Thermal Geons

EDWIN A. POWER, *Department of Mathematics, University College, London*

AND

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

1. INTRODUCTION AND SUMMARY

THE well-established classical theory of fields with zero rest mass has rich consequences, many of which still stand unexplored today (Table I), forty years after Einstein's formulation of the general relativity theory of electromagnetism and gravitation.

One new consequence appeared in a recent investigation.[1] An electromagnetic field, or a neutrino field, or a mixed field, of appropriate character and sufficient energy density can hold itself together, it was calculated, by its own gravitational attraction for a time long in comparison with the characteristic periods of the field oscillations. Or a the gravitation field, sufficiently strong, can guide an electromagnetic or neutrino wave and confine its energy to a bounded region of space. When the energy of the standing wave is great enough, it has enough mass to provide the guiding gravitational field all by itself. The wave holds itself together. Thus the gravitational and electromagnetic field equations of general relativity admit self-consistent solutions of great variety, many of which describe a reasonably stable concentration of energy. Such a "gravitational electromagnetic entity" or "geon" is endowed with mass, has a characteristic decay rate in the free state, moves through space like a Newtonian "body" when subjected to fields that vary sufficiently slowly in space and time, and under the influence of stronger fields undergoes transmutation. Geons of mass greater than 10^{38} g and radius over 10^{11} cm are subject to classical analysis. Smaller geons, even very much smaller geons, presumably exist, but are specifically quantum objects. Their properties have not yet been investigated. The large, classical geons have no known connection with observational science. Their interest lies in what they tell of the richness of a pure field theory. *The* pure classical field theory, it would perhaps be better to say: the only well tested theory that one has for fields of zero rest mass; the theory of elementary fields, not fields associated with particles that are themselves complex; the only field theory long established in its own right, not invented to account for selected particles or for special features of nuclear forces.

From the pure field theory of the fields of zero rest mass, and the geon concept, this article aims only at a detailed analysis of one particular kind of geon, a *thermal* geon, with the following properties:

(1) The gravitational field is static and spherically symmetric; i.e., in appropriate coordinates the element of proper distance, ds, or proper cotime, $d\tau$, has the form,

$$(ds)^2=-(d\tau)^2=g_{\alpha\beta}dx^\alpha dx^\beta=e^\lambda(dr)^2+r^2[(d\theta)^2+(\sin\theta d\phi)^2]-e^\nu(dT)^2, \quad (1)$$

where λ and ν are functions of r alone, found by numerical integration.

(2) Each independent mode of vibration of the electromagnetic field is idealized to fall into one or other of two sharply separated classes. The energy of modes of the first class remains trapped for all time. The second class of vibrations carry energy freely to infinity. The one class corresponds (in the language of photon orbits) to bounded null geodesics, the other to null geodesics that lead to infinity (Fig. 1). In actuality the energy of the "bounded" modes leaks off to infinity at a nonzero rate through a refractive index barrier, after the manner of alpha-particle penetration through a nuclear potential barrier. The rate of leakage falls off exponentially as the ratio of the dimensions of the geon to the wavelength of the disturbance in question. In a thermal geon the wavelength of the average mode is so small in comparison with the geon's radius that the leakage rate is effectively zero for all but a very restricted class of vibration modes. These few modes on the orbit picture belong to geodescis close to the crossover from bounded orbits to free orbits. The outward transport of energy resulting from such modes is relatively small and is legitimately neglected in a first analysis of a thermal geon.

(3) All free modes are assigned zero energy. Each bound mode of circular frequency Ωc is idealized to have an excitation

$$m_\Omega c^2=E_\Omega=\hbar c\Omega[-1+\exp(\hbar c\Omega/T)]^{-1}. \quad (2)$$

All properties of the thermal geon are specified by the single parameter T, which has the dimensions of energy and which we call the temperature. Occurrence of the quantum of angular momentum, $\hbar$, in this formula makes assignment of boundary conditions no different in principle from that in any classical problem: It is an act from outside, it may or may not have a quantum origin, but it alters in no way the purely classical character of all the *rest* of the analysis.

It is easy to compare and contrast thermal geons with other gravitational-electromagnetic entities. In all such objects effective confinement of the energy

[1] J. A. Wheeler, Phys. Rev. **97**, 511 (1955).

TABLE I. Qualitative comparison of the present state of charge- and mass-free gravitation and electromagnetism with hydrodynamic theory showing how many features of the Einstein-Maxwell field still remain to be explored.

Hydrodynamics	Gravitation plus electromagnetism
1. Divergence condition; pressure-density relation; and equations of motion.	1. Maxwell's source-free equations in curved space and Maxwell's stress-energy tensor as source term in Einstein's gravitational field equations.
2. Formulation in terms of a single action principle.	2. Formulation in terms of a single action principle.
3. Descriptions in alternative coordinate systems related by tensor analysis.	3. Description has same form in all coordinate systems, much as in group theory the laws of multiplication of the matrices that represent group elements are independent of the special choice of representation.
4. Hedlund and others have made a beginning at representing the laws of mechanics in an abstract coordinate-free form.[a]	4. Nothing known to have been done to represent the Maxwell-Einstein theory in a coordinate-free form as abstract as the abstract theory of groups.
5. Two alternative formulations of equations, according as one analyzes the time change of the hydrodynamic quantities at a given space point or at a given mass point	5. Coordinate system employed, x^i, may be arbitrary, or may be given an invariant significance (A. Komar) by identification with four of the fourteen invariants, I_s, of Géhéniau and Debever.[f]
6. Degree of arbitrariness in specification of initial conditions well understood.	6. Cartan-Lichnerowicz equations must be satisfied by initial conditions, but means are not yet known to generate the general solution of these initial value requirements.[g]
7. Expansions for hydrodynamic quantities near a typical point.	7. Expansion for g_{ik} and F_{ik} and R_{ijkl} near a typical point.
8. Behavior of fluid near a stagnation point or a triple point or a vortex center.	8. Behavior of electromagnetic and gravitational field near a point of special symmetry.
9. Fluid motion generated by an elemental source or sink.	9. Schwarzschild and Reissner-Nordström solution.[h]
10. Expansion of density, pressure and other quantities near a point in a series of spherical harmonics.	10. Similar expansion in spherical harmonics for small departures from a condition of spherical symmetry seems not yet to have been given.[i]
11. Sound waves; radiation pattern related to source geometry; radiation damping.	11. Electromagnetic and gravitational waves in small amplitude approximation and their radiation pattern as related to the geometry of the source. Radiation reaction analyzed for electromagnetic but not for gravitational radiation.
12. Shock waves; Mach triple point; slip stream.[b]	12. Behavior of waves of high amplitude not yet known, even qualitatively, except in very special cases.
13. Vortex ring moving in quiet fluid maintains its identity and integrity for a long time.	13. Collections of electromagnetic or gravitational waves or both under suitable conditions hold themselves together gravitationally for long periods of time ("geons").
14. Law of motion for the center of a vortex follows from the hydrodynamic equations themselves.[c]	14. Law of motion for a geon follows from the field equations themselves; the law of motion along a geodesic does not have to be introduced as a separate postulate.
15. Turbulence under appropriate conditions describable in statistical terms.	15. Radiation—isotropic or not—under appropriate conditions also describable in statistical terms.
16. Generation of sound waves by turbulence partly studied.[d]	16. Gravitation-induced interactions between electromagnetic waves, gravitational waves and geons, and cross sections for elementary types of encounter between these objects, hardly analyzed at all so far.
17. Many special solutions of the hydrodynamic equations are known from similarity arguments or group theory or other special methods of analysis.	17. Problem of homogeneous isotropic closed radiation-filled universe and its expansion and subsequent contraction, and a few other special problems have been analyzed. Field largely unexplored.[j]
18. Many types of hydrodynamic instability have been analyzed, among them Rayleigh-Taylor instability and Holmholtz instability and Bénard cells.[e]	18. One does not know how small departures from sphericity grow with time in the problem of the expanding universe, nor in the Schwarzschild-Reissner-Nordström solution. The problem of stability analysis is practically untouched.[i]
19. One has developed a set of secondary concepts adequate to describe many derived properties of the hydrodynamic field: turbulence; vorticity; acoustic impedance; radiation flux, etc.	19. Terminology for the electromagnetic field, with its six components, is extraordinarily rich; but for the gravitational field, with twenty R_{ijkl}, the present secondary conceptual structure is very rudimentary, having as yet no terms analogous to dielectric constant, permeability, Poynting vector, radiation pressure, Thomson scattering cross section, inductance, etc.

[a] G. A. Hedlund, Bull. Am. Math. Soc. **45**, 241 (1939); also G. A. Hedlund and M. Morse, Am. J. Math. **60**, 815 (1938), and W. H. Gottschalt and G. A. Hedlund, *Topological Dynamics*, Am. Math. Soc. Colloquium Publ. **36** (Providence, Rhode Island, 1955).
[b] See, for example, W. Bleakney and A. H. Taub, Revs. Modern Phys. **21**, 584 (1949); Fletcher, Bleakney, and Taub, Revs. Modern Phys. **23**, 271 (1951).
[c] See, for example, H. Lamb, *Hydrodynamics* (Cambridge University Press, London, 1953), sixth edition, Chap. VII.
[d] M. J. Lighthill, Proc. Roy. Soc. (London) **A211**, 564 (1952); *ibid.*, **A222**, 1 (1954).
[e] See, for example, reference c, Chap.XI; also Lord Rayleigh, *The Theory of Sound* (MacMillan and Company, Ltd., London, 1896), second edition, Chap. XXI; H. Bénard, Rev. gen. sci. pures appl. **12**, 1261 (1900); Ann. chim. phys. **23**, 62 (1901).
[f] A. Komar, Proc. Natl. Acad. Sci. U. S. **41**, 758 (1955); and Ph.D. thesis, Princeton, New Jersey, 1956 (unpublished); T. Y. Thomas, Proc. Natl. Acad. Sci. U. S. **31**, 306 (1945); J. Géhéniau and R. Debever, Acad. roy. de Belgique **42**, 114 (1956).
[g] A. Lichnerowicz, *Théories relativistes de la gravitation et de l'électromagnétisme* (Masson et Cie, Paris, 1955), Chap. II.
[h] H. Reissner, Ann. Phys. **50**, 106 (1916); G. Nordström, Proc. Amsterdam Acad. **21**, 68 (1918).
[i] *Note added in proof.*—The Schwarzschild solution has been shown stable against small perturbations in a paper submitted for publication by T. Regge and J. A. Wheeler. The stability of the expanding universe against small departures from sphericity has been partially treated by E. Lifshitz.
[j] See for example, L. Landau and E. Lifschitz, *The Classical Theory of Fields* (Addison-Wesley Press, Cambridge 42, Massachusetts, 1951), Chap. XI.

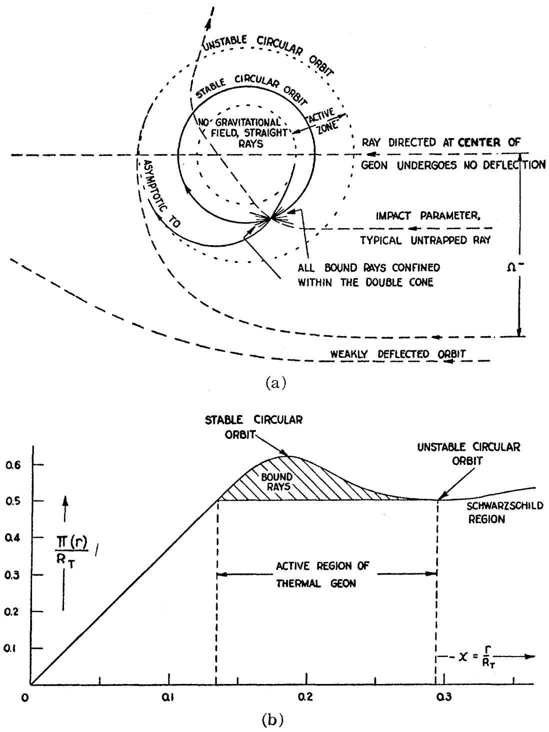


FIG. 1. (a) Schematic diagram of null geodesics or ray orbits in the gravitational field of a thermal geon. Smooth curves: bound orbits. Dashed curves: orbits that come from or go to infinity. Dotted curves: inner and outer boundaries of the active region of the geon; i.e., limits between which all bound rays circulate. (b) A ray which is to reach to the point r must have an impact parameter less than $\pi(r)$. Here $\pi(r)$ $(=re^{-\nu/2})$ is plotted as a function of r in dimensionless units. Rays that are bound must have impact parameters between $P_1=0.51R_T$ and $P_2=0.62R_T$, where R_T is a unit of length uniquely determined by the temperature.

demands localization of each mode of electromagnetic oscillation in a region where the dimensionless measure, $-g_{44}$, of gravitational potential $(=1-2\ GM/c^2r$ for Schwarzschild's point mass solution) is small compared with unity, and orientation of the flow of field energy in each mode normal to the gradient of $-g_{44}$, or in a direction not too far from normal. In contrast, a photon that travels outward along a radius vector in a spherical geon can always escape.

That region of space where $-g_{44}$ is small compared to unity, or where the rate of ticking of a standard clock is greatly slowed, is in some ways analogous to the closed container of the theory of blackbody radiation. There every mode is confined, while here only the bounded modes—or modes of semitangential energy transport—properly belong to the interior of the container. Except for this difference the varieties of geon are as numerous as the states of electromagnetic field disturbance in a hohlraum. In the general case many modes of oscillation of the field are excited, each with an amplitude of its own. Then the gravitational field resulting from the average energy distribution has no particular symmetry, and the mathematical analysis is complex. There are at least three particularly simple types of geons. (i) Toroidal geons. Here the electromagnetic energy is concentrated in the equatorial ring where the gravitational potential has its extreme value. The energy flows around the ring in equal or unequal measure in the positive and negative senses. The ratio of the fluxes in the two directions is unity or zero in the two special subcases of zero angular momentum and maximal angular momentum. Toroidal geons presumably have the greatest stability of all geons of a given mass and angular momentum. They have received partial mathematical analysis by F. J. Ernst.[2] (ii) Simple spherical geons. Here the gravitational field has nearly spherical symmetry. The electromagnetic energy is localized in a single mode of field oscillation with no radial nodes and has a large but definite number, l, of nodes in the angular variation of the field quantities. More accurately, the energy is divided over a nearly degenerate system of modes, corresponding to different spherical harmonics of the same l value but different m values. Such geons[1] require for specification two parameters: the mass and the azimuthal index number, l. The electromagnetic fields are intense in a relatively thin spherical shell, or active region. In the course of time nonlinear couplings build up the strength of originally unexcited modes, and the energy distribution slowly becomes more diffuse. (iii) Thermal geons. Here the gravitational field is again spherically symmetric, but the available electromagnetic energy is distributed over all confined modes according to the most natural of statistical laws. A thermal geon is in many ways the simplest to discuss because all its properties are fixed by a single parameter, the temperature.

Orders of magnitude and scaling laws for thermal geons are readily discussed. Let the energy density of blackbody radiation be denoted by $bT^4/\hbar^3c^3$, where $b=\pi^2/15$. The energy density in the active region is of the same order

$$\text{energy/volume}\sim Mc^2/R^3\sim bT^4/\hbar^3c^3. \tag{3}$$

The linear extension, R, of the active region is set by the requirement that $-g_{44}$ deviate substantially from unity inside:

$$GM/c^2R\sim 1. \tag{4}$$

From (3) and (4) the size is of the order

$$R\sim R_T\equiv(\hbar^3c^7/8\pi bGT^4)^{\frac{1}{2}}, \tag{5}$$

where the factor 8π has been included for later convenience; and the mass is of order

$$M\sim M_T\equiv(\hbar^3c^{11}/8\pi bG^3T^4)^{\frac{1}{2}}. \tag{6}$$

Higher temperature corresponds to geons of greater energy concentration, and of smaller mass and size.

There is a maximum temperature, and a smallest size, for which geons are free from electron pair creation and annihilation phenomena. (1) Thermal energies

[2] F. J. Ernst, Phys. Rev. 105, 1662 and 1665 (1957).

must be insufficient to creat pairs, or the temperature T must be significantly less than $2mc^2$. (2) The electric field $\mathcal{E}$ acting on an electronic charge over the characteristic localizability distance for an electron, $\hbar/mc$, must be insufficient to raise the particle to a state of positive energy:

$$\mathcal{E}<\mathcal{E}_{\text{crit}}\equiv m^2c^3/e\hbar=4.41\times10^{13}\ (\text{g/cm sec}^2)^{\frac{1}{2}}\quad(=4.41\times10^{13}\ \text{gauss}).\qquad(7)$$

Consequently the energy density of the field must be limited to values less than $\mathcal{E}_{\text{crit}}^2/8\pi$ to permit a simple analysis:

$$(\pi^2/15)T^4/\hbar^3c^3<(m^2c^3/e\hbar)^2/8\pi$$

whence

$$T<[(15/\pi^2)(137/8\pi)]^{\frac{1}{4}}mc^2=1.70mc^2(1.01\times10^{10}\ {}^\circ\text{K}).$$

Corresponding to the limit $T\sim mc^2$ on the temperature there are limits

$$\begin{aligned}&\sim(mc^2)^4/\hbar^3c^3=1.41\times10^{25}\ \text{erg/cm}^3\ \text{on energy density}\\&\sim m^4c^3/\hbar^3=1.57\times10^4\ \text{g/cm}^3\ \text{on mass density}\\&\sim(\hbar^3/Gm^4c)^{\frac{1}{2}}=9.25\times10^{11}\ \text{cm on size}\\&\sim(\hbar^3c^3/G^3m^4)^{\frac{1}{2}}=1.25\times10^{40}\ \text{g on mass.}\end{aligned}\qquad(8)$$

All analysis in this article confined to geons on the high mass side of these limits.

In the realm of sizes that are free of electron physics, geons satisfy the simple scaling law implied by (5) and (6): when one geon is hotter than another by a factor or two, it has a mass and radius four times as small; but apart from this difference of scale it has the same law of fall off of activity as does the cooler geon. For this reason the calculations are carried out in terms of the dimensionless scale independent variable

$$x=r/R_T,\qquad(9)$$

where R_T is the characteristic distance of (5). It is found that the mass has the value

$$M=0.099\ M_T,\qquad(10)$$

where M_T is the characteristic mass value of (6). The total mass $M(r)$ out to the distance r, and the dimensionless measure of this mass,

$$m(x)=M(r)/M_T,$$

are *defined* by

$$e^{-\lambda}=1-2GM(r)/c^2r=1-2m(x)/x\qquad(11)$$

The dimensionless measure of mass, $m(x)$, is shown in Fig. 2, along with the metric quantities $-g_{TT}(x)=e^\nu$ and $g_{rr}=e^\lambda$. Figures 1 and 2 contain the principal results of this paper.

The analysis is simple in outline. If the gravitational field is temporarily assumed known, then the information is at hand to set up Maxwell's equations for characteristic vibrations of the electromagnetic field. Each solution has a well-determined proper frequency, Ωc, and is therefore assigned by Planck's formula a definite amount of energy. Distribution of this energy in space follows from the form of the field eigenfunction. Summation over all characteristic modes then gives the total energy density—and similarly the total stress—in the field at any specified point in space. The stress-energy density so found—or its time average value—constitutes the entire source for the original static gravitational field. One arrives at a coupled system of equations for the gravitational field and for the characteristic vibrations of the electromagnetic field.

In practice the number of characteristic modes of appreciable excitation is so enormous that a statistical treatment of the characteristic modes is more appropriate than any detailed solution of Maxwell's equations. The Fermi-Thomas picture of electrons in an atomic field bears a relation to the Schrödinger equation like the connection between the statistical analysis of optical rays and the solutions of Maxwell's equations in a gravitational field. In both cases the JWKB approximation method links the wave point of view to the particle or ray picture. In Sec. 2 Maxwell's equations in a spherically symmetric metric are separated into an angular and a radial part. The solution of the radial part in the JWKB approximation is found to depend upon a single constant, the impact parameter P. For a ray which can escape to infinity this distance is defined as the asymptotic separation of the ray and a parallel ray which comes straight through the center of the geon without deflection. For a trapped ray, the concept of ray provides an idealized description of a mode of vibration of the electromagnetic field. The amplitude is large and oscillatory in the region of trapping, falls off exponentially outside the zone of confinement, but with still greater distance starts again to oscillate, corresponding to a weak leakage wave that runs off to infinity. This leakage wave, described in ray language, has an impact parameter, P, that is a property of the mode of vibration as a whole. Bound rays in a thermal geon have impact parameters between the two limits (Fig. 1),

$$P_1=3^{\frac{3}{2}}GM/c^2=0.51R_T$$

and

$$P_2=0.62R_T,\qquad(12)$$

and move always between the limits

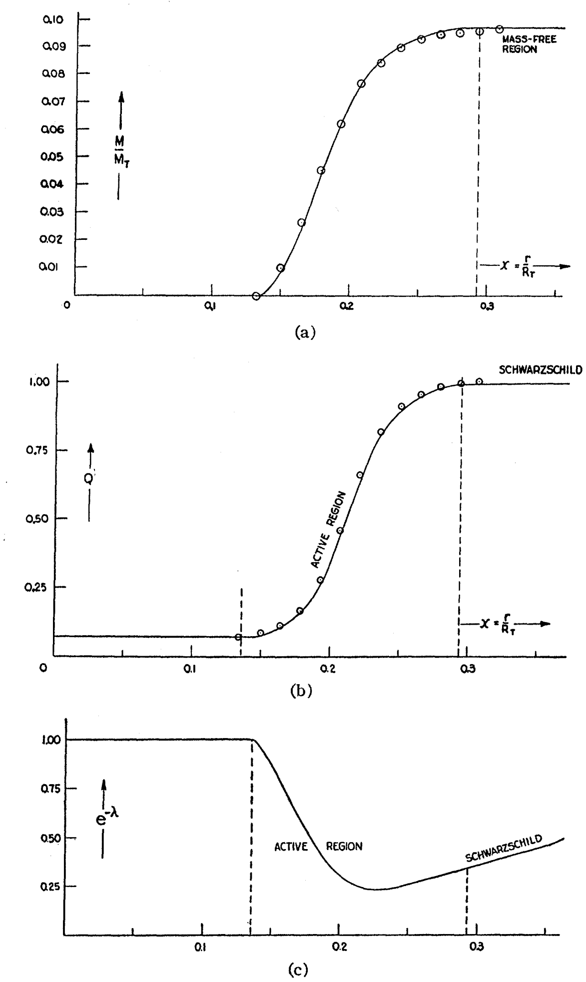
$$R_{\min}=0.14R_T\ \text{and}\ R_{\max}=0.30R_T.\qquad(13)$$

Every point in space is characterized by a critical impact parameter,

$$\pi(r)=r\exp[-\tfrac{1}{2}\nu(r)].\qquad(14)$$

No ray with impact parameter greater than $\pi(r)$ can arrive at r. Trapped rays are associated with the shaded region in Fig. 1. Among these, the bound null geodesics, there is exactly one which is circular. Its radius is

$$R_{\text{circle}}=0.19R_T\qquad(15)$$



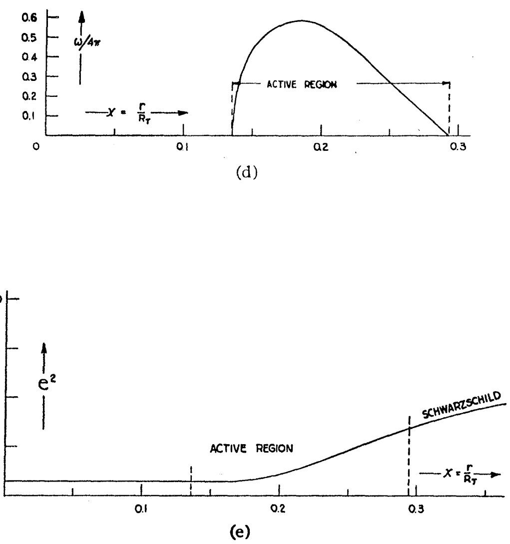


FIG. 2. Self-consistent solution of the problem of the thermal geon. The circles represent the results of numerical integration for the trial value of the eigenvalue parameter closest to giving a function $m(x)$ having the proper analytic behavior, as given by (86) and (89), at the two limits, and the smooth curves represent an estimate of the course of the functions for the correct eigenvalue. The quantity e^{ν} can be regarded as a measure of the gravitational potential that binds the photon in the orbits of Fig. 1. The ordinates in graphs (b) and (e) are Q^2 and e^{ν}, respectively.

and its impact parameter is

$$\pi(R_{\text{circle}}) = P_{\text{circle}} = P_2 = 0.62 R_T \tag{16}$$

Circular motion at the radius $R = R_{\max}$ is also possible, but is unstable. On receiving small disturbance the orbit spirals inward or outward from the circle, slowly at first and then with steeper pitch. This orbit, with impact parameter $P = P_1$, marks a boundary between orbits that can be trapped ($P_2 \geqslant P \geqslant P_1$) and those that can not.

The impact parameter $P = P_1$ gives the bound orbit with the greatest possible range of excursion in r, from $r = R_{\min}$ to $r = R_{\max}$. The bound orbit with the maximum possible impact parameter, $P = P_2$ has the minimum radial excursion: none at all.

It is easy to discuss the behavior of rays of impact parameter $P = P_1$ near the outer limit, $r = R_{\max}$, of the region of trapping. All the mass M of the system lies within the radius $R_{\max}$, and the total angular momentum of all the radiation mass-energy is zero. Consequently the metric from the point $r = R_{\max}$ to infinity has the Schwarzschild value,

$$e^{\nu} = e^{-\lambda} = 1 - (2GM/c^2 r); \tag{17}$$

and throughout this same region the critical impact parameter associated with any point r is

$$\pi(r) = r[1 - 2GM/c^2 r]^{-\frac{1}{2}}. \tag{18}$$

The minimum value of this quantity and the value of r at which it occurs have to be identified with the coordinates of the critical turning point, P_1, $R_{\max}$, in Fig. 1. Thus, a ray coming from outside with this critical impact parameter, like a ray coming from inside, must approach the point $r = R_{\max}$ without ever reaching it, in the manner of a spiral asymptotic to a circle. By differentiation of (18) we find

$$R_{\max} = 3GM/c^2$$
$$P_1 = \pi(R_{\max}) = 3^{\frac{3}{2}} GM/c^2. \tag{19}$$

The comparison between the two extreme bound orbits is striking (Fig. 1). One is exactly a circle of radius $r = R_{\text{circ}}$. The other, starting as a straight line tangent to the smaller circle $R_{\min}$, moves out to larger r, gradually curving as it goes, crosses the circle $r = R_{\text{circ}}$, and ends up at the circle $r = R_{\max}$ after an infinite number of convolutions.

Let one describe a sphere of a specific radius r between

$R_{\min}$ and $R_{\max}$ and ask which of the bound orbits cut this sphere. Not all of them, unless r happens accidentally to equal R_{circle}. A typical trapped ray ($P_2 \geqslant P \geqslant P_1$) will or will not reach the distance r according as its impact parameter is less than or greater than the critical impact parameter, $\pi(r)$, associated with the point r. The extremal ray $P=\pi(r)$, that can barely touch the sphere moves at *that* part of its orbit perpendicular to the axis. A ray of a slightly smaller impact parameter cuts across the sphere of radius r both on its outward excursion and on its return, both times with only a small velocity component normal to the surface. A bound ray of minimum impact parameter P_1, cuts this sphere closer to normality than does any other bound ray. Rays of this impact parameter observed at the point r will have all azimuths. Including both outgoing and reflected portions of orbits, these rays define a double cone. This double cone bounds the bundle of directions filled out by all the trapped rays ($\pi(r) \geqslant P \geqslant P_1$). This bundle embraces the largest solid angle when $r=R_{\text{circle}}$; i.e., when $\pi(r)=\pi_{\text{max bound}}=P_2$. The bundle narrows down to a flat disk perpendicular to the radius vector when r approaches either of the limiting values $R_{\min}$ or $R_{\max}$; i.e. $\pi(r)=P_1$.

For a quantitative measure of ray direction we define an angle of inclination, $\alpha=0$, for a ray (unbound!) that travels parallel to the radius vector. For any other ray the inclination, α, depends upon the impact parameter P that characterizes the whole course of that ray and the point r at which that ray is observed:

$$\tan\alpha=\frac{d\ (\text{proper distance}\perp r)}{d\ (\text{proper distance}\ \| r)}$$
$$=rd\theta/e^{\lambda/2}dr=[\pi^2(r)/P^2-1]^{-\frac{1}{2}}; \tag{20}$$
$$\sin\alpha=P/\pi(r).$$

The bundle of bound rays has, at the point r, angles of inclination that range from

$$\alpha=\arcsin[P_1/\pi(r)] \tag{21}$$

for outgoing rays of the minimum impact parameter P_1, through $\alpha=\pi/2$ for rays of the local extremal impact parameter $\pi(r)$, to

$$\alpha=\pi-\arcsin[P_1/\pi(r)] \tag{22}$$

for returning rays of the minimum impact parameter, P_1. The allowed rays fill a fraction of the entire solid angle given by

$$\omega/4\pi=\cos\alpha=[1-P_1^2/\pi^2(r)]^{\frac{1}{2}}. \tag{23}$$

Normal blackbody radiation would have the energy density

$$bT^4/\hbar^3c^3=(\pi^2/15)T^4/\hbar^3c^3. \tag{24}$$

The actual energy density is less because only a portion of the solid angle is filled with blackbody radiation (23) and because the effective temperature at the point r is not T, but

$$T/(-g_{44})^{\frac{1}{2}}=Te^{-\frac{1}{2}\nu(r)}, \tag{25}$$

according to Tolman.[3] Thus the energy per unit volume is

$$T_T{}^T=(bT^4/\hbar^3c^3)(\omega/4\pi)e^{-2\nu}. \tag{26}$$

The same result is obtained from first principles in §3 by evaluating the energy density of each individual bound mode of vibration of the electromagnetic field and summing over all bound modes. The JWKB approximation is used, together with the wave-ray correspondence. All summations are replaced by integrations in view of the enormous number of proper vibrations of appreciable energy content. A similar calculation gives for the radial component of the electromagnetic stress energy tensor

$$T_r{}^r=(1/3)(bT^4/\hbar^3c^3)(\omega/4\pi)^3e^{-2\nu}. \tag{27}$$

The circumstance that one has in principle to deal with a practically infinite number of eigenvalue problems in treating the trapped modes of vibration, far from causing difficulties, helps to express the stress-energy tensor in terms of purely geometrical quantities.

In Sec. 4 the self-consistent system of equations of the geon is analyzed and solved. Thus, (26) and (27) with (23) and (14) give the source of the gravitational field—the electromagnetic stress energy tensor—in terms of the gravitational field itself. In addition we require only the law to find the gravitational field from its sources—Einstein's generalization of the equation of Newtonian theory,

$$\nabla^2\varphi=4\pi G\rho, \tag{28}$$

where φ is Newtonian potential and ρ is the mass density.

The field equations of general relativity give the second derivatives of the g_{ik} in terms of the ten source strengths T_{ik}, of which the component, T_{44}, divided by c^2 is the mass density. The metric tensor, g_{ik}, is completely known except for the two dilatation functions, $\lambda(r)$ and $\nu(r)$. Consequently, two of the ten equations are sufficient to determine the potentials in terms of the source strengths:

$$R_T^2[e^{-\lambda}(r^{-1}d\nu/dr+r^{-2})-r^{-2}]=(8\pi GR_T^2/c^4)T_r{}^r$$
$$=(\tfrac{1}{3})(\omega/4\pi)^3e^{-2\nu}, \tag{29}$$

$$R_T^2[e^{-\lambda}(r^{-2}-r^{-1}d\lambda/dr)-r^{-2}]=(8\pi GR_T^2/c^4)T_T{}^T$$
$$=-(\omega/4\pi)e^{-2\nu}. \tag{30}$$

Here the expressions (26) and (27) for the energy stress tensor in terms of the solid angle have been used. In terms of the dimensionless independent variable,

[3] R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Clarendon Press, Oxford, 1934), p. 318. T corresponds to Einstein's "wahre Temperatur"; $Te^{-\frac{1}{2}\nu}$ to Ehrenfest's "Taschen-temperatur."

$x=r/R_T$; these are

$$e^{-\lambda}(x^{-1}d\nu/dx+x^{-2})-x^{-2}=(\tfrac{1}{3})(\omega/4\pi)^3e^{-2\nu}$$
$$=(\tfrac{1}{3})(1-x_1^2x^{-2}e^{\nu})^{\frac{3}{2}}e^{-2\nu} \quad (31)$$
$$e^{-\lambda}(x^{-2}-x^{-1}d\lambda/dx)-x^{-2}=-(\omega/4\pi)e^{-2\nu}$$
$$=-(1-x_1^2x^{-2}e^{\nu})^{\frac{1}{2}}e^{-2\nu}. \quad (32)$$

These are the two differential equations for the self-consistent solution of the thermal geon problem.

In (31) and (32) the quantity x_1 is an eigenvalue parameter which measures the critical impact parameter P_1 in dimensionless units. Also it determines what mass M is required to give stability to a thermal geon of temperature T:

$$x_1\equiv P_1/R_T=(3^{\frac{3}{2}}GM/c^2)/(GM_T/c^2)=3^{\frac{3}{2}}M/M_T. \quad (33)$$

This parameter has to be chosen in such a way as to yield a solution that is acceptable in the following sense:

(1) The metric dilatation parameters must join on smoothly to the Schwarzschild values at the outer boundary of the active region; (2) The metric must become flat at the inner boundary of the active region. Specifically, from Eqs. (17), (18), and (19), the conditions at the outer join point, $x=x_{\max}$, are

$$x_{\max}=x_1/3^{\frac{1}{2}},$$
$$e^{\nu}=\tfrac{1}{3};\quad d\nu/dx=3^{\frac{1}{2}}(2/x_1), \quad (34)$$
$$e^{\lambda}=3;\quad d\lambda/dx=-3^{\frac{1}{2}}(2/x_1).$$

At the inner join point, $x=x_{\min}$, the boundary conditions are

$$(xe^{-\frac{1}{2}\nu})\text{ at }x_{\min}=x_1\text{ (definition of }x_{\min})$$
$$d\nu/dx=d\lambda/dx=0. \quad (35)$$

For convenience in numerical integration of the self-consistent equations these boundary conditions on $\lambda(x)$ were expressed as boundary conditions on a variable, $m(x)$—the effective mass up to a distance x, defined by the question

$$e^{-\lambda}=1-2m(x)/x. \quad (36)$$

At $x_{\max}=x_1/3^{\frac{1}{2}}$ we have $m(x_{\max})=(\frac{1}{3})x_{\max}$ and $(dm/dx)=0$; at the inner point both $m(x)$ and (dm/dx) are zero. The differential equations were integrated with an electronic digital computer starting at a specific inner join point and proceding outwards until either $m(x)=x/3$ or $dm/dx=0$. In the former case the deviation from flatness was noted, the inner join point made larger, and the integration repeated from the start. When $m(x)$ became flat before reaching the value $(\frac{1}{3})x_{\max}$, the inner join point was made smaller and again the integration repeated. Figures 1 and 2 report the results of the integration for the best value of x_1 found with the limited machine time available

$$x_1=0.51. \quad (37)$$

Details of the present analysis of thermal geons follow.

2. THE PROPER VIBRATIONS OF THE ELECTROMAGNETIC FIELD

Analysis In Spherical Harmonics

The electromagnetic field $F_{jk}(F_{23}\sim H_x;\ F_{14}\sim E_x)$ satisfies the eight source-free Maxwell equations

$$(-g)^{-\frac{1}{2}}(\partial/\partial x^{\alpha})(-g)^{\frac{1}{2}}F^{i\alpha}=0 \quad (38)$$

and

$$\{i\alpha\beta\gamma\}\partial F_{\alpha\beta}/\partial x^{\alpha}=0. \quad (39)$$

Here the numbers $\{ijkl\}$ are defined by $\{1234\}=1$ and by $\{ijkl\}$ changing sign on reversal of any two indexes. Also g is the determinant of the g_{ik}: $g=-r^4\sin^2\theta\times\exp[\lambda(r)+\nu(r)]$. In the static spherically symmetric gravitation field the Maxwell equations are invariant with respect to a group of transformations built from the following elementary operations: (1) rotation of the space frame of reference; (2) translation of time; (3) inversion of the space frame in the center of symmetry; (4) reflection in a plane through this center; (5) reversal of time; (6) interchange of the roles of electric and magnetic fields, according to the substitution

$$F_{jk}{}^{II}=\tfrac{1}{2}(-g)^{\frac{1}{2}}\{ik\alpha\beta\}g^{\alpha\mu}g^{\beta\nu}F_{\mu\nu}{}^{I}, \quad (40)$$

such that $E^{II}\sim H^{I}$, $H^{II}\sim -E^{I}$. For a flat space time continuum one knows the irreducible representations of this group of transformations and the basic set of functions on which these transformations operate,[4] and it is very easy to modify the results for the present spherically symmetrical curved static space time continuum. The basic solutions are characterized by a circular frequency, Ω/c, by two angular momentum quantum numbers, l and m, and by the statement that the disturbance is of electric or magnetic multipole character, as the case may be, in the following sense: (1) Magnetic multipole field: the electric field is everywhere exactly perpendicular to the radius vector; the parity of the electric field with respect to space inversion is $(-1)^l$; the parity of the magnetic field is $(-1)^{l+1}$; and the field is expressed in terms of a four potential A_i

$$F_{jk}=\partial A_k/\partial x^j-\partial A_j/\partial x^k, \quad (41)$$

which lies always on the surface of a sphere with center at the origin, and points always in the direction of the electric field. Thus $A_r=A_T=0$, and the space vector $\mathbf{A}$ is a function of r multiplied by $\mathbf{r}\times\nabla$ acting on a spherical harmonic. The normalized spherical harmonic[4] is denoted by $Y_l{}^{(m)}(\theta,\phi)$. We write

$$Q\equiv Y_l{}^{(m)}(\theta,\phi)\exp(-i\Omega T); \quad (42)$$

then the nonzero components of the vector potential

[4] See, for example, J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (John Wiley and Sons, Inc., New York, 1952), pp. 799 ff. for formulas and references to the original literature.

for the mode in question have the form

$$A_\theta = a(r) imQ/\sin\theta + \text{complex conjugate},$$
$$A_\phi = -a(r)\sin\theta\partial Q/\partial\theta + \text{complex conjugate}, \quad (43)$$

where the radial function $a(r)$ has the dimensions $(\text{g cm}^3/\text{sec}^2)^{\frac{1}{2}}$ or gauss cm^2 or cm electrostatic volts. The field itself, F_{jk}, is expressed relative to the coordinate system that appears in the metric (1). However, the field is more easily visualized in a local orthonormal coordinate system that is aligned along the r, θ, ϕ, T axes. In this frame we shall call the field $(\mathbf{E},\mathbf{H})$. From (43) we derive

$$\begin{aligned} e^{\nu/2+\lambda/2}E_r &= F_{rT} = 0 \\ re^{\nu/2}E_\theta &= F_{\theta T} = -\Omega amQ/\sin\theta + \text{c.c.} \\ r\sin\theta e^{\nu/2}E_\phi &= F_{\phi T} = -\Omega ai\sin\theta\cdot\partial Q/\partial\theta + \text{c.c.} \\ r^2\sin\theta H_r &= F_{\theta\phi} = l(l+1)a\sin\theta Q + \text{c.c.} \\ r\sin\theta e^{\lambda/2}H_\theta &= F_{\phi r} = (da/dr)\sin\theta\cdot\partial Q/\partial\theta + \text{c.c.} \\ re^{\lambda/2}H_\phi &= F_{r\theta} = (da/dr) imQ/\sin\theta + \text{c.c.} \end{aligned} \quad (44)$$

(2) Electric multipole field: the magnetic field is everywhere exactly perpendicular to the radius vector; space inversion multiplies the electric field by $(-1)^{l+1}$, the magnetic field by $(-1)^l$. This type of field is obtained from the magnetic multipole field by two steps: (a) replace $a(r)$ in (43) by $b(r)$ and call the field F so calculated "F^{I}" (b) substitute this field into (40) and calculate the field "F^{II}." This is the desired electric multipole field:

$$\begin{aligned} F_{rT} &= l(l+1)\Omega^2 r^{-2} e^{\lambda/2+\nu/2} b(r) Q + \text{c.c.} \\ F_{\theta T} &= \Omega e^{-\lambda/2+\nu/2}(db/dr)(\partial Q/\partial\theta) + \text{c.c.} \\ F_{\phi T} &= \Omega e^{-\lambda/2+\nu/2}(db/dr) imQ + \text{c.c.} \\ F_{\theta\phi} &= 0 \\ F_{\phi r} &= e^{\lambda/2-\nu/2} b(r) mQ + \text{c.c.} \\ F_{\theta r} &= e^{\lambda/2-\nu/2} b(r) i(\partial Q/\partial\theta) + \text{c.c.} \end{aligned} \quad (45)$$

Radial Function

The radial factors $a(r)$ and $b(r)$ in the field expressions (44) and (45) satisfy a common differential equation,[5]

$$d^2f/dr^{*2} + [\Omega^2 - l(l+1)e^\nu/r^2]f = 0 \quad (46)$$

where

$$dr^* = e^{\lambda/2-\nu/2}dr.$$

We demand a solution of (46) that is regular at the origin. Near the origin λ and ν are constant, and λ is zero.[1] Consequently, f varies as r^{l+1} near the origin. At larger r, f increases roughly exponentially until the first zero of the square bracket in (46): $r=r_1$. Beyond this point the quantity

$$l(l+1)e^\nu/\Omega^2 r^2 \quad (47)$$

falls below unity, the square bracket is positive, and the solution is oscillatory. With further increase in r there are two possibilities which are familiar from the closely analogous problem of alpha decay: (1) The quantity (47) never rises above unity again. In this case the solution remains oscillatory to infinite r, ultimately approaching the behavior

$$f(r) \sim c_1 \sin(\Omega r + c_2). \quad (48)$$

Such a solution represents an electromagnetic wave that runs freely to infinity (analogous to an alpha particle with energy that exceeds the potential barrier). Since such a wave carries energy away from the geon, it is not of interest in constructing a relatively stable object. Such modes of oscillation of the electromagnetic field are assigned zero energy. (2) The quantity (47) rises above unity again at a point $r=r_2$ (analogous to inner radius of potential barrier for an alpha particle with energy below the top of the barrier). A sufficiently great increase in r will result in (47) once again falling below unity at some point, $r=r_3$. The solution, $f(r)$, of (46) rises monotonically from the origin to r_1, oscillates from r_1 to r_2, behaves between r_2 and r_3 as a linear combination of a function that falls roughly exponentially and another function that rises roughly exponentially, and resumes an oscillatory character from r_3 to ∞. The wave in this outer region transports energy away from the geon. The rate of transport is the smaller, the less is the ratio of the amplitudes of oscillation in the outer and inner regions. This ratio has a minimum value for certain characteristic values of Ω, designated as Ω_{nl}. At such a value of Ω the solution $f(r)$ decreases monotonically and roughly exponentially all the way from r_2 to r_3. For such characteristic solutions the region from $r_1(\Omega_{nl})$ to $r_2(\Omega_{nl})$ is called the region of activity. As a first approximation exponential fall-off is supposed to continue indefinitely beyond the point r_2.

Eigenvalues, Eigenfunctions, and Averages

The characteristic values, Ω_{nl}, of Eq. (46) in the sense just defined, are given in the JWKB approximation by the implicit equation,

$$\int_{r_1}^{r_2} [1 - l(l+1)e^\nu/\Omega_{nl}^2 r^2]^{\frac{1}{2}} \Omega_{nl} dr^* = (n+\tfrac{1}{2})\pi. \quad (49)$$

The solutions themselves[6] in the same approximation have the form

$$f_{nl}(r) = C_{nl}[1 - l(l+1)e^\nu/\Omega_{nl}^2 r^2]^{-\frac{1}{4}} \cdot \sin\left\{\int_{r_1}^{r} [1 - l(l+1)e^\nu/\Omega_{nl}^2 r^2]^{\frac{1}{2}} \Omega_{nl} dr^* + \frac{\pi}{4}\right\}. \quad (50)$$

---

[5] Reference 1, p. 520. See also A. Eddington, *The Mathematical Thoery of Relativity* (Cambridge University Press, 1923), p. 175, where the vector wave equation is written in the form $\Box A^m + R_\alpha{}^m A^\alpha = 0$, with $\Box\psi \equiv g^{\alpha\beta}\psi_{;\alpha\beta}$ (equivalent to our $(\partial/\partial x^\alpha) \times (-g)^{\frac{1}{2}} F^{m\alpha} = 0$).

[6] See, for example W. Pauli, *Handbuch der Physik* (Verlag Julius Springer, Berlin, 1933), second edition, Vol. 24, Part 2, p. 171.

The value of $|f|^2$, averaged over a region appreciable in comparison with one wavelength, but very small relative to the size of the region of activity, is

$$\langle |f|^2\rangle \doteq \tfrac{1}{2}|C_{nl}|^2[1-l(l+1)e^\nu/\Omega_{nl}^2r^2]^{-\frac{1}{2}} \qquad (51)$$

inside the active region, and practically zero outside. This measure of intensity, like the limits of activity, depends, not upon l and Ω individually, but only upon these two quantities in the single combination of an *impact parameter*,

$$P\equiv[l(l+1)/\Omega^2]^{\frac{1}{2}}. \qquad (52)$$

At a given point in the active region of the geon this quantity can have any value between the minimum value P_1 for binding (Fig. 1) and the maximum value, $\pi(r)$ $(=re^{-\nu/2})$, appropriate to that point.

The number of bound modes of vibration that contribute to the electromagnetic field intensity at the point r in the active region is obtained by taking the product of the following factors and summing:

2,	for polarizations
$2l+1$,	for values of m
dl,	for values of l
dn,	for values of n.

Summations over l and n are replaced by integrations and then by integrations over all the relevant values of the impact parameter P between P_1 and $\pi(r)$ and over the circular cofrequency Ω by virtue of equations (30), and (33). One finds that

$$dldn=\frac{\partial(l,n)}{\partial(P,\Omega)}=\pi^{-1}\Omega\int[1-P^2e^\nu/r^2]^{-\frac{1}{2}}dr^*dPd\Omega,$$

and consequently the total number of modes within specified limits of Ω and P is

$$dN=(2/\pi)\Omega^2d\Omega d(P^2)\int[1-P^2e^\nu/r^2]^{-\frac{1}{2}}dr^*. \qquad (53)$$

It is of interest to have not only the local average (51) of $|f|^2$, but also the local average of the square of its derivative

$$\langle|df/dr^*|^2\rangle\doteq\tfrac{1}{2}|C_{nl}|^2\Omega_{nl}^2[1-l(l+1)e^\nu/\Omega_{nl}^2r^2]^{\frac{1}{2}}. \qquad (54)$$

The symbol $\langle\langle\ \rangle\rangle$ denotes an average with respect to time and with respect to position over a spherical shell of thickness large compared to a typical wave length ($\sim10^{-11}$ cm) but small compared to the dimension of the active region ($>\sim10^{11}$ cm). We note

$$\langle\langle|Y_l^{(m)}|^2\rangle\rangle=\frac{1}{4\pi};$$

$$\langle\langle|\partial Y_l^{(m)}/\partial\theta|^2+m^2\sin^{-2}\theta|Y_l^{(m)}|^2\rangle\rangle=l(l+1)/4\pi. \qquad (55)$$

For a typical mode of the form (43) and (44),

$$\left.\begin{aligned}\langle\langle F_{\theta\phi}F^{\theta\phi}\rangle\rangle&=P^2e^\nu/r^2\\ \langle\langle F_{\phi r}F^{\phi r}+F_{r\theta}F^{r\theta}\rangle\rangle&=(1-P^2e^\nu/r^2)\\ \langle\langle F_{\theta T}F^{\theta T}+F_{\phi T}F^{\phi T}\rangle\rangle&=-1\end{aligned}\right\}\begin{aligned}&\text{times}\\ &|C_{nl}|^2\{l(l+1)/4\pi r^2\}\\ &\Omega^2e^{-\nu}[1-P^2e^\nu/r^2]^{-\frac{1}{2}}.\end{aligned}$$

The sum, $\frac{1}{2}F_{\alpha\beta}F^{\alpha\beta}=H^2-E^2$, of the quantities on the left represents the action invariant for the field oscillator. The action disappears on the time average for this oscillator as it does for every harmonic oscillator.[7]

3. ENERGY-DENSITY STRESS-DENSITY TENSOR

Radial Stress and Energy Density of One Mode

The source of the gravitational field is the stress-energy tensor of the electromagnetic field, with components

$$T_i^{\ k}=(4\pi)^{-1}F_{i\alpha}F^{k\alpha}-(16\pi)^{-1}\delta_i^{\ k}F_{\alpha\beta}F^{\alpha\beta} \qquad (56)$$

so that in flat space time $-T_{44}=T_4^{\ 4}=-(E^2+H^2)/8\pi$. In a geon with a negligible rate of leakage of energy we can assume time symmetry as well as spherical symmetry. In the r, θ, ϕ, T system of coordinates, all off-diagonal elements of the stress energy tensor vanish. Among the diagonal elements there exists the equality $T_\theta^{\ \theta}=T_\phi^{\ \phi}$, and also a general relation expressing these two tangential tensions in terms of the radial tensions—in terms of $T_r^{\ r}$, $T_T^{\ T}$, and $dT_r^{\ r}/dr$. This knowledge of $T_\theta^{\ \theta}$ and $T_\phi^{\ \phi}$ does not help in the determination of the gravitational field from the field equations

$$G_i^{\ k}=(8\pi G/c^4)T_i^{\ k} \qquad (57)$$

because there is an identity between $G_\theta^{\ \theta}$, $G_\phi^{\ \phi}$ and $G_r^{\ r}$, $G_T^{\ T}$, and $dG_r^{\ r}/dr$, which makes them automatically satisfy the same radial-tangential equilibrium condition as the corresponding components of the T's. For this reason it is enough to consider only the two remaining components of the stress-energy tensor, $T_r^{\ r}$ and $T_T^{\ T}$.

The intensity of the particular mode of field oscillation characterized in (55) results in this mode contributing to the energy density and radial stress:

$$\begin{aligned}&\langle\langle T_T^{\ T}\ \text{(one mode)}\rangle\rangle\\ &\quad=(8\pi)^{-1}\langle\langle F_{rT}F^{rT}+F_{\theta T}F^{\theta T}+F_{\phi T}F^{\phi T}\\ &\qquad\qquad-F_{\theta\phi}F^{\theta\phi}-F_{\phi r}F^{\phi r}-F_{r\theta}F^{r\theta}\rangle\rangle\\ &\quad=-(\Omega^2/16\pi^2r^2)e^{-\nu}l(l+1)|C_{nl}|^2[1-P^2e^\nu/r^2]^{-\frac{1}{2}}\\ &\langle\langle T_r^{\ r}\ \text{(one mode)}\rangle\rangle\\ &\quad=(8\pi)^{-1}\langle\langle F_{\phi r}F^{\phi r}+F_{r\theta}F^{r\theta}+F_{rT}F^{rT}\\ &\qquad\qquad-F_{\theta\phi}F^{\theta\phi}-F_{\theta T}F^{\theta T}-F_{\phi T}F^{\phi T}\rangle\rangle\\ &\quad=(\Omega^2/16\pi^2r^2)e^{-\nu}l(l+1)|C_{nl}|^2[1-P^2e^\nu/r^2]^{\frac{1}{2}}.\end{aligned} \qquad (58)$$

Many Modes; Total Mass

The electromagnetic stress energy tensor being bilinear in the field, the fields of all individual standing

[7] The integral of E^2-H^2 over all space for a field that is a superposition of two modes gives zero for the integral of the cross term in virtue of the orthogonality that follows from the eigenvalue equation (46) for radial functions of the same l and different n.

waves must be added before commencing the evaluation. For the time average value of the energy density, we assume that the elementary disturbances are incoherent[1]: the average contribution from cross terms between modes vanishes. Thus, we write

$$\langle T_T{}^T\rangle = \sum_{n,l,m,p} \langle\langle T_T{}^T(n,l,m,p\,;r,\theta,\phi)\rangle\rangle \qquad (59)$$

where the sum goes over all values of the angular index, l, and m; over all the values of the index number n of the radial proper function f_{nl} associated with the confrequency Ωnl; and over both types of polarization: p=electric or magnetic multipole. There is a similar expression for the radial stress.

There are two simple ways to get to the total mass.[8]

(1) Write

$$e^{-\lambda} = 1 - 2GM(r)/c^2r \qquad (60)$$

as a definition of the effective mass, $M(r)$; insert this expression into differential equation (30) to derive an expression for $dM(r)/dr$; integrate from $r=0$ to $r=\infty$; and find

$$Mc^2 = c^2M(r=\infty) = -\int_0^\infty \sum_{\text{modes}} \times \langle\langle T_T{}^T(\text{one mode})\rangle\rangle 4\pi r^2 dr. \qquad (61)$$

(2) Write down the expression for the local density of *gravitational plus electromagnetic* energy; integrate over all space; use the fact that the gravitational part is expressible in terms of a surface integral; use also the fact that this surface integral does not involve any details of the gravitational field when the metric is asymptotically flat; finally, use the fact that the trace of the electromagnetic stress-energy tensor is zero:

$$Mc^2 = \int\int\int \langle T_r{}^r + T_\theta{}^\theta + T_\phi{}^\phi - T_T{}^T\rangle(-g)^{\frac{1}{2}} dr d\theta d\phi$$

$$= -2\int \sum_{\text{modes}} \langle\langle T_T{}^T(\text{one mode})\rangle\rangle e^{\lambda/2+\nu/2} 4\pi r^2 dr. \qquad (62)$$

Expression (62) is more convenient than (61) because the radial integral is easier to calculate.

Contraction Effect

To define the energy of one mode, one can think of measuring the gravitational field far away from a geon before and after the stilling of that particular vibration. One might be tempted to write for this change in mass

$$c^2\Delta M = I_{\text{mode}}, \qquad (63)\ \text{(wrong)}$$

where

$$I_{\text{mode}} \equiv -2\int \langle\langle T_T{}^T(\text{mode})\rangle\rangle e^{\lambda/2+\nu/2} 4\pi r^2 dr. \qquad (64)$$

This would be wrong. As the radiation in question leaks out of the geon, less pressure is available to sustain the geon against gravitational forces. Consequently the geon contracts. Thus every other mode finds itself more tightly confined. Its frequency rises. Accordingly by the principle of adiabatic invariance there is a proportional *increase* in energy. Therefore, the mass of the geon will decrease by an amount less than given by (63).

The fractional change in mass due to stilling of a single mode is fantastically small. For a first primitive analysis the effect on the geon is primarily a scale transformation to smaller size. Mass and radius transform in parallel. Frequency goes up inversely as radius, and therefore inversely as mass. The same applies—according to adiabatic invariance—to the contribution, I_{mode}, to the energy from every mode except the selected mode:

$$\frac{\Delta I_{\text{mode}}\ \text{(increase)}}{I_{\text{mode}}} = \frac{\Delta M\ \text{(decrease)}}{M} \qquad (65)$$

(see qualification below). On this basis we compare the geon before,

$$Mc^2 = I_{\text{one}} + \sum I_{\text{others}} \qquad (66)$$

and after

$$(M-\Delta M)c^2 = 0 + \sum\left(1+\frac{\Delta M}{M}\right) I_{\text{others}}, \qquad (67)$$

the mode is stilled. Subtraction gives

$$c^2\Delta M = \frac{I_{\text{one}}}{1+(1/Mc^2)\sum I_{\text{others}}} = \tfrac{1}{2} I_{\text{one}}. \qquad (68)$$

We are considering a classical geon that has a great many proper modes, $s=1, 2, \cdots$, where s stands for the quartet (n,l,m,p). Each mode has a reduced action variable J_s (=action/2π) associated with it. The energy consists of two parts: electromagnetic energy of individual modes and gravitational energy of interaction between modes. The sum of both energies (61) is a function of the action variables of all these oscillators:

$$c^2M = c^2M\ (J_1, \cdots J_s, \cdots).$$

The circular frequency of a given oscillator is given by

$$c\Omega_s(J) = c^2\frac{\partial M(J)}{\partial J_s}. \qquad (69)$$

If a particular action variable decreases slowly from its normal value to zero, during this change all other action variables keep their original values, according to the principle of adiabatic invariance.

[8] R. C. Tolman, Phys. Rev. **35**, 875 (1930); see also L. Landau and E. Lifschitz, *The Classical Theory of Fields*, translated by M. Hammermesh (Addison-Wesley Press, Cambridge, Massachusetts, 1951), pp. 309 and 323.

Then the change in mass is

$$c^2\Delta M = c\Omega_s(J)dJ_s = \tfrac{1}{2}I_s$$
$$= -\int \langle\langle T_T{}^T \text{ (chosen mode } s)\rangle\rangle e^{\lambda/2+\nu/2}4\pi r^2 dr. \quad (70)$$

A weakness in the foregoing reasoning is that it assume that stopping of the given vibration produces a change in the field of gravitational force that is represented with sufficient accuracy simply by a change of scale. Perhaps the change in metric field will act in different ways on different modes, raising some in I value less than others. It is at least conceivable that such a differentation might take place. Since the quantity of interest is not the I value of one mode, but the sum of the changes in I values of all the modes, we need, not (65) for every mode individually, but only a weaker equation for this sum of changes:

$$\frac{\sum I_{\text{others}}(\text{increase})}{I_{\text{others}}} = \frac{\Delta M \text{ (decrease)}}{M}. \quad (71)$$

The question about correctness of (71) could be answered by an appeal back to first principles:

(1) Determine the effect on a single mode of an adiabatic change in the metric field. (a) Insert in the action principle of the electromagnetic field a trial function of the product form (43), where however the time factor, $\exp(-i\Omega T)$, is replaced by an arbitrary undetermined function of cotime, $\psi(T)$. (b) From the variational principle derive the second-order differential equation for this function. (c) Solve by the JWKB method. (d) This gives a general expression for properties of a single mode under arbitrary adiabatic changes in the Schwarzschild metric, containing one arbitrary amplitude constant, independent of both r and T.

(2) Calculate the stress energy associated with this mode, and find all over again the expressions (57), with the difference that not C_{nl}, but D_{nl}, is constant under adiabatic changes, where

$$C_{nl} = D_{nl}\Omega_{nl}{}^{-1}\left[\int [1-P^2e^{\nu}/r^2]^{-\frac{1}{2}}dr^*\right]^{-1} \quad (72)$$

(3) Write down the equations (29) and (30) of the self-consistent geon field for two problems: (a), with all modes s excited to specific metric-independent amplitudes D_s; and (b) with all modes but one so excited.

(4) By differencing these two sets of nonlinear equations derive linear equations for the small changes $\delta\lambda(r)$, $\delta\nu(r)$, produced in the metric by quenching one mode.

(5) Derive expressions for these changes in the metric.

(6) Find how these changes alter the I value of each mode individually and hence

(7) Compute the change in mass of the system as a whole due to stopping one vibration.

Test of Adiabatic Analysis against Ehrenfest-Tolman Formula

For thermal geons there is an extra simplification with which we can check the correctness of (68), without the more detailed analysis just outlined. Considering blackbody radiation in equilibrium in a gravitational field for an asymptotic temperature T, we know from Ehrenfest and Tolman[9] that the energy density has the value appropriate to the temperature $T(-g_{44})^{-\frac{1}{2}}$, rather than T itself. Here this means an energy density proportional to $T^4e^{-2\nu}$, multiplied by the fractional solid angle filled by trapped rays. This result comes from the normalization (70).

The energy leaked out of the geon by quenching one mode is equated to the thermal energy of a harmonic oscillator, as given by Planck's formula:

$$\hbar c\Omega_{nl}[-1+\exp(\hbar c\Omega_{nl}/T)]^{-1} = c^2\Delta M = \tfrac{1}{2}I_s$$
$$= -\int \langle\langle T_T{}^T \text{ (chosen mode)}\rangle\rangle e^{\lambda/2+\nu/2}4\pi r^2 dr$$
$$= (\Omega^2/4\pi)l(l+1)|C_{nl}|^2\int [1-P^2e^{\nu}/r^2]^{-\frac{1}{2}}dr^*. \quad (73)$$

Solving this for the amplitude $|C_{nl}|$, and inserting this value of $|C_{nl}|$ into Eqs. (38) we find *the energy density and radial stress due to one normalized mode*:

$$\langle\langle -T_T{}^T(\text{or}+T_r{}^r)\rangle\rangle = \hbar c\Omega_{nl}[-1+\exp(\hbar c\Omega_{nl}/T)]^{-1}$$
$$\cdot[1-P^2e^{\nu}/r^2]^{\{-\frac{1}{2}(\text{or}+\frac{1}{2})\}}(e^{-\nu}/4\pi r^2)$$
$$\times\left\{\int [1-P^2e^{\nu}/r^2]^{-\frac{1}{2}}dr^*\right\}^{-1}. \quad (74)$$

We multiply (74) by the number of modes, dN, as in (53), in the interval $d\Omega dP$, and sum over all modes by integrating over the wave number Ω and the equivalent impact parameter P. This gives the local value of the total energy density and radial stress:

$$\langle\langle -T_T{}^T(\text{or}+T_r{}^r)\rangle\rangle_{\text{local total}}$$
$$= \int \hbar c\Omega[-1+\exp(\hbar c\Omega/T)]^{-1}\Omega^2 d\Omega(\tfrac{1}{2}e^{-\nu}/r^2)$$
$$\times\int_{P_1}^{P_{\max}=re^{-\nu/2}}[1-P^2e^{\nu}/r^2]^{\{-\frac{1}{2}(\text{or}+\frac{1}{2})\}}d(P^2)$$
$$= (\pi^2T^4/15\hbar^3c^3)e^{-2\nu}\{[1-P_1{}^2e^{\nu}/r^2]^{\frac{1}{2}}$$
$$(\text{or } \tfrac{1}{3}[1-P_1{}^2e^{\nu}/r^2]^{\frac{3}{2}})\}. \quad (75)$$

In terms of the solid angle, ω, spanned by the trapped rays, the complete stress energy tensor has the form

[9] R. C. Tolman and P. Ehrenfest, Phys. Rev. **36**, 1791 (1930); see also Tolman, reference 3, p. 318.

$$\langle\langle T_j{}^i\rangle\rangle_{\text{local total}}=(\pi^2T^4e^{-2\nu}/15\pi^3c^3)\times\begin{vmatrix}\tfrac{1}{3}(\omega/4\pi)^3 & & & \\ & \tfrac{1}{2}(\omega/4\pi)-\tfrac{1}{6}(\omega/4\pi)^3 & & \\ & & \tfrac{1}{2}(\omega/4\pi)-\tfrac{1}{6}(\omega/4\pi)^3 & \\ & & & -(\omega/4\pi)\end{vmatrix}. \quad (76)$$

The stress-energy tensor calculated here by consideration of individual modes agrees with the usual blackbody value, corrected properly as demanded by the Ehrenfest-Tolman argument and by the solid angle factor. No such agreement would have resulted if the factor $\frac{1}{2}$ in (70) for the loss of mass on stilling of one mode is left out. Thus the contraction effect is essential.

Expressions (76) as derived here were used as described in the introduction to set up (29) and (30) for the self-consistent thermal geon field.

To restate the results of the present analysis, the total mass of the geon is given by the superficially paradoxical formula,

$$Mc^2=\frac{1}{2}\sum_{\text{all modes}}\begin{pmatrix}\text{Planck formula for energy}\\ \text{of one mode, as in (2)}\end{pmatrix}. \quad (77)$$

4. NUMERICAL SOLUTION OF THE DIFFERENTIAL EQUATIONS

In the dimensionless variable $x=r/R_T$ the equations for the geon field are

$$e^{-\lambda}(x d\nu/dx+1)-1=(x^2/3)(1-x_1{}^2e^\nu/x^2)^{\frac{3}{2}}e^{-2\nu} \quad (78)$$

$$e^{-\lambda}(1-x d\lambda/dx)-1=-x^2(1-x_1{}^2e^\nu/x^2)^{\frac{1}{2}}e^{-2\nu}. \quad (79)$$

Let $m(x)$ represent the effective mass, in units M_T, out to the distance $r=xR_T$, as *defined* by

$$m(x)=x(1-e^{-\lambda})/2. \quad (80)$$

Thus the mass of the geon is

$$M=m(\infty)M_T.$$

Also define a scale factor, Q, by the equation

$$Q^2=e^{\lambda+\nu}. \quad (81)$$

These variables behave as follows (Fig. 2). From $x=0$ to the radius $x=x_{\min}$, the dimensionless mass variable, m, is zero. Then it increases up to

$$m=m(\infty)=x_{\max}/3=x_1/3^{\frac{3}{2}} \quad (82)$$

at the outer boundary

$$x_{\max}=x_1/3^{\frac{1}{2}} \quad (83)$$

of the active region. Here x_1 is a measure of the critical impact parameter for trapped rays:

$$x_1=P_1/R_T. \quad (84)$$

Thereafter it remains constant at the value $m(\infty)$. The curve for m as a function of x has a horizontal slope at $x=x_{\min}$ and $x=x_{\max}$. A similar description applies to Q, with these exceptions: Q is not zero, but has a constant value between 0 and 1 for $x<x_{\min}$; and Q has the constant Schwarzschild value of unity for $x>x_{\max}$. Between these limits these quantities satisfy the equations

$$dm/dx=x^2[1-x_1{}^2x^{-2}Q^2(1-2m/x)]^{\frac{1}{2}}\div 2Q^4(1-2m/x)^2 \quad (85)$$

$$dQ^2/dx=x\{[1-x_1{}^2x^{-2}Q^2(1-2m/x)]^{\frac{1}{2}}+[1-x_1{}^2x^{-2}Q^2(1-2m/x)]^{\frac{3}{2}}/3\}\div Q^2(1-2m/x)^3. \quad (86)$$

Near the outer limit, $x_{\max}=x_1/3^{\frac{1}{2}}$, of the zone of trapping, write

$$x=x_{\max}(1+s), \quad (87)$$

where s is understood to be a small *negative* quantity. Then, from the differential equations and the boundary conditions the following behavior follows for the physically relevant quantities:

$$\begin{aligned} m(x)&=(x_1/3^{\frac{3}{2}})-(3x_1{}^3/4)s^2+\cdots\\ Q^2(x)&=1-(3^{\frac{5}{2}}/2)x_1{}^2s^2+\cdots\\ e^{-\lambda}&=(\tfrac{1}{3})+(\tfrac{2}{3})s+(\tfrac{2}{3})[(3^{\frac{5}{2}}x_1{}^2/4)-1]s^2+\cdots \quad (88)\\ e^{\nu}&=(\tfrac{1}{3})+(\tfrac{2}{3})s-(\tfrac{2}{3})s^2+\cdots\\ \omega/4\pi&=-3^{\frac{1}{2}}s+\cdots. \end{aligned}$$

The power series expressions for $e^{-\lambda}$ and e^ν join on smoothly to the accurate exterior values,

$$e^\nu=e^{-\lambda}=(1-2x_1/3^{\frac{3}{2}}x). \quad (89)$$

Near the inner limit, for $x\sim x_{\min}$, write

$$x=x_{\min}(1+u). \quad (90)$$

Then

$$\begin{aligned} m(x)&=(2^{\frac{1}{2}}/3)(x_1{}^4/x_{\min})u^{\frac{3}{2}}+\cdots\\ Q^2(x)&=(x_{\min}{}^2/x_1{}^2)+(2^{\frac{3}{2}}/3)x_1{}^2u^{\frac{3}{2}}+\cdots\\ e^{-\lambda}&=1-(2^{\frac{3}{2}}/3)(x_1{}^4/x_{\min}{}^2)u^{\frac{3}{2}}+\cdots \quad (91)\\ e^\nu&=(x_{\min}{}^2/x_1{}^2)+(2^{7/2}/15)x_1{}^2u^{5/2}+\cdots\\ \omega/4\pi&=2^{\frac{1}{2}}u^{\frac{1}{2}}+\cdots \end{aligned}$$

and, for $x_1\leqslant x_{\min}$,

$$e^{-\lambda}=1;\quad e^\nu=(x_{\min}{}^2/x_1{}^2);\quad m=0.$$

It is reasonable to assume a trial value for the eigenvalue parameter x_1, and start a numerical integration working inward from $x_{\max}=x_1/3^{\frac{1}{2}}$ with the starting series (88). The solid angle will first increase, then decrease. When it goes to zero, one wants $e^{-\lambda}$ to be unity. This condition is not satisfied in general. Accordingly, a new choice for x_1 is made. One proceeds by trial and error until $e^{-\lambda}$ goes to 1 as the solid angle goes to zero. This procedure was not adopted because the series expansion (88) was not available when the numerical work was done.

In the procedure that was used the coupled Eqs. (85) and (86) were integrated from the interior boundary $x_{\min}$ although the initial value of Q^2 depends not on a chosen value of x_1 alone, but also on the value of $x_{\min}$. To make the integration depend on a single parameter, the invariance of the equations to scale change, similar to that used in reference 1, was utilized. Defining $\bar{x}$, $\bar{m}$ and $\bar{Q}^2$ and c by the relations

$$x=b\bar{x},\quad Q^2=b\bar{Q}^2,\quad m=b\bar{m},\quad x_1{}^2=bc^2; \tag{92}$$

one has for $\bar{m}$ and $\bar{Q}^2$

$$d\bar{m}/d\bar{x}=(\bar{x}^2/2\bar{Q}^4(1-2\bar{m}/\bar{x})^2(\omega/4\pi)$$
$$d\bar{Q}^2/d\bar{x}=(\bar{x}/\bar{Q}^2(1-2\bar{m}/\bar{x})^3)\{(\omega/4\pi)+(\omega/4\pi)^3/3\},$$

where

$$\omega/4\pi=[1-c^2\bar{x}^{-2}\bar{Q}^2(1-2\bar{m}/\bar{x})]^{\frac{1}{2}}. \tag{93}$$

The choice $b=x_1{}^2/\bar{x}_{\min}{}^2$ or, equivalently, $c=\bar{x}_{\min}$, gives the simple boundary conditions at $\bar{x}_{\min}$:

$$\begin{aligned}\bar{Q}^2(c)=1,\quad & d\bar{Q}^2/d\bar{x}|_c=0,\\ \bar{m}(c)=0,\quad & d\bar{m}/d\bar{x}|_c=0;\end{aligned} \tag{94}$$

and the integration can be carried through for any c. At the outer boundary, $\bar{x}_{\max}=\Gamma$, the correct boundary values, following from (88) are

$$\begin{aligned}\bar{Q}^2(\Gamma)=3(\Gamma/c)^2,\quad & d\bar{Q}^2/d\bar{x}|_\Gamma=0,\\ \bar{m}(\Gamma)=\Gamma/3,\quad & d\bar{m}/d\bar{x}|_\Gamma=0;\end{aligned} \tag{95}$$

but in general these will not be satisfied for a particular choice of c. Starting at an arbitrary c the equations were integrated, on a punched card programed electronic computer, using the Kutta-Runge method; $\bar{m}$ was plotted as a function of $\bar{x}$. If a c is chosen smaller than the actual eigenvalue then, when $\bar{m}(\bar{x})=\bar{x}/3$, the slope $(d\bar{m}/d\bar{x})$ will not vanish. In such cases the integration was stopped at the point where $\bar{m}(\bar{x})=\bar{x}/3$ and the slope examined. A larger c, still less than the eigenvalue, will reduce the slope at the critical point. If c is chosen greater than the eigenvalue then $d\bar{m}/d\bar{x}$ vanishes before $\bar{m}(\bar{x})$ has decreased to $\bar{x}/3$. Thus, the correct eigenvalue for c can be approached both from above and below. With the limited machine time available the nearest value of c to its true eigenvalue for which a numerical integration was carried out was 1.875: the corresponding values for Γ and $\bar{m}(\infty)$ are $\Gamma=4.05$ and $\bar{m}(\infty)=1.35$. In Fig. 2 $m(x)$ and $Q^2(x)$ are plotted, by a dotted line, for this value of c, for which $b=0.071$ (and therefore $x_{\min}=0.134$ and $x_{\max}=0.289$). The smooth curves represent an estimate of $m(x)$, $Q^2(x)$, $e^{-\lambda(x)}$, $e^{\nu(x)}$ and $\omega(x)/4\pi$ for the correct eigenvalue.

5. VARIATIONAL PRINCIPLE FOR THE THERMAL GEON

All the differential equations of the Maxwell-Einstein theory, can be derived from the action principle

$$\delta I=0, \tag{96}$$

where

$$I=\int\int\int\int \mathfrak{L}(-g)^{\frac{1}{2}}dx^1dx^2dx^3dx^4 \tag{97}$$

and where

$$16\pi c\mathfrak{L}=c^4/GR_\alpha{}^\alpha-F_{\alpha\beta}F^{\alpha\beta} \tag{98}$$

and where the independent variables are the ten metric quantities g^{ik} and the four electromagnetic potentials Aj. Similarly, one can derive the differential equations of the simple spherical geon,[1]

$$\begin{aligned}d^2\phi/dx^2+jk\phi&=0\\ dk/dx+\phi^2&=0\\ dj/dx&=3-[1+(d\phi/dx)^2]/k^2\end{aligned} \tag{99}$$

from the variational principle

$$\begin{aligned}&\delta I=0\\ &I=\int\mathfrak{L}(u)du\\ &du=k-dx\\ &\mathfrak{L}=k^{-1}+3k+jkdk/du+j\phi^2-k-(d\phi/du)^2,\end{aligned} \tag{100}$$

as recently shown by Ernst.[2] Encouraged by this result, we found that we could derive the equations for the thermal geon from the principle

$$\begin{aligned}&\delta I=0\\ &I=\int\mathfrak{L}(x)dx\\ &\mathfrak{L}(x)=(1-xd\lambda/dx)e^{\nu/2-\lambda/2}-e^{\nu/2+\lambda/2}\\ &\qquad -(x^2/3)e^{\lambda/2}(e^{-\nu}-x_1{}^2x^{-2})^{\frac{3}{2}}.\end{aligned} \tag{101}$$

However, here the eigenvalue parameter x_1, which we would like to know, already enters the variational principle. The situation resembles that in which we would find ourselves if we tried to express the content of the Schrödinger equation in a variational principle built upon the function

$$\mathfrak{L}(x)=(\nabla\psi)^2+(V-\epsilon)\psi^2. \tag{102}$$

We will always find the trivial solution $\psi(x)=0$ for this variational problem unless ϵ happens to be an eigenvalue: not a very happy way to find eigenvalues! Much more appropriate is the more familiar Ritz variation principle

$$\epsilon\equiv\frac{\int[(\Delta\psi)^2+V\psi^2]dx}{\int\psi^2dx}=\text{minimum}, \tag{103}$$

into which we can substitute *any* well-behaved trial function and obtain an upper limit on ϵ.

Correspondingly, we take for variation principle

$$\delta(A^*B^*)=0, \tag{104}$$

$$A^*=\int_b^1 e^{\lambda^*/2+\nu^*/2}(d/dy)[y(1-e^{-\lambda^*})]dy, \tag{105}$$

$$B^*=\int_b^1 (y^2/3)e^{\lambda^*/2}(e^{-\nu^*}-3y^{-2})^{\frac{3}{2}}dy. \tag{106}$$

The independent variable here is y:

$$\begin{aligned}x=sy;\quad \lambda(x)=\lambda^*(y);\quad \nu(x)=\nu^*(y);\\ e^{-\lambda^*}=1-2m^*/y;\quad s=x_{\max}=x_1/3^{\frac{1}{2}};\\ x_{\min}=sy_{\min}=sb.\end{aligned} \tag{107}$$

The boundary conditions are

$$e^{-\lambda^*}=m^*=e^{\nu^*}=\tfrac{1}{3} \quad \text{at} \quad y=1 \tag{108}$$

$$\left.\begin{matrix} e^{\lambda^*}=1, & m^*=0, & e^{\nu^*}=b^2/3 \\ d\lambda^*/dy=0, & dm^*/dy=0 & \end{matrix}\right\} \quad \text{at} \quad y=b. \tag{109}$$

From the variational principle (104) follow differential equations for λ^* and ν^* as functions of y which are identical in form to Eqs. (78) and (79) for λ and ν as functions of x *except* for the introduction on the right-hand side of (78) and (79) of an extra factor

$$s^2=A^*/B^*. \tag{110}$$

The logic proceeds so: (1) Introduce into (105), (106) trial functions $\lambda^*(y)$ and $\nu^*(y)$ that satisfy (108), (109) and that contain one or more adjustable parameters, among them the lower limit b. (2) Calculate (A^*B^*) and extremize with respect to choice of these parameters. (3) From (110) calculate s^2 and hence x_1 and the geon mass

$$M=M_T x_1/3^{\frac{3}{2}}=M_T s/3. \tag{111}$$

If instead the quantities λ and ν are calculated as functions of x by solution of *differential* equations (78) and (79) for a trial eigenvalue x_1, an improved estimate of the eigenvalue is given by

$$x_1/3^{\frac{1}{2}}=s=(AB)^{\frac{1}{4}},$$

where A and B are the analogs of (105) and (106) for $\lambda(x)$ and $\nu(x)$.

6. PHOTON-PHOTON COLLISIONS IN A THERMAL GEON

The electromagnetic energy content of a thermal geon decreases slowly not only by the monomolecular process of barrier penetration but also by bimolecular processes in which two photons collide, either to produce a pair of electrons, or to go off as photons in new directions.[1] We wish to estimate the rate of these two processes in a very active part of a thermal geon: at the radius of the stable circular orbit (Fig. 1), where the solid angle occupied by bound rays is the largest,

$$(\text{solid angle}/4\pi)\doteqdot 0.59 \tag{112}$$

and where the effective temperature is

$$T_{\text{eff}}=Te^{-\nu/2}\doteqdot 3.3T. \tag{113}$$

Here the number of photons Ω_1 per unit volume in the interval of circular wave number $d\Omega_1$, and in the interval of solid angle $d\omega$ (within the allowed cone) will be

$$dn_1=(d\omega/4\pi)\Omega_1^2 d\Omega_1[-1+\exp(\hbar c\Omega_1/T_{\text{eff}})]^{-1}. \tag{114}$$

The collision between two photons[10] of wave numbers Ω_1 and Ω_2 whose directions of motion make an angle θ will look like the collision between two photons of equal wave number Ω^* and opposite direction, in a suitably selected local Lorentz system, where

$$\Omega^*=(\Omega_1\Omega_2)^{\frac{1}{2}}\sin\tfrac{1}{2}\theta. \tag{115}$$

[10] G. Breit and J. A. Wheeler, Phys. Rev. 46, 1087 (1934).

The total collision cross section, integrated over all angles of the emergence of the pair (for process 1) or of the scattered photons (for process 2) will be

$$\sigma=\sin^2\tfrac{1}{2}\theta\sigma^*(\Omega^*), \tag{116}$$

where Ω^* is the cross section calculated for equally energetic but oppositely moving photons. For the pair process the cross section σ^* vanishes below the threshold, $\Omega^*=mc/\hbar$, and reaches a maximum value of the order (e^2/mc^2) for a wave number that is a small multiple of this threshold value. For temperatures small compared to mc^2 it follows from the Planck formula that the number of pair production processes is exponentially small, with an exponential factor qualitatively of the form

$$\exp-(\mu mc^2/T), \tag{117}$$

where μ is of the order of unity.

The cross section for elastic photon-photon collisions also reaches a peak for wave numbers, Ω^*, of the order $mc/\hbar$. The peak value of σ^* for this process is of the order $(e^2/\hbar c)^2(e^2/mc^2)^2$, much smaller than that for the pair production mechanism. However, the cross section has no threshold and varies at low wave numbers in accordance with the formula[11]

$$\sigma^*=(52/1125\pi)(e^2/\hbar c)^2(e^2/mc^2)^2(\Omega^*\hbar/mc)^6. \tag{118}$$

For low temperatures the effective cross section for collision between two photons consequently varies as T^6, and therefore dominates over an exponentially small factor of the form (117). For this reason we can disregard pair production processes, relative to elastic collisions, so long as the temperature is considerably less than mc^2.

Not all photon collision processes result in loss of photons from the system. In some cases one or both of the new quanta still move in bound orbits. Only those are lost whose directions are thrown outside the cone of trapping angles. In particular a new photon escapes from the system if it is moving along a radius.

For an order-of-magnitude estimate of the rate of loss of energy we disregard details of the differences between trapping and escaping directions and consider the product of the following factors:

number of photons Ω_1 per cm^3	$\sim T^3/\hbar^3c^3$
number of photons Ω_2 per cm^3	$\sim T^3/\hbar^3c^3$
collision velocity	c
cross section	$\sim(e^2/\hbar c)^2(e^2/mc^2)^2(T/mc^2)^6$
energy loss from geon on collision	$\sim T$

(119)

product, energy loss per cm^3 and per sec $\sim(e^2/\hbar c)^4(mc^2/\hbar)\times(mc^2)^4\hbar^{-3}c^{-3}(T/mc^2)^{13}$

[11] H. Euler and B. Kockel, Naturwissenschaften 23, 246 (1935); W. Heisenberg and H. Euler, Z. Physik 98, 714 (1936).

In contrast, the energy on hand per unit volume is of the order $T^4/\hbar^3c^3$. This quantity, divided by (119), fixes a characteristic scale of time, τ, for depletion of the geon by photon-photon collisions:

$$\begin{aligned}\tau^{-1} &\sim \frac{(\text{energy loss/cm}^3\text{ sec})}{(\text{energy/cm}^3)} \\ &\sim (e^2/\hbar c)^4(mc^2/\hbar)(T/mc^2)^9 \\ &\sim (10^{12}\text{ sec}^{-1})(T/mc^2)^9.\end{aligned} \tag{120}$$

For a characteristic time as long as a year it is sufficient to have a temperature of the order

$$T \sim 10^{-20/9}mc^2 \sim 10^{-2}mc^2$$

corresponding to a geon mass of the order

$$M_T \sim 10^{44}g$$

and radius

$$R_T \sim 10^{16}\text{ cm}.$$

As the energy loss continues, the thermal geon shrinsk, grows denser and hotter, and loses energy at a rapidly increasing rate. As the temperature rises to the neighborhood of mc^2, pair production processes rapidly increase in importance. Then the physics of the system takes on quite a different character which we do not analyze here.

7. ZERO-POINT ENERGY

We have consistently disregarded quantum effects, or rather have consistently attempted to choose conditions where quantum effects are unimportant, in all except the considerations of very hot geons in the last section. However, at all temperatures one has to reckon with zero-point fluctuations in the electromagnetic field, as well as with the fluctuations due to the thermal radiation itself. One can formally associate these zero-point fluctuations with a zero-point energy, $\frac{1}{2}\hbar c\Omega$, that goes with each field oscillator. Usually this energy is left out in the bookkeeping of the energy of the electromagnetic field. The subtracted density of zero point energy of the vacuum is infinite. Normally one deals with field physics at the level of special relativity, where such an infinite quantity can be disregarded. However, in general relativity there is no such thing as an arbitrary additive constant in the density of field energy. Energy density curves space, and an increase in energy density produces an increase of curvature. Our analysis of thermal geons is based on the tacit assumption that the zero-point energy does not have to be counted, either as energy or as a source of curvature. We hope this point of view is correct, as it gives reasonable results in familiar situations. Nevertheless, a deeper approach to the problem of zero-point energy is needed—a problem whose overriding importance to all of field physics Niels Bohr has often stressed.

APPENDIX ON RAY-WAVE EQUIVALENCE

There is a close correlation between the field theory solutions of (44), (45), and (50) and the motion of a classical corpuscle of light or "pseudophoton" along a geodesic in the same gravitational field. Superposition of standing waves of slightly different Ω_{nl}, l, and m values allows one to build up a wave packet. This concentration of energy will remain the better defined in space and time the larger are the relevant values of l and n—that is, the shorter the wavelength of the disturbances of significant amplitude in comparison with the scale of distances over which the gravitational field changes appreciably. The correspondence between waves and pseudophoton orbits in the idealized limit which disregards the spreading of a wave packet may be analyzed this way: (1) We pass from the wave itself to its phase,

$$(\text{vector potential}) = \left\{\begin{matrix}\text{slowly varying} \\ \text{vector function} \\ \text{of position}\end{matrix}\right\} e^{i\cdot\text{phase}} \tag{A1}$$

where $e^{i\cdot\text{phase}}$ varies rapidly with position; (2) we approximate the differential equation for the phase by neglecting the slow change of the amplitude with distance in comparison with the rapid change of the phase with distance. Thus we pass from the accurate phase of (A1) to the pseudophase (or eikonal)—the central concept in William Rowan Hamilton's method of treating problems in geometrical optics as well as in mechanics.* The equation for the pseudophase has the Hamilton-Jacobi form

$$g^{\alpha\beta}(\partial\Phi/\partial x^\alpha)(\partial\Phi/\partial x^\beta) = 0. \tag{A2}$$

[Instead of deriving this from the wave equation by the substitution (A1), one can start directly with the picture of a surface propagating parallel to itself with the speed of light. The surface and its propagation with time can be expressed in the form

$$\Phi(x^1, \cdots x^4) = 17.2. \tag{A3}$$

Let x^4 be one value of the cotime, and x^1, x^2, x^3 the coordinates of one point on the surface at that time. Let $x^i + dx^i$ denote a neighboring point on the same moving surface, $\Phi = 17.2$, at a slightly later time. The space and time separations of the two points are connected by the relation

$$(\partial\Phi/\partial x^\mu)dx^\mu = 0. \tag{A4}$$

The requirement that the surfaces of (A4) move with the speed of light leads directly to (A2).] (3) We recognize not only single solutions, $\Phi(x_1' \cdots x^4)$ of the pseudophase equation, but also a typical family of

* See, for example, J. L. Synge, *Geometrical Mechanics and de Broglie Waves* (Cambridge University Press, 1954), Chap. II; or L. Landau and E. Lifschitz, *The Classical Theory of Fields*, reference 8, p. 271.

pseudophase functions,

$$\Phi=\text{function of } (x^1,\cdots x^4;\beta^1,\beta^2,\beta^3)+\beta^4(\beta^1,\beta^2,\beta^3), \quad \text{(A5)}$$

where each number of the family is characterized by a specific set of values for the three numbers β^1, β^2, β^2. These numbers might correspond in the case of physical optics to components of the wave vector for a monochromatic wave, but need not have such an immediate interpretation. The fourth constant of integration β^4, represents an additive phase constant, always allowed because (A2) contains only derivatives of the pseudophase, never Φ itself. (4) We impose the requirement of constructive interference. Several neighboring solutions of the wave equation of the form (A1) are superposed in such a way as to build up a wave packet. To make this requirement clean cut, we go to the pseudophase picture, and enquire how a pseudophoton must move in order that it shall always lie at the point where four pseudophase functions of nearly identical β values have a common value. The pseudophase need not be constant along the path of the pseudophoton. It is only required that at each point of the path, $x^i(\sigma)$, the four pseudophase functions should have a common value:

$$\begin{aligned}&\Phi(x(\sigma);\ \ \beta^1,\beta^2,\beta^3;\ \ \beta^4(\beta^1,\beta^2,\beta^3))\\ &\quad=\Phi(x(\sigma);\ \ \beta^1+d\beta^1,\beta^2,\beta^3;\ \ \beta^4(\beta^1+d\beta^1,\beta^2,\beta^3))\\ &\quad=\Phi(x(\sigma);\ \ \beta^1,\beta^2+d\beta^2,\beta^3;\ \ \beta^4(\beta^1,\beta^2+d\beta^2,\beta^3))\\ &\quad=\Phi(x(\sigma);\ \ \beta^1,\beta^2,\beta^3+d\beta^3;\ \ \beta^4(\beta^1,\beta^2,\beta^3+d\beta^3))\end{aligned}$$

or

$$\partial\Phi(x;\beta)/\partial\beta^s+(\partial\Phi/\partial\beta^4)(\partial\beta^4/\partial\beta^s)=0\ (=1,2,3). \quad \text{(A6)}$$

These three equations connect the three position coordinates of the pseudophoton with the time, and therefore suffice completely to determine its motion. From (A2) and its first derivatives with respect to the β's, this motion proceeds along a null geodesic,

$$d^2x^k/d\sigma^2+\Gamma_{\mu\nu}{}^k(dx^\mu/d\sigma)(dx^\nu/d\sigma) -\begin{pmatrix}\text{undetermined}\\ \text{function of }\sigma\end{pmatrix}\cdot(dx^k/d\sigma)=0$$

as expected. As an illustration how one gets all details of the pseudophoton motion from (A6), one can ask for the velocity components of the motion. For this one differentiates (A6) once with respect to σ and solves for the $dx^i/d\sigma$ by the method of determinants, finding

$$dx^i/d\sigma=f(\sigma)\{i\lambda\mu\nu\}(\partial^2\Phi/\partial x^\lambda\partial\beta^1)(\partial^2\Phi/\partial x^\mu\partial\beta^2) \times(\partial^2\Phi/\partial x^\lambda\partial\beta^3). \quad \text{(A7)}$$

Here the arbitrary functions $f(\sigma)$ in the solution appear because σ itself represented an arbitrary parametrization of the path of the pseudophoton. The same method gives the acceleration and other details of the motion.

For the spherically symmetric metric of the thermal geon the pseudophase propagation equation takes the form

$$e^{-\lambda}(\partial\Phi/\partial r)^2+r^{-2}(\partial\Phi/\partial\theta)^2 +(r\sin\theta)^{-2}(\partial\Phi/\partial\phi)^2-e^{-\nu}(\partial\Phi/\partial T)^2=0 \quad \text{(A8)}$$

which possesses separable solutions of the form

$$\Phi=\int_{r\min}^{r}[1-l(l+1)e^\nu/\Omega^2r^2]^{\frac{1}{2}}\Omega e^{\lambda/2-\nu/2}dr +\int_{\theta\min}^{\theta}[l(l+1)-m^2\sin^2\theta]^{\frac{1}{2}}d\theta +m\phi-\Omega T+\beta^4(\beta^1,\beta^2,\beta^3) \quad \text{(A9)}$$

Here the three constants of integration, β^1, β^2, β^3, have been expressed in the form Ω, l, and m to bring out the identity between the pseudophase and the phase of the JWKB approximation to the solution of the wave equation. We can superpose waves of different l, m, and Ω values to build up a wave packet that will trace out the designated geodesic.

The nature of the geodesic curve is independent of its orientation; consequently it is sufficient to consider m values close to zero, corresponding to motion in a meridian plane. *Which* meridian can be specified in the wave picture by the relative phase or β^4 values with which one superposes waves of slightly different m values; in the eikonal formulation, by setting equal to zero the derivative of the pseudophase with respect to m; thus $\partial\Phi/\partial\beta^3=\partial\Phi/\partial m=0$ gives an equation for ϕ. However, as there is no interest in this angle, it is appropriate to overlook this relation, and set m equal to a fixed value: $m=0$. Likewise the first equation of stationary pseudophase, $\partial\Phi/\partial\beta^1=\partial\Phi/\partial\Omega=0$, is also irrelevant for our purpose; we do not care *when* the pseudophoton arrives at a given point in its orbit. We are left with the second equation of constructive interference, $\partial\Phi/\partial\beta^2=\partial\Phi/\partial l=0$, to determine the *shape* of the geodesic:

$$0=-\int_{r\min}^{r}[r^2/e^\nu P^2-1]^{-\frac{1}{2}}e^{\lambda/2}dr/r+\theta+\text{const.} \quad \text{(A10)}$$

Here the quantity P is an abbreviation for the expression $[l(l+1)]^{\frac{1}{2}}/\Omega$ and represents the *impact parameter* of the pseudophoton—the distance of closest approach in the absence of gravitational forces. Evidently *this single constant determines the shape of the geodesic*, not Ω or l individually. From (A10) we derive the properties of the geodesics already discussed in the main text and pictured in Fig. 1. All the information gained about geodesics carries over to the characteristic solutions of the wave equation. The limits of motion, $r_1(P)$ and $r_2(P)$, of the rays correspond to the points where the field amplitudes change from oscillation to exponential fall off. Inclination of the nodal surfaces of the field is closely related to the inclination of the rays; and the energy carried by the rays or the wave fields has to supply the gravitating mass that holds the geon together.