Causality and Multiply Connected Space-Time

ROBERT W. FULLER*
Pupin Physical Laboratories, Columbia University, New York, New York

AND

JOHN A. WHEELER†
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received May 16, 1962)

With the introduction of multiply connected topologies into physics, a question of causality arises. There are alternative routes between two points in a multiply connected space. Therefore, one may ask if a signal traveling at the speed of light along one route could be outpaced by a signal which has traveled a much shorter path through a handle or "wormhole." This paper examines one such situation and shows that in this example causality is preserved. It proves essential in the analysis to distinguish between those regions of space-time which are catastrophic and those which are not. A catastrophic region is composed of catastrophic points. A catastrophic point in space-time is so located with respect to eventual singularities in the intrinsic geometry that every time-like geodesic through it necessarily runs into a region of infinite curvature at some time in the future—or was born out of a region of infinite curvature at some time in the past—or both. If a classical analysis of nature were possible—which it is not—then it would be natural to postulate that laboratory physics is carried out in noncatastrophic regions of space-time. Two such regions are shown to exist in the example considered in the paper. It is shown that no signal can ever be sent from one to the other. The key point in preventing any violation of causality is simple: The (Schwarzschild) throat of the wormhole pinches off in a finite time and traps the signal in a region of infinite curvature. This investigation also displays some of the unusual geometric features of the Schwarzschild solution of Einstein's equations for a spherically symmetrical center of attraction. Radial spacelike geodesics passing through the throat are calculated and it is shown that there exist regions of space-time unreachable by any radial geodesics that issue from a given point. Also, there exist points in space-time from which light signals can never be received no matter how long one waits.

I. INTRODUCTION

THE Einstein field equations, being differential equations, are purely local in character. They tell nothing about the topology of the space with which one is dealing. Recent investigations of multiply connected space show that this kind of topology is of great interest in giving a natural place for electricity in a curved empty space that is free of all "real electric charge." Charge—in this purely classical model of physics which has nothing directly to do with the real world of quantum physics—is interpreted as lines of force trapped in the topology of space,[1] However, it might appear that a multiply connected topology violates elementary principles of causality. There are alternative routes for a disturbance to pass from a point A to a point B. A disturbance going by one of these routes as fast as it can—at the speed of light—may arrive only to find itself outpaced by a disturbance which has gone through a handle or "wormhole" and a much shorter path. This paper examines in detail one such situation, and shows that in this example causality, despite first expectations, is not violated. The key point in preventing any violation of causality is this: that the "throat" of the wormhole pinches off a finite time. This investigation incidentally also gives some insight into the unusual geometric properties of the Schwarzschild solution.

It should be emphasized that the world of geometrodynamics—of the geometry of curved empty space developing in time in accordance with Einstein's equations—is only a model. The connection, if any, between this model and the real physical world is a matter about which we are still in great ignorance. Therefore, it should be stressed that the masses to be discussed here have not the least direct connection with either real stars or, still less elementary particles. Any consequences that come out of the analysis are of value chiefly for what they tell about the spirit and way of working of general relativity, not for anything they might conceivably tell about the physical world.

II. TRANSMISSION OF SIGNALS THROUGH WORMHOLES

The most elementary example of a wormhole—and the one at the basis of the discussion in this paper—is shown in Fig. 1. This illustrates schematically the geometry of a three-dimensional space with a positive definite metric, though, of course, only two dimensions can be shown in the diagram. This three-dimensional "spacelike hypersurface" is to be conceived as a slice through the four-dimensional geometry; in other words a snapshot showing the 3-geometry at a particular time. The multiple connectedness is a multiple connectedness in space, not in time. A wire can thread through the

* This work was begun while the author (RWF) was at the University of California, Berkeley, California, and it was supported by the U. S. Atomic Energy Commission through contract with Princeton University.

† Part of this work was done at the University of California, Berkeley, California while the author (JAW) was on leave of absence from Princeton University.

[1] J. A. Wheeler, *Geometrodynamics* (Academic Press Inc., New York, 1962).

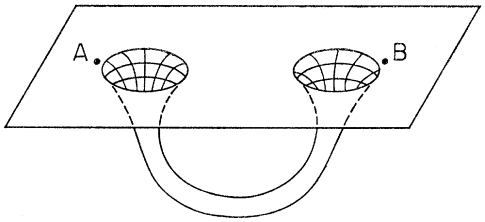

FIG. 1. Wormhole topology. The diagram shows the geometry of space at a particular moment of time. In other words, it represents the geometry on a particular slice of space-time. For convenience of representation one of the three space coordinates has been dropped. A 2-dimensional space is used here to epitomize the properties of this 3-dimensional space. This 2-dimensional space is curved. This curvature is most conveniently shown by imbedding the 2-dimensional space in a 3-dimensional Euclidean space (above diagram). The third dimension in this diagram has no physical meaning. Only distances within the surface have significance. This surface is endowed with an unusual topology but is everywhere free of any geometrical singularity. The topology shows up in the existence of two inequivalent ways to connect point A with point B. One connection passes through the throat of the wormhole. The other remains entirely in the quasi-Euclidean space exterior to the mouths of the wormhole. The lengths of the two connections happen to be comparable in the figure so drawn. However, it is perfectly possible for the connection through the wormhole to be shorter by many orders of magnitude than the "normal" route from A to B. This possibility is most readily visualized on bending the surface so the two mouths of the wormhole come almost back-to-back. Then the throat becomes almost negligibly short. The possibility of sending signals through such a shortened route is the problem of concern in this paper.

throat in this diagram. Then it supplies a connection between the two mouths of the wormhole. Another wire can be strung out which also goes from one mouth to the other without ever going through the throat. The two types of connection are topologically distinct. It is impossible by a continuous sequence of small deformations to transform one route into the other. Moreover, the lengths of the two routes are metrically distinct. The wire that stays in the nearly Euclidean space fringing the throat—and remote from the throat—may, for example, have to have a length of 1 km to connect one mouth with the other. Yet in the same illustration the other wire can run directly through the throat from one mouth to the other with a length of only 1 m.

There is even a simple example of a connection in which the length of the throat is zero. Start with a three-dimensional Euclidean space. At point $(x,y,z)=(0,0,b)$ construct a sphere of radius a small compared to b. Construct a similar sphere of the same size at the point $(0, 0, -b)$. "Remove the points" that lie within the two spheres. Identify the points that lie on the surfaces of the two spheres, pairwise between one sphere and the other. A test particle approaching the upper sphere and crossing the critical boundary finds itself suddenly emerging from the lower sphere. Thus the test particle—or a ray of light—appears to have the possibility to pass from one point in space to a point in space which may be many miles away in a negligible amount of time.

Such a rapid communication from one place A to another place B would seem to violate elementary notions of causality. To say that any method of signaling is impossible which would exceed the speed of light is only the weakest way to state the conflict. Einstein long ago stressed that such a signaling process, viewed from an inertial system moving at a sufficiently great but still allowable speed ($v<c$), will show that B *receives* the message before A *transmits* it.

To examine whether this difficulty will arise for signals sent through a simple type of wormhole is the purpose of this paper. We conclude in this special case that the throat pinches off before the signal can get far enough to violate causality. We leave untouched the question of a general analysis suitable for the case of an arbitrary multiply connected topology.

It is perfectly possible to write down a mathematical expression for the metric of a space which has a simple wormhole geometry just discussed. This geometry can have a high degree of symmetry. It can be invariant with respect to rotations about the z axis. Nevertheless, the problems of the mathematical analysis are sufficient so that one would like to have a still higher degree of symmetry if he could in order to simplify the analysis further. This extra symmetry can be obtained if one will give up looking at the uninteresting aspects of the problem—that is to say the passage of light by the "long way" from one wormhole mouth to the other—and concentrate instead on the decisive part of the problem, the passage of a light ray or material particle through the throat itself.

With this idea of simplification in mind go to the ideal limit where the two wormhole mouths are indefinitely far apart in comparison to their own proper dimensions. In this case, a material particle or light signal coming out of one of the mouths and traveling even for a very long time will still not be able to arrive in the vicinity of the other mouth and will not be aware that it is there. The same will apply for a material particle or light signal emerging out of the other mouth. Therefore, for all practical purposes the single Euclidean space can be regarded as two separate Euclidean spaces. The only easy way to get from one to the other is by way of the throat itself. Thus, the limiting case of a single wormhole with its mouths indefinitely far apart may be regarded, or may be idealized, as a pair of Euclidean spaces connected by a "bridge," Fig. 2.

III. GEOMETRY OF THE SCHWARZSCHILD THROAT

This kind of geometry does not have to be invented; it was already found long ago by Schwarzschild in his famous solution of Einstein's equations for a spherically symmetrical center of attraction, though it was only recently through the work of Fronsdal[2] and Kruskal[3] that one has come to understand the unusual nature of the topology implied by the Schwarzschild solution.

In most physical applications the center of attraction is portrayed as having "real mass" as, for example,

[2] C. Fronsdal, Phys. Rev. **116**, 778 (1959).
[3] M. D. Kruskal, Phys. Rev. **119**, 1743 (1960).

in the case of a sphere of incompressible fluid. In this example the Newtonian gravitational potential outside the sphere has a $1/r$ behavior and inside the sphere has a harmonic oscillator character, the two functions matching smoothly in magnitude and slope at the boundary. In the Einsteinian description of the same situation different mathematical expressions are likewise used for the geometry inside the sphere and outside: the "interior Schwarzschild solution" and the "exterior Schwarzschild solution." The mass is measured both in the Newtonian case and in the Schwarzschild metric by the coefficient of the $1/r$ term in the relevant mathematical expression far away from the center of attraction.

For a given value of this mass it is, of course, possible to have one or another value for the radius at which the mass distribution is encountered as one goes inward. However it is also possible to investigate the character of the geometry when there is no "real mass" present at all. Kruskal and Fronsdal have shown that this geometry, depicted at an appropriate moment of time, is perfectly regular and free of all singularity. It describes a throat of circumference

$$2\pi(\text{"Schwarzschild radius"}) = 2\pi m^* = 2\pi[G(\text{mass as seen far away})/c^2].$$

The mass in such a situation arises entirely from the equivalent mass-energy of the gravitational disturbance (curving of space) itself.

In terms of the geometry that is now at hand, the question becomes the following: "Can a disturbance start from a point on the 'lower space' and get through to a point on the 'upper space'?" Translated back to the kind of wormhole geometry originally depicted in Fig. 1, this question translates over into the question: "Can a signal, instead of having to travel the great distance from the neighborhood of one wormhole mouth through the quasi-Euclidean space to the other wormhole mouth, get through directly from one point to the other by the short circuit through the throat of the wormhole?"

IV. GEOMETRODYNAMICS OF THE SCHWARZSCHILD THROAT

It might appear that the answer is plainly yes. The geometry of the Schwarzschild throat is perfectly regular and light propagates there just as it does anywhere else. However, an account has not yet been taken of the geometrodynamical character of the Schwarzschild geometry. Usually one thinks of the geometry as being static because there is no time dependence of the metric coefficients in Schwarzschild's familiar way of writing his solution of Einstein's equations. However, a proper analysis deals with the intrinsic geometry as distinguished from the coordinates. Such an analysis shows that the geometry is changing with time. This change proceeds in such a way that the throat of the wormhole, like the shutter of a camera, is open only

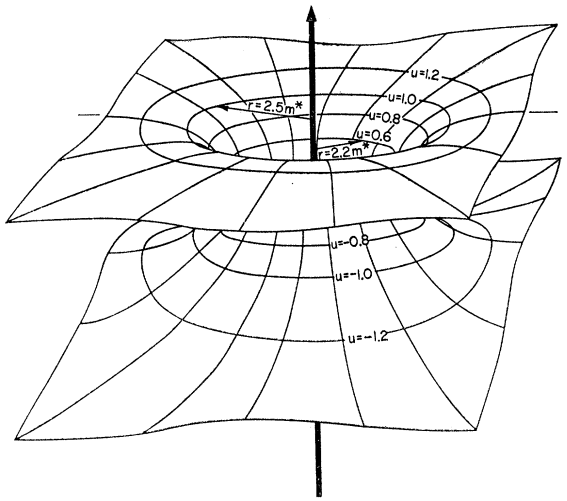

FIG. 2. Schematic representation of Einstein-Rosen bridge between two nearly Euclidean spaces. A slice has been taken through space-time at the time $T=0$ (Kruskal's $v=0$). In addition the Schwarzschild polar angle θ has been given the equatorial value $\theta=\pi/2$. Therefore, only a two-dimensional space comes into consideration: $ds^2=(1-2m^*/r)^{-1}dr^2+r^2d\phi^2$. Here r runs from $2m^*$ to ∞. This representation hides the fact that the proper space is two-sheeted. This two-sheeted structure is shown in the diagram. Analytically, the double structure is most conveniently shown by introducing the new coordinate u. Values of this coordinate are marked on the surface. The quantity u is defined by the equation $u^2=[(r/2m^*)-1]\exp(r/2m^*)$. Then u has positive values in the upper sheet and negative values in the lower sheet. The metric, expressed in terms of Kruskal's[3] variable u (transcendental function!) is completely regular over the entire 3-space. The apparent singularity at $r=2m^*$ was due entirely to an unfortunate choice of coordinates.

for a limited proper time. This circumstance by itself would not interpose an insuperable obstacle in the way of sending a signal from the lower space in Fig. 2 to the upper space. If this were the only problem it would only be necessary to time the start of the signal towards the first wormhole mouth to have it get through safely during the short interval "while the shutter is open" and emerge from the second mouth. However, the dynamics of the geometry has another critical consequence for the propagation of the signal which is best described by way of an analogy.

Compare the surface in Fig. 2 (which actually epitomizes a curved 3-geometry) with a sheet of rubber. Around the upper and lower rims of this sheet at great distances radial tensions can be applied or removed. In consequence, the throat can be made to stretch out and reach a maximum circumference. By release of the tensions at the two rims it can be allowed to collapse, Fig. 3. This model of a rubber sheet provides a remarkably useful insight into the behavior of the Schwarzschild-Fronsdal-Kruskal geometry. The moment of maximum extension corresponds to the situation shown in Fig. 2. Consider, in contrast, the phase of more and more complete release of tension as T goes to infinity (and the time-symmetric situation as T is traced back to minus infinity). Then the rubber is piled up more and more around the axis as depicted in Fig. 3. It follows

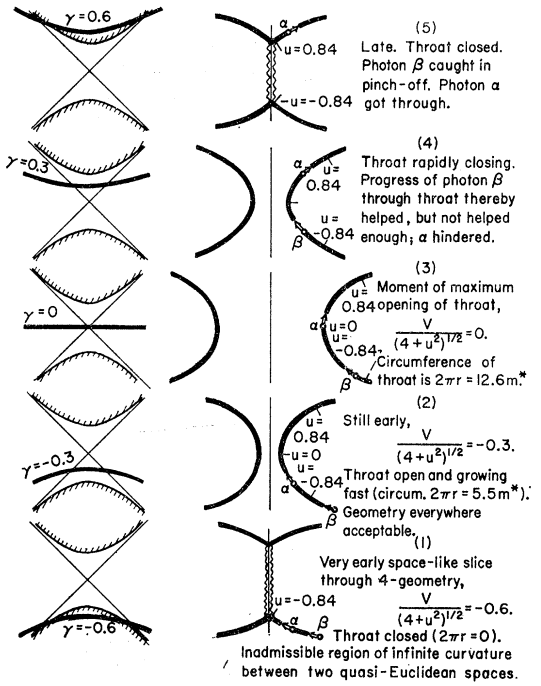

FIG. 3. Change of geometry with time as Schwarzschild throat opens, reaches maximum diameter, contracts, and pinches off (Sequence 1 through 5). The diagrams at the left depict space-time in terms of Krushal's coordinates u and v. They show the particular, arbitrary, slices through space-time, the 3-geometry of which is schematically represented in the drawings at the right. The slices were quite arbitrarily selected by values of a coordinate γ defined, as a matter of pure option, by the equation $\gamma=v/(4+u^2)^{1/2}$. What goes on is most easily visualized in terms of a sheet of rubber outward by a pull around two rims (phase of maximum extension depicted more fully in Fig. 2) and then allowed to collapse again. A typical photon that was moving initially radially inward progresses with the speed of light on this surface which is itself however best visualized as moving inward. Points on this surface are designated by u values to provide reference marks to measure the progress of two sample photons, α and β. One has got caught in the pinch off. Its fate is as mysterious as is the condition of geometry (rippled region near axis in diagrams 1 and 5) which has had a wave of infinite curvature sweep over it (geometry which has turned into something like a foam of scale, $(\hbar G/c^3)^{1/2}=1.6\times10^{-33}$ cm, comparable to the dimension characteristic of quantum fluctuations in the metric? See reference 1, pp. 67–77. The question still remains (see discussion in text) whether photon α is safely through or whether it will get caught in the collapse of the rubber sheet down to a circumference $2\pi r=0$.

from this way of speaking that the dynamics of the geometry is not confined to the immediate vicinity of the throat and the opening and closing of the "shutter". Instead the dynamics affects the geometry out to indefinitely great distances. In effect the photon is traveling in a medium which itself is moving—moving away from the throat during the phase of expansion; moving toward the throat during the phase of contraction. Consequently, there is no hope for a photon, which starts far away from the throat to ever get to and through this passage no matter how early it starts. During all the critical time before the quick opening and closing of the shutter it is struggling against an irresistible tide. When the tide turns the assistance comes too late. Therefore, only those photons which start close enough to the throat have any hope to go through the throat.

V. CAUSALITY PRESERVED BY THE PINCH OFF OF THE SCHWARZSCHILD THROAT

These qualitative considerations can be stated quite precisely (Fig. 4). The space-time in the neighborhood of the Schwarzschild throat is sharply divided by light cones through an invariantly defined origin (singular point of Killing vector field) into two "catastrophic regions," II and IV, and two "noncatastrophic regions" I and III. The track of any material particle, any atom, or any light source is a timelike geodesic inclined to the vertical in Fig. 4 less than 45°. Such a geodesic in region II runs into catastrophic conditions after a finite and short proper time. Similarly, any timelike track in IV originates in such conditions. These are not the conditions in which any laboratory experiments are ever carried out. Therefore, it is natural to postulate that all normal observations and experimentation are carried out in regions like I and III. But simple inspection of Fig. 4 shows that it is impossible to get a light ray across from region III to region I. On this basis it is necessary to conclude that no experimenter in one noncatastrophic region can send a signal through the throat to an observer who is in the other noncatastrophic region. Therefore, in this sense, any violation of causal principle such as discussed in the beginning can not occur.

It is comforting to have this confirmation of causality for sources and receptors which can escape running into regions of infinite curvature. One would still like to characterize more fully points of emission or reception which lie in the catastrophic region, the region whose properties are so foreign to usual ways of thinking. There is an alternative characterization of these points which is very simple. Their Schwarzschild coordinate r is less than $2m^*$. These points can therefore be pictured in Figs. 2 and 3 as points which would be cut out of the rubber sheet were a cookie cutter of this radius punched through it. Every point outside this cookie cutter on the lower sheet lies in region III. No signal can ever get from outside the cookie cutter on this lower sheet to outside it on the upper sheet. The sources and receptors with a life record of catastrophe which (region IV) can send signals both to I and III or receive (region II) from both I and III lie entirely within the cookie cutter.

VI. UNSOLVED PROBLEMS: THE CHALLENGE OF THE GENERAL CASE

Interesting as it is to analyze causality in the particularly simple case of the Schwarzschild throat, one is unhappy not to have an analysis more far-reaching and general in its scope. It appears likely that additional insights could be gained by analyzing more general situations. Among these, one of the more interesting, is that of a Schwarzschild throat through which thread

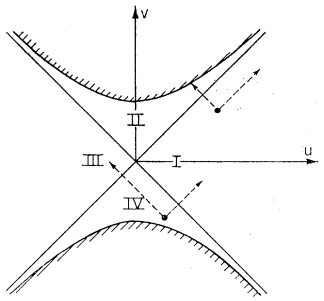

FIG. 4. Classification of regions of space-time associated with the Schwarzschild geometry and of the possibilities for communicating between them. Points in space-time are designated by Kruskal's coordinates u (space-like) and v (time-like). All that connected region of space-time within which the intrinsic geometry is free of singularity is comprised between the two cross-hatched hyperbolas. Any time-like geodesic is inclined closer to the vertical than 45°. Any such geodesic in region II is fated to run into a region of catastrophically high curvature. Similarly, any time-like geodesic in region IV has been born out of a condition of infinite curvature. It is postulated (insofar as it makes sense to use this or any related terminology out of classical physics to discuss what is really a quantum world) that laboratory physics is not carried out in such regions. On this basis it is concluded that no experimenter can send a message, as can a source in region IV, to both regions I and III. Laboratory observers, confined to noncatastrophic regions of space-time such as I and III are deprived of all possibility to send signals through the throat of the wormhole. The means which they would otherwise have to bring noncausal phenomena into evidence is thereby taken away from them (separation of III and I by intersecting light rays, boundaries which no other light rays can ever cross).

electric lines of force. As the throat starts to pinch off, the electric pressure rises and the contraction stops short of pinch off. The radius of the throat oscillates periodically with time as is shown in most interesting detail by Graves and Brill.[4] Under these conditions it might at first sight appear obvious that causality would be violated. An indefinitely long time appears to be available for signaling from one quasi-Euclidean space to the other. However, the topological considerations which are called for in this problem are even more subtle than those encountered in the present instance. Therefore, it would seem rash to conclude without investigation that any question can remain in the end as to the validity of the causal principle.

In this connection mention should be made of the recent work of Smith.[5] He notes that the existence of the light cone in a physical world gives one a means to characterize a manifold which is additional to the usual topological indices (number of bundles or wormholes; degree of multiple connectedness; etc.) and which still falls short of all the detail which is demanded by a full knowledge of the metric. As an example of the classificatory tools provided by the combination of topology and light cone it is enough to point to the division of the regular part of the Schwarzschild-Fronsdal-Kruskal metric into the regions which have been designated above as I, II, III, IV. It is conceivable that one has only to spell out the consequences of the classification scheme of Smith in order to have a complete analysis of causal relations in the most general manifold allowed by geometrodynamics.

In the course of investigating motion in the immediate vicinity of the Schwarzschild throat we found ourselves asking how the geodesics could be continued across from one of the four regions (I, II, III, IV) to another. Darwin[6] has already made a detailed analysis of motion in the familiar region I. However, the problem of matching coordinates between this region and its neighbors presented certain sophistications. For this reason we found it interesting to analyze this problem in a little detail with the results illustrated in Fig. 6. The details which went into this analysis are reported in the following Appendix. They add particulars to the foregoing general conclusions about causality without in any way modifying them.

ACKNOWLEDGMENTS

We wish to express our appreciation to Professor Niels Bohr who—after listening to an account of the concept of electricity as lines of force trapped in the topology of a multiply connected space, given by one of us at Copenhagen on March 17, 1958—raised the problem of causality and stimulated the investigation reported here. We are also indebted to Dr. Peter Putnam and Professor Martin Kruskal for discussions. Finally, we are grateful to Professor Valentine Bargmann, Professor David Finkelstein, Professor Charles Misner, and Professor Wolfgang Smith for often emphasizing that the light cone provides a mathematical tool for the analysis of geometrodynamics additional to the usual tools of metric geometry. We believe that this tool still remains to be put to full use, and that causality is the physical principle which will guide this future development.

APPENDIX

I. Radial Geodesics Through the Schwarzschild Throat

Fronsdal[2] and Kruskal,[3] following earlier work of Lemaître,[7] Einstein and Rosen,[8] Synge,[9] and Finkelstein,[10] have shown that Schwarzschild's unique[11] solution of Einstein's field equations for curved empty space,

$$ds^2 = -(1-2m^*/r)dT^2 + (1-2m^*/r)^{-1}dr^2 + r^2(d\theta^2+\sin^2\theta d\phi^2), \quad (1)$$

is expressed in a defective coordinate system and have

[4] J. C. Graves and D. R. Brill, Phys. Rev. **120**, 1507 (1960).

[5] J. Wolfgang Smith, Proc. Natl. Acad. Sci. U. S. **46**, 111 (1960). Also, Y. H. Clifton and J. W. Smith, Proc. Natl. Acad. Sci. U. S. **47**, 190 (1961).

[6] C. G. Darwin, Proc. Roy. Soc. (London) **A249**, 180 (1958); **A263**, 39 (1961). See pp. 49–50 for discussion of $r \leq 2m^*$.

[7] G. Lemaître, Ann. soc. sci. Bruxelles **53A**, 51 (1933).

[8] A. Einstein and N. Rosen, Phys. Rev. **48**, 73 (1935).

[9] J. L. Synge, Proc. Roy. Irish Acad. **A53**, 83 (1950).

[10] D. Findelstein, Phys. Rev. **110**, 965 (1958).

[11] That this is the only spherically symmetric solution up to a coordinate transformation, of the equations for empty space was proven by G. D. Birkhoff, *Relativity and Modern Physics* (Harvard University Press, Cambridge, Massachusetts, 1923), p. 253.

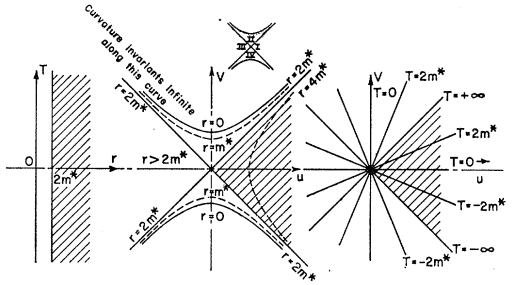

FIG. 5. Corresponding regions of the (r,T) and (u,v) planes. In the (u,v) plane curves of constant r are hyperbolas asymptotic to the lines $r=2m^*$. Curves of constant T are straight lines through the origin. Older treatments of the Schwarzschild metric limited attention to the shaded areas, defined by $r>2m^*$ (or $|v|<u$ in Kruskal coordinates). This region (middle diagram) lies between the lines $T=-\infty$ and $T=+\infty$. Isn't this time interval enough for anyone? No! T is only a coordinate, not any direct measure of proper time. It is ill suited to cover the whole of that region of 4-space where the geometry is free of singularity and where geodesics can run freely. That the coordinate r is also defective is seen even more easily. Example: Consider the hypersurface $T=0$ (or $v=0$ in Kruskal coordinates). The 3-geometry of this hypersurface is that of the Einstein-Rosen bridge depicted in Fig. 2. To trace out this surface it is not sufficient merely to let the coordinate u decrease from ∞ down to 0 (r from ∞ down to $2m^*$); it is necessary that u continue on to $-\infty$ (r going back on the second sheet from $2m^*$ to $+\infty$). The circumference of the throat at the time in question is evidently $2\pi r_{\min}=4\pi m^*$. At a later time (a greater v, for example; not a greater T, because T is a defective coordinate) the 3-geometry has changed and the throat has shrunk. This is seen by running across the central diagram at constant v with u decreasing from $+\infty$ to $-\infty$. The coordinate r drops from $+\infty$ to a value less than $2m^*$ and rises again to $+\infty$. Therefore, the circumference of the throat is less than $4\pi m^*$. As time proceeds, the size of the throat—measured by its circumference—shrinks further and ultimately goes to zero. This pinch off shows itself also in this, that the intrinsic curvature invariants of the geometry (the quantities whose values are independent of the coordinates in terms of which they are calculated) go to infinity along the two hyperbolas $r=0$ in the central diagram. In the regions marked I, II, III, IV in the insert the equations for geodesics are solved in terms of distinct parameters. (See Table II for the parameters and Fig. 6 for the geodesics.)

supplied new coordinates to take their place. The coordinates r and T cover only a part of the space-time geometry defined by the field equations. New coordinates can be given which cover in a regular way the entirety of that region of spacetime where the intrinsic geometry—as distinct from any particular sets of coordinates—is regular. Figure 5 illustrates the character of the coordinates (u,v) of Kruskal. In terms of these coordinates the metric takes the form

$$ds^2=f^2(-dv^2+du^2)+r^2(d\theta^2+\sin^2\theta d\phi^2). \tag{2}$$

Here, f and r are transcendental functions of the quantity u^2-v^2. Their dependence on u^2-v^2 is defined by the equations

$$[(r/2m^*)-1]\exp(r/2m^*)=u^2-v^2 \quad (\text{with } r\geqslant 0), \tag{3}$$

$$f^2=(32m^{*3}/r)\exp(-r/2m^*). \tag{4}$$

Kruskal's coordinates are well adopted to discuss the geodesics of the space because they make all radial lightlike geodesics appear as straight lines in the (u,v) plane with slope $\pm 45°$.

We have looked into the nature of radial geodesics in this space because we wanted to know the answer to the following question: Can every point A in the continuum be connected with every other point B, endowed with the same polar coordinates, θ and ϕ, by a geodesic? This question arose because the intrinsic geometrical singularity in the metric at $r=0$ in the Schwarzschild coordinates, or on the hyperbola

$$v^2-u^2=1 \tag{5}$$

in the Kruskal coordinates, might under certain circumstances be conceived to interpose an obstacle to making a geodesic connection between A and B. It is not clear that the question has any deeper significance than this, that it aims at understanding better some of the new features of the geometry in the Schwarzschild space-time manifold. It has recently been conjectured that all closed space solutions of Einstein's equations with the topology S^3 of the 3-sphere will develop an intrinsic geometrical singularity after the lapse of a finite proper time just as does the open universe of Schwarzschild.[12] Therefore some insight into the consequences of "curvature barriers" in the one case may some day be of use in understanding a little better what can happen in the other case.

II. Calculation of the Radial Geodesics

For simplicity, we use hereafter the reduced variables

$$\begin{aligned} r^*&=r/m^*,\\ T^*&=T/m^*,\\ ds^*&=ds/m^* \quad (\text{proper distance}),\\ d\tau^*&=(-ds^2)^{1/2}/m^* \quad (\text{proper time}),\\ a&=R/m^* \quad (\text{initial value of } r^* \text{ in certain formulas}); \end{aligned} \tag{6}$$

and for further simplicity we omit the stars in all that follows. It would be reasonable to carry out the analysis of geodesics entirely in terms of the variables u and v. However, this course presents difficulties because the metric coefficients are defined in terms of u and v only through implicit transcendental relations. Therefore, another plan is followed: (1) We write the equations for geodesics in the variables r and T. (2) We note that any direct integral of these equations is unobtainable because it would itself involve implicit transcendental relations between r and T. (3) Therefore, we introduce a well-known supplementary parametric angle w such that r and T become simple trigonometric or hyperbolic functions of w. (4) Finally, we translate the resulting information about r and T as functions of w into information about u and v as functions of w.

The equations for spacelike radial geodesics[12] have the form

$$(d^2r/ds^2)+(1-2/r)(1/r^2)(dT/ds)^2 - (1/r^2)(1-2/r)^{-1}(dr/ds)^2=0, \tag{7}$$

[12] See, for example, R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Clarendon Press, Oxford, 1934).

TABLE I. The three kinds of radial geodesics. For time-like geodesics with energy less than mc^2, a represents the maximum value of r reached by the object before it starts falling back towards the center of attraction. When a takes on any one of the limiting values indicated in the table by inequalities, then the geodesic in question becomes light-like.

Kind of geodesic	Space-like	Time-like: Energy $<mc^2$	Time-like: Energy $>mc^2$
Value of a in Eq. (11)	$0<a<2$	$2<a<\infty$	$-\infty<a<0$
$\dfrac{\text{(kinetic energy)}}{mc^2}=\left[-g_{TT}\left(\dfrac{dT}{d\tau}\right)^2\right]^{1/2}$ for $r>2$	Imaginary	$\left(\dfrac{1-2/a}{1-2/r}\right)^{1/2}$	$[(1-2/a)/(1-2/r)]^{1/2}$
$\dfrac{\text{(kinetic momentum)}}{mc}=\left[-g_{rr}\left(\dfrac{dr}{d\tau}\right)^2\right]^{1/2}$ for $r>2$	Imaginary	$\left(\dfrac{2/r-2/a}{1-2/r}\right)^{1/2}$	$\left(\dfrac{2/r-2/a}{1-2/r}\right)^{1/2}$
$(E_{\text{kin}}/mc^2)^2-(P_{\text{kin}}/mc)^2$ for *all* r	-1	1	1
$(dT/d\tau)_{\text{"}r=\infty\text{"}}=$ (total energy)$/mc^2$	Imaginary	$(1-2/a)^{1/2}<1$	$(1-2/\alpha)^{1/2}>1$
Course of geodesic	Starts at infinite distance, goes through throat of wormhole and goes off again to infinite distance. Asymptotic behavior: $(dT/ds)_\infty=\pm(2/a-1)^{1/2}$; $(dr/ds)_\infty=+(2/a)^{1/2}$	Starts at $r=0$ at point (u,v) on curve $v^2-u^2=1$; v increases steadily; r rises to a and falls back again to $r=0$; geodesic ends on curve to $v^2-u^2=1$.	Starts at $r=0$ at point (u,v) on curve $v^2-u^2=1$; v increases steadily; so does r; r goes to ∞; all this for an outgoing geodesic; there also exist incoming geodesics of the converse character.
Expression for a used in further analytic treatment of motion	$a=2/\cosh^2 w_0$ $=r_{\min}$	$a=2/\cos^2 w_0$ $=r_{\max}$	$a=-2/\sinh^2 w_0$
Expression for (total energy$/mc^2)^2=(1-2/a)$ in terms of w_0	$-\sinh^2 w_0<0$	$\sin^2 w_0<1$	$\cosh^2 w_0>1$

and

$$(d^2T/ds^2)+(2/r^2)(1-2/r)^{-1}(dr/ds)(dT/ds)=0. \quad (8)$$

For time-like geodesics ds^2 is replaced everywhere by $-d\tau^2$. One first integral of (7) and (8) is the equation normally used to define proper distance or proper time:

$$-(1-2/r)dT^2+(1-2/r)^{-1}dr^2=ds^2=-d\tau^2. \quad (9)$$

Another first integral is the relation

$$-(dT/ds)^2=(dT/d\tau)^2=(1-2/\alpha)(1-2/r)^{-2}. \quad (10)$$

or, better, a combination of (9) and (10):

$$dT^2=\frac{(1-2/a)dr^2}{(1-2/r)^2(2/r-2/a)}. \quad (11)$$

The quantity a determines the character of the geodesic (Table I).

To proceed further it is convenient to divide up the situations under analysis in Table I and introduce the angular parameters of Tables II and III.

The constants of integration are determined in such a way that the geodesic shall go through a prescribed point, such as the point A in Fig. 6. The calculation proceeds as follows:

(1) From the prescribed values of u and v at the point A, values are calculated for r and T. (2) An arbitrary choice is made for the parameter w_0 (which in the case of timelike geodesic determines the energy). (3) Then, w (which determines the phase of the motion) is selected to make r equal to the prescribed r. (4) Finally, the value of T is calculated for this w and for the prescribed w_0, in a form such as $T=T_0+$"1.913"; and then the value of the constant T_0 is so selected as to make the time take on the assigned value.

The geodesic is then computed for a succession of reasonably spaced values of w extending throughout the zone in which the chosen formulas for $r(w,w_0)$ and $T(w,w_0,T_0)$ apply. When the geodesic reaches infinity or impinges on the surface of true geometrical singularity, $v^2-u^2=1$, the problem at that end of the geodesic is ended. When the other end comes to the boundary of two zones, the time variable $T(w,w_0,T_0)$ becomes infinite, though u and v individually remain finite. Then it becomes necessary to join on to the appropriate expression valid in the next zone, by suitable matching of the two constants of integration, in such a way as to preserve continuity of function and derivative in the u, v plane. In this way, the curves were constructed which are presented in Fig. 6.

The courses of rays 2, 3, 4, and their mirror images 8, 7, 6 of Fig. 6 were evaluated by hand computer using the formulas in Tables II and III.

The space-like geodesic 3, for example, starts at

$$u=\bar{u}, \qquad v=\bar{v}=0, \quad (12)$$

TABLE II. Classification of regimes in analysis of geodesics in Schwarzschild metric in accordance with the numbering of the zones in the insert in Fig. 5. Light-like goedesics are not included because they are represented in Kruskal's coordinates as simple straight lines inclined at 45° to the u and v axes.

Zone in Fig. 5	Value of r	Sheet in Fig. 2 (at $T=0$ or $v=0$)	Sign of u^2-v^2	Sign of u	Sign of v	Goedesics: Space-like	Goedesics: Time-like	Formula for $\begin{cases}u\\v\end{cases}$
I	$r>2m^*$	upper sheet	+	+	±	I_s	I_t	$(r/2-1)^{1/2}e^{r/4}\begin{cases}\cosh(T/4)\\\sinh(T/4)\end{cases}$
II	$2m^*>r>0$	not in sheet	−	±	+	II_s	II_t	$(1-r/2)^{1/2}e^{r/4}\begin{cases}\sinh(T/4)\\\cosh(T/4)\end{cases}$
III	$r>2m^*$	lower sheet	+	−	±	III_s	III_t	$-(r/2-1)^{1/2}e^{r/4}\begin{cases}\cosh(T/4)\\\sinh(T/4)\end{cases}$
IV	$2m^*>r>0$	not in sheet	−	±	−	IV_s	IV_t	$-(1-r/2)^{1/2}e^{r/4}\begin{cases}\sinh(T/4)\\\cosh(T/4)\end{cases}$

and crosses from zone I into zone IV, as w reaches w_0, at the point

$$u=-v=\sinh w_0 \exp(\tfrac{1}{2}-\tfrac{1}{4}T_0-D), \tag{13}$$

where

$$D=(\sinh w_0/2\cosh^3 w_0)\times[\sinh w_0\cosh w_0+(1+2\cosh^2 w_0)w_0], \tag{14}$$

with the slope

$$dv/du=\tanh(\tfrac{1}{4}T_0+D). \tag{15}$$

The formulas for u and v valid in the next zone, IV, yield the same formulas for the value and slope so there is no problem of matching constants of integration at the boundary; they match automatically. This geodesic then continues through zone IV until it reaches the boundary to zone III. Here, w reaches $-w_0$ and the coordinates on the geodesic are

$$-u=-v=\sinh w_0 \exp(\tfrac{1}{2}+\tfrac{1}{4}T_0-D). \tag{16}$$

The slope at this crossover point is

$$dv/du=\tanh(\tfrac{1}{4}T_0-D). \tag{17}$$

Once in zone III this geodesic goes off to infinity.

For example, Table IV lists a few points on geodesic 3.

The limiting space-like geodesics were even easier to work out. One limit is formed by light-like geodesics 1 and 9. The light-like geodesic 1 starts at

$$\begin{aligned}u&=\bar{u},\\v&=\bar{v}=0,\end{aligned} \tag{18}$$

and follows the straight line equation,

$$u=\bar{u}+v, \tag{19}$$

TABLE III. Integration of equations of geodesic in appropriate parameters.

Substitution A (Time-like geodesics)	Total energy$/mc^2=(dT/d\tau)_{r=\infty}=\sin w_0<1$
$r=2\cos^2 w/\cos^2 w_0$	$dr=-4\sin w\cos w dw/\cos^2 w_0$
$\tau=\tau_0+2(w+\sin w\cos w)/\cos^3 w_0$	$d\tau=4\cos^2 w dw/\cos^3 w_0$
$T=T_0+2\sin w_0\sin w\cos w/\cos^3 w_0+2\sin w_0(1+2\cos^2 w_0)w/\cos^3 w_0$ $+2\ln\lvert\sin(w_0+w)/\sin(w_0-w)\rvert$	$dT=\dfrac{4\sin w_0\cos^4 w dw}{\cos^3 w_0(\cos^2 w-\cos^2 w_0)}$
Substitution B (Time-like geodesics)	**Total energy$/mc^2=(dT/d\tau)_{r=\infty}=\cosh w_0>1$**
$r=2\sinh^2 w/\sinh^2 w_0$	$dr=4\cosh w\sinh w dw/\sinh^2 w_0$
$s=s_0+2(w+\sinh w\cosh w)/\sinh^3 w_0$	$ds=4\cosh^2 w dw/\sinh^3 w_0$
$T=T_0+2\cosh w_0\sinh w\cosh w/\sinh^3 w_0+2\cosh w_0(2\sinh^2 w_0-1)w/\sinh^3 w_0$ $-2\ln\lvert\sinh(w_0+w)/\sinh(w_0-w)\rvert$	$dT=\dfrac{4\cosh w_0\sinh^4 w dw}{\sinh^3 w_0(\sinh^2 w-\sinh^2 w_0)}$
Substitution C (Space-like geodesics)	**(Total energy$/mc^2)^2=(dT/d\tau)^2_{\text{asym}}=-\sinh^2 w_0<0$**
$r=2\cosh^2 w/\cosh^2 w_0$	$dr=4\cosh w\sinh w dw/\cosh^2 w_0$
$s=s_0+2(w+\sinh w\cosh w)/\cosh^2 w_0$	$ds=4\cosh^2 w dw/\cosh^3 w_0$
$T=T_0+2\sinh w_0\sinh w\cosh w/\cosh^3 w_0+2\sinh w_0(1+2\cosh^2 w_0)w/\cosh^3 w_0$ $-2\ln\lvert\sinh(w_0+w)/\sinh(w_0-w)\rvert$	$dT=\dfrac{4\sinh w_0\cosh^4 w dw}{\cosh^3 w_0(\cosh^2 w-\cosh^2 w_0)}$

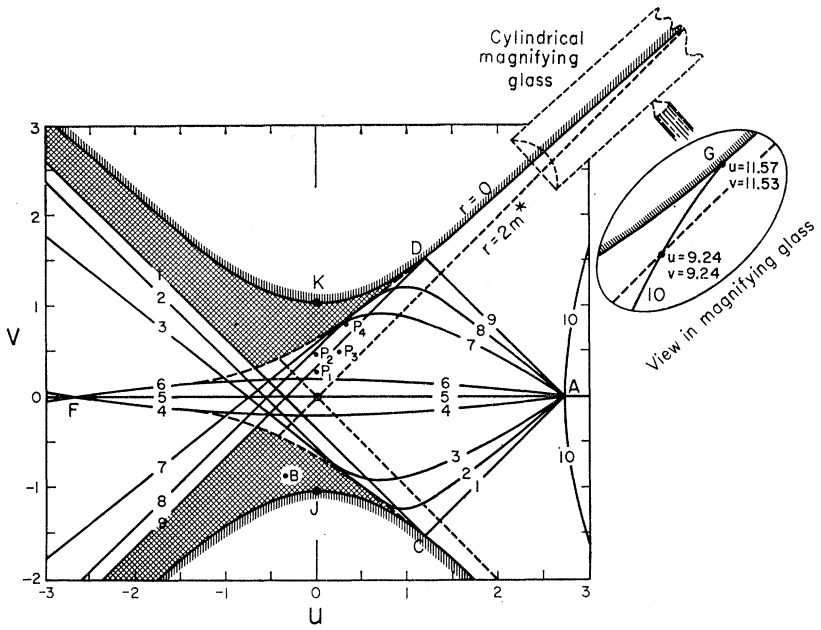

FIG. 6. Shadow zone (cross-hatched) unreachable by any radial geodesic that issues from A. Radial geodesics are plotted here in terms of Kruskal's space-like and time-like coordinates u and v, such that the Schwarzschild metric is $ds^2=f^2(du^2-dv^2)$ with $f^2=(32m^{*3}/r)\exp(-r/2m^*)$ and $[(r/2m^*)-1]\times\exp(r/2m^*)=u^2-v^2$. Space-like geodesics: 2, 3, 4, 5, 6, 7, 8. Limiting cases of space-like geodesics give light-like geodesics 1 and 9. There exists no radial geodesic that connects A and a typical point B in the shadow zone. Space-like geodesics from A never strike the region $r=0$ (shaded hyperbolas $v^2-u^2=1$) where the intrinsic curvature becomes infinite. A time-like geodesic through A, on the other hand (path of ideal test particle) always runs into this barrier on at least one end, and runs into this barrier at both ends when the energy of the test particle is insufficient for escape from the center of attraction. Geodesic 10 corresponds to the special case where the test particle reaches its maximum Schwarzschild coordinate r just at A itself. This geodesic ends at point G not tangent to the hyperbola in the (u,v) plane. An observer far away in the region to the right will never receive light signals from P_1, P_2, P_3, P_4 no matter how long he waits. The interval of proper time from J to K is however perfectly finite $(2\pi m^*)$.

down to its intersection with the hyperbola

$$v^2-u^2=1. \tag{20}$$

The intersection occurs at

$$\begin{aligned} u&=\tfrac{1}{2}(\bar{u}-\bar{u}^{-1}),\\ v&=\tfrac{1}{2}(-\bar{u}-\bar{u}^{-1}). \end{aligned} \tag{21}$$

The new light-like geodesic issuing from this point with slope -1 in the (u,v) plane has the equation

$$u=-\bar{u}^{-1}-v, \tag{22}$$

and crosses the horizontal line $v=0$ at the "focus"

$$u=-1/\bar{u}. \tag{23}$$

TABLE IV. A few points on geodesic 3 ($w_0=0.6$).

w	-1.00	-0.70	-0.60	0	$+0.60$	1.106	1.50
Special feature	Above "focus"	Below "focus"	Cross $r=2m^*$	$r_{\min}$	Cross $r=2m^*$	A	Above u axis
r/m^*	3.39	2.24	2.00	1.42	2.00	4.00	7.87
T/m^*	-2.50	1.18	$+\infty$	-1.71	$-\infty$	0.00	3.76
u	-2.34	-0.64	-0.38	0.34	0.90	2.72	18.08
v	1.30	-0.18	-0.38	-0.84	-0.90	0.00	13.30

The other limit is associated with geodesics which deviate only very little from the horizontal line $v=0$. These geodesics are described by values of the parameter w_0 in Table III Sec. C which are very close to zero:

$$w_0=\epsilon\ll 1. \tag{24}$$

In this limit the equations for a geodesic become

$$u=(\sinh w)\exp(\tfrac{1}{2}\cosh^2 w), \tag{25}$$

$$\begin{aligned} v=\epsilon(\sinh w)\exp(\tfrac{1}{2}\cosh^2 w)&\\ \times[(\tfrac{1}{2})\sinh w\cosh w+(3w/2)-\coth w&\\ +(\tfrac{1}{2})\sinh\bar{w}\cosh\bar{w}+(3\bar{w}/2)-\coth\bar{w}]. \end{aligned} \tag{26}$$

Here $\bar{w}$ is defined by the starting point A of the geodesic:

$$\begin{aligned} \bar{u}&=\sinh\bar{w}\exp(\tfrac{1}{2}\cosh^2\bar{w}),\\ \bar{v}&=0, \end{aligned} \tag{27}$$

or

$$\begin{aligned} \bar{r}&=2\cosh^2\bar{w},\\ \bar{T}&=0. \end{aligned} \tag{28}$$

The focus of these nearly horizontal geodesics is defined by that value of the parameter w which annuls the quantity in square brackets in Eq. (26). In the special

case where the source A lies at

$$\bar{r}=4m^*, \quad \bar{T}=0, \tag{29}$$

or at

$$\bar{u}=e=2.718\cdots, \quad \bar{v}=0, \tag{30}$$

the focus F is symmetrically located to the left of the origin:

$$u_F=-e, \quad v_F=0.$$

For values of the "source" coordinate $\bar{r}$ greater than zero and less than ∞ there is also always exactly one focus, but only when the "source" coordinate is equal to $4m^*$ is the proper distance from the origin to the "image" equal to the proper distance from the origin to the "source."

Also shown in Fig. 6 is a time-like geodesic, the track of an infinitesimal test particle, for the special case where the test particle reaches its maximum Schwarzschild coordinate r just at A itself (parameter T_0 in Table III Sec. A equal to 0):

$$\begin{aligned} r&=2/\cos^2 w_0, \\ T&=0, \\ \tau&=0, \\ u&=\tan w_0 \exp(\tfrac{1}{2}\cos^2 w_0), \\ v&=0. \end{aligned} \tag{31}$$

This geodesic crosses the 45° line $u=v$ at

$$\begin{aligned} r&=2, \\ T&=\infty, \\ u=v&=\sin w_0 \exp(\tfrac{1}{2}+\tfrac{1}{4}T_0+D). \end{aligned} \tag{32}$$

$$dv/du=\coth(\tfrac{1}{4}T_0)+D),$$

where

$$D=[\sin w_0/2\cos^3 w_0] \times[\sin w_0 \cos w_0+(1+2\cos^2 w_0)w_0]. \tag{33}$$

It continues on into the region $r<2$ and arrives at $r=0$, at a point in the (u,v) plane on the hyperbola $v^2-u^2=1$, but *not* tangent to this hyperbola. Simplest of all time-like geodesics is the one that runs from $(u,v)=(0,-1)$ to $(u,v)=(0,1)$. This geodesic is described parametrically by the equations

$$\begin{aligned} &r=2\cos^2 w, \quad (w \text{ from } -\pi/2 \text{ to } +\pi/2) \\ &\tau=2(w+\sin w\cos w), \quad (\text{goes from } -\pi \text{ to } +\pi) \\ &T=0, \\ &u=0, \\ &v=\sin w \exp(\tfrac{1}{2}\cos^2 w), \quad (\text{goes from } -1 \text{ to } +1). \end{aligned} \tag{34}$$

In other words, a particle at the throat of the wormhole can live only for a proper time equal to $2\pi m^*$.

III. Envelope of the Family of Geodesics

From Fig. 6, it is apparent that the family of all space-like geodesics which pass through a given point $(\bar{r},\bar{T})$ or $(\bar{u},\bar{v})$ have an envelope. This envelope starts on the hyperbola $v^2-u^2=1$ at the point of intersection with the extreme or light-like geodesic through $(\bar{u},\bar{v})$. A typical point on this envelope is common to two neighboring geodesics, characterized by the fact that both go through $(\bar{u},\bar{v})$, one with the constant of integration w_0, the other with the constant of integration w_0+dw_0. Let the expressions for $r(w_0,w)$ and $T(w_0,w)$ in Table III be written in the form

$$r(w_0,w)=2f(w)/f(w_0), \tag{35}$$

and

$$T(w_0,w)=T_0+T_1(w_0,w). \tag{36}$$

Then the requirement that a geodesic with energy parameter w_0 shall pass through $(\bar{T},\bar{r})$ determines a value of the position parameter $w=\bar{w}$, via the equation

$$\bar{r}=2f(\bar{w})/f(w_0), \tag{37}$$

whence

$$f(\bar{w})=f(w_0)\bar{r}/2, \tag{38}$$

an implicit equation for $\bar{w}$. At the point $\bar{w}$ on the geodesic w_0 the time coordinate must have the value $\bar{T}$. This requirement fixes T_0 (Eq. 36). One finds

$$T(w_0,w)=\bar{T}-T_1(w_0,\bar{w})+T_1(w_0,w). \tag{39}$$

Consider the geodesic w_0 at the point w where it is tangent to the envelope. Tangency means that on another geodesic w_0+dw_0 there is another point $w+dw$ which has the same r, T coordinates:

$$0=dr=\frac{\partial r(w_0,w)}{\partial w}dw+\frac{\partial r(w_0,w)}{\partial w_0}dw_0, \tag{40}$$

$$0=dT=\frac{\partial T_1(w_0,w)}{\partial w}dw+\frac{\partial T_1(w_0,w)}{\partial w_0}dw_0 - \frac{dT_1(w_0,\bar{w}(w_0))}{dw_0}dw_0. \tag{41}$$

Multiply the first equation by $\partial T_1/\partial w$, the second by $-\partial r/\partial w$, and add. We obtain the result

$$\frac{\partial T_1}{\partial w}\frac{\partial r}{\partial w_0}-\frac{\partial r}{\partial w}\left[\frac{\partial T_1(w_0,w)}{\partial w}-\frac{\partial T_1(w_0,\bar{w})}{\partial w_0} -\frac{\partial \bar{w}}{\partial w_0}\frac{\partial T_1(w_0,\bar{w})}{\partial \bar{w}}\right]=0. \tag{42}$$

This is an equation for w; for that point on the geodesic w_0 which touches the envelope. The coordinates of this point are found from (35) and (36) after w itself is found. This equation is written out below for the case of space-like geodesics, in the notation

$$\begin{aligned} s&=\sinh w, & c&=\cosh w,\\ s_0&=\sinh w_0, & c_0&=\cosh w_0,\\ \bar{s}&=\sinh \bar{w}, & \bar{c}&=\cosh \bar{w}. \end{aligned} \tag{43}$$

Equation (42) for w becomes

$$\frac{-s_0^2\bar{c}^5}{\bar{s}(\bar{c}^2-c_0^2)}+\frac{c_0^2 w}{2}-\tfrac{3}{2}s_0^2 w+c_0^4 w-s_0^2c_0^2 w+\frac{c^3-3c}{2s}-\frac{c_0^2\bar{s}\bar{c}}{2}+\tfrac{3}{2}s_0^2\bar{s}\bar{c}-\frac{c_0^2\bar{w}}{2}+\tfrac{3}{2}s_0^2\bar{w}-c_0^4\bar{w}+s_0^2c_0^2\bar{w}+\frac{\bar{s}\bar{c}c_0^4}{\bar{c}^2-c_0^2}=0. \tag{44}$$

No method other than trial and error is evident for solving this equation for w. The envelope of geodesics shown in Fig. 6 was determined, not by solving this equation, but by inspection of the geodesics themselves.

PHYSICAL REVIEW VOLUME 128, NUMBER 2 OCTOBER 15, 1962

Nonideal Bose Gas at Nonzero Temperatures

W. E. PARRY AND R. E. TURNER*
The Clarendon Laboratory, Oxford, England
(Received April 24, 1962)

An extension to nonzero temperatures of the Hugenholtz-Pines procedure for the degenerate Bose gas is provided. It is shown that under certain conditions the elementary excitation spectrum must approach zero for zero momentum, at nonzero, as well as at zero temperatures. The apparent discrepancy of this result with the particular case of a charged boson gas with a uniform positive background is discussed. Finally, a suggestion is made for modifying the Bogoliubov approximation, which is particularly relevant at nonzero temperatures.

1. GRAND PARTITION FUNCTION

IT was pointed out by Hugenholtz and Pines[1] that one cannot apply, straightforwardly, field-theoretic techniques to the degenerate gas of interacting Bose particles. Because of the high probability of occupation of the zero-momentum single free-particle state, the cancellation which usually occurs up to terms of order (volume)$^{-1}$ between disconnected diagrams and renormalization terms in the perturbation series for, say, the ground-state energy per particle no longer takes place. They have also provided a procedure for determining the ground state which is an extension of the Bogoliubov[2] method for low-density systems. If $a_{\mathbf{k}}^\dagger$ and $a_{\mathbf{k}}$ are the creation and annihilation operators of particles with momentum $\hbar\mathbf{k}$, we replace $a_0^\dagger$ and a_0 by a parameter $\sqrt{N_0}$, and thus obtain a Hamiltonian $H(N_0)$ which is a function of N_0. One then determines the eigenstate of $H(N_0)$ with the lowest eigenvalue E_0 subject to the subsidiary condition

$$\bar{N}'=N-N_0, \tag{1.1}$$

where $N'=\sum_{\mathbf{k}}{}' a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$; the prime on the summation symbol indicating the value $k=0$ is to be omitted; the bar indicates a quantum-mechanical average; and N is the total number of particles in the original system. The parameter N_0 must then be determined in such a way that E_0 is minimal.

Since $H(N_0)$ does not commute with N', it is insufficient to impose this condition on the unperturbed wave functions. The simplest way of including it is to introduce an undetermined multiplier $\mu_{\text{H.P.}}$, to find the ground state of the Hamiltonian

$$H'=H(N_0)-\mu_{\text{H.P.}}N', \tag{1.2}$$

without any subsidiary condition, and to determine $\mu_{\text{H.P.}}$ from

$$\bar{N}'=N-N_0.$$

N_0 is then again determined by the condition

$$(d/dN_0)(E_0)=0. \tag{1.3}$$

Although it has not been proved rigorously that this whole procedure leads to correct results, Misawa[3] has shown that to third order in the interaction potential it gives correct results for E_0, and one of us (W.E.P.) has extended this calculation to fourth order, with the same conclusion. We also note that by this method one obtains the same results as Beliaev,[4] who used a different procedure.

One would next like to extend this idea to the determination of thermodynamic functions at nonzero

* Pressed Steel Research Fellow.
[1] N. M. Hugenholtz and D. Pines, Phys. Rev. **116**, 489 (1959).
[2] N. N. Bogoliubov, J. Phys. (U.S.S.R.) **11**, 23 (1947).
[3] S. Misawa, Prog. Theoret. Phys. (Kyoto) **24**, 1224 (1960).
[4] S. T. Beliaev, J. Exptl. Theoret. Phys. (U.S.S.R.) **34**, 417 (1958) [translation: Soviet Phys.—JETP **7**, 289 (1958)].