REVIEWS OF

MODERN PHYSICS

VOLUME 43, NUMBER 3 JULY 1971

Terrestrial and Extraterrestrial Limits on The Photon Mass

ALFRED S. GOLDHABER*
Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, New York 11790
MICHAEL MARTIN NIETO† ‡
The Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
Department of Physics, University of California, Santa Barbara, California 93106

We give a review of methods used to set a limit on the mass μ of the photon. Direct tests for frequency dependence of the speed of light are discussed, along with more sensitive techniques which test Coulomb's Law and its analog in magnetostatics. The link between dynamic and static implications of finite μ is deduced from a set of postulates that make Proca's equations the unique generalization of Maxwell's. We note one hallowed postulate, that of energy conservation, which may be tested severely using pulsar signals. We present the merits of the old methods and of possible new experiments, and discuss other physical implications of finite μ. A simple theorem is proved: For an experiment confined in dimensions D, effects of finite μ are of order $(\mu D)^2$—there is no "resonance" as the oscillation frequency ω approaches μ ($\hbar=c=1$). The best results from past experiments are (a) terrestrial measurements of c at different frequencies

$$\mu \leq 2\times 10^{-43}\ \text{g} \equiv 7\times 10^{-6}\ \text{cm}^{-1} \equiv 10^{-10}\ \text{eV};$$

(b) measurements of radio dispersion in pulsar signals (whistler effect)

$$\mu \leq 10^{-44}\ \text{g} \equiv 3\times 10^{-7}\ \text{cm}^{-1} \equiv 6\times 10^{-12}\ \text{eV};$$

(c) laboratory tests of Coulomb's law

$$\mu \leq 2\times 10^{-47}\ \text{g} \equiv 6\times 10^{-10}\ \text{cm}^{-1} \equiv 10^{-14}\ \text{eV};$$

(d) limits on a constant "external" magnetic field at the earth's surface

$$\mu \leq 4\times 10^{-48}\ \text{g} \equiv 10^{-10}\ \text{cm}^{-1} \equiv 3\times 10^{-15}\ \text{eV}.$$

Observations of the Galactic magnetic field could improve the limit dramatically.

I. INTRODUCTION

One of the great triumphs of classical physics was the formulation of the Maxwell electromagnetic field equations. A fundamental prediction of these equations is that all electromagnetic radiation in vacuum travels at a constant velocity c. The most recent experiments have confirmed this prediction with an accuracy near to one part per million, over a wide range of frequencies (Froome and Essen, 1969; Taylor, Parker, and Langenberg, 1969).

In the context of quantum theory, a relativistic, quantized electromagnetic field of frequency ν is recognized as an assembly of photon particles with energy $h\nu$. These light quanta travel with velocity c, and hence have zero rest mass. The success of quantum electrodynamics in predicting experiments to six or more decimal places has made the massless photon a tacit axiom of physics. A sign of this is that as late as 1968 the Particle Data Group tables gave experimental limits on the neutrino masses, but just a zero for the photon mass (Rosenfeld, *et al.* 1968). This is not too surprising since QED is our only "exact" quantum theory. Nuclear and particle quantum theories do not even approach such accuracy.

The tacit axiom of masslessness corresponds to the belief that if the photon has an effective mass μ, it does so only because it is slightly off the mass shell. Using an uncertainty argument, we would estimate

$$\mu \approx \hbar/(\Delta t)c^2 = 3.7\times 10^{-66}\ \text{g}/T, \tag{1.1}$$

where T is the age of the universe in units of 10^{10} years.

* Supported in part under the auspices of the United States Atomic Energy Commission.
† Supported in part by the National Science Foundation.
‡ Address after September 1, 1971: Department of Physics, Purdue University, Lafayette, Ind. 47907.

Alternatively, one could get a similar number, following de Broglie (1954)[1] by considering a spherical de Sitter cosmology. In this model the cosmological constant K is given by the two equations

$$K=3/(cT)^2;\qquad K=\tfrac{1}{2}[\mu c/\hbar]^2, \tag{1.2}$$

or

$$\mu=6^{1/2}\hbar/Tc^2. \tag{1.3}$$

Equations (1.1) and (1.3) give an ultimate limit for a meaningful experimental measurement of the photon mass.

Since the time of Cavendish, certain critical physicists have not been satisfied with speculative assertions on this subject, and have periodically re-examined the question (or an equivalent one in the language of their time) to determine what valid experimental limit could be placed on the photon mass. In this paper we shall give a review of methods devised to improve the limit.

In Sec. II, we develop the theory of classical electrodynamics from postulates of special relativity, plus the assumption of a well-defined, locally conserved energy density associated with electromagnetic fields. We indicate how this assumption can be tested with pulsar signals. We proceed in Sec. III to discuss limits that have been set on the mass by terrestrial methods. These include determinations of the constancy of the velocity of light for all wavelengths, and testing the exactness of Coulomb's Law. The latter method yields the best laboratory mass limit to date, $\mu\leq 2\times 10^{-47}$ g.

In the next section extraterrestrial methods are reviewed. The first method is a variation on the terrestrial velocity of light experiments. Dispersion in the speed of starlight is inferred from the difference in arrival times of different colors of light from the same astronomical event. We then discuss the limits that can be obtained by studying the effects that a massive photon would have on the earth's magnetic field. This yields the lowest limit to date, $\mu\leq 4\times 10^{-48}$ g. Another technique considered is the study of long period hydromagnetic waves in plasma. If the photon has a finite mass, then such waves are damped below a critical frequency depending on μ and the plasma characteristics.

In the next section the physical effects of longitudinal photons are derived. We close in Sec. VI with a discussion of possible future experiments, their efficacy in improving present limits, and the physical implications of the results.

II. ELECTRODYNAMICS WITH FINITE μ

A. Heuristic Discussion

The assertion of a definite nonzero photon mass is equivalent to the specification of a free-electromagnetic wave by

$$\mathbf{E}=\operatorname{Re}\mathbf{E}_0\exp[-i(\omega t-\mathbf{k}\cdot\mathbf{x})],$$

$$\mathbf{H}=\operatorname{Re}\mathbf{H}_0\exp[-i(\omega t-\mathbf{k}\cdot\mathbf{x})], \tag{2.1}$$

$$(\omega/c)^2-\mathbf{k}^2=\mu^2, \tag{2.2}$$

where the last line defines μ in units of wavenumber, or inverse length. Standard arguments (Goldberger and Watson, 1964) then yield the desired expression for group velocity of a wave packet

$$v_g=d\omega/d|\mathbf{k}|=c|\mathbf{k}|/\omega=c|\mathbf{k}|/(\mathbf{k}^2+\mu^2)^{1/2}$$
$$=c(\omega^2-\mu^2c^2)^{1/2}/\omega. \tag{2.3}$$

This expression corresponds to a frequency dispersion of the velocity of light, the first and most direct consequence of a finite photon mass. [Note that here and in what follows, giving μ in units of wavenumber is using units of $c/\hbar$.]

Going to the Lorentz frame in which the photon is at rest, i.e., $\mathbf{k}=0$, we see that there must be three independent polarization directions for a massive photon, since the plane transverse to $\mathbf{k}$ is undefined in this frame. The argument fails for a massless photon because it can never have $\mathbf{k}=0$. In the photon rest frame the electric field energy density $\mathbf{E}^2$ is proportional to photon intensity. However, the well-known law of Lorentz transformations tells us that the fields in a frame with photon frequency ω and momentum $\mathbf{k}$ will be very different for photons polarized $\perp$ or $\|$ to $\mathbf{k}$ (Jackson, 1962; unreferenced assertions on electromagnetism in this paper may be found in Jackson's book):

$$\mathbf{E}_{0\perp}=(\omega/\mu c)\mathbf{E}_{0\perp\,\text{rest}},$$

$$\mathbf{H}_{\perp}=(\mathbf{k}/\mu)\times\mathbf{E}_{0\perp\,\text{rest}},$$

$$\mathbf{E}_{0\|}=\mathbf{E}_{0\|\,\text{rest}}. \tag{2.4}$$

If μ is much smaller than $|\mathbf{k}|$, the field of a longitudinal ($\|$) photon will be smaller than that of a transverse ($\perp$) photon by the factor $\mu c/\omega$. Since power absorbed by electric charges is proportional to $\mathbf{E}^2$, we infer that scattering cross sections of longitudinal photons will be suppressed compared to those of transverse photons by a factor $(\mu c/\omega)^2$; this weak coupling explains how the longitudinal polarization, if it exists, could have escaped detection up to the present. The phantom longitudinal photon is the second consequence of nonzero μ.

Finally, we consider the limit of static fields. For these fields, we have $\omega=(\mathbf{k}^2+\mu^2)^{1/2}=0$, implying $|\mathbf{k}|=i\mu$, hence, exponential decay of static fields with a range μ^{-1}. This behavior is familiar from Yukawa's model for interaction of nucleons through pion exchange. The exponential deviation from Coulomb's law, and its magnetic analog, provide the most sensitive current test for a photon mass. In the next section we find the postulates required to link this third effect rigorously with the previous consequences of finite μ.

---

[1] A remarkably similar discussion was given by Cap (1953). (See also Marochnik, 1968).

B. Deductive Approach

We adopt the following postulates:

(1) The electromagnetic field is defined through its action on a test charge q by the Lorentz force law,

$$\mathbf{F}=q[\mathbf{E}+(\mathbf{v}/c)\times\mathbf{H}]. \tag{2.5}$$

This law determines the behavior of $\mathbf{E}$ and $\mathbf{H}$ under Lorentz transformations: they may be identified as independent components of the antisymmetric 4-tensor $F_{\alpha\beta}$ by

$$F_{0i}=E_i,$$
$$F_{ij}=\epsilon_{ijk}H_k. \tag{2.6}$$

The force law in standard notation becomes

$$dp_\beta/d\tau=qu^\alpha F_{\alpha\beta}. \tag{2.7}$$

(2) The electromagnetic field at point x in space-time is linear in the charge and current densities, and in the derivatives of these densities, all evaluated at earlier points x'. Further, this linear relationship is Poincaré covariant (translation invariant and Lorentz covariant):

$$F_{\alpha\beta}(x)=\int d^4x D_{\alpha\beta/\gamma\delta}(x-x')\partial_\gamma J_\delta(x')$$
$$+\text{ terms with higher derivatives.} \tag{2.8}$$

This latter requirement is applied to assure invariance of the theory under the transformations of special relativity. The quantity $D_{\alpha\beta/\gamma\delta}$ must be an invariant tensor. There are only two possibilities:

$$D_{\alpha\beta/\gamma\delta}(x)=D(x)(g_{\alpha\gamma}g_{\beta\delta}-g_{\alpha\delta}g_{\beta\gamma})+\bar{D}(x)\epsilon_{\alpha\beta\gamma\delta}, \tag{2.9}$$

where ϵ is the completely antisymmetric 4-tensor. The presence of $\bar{D}$ implies parity violation or magnetic sources, depending on the point of view. The reason is that $\bar{D}$ produces a pseudovector $\mathbf{E}$ field, and a vector $\mathbf{H}$ field.

(3) We shall assume there are no magnetic sources or parity-violating terms in the theory. This eliminates terms like $\bar{D}$.

(4) Finally, we insist that the dependence of the theory on a small photon mass, μ, be such that as $\mu\to 0$ there is a smooth transition to the Maxwell theory.

It is easiest to find the consequences of these postulates in "momentum space". Define (k a 4-vector)

$$\tilde{F}_{\alpha\beta}(k)=\int d^4x \exp(ik\cdot x)F_{\alpha\beta}(x),$$
$$\tilde{D}_{\alpha\beta/\gamma\delta}(k)=\int d^4x \exp(ik\cdot x)D_{\alpha\beta/\gamma\delta}(x),$$
$$\tilde{J}_\alpha(k)=\int d^4x \exp(ik\cdot x)J_\alpha(x). \tag{2.10}$$

Then, the convolution integral Eq. (2.8) becomes

$$\tilde{F}_{\alpha\beta}=\tilde{D}_{\alpha\beta/\gamma\delta}(-ik_\gamma)\tilde{J}_\delta$$
$$+\text{ terms with more factors of the 4-vector } k. \tag{2.11}$$

If we ignore parity-violating terms as required by Postulate 3 above, we may write Eq. (2.11) more simply as

$$\tilde{F}_{\alpha\beta}(k)=-i\tilde{D}(k)(k_\alpha\tilde{J}_\beta-k_\beta\tilde{J}_\alpha), \tag{2.12}$$

where $\tilde{D}$ is an invariant function of k, and the right-hand side is the most general antisymmetric tensor built out of $\tilde{J}$ and its derivatives, i.e., linear in $\tilde{J}_\alpha$ and an arbitrary function of k_α. Thus, the requirements of Poincaré invariance (including parity) are sufficient to deduce the homogeneous Maxwell equations, which may be written

$$k^\alpha\epsilon_{\alpha\beta\gamma\delta}\tilde{F}^{\gamma\delta}(k)=0, \tag{2.13}$$

and are obviously satisfied by the above form Eq. (2.12). To state this another way, we have now shown from invariance requirements alone that the fields may be derived from a 4-vector potential:

$$\tilde{F}_{\alpha\beta}(k)=-i[k_\alpha\tilde{A}_\beta(k)-k_\beta\tilde{A}_\alpha(k)],$$
$$\tilde{A}_\alpha(k)=\tilde{D}(k)\tilde{J}_\alpha(k). \tag{2.14}$$

Next, we study the properties of $\tilde{D}(k)$. Since $\tilde{D}$ is Lorentz invariant, we shall assume that it is a function only of the invariant quantity $k^2\equiv k_\alpha k^\alpha$, even for complex k_α, giving $\tilde{D}(k)=\tilde{D}(k^2)$. This can be proven from our postulates.[2] Let us consider $\mathbf{k}=0$. The condition $D(t<0)=0$ implied by Postulate 2 in turn implies that if the inverse Fourier transform $D(t)=(2\pi)^{-4}\int d^4k \exp(-ik\cdot x)\tilde{D}(\omega, \mathbf{k}=0)$ exists, then $\tilde{D}(\omega,\mathbf{k}=0)$ is analytic in the upper-half complex ω plane. Further, the requirement that D is real implies $\tilde{D}(\omega)=\tilde{D}^*(-\omega^*)$. Translated into the variable $k^2\equiv\omega^2/c^2-\mathbf{k}^2=\omega^2/c^2$, these results imply that $\tilde{D}(k^2)$ is analytic in the entire complex k^2 plane except for the positive real k^2 axis, and any discontinuity across this axis is imaginary. Unless there is a purely local current–current interaction, $\tilde{D}(k^2)$ must go to zero as k^2 goes to infinity. We exclude the local interaction since it is not present in the Maxwell theory.

We then may use Cauchy's theorem to write a dispersion relation for $\tilde{D}$ by integrating over its imaginary discontinuity

$$\tilde{D}(k^2)=\pi^{-1}\int_0^\infty\frac{d\mu^2\,\mathrm{Im}\,\tilde{D}(\mu^2)}{\mu^2-k^2}. \tag{2.15}$$

If Im $\tilde{D}$ has a delta function, then $\tilde{D}$ has a pole.

Before considering the most general case, let us specialize by assuming Im $\tilde{D}$ consists of a single delta function at a particular value μ^2, giving

$$(-k^2+\mu^2)\tilde{F}_{\alpha\beta}=(4\pi/c)(-i)(k_\alpha\tilde{J}_\beta-k_\beta\tilde{J}_\alpha)$$

or

$$(\Box+\mu^2)F_{\alpha\beta}=(4\pi/c)(\partial_\alpha J_\beta-\partial_\beta J_\alpha). \tag{2.16}$$

[2] This can be shown as a trivial example of the discussion in Streater and Wightman (1964).

This may be recognized as the ordinary Maxwell equation, modified by the addition of μ^2 to the D'Alembertian operator. [The free ($J_\alpha=0$) solutions of this equation obey the relation $\omega/c=(\mu^2+\mathbf{k}^2)^{1/2}$.] We may rearrange Eq. (2.16) by introducing the vector potential A_α satisfying

$$(\Box+\mu^2)A_\alpha=(4\pi/c)J_\alpha,$$
$$\partial_\alpha J^\alpha=0,$$
$$F_{\alpha\beta}=\partial_\alpha A_\beta-\partial_\beta A_\alpha. \tag{2.17}$$

Rewriting further gives us the famous Proca equation (Proca, 1930a, b, c; 1931; 1936a, b, c, d, e) for a massive vector field coupled to a conserved current,

$$\partial^\alpha F_{\alpha\beta}+\mu^2 A_\beta=(4\pi/c)J_\beta,$$
$$F_{\alpha\beta}=\partial_\alpha A_\beta-\partial_\beta A_\alpha. \tag{2.18}$$

The whole effect of finite photon mass is to introduce at each point x a spurious current proportional to the vector potential and, therefore, a function of the true current at many earlier points x'. In three-dimensional notation the massive Maxwell equations become

$$\nabla\cdot\mathbf{E}=4\pi\rho-\mu^2 V,$$
$$\nabla\times\mathbf{E}=-(1/c)(\partial\mathbf{H}/\partial t),$$
$$\nabla\cdot\mathbf{H}=0,$$
$$\nabla\times\mathbf{H}=(4\pi/c)\mathbf{J}-\mu^2\mathbf{A}, \tag{2.19}$$

with $\mathbf{A}$ and V the space and time components of the 4-vector potential A_μ.

It is worth noting that the freedom of gauge invariance found in conventional electrodynamics is completely lost here. First of all, the Lorentz gauge must be used, i.e., $\partial_\alpha A^\alpha=0$. Within that restriction, one might imagine adding to A_α a term $\partial_\alpha\Lambda$, where Λ is a scalar function. This does not change $F_{\alpha\beta}$, of course, but the Lorentz gauge condition implies $\Box\Lambda=0$. Therefore, if A_α is already a solution of the Proca equation we have the contradictory requirements $\Box\partial_\alpha\Lambda=0$ and $(\Box+\mu^2)\partial_\alpha\Lambda=0$, satisfied only if Λ is constant. Hence, all freedom of gauge change is lost.

It is easy to verify, for free fields, that there exists a conserved energy–momentum density (de Broglie, 1957; Bass and Schrödinger, 1955) such that

$$\mathcal{E}_{\mathrm{EM}}=[\mathbf{E}^2+\mathbf{H}^2+\mu^2(\mathbf{A}^2+V^2)]/8\pi,$$
$$\mathbf{p}_{\mathrm{EM}}=[\mathbf{E}\times\mathbf{H}+\mu^2 V\mathbf{A}]/4\pi c, \tag{2.20}$$

where the conservation we refer to is the equation of continuity

$$(1/c)(\partial\mathcal{E}_{\mathrm{EM}}/\partial t)+\nabla\cdot\mathbf{p}_{\mathrm{EM}}c=0. \tag{2.21}$$

When charges and currents are present we obtain

$$d\mathbf{P}/dt=0, \tag{2.22}$$

with

$$\mathbf{P}=\int d^3x(\mathbf{p}_{\mathrm{EM}}+\mathbf{p}_{\mathrm{matter}}), \tag{2.23}$$

and

$$(d\mathbf{p}/dt)_{\mathrm{matter}}=\rho\mathbf{E}+(\mathbf{J}/c)\times\mathbf{H}, \tag{2.24}$$

the Lorentz force density.

The vector potential is never measured directly, but it is determined uniquely, and is required for construction of a locally conserved electromagnetic energy and momentum density.

Let us elevate the principle just mentioned to a fifth postulate:

(5) There exists a locally conserved energy–momentum density, such that the total energy and momentum of a system of charges and fields is conserved.

We shall now consider the restrictions implied by this postulate on $\mathrm{Im}\,\tilde{D}(\mu^2)$.

Clearly, a minimal requirement on $\mathrm{Im}\,\tilde{D}(\mu^2)$ is that it be integrable, i.e., a bounded continuous function falling faster than $1/\ln\mu^2$ at high masses, plus a sum of delta functions and derivatives of delta functions. Therefore, $\tilde{D}(k^2)$ will be a sum of pole terms

$$\{\sum_i d_i/(\mu_i^2-k_i^2)\}$$

plus a continuous integral over pole terms (a cut)

$$\{\int[d(\mu^2)/(\mu^2-k^2)]\}$$

plus second or higher order poles $[d/(\mu^2-k^2)^2$, etc.]. All these terms can be written as simple poles or limits of sums of simple poles.

Consider the case of two pole terms $[\tilde{D}=d_1/(\mu_1^2-k^2)+d_2/(\mu_2^2-k^2)]$. This leads to the possibility of arbitrary free fields with either $\omega=c(\mu_1^2+k^2)^{1/2}$ or $\omega=c(\mu_2^2+k^2)^{1/2}$. Take the case $\mathbf{k}=0$. One may have an electric field $\mathbf{E}=\mathbf{E}_0(\cos\mu_1 ct-\cos\mu_2 ct)$, with $\mathbf{A}=\mathbf{E}_0(\mu_2^{-1}\sin\mu_2 ct-\mu_1^{-1}\sin\mu_1 ct)$. At $t=0$, both $F_{\alpha\beta}$ and A_α are zero everywhere, so that any energy density quadratic in F and A must vanish. However, an instant later this is no longer true. Therefore, there is no conserved electromagnetic energy built simply from F and A. For free fields, a conserved energy density can be constructed by projecting the parts of E corresponding to each mass

$$\mathbf{E}_1=[(\mu_2^2+\Box)/(\mu_2^2-\mu_1^2)]\mathbf{E},$$
$$\mathbf{E}_2=[(\mu_1^2+\Box)/(\mu_1^2-\mu_2^2)]\mathbf{E}. \tag{2.25}$$

With the obvious definitions of $\mathbf{A}_1$ and $\mathbf{A}_2$, etc., we get the conserved energy density

$$8\pi\mathcal{E}=c_1[\mathbf{E}_1^2+\mathbf{H}_1^2+\mu_1^2(V_1^2+\mathbf{A}_1^2)]+c_2[\mathbf{E}_2^2+\mathbf{H}_2^2+\mu_2^2(V_2^2+\mathbf{A}_2^2)]. \tag{2.26}$$

In the presence of sources, however, our arbitrary but simple definition of $\mathcal{E}$ may be seen to fail. For example, by calculating the potential energy of a charge distribution and comparing it with the total electromagnetic energy E one finds that the two are not equal.

The only way to maintain energy conservation is to insist that the fields associated with μ_1 and μ_2 are independent contributors to the energy, even though there is no general operational distinction between them. In particle language, we would say there are two different photons, though they act on charges in the same way.

Once this is accepted, it is straightforward to deduce

$$(d/dt)\int d^3x\mathcal{E}(x) = -\int d^3x\mathbf{J}\cdot(c_1d_1\mathbf{E}_1+c_2d_2\mathbf{E}_2). \quad (2.27)$$

In order that total energy be conserved, this must balance the effect of the Lorentz force on charges. This means that

$$c_1d_1=+1, \qquad c_2d_2=+1. \quad (2.28)$$

If $\mathcal{E}(x)$ is positive definite, $(c_i\geq 0)$, then the residues d_i must both be positive. This excludes higher order poles, which are obtained in a limit as simple poles with residues of both signs approach each other. Another way to express the difficulty with higher order poles is to observe that they lead to fields which grow in time, e.g., $\mathbf{E}_0 t\cos\mu t$ for a second-order pole at μ. This is a solution of $(\Box+\mu^2)^2\mathbf{E}=0$. A cut in $\tilde{D}(k^2)$ may be produced as a limit as the number of poles in a certain interval diverges and the residue d_i of each pole goes to zero. From Eq. (2.28) this means that the coefficient c_i of the corresponding field energy density diverges, so that in the limit $\mathcal{E}(x)$ is undefined. Thus, it is impossible to produce a cut by exciting an infinite number of photonlike degrees of freedom, and still preserve energy–momentum conservation: a "continuous-mass" photon is excluded.

If Postulate (5) holds, we may introduce one or n new poles in D at a price of the admission of one or n new photons each with three degrees of freedom. This would contradict well-known information about blackbody radiation (de Broglie, 1957; Bass and Schrödinger, 1955), and elementary particle reactions (Brodsky and Drell, 1971) unless either the new photons all have a mass greater than many GeV, or else their coupling to charge d_i is so small that their degrees of freedom are not appreciably excited during times of practical interest. In either case, their existence would have no significant effect on a search for effects of a possible finite mass of the everyday photon. In fact, there *are* known weak cut contributions to $\tilde{D}$ derivable in quantum electrodynamics and indeed, associated with new degrees of freedom. For example, at values of $\mu^2>4m_e^2$, a virtual photon can dissociate into an e^+e^- pair. This leads to a contribution to D suppressed by at least a factor of the fine structure constant $\alpha\approx 1/137$, and of very short range (10^{-11} cm) for static fields (Bjorken and Drell, 1965).[3] An even weaker cut beginning at $\mu^2=0$ but very small below $\mu^2=m_e^2$ and suppressed at least by α^2, is produced by the dissociation of a virtual photon into three correlated photons. These cuts are not associated with free photonlike degrees of freedom and do not violate our earlier conclusion forbidding a continuous mass photon.

It is amusing to consider in this classical context the modified electrodynamics of Lee and Wick (1969). In order to eliminate the small distance divergence in Maxwell's theory and its quantized version, they introduce a D with two poles; one at zero mass, and one at very large mass with

$$d_1=-d_2=4\pi. \quad (2.29)$$

In consequence, the electric potential between two point charges is bounded at small distances

$$V(r)=(qq'/r)(1-e^{-\mu r_2})\rightarrow\mu_2 qq' \quad \text{as} \quad r\rightarrow 0, \quad (2.30)$$

where r is the distance between q and q'. However, since d_2 is negative, so is c_2. Therefore, a wave packet of type 2 photons will carry negative energy. This creates the problem that by producing more and more type 2 photons one can gain more and more energy. In a quantum context the problem may be stated as a violation of unitarity (conservation of probability). Lee and Wick circumvent this by indicating a calculational scheme in which free type 2 photons are never produced, and energy densities are always positive definite.

Since Postulate (5) is of a different character from the other four postulates, we may ask what complications arise if it fails and there are several very low mass poles or even a cut restricted to low mass. Now we expect violations of local energy–momentum conservation, but these would be conspicuous only for fields with very small ω and k. There could be a fortuitous cancellation of the lowest order effect in electro- or magneto-statics. For example, with the two poles:

$$\begin{aligned}\tilde{D}&=4\pi[2/(\mu^2+k^2)-1/(2\mu^2+k^2)]\\&=+(4\pi/k^2)[1+\mathcal{O}(\mu^2/k^2)^2],\end{aligned} \quad (2.31)$$

one would have much smaller deviations from Coulomb's law than with one pole

$$\begin{aligned}\tilde{D}&=4\pi/(\mu^2+k^2)\\&=+(4\pi/k^2)[1-(\mu^2/k^2)+\mathcal{O}(\mu^2/k^2)^2].\end{aligned} \quad (2.32)$$

However, the resulting spreading of light pulses (an energy nonconserving effect) could be looked for as a phenomenon distinct from frequency dispersion of v_g, since it could be observed at a single given frequency. Pulsar signals can be used to give a limit on such violations of Postulate (5), but, even if they exist, special cancellations must occur if the effect on static fields is to be masked to any given order in $\mu/|\mathbf{k}|\approx\mu D$ (where D is the dimension of the experimental apparatus).

We conclude that, in addition to the basic symmetry

[3] An amusing line of speculation is indicated in a series of papers culminating in that of Bandyopadhyay, Chaudhuri, and Saha (1970). They suggest that the photon may couple to a neutrino–antineutrino pair, producing an effective photon mass which is different in different Lorentz frames, because there is a filled neutrino sea which is "at rest" only in the "rest frame" of the Universe.

principles of special relativity and the assumption that fields are linear functions of currents, an eminently reasonable postulate of energy conservation is required to deduce that classical electromagnetic theory can be modified in only one way—replacing the Maxwell equation by the Proca equation to account for a possible small photon mass. If the energy postulate is omitted there is no simple prediction for the effect on statics, but there are two remarkable effects on light radiation:

1. A narrow-band pulse of light may spread in duration or separate into a number of discrete components in a time $(c/\Delta v)\tau$, where τ is the original pulse length, and Δv is the range of light velocities associated with the range K of μ values in $\tilde{D}$.
2. As the spreading occurs, the classical integrated intensity $\int d^3x(\mathbf{E}^2+\mathbf{H}^2)/8\pi$ may increase or decrease dramatically depending on the variation of the sign of $\mathrm{Im}\,\tilde{D}(\mu^2)$ in K.

One could also question the postulates of special relativity and linearity. We don't do so here for two reasons. First, we have no simple way to parameterize deviations from these postulates—too many possibilities would be opened by discarding them. Second, both postulates have been very successful in quantum electrodynamics, where the accuracy of perturbation theory validates the use of linearity. Thus, any violations must appear only at very long times or distances. This was easy to arrange for our particular version of energy nonconservation, but seems nontrivial for the other assumptions. Not surprisingly, we feel that the effects (1) and (2) above are also unlikely. In the remainder of this paper we shall assume there is a single fixed value of $\mu\geq 0$, except where indicated explicitly.

It is worth noting that the most common technique for deriving massive electrodynamics is the use of a Lagrangian density (cf. for example, Gintsburg, 1963). One simply adds to the $\mu=0$ Lagrangian a "photon mass term" proportional to $\mu^2A_\alpha A^\alpha$. This is the most general modification which vanishes as $\mu\rightarrow 0$ and involves only local coupling (all fields evaluated at the same point in spacetime). The Lagrangian approach embodies all of our postulates (1–5), but we hope the reader has found it instructive to examine these assumptions separately.

III. TERRESTRIAL LIMITS

A. Measurement of c

The most straightforward way to obtain a limit on the photon mass is to look for a variation in c over the spectrum. Equation (2.3) shows that long-wavelength light will have a velocity differential from the short-wavelength limit ($v=c$) of

$$-(\Delta v/c)=\mu^2c^2/2\omega^2+\mathcal{O}[(\mu c/\omega)^4]$$
$$=(\mu^2\lambda^2/8\pi^2)+\mathcal{O}[(\mu\lambda)^4]. \quad (3.1)$$

The velocity c is measured to an accuracy of one to ten parts in 10^6 over much of the electromagnetic spectrum.[4] The lowest frequency measurement with such precision is that at $\nu=173$ MHz ($\lambda=1.73$ m) (Florman, 1955). The one in 10^5 accuracy of this measurement implies $\mu\leq 2\times10^{-4}$ cm^{-1} $\equiv 3\times10^{-9}$ eV $\equiv 6\times10^{-42}$ g. Because the effect is quadratic in wavelength, one can improve considerably on this number by going to lower frequency, even if the measurement is less accurate. In the 1930's, Mandel'shtam and Papalexi (1944) and their collaborators developed a technique for measuring the velocity of long radio waves.[5] A radio wave of frequency ν is sent from a transmitter to a receiving station far away. At the receiver, a wave of frequency $(\frac{3}{2})\nu$, for example, is synchronized with the received wave transmitted back to the original station. The phase lag of the return wave with respect to the original signal has calculable contributions, including effects of the apparatus at both ends, plus a term proportional to the time of travel. Al'pert, Migulin and Ryazin (1941)—and earlier work cited therein—used this technique to measure the dispersion of long ($\gtrsim 10^2$ m) waves travelling over land and sea. Over land, the dispersion was quite large ($\approx 1\%$), but over sea they measured a velocity shift of 7×10^{-4} between 300–450 m. If this is interpreted as a photon mass effect, it corresponds to

$$\mu\leq 2\times10^{-43}\text{ g}\equiv 7\times10^{-6}\text{ cm}^{-1}\equiv 10^{-10}\text{ eV}. \quad (3.2)$$

It is possible that the result of Al'pert, *et al.* was due to instrumental error. However, it would appear difficult to improve enormously on their work because of irregularities in the medium through which the wave propagates (the Earth and its atmosphere). We dismiss

TABLE I. Experimental limits on deviations from Coulomb's law.

Authors	Date	ν (Hz)	q	μ (cm^{-1})
Coulomb	1785	0	10^{-1}	$\approx 10^{-1}$
Robison	1769	0	6×10^{-2}	$\approx 10^{-2}$
Cavendish	1773	0	3×10^{-2}	$\approx 10^{-2}$
Maxwell	1873	0	5×10^{-5}	$\approx 10^{-3}$
Plimpton *et al.*	1936	2	2×10^{-9}	10^{-6}
Cochran *et al.*	1968	10^2–10^3	9×10^{-12}	9×10^{-8}
Bartlett *et al.*	1970	2.5×10^3	10^{-13}	10^{-8}
Williams *et al.*	1971	4×10^6	6×10^{-16}	5×10^{-10}

[4] The most complete recent summary of precise measurements of c is Froome and Essen (1969).

[5] Brief summaries in English are given by Smith-Rose (1942a, b).

in Sec. VI the possibility of further improvement using a more controlled environment of smaller dimensions.

B. Deviations from Coulomb's Law

The inverse square force law was first announced in 1785 by Coulomb (1788), who used a torsion balance to measure directly the repulsive force between two like charges. A significantly better test of Coulomb's law was devised by Cavendish in 1773 (sic!).[6] Cavendish set up two concentric conducting spheres connected by a wire. He charged the outer sphere, and then disconnected the wire. If there were a deviation from the inverse square law, then upon removing the outer sphere, one would find a calculable charge on the inner sphere. Cavendish's experimental limits on such a charge allowed him to say that, if the correct law is

$$F=e_1e_2/r^{2+q}, \qquad (3.3)$$

then $|q|$ is less than 1/50. Surprisingly, Cavendish never published this result. A public description had to wait until Maxwell (1873) included it in his great treatise.[7]

Maxwell improved the result in a new experiment. The only modification was that the outer shell was grounded instead of removed, and the inner globe was tested for charge through a small hole. Maxwell also derived the theory of such an experiment for an arbitrary central force law. From his null result he obtained a limit (Maxwell, 1873)

$$q<1/21\,600. \qquad (3.4)$$

Plimpton and Lawton (1936) performed an improved version of the Cavendish–Maxwell experiment. They took two concentric conducting spheres of radii $a=2.5$ ft, and $b=2.0$ ft, grounded them, and then charged the outer sphere to $V=3000$ V. Actually, for technical reasons, the voltage was quasistatic, having a frequency of 130 cycles/min. A galvanometer which connected the two spheres and which could be observed through a conducting window indicated $\Delta V=\Phi(a)-\Phi(b)<10^{-6}$ V. Using the theory of Maxwell (1873), this meant

$$q<(\Delta V/V)F(a,b)\simeq 2\times10^{-9}, \qquad (3.5)$$

where

$$F(a,b)=\tfrac{1}{2}n\ln[(n+1)/(n-1)]-\tfrac{1}{2}\ln[4n^2/(n^2-1)], \qquad n=a/b. \qquad (3.6)$$

To convert this result to a limit on the photon mass, we note that from Eq. (2.17) the potential between the two spheres is given in the static limit by

$$(\nabla^2-\mu^2)\Phi=0, \qquad (3.7)$$

or

$$\Phi(r)\propto(e^{\mu r}-e^{-\mu r})/2\mu r. \qquad (3.8)$$

Equation (3.8) is normalized by taking

$$V=\Phi(a) \qquad \Delta V=\Phi(a)-\Phi(b). \qquad (3.9)$$

Then an expansion in powers of (μa) and (μb) yields

$$\Delta V/V=\tfrac{1}{6}\mu^2(a^2-b^2)+\mathcal{O}[(\mu a)^4]. \qquad (3.10)$$

Substitution of Plimpton and Lawton's experimental results into Eq. (3.10) gives

$$\mu\lesssim10^{-6}\ \mathrm{cm}^{-1} \equiv 2\times10^{-11}\ \mathrm{eV} \equiv 4\times10^{-44}\ \mathrm{g}, \qquad (3.11)$$

which was the best laboratory limit until recently.

Within the past four years, Cochran and Franken (1967, 1968); Bartlett, Goldhagen, and Phillips (1969, 1970); and Williams, Faller and Hill (1970a, b; 1971) have surpassed the Plimpton–Lawton result by one, two, and better than three orders of magnitude, respectively. The three experiments are similar to the older one in principle. Aside from advances in quality of available equipment, the first essential improvement is the use of a "lock-in" detector to observe oscillations in ΔV in synchronism with oscillations in the applied potential V. The second improvement is to increase the oscillation frequency, reducing thermal, or Johnson, noise in the relevant frequency band: the Johnson (1928)[8] noise

[6] Actually, the discovery of the inverse-square law for electricity precedes Coulomb by quite a number of years. In 1755, Benjamin Franklin noted that a cork lowered inside a charged silver can was not attracted to the side of it, as he thought it would be. He wrote to John Lining that, "You require the reason; I do not know it." After Franklin later wrote to Joseph Priestley, Priestley repeated the experiments and reported them at the end of his great work of 1767. There Priestley made the brilliant deduction that the experiment implied that the electric force law was the same as the law of gravitational attraction, i.e., it was an inverse square law. Two years later in 1769 the Scotsman John Robison made the first experimental determination of the law. Robison had been inspired by the speculation of AEpinus that there was an inverse square law, and AEpinus in turn had been inspired to this speculation by the two charge theory of Franklin. By balancing the electrical and gravitational forces acting on a sphere, Robinson obtained a result of $q=2.06$. Thus, Robison preceded the experiments of Cavendish that we mention below. But, except for an unremembered lecture, he did not make public his results until 1803. Consult: Franklin (1774), Priestley (1767), AEpinus (1759), and Robison (1803).

Finally, we mention that the inverse square law for magnetic forces was discovered by Johann Tobias Mayer in 1760, by Johann Heinrich Lambert in 1766-1776, and in its fullness by Coulomb in 1785. Consult Mottelay (1922).

[7] A description of Cavendish's experiments, taken from his manuscripts, is contained in Cavendish (1879). A. D. Dolgov and V. I. Zakharov (1971) have pointed out that the Cavendish technique was more sensitive to $\mu\neq0$ than the raw data indicate. The reason is that, as we now know, conductors on the Earth's surface are at an absolute potential of about 10^5 V because of charge separation between earth and ionosphere. Therefore, the V in Eq. (3.10) is 10^5 V for the static experiments of Cavendish and Maxwell. The corresponding limits are improved by a considerable factor, though remaining inferior to the nonstatic result of Plimpton and Lawton (1936).

[8] The theory of Johnson's experimental discovery was presented by Nyquist (1928).

in the input to the amplifier is given by

$$\langle \Delta V_{\text{noise}}^2 \rangle = 4kT\Delta\nu \text{ Re } Z, \quad (3.12)$$

where k is Boltzmann's constant, T is the absolute temperature, $\Delta\nu$ is the bandwidth (or reciprocal of observation time), and Re Z is the real part of the impedance,

$$Z^{-1} = R^{-1} + i\omega C. \quad (3.13)$$

Here R is the input resistance, C is the parallel capacitance, and $\omega/2\pi$ is the frequency. For large ω we have

$$\text{Re } Z \approx [R(\omega C)^2]^{-1}, \quad (3.14)$$

and the root mean square noise voltage is inversely proportional to frequency.

Rather than spoil the reader's pleasure by a second-hand description of these experiments, we refer him to the original papers, contenting ourselves with the (tabular) summary presented in Table I, modified and expanded from that of Bartlett, *et al.* (1970).

The value of Williams *et al.* may still improve. It represents the best laboratory limit to date, an improvement in voltage sensitivity of more than 10^6 over Plimpton and Lawton and 10^{13} over Cavendish!

IV. EXTRATERRESTRIAL LIMITS

A. Dispersion in the Speed of Starlight

A limit on the photon mass can be obtained by measuring the difference in time of arrival of radiation of different frequencies with the same origin. For example, if blue and red light rays come from the same event, the difference in time of arrival is

$$\delta t = \int dl[(1/v_R) - (1/v_B)] \simeq (L/c^2)(v_B - v_R), \quad (4.1)$$

which, from Eq. (3.1), is

$$\delta t = (8\pi^2 c)^{-1} \mu^2 L(\lambda_B^2 - \lambda_R^2). \quad (4.2)$$

De Broglie (1940) suggested that this method could yield a mass limit by using light from a star emerging from behind its dark binary companion. De Broglie considered the case $(\lambda_B^2 - \lambda_R^2) = 0.5 \times 10^{-8}$ cm^2 (for example, $\lambda_R \sim 8000$ Å, $\lambda_B \sim 4000$ Å), $L = 10^3$ light years, and $\delta t \leq 10^{-3}$ sec. Then one gets

$$\mu \leq 0.78 \times 10^{-39} \text{ g}. \quad (4.3)$$

Interestingly, in the last step of his numerical calculation, de Broglie made a mistake of order 10^5, quoting a limit of 10^{-44} g. What is more amusing, however, is that this number was quoted (de Broglie, 1957, p. 59) and requoted (Bass and Schrödinger, 1955), but for 28 years no one publicly took the trouble to check de Broglie's simple numerical calculation.[9] The first published correction appeared in Kobzarev and Okun' (1968). We noticed it in preparing Goldhaber and Nieto (1968).

However, even if one takes advantage of the wider spectrum observations and faster electronics now available, this method is intrinsically limited. As Gintsburg (1963) has nicely summarized, this is true for a number of reasons, but the crucial one is the natural dispersion of light traveling through the interstellar plasma in a magnetic field. The dispersion equation for such a system is

$$k^2 = (\omega^2/c^2)[1 - \omega_p^2/(\omega^2 \pm \omega\omega_B)], \quad (4.4)$$

$$\omega_p^2 = 4\pi ne^2/m; \qquad \omega_B = (eB/mc) \cos\alpha, \quad (4.5)$$

where n is the electron (mass $= m$) density, and α is the angle between the magnetic field (B) and the direction of propagation. If B is small, Eq. (4.5) yields a dispersion similar to the photon mass effect:

$$v = d\omega/dk = c[1 - (\omega_p^2/\omega^2)]^{1/2} \simeq c[1 - \tfrac{1}{2}(\omega_p^2/\omega^2) + \cdots]. \quad (4.6)$$

Eqs. (3.1) and (4.5) imply that we need to compare the "rest frequency" of the photon in cgs units with ω_p:

$$(\mu c^2/\hbar) = 8.2 \times 10^{47} [\mu(g)] \text{ sec}^{-1}, \quad (4.7)$$

$$(4\pi ne^2/m)^{1/2} = 5.6 \times 10^4 [n(\text{cm}^{-3})]^{1/2} \text{ sec}^{-1}. \quad (4.8)$$

---

[9] Part of this may be due to the inaccessibility of the original work (de Broglie, 1940). It was published in occupied France during World War II, and is difficult to find. In his later works, e.g., de Broglie (1957), he quotes the result, but the calculations are not repeated. Here it is appropriate to point out that de Broglie has had a life-long interest in the question of a photon mass. He first proposed a set of massive photon equations in de Broglie (1934). After that, besides numerous articles, he has written an extraordinary number of original and revised books that discuss the subject, some of which we have mentioned. For a complete bibliography, the interested reader is referred to the Library of Congress listings under has name. See de Broglie (1969–1970).

During the 1930's de Broglie had a strong influence on many young theorists in Paris including Proca and Petiau. The connection between de Broglie's photon equations and the work of others can be seen by looking at the fundamental (reducible) 16-dimensional representation of the Duffin–Kemmer–Petiau wave equation for spin-zero and spin-one particles. The 16-dimensional representation can be defined as a symmetric product space of two Dirac spaces (that is to say, a composite of two Dirac particle spaces). Contrariwise, the de Broglie equations come from the product of a Dirac particle space with a Dirac antiparticle space. When one reduces the Duffin–Kemmer–Petiau equation into irreducible representations, and makes the added assumption to set the parity that the wave function transforms as the product of two Dirac wave functions, one obtains an identically zero, one-dimensional scalar equation, a five-dimensional pseudoscalar (which for plane-wave and diagonal matrix elements is equivalent to the Klein–Gordon equation), and a 10-dimensional spin-one equation (which is equivalent to the Proca equation). As one could surmise from the parity change in going to the Dirac antiparticle space, with the added assumption of transforming as the product of a Dirac particle and an antiparticle wave function the de Broglie equation decomposes into a one-dimensional pseudoscalar, a five-dimensional scalar, and a 10-dimensional axial vector equation. See also Duffin (1938), Kemmer (1939), and Petiau (1936, 1949).

Equations (4.7) and (4.8) show that the dispersion due to a plasma of one electron per cm^3 would equal the dispersion of a photon with the Plimpton–Lawton mass. Despite this limitation, as Feinberg (1969) has pointed out, the observed dispersion in arrival time of radio signals from pulsars provides the most stringent "dynamic" test of the photon mass to date. These data (assuming $\mu=0$) may be used to deduce an average interstellar plasma density of <0.028 electron/cm^3 for radiation from the Crab pulsar NP0532. If the dispersion is partly a photon mass effect, then we have

$$\mu \lesssim 10^{-44}\text{ g} \equiv 3\times 10^{-7}\text{ cm}^{-1} \equiv 6\times 10^{-12}\text{ eV}. \quad (4.9)$$

Feinberg also makes the interesting point that pulse arrival times show no sign of any dispersion, except that implied by the simple quadratic formula Eq. (4.6), over the whole range of frequency from radio to optical. For the Crab pulsar the departure from Eq. (4.6) is $\Delta v/c < 10^{-14}$ (Δt arrival $<10^{-3}$ sec).

We may ask for limits on the kind of low mass "structure" of the photon associated with violation of energy conservation, for example, two poles in $\tilde{D}(k^2)$ separated by $\Delta\mu^2$. As discussed at the end of Sec. II.B, such structure could show itself in two ways. The first is by a spreading in duration of low frequency pulses $\Delta t/T = \Delta\mu^2[2(\omega/c)^2]^{-1}$. Here T is the flight time from source to receiver of the radiation. The most accurate test on this phenomenon is again supplied by the Crab pulsar, with its very narrow (<1 msec) pulse peaks. In fact, there *is* an observed pulse broadening of about 10 msec at 74 MHz frequency (Rankin *et al.*, 1970; Rankin, private communication). This broadening is believed due to "scintillations" or fluctuating irregularities in the interstellar medium. If one assumes that part of it is due to a photon mass spread $\Delta\mu^2$, then we learn that

$$(\Delta\mu^2)^{1/2} \lesssim 5\times 10^{-9}\text{ cm}^{-1}\text{ (pulse broadening)}. \quad (4.10)$$

The second effect of structure in $\tilde{D}$ is variation of intensity of the signal at a given frequency, as a function of time after emission. If the power in the signal is distributed smoothly over a broad range of frequencies, then this effect can also be detected by an oscillation of intensity as a function of frequency at a fixed receiver. For the case of two poles in $\tilde{D}$, one at μ_1^2 with residue $(1+\epsilon)^{-1}$, and one at μ_2^2 with residue $(1+\epsilon)^{-1}\epsilon$, one may derive the modulation of intensity as a function of frequency:

$$I(\omega) = 1 - \{1 - [(1-\epsilon)/(1+\epsilon)]^2\}\sin^2(\Delta kr/2),$$
$$\Delta k = (\mu_1^2 - \mu_2^2)[2(\omega/c)]^{-1} + \mathcal{O}(\mu^4). \quad (4.11)$$

For the Crab pulsar, the lack of conspicuous oscillations down to $\nu=\omega/2\pi = 74$ MHz implies $\epsilon\Delta kr < 4\pi$ at $\nu=100$ MHz or, if ϵ is of order unity,

$$(\Delta\mu^2)^{1/2} < 10^{-11}\text{ cm}^{-1}\text{ (intensity oscillations)}. \quad (4.12)$$

An amusing particular case of Eq. (4.11) arises for $\epsilon = -\mu_1^2/\mu_2^2$. This case is intriguing because, as discussed in Sec. II.B, such a form of $\tilde{D}$ would produce exact cancellation of the lowest-order effects (proportional to μ^2) in electrostatics or magnetostatics. For this special form, the limit Eq. (4.12) on $\Delta\mu^2$ becomes a limit on $\mu_1\mu_2$. If combined with Eq. (4.10), this result implies $\mu_1 \lesssim 10^{-22}\text{ cm}^{-2}/\mu_2$, $\mu_1<\mu_2<5\times 10^{-9}\text{ cm}^{-1}$.

We conclude that pulsar data do more than give the best dynamic (velocity–dispersion) limit on μ. They also provide stringent limits on violation of energy conservation associated with a "multicomponent" photon having two or more different small masses.

B. Magnetostatic Effects

1. Schrödinger's External Field Method

Schrödinger (1943b), following an observation of McConnell, proposed a method using the earth's static magnetic field that has yielded the best photon mass limit to date. Let us begin with a discussion of the principles on which the method is based. As mentioned in Sec. II.A, the qualitative effect of a photon of mass μ on static fields is to cause an extra "Yukawa" decrease in field strength as $e^{-\mu r}$, where r is distance from the source. However, we also have seen that even in massive electrodynamics, the divergence of the magnetic field $\mathbf{H}$ must vanish. This is simply a consequence of reflection symmetry in electrodynamics, and the absence of magnetic sources: Consider the magnetic field at a point $\mathbf{r}$ produced by electric currents $\mathbf{J}$ at points $\mathbf{r}'$. Any divergence of $\mathbf{H}$ would be a pseudoscalar function of $\mathbf{J}$ and $\mathbf{r}-\mathbf{r}'$, but there is no such function. The dependence only on $\mathbf{r}-\mathbf{r}'$ is a consequence of assuming that the equations of physics are independent of the choice of origin of coordinates.[10] Applying these thoughts to the magnetic dipole field of the Earth, we note $\nabla\cdot\mathbf{H}=0$ means that the flux in each field line is conserved. Now a field line is farther out at the magnetic equator than it is near the pole. Hence, the Yukawa exponential decrease affects the field line most at the equator. To keep constant flux, the field pattern must *change shape*, allowing flux lines to move in somewhat at the equator. This compression of the equatorial field lines has the effect, on a sphere of fixed radius, of increasing the field at the equator relative to the field at the pole. The effect is the same as that of a constant external field parallel to $\mathbf{H}_{\text{equatorial}}$. Of course, the field of the Earth is not pure magnetic dipole. However, for the massless Maxwell theory it is a theorem that only a true external current can produce a uniform field over the surface of a sphere.[11] In the absence of such currents,

[10] These comments simply express in familiar three-dimensional terms the arguments of Sec. II.B, which were given in relativistic, four-dimensional notation.

[11] This is most easily proven by noting that if there are no external currents, $\mathbf{H}$ outside the sphere is the gradient of a solution of the Laplace equation, and then expanding that solution in spherical harmonics.

the average of any component of **H** over the sphere must vanish. Therefore, if true external currents be estimated, detection of any anomalous uniform "external" field on the surface of the Earth would demonstrate a violation of Maxwell's theory of just the form produced by a finite photon mass.

We turn to a quantitative derivation of the external field effect. Recalling the definition of a magnetic dipole moment

$$D\hat{z} \equiv \mathbf{D} = (-\tfrac{1}{2}) \int d^3r \mathbf{J} \times \mathbf{r}/c, \tag{4.13}$$

and the static limit of Proca's Eq. (2.18)

$$(-\nabla^2 + \mu^2)\mathbf{A} = \mathbf{J}(4\pi/c), \tag{4.14}$$

we have the dipole vector potential

$$\mathbf{A} = \nabla \times \mathbf{D}[e^{-\mu r}/r]. \tag{4.15}$$

Since **H** is $\nabla \times \mathbf{A}$, the dipole contribution to the earth's field, measured from coordinates centered at the dipole,[12] is

$$\mathbf{H} = (De^{-\mu r}/r^3)[(1+\mu r+\tfrac{1}{3}\mu^2 r^2)(3\hat{z}\cdot\hat{r}\hat{r}-\hat{z}) - \tfrac{2}{3}\mu^2 r^2 \hat{z}]. \tag{4.16}$$

This is to be compared to the ordinary dipole field,

$$\mathbf{H}_D = (D/r^3)(3\hat{z}\cdot\hat{r}\hat{r}-\hat{z}). \tag{4.17}$$

If observations are made near the surface of a sphere centered at the dipole, (here it will be the surface of the earth $r \cong R =$ const), the factors in the first term of Eq. (4.16) would just make D appear to have a slightly different value when compared to Eq. (4.17). However, the last term in Eq. (4.16) is new. It will be observed as an apparent external magnetic field *antiparallel* to the direction of the dipole. The ratio of the "external field" (H_{ext}) to the dipole field at the equator (H_{DE}) is

$$H_{\text{ext}}/H_{\text{DE}} = \tfrac{2}{3}(\mu R)^2/[1+\mu R+\tfrac{1}{3}(\mu R)^2] \tag{4.18}$$

Using the 1922 magnetic surveys discussed by Schmidt (1924),[13] Schrödinger obtained a ratio of $(H_{\text{ext}}/H_{\text{DE}}) = (539\gamma)/(31089\gamma)$, where $1\gamma = 10^{-5}$ G.[14] Putting this into Eq. (4.18) yields $\mu = 2.76 \times 10^{-10}$ cm^{-1}. Later, Bass and Schrödinger (1955) argued that multiplying this estimate by a factor of 2 would give a reliable upper limit $\mu \lesssim 5.5 \times 10^{-10}$ cm$^{-1} \equiv 1.9 \times 10^{-47}$ g.

The critical reader will have noticed that Schrödinger's results did not use extraterrestrial measurements at all. In fact the ingenuity of the external field method is precisely that it requires only ground observations. However, the more recent work on the geomagnetic limit on μ does exploit satellite data as well as earth-based measurements. Furthermore, the Schrödinger method might be applied to other planets in the not too distant future. These facts seem a sufficient defense for our classification.

Recently the present authors (1968) improved on Schrödinger's results by using Cain's fit[15] (Cain, 1966; Cain *et al.* 1965, 1968; Hendricks and Cain, 1965) to geomagnetic data from earthbound and satellite measurements. For epoch 1960.0 Cain obtains values[16] of

$$\begin{aligned} D &= 31044\gamma R^3, \\ \mathbf{H}_{\text{ext}}\cdot\hat{z} &= (21\pm5)\gamma, \\ \mathbf{H}_{\text{ext}}\cdot\hat{w} \equiv \mathbf{H}_{\text{ext}}\cdot(\hat{s}\times\hat{z})/|\hat{s}\times\hat{z}| &= (14\pm5)\gamma, \\ \mathbf{H}_{\text{ext}}\cdot\hat{w}\times\hat{z} &= (8\pm5)\gamma, \end{aligned} \tag{4.19}$$

where $\hat{s}$ points towards the south *geographic* pole.

To obtain our mass limit from this number, we had to take into account the "true external" sources, which were unknown when Schrödinger wrote his paper. As discussed in Goldhaber and Nieto (1968) these include $\sim 9\gamma$ from the quiet-time proton belt, perhaps 15–30γ due to currents in the geomagnetic tail, and $\sim 15\gamma$ from the hot component of the plasma in the magnetosphere, which are all *parallel* to the dipole moment. In addition, there is a true external field antiparallel to the dipole of $\sim 20\gamma$ at the equator, which is due to the compression of the geomagnetic field by the solar wind. Finally, the interplanetary field of $\sim 5\gamma$ points in an unknown direction at the earth's surface.[17]

Thus, the total external field parallel to **D** due to known sources is $\lesssim 40\gamma$. Subtracting this from $\hat{z}\cdot\mathbf{H}_{\text{ext}}$ in Eq. (4.18) gives an upper limit on the *antiparallel* external field which could be due to a finite photon mass, $\mathbf{H}_{\text{ext}}$(antiparallel) $\lesssim 20\gamma$.

The significance of this limit depends crucially on the reliability of the fit of Cain to the geomagnetic field. In Goldhaber and Nieto (1968) we made a number of observations concerning this reliability: (1) The existence in Eq. (4.19) of external components *perpendicular* to D is hard to explain in any known physical model; (2) There appears to be an irreducible "noise" in data from earthbound observatories of about 100 γ, in large part due to magnetic anomalies in the earth's crust; (3) Earthbound data are very sparse in Asia and much of the southern hemisphere. Despite this, the fit is made by an expansion in spherical harmonics of a potential whose gradient gives **H**, even though these

[12] The orientation of the earth's dipole moment is given by Finch and Leaton (1957). A more recent value is to be found in Cain and Hendricks (1968). In the usual physics convention, this dipole points to the southern hemisphere.

[13] See also Chapman and Bartels (1940).

[14] The equations we are using are in cgs electrostatic units, so that magnetic fields are not measured in the Gauss of cgs electromagnetic units. In Sec. IV.B.1, this does not matter since we are calculating ratios of fields. However, in Sec. IV.B.2, one must be sure to insert the correct unit conversion factor (which is essentially c).

[15] Other reports on this study are contained in Hendricks and Cain (1966), and Cain *et al.* (1965).

[16] In consulting the literature, care should be taken to differentiate between the sign conventions used for the spherical harmonics, and between north-seeking and north poles. See, for example, Table 3 in Fougère (1965), as well as Cain *et al.* (1967), Appendix. This is a later fit than that of Cain (1966), but it does not calculate possible external terms.

[17] A more detailed discussion on these points and pertinent references can be found in Goldhaber and Nieto (1968).

harmonics form a complete orthonormal set only over an *entire* sphere; (4) Even low noise satellites cannot clear up the picture because there are peculiar secular (time) variations of the fits (Cain, private communication) to the satellite data of the order of tens of γ, and also because up to now only the magnitude but not the direction of **H** has been measured with satellites.

For these reasons, the possibility of systematic errors in the fits to a particular spherical-harmonic coefficient of many tens of γ cannot be excluded. To take account of this, and any errors in the estimates of true external fields, we added 100 γ to our estimate, meaning that the left-hand side of Eq. (4.18) is less than 4×10^{-3}. Simple numerical work then gives us a limit on the mass of the photon of[18]

$$\begin{aligned}\mu &\leq 10^{-10}\ \text{cm}^{-1}\\ &\equiv 3\times10^{-15}\ \text{eV}\\ &\equiv 4\times10^{-48}\ \text{g},\end{aligned} \tag{4.20}$$

which is five times better than the number that Schrödinger suggested on the basis of much less precise and very much less reliable data. To our knowledge this is the smallest energy or energy limit yet measured for anything.[19] [The limit in Eq. (4.20) corresponds to a photon Compton wavelength[20] of $80R$.]

2. *Altitude-Dependent Method*

One might hope to use satellite measurements at varying altitudes to detect the exponential decay of **H**. From Eq. (4.16) the magnitude of the dipole field for $\mu\neq0$ is

$$\begin{aligned}H_D(\mu) &= \{1+\tfrac{1}{2}(\mu r)^2[(1-5\cos^2\theta)/(1+3\cos^2\theta)]\}H_D,\\ &\equiv F(\mu, r, \theta)H_D,\end{aligned} \tag{4.21}$$

where $\cos\theta=\hat{D}\cdot\hat{r}$, H_D is the magnitude of the orthodox dipole field of Eq. (4.19), and cubic terms in μr are neglected.

Gintsburg (1963), the first to apply the altitude-dependent method, used the assumption $F=1-(\mu r)^2$ to obtain a mass limit from magnetic measurements at varying altitudes by Vanguard, Explorer, and Pioneer satellites. He gives a limit $\mu<3\times10^{-48}$ g, and states that this may be too low by a factor of 2 or 3. After reconsidering his numbers in the light of Eq. (4.21) we [Goldhaber and Nieto (1968)] felt that the same conservative error estimation that we had used would make Gintsburg's limit $(8\text{–}10)\times10^{-48}$ g. Thus, it is nearly a geometric mean between the old and new results of the Schrödinger method. The main limitation on the altitude-dependent method is that external perturbations become quite significant beyond $\sim3R$ (Frank, 1967).

3. *Eccentric Dipole or "Vertical Current" Effect*

In his early paper, Schrödinger (1943b) pointed out an effect due to the displacement **b** of the magnetic dipole origin from the geocenter. From Eq. (4.15), the lines of vector potential are circles around the magnetic dipole axis, and hence would intersect the surface of the Earth at a small angle of order $|\mathbf{b}|/R\approx1/19$. From Eq. (2.19), this means that there is a "feigned" vertical current from the component of **A** perpendicular to the Earth's surface. That is, the integral of **H** along a closed path on the surface would fail to vanish—**H** could not be written as the gradient of a scalar potential. To look for this effect, Schrödinger consulted Schmidt's (1924) vertical current measurements for the independent surveys of 1885 and 1922. Schmidt had generally found ascending currents in the Eurasian-African hemisphere, and descending currents in the American–Pacific one in 1885. This pleased Schrödinger because these signs are implied by a finite μ. However, the later survey indicated that the currents were smaller and fluctuated in sign from point to point within the same hemisphere. This eventually caused Schmidt to doubt his earlier "hemispheric current effect" (Schmidt, 1924, 1939).[21]

Still, from the data, Schrödinger was able to speculate at a figure of $\pm50.8\times10^{-3}$ A/km² for the maximum and minimum vertical currents. To compare this to his theory he used his then quoted result $\mu=2.76\times10^{-10}$ cm^{-1}, from the "Schrödinger Method" of Sec. IV.B.1. He obtained

$$J_{\text{max theoretical}}=6.4\times10^{-3}\ \text{A/km}^3. \tag{4.22}$$

Although this was smaller than the "measured" value of 50.8×10^{-3} A/km², it was within an order of magnitude, and of the correct sign. Schrödinger took hope from this since at that time he had field-theoretic reasons for believing in a photon mass of about this

[18] After our comments on numerical errors by others, we must candidly acknowledge that in Goldhaber and Nieto (1968), we gave a value 10% too low for μ (1.15 instead of 1.25×10^{-10} cm^{-1}); this was chiefly due to using an incorrect value for the radius of the Earth!

[19] The K_1-K_2 mass difference, for example, is 0.63×10^{-38} g. We observe, however, that from observations on the sizes of clusters of galaxies one could set a limit to the mass of the "graviton" possibly as low as 10^{-24} cm$^{-1}=4\times10^{-62}$ g. But this ignores the troublesome point that no one has ever observed a real live single graviton (nor should we expect that anyone will soon). Also, it is not totally clear that the gravitational field should be quantized in the first place. [See, e.g., Lecture 1, p. 12 of Feynman (1962–1963).] We shall comment further on these points elsewhere.

[20] Note that in the units we are using, the photon Compton wavelength is given by $\lambda_\gamma=2\pi/\mu$.

[21] With hindsight we can say that perhaps Schmidt was too pessimistic in dismissing his early observations. As we now know, sunspot activity plays a heavy role in the currents circulating in the radiation belts. Schmidt's 1885 survey was in a year of heavy sunspot activity (52.2) near the sunspot maximum of 1883.9, whereas the 1922 survey was in a year of low activity (14.2) near the sunspot minimum of 1923.6, which could have caused the change in his results. See Waldmeier (1941), who discusses sunspot observations. Waldmeier's figures are also given by Kiepenheuer (1951).

size.[22] Realistically, though, the observed larger currents and the huge fluctuations within the same hemisphere certainly meant that one would have had to weed out the "real" currents to obtain as low a mass limit as that found via the "external field."

However, there is a more elementary reason for giving up this approach. The effect produced by a massive photon via its "eccentric dipole" compared to that produced via its "external field" is intrinsically smaller by a number of order (b/R), which for the earth is $\simeq 1/19$. Thus, the "eccentric dipole effect" will not yield as good a mass limit as that which could be obtained by a similar experimental effort using the "external field effect," even if the real currents could be separated out.

C. Long-Period Magnetic Waves

Gintsburg (1963) proposed a method to limit the photon mass that deals with the propagation of long period magnetic waves. The idea focuses on the form of the index of refraction $(N=|\mathbf{k}|c/\omega)$ of electromagnetic radiation in a cold, nondissipative, magnetized plasma (cf. Thompson, 1962, Chap. 7; and Jackson, 1962, Sec. 10.8). There are two types of magnetic waves in such a plasma. In the first, or acoustic type, the wave propagates perpendicular to the direction of the static magnetic field $\mathbf{H}_0$, and the oscillating field points in the same direction as $\mathbf{H}_0$. The dispersion equation for these magnetosonic waves is

$$\mathbf{k}^2=\omega^2/V_A^2, \tag{4.23}$$

where V_A is the Alfvén velocity $V_A=H_0/(4\pi\rho)^{1/2}$, and ρ is the mass density. If the plasma were not cold, so that the speed of sound s was no longer negligible compared to V_A, then Eq. (4.23) would be modified by the replacement $V_A\rightarrow(V_A^2+s^2)^{1/2}$.

The second, and purely magnetic, wave is the Alfvén wave, in which the magnetic field oscillates perpendicular to $\mathbf{H}_0$, and the wave travels in the plane perpendicular to the oscillating field, at an angle θ to $\mathbf{H}_0$. Now the dispersion equation is

$$\mathbf{k}^2\cos^2\theta=\omega^2/V_A^2. \tag{4.24}$$

The factor $\cos^2\theta$ arises from the powerful constraints relating electric and magnetic fields in a highly conductive plasma. It is natural to describe the quantity V_A^{-2} in Eqs. (4.23) and (4.24) as ϵc^{-2}, where ϵ is the effective dielectric constant. However, this is a highly sophisticated use of the term, as one may see from the fact that, as long as the plasma is motionless, the electric field must vanish because of the nearly infinite conductivity of the plasma, so that ϵ is infinite in this limit. Nevertheless, one is enticed by the form of these equations to introduce finite photon mass simply by subtracting μ^2 from the right-hand side. In fact this procedure can be justified by introducing the Proca Eqs. (2.19) in place of Maxwell's, and following carefully the modifications in the standard derivations of the dispersion equations (Jackson, 1962, Sec. 10.8). We are left with the new relations

$$\mathbf{k}^2=\omega^2/V_A^2-\mu^2 \quad \text{(magnetosonic)}, \tag{4.25}$$

and

$$\mathbf{k}^2\cos^2\theta=\omega^2/V_A^2-\mu^2 \quad \text{(Alfvén)}. \tag{4.26}$$

For $\mu\neq 0$, $\mathbf{k}^2$ would become negative for low enough ω, so that the waves would be damped. In principle, if the critical frequency at which the waves did not propagate could be found, the photon mass could be measured directly. Of course, the lowest mass that could be measured for any given frequency or period (T) would be

$$\mu=\omega/c=(2.09\times 10^{-10}\ \text{cm}^{-1})/T(\text{sec}), \tag{4.27}$$

no matter what the plasma density or field strength.

Patel (1965) proposed a mass limit from this idea by using the data of Patel and Cahill (1964). These authors had combined ground magnetometer and Explorer XII results to surmise that hydromagnetic waves having a period of 200 sec were generated at about 50 000 km from the center of the earth and then traveled along the magnetic field lines to the earth's surface in about 1.5 min, during which time their amplitude was attenuated by one-third. From this Patel conjectured that it was reasonable to consider these waves as being at or above the critical frequency. By taking values for ρ of 50 $M_{\text{proton}}/\text{cm}^3$ and for $|\mathbf{H}_0|$ of 10^{-3} G, and using these values in Eqs. (4.25)–(4.26) for $\mathbf{k}=0$, Patel obtained the mass value (restoring an omission of π)

$$\mu=10^{-9}\ \text{cm}^{-1}. \tag{4.28}$$

He then took this value as the upper limit to the photon mass, meaning

$$\mu\leq 10^{-9}\ \text{cm}^{-1} = 4\times 10^{-47}\ \text{g}. \tag{4.29}$$

However, there are numerous reasons for questioning this value. Patel himself pointed out that it could easily be off by an order or two of magnitude because of uncertainties in the values assumed for the plasma density and the magnetic field strength. Along the path of the waves there could have been very large fluctuations in these parameters, which would have produced a marked effect on the wave propagation characteristics.

To add more confusion, there obviously is huge uncertainty in the assumption that the damping came from a massive photon, and not from other dissipative

[22] Schrödinger (1943a) [and a note added in proof in his later paper (Schrödinger, 1943b)] had developed a "Unitary Field Theory" which predicted

$$\mu^{-1}=\tfrac{1}{6}\Gamma(\tfrac{1}{4})^4(6\pi)^{-3/2}(e/mc)^2(hc/k)^{1/2}=29\ 715\ \text{km},$$

which agreed well with his "Schrödinger Method" mass limit $\mu^{-1}=36\ 300$ km.

mechanisms. For example, if one takes the order-of-magnitude lower mass limit that is obtained from the Schrödinger method, then there was no reason, in Patel's analysis, for these waves to be damped.

In order to establish these waves as being near the critical frequency, one would have to observe higher frequency waves propagating for at least a few wavelengths, and lower frequency waves not propagating.

Also, the damping distance is greater than μ^{-1} for Alfvén waves propagating along $\mathbf{H}_0$, so that meaningful values for a mass limit would require propagation distances sufficient for attenuation to occur. This would make 2×10^{-10} cm^{-1} about the best limit obtainable with this type of data, assuming no uncertainties about the medium of propagation.

To sum up, although this idea is very interesting, a convincing mass limit would require the solution of many difficulties.

V. LONGITUDINAL PHOTONS

As mentioned briefly in Sec. II.A., $\mu\neq0$ implies the existence of a third degree of freedom for the photon: longitudinal polarization. This immediately suggests potentially significant effects, of which the most dramatic is a fifty per cent increase in the stored energy of a system of photons in thermodynamic equilibrium with a reservoir at temperature T. Since Planck's radiation law is known to be quite precise, this effect is not present. Can we not conclude that μ must vanish? The answer was given most vividly by Bass and Schrödinger (1955) [See also Stueckelberg (1957)] who pointed out that the approach to equilibrium of longitudinal photons in a cavity is very slow, having a time scale comparable with the age of the Universe. They considered a closed box with perfectly conducting walls, and asked how long it would take for energy originally stored in transverse waves to be partitioned equally among all three polarizations. While they came up with a very long time for any reasonably big container, the basis of their calculation is open to question, since even the best known conductors are not good enough to reflect longitudinal waves. The key concept here is the skin depth of the conductor. It is straightforward to show that the skin depth for longitudinal waves of angular frequency ω is given by (Gertsenshtein and Solovei, 1969; Kroll, 1971)

$$\delta_L\geq(\omega/\mu c)^2\delta_T, \tag{5.1}$$

where δ_T is the skin depth for transverse waves. The strict equality holds only for low conductivity and, in fact, S_L becomes infinite in the limit of infinite conductivity (Kroll, 1971)! Let us consider photons of typical frequency for a system at a temperature of 10^{-3} °K, near the lowest attainable at present. This would mean $\omega>10^8$ Hz, a skin depth δ_T certainly bigger than 10^{-5} cm, and therefore $\delta_L>10^{10}$ cm, comparable with the radius of the sun! For temperatures of interest in astronomical processes, the factor $(\omega/\mu c)^2$ is sufficient to guarantee the complete transparency of stellar material to longitudinal photons, even assuming a mean free path of 10^{-13} cm for transverse photons.

Longitudinal photons would not supply an important mechanism for energy loss by stars because the production rate is far too small. A nucleus which can emit a 1 kV x ray, for example, would emit a longitudinal x ray with probability $\frac{1}{2}(\mu c/\omega)^2\approx10^{-36}$ and absolute rate of perhaps 10^{-20}/sec, again taking longer than the age of the Universe. Thus, even with the highest densities and longest time scales known, longitudinal photons would have negligible effect on thermodynamic systems.[23]

The possibility of detecting longitudinal photons is an intriguing one. For example, a capacitor might be placed inside a conducting sphere: Any incident longitudinal wave would pass through the sphere and deposit a small fraction of its energy in the capacitor. Capacitors outside the sphere would receive comparable energy, but for transverse waves they would receive much more than the shielded capacitor. The difficulty with this suggestion lies in conceiving of a sufficiently powerful source for L waves. The sun is the most obvious candidate. From the earlier discussion, the sun might be opaque to longitudinal photons with $\omega=10^8$ Hz. Bass (1956, 1963) has pointed out that if the usual thermodynamic considerations are applied to black body radiation from a slab of thickness X, and the mean free path of longitudinal photons of a given frequency is Λ, then the L-wave radiation from the slab will have an intensity

$$I_L=[1-\exp(-X/\Lambda)]I_T, \tag{5.2}$$

where I_T is the intensity of either transverse polarization. Let us assume for the application at hand that the factor in parentheses is unity for $\omega=10^8$ Hz; then I_L will be just half of the ordinary black body radiation. A rough estimate indicates that the L-wave radiation arriving at the earth would produce a root mean square voltage across a capacitor with 1-m^2 area about 10^5 times smaller than the Johnson noise across the capacitor (with a parallel resistance of $10^8\,\Omega$). Thus the chances of observing thermal longitudinal photons are remote indeed. Instead, one must hypothesize an implausible coherent charge oscillation in the sun at low frequency ($\omega\lesssim10^8$ Hz) in order to predict an observable effect. It might be worthwhile to scan the sky in search for such a coherent source of L waves.

One may ask whether the statement that reaction or emission rates for longitudinal photons are suppressed by a factor $(\mu c/\omega)^2$ compared to those for transverse photons remains true for all quantum transitions in atomic or smaller systems. In fact it is not strictly true.

---

[23] An early criticism along these lines was voiced by Wigner (1956).

For example, it is well known that a transition between two states, each with angular momentum zero, cannot proceed by emission of a single transverse photon. However, it can go by L-wave emission, albeit so slowly that double-photon emission is much more likely. The rate for such a transition is suppressed by a factor $(\mu r)^2$ compared with an allowed electric dipole transition of the same frequency. Here r is a characteristic (length) dimension of the quantum system in question. For a photon-induced transition to an arbitrary state f, the quantum mechanical matrix element is

$$T_{fi} = -\epsilon^{\alpha}\langle f \mid J_{\alpha} \mid i\rangle, \tag{5.3}$$

where i is the initial target state. The polarization vector ϵ may be normalized by the condition $\epsilon_{\alpha}\epsilon^{\alpha} = -1$. For transverse polarizations this means

$$\epsilon_T^{(1)} = (0, 1, 0, 0)$$
$$\epsilon_T^{(2)} = (0, 0, 1, 0), \tag{5.4}$$

and for the longitudinal polarization

$$\epsilon_L = (\mid \mathbf{k} \mid/\mu, 0, 0, \omega/\mu c), \tag{5.5}$$

where we take the photon momentum in the z direction, and an arbitrary coordinate 4-vector would be written as (ct, x, y, z). Current conservation implies the vanishing of T_{fi} if the photon 4-momentum k^{α} is substituted for ϵ^{α}. This immediately gives

$$T_{fi}^{(L)} = (-\mathbf{k}c/\mu\omega + \omega/\mu c)\langle f \mid J_z \mid i\rangle$$
$$= (\mu c/\omega)\langle f \mid J_z \mid i\rangle \tag{5.6}$$

but, for transverse photons,

$$T_{fi}^{(X,Y)} = \langle f \mid J_{x,y} \mid i\rangle. \tag{5.7}$$

Thus, if transverse and longitudinal matrix elements of J_{α} are comparable, then amplitudes for absorbing or emitting longitudinal photons are suppressed by $(\mu c/\omega)$, and rates or cross sections are suppressed by the square of this factor. In the classical (Thomson) limit, matrix elements of J_{α} are independent of direction, and the suppression factor is exactly $(\mu c/\omega)^2$. The only way to get large longitudinal cross sections is to have much bigger matrix elements for longitudinal currents, or else for more longitudinal- than transverse-photon induced transitions.

A specific model for high-energy cross sections of longitudinal photons on hadrons is supplied by the "vector meson-current identity," (see, for example, Sakurai, 1969, 1969a), which states that the electromagnetic current operator is a linear combination of the fields which create the ρ, ω, and ϕ mesons. Since ρ-hadron cross sections appear to approach a constant spin-independent value at high energies, we require for consistent normalization on the scattering amplitudes

$$\langle A \mid J_z \mid B\rangle = (\omega/m_{\rho}c)\langle A \mid J_x \mid B\rangle, \tag{5.8}$$

where m_{ρ} is the ρ-meson mass in inverse length units. The ω and ϕ make small contributions, but their masses are even bigger, and so do not alter the point made now. The result obtained for J_z/J_x implies for the total cross sections of photons

$$\sigma_L/\sigma_T = (\mu/m_{\rho})^2. \tag{5.9}$$

Thus, the ratio of mean free paths Λ_L/Λ_T goes as $(\omega/\mu c)^2$ for low frequencies, but as $(m_{\rho}/\mu)^2$ for high $\omega(\gg 10^{23}$ Hz) in this model.[24] It would require an *extreme* modification to alter the conclusion that longitudinal cross sections are infinitesimal at all frequencies except ω of order μc.

We conclude that, if longitudinal photons exist, their observation will be difficult if not impossible, and their effects are negligible, even on an astronomical scale.[25]

VI. CONCLUSIONS

A. Future Experiments

Let us re-examine each of the categories discussed so far and see if the results could be improved.

1. *Terrestrial Measurements of c*

Comparison of past results here with those of other methods is not encouraging. To give a feeling for the difficulty in making improvements, consider the classic technique of measuring c by observing the normal mode frequency of a cavity with conducting walls. If the cavity is a cube of side a then the lowest mode frequency (which would be transverse in the $\mu=0$ case) is completely independent of μ! The solution for the electric field $\mathbf{E}$ in this $\mu\neq 0$ mode is

$$\mathbf{E} = \hat{\mathbf{z}}E_0 \cosh \mu(z-a/2) \sin(\pi x/a) \sin(\pi y/a) \cos \omega t, \tag{6.1a}$$

with

$$\omega = \sqrt{2}\pi c/a. \tag{6.1b}$$

The only effect of μ is an almost imperceptible bulge in the magnitude of $\mathbf{E}$ in the middle of the box. If we think of this field as a superposition of travelling waves, then these waves still have phase velocity c, independent of μ. But they are no longer truly transverse since $\mathbf{E}$ obeys the divergence relation

$$\nabla\cdot\mathbf{E} = -\mu^2 V = \mu E_0 \sinh \mu[z-(a/2)]$$
$$\times \sin(\pi x/a) \sin(\pi y/2) \cos \omega t \neq 0. \tag{6.2}$$

The reason for this remarkable result, which tends to go unnoticed in $\mu=0$ electrodynamics, is that the phase velocity of a wave enclosed by conducting walls is determined by the parameter c that appears when

[24] Data from inelastic electron scattering do not decisively determine the asymptotic behavior of $\sigma_L/\sigma_T = (0.2\pm 0.25)$. See Taylor (1969).

[25] In particular, we do not agree with Bass (1956, 1963) who proposed longitudinal photon emission as a mechanism for cooling of the earth's interior. A crucial element in his calculation is the mean free path of transverse photons in the earth. To obtain this he uses a formula of Eddington designed to apply to stellar plasma. The calculated path is about 10^{10} times too short, since the most elementary considerations imply a path of about one micron.

Ampere's and Faraday's laws are expressed in units of charge and velocity. These waves are *not* free electromagnetic waves. Free and bound waves have the same phase velocities for $\mu=0$, but for $\mu\neq 0$ the free waves are slower! The bound waves are really an effect of charge and current oscillations nearby, completely distinct in origin from electromagnetic radiation. In fact, the bound fields propagate parallel to the current, while radiation is emitted perpendicular to an oscillating current source (Goldhaber and Nieto, 1971; Kroll, 1971).

A related difficulty arises even for long radio waves in the atmosphere. Aside from all other uncertainties, if the wave can propagate in the "dissipative waveguide" made by the Earth's surface and the ionosphere, then it can propagate with phase velocity c even if μ is nonzero. A μ-dependent dispersion of velocity can appear only for wavelengths λ obeying the condition

$$\lambda/h \ll 1, \tag{6.3}$$

where h is the height of the ionosphere or the width of the "waveguide." Such a short wave propagates as ordinary electromagnetic radiation (except for dispersion introduced by the atmosphere and by the irregular surface of the Earth). Thus, although it is possible that the long-wavelength radio dispersion measurements of Mandel'shtam *et al.* (1944) could be improved somewhat, the above argument indicates that a considerable increase in wavelength would make the Mandel'shtam method irrelevant to photon mass limits (see also Kroll, 1971).

2. *Deviations from Coulomb's Law*

Since the greatest promise for improvement in the Cavendish experiment is associated with high frequency, it is worth discussing the theory for nonstatic fields at some length. Let us consider a spherical cavity whose boundary is an equipotential surface, and ask for the electric field at angular frequency ω. From the symmetry, the potential V is a function only of r, and $\mathbf{E}$ points in the radial direction. We have

$$[\nabla^2+\omega^2/c^2-\mu^2]V=0 \tag{6.4}$$

yielding

$$V=[V_0 \sin|\mathbf{k}|r/|\mathbf{k}|r]\cos\omega t,$$
$$|\mathbf{k}|=[(\omega^2/c^2)-\mu^2]^{1/2}. \tag{6.5}$$

The shortcut to finding $\mathbf{E}$ is to use Eq. (2.19)

$$\nabla\cdot\mathbf{E}=-\mu^2 V \tag{6.5'}$$

to get

$$\mathbf{E}=\frac{-\hat{r}}{r^2}\int_0^r r'^2\, dr' \mu^2 V(r',t). \tag{6.6}$$

A spurious charge density produces a field within the cavity. What interests us is the electric potential

$$\Phi=-\int_0^r dr' E_r(r'),$$

which would be measured by a voltmeter. This means

$$\nabla^2\Phi=-\nabla\cdot\mathbf{E}=\mu^2 V. \tag{6.7}$$

Since $\mathbf{E}$ obeys the wave equation in space free of charges and currents, so must Φ, except for a radially constant term, giving us

$$\nabla^2\Phi(r,t)=-\mathbf{k}^2\Phi(r,t)+g(t), \tag{6.8}$$

$$\Phi=-(\mu^2/\mathbf{k}^2)V+(\mu^2/\mathbf{k}^2)V(0,t)$$
$$=(\mu^2/\mathbf{k}^2)V_0[1-(\sin|\mathbf{k}|r/|\mathbf{k}|r)]\cos\omega t$$
$$=(\mu r)^2 V_0\cos(\omega t) f(|\mathbf{k}|r), \tag{6.9}$$

with

$$f(x)=(x-\sin x)/x^3,$$
$$f(\pi/4)\simeq\tfrac{1}{6}. \tag{6.10}$$

For any frequencies such that $\omega r/c$ is much less than unity, this expression for Φ agrees with that calculated in the static case, except that the next correction is not of order $(\mu r)^4$, but of order $(\mu r)^2(|\mathbf{k}|r)^2$. In order to complete this discussion we must know the physical significance of V_0 in Eq. (6.9). Let us consider an ideal case in which the charge density is spherically symmetric and oscillates with frequency ω. We want the fields in the region $r<r_0$, with the condition

$$\rho(r,t)=\rho(r)\cos\omega t\equiv 0, \qquad r<r_0. \tag{6.11}$$

We obtain V from Eq. (2.17)

$$(\nabla^2+\mathbf{k}^2)V=-4\pi\rho, \tag{6.12}$$

$$V(r<r_0)$$
$$=\frac{4\pi}{|\mathbf{k}|r}\int_{r_0}^{\infty} r'\,dr'\rho(r')\cos|\mathbf{k}|r'\sin|\mathbf{k}|r\cos\omega t. \tag{6.13}$$

Suppose that there is simply a shell of total charge $Q\cos\omega t$ at radius a. Then we get

$$V_0=(Q/a)\cos|\mathbf{k}|a, \tag{6.14}$$

and V_0 is very close to what we would have naively expected as the potential on the sphere of radius a. Of course, to maintain such an oscillating charge shell there must be a radial current out to infinity, or equivalently, to a second sphere of radius b such that $\cos|k|b$ vanishes. This idealized case is not easily realized in practice. A more realistic case is one in which a perfect conducting sphere lies at radius a, and there is a point charge $Q\cos\omega t$ somewhere just above the surface. As long as the conductor is perfect, so that $\mathbf{E}\parallel\hat{r}$ inside, Φ and V_0 will be given by the same expressions as before. This may be verified by expanding the solution in Legendre polynomials and imposing the spherical boundary condition on Φ. In practice, if all external sources are outside a conducting shell whose thickness is large compared to its skin depth at frequency ω, then Eqs. (6.9) and (6.14) should be excellent approximations inside the shell. The optimum frequency ω is one for which $f(x)$ and V_0 are near their maxima, but

ω is as big as possible to minimize thermal noise compared to the signal. We choose arbitrarily $2a_{\text{shell}}=1$ m, $|\mathbf{k}|\, a=x=\pi/4$. This gives $f(x)\approx 0.16$ (almost $\frac{1}{6}$), $\cos|\mathbf{k}|\, a=2^{-1/2}$, and

$$\Phi(r)\approx 0.12(\mu r)^2 Q/a, \qquad \omega\approx 2\times 10^8 \text{ Hz}. \quad (6.15)$$

One may ask if some other geometry would have superior sensitivity to the spherical arrangement used in most previous experiments. To deal with this, we prove the following:

Theorem: If field generating apparatus and field detectors are confined to a region with maximum dimensions of length D, then effects of a finite photon mass μ are of order $\mathcal{O}[(\mu D)^2]$ or smaller, regardless of field oscillation frequency.

Proof: We recall from Sec. II the expression for the field in momentum space

$$\tilde{F}_{\alpha\beta}(k)=-(4\pi/c)i[1/(\mu^2-k^2)](k_\alpha \tilde{J}_\beta(k)-k_\beta \tilde{J}_\alpha(k)), \quad (6.16)$$

and rewrite this as

$$\tilde{F}_{\alpha\beta}(k)=[1-\mu^2/(\mu^2-k^2)]\tilde{F}_{\alpha\beta}^{(0)}(k), \quad (6.17)$$

where $\tilde{F}_{\alpha\beta}^{(0)}(k)$ is the field produced with $\mu=0$. In coordinate space, with fixed frequency ω, we get

$$F_{\alpha\beta}(\omega,\mathbf{r})=F_{\alpha\beta}^{(0)}(\omega,\mathbf{r})+\delta F_{\alpha\beta}(\omega,\mathbf{r}),$$
$$\delta F_{\alpha\beta}(\omega,\mathbf{r})=-(\mu^2/4\pi)\int d^3r' \exp(i|\mathbf{k}|\,|\mathbf{r}-\mathbf{r}'|) \times F_{\alpha\beta}^{(0)}(\omega,\mathbf{r}')\,|\mathbf{r}-\mathbf{r}'|^{-1}. \quad (6.18)$$

As long as the integral converges well it provides an excellent means of estimating corrections due to finite photon mass. In particular, if charges and currents are confined to a region of dimensions $<D$ in all directions, and observations are made in that region, then we have

$$\delta F_{\alpha\beta}=\mathcal{O}(\mu^2 D^2)\bar{F}_{\alpha\beta}^{(0)}, \quad (6.19)$$

where $\bar{F}_{\alpha\beta}^{(0)}$ is a typical value of $F_{\alpha\beta}^{(0)}$ in the region. This result is not altered by the fact that radiation may be emitted from the region, since it is easy to verify that the integral over the radiation fields is of the same order (even smaller if $|\mathbf{k}|\, D<1$).

This theorem contradicts the conclusion of Gertsenshtein and Solovei (1969) that a limit on μ 1000 times smaller than the geomagnetic limit could be obtained by a proposed electrical experiment with a geometry of parallel planes. A capacitor is driven with an oscillating voltage. A second capacitor is shielded from the first by a grounded conducting sheet parallel (geometrically) to both capacitors. The voltage across the second capacitor is interpreted as due to the passage of a longitudinal electric wave through the conducting shield. They correctly deduce what in our notation would be the statement that the potential across the second capacitor is $\Phi=(\mu^2/\mathbf{k}^2)V$. However, they assume that V is the potential that would exist in the absence of a conducting shield. When this assumption is corrected to make V obey proper boundary conditions, Φ is reduced by a factor $\mathcal{O}(\mathbf{k}^2 d^2)$, where d is the distance between plates of the capacitors, and between either capacitor and the grounded shield. As a result, sensitivity to μ is reduced by a factor $\approx|\mathbf{k}|\, d\approx 10^{-5}$–$10^{-6}$, meaning that even if the postulated detection efficiency could be achieved, it would still lead to a result no better than that of Bartlett *et al.* (1970).

Our theorem, Eq. (6.19), also refutes the suggestion of Franken and Ampulski (1971) that a "table-top" apparatus (of dimension D) could be used to detect a shift of $\delta(\omega^2)=(\mu c)^2$ in the resonance frequency in an LC circuit because ω^2 should really be given by $\omega^2=1/LC+(\mu c)^2$. The problems inherent in this suggestion are analyzed by Goldhaber and Nieto, (1971), but the crucial point is that the natural resonant frequency $\omega_E=(LC)^{-1/2}$ cannot be shifted by more than a fraction of order $(\mu D)^2$, since the fields acting on charges and currents in the circuit are only changed by this amount. Therefore, in particular, a resonance at $\omega=\mu c$ will occur for $\omega_E\cong\mu c$, *not* $\omega_E=0$. The photon mass effect on the resonant frequency is practically negligible. The same conclusion has also been reached by Kroll (1971), Park and Williams (1971), Meyer (1971), and Boulware (1971).

Armed with these results we may assert confidently that the spherical geometry is superior to any other for Coulomb's law tests. The reason is alluded to by Bartlett *et al.* (1970), and Bartlett and Phillips (1969), and was doubtless familiar to earlier workers. In conventional ($\mu=0$) electrodynamics, an oscillating spherically symmetric potential cannot occur within radius $r<R$, unless there is an oscillating charge density within that region. In performing an experiment to measure such a potential, it is always hard to exclude stray inductance effects with plane parallel geometry, but only an actual charge leakage could produce a spurious effect in the spherical geometry. So, since, as we have just seen, the expected effect of $\mu\neq 0$ is of similar size for apparatus of similar dimensions, the spherical arrangement is preferable because it will have a comparable signal but a smaller synchronous background voltage than any other geometry.

What are the ultimate limits on the Cavendish method? Let us consider possible improvements on the experiment of Williams *et al.* (1971). First of all, one can reduce the Johnson noise voltage (Johnson, 1928; Nyquist, 1928) by increasing the time of observation. Williams *et al.* would be able to obtain a limit comparable to the geomagnetic value (and five times better than their present result) by running their experiment for a year—about the maximum time one could practically consider (Faller, private communication).[26] Secondly,

[26] In connection with lowering noise voltage, we should note the possibility of using low temperatures to make a superconducting apparatus. R. Y. Chiao has pointed out to us that voltage sensitivity of better than 10^{-15} V can be achieved, about a thousand times better than the sensitivity of Williams *et al.* However, cooling such a large apparatus to liquid helium temperatures is a formidable task.

one could increase the frequency to the magic value $|\mathbf{k}|\, a_{\text{shell}} \approx \pi/4$, reducing Johnson voltage in proportion. This would give a factor ≈ 15 improvement, in $\Delta V/V$, for Williams, or a factor ≈ 4 in μ. One could imagine other improvements: Increasing the applied voltage, increasing the input resistance (this reduces Johnson voltage as the square root of the increased resistance), and increasing the dimensions of the apparatus. The first and third of these cause large increases in power requirements. The third (increasing a_{shell}) would appear most promising because the limit on μ goes inversely as the diameter of the spheres. While the optimum frequency goes inversely with a_{shell}, the Johnson noise $\delta V \propto (\omega C)^{-1}$ remains constant since C increases. Hence, the attainable limit really does go inversely with the diameter. An optimistic guess would seem to be a further improvement by a factor 100 in $\Delta V/V$, leading to a final limit 40 times smaller than the present geomagnetic value.

3. *Pulsar Signals*

Feinberg (1969) has mentioned the possibility of observing pulsar signals from another galaxy. For two reasons, however, this is not very promising. The first is because the distance to the nearest galaxies (the Magellenic Clouds) is of the order of 50×10^3 kpc (about 50 times the distance to the Crab nebula pulsars) (Allen, 1963). This means that signals will be down by the factor of 2500. Thus, a successful pulsar search will probably have to depend on a statistical analysis of radio signals rather than on the direct observations that are needed for a mass limit.

Further, even if such a pulsar could be observed directly, on the average the signal would have to travel through about 1,000 kpc in each galaxy. Thus, even if intergalactic space were empty of plasma, the plasma dispersion in the galaxies would limit the improvement in the mass limit from pulsars to a factor of only $(50/2)^{1/2}=5$, so that the ultimate mass limit from pulsars is about 500 times worse than the present geomagnetic value.

It would be interesting to see if the low frequency pulse broadening and/or intensity oscillations with ω became more noticeable for such a distant pulsar.

4. *Other Magnetostatic Tests*

From the discussion of the Schrödinger method applied to the Earth's magnetic field, it seems unlikely that an improvement of more than a factor 2 in the limit on μ could be achieved, and a factor 4 would be the most optimistic estimate. The two most conspicuous alternative objects for magnetic measurement are the dipole fields of the Sun and of Jupiter. The Sun has the advantage of large size ($R_{\odot}=109R_{\oplus}$), but the disadvantage of enormous and rapidly varying plasma currents which make accurate magnetic surveys quite difficult. Also, the relevant region is very hot, adding substantially to the challenge!

Jupiter is about 11 times larger in radius than the Earth, and has a magnetic field ten to 1,000 times as strong at the planetary surface (Michaux, 1967). Further, the Jovian magnetosphere appears to be 40–50 Jovian radii from the planet, while the Earth's magnetosphere is at 8–10 Earth radii (Frank, 1967; Parker, 1967, 1970). As a consequence, one may expect a low charge and current density out to a much larger distance, in Jovian radii, than holds for the Earth, measuring in Earth radii. The higher field, the larger radius, and the enormous magnetosphere should make an improvement of a factor ten on the geomagnetic limit quite easy with orbiting satellite magnetometer data, and a factor 100 improvement possible. The latter would imply greater than 10% corrections to the Jovian field at $30R$, where it is between 10^{-4} and 10^{-2} G. The main potential difficulty would be rapid fluctuations in field during satellite orbits. The large magnetosphere would make the altitude-dependent method more useful than it is for the geomagnetic limits.

5. *Magnetic fields in the Galaxy*

Yamaguchi (1959) proposed that a photon mass limit could be obtained from magnetic fields detected in astronomical observations. His idea was that if fields extend over a distance D then the photon mass must be $\mu \lesssim D^{-1}$. Using the dimensions of the Crab nebula, he deduced $\mu \lesssim 10^{-17}$ cm^{-1}. In a recent private communication to the present authors, he pointed out that the same technique applied to the field in one of the spiral arms of our galaxy could yield a limit $\mu \lesssim 10^{-21}$ cm^{-1}.

There is an immediate naive objection to this reasoning. Even for $\mu=0$, magnetic fields extend over a great distance *only* in the presence of a conducting medium, e.g., plasma. Therefore, the relevance of free space solutions of the Proca equations, with their Yukawa falloff factor, is not at all manifest. A closer look is required.

If a magnetic field has average axial component $\langle H_z \rangle$ in a cylinder of radius R, then the average azimuthal vector potential satisfies the relation

$$\langle \mu^2 \mathbf{A}^2 \rangle \gtrsim (\mu^2 R^2/8) \langle H_z^2 \rangle, \qquad (6.20)$$

assuming H_z varies slowly. Thus, if μR is large, the $\mathbf{A}$ energy is much greater than the ordinary magnetic energy. If more could be learned about the source of galactic fields, and if it could be shown that the ordinary $\mathbf{H}$ energy saturates the capacity of the source, then it would be clear that the Yamaguchi proposal $\mu R \lesssim 1$ is correct.

However, at present the source of the fields is unknown. It may have been an event in the early development of the galaxy, or there may be a continuous pumping of energy into the fields. Further, even the characteristic survival time for the fields is not well determined. Old ideas that cosmic ray flux gives a lower bound to the field variation time are less convincing, now that pulsars have appeared as possible sources for cosmic rays (Lingenfelter, 1969). There *is* convincing evidence that part of the Galactic field is slowly varying across distances of the order of the thickness of a spiral arm. This comes from Faraday rotation measurements for radio waves from sources outside the Galaxy. The Faraday rotation shows a systematic dependence on Galactic latitude and longitude, albeit with some violent fluctuations (Berge and Seielstad, 1967; Parker, 1969, 70). However, the actual magnitude of **H**, and the relative size of any more rapidly varying field, are still subject to large uncertainty (Verschuur, 1970).

Nevertheless, it is reasonable to suppose that these great uncertainties will be partly or wholly resolved, and to speculate on the consequences for a photon mass limit. An admirable effort in this direction, and the only one of which we are aware, is a recent paper by Williams and Park (1971). Adopting plausible parameters for the plasma and **H** field in a spiral arm of the galaxy, they make the very reasonable assumption that a "straightened" spiral arm—a cylinder of similar dimensions—would behave similarly. The critical point in the calculation is that the plasma has a very high resistivity for currents traveling perpendicular to **H**. They suppose that **H** is parallel to the axis of the arm and decreases slowly away from the axis. These assumptions imply a large azimuthal vector potential $A_\varphi \approx rH/2$, where r is the distance from the axis. To keep **H** slowly varying there must be a large current perpendicular to **H**, cancelling the Proca pseudocurrent $\mathbf{J}_P = -(c/4\pi)\mu^2\mathbf{A}$. Because this current is perpendicular, it suffers high resistivity $\sigma_\perp^{-1} \approx 10^3$ sec, leading to a dissipation of the stored energy at a rate $(c^2/2\pi)\mu^2\sigma_\perp^{-1}$. Insisting on a field lifetime of $10^6 y$, Williams and Park obtain a limit $\mu \lesssim 10^{-18}$ cm^{-1}. It is amusing to note that in their cylinder model a special modification could drastically alter the result. Adding a smoothly varying azimuthal field H_φ comparable to H_z in magnitude, one can obtain a field pattern obeying the remarkable condition $\mathbf{A} \| \mathbf{H} = \nabla \times \mathbf{A}$! In this case the current required to cancel the Proca pseudocurrent travels parallel to **H**, and faces much less resistance. Using the approach of Williams and Park, one obtains a limit $\mu \lesssim 10^{-11}$ cm^{-1}. For this case, a better limit can be got in another way. There is a maximum current the plasma can support, $J_{\max} = Vne$, where V is the velocity of the charges, and ne the density of electron or ion charge in the plasma. For $H \approx 10^{-6}$ G, $V \lesssim 10^5$ cm/sec, and $n \lesssim 1/\text{cm}^3$ (Allen, 1963; Verschuur, 1970), we would get $\mu \lesssim 10^{-15}$ cm^{-1}, assuming $R \approx 10^{21}$ cm.

It is probable that a conservative treatment of data now available would produce a convincing limit of two or three orders of magnitude better than the geomagnetic value at far less cost than previously mentioned approaches. In any case, future "experiments" in Galactic magnetohydrodynamics would open the door to a hierarchy of improved μ limits, depending on the detailed character of the results, ranging from $\approx 10^{-14}$ cm^{-1} through the Williams–Park 10^{-18} cm^{-1} to the Yamaguchi limit 10^{-21} cm^{-1}. The last number represents the ultimate possibility for an improved limit among suggestions we have seen.

B. Discussion

1. Special Relativity with Finite Photon Mass

Before summarizing the status of the search for a photon mass, let us digress briefly to consider a pedagogical point. We have made critical use of special relativity in arguing that the Proca Eqs. (2.19) represent practically unique modifications of Maxwellian electrodynamics. However, special relativity itself was first presented as a consequence of the constancy of the speed of light.[27] What new postulate must we adopt in order to restore the deductions of special relativity theory, in a world with $\mu > 0$!

One answer is quite straightforward, if less appealing in its simplicity than $c =$ constant: *Postulate*: Given any two inertial frames, the first traveling at velocity **v** with respect to the second, there exists a frequency ν_0, depending on $|\mathbf{v}|$ and the desired accuracy ϵ, such that any light wave of frequency greater than ν_0 in either frame will have a speed between c and $c-\epsilon$ in both frames.

We leave it to the reader to verify this postulate by explicit construction of the function $\nu_0(v, \epsilon)$. The usual derivation of special relativity can now be applied, using the Lorentz invariant limit $\nu_0 \to \infty$.

The above postulate is equivalent, in the special case of light, to the Einstein assumption that there is an *unique* limiting velocity (c) for *all* phenomena. In a world with $\mu \neq 0$ but the neutrino mass equal to zero, then the velocity of the neutrino would be equal to c. This view makes the structure of space-time more fundamental than interactions. For example, it rules out the possibility that strongly interacting particles could have one limiting velocity while electromagnetically interacting particles had another. (Otherwise acausal effects could occur.)

2. Monopoles and $\mu = 0$

Another amusing side issue is the combination of two fascinating speculations; a finite μ and the existence of magnetic monopoles. If both these dreams come true, one could not apply the ideas of gauge invariance (Dirac, 1931, 1948) or rotational invariance (Goldhaber, 1965) to derive the famous Dirac quantization condition that $eg/\hbar c$ is a half-integer (with e and g the elementary electric and magnetic charges, respectively). In fact, the approach of Dirac would not allow any gauge-invariant theory including a particle which is a source for a $1/r^2$ magnetic field with additional "Yukawa" falloff. It is another interesting exercise for the reader to verify this difficulty. Along these lines, it is worth noting that finite μ does not disturb another quantization condition, the quantization of magnetic flux

[27] Our discussion of this matter is similar to that of de Broglie (1957), p. 61.

through a superconducting loop. This condition is derived from the fact that the line integral of the vector potential around the loop is equal to the flux through the loop.[28] That fact is not changed by finite μ. What is changed is the equality between the line integral of **H** around a loop, and the total current through the loop. Thus, "current quantization" conditions, if ever enunciated, could be modified by $\mu\neq 0$.

3. *Prospects for μ Hunters*

We have seen that improvements of more than an order of magnitude in static photon mass limits may be achieved by known methods. The theorem expressed in Eq. (6.19) demonstrates that such experiments require not only low field frequency, but also large dimensions of the "experimental region," with the sole exception of the phenomenally accurate null test for deviations from Coulomb's law. Chances for improvement in dynamic mass limits are less clear. For both types of limit, terrestrial and extraterrestrial observations are competitive with each other, although completed terrestrial measurements are about an order of magnitude worse than their extraterrestrial counterparts.

There remains a class of promising extraterrestial static "experiments" with no competition from the laboratory. As Yamaguchi (1959) and Williams and Park (1971) have noted, observations of magnetic fields over galactic dimensions could lead to a far better limit on μ, conceivably ten orders of magnitude below the geomagnetic value.

Despite the possibilities for improvement, it is quite unlikely that the dynamic measurements will become comparable in accuracy to the static ones, except in detecting energy nonconserving effects in pulsar signals. In fact, if the next improvement in static experiments demonstrates a real effect, it will probably be the only effect of nonzero photon mass observed in the foreseeable future.

Among the remote possibilites for other observable effects are: (a) detection of longitudinal waves from coherent astronomical sources, (b) production and scattering of very high energy ($\gg$100 GeV) longitudinal photons through unexpectedly enormous longitudinal current matrix elements, and (c) most "remote" of all, measurement of velocity dispersion for low frequency radiation in intergalactic space—which means taking some spaceships a long, long way from home.[29]

ACKNOWLEDGMENTS

In composing this paper, we have been inspired and instructed by a large number of people, both through published articles and private conversations. Among publications of the last few years, the works of Bartlett, Goldhagen and Phillips (1970); Feinberg (1969); Kobzarev and Okun' (1968); and Williams and Park (1971) have been most helpful.

We have received valued comments from D. F. Bartlett, R. Y. Chiao, E. R. Cohen, J. E. Faller, G. Feinberg, D. Z. Freedman, P. A. M. Gram, J. L. Heilbron, R. W. P. King, N. M. Kroll, B. Lautrup, B. W. Lee, L. B. Okun', S. J. Peale, E. A. Phillips, J. M. Rankin, K. Riegel, A. E. Sanderson, L. Van Hove, C. N. Yang and D. Yennie.

A. S. G. wishes to thank Drs. Louis Rosen and B. Zumino for hospitality at Los Alamos Scientific Laboratory and CERN, respectively, while this work reached final form.

M. M. N. wishes to thank Prof. Aage Bohr for hospitality at the Niels Bohr Institute, during the period of a challenging transAtlantic collaboration.

REFERENCES

AEpinus, F. U. T., 1759, *Tentamen Theoriae Electricitatis et Magnetismi Accedunt Dissertationes Duae, Quarum Prior, Phaenomenon Quoddam Electricum, Altera, Magneticum, Explicat* (Imperial Scientific Academy, St. Petersburg), pp.137–143.

Allen, C. W., 1963, *Astrophysical Quantitites* (Athlone Press, London), 2nd ed.

Al'pert, Y. L., V. V. Migulin, and P. A. Ryazin, 1941, Zh. Tech. Fiz. **11,** 29.

Bandyopadhyay, P., P. R. Chaudhuri, and S. K. Saha, 1970, Phys. Rev. D **1,** 377.

Bartlett, D. F., P. E. Goldhagen, and E. A. Phillips, 1970, Phys. Rev. D **2,** 483.

——, and E. A. Phillips, 1969, Bull. Am. Phys. Soc. **14,** 17.

Bass, L., 1956, Nuovo Cimento **3,** 1204.

——, 1963, Nature **198,** 980.

——, and E. Schrödinger, 1955, Proc. Roy. Soc. (London) **A232,** 1.

Berge, G. L. and G. A. Seielstad, 1967, Astrophys. J. **148,** 367.

Bjorken, J. D., and S. D. Drell, 1965, *Relativistic Quantum Fields* (McGraw-Hill, New York).

Boulware, D. G., 1971, "Lumped Circuits and the Photon Mass," Phys. Rev. Letters (to be published).

Brodsky, S., and S. D. Drell, 1971, "The Present Status of Quantum Electrodynamics." *Annual Review of Nuclear Science* **21** (to be published).

Cain, J. C., 1966, in *Radiation Trapped in the Earth's Magnetic Field*, edited by B. M. McCormac (D. Reidel Publ. Co., Dordrecht, Holland), p. 7.

——, W. E. Daniels, S. J. Hendricks, and D. C. Jensen, 1965, J. Geophys. Res. **70,** 3647.

——, and S. J. Hendricks, 1968, N. A. S. A. Tech. Note No. D-4527 (unpublished).

[28] For references, see the article of Taylor *et al.* (1969). The statements in our text are based on the assumption that the electromagnetic interaction is of the usual "minimal" form obtained by the substitution $\mathbf{p}\rightarrow\mathbf{p}-(e/c)\mathbf{A}$. It might seem impossible to "derive" the minimal interaction from gauge invariance, if μ is nonzero. However, B. Lautrup has pointed out to us that there exists a formal gauge invariance which has no significance in classical electrodynamics, but which can be important in quantum mechanics. Namely, one may replace the electromagnetic current in Proca's Equation by $J_\alpha'=J_\alpha+\partial_\alpha\Lambda$ and have the new condition $\Box\Lambda=\mu^2\partial^\alpha A_\alpha$. Only the combination $(\mu^2A_\alpha-\partial_\alpha\Lambda)$ enters in any classical measurable quantity, but this trick permits one to introduce a formal gauge invariance that is helpful in a Lagrangian formulation of quantum electrodynamics. It is equivalent to an explicit separation of spin-one longitudinal photons from spin-zero photons. Gauge invariance, or current conservation, is equivalent to the statement that this separation has no physical significance, even for finite photon mass.

[29] Ending on this wistful note, as we must in discussing a still unfinished search, we find it cheering to recall an amusing anecdote by R. P. Feynman (1962–1963) in his Lecture Notes on Gravitation. Feynman tells the story of a cocktail party in Paris, where, challenged by a French professor, he managed in a short time to derive at least crude mass limits in almost all the ways indicated here. One can only regret the discussion was not longer, for a new method might have been unearthed.

——, S. J. Hendricks, R. A. Langel, and W. V. Hudson, 1967, J. Geomag. and Geoelect. **19,** 335, Appendix.
Cap, F., 1953, J. de Physique et le Radium **14,** 213.
Cavendish, H., 1879, *The Electrical Researches of the Honourable Henry Cavendish, F.R.S.*, edited by J. C. Maxwell (Cambridge U.P., Cambridge, reprinted by Frank Cass & Co., London, 1967), pp. 104–113.
Chapman, S. and J. Bartels, 1940, *Geomagnetism* (Clarendon, Oxford, England), Vol. 2, p. 710.
Cochran, G. D., 1967, thesis, University of Michigan.
——, and P. A. Franken, 1968, Bull. Am. Phys. Soc. **13,** 1379.
Coulomb, C. A., 1788, Acad. Roy. Sci., Paris, Histoire 1785, p. 569, 578. Translated in part into English in W. F. Migie, 1965, *A Source Book in Physics* (Harvard U. P., Cambridge, Mass.), pp. 408–420.
de Broglie, L., 1934, Compt. Rend **199,** 445.
——, 1940, *La Méchanique Ondulatoire du Photon, Une Nouvelle Théorie de la Lumière*, Vol. 1, pp. 39–40, Hermann, Paris.
——, 1954, *Théorie Générale des Particules à Spin*, (Gauthier-Villars, Paris) 2nd. ed. p. 190. (The first edition was published in 1943.)
——, 1957, *Méchanique Ondulatoire du Photon et Théorie Quantique des Champs*, (Gauthier-Villars, Paris) 2nd ed., p. 62. (First edition was published in 1949).
——, 1969–1970. The following are Library of Congress listings under de Broglie's name: *The National Union Catalog.* Pre-1956 Imprints (Mansell Information Pub. Ltd., London, 1970), Vol. 77, pp. 434–9. *Library of Congress and National Union Catalog Author Lists*, 1942–62 (Gale Research Company, Detroit 1969), Vol. 21, pp. 366–8. *The National Union Catalog.* A Cumulative Author List 1963–1967, (J. W. Edwards, Ann Arbor, Michigan 1969), Vol. 8, pp. 118–119.
Dirac, P. A. M., 1931, Proc. Roy. Soc. (London) **A133,** 60.
——, 1948, Phys. Rev. **74,** 817.
Dolgov, A. D., and V. I. Zakharov, 1971, Zh. Eksperm. Teor. Fiz. (to be published).
Duffin, R. J., 1938, Phys. Rev. **54,** 1114.
Feinberg, G., 1969, Science **166,** 879.
Feynman, R. P., 1962–1963, *Lecture Notes on Gravitation* (California Institute of Technology, unpublished), Lecture 2, p. 24.
Finch, H. F., and B. R. Leaton, 1957, Geophys. J. Suppl. **7,** 314.
Florman, E. F., 1955, J. Res. Natl. Bur. Std. **54,** 335.
Fougère, P. F., 1965, J. Geophys. Res. **70,** 2171.
Frank, L. A., 1967, J. Geophys. Res. **72,** 3753.
Franken, P. A., and G. W. Ampulski, 1971, Phys. Rev. Letters **26,** 115.
Franklin, B., 1774, *Benjamin Franklin's Experiments (1774)*, a new edition of Franklin's *Experiments and Observations on Electricity*, edited by I. Bernard Cohen, (Harvard U. P., Cambridge, Mass., 1941), p. 72, 336.
Froome, K. D., and L. Essen, 1969, *The Velocity of Light and Radio Waves* (Academic, New York).
Gertsenshtein, M. E., and L. G. Solovei, 1969, Zh.ETF Pis. Red. **9,** 137 [JETP Lett. **9,** 79 (1969)].
Gintsburg, M. A., 1963, Astron. Zh. **40,** 703 [Sov. Astron. AJ **7,** 536 (1964)].
Goldberger, M. L., and K. M. Watson, 1964, *Collision Theory*, (Wiley, New York) p. 63.
Goldhaber, A. S., 1965, Phys. Rev. **140,** B1407.
——, and M. M. Nieto, 1968, Phys. Rev. Letters **21,** 567.
——, and M. M. Nieto, 1971, "How to Catch a Photon and Measure its Mass" Phys. Rev. Letters **26** (to be published).
Hendricks, S. J., and J. C. Cain, 1966, J. Geophys. Res. **71,** 346.
Jackson, J. D., 1962, *Classical Electrodynamics* (Wiley, New York), Chap. 11.
Johnson, J. B., 1928, Phys. Rev. **32,** 97.
Kemmer, N., 1938, Proc. Roy. Soc. **A173,** 91.
Kiepenheuer, K. O., 1951, in *The Solar System*, edited by G. P. Kuiper (University of Chicago, Chicago), Vol. I, *The Sun*, p. 330ff.
Kobzarev, I. Yu., and L. B. Okun', 1968, Usp. Fiz. Nauk **95,** 131 [Sov. Phys. Uspe. **11,** 338 (1968)].
Kroll, N. M., 1971, "Theoretical Interpretation of a Recent Experimental Investigation of the Photon Rest Mass" Phys. Rev. Letters **26** (to be published).
Lee, T. D., and G. C. Wick, 1969, Nucl. Phys. **B9,** 209.
Lingenfelter, R. E., 1969, Nature, **224,** 1182.
Mandel'shtam, L. I., and N. D. Papalexi, 1944, Usp. Fiz. Nauk **26,** 144. [reprinted in *The Complete Collected Works of L. I. Mandel'shtam*, Soviet Academy of Sciences, Moscow, 1949)] Vol. 3, p. 238.
Marochnik, L. S., 1968, Astron. Zh. **45,** 213 [Sov. Astron.-AJ **12,** 171 (1968)].
Maxwell, J. C., 1873, *A Treatise on Electricity and Magnetism* (Oxford U. P.), Vol. 1. 3rd ed. of 1891 reprinted by (Dover, New York, 1954). See pp. 80–86.
Meyer, P., 1971, Phys. Rev. Letters (submitted).
Michaux, C. M., 1967, *Handbook of the Physical Properties of the Planet Jupiter*, NASA SP-3031. (NASA, Washington, D.C.) Chap. 8, p. 47.
Mottelay, P. F., 1922, *Bibliographical History of Electricity and Magnetism* (Charles Griffin & Co., London).
Nyquist, H., 1928, Phys. Rev. **32,** 110.
Park, D., and E. R. Williams, 1971, "Comments on a Proposal for Determining the Photon Rest Mass," Phys. Rev. Letters **26** (to be published).
Parker, E. N., 1967, in *Physics of Geomagnetic Phenomena II*, edited by S. Matsushita and W. H. Campbell (Academic, New York).
Parker, E. N., 1969, Astrophys. J. **157,** 1129.
——, 1970, Astrophys. J. **160,** 383.
Patel, V. L., 1965, Phys. Letters **14,** 105.
——, and L. J. Cahill, Jr., 1964, Phys. Rev. Letters **12,** 213.
Petiau, G., 1936, "Contribution à la Théorie des Équations d'Ondes Corpusculaires," thesis University of Paris. This was also published in Acad. Roy. de Belg., Classe Sci., Mém. in 8° **16,** No. 2.
——, 1949, J. Phys. Radium Ser. VIII, **10,** 215.
Plimpton, S. J., and W. E. Lawton, 1936, Phys. Rev. **50,** 1066.
Priestley, J., 1767, *The History and Present State of Electricity with Original Experiments*, (London, 1767) Vol. 2, Part. VIII, Sec. XVI, Sub. XV. The third edition has been reprinted by Johnson Reprint Corp., New York (1966). See. Vol. II, p. 374.
Proca, A., 1930a, Compt. Rend. **190,** 1377.
——, 1930b, *ibid.* **191,** 26.
——, 1930c, J. Phys. Radium Ser. VII, **1,** 235.
——, 1931, Compt. Rend. **193,** 832.
——, 1936a, Compt. Rend. **202,** 1366.
——, 1936b, Compt. Rend. **202,** 1490.
——, 1936c, Compt. Rend. **203,** 709.
——, 1936d, J. Phys. Radium Ser. VII, **7,** 347.
——, 1936e, J. Phys. Radium Ser. VII, **8,** 23.
Rankin, J. M., J. M. Comella, H. D. Craft, Jr., D. W. Richards, D. B. Campbell, and C. C. Counselman III, 1970, Astrophs. J., **162,** 707. Also, private communication from J. M. Rankin.
Robison, J., 1803, "Electricity," article in *Encyclopaedia Britannica* (Thompson Boner, Parliament Square, Edinburgh), 3rd ed., Vol. 19, Supp. Vol. I, p. 558. See pp. 578–580.
Rosenfeld, A. H., *et al.*, 1968, Rev. Mod. Phys. **40,** 77.
Sakurai, J. J. 1969, *Currents and Mesons* (University of Chicago Press, Chicago), Chap. 3.
——, 1969b in *Proc. Intern. Symp. Electron Photon Interactions High Energy.*
Schmidt, Ad., 1924, Z. f. Geophysik **1,** 3.
——, (with an Appendix by J. Bartels), 1939, Gerland's Beitr. Geophysik **55,** 292.
Schrödinger, E., 1943a, Proc. Roy. Irish Acad. **A49,** 43.
——, 1943b, Proc. Roy. Irish Acad. **A49,** 135.
Smith-Rose, R. L., 1942a, Nature **160,** 477.
——, 1942b, Electrician **129,** 415.
Streater, R. F., and A. S. Wightman, 1964, *PCT, Spin, Statistics, and All That* (Benjamin, New York), Chap. 2.
Stueckelberg, E. C. G., 1957, Helv. Phys. Acta **30,** 209.
Taylor, B. N., W. H. Parker, and D. N. Langenberg, 1969, Rev. Mod. Phys. **41,** 375.
Taylor, R. E., 1969, in *Proc. 4th Intern. Symp. Electron Photon Interactions at High Energy.*
Thompson, W. B., 1962, *An Introduction to Plasma Physics* (Pergamon, Oxford), Chap. 7.
Verschuur, G. L., 1970, Astrophys. J., **161,** 867.
Waldmeier, M., 1941, Astr. Mitt. Zürich **14,** No. 140, 551.
Wigner, E. P., 1956, Nuovo Cimento Suppl. **4,** 826.
Williams, E. R., 1970, Ph. D. thesis, Wesleyan University.
——, J. E. Faller, and H. Hill, 1970, Bull. Am. Phys. Soc. **15,** 586.
——, J. E. Faller, and H. Hill, 1971, Phys. Rev. Letters **26,** 721.
——, and D. Park, 1971, *Photon Mass and Galactic Magnetic Field* (unpublished).
Yamaguchi, Y., 1959, Prog. Theor. Physics (Japan), Supp. **11,** 33.