free field equations hold everywhere without exception. Then electric charge becomes possible only if space is multiply connected. Therefore search nature for any evidence of electric charge. Find it – and conclude that space must, indeed, be multiply connected in the small. From this point of view, the existence of electric charge is the most compelling evidence we have today for Planck-scale fluctuations taking place in geometry and connectivity throughout all space. These are the fluctuations that say "No!" to spacetime and to time at small distances.

Gravitational Collapse Reexamined within the Framework of Superspace

If Hilbert's variational principle leads to superspace, and superspace leads to fluctuations and two decisive negatives, may not superspace also lead to an important positive? It furnishes an arena in which to take a fresh look at gravitational collapse, the greatest crisis in the theoretical physics of our times (Fig. 12).

COLLAPSE	MATTER (1911)	SPACE (1970's)
DYNAMIC SYSTEM	e^- and + CHARGE	GEOMETRY (HUBBLE)
CLASSICAL PREDICTION	∞ KINETIC ENERGY IN A FINITE TIME	∞ COMPACTION OF MATTER AND GEOMETRY IN A FINITE TIME
ONE REJECTED SOLUTION	GIVE UP COULOMB LAW (10^{-8}cm, 5×10^{-15}cm)	GIVE UP EINSTEIN'S EQUATION
ANOTHER ATTEMPT AT A "CHEAP WAY OUT"	ABANDON IDEA THAT ACCELERATED CHARGE RADIATES	ABANDON IDEA THAT MATTER CAN BE COMPACTED INDEFINITELY
PRINCIPLE OF CAUSALITY RULES THIS OUT	J. J. THOMPSON E. PURCELL IN BERKELEY PHYS.	$V^2_{SOUND} = \frac{dP}{d\rho} > C^2$
IMPLICATION OF PLANCKS QUANTUM PRINCIPLE	QUANTUM SPREAD IN SPACE $\Delta p \sim \hbar / \Delta x$	QUANTUM SPREAD IN SUPERSPACE

Fig. 12. Parallels between past and present crisis.

The electric collapse of matter, the great problem of the early 1910's, found its solution in the quantum principle. According to classical theory, the electron headed for the point centre of attraction arrived in a finite time at a condition of infinite kinetic energy. One had only to translate the classical Hamilton-Jacobi equation of motion of this particle to the Schrödinger wave equation to see deterministic collapse turned into probabilistic scattering (Fig. 13).

A classical leaf of history shows the universe expanding, reaching a maximum volume, and finally collapsing in a finite proper time to a state of infinite compaction. Turn from classical determinism to a probability wave propagating in superspace. Can this wave not also undergo scattering at the point in superspace where otherwise collapse would have been expected? And if the electron scattered by the nucleus goes off on a quite new worldline, cannot the wave scattered in superspace go off on a quite new leaf of history?

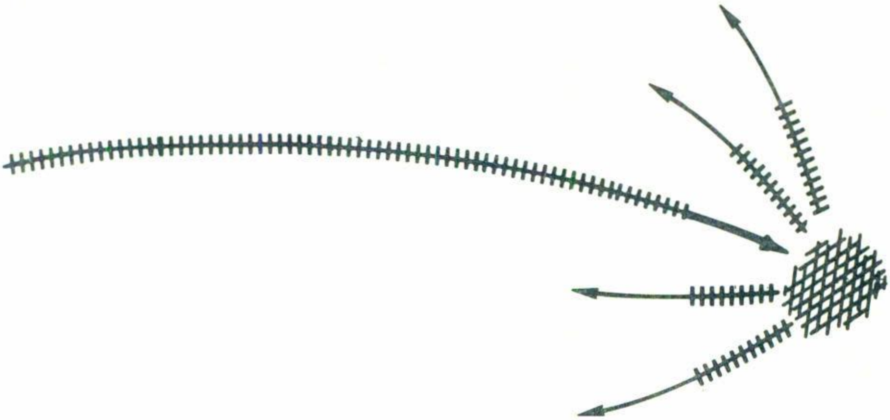


Fig. 13. Not deterministic collapse (in the cross-hatched "zone of collapse") but probabilistic scattering, is the outcome of the encounter (motion in 3-space) of the negatively charged electron with the positively charged centre of attraction; and is also on outcome natural to consider seriously for the gravitational collapse of the universe itself (motion of representative point in superspace described, not by a deterministic leaf of history, but by a probability wave; and this wave undergoing scattering in superspace; not a deterministic new cycle of the universe, but a "probability distribution" of new cycles of the universe).

In What Sense Do Other Leaves of History "Coexist" with Our Own?

We have only to ask questions such as these to find ourselves facing a still deeper question. With two or more quite different leaves of history located in one and the same superspace, what strange kind of "coexistence" of two universes are we confronting? It is not absolute nonsense to speak of another universe coexisting with our own, no matter in how attenuated and ethereal a way we use the word "coexist"? Almost a century ago Auguste Comte[58] also decried as absolute nonsense the idea of attributing a chemical composition to a distant star. It may have a sense to speak of the chemical composition of the Sun, he was willing to admit; but certainly not the composition of a star to which there is not the slightest possibility of anyone ever travelling. Of course, in the meantime, half a dozen ways have been found to get at the composition of a star, and many a satisfactory check has been made of one method against another. No one would think of dispensing with this concept.

There is also not the slightest possibility to travel to another leaf of history. Gravitational collapse places an impenetrable barrier between one leaf and another. Life cannot get through. Even such ideas as "before" and "after" lose their relevance in the final state of collapse, thus altogether forbidding any direct comparison of time between one leaf of history and another.

Consider more closely this question of "coexistence" of alternative histories of the universe. Quantum spread moves the representative point that describes the universe little way off one classical leaf in history in superspace. A larger movement takes it to another classical leaf of history. There is no difference of principle between the two. There is only a difference of degree. No one can deny the "coexistence" of alternative histories of the universe who accepts the existence of quantum fluctuations in the geometry of space.

One has only to recall the famous double slit electron interference experiment to see the same principle in a simpler context. The "coexistence of two histories" of the electron is the very heart of the observed interference. No one has ever successfully contested it.

"Scattering in Superspace" as the Final Phase of Collapse

Between "fluctuations" and "scattering" there is a difference only in degree, not in kind. Do then the predicted final stages of collapse of the universe lead, not to the deterministic catastrophe of classical theory, but to a probabilistic scattering in superspace? If the electron, moving faster and faster towards the disaster, experiences scattering, not catastrophe, does the universe do the same? The arena for the dynamics of the electron is Minkowski spacetime; the arena for the dynamics of geometry is superspace; but is there otherwise any reason why scattering into a new history is not as truly the outcome in the one case as in the other?

Why collapse may not be final, why it may be followed by a new history or, rather, by a probability distribution of new histories, when the dimensions of the universe get down to a value governed by the Planck length, may be put in still other words. Already here and now, according to quantum geometrodynamics, violent fluctuations are going on in geometry as viewed at the Planck scale of distances. On such a worm's eye view a fluctuation is hardly distinguishable from the collapse of the universe itself. In effect, gravitational collapse of the "local universe" is *already* over and over taking place and being undone. Moreover, this doing and undoing of collapse is going on everywhere in space and all the time without catastrophe. So why anticipate catastrophe from the collapse of the universe itself?

If one can foresee along these lines the answer to the paradox of collapse, why not work it out and demonstrate it by calculation? In the problem of the electron one goes easily from the classical Hamilton–Jacobi equation to the Schrödinger wave equation and from that to the Rutherford law for the probability distribution of scattering angles. Why not proceed similarly here, where one already has the Hamilton–Jacobi equation?

First, there are unsolved problems of factor ordering in translating the *H-J* equation into a Schrödinger equation. Second, both equations presume classical differential geometry. Classical differential geometry leaves no room for changes in topology. Yet it is an inescapable characteristic of quantum field theory that, in the phrase of John Klauder,[59] unruly configurations predominate. From unruly configurations of 3-geometry like those symbolized in Fig. 11 it is a small step in the imagination to go to a doubly connected 3-geometry, as would also seem to be required by the existence of electric charges. But classical differential geometry says "No" to this step. If that mathematics applied to nuclear matter, it would also say "No" to nuclear fission, with *its* change in connectivity. But the nucleus elongates and divides, all prohibitions of differential geometry notwithstanding. For the description of this change in connectivity today's nuclear physics has the right mathematical machinery. Today's geometrodynamics does not.

Lastly, no one can believe any purportedly quantitative treatment of the final stages of collapse as "scattering in superspace" that assigns no role to Fermion fields and particles, has no explanation[60] for their spin, and pays no heed to their fate.

Scattering and Superspace as Waystations on the Road to Deeper Views

"Scattering in superspace" contains two concepts. One is "scattering" as the terminal phase of collapse. The other is "superspace" as the arena for that collapse. Both concepts, it is possible to believe, are way stations, useful way stations, but nevertheless only way stations, on the road to still deeper penetration. Therefore take a second look: at "scattering", later; at superspace, now.

Superspace is a point of farthest advance in the understanding of relativity. In no arena does the dynamics of geometry express itself more compactly. From no vantage point do collapse and quantum fluctuations in metric appear more clearly as two aspects of the same geometrodynamics. Superspace is here to stay.

The mathematics of superspace nevertheless seems at first sight in two ways too frozen to expose to view any still deeper level of physics. (1) The dynamic law, the Einstein–Hamilton–Jacobi law[32] for the propagation of wavecrests in superspace, looks as if handed down from on high and beyond further explanation. Riemann[61] fought to make geometry a part of physics. How could he have counted it a victory to see the God-given geometry of Euclid replaced by the God-given geometry of superspace?[62] To have superspace instead of space is no advance towards the explanation of space. (2) The representative point in superspace is a 3-geometry. Three-geometry appears as the one and only dynamic entity: What about the rest of physics? Has one locked himself, unawares, into the view that particles and fields are all derivative, somehow constructed from geometry as from a "magic building material"? Has one adopted the Clifford–Einstein "space theory of matter"? Not at all.

Only to minimize detail has one limited attention to pure geometry: to gravitational waves, geons made out of gravitational waves, and black holes made by collapse of such geons, all in a universe curved up into closure by its content of black holes, geons and gravitational waves. How then does one give an account of electromagnetic fields and effects? One augments the variables that appear in the state functional, ψ, from ${}^3G$ to ${}^3G$ plus B, where B is a divergence-free magnetic field defined everywhere thoughout the manifold ${}^3G$. Similarly for other fields: the field coordinate, or the field momentum, but not both, grace to complementarity, also appears in ψ: or, in the semiclassical approximation, appears in the Hamilton–Jacobi functional S. Accordingly "augmented superspace", the configuration space of the dynamics, contains additional and non-geometrical coordinates.

Deeper questions do not arise. Are electromagnetism and particle fields a manifestation of pure geometry? Or is geometry a mere bookkeeping for relations between particles? Or are particles and geometry both primordial? Or are they both derived from something more primordial than either, call it pregeometry or call it what one will?

No immediate help does one get from the previous four derivations of relativity (Einstein, Cartan, compact, Hilbert) in penetrating deeper into such questions, either to understand why superspace has the special Hamilton-Jacobi structure (32), or to suggest what particles have to do with geometry. Guidance into these issues comes first from the final two derivations of relativity. Number five has for key idea that "dynamically changing space must be imbeddable in spacetime". Number six, epitomized, says "space acquires its resistance to curvature from the curvature-dependence of the zero-point energy of particles and fields". In penetrating to a stratum of ideas deeper than those encountered in previous derivations, these approaches begin to recognize that geometry may be a derivative rather than a primordial concept.

Rejection of the View that Space is the Primordial Dynamic Entity

The directly contrary vision, going back to Clifford and Einstein, that geometry is the primordial entity, and everything else is derived or constructed from geometry, deserves its assessment before one turns to these final two derivations of general relativity.

As early as February 21, 1870, in a paper before the Cambridge Philosophical Society *On the Space Theory of Matter*[63] W.K. Clifford (1845–1879; Clifford algebras), inspired by the 1854 lecture of B. Riemann (1826–1866), had proposed that a particle is a "hill" built out of the geometry of space rather than a foreign and physical object immersed in the geometry of space. Einstein himself was animated by the vision of a purely geometrical account of physics. Many a worker since who has occupied himself at all with general relativity has found himself little by little caught up in the same Clifford-Einstein vision. In such cases it is not rare to arrive at a little new understanding of Einstein's general relativity, a great appreciation of the crisis of gravitational collapse, and also, in the end, the conviction that the quantum principle is even more fundamental than geometrodynamics to the make-up of physics and the elucidation of collapse. A sample case history, for one of the many workers in the field, will illustrate this course of evolution of ideas:

1953: Accept gravitational collapse as a central issue. Simplify equation of state of the collapsing object by taking radiation alone as the source of its mass-energy.

1954: Insist this radiation shall travel perpendicular, or nearly perpendicular to r. Arrive at a "geon". It holds itself together by its own gravitational attraction for a time long in comparison to periods of individual quanta. Attracts as a mass, moves as a mass, but nowhere contains any "real" mass. Model for "mass without mass". A classical object. No direct relation whatsoever to a particle.

1955: "Charge without charge": electricity as lines of force trapped in a multiply connected space. Existence of charge in nature taken as evidence that space in the small is multiply connected. "Electromagnetism without electromagnetism". 2nd order Maxwell equations and 2nd order Einstein equations put together in 4th order Rainich equations dealing with geometry and nothing but geometry.

1956: A particle —that looks impressive — is as unimportant relative to the quantum fluctuation physics of the vacuum as a cloud —that looks impressive— is unimportant to the physics of the sky. Particle physics is not the right starting point for dealing with particle physics. Vacuum physics is. Space, owing to quantum fluctuations in geometry and connectivity at small distances, of necessity has a "foamlike structure".

1964: Superspace: allows one to see the structure of geometrodynamics at a glance, and see collapse and quantum fluctuations of geometry within the same dynamic framework.

1968: When an orientable 3-geometry is multiply connected, (n handles or "wormholes") superspace has 2^n sheets. Each sheet corresponds to a topologically distinct continuous field of triads that can be laid down on the 3-geometry. There are 2^n distinct probability amplitudes associated with the same 3-geometry, or, per wormhole, one "non-classical two-valuedness" or spinor degree of freedom ("spin without spin"). Question raised, can a particle be regarded as a "geometrodynamical exciton"? And can neutrino fields, pion fields, hyperon fields and other fields likewise be interpreted in terms of "modes of excitation" of multiply connected geometry?

1969: Continuing analysis of black hole physics.

1970: Outcome of gravitational collapse of the universe itself discussed in terms of "scattering in superspace".

1971: No dynamics of topological spin and no quantum fluctuations in topology —and therefore, one can believe, no proper treatment of collapse as scattering in superspace— without change in connectivity; but no continuous change in connectivity is allowed by differential geometry. Differential geometry presupposes a concept of "point neighborhood" that cannot be a correct description of the physics at small distances. The thinning and breaking of a handle makes points that were near suddenly become far. Conversely, far away points have a potentiality for becoming immediately adjacent that is incompatible with the ideas of differential geometry. Even the concept of dimensionality cannot be applicable at small distances. With the failure of differential geometry, general relativity also

> fails; it cannot provide anything more than a crude approximation to what goes on at the smallest distances. Geometry "is not crazy enough" to describe all of physics. But particle physics also does not provide any "magic building material". No account of particles that deals only with particles will ever explain particles. There must exist an entity ("pregeometry") more primordial than either geometry or particles on the foundation of which both are built. The nature of pregeometry will first become cleat when one sees the quantum principle in all completeness, not as something strange and foreign imposed on the world, but as the central principle without which the world could not even come into being.

Out of a case history such as this, and many another, each with its pluses and minuses, what is the conclusion?

To those who have labored in the garden of geometrodynamics, or watched its development, it has been a reward to see the blossoming of neutron-star astrophysics and the budding of black-hole astrophysics. It has been a satisfaction to observe that new dimension come to life that Einstein's theory gives to all of physics-*geometry*, from tidal acceleration as Riemannian curvature to superspace as the arena for geometrodynamics. It has been tantalizing that electricity lets itself be interpreted as lines of force trapped in a multi-wormhole geometry, with one spin $\frac{1}{2}$ tied to each wormhole. It has been both a disappointment and an inspiration to sense at last that one must look beyond geometry for the understanding of geometry— and of collapse.

The view that "everything is geometry" has shown itself in the end a view "too finalistic to be final". The very surprise of the predictions of general relativity (expansion of the universe predicted, and predicted correctly, and predicted against all expectation; gravitational collapse; black hole) and the scope of its explanations (gravitation as a manifestation of geometry; conservation of mass-energy guaranteed by the principle that the boundary of a boundary is zero) created a new standard for the surprise of a prediction and for the scope of an explanation. The standard has meantime risen, not least because of the beautiful regularities uncovered in particle physics. General relativity has not kept up with the rise.

The "Surface Geology" and "Underground Geology" of the Vacuum.

The student who first takes up geology finds no feature of the landscape more interesting than its topography. Later he sees cores drilled out from widely separated locations with identical strata. He comes to think of the stratum as the primary concept. Finally he begins to appreciate that underground strata and surface topography are manifestations of one and the same dynamic geology.

To the student who first learns about relativity no feature of the vacuum attests more clearly its power to take part in physics than its curvature. Then he sees energy slammed into the vacuum here, and discovers particles spray out with a characteristic spectrum of masses. He observes energy poured in at a remote point and finds that the same spectrum emerges from the vacuum. Particles seem to look like the central feature of vacuum physics. But further study makes him believe that both the geometry, "the surface geology of space", and the particles, "the underground strata" are manifestations of a something more primordial than either. This is the point of view we adopt in looking into the structure of relativity through the last two windows.

6. THE STRUCTURE OF SPACETIME DERIVED FROM THE "GROUP" OF DEFORMATIONS

General Relativity as Representation of the "Group" of Deformations

The many-fingered time of Einstein's general relativity is a concept so simple that its sophistication does not immediately surface: its central presupposition that *space is imbedded in spacetime*. Let a band of observers explore the dynamics of geometry and other fields. Like a line of soldiers, they can advance faster on one front, slower on another, and later push ahead more rapidly in the second region, slower in the first, until they come to the preassigned "river line", or spacelike hypersurface. What they find there must be the same whether the moving hypersurface surged ahead first on the left or on the right. The change in the physics from the initial simultaneity to the final simultaneity must be independent of the choice of simultaneities in between. In mathematical terms, the dynamics must provide a representation of the "group" of deformations of a spacelike hypersurface. This requirement, Hojman, Kuchař and Teitelboim show[64], fixes the Hamiltonian of general relativity as of the form (32), up to an arbitrary canonical transformation, and up to the permitted addition of the cosmological term that Einstein first introduced and later rejected.

If the structure of Euclidean geometry ever seemed arbitrary, its general-relativity substitute, the law (32) of propagation of wave fronts in superspace, must have appeared as still more arbitrary; but it is not, one now sees. Superspace turns out to follow the only law that one can easily imagine, a law so simple in its principle that anything simpler could hardly be a law.

Were the Hamiltonian different, one would still have the geometrodynamical field coordinates, $g_{ij} = g_{ij}(x, y, z)$, and the geometrodynamical field momenta,

$$\pi^{ij} = \delta S/\delta g_{ij} \; ; \tag{36}$$

and an acceptable "initial value set" of these $6+6 = 12$ functions of position would still determine an entire leaf of history in superspace. However, the 3-geometries making up this leaf of history would no longer fit into any one spacetime.

The band of observers would still have the freedom to push forward "many-fingered time" with all the individual free choice that that term implies. However, these time increments would no longer let themselves be described as increments of a time coordinate in any manifold that in any way whatsoever constituted a spacetime.

Demand Hamiltonian theory in superspace and demand that that Hamiltonian theory shall yield spacetime, and automatically end up with the Einstein-Hamilton-Jacobi equation —that is the beautiful route to general relativity opened up by Hojman, Kuchař and Teitelboim. When a vector field is added, electromagnetism also emerges. When other fields are included, their dynamics similarly comes out of the condition of imbeddability.

Relativity Compared to Elasticity

The very austerity of "relativity out of imbeddability" shows how little of a fundamental nature goes into the derivation of Einstein's law of gravity, and how little of the inner working of physics one really can read out of relativity. One is led to compare relativity with elasticity. The elastic energy-per-unit-volume of the small deformation $x^i \to x^i + \xi^i$, of a homogeneous isotropic solid, expressed in terms of the strain tensor θ with components

$$e_{ij} = \tfrac{1}{2}\left(\frac{\partial \xi^i}{\partial x^j} + \frac{\partial \xi^j}{\partial x^i}\right) \tag{37}$$

is

$$c_1(\mathrm{Tr}\ \mathbf{e})^2 + c_2\, \mathrm{Tr}\ \mathbf{e}^2 \tag{38}$$

according to reasoning based upon considerations of symmetry and group theory alone. The binding of a mixed solid arises from bonds between a multitude of different atoms. Each bond has its own potential energy curve and resistance to bending. However, only the sums of the second derivatives of these many potentials appear in the elastic constants c_1 and c_2. Not one hint do these two totals give about the size of the individual atoms, the composition of the solid, or the origin of a potential energy curve.

In general relativity there appears only the one constant, the Newtonian constant of gravity G. The existence of such a constant again follows, as shown by Hojman, Kuchař and Teitelboim, from group theory alone (the "group" of deformations of a spacelike hypersurface). Nevertheless the origin and nature of any individual contributions to G are again totally concealed from view.

7. SAKHAROV'S DERIVATION: GRAVITATION AS THE "METRIC ELASTICITY OF SPACE"

Nothing forces the student of elasticity to rely on measurement alone for values of the elastic constants of the solid. He can evaluate them from spectroscopic or calculated or estimated values of the stiffness parameters of the individual bonds. Sakharov[65] (see also Zel'dovich and Novikov[66]) similarly proposes to view the gravitation constant (1) as measuring the "metric elasticity of space", and (2) as given by the sum of individual contributions, each of which in principle can be estimated. On this view space is like an empty sausage skin, which is "floppy" and deprived of all resistance to bending until it has been filled with sausage meat. The "sausage meat" is the zero-point energy of particles and fields.

In undeformed space the electromagnetic field, as an example, has per unit of volume a zero-point energy that is obtained by integrating the product of the following factors:

Number of independent modes in interval of circular wave numbers from k to $k + dk$,	$4\pi^2 dk/(2\pi)^3$
Number of states of polarization per mode,	2
Zero point energy per mode,	$\hbar ck/2$
Product,	$\hbar ck^3 dk/2\pi^2$

(39)

The result diverges. It has to be renormalized to zero to be compatible with experience. The result is similar for other fields. Moreover, the result is qualitatively the same whether one deals with energy (one component of a 4-vector) or the Lagrangian (invariant density). However, when space is curved, correction terms arise in the renomalized invariant Lagrangian density for each field proportional to the 4-dimensional Riemann scalar curvature invariant:

$$\delta L_{\text{one field}} \sim \hbar c\,{}^{(4)}R \int k\,dk \tag{40}$$

(see Berger, Gauduchon and Mazet[67] for more on the effect of curvature on the spectrum of standing waves). This integral is still divergent. Sakharov reasons that there is a highest circular wave number, $k = k_{\text{crit}}$, for which the calculation makes sense. Here k_{crit} for all fields alike, he proposes, is of the order of the reciprocal of the Planck length,

$$k_{\text{crit}} \sim (\hbar G/c^3)^{-\frac{1}{2}} \sim 10^{33}\ \text{cm}^{-1}\ . \tag{41}$$

It follows that the contribution to the Lagrangian of the vacuum from the curvature-dependent zero-point energies of all fields together has the same form and order of magnitude as the Lagrangian of Einstein's theory of gravity;

$$L_{\text{grav}} = (c^4/16\pi G)\,{}^{(4)}R \sim \hbar c\, k_{\text{crit}}^2\,{}^{(4)}R \sim \sum_{\text{fields}} \delta L_{\text{one field}} \tag{42}$$

This is the sense in which Sakharov considers gravitation to be the metric elasticity of the vacuum.

The constant of gravitation as estimated in this way can be given almost any value one chooses by appropriate choice of the cutoff wave number k_{crit}. Sakharov tailors k_{crit} to give the known value of G. No one sees how to get k_{crit} from first principles. Nevertheless, Sakharov reasons, the proper order of ideas is not, first gravitation and then fields and particles, but first fields and particles and then gravitation, as a derivative effect.

From Sakharov's "particle first" point of view, gravitation is as much derivative from particle physics as elasticity is derivative from molecular physics. If one accepts his point of view one does wrong to try to build particles out of geometry. One does wrong whether one speaks of 1870 Clifford "hills" in space or 1970 "geometrodynamic excitons". One might as well try to build atoms out of elasticity! Atoms come first, and only then elasticity; particles first, and only then geometrodynamics.

8. MUTABILITY AND BEYOND

Pregeometry as More Primordial than Either Particles or Geometry.

The last two derivations of relativity, different though they are, suggest that gravitation is as far removed as elasticity from being primordial. But does that mean that particles are primordial? Hardly. The derivative character of elasticity by no means implies that atoms are the primitive entities. On the contrary, it was the first and smallest advance in the study of solids to understand the two elastic constant terms of scores of molecular potential energy curves, many of them not known in any detail. Only when those scores of interactions found explanation in terms of a system of electrons and positively charged nuclei and Schrödinger's equation and nothing more did the decisive advance in understanding come. Likewise it may be only the first step forward to interpret the one "constant of gravitation as the sum of the coefficients of curvature dependency of the vacuum energies of all the fields and particles of physics". Yet to come would seem a second and far larger step: to see both geometry and all these fields and particles as manifestations of something more basic ("pregeometry") than any of them.

Constants and Dynamic Law not as Immutable but as "Frozen in" in the First Stage of the Big Bang

What difference does it make if geometry and fields and particles are built up from something more primordial? Does not one then have to ask, when were conditions ever intense enough to form, and when were these conditions ever released fast enough to freeze, this structure of geometry and fields and particles into its present set of laws? When else than in the "big bang"?

This piece of wood, this solid, is a "fossil" from a photochemical reaction in a tree twenty years ago at a few hundred degrees Kelvin. One has only to subject it to higher temperatures to alter drastically its molecular constitution and switch it over to a new "fossil".

The iron nuclei in this steel pen nib are "fossils" from a thermonuclear reaction in a star some billions of years ago at a temperature of some tens of millions degrees Kelvin. One has only to put these nuclei back into a star where conditions are sufficiently intense to transmute them into still heavier nuclei, which upon removal, rate as new "fossils".

Can particles themselves (and fields and geometry) be anything but "fossils" from the most violent conditions of all, those encountered in the very earliest phase of the "big bang", that mirror of gravitational collapse?

That there was a big bang (see for example the review of Peebles[68]) is evidenced not least by the recession of the galaxies, the proportion between primordial helium and hydrogen, and the primordial cosmic fireball radiation. The inevitability of gravitational collapse of every closed model universe, no matter how irregular, is by now as well established prediction of standard relativity.[69, 70, 71] Both at big bang and at collapse, calculated temperatures and pressures rise without limit. Between these times of conditions unprecedented in their extremity, physics is fossilized. No change with time has ever been found in the fine structure constant (see the impressive evidence adduced by Dyson[72]), in the mass of any particle, or in any other constant of physics.

One used to think of someday finding a "theory" of the fine structure constant, of the basic constants of particle physics and of the "big number scale",

$$\begin{aligned}
&\begin{bmatrix}\text{number of photons per}\\ \text{baryon in the universe}\end{bmatrix} && \sim 10^{10}\\
&\begin{bmatrix}\text{particle dimensions, } 10^{-13}\text{ cm,}\\ \text{relative to Planck length}\end{bmatrix} && \sim 10^{20}\\
&\begin{bmatrix}\text{estimated radius of universe}\\ \text{at full tide relative to}\\ \text{nuclear dimensions}\end{bmatrix} && \sim 10^{40} \qquad (42)\\
&\begin{bmatrix}\text{electric force between}\\ \text{two particles relative to}\\ \text{gravitational force}\end{bmatrix} && \sim 10^{40}\\
&\begin{bmatrix}\text{estimated number of}\\ \text{baryons in universe}\end{bmatrix} && \sim 10^{80}
\end{aligned}$$

Today, forty years later, such a dream is as far from realization as ever. One is open to believe that one has been looking for the right answer to the wrong question. A century and a half ago Laplace dramatized the difference between initial conditions and dynamic law. The intervening decades have seen new laws uncovered, but not a single discovery about what fixes the initial conditions. The time has come to ask if the constants and the scale of the big numbers belong in the realm of law at all. Are they not more reasonably to be understood as initial conditions?

Mutability as Central Feature of Physics

"Constants" and laws alike "frozen in" at the very earliest stage of the big bang, and rubbed out in the very last stage of gravitational collapse: that is the picture that one is led to examine seriously. On this picture physics is a staircase. Each tread registers a law (e.g., law of chemical valence). Each riser marks the transcendence of that law (e.g., temperatures and pressures so high that valence loses its significance). The staircase climbs from step to step: density, and density found alterable; valence law, and valence law melted away; conservation of net baryon and net lepton number, and these conservation laws transcended; conservation of energy and angular momentum, and these laws likewise overstepped; and then the top tread displaying all the key constants and basic dynamic laws — but above a final riser leading upward into nothingness. It bears a message: With the collapse of the universe, the framework falls down for every law of physics. There is no dynamic principle that does not require space and time for its formulation; but space and time collapse; and with their collapse every known dynamic principle collapses.

If the laws of conservation of particle number are transcended in black hole physics; if all dynamic laws are transcended in the collapse of the universe; if laws and constants of physics are first imprinted as initial conditions in the earliest phase of the big bang and erased in the final stage of gravitational collapse, then dimensionality itself can hardly be exempt from the universal mutability.

The review one by one of fixed points of physics has left not a single one unquestioned, neither "constant" nor principle. It is difficult to find any other way to summarize the situation as it now appears than this: "There is no law except the law that there is no law;" or more briefly, "Ultimate MUTABILITY is the central feature of physics".

Beyond Mutability

One is led to think of a universe more ephemeral than would be admitted by any "bootstrap particle model", or any model based upon a "fundamental field", or any model that considers geometry to be the "magic building material" of existence. Only by giving up almost everything, it would seem, can one be truly responsive to the imperatives of collapse.

In all the marvellous history of physics nothing stands out more impressively than the step-by-step transcendence of categories. "Green" was adequate as description of the color of a mineral, but "green" disappeared when one came to the motion of the electron around the nucleus. The planetary circle of Copernicus faded from view before the differential equation of Newton and Euler. Gravitation disappeared and geometry took its place. The classical orbit made its exit when the wave of de Broglie and Schrödinger made its entrance. The fantastic wealth of chemical fact and force boiled down to electrons and nuclei and Schrödinger's equation. Each complication of the evidence was not matched by a corresponding complication of principle. The more one gave up the more one gained; and the more one gained the more one gave up.

Dynamic Laws Transcended

If mutability demands the giving up of almost everything, what goes, what comes, and what stays?

Superspace is the quintessence of relativity; and in the context of this arena one has been led to think of the outcome of gravitational collapse as "probabilistic scattering in superspace". On this view collapse is followed, not by a unique new cycle of big bang, expansion, recontraction and collapse, but by a probability distribution of such histories, each (because of transcendence of conservation laws) with its own new number of particles and own new time from big bang to collapse. This picture now appears inadequate because it presumes, not too much to change, but too little. When one began to consider particle number and particle masses and the dimensionless constants of physics as altering from one cycle of the universe to the next, one also started to view the dynamic laws themselves as like the laws of valence, wiped out by conditions sufficiently extreme, and therefore extinguished in collapse. One had already found it impossible to calculate his way through the quantum mechanics of collapse within the context of superspace. Also one had already realized that the superspace of general relativity is an incomplete arena. But to count relativity as wiped out in collapse is to destroy superspace, and therefore take away the foundations for any picture of collapse as "scattering in superspace".

That "scattering" is way station to a larger picture, in which all the constants and dynamic laws get established only in the first stage of the big bang itself. Thus superspace goes and law goes. What comes?

Chaos Accepted, and Law Built on Chaos

If law goes, what can take its place but chaos? Chaos is not new for physics to encounter. Physics has mastered chaos before and translated it into the order of law. One can solve the two-body problem easily and the three- and four-body problem with greater and greater difficulty; but the N-body problem, with $N \geqslant 5$, is intractable. Nevertheless, when N grows and grows, the curtain rises to reveal temperature and entropy, new concepts unimagined and unimaginable at an earlier phase of physics. Moreover, the molecular chaos underneath in no way deprives the resulting laws of thermodynamics of the most impressive precision.

A "pregeometry" that is primordial chaos, and law built upon this chaos: that is the vision of physics that we are led to examine.

How is one to find the key element of this underlying chaos or "pregeometry"?

Nothing did one learn from a hundred years of elasticity about chemical forces; and a hundred years of chemistry unfolding all its wonderful regularities, provided not one clue to Schrödinger's equation. The order of understanding ran, not down, but up. One had to have quantum mechanics to understand chemical forces; and one had to know chemical forces to understand elasticity. Likewise a half century of gravitation –as– geometry has revealed nothing of the constitution of particles; and a half century of particle physics, laying open so many beautiful symmetries, has given not one hint of what lies beneath. Not down, but up; not down from particle physics or geometrodynamics, but up from the quantum principle would appear the right route to the primordial element, the "pregeometry" that we visualize as chaos.

The Quantum Principle as the Only Principle

With law going and chaos arriving, one principle remains, the quantum principle. With all other laws of physics rated as mutable, it is the only principle. If no one ignorant of evolution has the first idea about the origin lf life, it is also true that no one who is unacquainted with the quantum principle has the first idea how nature works. Physics without the quantum is medieval physics.

The quantum principle might almost be called the Merlin principle. Merlin the magician, on being pursued, changed first to a fox, then a rabbit, then a bird fluttering on one's shoulder. The quantum concept underwent still more spectacular changes in outward appearance: Mendeléev's[73] "individuality amid uniformity"; Planck's law for the energy of an oscillator; the law of Rutherford and Einstein for radioactive decay and atomic transitions; Bohr's quantization of angular momentum; the non-commuting observables of Heisenberg and Dirac; Heisenberg's uncertainty principle; Bohr's principle of complementarity; Feynman's principle of the democracy of all histories; Everett's "many-universes" formulation[74, 75], and the lattice of propositions of von Neumann and Birkhoff.[76]

Nothing is more surprising about quantum mechanics than this, that it still comes to us as a surprise. We have not yet discovered the most central consideration of all, the consideration that would tell us that the universe could not even have come into being had there been no quantum principle. We have no answer to the great Leibniz, "Why is there something rather than nothing?"

Nothing is more important about the quantum principle than this, that it destroys the concept of the world as "sitting out there", with the observer safely separated from it by a 20 centimeter slab of plate glass. Even to observe so miniscule an object as an electron, he must shatter the glass. He must reach in. He must install his chosen measuring equipment. It is up to him to decide whether he shall measure position or momentum. To install the equipment to measure the one prevents and excludes his installing the equipment to measure the other. Moreover, the measurement changes the state of the electron. The universe will never afterwards be the same. To describe what has happened, one has to cross out that old word "observer" and put in its place the new word "participator". In some strange sense the universe is a participatory universe.

Is this instance of participation the tiny tip of a giant iceberg? Molecular chaos leads to concepts like temperature and entropy only when limitations are imposed, such as fixity of volume and total energy. Otherwise chaos is chaos. Does the chaos, the "pregeometry", that we think of as underlying the universe, also fail to yield any law until it is analogously limited? Do we ourselves supply this limitation, we who have been forcibly elevated from observers to participators? Are we, in the words of Thomas Mann[77] "actually bringing about what seems to be happening"? Are we destined to return to the deep conception of Parmenides[78], precursor of Socrates and Plato, that, "what is, . . . , is identical with the thought that recognizes it"?

Leibniz[79] reassures us that, "Although the whole of this life were said to be nothing but a dream and the physical world nothing but a phantasm, I should call this dream or phantasm real enough if, using reason well, we were never deceived by it". Never was the call "use reason well" more timely than today. Collapse and mutability make unprecedented demands on imagination and judgment. Now more than ever one is certain that no approach to physics that deals only with physics will ever explain physics.

No proud tower of human thought can remain unshaken by the greatest crisis one can name in the history of science: neither mathematics nor logic, neither philosophy nor physics. The budget officer may be able to parcel out money neatly to those areas of thought; but "the good Lord" did not appreciate these fine distinctions and mixed them all up in the founding of the world. It will take the power of all of thought together if we are ever to understand why we have "something rather than nothing". We can believe that we will first understand how simple the universe is when we recognize how strange it is.

NOTES AND REFERENCES

1. T.S. Eliot, *Little Gidding*, (Faber and Faber, London, 1942).
2. E.P. Wigner, July 14, 1963 lecture, reprinted in *Symmetries and reflections: scientific essays of Eugene P. Wigner*, (Indiana University Press, Bloomington, Indiana, 1967) p. 32.
3. C.W. Misner, K.S. Thorne and J. A. Wheeler, *Gravitation*, (Freeman, San Francisco, 1973).
4. A. Einstein as quoted in W. Sullivan, *The Einstein Papers,* New York Times, March 28, 1972, p. 1.
5. A. Einstein letter to E. Mach of June 25, 1913, Einstein archives, Institute for Advanced Study, Princeton, New Jersey, courtesy of the Einstein estate.
6. A. Einstein, *Essays in science*, (Philosophical Library, New York, 1934) p. 68.
7. A. Einstein, *Preuss. Akad. der Wissenschaften*, Berlin, *Sitzungsberichte*, pp. 778-786 (session of Nov. 4, 1915, published Nov. 11); pp. 799-801 (session of Nov. 11, published Nov. 18); and (the field equation in its final form) pp. 844-847 (session of Nov. 25, published Dec. 2).
8. See for example reference 3, Fig. 29.2.
9. A. Einstein, *The meaning of relativity*, (Princeton University Press, Princeton, N.J., 3rd ed., 1950) p. 107.

10. A. Einstein, *Essays in science*, (Philosophical Library, New York, 1934) p. 52. Einstein acknowledged a debt of parentage for his theory to Mach's principle. It is therefore only justice that Mach's principle should in turn owe its elucidation to Einstein's theory, in the sense that (a) the "inertial properties of matter" are today understood in terms of, and as equivalent to, the straight lines of the local tangent space (local Lorentz geometry) of the curved spacetime manifold; (b) the 4-geometry of this manifold (and therefore these inertial properties of matter everywhere and at all times) is uniquely determined by appropriate arbitrarily chosen initial value data on an initial spacelike hypersurface of this manifold (for greatest convenience taken to be a hypersurface on which the locally defined local expansion rate or "local Hubble time" has everywhere the same value); (c) specify this initial hypersurface arbitrarily but to this extent completely, that one gives everywhere the conformal part of its 3-geometry (otherwise one would not be able to specify "where" one is putting the mass-energy that is to determine inertia); (d) give everywhere on this 3-geometry the appropriate measures of density and flow of mass-energy plus (e) the appropriate "wave coordinates" and "wave momenta" of the gravitational field (because the gravitational waves and fields are effectively a source of mass-energy, too); (f) solve the appropriate initial value equation of geometrodynamics, a non-linear second-order partial differential equation, for the conformal scale factor of the 3-geometry (the existence and uniqueness of this solution being assured by the work of N. O'Murchadha and J. W. York, Jr.); (g) thus have in hand complete initial value data that satisfy the initial value equations of general realtivity; (h) solve the six dynamic equations of Einstein's theory for the entire 4-geometry; and in the end (i) have in this 4-geometry the full description of the inertial properties of every test particle, everywhere and at all times, as determined by the initial distribution of mass, as called for by Mach's principle. For details, see for example reference 3, pp. 543-549.

11. R. P. Geroch, *Singularities in the spacetime of general relativity: their definition, existence, and local characterization*, doctoral dissertation, (Princeton University, Princeton, New Jersey, 1967).

12. S. W. Hawking and R. Penrose, Proc. Roy. Soc. London A134, 529 (1969).

13. A. Einstein as quoted by G. Gamow, *My world line*, (Viking Press, New York, 1970).

14. W. Baade, Pub. Astron. Soc. Pacific, 28 (1956) 5.

15. A. Sandage in Proceedings of the Symposium on the Galaxy and the Distance Scale, Essex, England, in press, (1972).
16. J.H. Oort in *Onzième Conseil de Physique Solvay: La structure et l'évolution de l'univers*, (Editions Stoops, Brussels, 1958).
17. J.H. Oort, Science, 170 (1970) 1363.
18. D.W. Sciama, *Modern cosmology*, (Cambridge University Press, Cambridge, England, 1971) and earlier work cited there.
19. See for example reference 18.
20. See for example reference 18.
21. A.A. Penzias and R.W. Wilson, Astrophys. J., 142 (1965) 419
22. R.H. Dicke, P.J.E. Peebles, P.G. Roll, and D.T. Wilkinson, Astrophys. J., 142 (1965) 414.
23. For a summary and literature, see for example reference 3, p. 1214; pp. 457–459 and L. Landau and E. Lifshitz, *The Classical Theory of Fields*, translated by M. Hamermesh, (Addison-Wesley, then Cambridge, now Reading, Massachusetts, 1951) p. 320.
24. Theory reviewed in reference 3, pp. 978–979.
25. W. Baade and F. Zwicky, Proc. Nat. Acad. Sci. U.S., 20, 254–259 and 259–263 (1934) and Phys. Rev., 45 (1934) 138 (abstract).
26. A. Hewish, S.J. Bell, J.D.H. Pilkington, P.F. Scott, and R.A. Collins, Nature, 217 (1968) 709.
27. W.H. Press and K.S. Thorne, Ann. Rev. Astron. Astrophys., 10 (1972) 335.
28. J. Weber, Phys. Rev., 117 (1960) 307; also Phys. Rev. Lett., 24, (1970) 276 and papers therein cited.
29. J.M. Bardeen, Nature, 226 (1970) 64.
30. R. Penrose, Nuovo Cimento, 1 special number, (1969) 252.
31. S.W. Hawking, Phys. Rev. Lett., 26 (1971) 1344 and Commun. Math. Phys., 25 (1972) 152.
32. D. Christodoulou, Phys. Rev. Lett., 25 (1970) 1596.
33. D. Christodoulou and R. Ruffini, Phys. Rev. D4, (1971) 3552.
34. Ya. B. Zel'dovich and I.D. Novikov, *Relativistic Astrophysics, Vol. I: Stars and Relativity*, (University of Chicago Press, Chicago, Illinois, 1971).
35. J.A. Wheeler, "Mechanisms for jets" in *Study week on nuclei of galaxies, Pontificiae Acadameiae Scientarum Scripta Varia* no. 35, (1971) 539.
36. S.W. Hawking, Commun. Math. Phys., 25 (1972) 152 and W.H. Press and S. Teukolsky, Nature 238 (1972) 211.

37. C. DeWitt and B.S. DeWitt, eds., *Black holes*, Proceedings of 1972 session of Ecole d'été de physique théorique, (Gordon and Breach, New York, 1973).
38. R.W. Leach and R. Ruffini, Astrophys. J., 180 (1973) L15.
39. E. Cartan, *Leçons sur la géometrie des espaces de Riemann*, (Gauthier-Villars, Paris, 1928 and 1946); for more details, see J.A. Wheeler, "Gravitation as geometry, II" in H.-Y. Chiu and W.F. Hoffman, eds., *Gravitation and relativity*, (W.A. Benjamin, New York, 1964) and
C.W. Misner and J.A. Wheeler, "Conservation laws and the boundary of a boundary" in V.P. Shelest, ed., *Gravitatsiya: problemi i perspectivi: pamyati Alekseya Zinovievicha Petrova posvashaetsya*, (Naukova Dumka, Kiev, 1972), and reference 3, chapter 15.
40. D. Hilbert, Konigl. Gesell. d. Wiss. Göttingen, Nachr., Math. Phys., (1915; presented Nov. 20, 1915) 395.
41. R. Arnowitt, S. Deser and C.W. Misner, "The dynamics of general relativity" in L. Witten, ed., *Gravitation: an introduction to current research*, (Wiley, New York, 1962) pp. 227-265; see also ref. 3, chapter 21.
42. J.A. Wheeler, "Geometrodynamics and the issue of the final state" in C. DeWitt and B.S. DeWitt, eds., *Relativity, Groups and Topology*, (Gordon and Breach, New York, 1964), see also ref. 3, chapter 21.
43. J.W. York, Jr., Phys. Rev. Lett., 28 (1972) 1082 and J. Math. Phys., 14 (1973) 456.
44. A. Lichnerowicz, J. Math. Pures et Appl., 23 (1944) 37.
45. Y. Choquet-Bruhat, Comptes Rend., Acad. Sci. Paris, A274 (1972) 682 and Gen. Rel. and Grav. (1973), in press.
46. B.S. DeWitt, Phys. Rev., 160 (1967) 1113; 162 (1967) 1195; and 162 (1967) 1239.
47. A. Peres, Nuovo Cimento, 26 (1962) 53.
48. U. Gerlach, Phys. Rev., 177 (1969) 1929.
49. T. Regge, Nuovo Cimento, 19 (1961) 558.
50. W. Pauli, Helv. Phys. Acta Suppl., 4 (1956) 69.
51. J.A. Wheeler, *Geometrodynamics*, (Academic Press, New York, 1962) pp. 75-83.
52. Summarized for example in P.J. Dyson, *Advanced quantum mechanics*, (Cornell University, Ithaca, New York, 1954) (mimeographed), p. 54.
53. Ref. 51, pp. 71-75.
54. K. Kuchař, J. Math. Phys., 11 (1970) 3322.

55. W.K. Clifford, *Lectures and essays*, L. Stephen and F. Pollock, eds., (MacMillan, London) vol. 1, p. 244 and p. 322, 1879; see also *Mathematical papers*, R. Tucker, ed., (MacMillan, London), p. 21, 1882.
56. H. Weyl, *Was ist Materie*, (Springer, Berlin, 1924), pp. 57–58.
57. C.W.Misner and J.A. Wheeler, Ann. of Phys., 2 (1957) 525.
58. A. Compte, *Cours de philosophie positive*, (Paris, 1835).
59. J.R. Klauder, Ann. of Physics 11 (1960) 123.
60. Regarding the connection between topology, superspace, and spin, see for example J.A. Wheeler, "Superspace and the nature of quantum geometrodynamics" in C.M. DeWitt and J.A. Wheeler, eds., *Battelle rencontres: 1967 lectures in mathematics and physics*, (W.A. Benjamin, New York, 1968) 242–307, esp. pp. 284–289.
61. B. Riemann, 1854 "Habilitationsvorlesung" in H. Weber, ed., *B. Riemann: Gesammelte mathematische Werke*, 2nd ed., paperback reprint, (Dover, New York, 1953); translated into English by W.K. Clifford, Nature, 8 (1873) 14.
62. Riemann himself was the first to introduce superspace. In his case it was a superspace built of conformal 2-geometries, $^{(2)}\mathcal{C}$. This superspace has since been the object of research of more than a hundred papers. Today it is known as Teichmüller space.
63. See ref. 55.
64. S.A. Hojman, K. Kuchař, and C. Teitelboim, "New approach to general relativity", April 1973 preprint.
65. A.D. Sakharov, *Doklady Akad. Nauk S.S.R.*, *177*, (1967) 70; English translation in Sov. Phys. Doklady, 12 (1968) 1040.
66. See ref. 34.
67. M. Berger, P. Gauduchon, and E. Mazet, *Le spectre d'une variété Riemannienne*, (Springer, Berlin and New York, 1971).
68. P.J.E. Peebles, *Physical cosmology*, (Princeton University Press, Princeton, New Jersey, 1971).
69. A. Avez, Comptes Rend., Acad. des Sciences, Paris, 250 (1960) 3583.
70. See ref. 11.
71. See ref. 12.
72. F.J. Dyson, "The fundamental constants and their time variation" in A. Salam, ed., *Dirac Festschrift*, (Cambridge University Press, Cambridge, England, 1972).
73. D.I. Mendeléev, selection of quotations from, in J.A. Wheeler, "From Mendeléev's atom to the collapsing star", 189–233 in *Atti del Convegno Mendeléeviano*, (Accademia delle Scienze, Torino, 1971); reprinted in Trans. New York Acad. Sci., 33 (1971) 745.

74. H. Everett III, Rev. Mod. Phys., 29 (1957) 454.
75. B.S. DeWitt, "The many-universes interpretation of quantum mechanics", 211-262 in B. d'Espagnat, ed., *Foundations of quantum mechanics*, (Academic Press, New York, 1971).
76. J.M. Jauch, *Foundations of quantum mechanics*, (Addison-Wesley, Reading, Massachusetts, 1968), especially chapters 5 and 8.
77. T. Mann, *Freud, Goethe, Wagner*, (Knopf, New York, 1937). Essay on Freud published originally as *Freud und die Zukunft (Vortrag gehalten in Wien am 8 Mai 1936 zur Feier von Sigmund Freuds 80 Geburtstag)*, (Bormann-Fischer, Wien, 1936).
78. Parmenides of Elea, section on *Truth* in his poem, as paraphrased in *Encyclopedia Brittanica* article on Parmenides, p. 327, column 2, lines 9 and 16-17.
79. G.W. von Leibniz (1646-1716) as quoted by J. R. Newman, *The world of mathematics*, (Simon and Schuster, New York, 1956) p. 1832.