THE PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 66, NOS. 7 AND 8 — OCTOBER 1 AND 15, 1944

Theory of Diffraction by Small Holes

H. A. BETHE

Department of Physics, Cornell University, Ithaca, New York

(Received January 26, 1942)

The diffraction of electromagnetic radiation by a hole small compared with the wave-length is treated theoretically. A complete solution is found satisfying Maxwell's equations and the boundary conditions everywhere (Section 4). The solution holds for a circular hole in a perfectly conducting plane screen, but it is believed that the method will be applicable to much more general problems (Section 8). The method is based on the use of fictitious magnetic charges and currents in the diffracting hole which has the advantage of automatically satisfying the boundary conditions on the conducting screen. The charges and currents are adjusted so as to give the correct tangential magnetic, and normal electric, field in the hole. The result (Section 5) is completely different from that of Kirchhoff's method, giving for the diffracted electric and magnetic field values which are smaller in the ratio (radius of the hole/wave-length) (Section 6). The diffracted field can be considered as caused by a magnetic moment in the plane of the hole, and an electric moment perpendicular to it (Section 6). The theory is applied to the problem of mutual excitation of cavities coupled by small holes (Section 9). This leads to equations very similar to those for ordinary coupled circuits. The phase and amplitude relations of two coupled cavities are not uniquely determined, but there are two modes of oscillation, of slightly different frequency, for which these relations are opposite (Section 10). The problem of stepping up the excitation from one cavity to another is treated (Section 11).

1. THE PROBLEM

In microwave work it is often important to know the effect of a small hole in a cavity upon the oscillation of that cavity. For instance, two cavities may be coupled by a small hole in their common boundary (Fig. 1); in this case, we wish to know the characteristic frequencies and the phase relations for the oscillations of the coupled system. Or a hole in a cavity may serve the purpose of getting radiation out of it; then we want to calculate the amount and the spatial distribution of the emitted radiation. Another similar problem would be to calculate the effect of a small gap in a wave guide upon the propagation of waves along that guide.

A less practical problem but probably the simplest one of the same type, is the *diffraction of electromagnetic waves by a small hole in an infinite plane conducting screen.* This is the problem which we are going to solve first

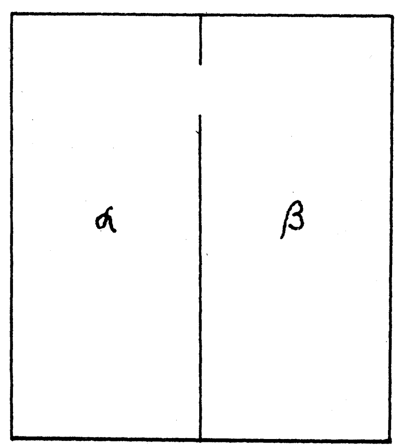

FIG. 1. Two cavities, α and β, coupled by a small hole.

(Sections 3–7); the application of the solution to the practical problems mentioned is then rather straightforward (Sections 9–11).

The available theoretical methods are entirely inadequate for the treatment of our problem. In the usual Kirchhoff method, the diffracted field is expressed in terms of the incident field in the hole. However, the Kirchhoff solution does not satisfy the boundary conditions, *viz.*, it does not give zero tangential component of the electric field on the screen. In most textbooks, the pious hope is expressed that Kirchhoff's method will give at least the first term of a convergent series. This is probably true for the diffraction by an opening, large compared with the wave-length, because then the diffracted field will be relatively small on the screen, thus "almost" fulfilling the boundary conditions. But it is certainly not true for a small hole; in fact, our exact solution of the problem will turn out to be entirely different from Kirchhoff's. The failure of Kirchhoff's theory will be demonstrated mathematically in Section 2.

Kirchhoff's method has the additional defect of being a scalar theory while the electromagnetic field is essentially vectorial. This shortcoming has been remedied in the last forty years by a number of writers; a very good account of the vector equivalent of Kirchhoff's theory is given in Stratton's book,[1] to which we shall frequently refer in this paper. The vectorial theory ensures the fulfillment of the divergence conditions, $\operatorname{div} E = \operatorname{div} H = 0$; i.e., it gives transverse waves in the wave zone which would not necessarily be the case in Kirchhoff's scalar formulation. However, the vector formulation in no way improves the situation regarding the fulfillment of the boundary conditions on the conducting screen.

The only rigorous solution of any diffraction problem known to me is Sommerfeld's solution[2] of the diffraction by a conducting semi-infinite plane, or by a wedge. As is well known, this solution is rather complicated although the problem is the simplest imaginable, being two- rather than three-dimensional. It appears hopeless to look for a rigorous solution of our problem along the lines of Sommerfeld's solution. A slight similarity between his and our method will be pointed out later (Section 6d).

The main simplifying assumption we are going to use is that the hole is small compared with the wave-length. This means that the incident electromagnetic field is almost constant over the hole. We believe, however, that our method can be generalized to holes of larger size (Section 8).

2. THE FAILURE OF KIRCHHOFF'S THEORY

The failure is most easily seen in the simplest formulation of the theory, *viz.*, Kirchhoff's own, scalar formulation. Let u be a scalar wave function satisfying the wave equation

$$\nabla^2 u + k^2 u = 0. \tag{1}$$

We may identify u, e.g., with one of the components of the electric field, say E_y. The derivatives of u will then be connected with the magnetic field. If the conducting screen is placed at $x=0$, the boundary condition is

$$u = 0 \quad \text{at} \quad x = 0. \tag{2}$$

Let electromagnetic waves come in from the left ($x<0$), the corresponding wave function may be u_0. Then according to Green's theorem, the wave function at any point r on the right of the screen is

$$u(r) = \int d\sigma \left[-\frac{\partial u_0}{\partial x'}(r')\varphi(|r-r'|) + u_0(r')\frac{\partial \varphi}{\partial x'} \right] \tag{3}$$

with the Green's function

$$\varphi(r) = e^{ikr}/r. \tag{3a}$$

The integration variables are y', z'. The coordinates r of the "field point" and x' ($=0$) are kept constant. The solution (3) is rigorous if the integral is extended over the entire surface $x'=0$, with the correct values for $u_0(0, y', z')$. However, the distribution of u_0 on the plane $x'=0$ is not known. Kirchhoff's method consists now in putting $u_0 = \partial u_0/\partial x' = 0$ everywhere on the screen outside the hole, and replacing u_0 by the field of the incident wave *in* the hole.

Suppose now the hole is very small compared with the wave-length, then u_0 may be assumed to

[1] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), p. 464 ff.

[2] A. Sommerfeld, Riemann-Weber's *Differentialgleichungen der Physik*, seventh edition, p. 433.

be constant over the hole and we obtain, remembering that φ is a symmetrical function of r and r',

$$u(r) = -A\left[\frac{\partial u_0}{\partial x'}\varphi(r) + u_0\frac{\partial \varphi(r)}{\partial x}\right], \tag{4}$$

where A is the area of the hole. We can now make two alternative assumptions, *viz.:*

(*a*) u_0 in the hole has the value given by the unperturbed incident wave. Then both u_0 and $\partial u_0/\partial x'$ are different from zero. If we have an incident plane wave traveling at an angle ϑ with the x direction, we have

$$\partial u_0/\partial x' = iku_0 \cos\vartheta. \tag{4a}$$

(*b*) We may take into account the reflection of the incident wave by the conducting screen and put u_0 equal to incident plus reflected wave. Then we have in the hole $u_0 = 0$ while the value of $\partial u_0/\partial x'$ is doubled.

Now consider the expression (4) on the screen ($x=0$). Then it follows immediately from (3a) that $\partial\varphi/\partial x = 0$ so that the second term in (4) satisfied the boundary condition (2). The first term in (4), however, does not vanish by any means. Unfortunately, it is just this first term which remains present when assumption (*b*) is made. Therefore, for either of the assumptions (*a*) and (*b*), the boundary condition (2) is violated on the "right-hand" (back) side of the screen. This would only be different if we could represent the field by the second term in (4) alone.

The vectorial theory has the same defects. If there are no currents or charges in the half-space $x>0$, the electric field at a point r in that half-space is given by Eq. (19), p. 466, in Stratton's book, which is

$$E(r) = \frac{1}{4\pi}\int d\sigma[ikn\times H(r')\varphi - (n\times E(r'))\times\text{grad}\,\varphi - n\cdot E(r')\,\text{grad}\,\varphi]. \tag{5}$$

In this n is a unit vector in the direction of the inward normal to the surface, i.e., in our case the positive x direction. The first integral in Stratton's formula is omitted because there are assumed to be no charges or currents. The notation is slightly changed (1) because we use Gaussian units so that Stratton's $\omega\mu$ is replaced by k, (2) because we use the gradient with respect to the coordinates of the field point r (point at which the field is to be calculated) rather than of the source point r' (point on the surface where the field is given), (3) because we have interchanged the notations r and r', and (4) because we have reversed the sign of n.

If the integral is extended over the entire surface $x'=0$, Eq. (5) is of course correct. However, according to Kirchhoff's method, the integral is only extended over the hole, and for E and H we insert the field of the wave incident from the left, i.e., E_0, H_0. These quantities may be again considered as constant over the hole which permits integration as in Eq. (4).

Now let us again consider the electric field at a point r on the screen. Then grad φ will be a vector in the plane of the screen, and the same is true of $n\times H$ and of $n\times E$. Therefore the first and last term in (5) give tangential components of E, thus violating the boundary condition $E_{\tan}=0$. Only the second term satisfies the boundary condition, giving an electric field normal to the surface. Again, this second term is the only one which drops out if we include the reflected wave in calculating $E(r')$ and $H(r')$ in the hole.

3. MATHEMATICAL FORMULATION

Let H_0, E_0 be the field on the left-hand side of the screen *if there is no hole*. This field fulfills the boundary condition for $x=0$ (plane of the screen)

$$E_{0\,\tan} = 0 \tag{6}$$

which may also be written in the form

$$n\times E_0 = 0. \tag{6a}$$

This makes automatically

$$H_0\cdot n = 0. \tag{6b}$$

$H_{0\,\tan}$, and E_{0n}, are in general different from zero. On the right-hand side of the screen, the field in zero approximation vanishes identically. Then the zero approximation field satisfies the boundary conditions everywhere on the screen but not in the hole: In the hole, H and the normal component of E are discontinuous.

We write now the actual field

$$H=H_0+H_1 \quad \text{for } x<0, \qquad H=H_2 \quad \text{for } x>0, \tag{7}$$

and similarly for the electric field. Then we have the boundary conditions:

$$E_{1\,\tan}=E_{2\,\tan} \quad \text{in the hole,} \tag{8}$$

$$E_{1\,\tan}=E_{2\,\tan}=0 \quad \text{for } x=0 \text{ outside the hole,} \tag{8a}$$

$$H_{2\,\tan}-H_{1\,\tan}=H_{0\,\tan} \quad \text{in the hole.} \tag{8b}$$

The boundary conditions for the normal components are automatically fulfilled if those for the tangential components are satisfied.

It is easily seen that all boundary conditions for E_1 and H_1 are satisfied if E_2 satisfies the boundary condition (8a) and if we put, for any $x>0$, and any y and z:

$$E_{1y}(-x,y,z)=E_{2y}(x,y,z), \tag{9}$$

$$H_{1y}(-x,y,z)=-H_{2y}(x,y,z), \tag{9a}$$

and correspondingly for the z components, Maxwell's equations are consistent with (9), (9a) and will further make

$$H_{1x}(-x,y,z)=H_{2x}(x,y,z), \tag{9b}$$

$$E_{1x}(-x,y,z)=-E_{2x}(x,y,z). \tag{9c}$$

Inserting (9a) into (8b) we find that H_2 must satisfy the boundary condition

$$H_{2\,\tan}=\tfrac{1}{2}H_{0\,\tan} \quad \text{in the hole.} \tag{10}$$

Likewise, from (9c) we get the similar condition

$$E_{2x}=\tfrac{1}{2}E_{0,x} \quad \text{in the hole.} \tag{10a}$$

H_0 and $E_{0,x}$ may be considered as known.

The problem is then to calculate the field E_2, H_2 subject to the boundary conditions (8a), (10), and (10a). These conditions are valid irrespective of size and shape of the hole. However, in the following, we shall assume the hole to be small compared with the wave-length so that H_{0y}, H_{0z}, and E_{0x} may be considered as constant over the hole, and we shall take the shape of the hole to be circular, of radius a.

4. SOLUTION

As we have seen in Section 2, only one of the terms in (5) leads to an acceptable solution, *viz.*, the second. Such a term would be produced by a distribution of "magnetic currents" over the hole (cf. Stratton[1]). We shall therefore *assume a distribution of magnetic currents in the plane of the hole*, but instead of assuming the *current density* to be proportional to $n\times E$, we shall determine it so as to satisfy the boundary conditions (10), (10a).

The magnetic current density J^* and charge density ρ^* can be introduced into Maxwell's equations in the same way as the electric charge and current, *viz.*,

$$\operatorname{div} H=4\pi\rho^*, \tag{11}$$

$$\operatorname{curl} E+\frac{1}{c}\frac{\partial H}{\partial t}=-4\pi J^*. \tag{11a}$$

These equations are identical with Stratton's Eqs. I and III, reference 1, p. 464, except for the units used: We are using Gaussian, non-rational units, and we measure ρ^* in "magnetostatic" units, J^* in "magnetoelectric" units. The continuity equation corresponding to (11), (11a) is

$$\operatorname{div} J^*+\frac{1}{c}\frac{\partial\rho^*}{\partial t}=0. \tag{12}$$

We shall assume the time dependence of all quantities to be as

$$e^{-i\omega t} \tag{12a}$$

so that (3a) represents an outgoing spherical wave. Then (12) becomes

$$\operatorname{div} J^*=ik\rho^* \tag{12b}$$

with

$$k=\omega/c. \tag{12c}$$

It need hardly be pointed out that J^* and ρ^* have no physical meaning.

We shall not use magnetic volume currents and charges but only magnetic surface currents (density K) and surface charges (density η); η corresponds to a discontinuity of H_n at the surface [cf. (11)] and K to a discontinuity of $E_{\tan}$ [cf. (11a)]. The quantities K and η satisfy a continuity equation exactly like (12b), *viz.*,

$$\operatorname{div} K=ik\eta. \tag{13}$$

The electric and magnetic field can be expressed in terms of K and η most conveniently with the help of a scalar and vector potential, *viz.*,

$$E = \text{curl}\, F, \tag{14}$$

$$H = -\frac{1}{c}\frac{\partial F}{\partial t} - \text{grad}\, \psi. \tag{14a}$$

Equations (14) and (14a) automatically satisfy the Maxwell equations

$$\text{curl}\, H = -\frac{1}{c}\frac{\partial E}{\partial t}, \tag{15}$$

$$\text{div}\, E = 0. \tag{15a}$$

From Eqs. (11), (11a) we find ψ and F in exact analogy to the electric case. If only surface charges are present, we obtain

$$F(r) = -\int K(r')\varphi(|r-r'|)dr', \tag{16}$$

$$\psi(r) = \int \eta(r')\varphi(|r-r'|)dr'. \tag{16a}$$

Inserting these results, and the time dependence (12a), into (14a) we find the magnetic and electric fields explicitly [cf. Stratton, reference 1, p. 466, Eqs. (19) and (23)]

$$E(r) = \int K(r')\times \text{grad}\, \varphi\, d\sigma \tag{17}$$

and [Stratton, reference 1, Eqs. (20) and (23)]

$$H(r) = \int (ikK(r')\varphi - \eta(r')\, \text{grad}\, \varphi)d\sigma. \tag{18}$$

The gradients are taken with respect to the coordinates of the field point r (cf. Section 2 for differences between our and Stratton's notation); the integral goes over the area of the hole.

It is seen immediately that (17) satisfied the boundary condition (8a), *viz.*, $E_{\tan}$ vanishes on the screen everywhere outside the hole. In the hole, $E_{\tan}$ is of course not zero but directly related to K [cf. Section 7, Eq. (61)]. However, the problem remains to determine K in such a way as to satisfy the boundary conditions for the magnetic field in the hole, *viz.*, (10), (10a), and this is, of course, a more difficult task.

a. Determination of η from $H_{\tan}$

In the solution of our problem, we are greatly helped by the fact that the hole is small. Then the retardation may be neglected, and H may be considered essentially as the magnetostatic field corresponding to the charge distribution η. The first term in (14a) or (18) is small; indeed, from (13) we find that K is of order ηka where a is the radius of the hole; moreover, grad φ is of the order φ/a so that the first term of (18) is of the order of $(ka)^2$ times the second. Since the hole is assumed to be small compared with the wave-length ($ka \ll 1$), we shall neglect the first term of (18) altogether. Then (14a) reduces to

$$H = -\text{grad}\, \psi, \tag{19}$$

and since we also neglect the retardation in φ, (16a) reduces to

$$\psi(r) = \int \eta(r')\frac{dr'}{|r-r'|}. \tag{20}$$

In our approximation we may assume $H_{\tan}$ to be constant over the hole and equal to $\frac{1}{2}H_0$ [cf. (10)]. Then we have from (19)

$$\psi = -\tfrac{1}{2}H_0\cdot r. \tag{21}$$

The potential problem given by (20), (21) is fairly well known from electrostatics;[3] we seek a two-dimensional charge distribution which gives a constant field, $\frac{1}{2}H_0$, inside the region occupied by the charges. It is known that a constant inside field *is* produced by a uniform distribution of dipoles in an ellipsoid, the dipoles having (in simple cases) the same direction as the field. If we now assume the x axis of the ellipsoid $2h$ to be very small, the ellipsoidal charge distribution will be equivalent to a surface charge distribution. The cross section of the ellipsoid in the YZ plane should, of course, be taken equal to the hole so that we obtain a rotational ellipsoid with semi-axes, a, a, h. The surface density of dipoles is proportional to the ordinate of the ellipsoid, i.e., proportional to

$$\mu = (a^2 - r'^2)^{\frac{1}{2}}. \tag{21a}$$

[3] The following argument was taken from H. Hertz's solution of the elastic problem of the contact between two spheres. See H. Hertz, Crelle's J. 92 (1881) or A. E. H. Love, *Theory of Elasticity*.

The surface *charge* density is then proportional to

$$H_0\cdot\text{grad}\,\mu = -\frac{H_0\cdot r'}{(a^2-r'^2)^{\frac{1}{2}}}. \qquad (22)$$

It would also be possible to determine the coefficient of (22) from electrostatics.

We shall, however, determine this coefficient by direct calculation, thus at the same time verifying the solution (22). We put therefore

$$\eta = -C\frac{H_0\cdot r'}{(a^2-r'^2)^{\frac{1}{2}}}. \qquad (23)$$

In the integral (20), we introduce the coordinates $\rho=|r-r'|$ and β, the angle between the vectors $r'-r$ and r (cf. Fig. 2). Then if α is the angle between r and H_0, we have

$$H_0\cdot r' = H_0\cdot r + H_0\rho\cos(\alpha-\beta). \qquad (23a)$$

The integral Eq. (20) becomes, when we insert (21) and (23):

$$C\int_0^{2\pi} d\beta\int_0^{\rho(\beta)}\frac{d\rho}{(a^2-r'^2)^{\frac{1}{2}}} \times[H_0\cdot r + H_0\rho\cos(\alpha-\beta)] = \tfrac{1}{2}H_0\cdot r. \qquad (24)$$

The factor ρ from the element of area cancels the factor $\varphi=1/|r-r'|$ in (20). It is more convenient to integrate in (24) along a whole chord such as RS in Fig. 2, instead of separately over the parts RP and PS. Then the integral over β goes only from 0 to π, and assumes positive (on PS) as well as negative (on RP) values.

Now we find from Fig. 2,

$$a^2-r'^2 = a^2 - OQ^2 - QT^2 = s^2 - x^2, \qquad (25)$$

where $s=QS$ is half the length of the chord and

$$x = QT = \rho + QP = \rho + r\cos\beta. \qquad (25a)$$

x goes from $-s$ to $+s$, and we have

$$\int_{-s}^{s}\frac{dx}{(a^2-r'^2)^{\frac{1}{2}}} = \pi, \qquad (26)$$

$$\int_{-s}^{s}\frac{\rho dx}{(a^2-r'^2)^{\frac{1}{2}}} = -\pi r\cos\beta. \qquad (26a)$$

Then (24) becomes

$$C\pi\int_0^{\pi} d\beta(H_0\cdot r - H_0 r\cos\beta)\cos(\alpha-\beta) = \tfrac{1}{2}H_0\cdot r. \qquad (27)$$

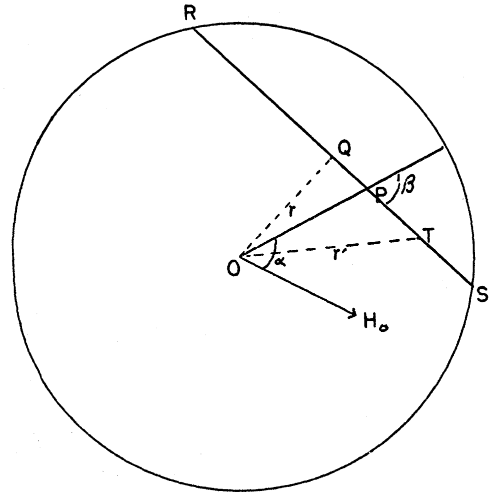

FIG. 2. Illustrating the integration, Eq. (24) to (26). O is the center of the hole, $r=0$, P is the field point r, RS any chord through P, T any point r' on this chord.

Integration over β gives

$$C\pi\cdot\pi(H_0\cdot r - \tfrac{1}{2}H_0 r\cos\alpha) = \tfrac{1}{2}C\pi^2 H_0\cdot r \qquad (27a)$$

so that

$$C = 1/\pi^2. \qquad (27b)$$

The assumption (23) therefore actually solves the integral equation, and the magnetic charge density is

$$\eta = -\frac{1}{\pi^2(a^2-r'^2)^{\frac{1}{2}}}H_0\cdot r'. \qquad (28)$$

From the charge density we can obtain the current density K, using (13). We find

$$K = \frac{ik}{\pi^2}(a^2-r'^2)^{\frac{1}{2}}H_0. \qquad (29)$$

Remembering that H_0 has vanishing normal component everywhere on the plane $x=0$, we see that K is entirely in the plane of the hole, as it should be.

b. Boundary Condition for E_n

The correct normal component of E_0, *viz.*, $E_{2n}=\frac{1}{2}E_{0n}$, must also be obtained by a suitable distribution of magnetic currents K, from (17). It can easily be seen that the distribution (29) does not contribute appreciably to (17). Expression (29) is of the order kaH_0, grad φ is of order $1/a^2$, and the integral (17) goes over an area of order a^2. Therefore the contribution of (29) to (17) is at most of order kaH_0 while actually the normal component of E_0 is of the same order as H_0. [In reality, the spatial dependence introduced by the grad operator in (17) makes the contribution of (29) of the order $(ka)^2H_0$ only.]

Therefore we must obtain an additional magnetic current distribution K_E to fit the boundary condition for E_n. However, we must take care that the new current distribution does not give rise to an additional magnetic charge density η which might destroy the agreement obtained above for the magnetic field. This is most easily achieved by letting the magnetic current lines be closed circles. Mathematically we must have

$$\operatorname{div} K_E = 0.$$

For the actual evaluation it is most convenient to use (14), (16). The given electric field E is normal to the surface and had the value $\frac{1}{2}E_{0,x}$ which may be considered constant over the hole. Therefore (14) gives

$$F = \tfrac{1}{4}E_0 \times r. \tag{30}$$

The integral equation for the components of K is the same as for η in Section 2, e.g.,

$$\int K_y \frac{dy'dz'}{|r-r'|} = -F_y = \tfrac{1}{4}E_{0,x}z, \tag{31}$$

which corresponds to (20), (21). Therefore we have, in analogy to (28),

$$K_y = \frac{1}{2\pi^2(a^2-r'^2)^{\frac{1}{2}}}E_{0,x}z' \tag{31a}$$

and altogether for the current giving the required value of E.

$$K_E = \frac{1}{2\pi^2(a^2-r'^2)^{\frac{1}{2}}}r' \times E_0. \tag{32}$$

It is easily seen that this solution satisfies the condition (30a). Moreover, since E_0 is normal to the plane, the vector K_E is entirely in the plane of the hole as it should be. The contribution of K_E to the first term in (18) can be shown to be of order $(ka)^2E_{0n}$ which is negligible.

c. Final Formula

The total magnetic current and charge density in the hole is now

$$K = K_H + K_E = \frac{1}{\pi^2}\left(ik(a^2-r'^2)^{\frac{1}{2}}H_0 + \frac{1}{2(a^2-r'^2)^{\frac{1}{2}}}r' \times E_0 \right), \tag{33}$$

$$\eta = -\frac{1}{\pi^2(a^2-r'^2)^{\frac{1}{2}}}r' \cdot H_0. \tag{33a}$$

5. CALCULATION OF THE DIFFRACTED FIELD

The diffracted field E, H for $x>0$ can be calculated by inserting (33), (33a) into (17) and (18). According to its construction, this field will satisfy Maxwell's equations and all boundary conditions.

We shall carry out the integration at large distance from the hole, $kr \gg 1$. Let κ be a unit vector in the direction of r, i.e., in the direction of propagation of the diffracted wave. Then we have

$$\operatorname{grad} \varphi = ik\kappa\varphi. \tag{34}$$

In some of the integrals we may replace φ by its value for $r'=0$, *viz.*,

$$\varphi_0 = e^{ikr}/r. \tag{35}$$

In other cases, it is necessary to expand φ in powers of r' and keep the linear term; then we obtain

$$\varphi = \varphi_0(1 - ik\kappa \cdot r'). \tag{35a}$$

a. Evaluation of the Electric Field

According to (17) and (34), the electric field is

$$E(r) = -ik\kappa \times \int K\varphi dy'dz'. \tag{36}$$

It is convenient to calculate separately the contributions of K_H and K_E [cf. (33)]. We have

$$\int K_H \varphi dy'dz' = \frac{ik}{\pi^2}H_0\varphi_0 \int_0^a (a^2-r'^2)^{\frac{1}{2}} 2\pi r'dr' = \frac{2i}{3\pi}ka^3H_0\varphi_0. \tag{37}$$

The contribution of K_E is slightly more difficult to calculate because K_E contains the factor r' and therefore φ must be expanded in the form (35a). The first term, φ_0, does not give any contribution so that we obtain

$$\int K_E \varphi dy'dz' = -\frac{ik\varphi_0}{2\pi^2} \times \int_0^{2\pi} d\beta \int_0^a \frac{r'dr'}{(a^2-r'^2)^{\frac{1}{2}}} \kappa\cdot r'\, r'\times E_0, \quad (37a)$$

where r' and β are polar coordinates in the plane of the hole. Remembering that E_0 is in the x direction (normal to the hole), we find

$$\int d\beta \kappa\cdot r' r' \times E_0 = -\pi r'^2 E_0 \times \kappa \quad (37b)$$

and

$$\int K_E \varphi dy'dz' = \frac{i}{3\pi} k a^3 E_0 \times \kappa \varphi_0. \quad (38)$$

Inserting (37) and (38) into (36) we find

$$E = \frac{1}{3\pi} k^2 a^3 \varphi_0 \kappa \times (2H_0 + E_0 \times \kappa). \quad (39)$$

We see that E is always perpendicular to κ, i.e., the waves are transverse. Moreover, for a point r on the conducting screen, κ lies in the plane of the screen and H_0 and $\kappa\times E_0$ do likewise. Therefore E is normal to the screen thus satisfying the boundary conditions.

b. Magnetic Field

The first term of (18) follows directly from (37), (38):

$$H^{(1)} = ik\int K\varphi d\sigma = -\frac{1}{3\pi} k^2 a^3 \varphi_0 (2H_0 + E_0 \times \kappa). \quad (40)$$

This field is not transversal; only the second term $(E_0\times\kappa)$ has this property. From (39) and (40) we see that

$$E = -\kappa \times H^{(1)}. \quad (40a)$$

The second term of (18) gives

$$H^{(2)} = -\int \eta \operatorname{grad} \varphi d\sigma = -k^2 \kappa \varphi_0 \int \eta \kappa \cdot r' d\sigma. \quad (41)$$

Inserting (33a) we get

$$H^{(2)} = \frac{k^2}{\pi^2} \kappa \varphi_0 \int_0^{2\pi} d\beta \int_0^a \frac{r'dr'}{(a^2-r'^2)^{\frac{1}{2}}} \kappa\cdot r' H_0 \cdot r'. \quad (41a)$$

Now

$$\int d\beta \kappa\cdot r' H_0 \cdot r' = \pi r'^2 H_0 \cdot \kappa \quad (41b)$$

and therefore

$$H^{(2)} = \frac{2}{3\pi} k^2 a^3 \varphi_0 \kappa\, H_0 \cdot \kappa. \quad (42)$$

Remembering the vector identity,

$$H_0 - H_0\cdot\kappa\, \kappa = \kappa \times (H_0 \times \kappa), \quad (42a)$$

we find the entire magnetic field, (40)+(42), becomes

$$H = H^{(1)} + H^{(2)} = -\frac{1}{3\pi} k^2 a^3 \varphi_0 \kappa \times (2H_0 \times \kappa - E_0). \quad (43)$$

This expression is a transverse wave as it should be. Thus the field $H^{(2)}$ serves to satisfy the transversality condition. In fact, since (42) is entirely longitudinal (in direction of κ), $H^{(2)}$ *serves only to eliminate the longitudinal component of* $H^{(1)}$ *and leaves the transverse component unchanged.* This result will be important in the application to the cavity problem (Section 9).

Comparing (43) and (39) we see that

$$H = \kappa \times E \quad (44)$$

$$E = -\kappa \times H \quad (44a)$$

as is required by Maxwell's equations for plane waves.

c. Total Radiation

Poynting vector of the diffracted field is

$$S = \frac{c}{4\pi} E \times H = \frac{c}{36\pi^3} \frac{k^4 a^6}{r^2} \kappa (2\kappa \times H_0 - \kappa \times \kappa \times E_0)^2. \quad (45)$$

Now let θ denote the angle between the propagation vector κ and the vector n normal to the screen (x direction) and α the azimuth of κ, i.e., the angle between the plane of κ and n and the plane of H_0 and n. Then the intensity of radiation in the direction κ is, per unit solid angle,

$$r^2|S| = \frac{c}{36\pi^3}k^4a^6[4H_0^2\cos^2\theta\cos^2\alpha + (\sin\theta E_{0,x} - 2H_0\sin\alpha)^2]. \quad (46)$$

The radiation is thus not symmetrical about any axis. The total radiation in all directions is

$$S_{\text{tot}} = \int_0^{\pi/2}\sin\theta d\theta\int_0^{2\pi}d\alpha r^2|S| = \frac{c}{27\pi^2}k^2a^6(4H_0^2+E_0^2). \quad (47)$$

The expressions E_0^2 and H_0^2 on the right-hand side denote the time averages.

6. DISCUSSION OF THE RESULT

a. Comparison with the Kirchhoff Solution

The result for the diffracted field is entirely different from that of the Kirchhoff theory. Quite obviously, the polarization of the radiation is different because our solution satisfies the boundary condition while Kirchhoff's does not. Also the angular distribution of the total radiation is different (cf. Subsection c, below). However, the most striking difference is the absolute value of the field. As is easily seen from expression (5), the Kirchhoff solution gives magnetic and electric fields of the order

$$H_K \sim ka^2H_0.$$

Our solution gives instead [cf. (39), (43)]

$$H_B \sim k^2a^3H_0$$

which differs from H_K by a factor ka. In addition, our solution has a smaller numerical factor. Therefore, for small holes, the radiation transmitted through the hole is very much smaller than Kirchhoff's theory would indicate. The fields E and H are reduced roughly in the ratio a/λ where λ is the wave-length of the radiation. The radiation intensity is therefore reduced by a factor of the order $(a/\lambda)^2$.

b. Representation by Electric and Magnetic Dipoles

Turning now to our solution itself, we note that the field may be considered as owing to an electric and a magnetic dipole. The field of a magnetic dipole of moment M at the origin is, according to usual electromagnetic formulas:

$$H = k^2\varphi_0\kappa\times(M\times\kappa), \quad (48)$$

$$E = -k^2\varphi_0\kappa\times M. \quad (48a)$$

Similarly, the field of an electric dipole of moment P is

$$E = k^2\varphi_0\kappa\times(P\times\kappa) \quad (49)$$

$$H = k^2\varphi_0\kappa\times P. \quad (49a)$$

Comparison with (39) or (43) shows that the hole is equivalent to a magnetic dipole

$$M = -\frac{2}{3\pi}a^3H_0 \quad (50)$$

and an electric dipole

$$P = -\frac{1}{3\pi}a^3E_0. \quad (50a)$$

Each of the dipoles is antiparallel to the respective field in the hole as might be expected. Accordingly, the magnetic dipole is in the plane of the hole and the electric one in the direction of the normal. These directions are just those required to fulfill the boundary conditions for the electric field on the screen. The magnetic moment (50) can easily be calculated from the expression (33a) for the magnetic charge density; we have

$$M = \int\eta r'dr' = -\frac{1}{\pi^2}\int r'\cdot H_0\frac{r'}{(a^2-r'^2)^{\frac{1}{2}}}r'dr'd\beta = -\frac{1}{\pi}H_0\int r'dr'\frac{r'^2}{(a^2-r'^2)^{\frac{1}{2}}} = -\frac{2}{3\pi}H_0a^3. \quad (51)$$

It will be remembered that the solution in Section 4a was actually obtained by assuming a distribution of magnetic dipoles in a flat ellipsoid. As is well known from electrostatics, the inner field produced by such a distribution is proportional to the polarization per unit volume (for any given shape of the ellipsoid); the total moment is therefore proportional to H_0 times the volume which in turn is proportional to a^3.

Dimensional considerations, in combination with the fact that the relation between M and H_0 is a purely magnetostatic problem, also show that the moment M must be proportional to $H_0 a^3$.

The different factor (1 instead of 2) in the electric moment (50a) is no doubt owing to the fact that the electric moment is in the direction of the small axis of the ellipsoid while the magnetic moment is in the direction of one of the long axes.

The order of magnitude of M and P is in general the same although there are cases for which one of them vanishes. This is frequently the case for the electric moment P: It is only necessary to assume incident plane waves with the electric field perpendicular to the plane of incidence. Then the electric field E_0 will have no component perpendicular to the screen, and only H_0 will be different from zero. The radiation intensity is then proportional to $\sin^2 \chi$ where χ is the angle between the direction of H_0 and the direction of propagation of the diffracted waves κ. This radiation is large in the normal direction, and also on the screen in the direction perpendicular to H_0, i.e., in the direction of the electric vector of the incident wave.

To make H_0 vanish and $E_0 \neq 0$, it is necessary to assume standing waves on the "left-hand side" of the screen. For example, in a rectangular cavity with the sides $L_1L_2L_3$, the relevant components of the electric and magnetic fields are given by

$$E_x = A_x \cos\frac{\pi m_1 x}{L_1} \sin\frac{\pi m_2 y}{L_2} \sin\frac{\pi m_3 z}{L_3}, \tag{52}$$

$$H_y = B_y \cos\frac{\pi m_1 x}{L_1} \sin\frac{\pi m_2 y}{L_2} \cos\frac{\pi m_3 z}{L_3}, \tag{52a}$$

$$H_z = B_z \cos\frac{\pi m_1 x}{L_1} \cos\frac{\pi m_2 y}{L_2} \sin\frac{\pi m_3 z}{L_3}, \tag{52b}$$

where A_x, B_y, and B_z are constants and m_1, m_2, m_3 integers. Then if we place the hole on the wall $x=0$ at a point where

$$\frac{m_2 y}{L_2} = n_2 + \tfrac{1}{2} \quad \text{and} \quad \frac{m_3 z}{L_3} = n_3 + \tfrac{1}{2} \tag{52c}$$

(n_2 and n_3 integers), we have $H_{0y} = H_{0z} = 0$ and $E_{0x} \neq 0$. Then the radiated intensity [cf. (46)] is proportional to $\sin^2 \theta$ where θ is the angle between the normal to the screen (direction of E_0) and the direction of propagation. This intensity is a maximum on the screen and zero in the normal direction, quite different from the previous case.

c. Diffraction of a Plane Wave

It is of some interest to discuss the case of an incident plane wave. Let H_i, E_i, κ_i be, respectively, the magnetic field, the electric field, and a unit vector in the direction of propagation for the incident wave. Analogously to (44), (44a) we have

$$H_i = \kappa_i \times E_i, \tag{53}$$

$$E_i = -\kappa_i \times H_i, \tag{53a}$$

and the Poynting vector is

$$S_i = \kappa_i \frac{c}{4\pi} E_i^2 = \kappa_i \frac{c}{4\pi} H_i^2. \tag{54}$$

We shall denote by ϑ the angle of incidence, i.e., between κ_i and n (the vector n is in the x direction); then $0 < \vartheta < \pi/2$. We have to distinguish two cases.

α. Electric Field of Incident Wave Perpendicular to Plane of Incidence

For this case we have E_i perpendicular to the plane of κ_i and n. Then E_i has no normal component, and the tangential component of H_i is

$$H_{i\,\tan} = H_i \cos\vartheta. \tag{55}$$

This tangential component is doubled upon reflection so that $H_0 = 2H_{i\,\tan}$; it lies in the plane of incidence. If, then, β is the angle between the plane of κ and n and the plane of κ_i and n, we have for the diffracted radiation intensity per unit solid angle

$$r^2|S| = \frac{4c}{9\pi^3} k^4 a^6 H_i^2 \cos^2\vartheta (1 - \sin^2\theta \cos^2\beta)$$

$$= \frac{16}{9\pi^2} k^4 a^6 S_i \cos^2\vartheta (1 - \sin^2\theta \cos^2\beta). \tag{56}$$

The "diffraction cross section" of the hole for this polarization is then

$$A_{\perp}=\frac{\int |S|r^2 \sin\theta d\theta d\beta}{S_i}=\frac{64}{27\pi}k^4a^6\cos^2\vartheta. \quad (57)$$

β. Electric Field in Plane of Incidence

If we choose the plane of incidence as XY plane, then κ_i has the components $(\cos\vartheta, \sin\vartheta, 0)$ and E_i has the components $E_i(-\sin\vartheta, +\cos\vartheta, 0)$. According to (53), H_i has then the components $E_i(0, 0, 1)$. The components of κ are $(\cos\theta, \sin\theta\cos\beta, \sin\theta\sin\beta)$. Taking into account the reflection and the fact that only the normal component of E_i is effective, we get

$$E_0=-2E_i(\sin\vartheta, 0, 0),$$

$$H_0=2E_i(0, 0, 1), \quad (57a)$$

and [cf. (45)]

$$\begin{aligned}(2\kappa\times H_0-\kappa\times\kappa\times E_0)^2 &=4(\kappa\times H_0)^2+(\kappa\times E_0)^2+4\kappa\cdot(H_0\times E_0)\\ &=4E_i^2[4(1-\sin^2\theta\sin^2\beta)+\sin^2\theta\sin^2\vartheta\\ &\quad -4\sin\theta\sin\vartheta\cos\beta]. \quad (57b)\end{aligned}$$

Inserting into (45) we find for the differential cross section

$$\begin{aligned}\frac{dA}{\sin\theta d\theta d\beta}&=\frac{r^2|S|}{S_i}\\ &=\frac{16}{9\pi^2}k^4a^6[\cos^2\theta+\sin^2\theta(\cos^2\beta\\ &\quad+\tfrac{1}{4}\sin^2\vartheta)-\sin\theta\sin\vartheta\cos\beta] \quad (58)\end{aligned}$$

and for the total cross section

$$A_{\parallel}=\frac{64}{27\pi}k^4a^6(1+\tfrac{1}{4}\sin^2\vartheta). \quad (59)$$

The most remarkable feature of the angular distribution (58) is the dependence on the azimuth β: For a given θ, the radiation is smaller for $\beta=0$ (i.e., in the direction closest to that of the incident wave), than in the "reflected" direction $\beta=\pi$. We suspected an error in sign to be the cause of this result, but we have found none. Perhaps the most convincing check on the sign is provided by Eqs. (48) to (50a). The result mentioned is of course exactly the opposite of that expected from any elementary considerations based on the Huygens principle. For any value of θ, the radiation is a minimum in the azimuth given by

$$\cos\beta_0=\tfrac{1}{2}\sin\vartheta/\sin\theta. \quad (58a)$$

In the plane of the screen, the radiation is zero for $\beta=\beta_0$, which in this case lies between 60° and 90°.

The total cross section (59) is greater than that for the α-direction of polarization, (57), except for normal incidence when the two expressions are equal. For unpolarized light,

$$A'=\frac{64}{27\pi}k^4a^6(1-\tfrac{3}{8}\sin^2\vartheta). \quad (59a)$$

All cross sections are proportional to λ^{-4} as in Rayleigh's theory of the scattering by small objects. Also the proportionality with a^6 is the same as in the scattering theory. The cross section is of the same order as the scattering cross section of a dielectric sphere or disk of radius a and dielectric constant of the order of 2.

d. Comparison with Sommerfeld's Solution

One characteristic feature of our solution is that the current and charge distribution K, η determine the field only in the half-space on the right of the screen. The perturbation of the field on the left, H_1 and E_1 [cf. (7), (8)] would be obtained by a current and charge distribution $-K$, $-\eta$. This reversal of sign is unusual in electrodynamic theory. But a somewhat similar procedure had to be used by Sommerfeld in the wedge problem. In order to take into account that the field is entirely different on the two sides of the screen, Sommerfeld put a "branch plane" in the screen and obtained the solution of the problem in a Riemann space. This is the mathematical expression for the fact that the field is not given by the same integral representation on the two sides of the screen.

7. FIELD IN AND NEAR THE HOLE

a. Inside the Hole

From the conditions (10), (10a), (7), (9a), (9c), we see immediately that in the hole

$$H_{\tan}=\tfrac{1}{2}H_{0\,\tan}, \tag{60}$$

$$E_n=\tfrac{1}{2}E_{0,\,n}. \tag{60a}$$

$H_{\tan}$ and E_n are thus halfway between their values in the "unperturbed" fields on the right and on the left side of the screen. On the other hand, those field components which would be zero if the hole were absent, i.e., the normal component of H and the tangential component of E, are determined directly by η and K. Since $4\pi\eta$ measures the discontinuity[4] of H_n, we have, at a point just to the right of the hole,

$$H_x(r)=2\pi\eta=-\frac{2}{\pi}\frac{H_0\cdot r}{(a^2-r^2)^{\frac{1}{2}}}. \tag{60b}$$

In the hole, therefore, the normal component of H is of the same order of magnitude as the tangential component. This does not in any way invalidate our calculations because we have not used a perturbation method in which H_x was considered small compared with $H_{\tan}$, but we have satisfied the boundary conditions exactly.

In contrast to $H_{\tan}$ which was assumed to be constant, H_x varies rapidly over the surface of the hole. Near the boundary $r=a$, H_x becomes infinite as $1/(a-r)^{\frac{1}{2}}$. An integrable infinity of this type occurs frequently at edges. It occurs, e.g., in Summerfeld's solution of the diffraction by a half-plane, in the neighborhood of the diffracting edge.

The tangential component of the electric field is given by [cf. Stratton, p. 467, Eq. (23)]

$$E_{\tan}=2\pi n\times K, \tag{61}$$

where n is in the positive x direction (Stratton's n is in our negative x direction). Inserting (33) and remembering that in the hole $r\cdot n=0$, we have

$$E_{\tan}=-\frac{1}{\pi}\frac{\mathbf{r}}{(a^2-r^2)^{\frac{1}{2}}}E_{0,x}+\frac{2}{\pi}ik(a^2-r^2)^{\frac{1}{2}}n\times H_0. \tag{61a}$$

[4] The discontinuity of H_n at $x=0$ is, of course, only mathematical. As explained in Section 6d, the current distribution K, η determines the field only for $x>0$. The physical field H_n is continuous at $x=0$, according to (9b).

The second term is in general negligible compared with the first since H_0 and E_{0x} are usually of the same order of magnitude. The first term is again of the same order as E_{0x} [cf. (60b)] and is singular near the edge of the hole. The tangential electric field is directed radially outwards from the center of the hole.

b. E_{norm} and $H_{\tan}$ near the Hole

Outside the hole, $E_{\tan}$ and H_n are of course zero. The other field components have singularities near the edge, behaving as $1/(r^2-a^2)^{\frac{1}{2}}$. At larger distance from the hole, $H_{\tan}$ and E_n behave as H_0a^3/r^3, up to distances of the order λ. For still larger distances, the solutions of Section 5 hold. The field in the neighborhood of the hole has not yet been calculated in detail. It is clear, however, that $H_{\tan}$ and E_n are very far from zero on the screen in contrast to the assumption made in Kirchhoff's theory.

The tangential component of H is especially important because it determines the dissipation of energy (joule heat) in the screen if the latter has finite conductivity. The energy dissipation per unit time is

$$P=\tfrac{1}{4}\nu D\int H_{\tan}^2 dA, \tag{62}$$

where $\nu=\omega/2\pi$ is the frequency,

$$D=\frac{1}{2\pi(\sigma\nu)^{\frac{1}{2}}} \tag{62a}$$

is the thickness of the current carrying sheath in the conductor, σ is the conductivity in electromagnetic units ($=10^{-9}$ times conductivity in mho), and the magnetic permeability has been assumed to be 1. The integral in (62) goes over the whole surface of the conductor. Since $H_{\tan}$ behaves as $1/(r^2-a^2)^{\frac{1}{2}}$ near the edge of the hole, (62) diverges logarithmically; it must be cut off at a value of $r-a$ of the order of D. The power dissipation on the right-hand surface of the screen becomes then of the order

$$P\sim H_0^2\nu Da^2\ln(a/D). \tag{62b}$$

An accurate calculation of this quantity would be useful for the theory of cavities with small holes.

8. POSSIBLE EXTENSIONS OF THE THEORY

In its present form, the theory is only applicable to holes small compared with the wave-length. However, it seems possible to extend it to holes comparable in size with the wave-length. In this case, the given fields H_0 and E_0 will contain factors of the type $e^{ik\cdot r'}$. Similarly, the variation of the factor $e^{ik|r-r'|}$ in the Green's function φ must be considered. From an approximate consideration, it seems that the correction terms will be of relative order $(ka)^2$ rather than ka.

It seems certainly possible to obtain solutions in terms of power series in y' and z'. The integral (20), for instance, can be evaluated if η is any odd power of $(a^2-r'^2)^{\frac{1}{2}}$, giving for H_{0y} a power series in r^2. From this type of solution, others can be obtained by differentiation with respect to y and z. We believe that in this way a solution for an arbitrary function $H_{0y}(y, z)$ can be obtained. However, a more elegant solution may perhaps be found in terms of the electric oscillations of an ellipsoid.

A second question is the dependence on the shape of the hole. From the way in which our solution was obtained in Section 4a, it is clear that a solution can also be found for an elliptical hole. The solution for an ellipse of arbitrary eccentricity should give a sufficient idea about the dependence of the diffraction on the shape of the hole. However, the case of a rectangle may also be solvable.

A more difficult question is the extension to screens and holes which are not plane. The case of a small hole in a curved surface will probably still give a result very similar to ours, as long as the surface may be considered plane over a region large compared with the hole. This is usually the case since the radius of curvature of cavities is ordinarily of the order of the wave-lengths. If this condition is fulfilled, the boundary condition on the surface will be violated only in the wave zone where the diffracted field is no longer very large. Perhaps a method of successive approximations will be applicable to this case.

However, a different situation arises when the hole itself is essentially curved. A simple example is a gap in a cylindrical wave guide; in this case, the gap itself is the surface of a short cylinder. Such a problem will presumably require an entirely new solution, but it is likely that at least the same principle will work as in our case, and that the symmetry of the cylinder will be helpful for obtaining the solution.

9. APPLICATION TO THE THEORY OF CAVITIES

Condon[5] has given a most convenient theory of the excitation of cavities by electric current loops. In this theory, the vector potential A is developed in terms of normal modes A_m, thus:

$$A=\sum_m p_m(t)A_m(r). \tag{63}$$

By inserting this expression into Maxwell's equations, a differential equation is obtained for p_m in terms of the current I exciting the cavity, *viz.:*

$$\frac{d^2p_m}{dt^2}+\frac{\omega_m}{Q_m}\frac{dp_m}{dt}+\omega_m{}^2p_m=\frac{4\pi I}{V}\int A_m\cdot ds, \tag{64}$$

where ω_m is the frequency of the mth normal mode, Q_m is the dissipation constant, V is the volume of the cavity, and the integral is taken along the conductor carrying the current I. The current is measured in electromagnetic units.

Since our solution is obtained in terms of magnetic currents and charges, it is convenient to expand the field in a slightly different way, by using a magnetic vector potential F as defined in (14), (14a). The scalar potential ψ will be unimportant because the magnetic field is transverse. We expand:

$$F=\sum_m q_m(t)F_m(r) \tag{65}$$

and have

$$H=\frac{1}{c}\sum_m \frac{dq_m}{dt}F_m(r), \tag{65a}$$

$$E=\sum_m q_mk_mA_m(r) \tag{65b}$$

with

$$A_m=\text{curl } F_m/k_m. \tag{65c}$$

The F_m form an orthonormal set:

$$\int F_l\cdot F_m dV=V\delta_{lm}, \tag{65d}$$

where V is the volume of the cavity. Our q_m and

[5] E. V. Condon, J. App. Phys. **12**, 129 (1941).

Condon's p_m stand in the relation of coordinate and momentum.

The solution (65) satisfies automatically the first Maxwell equation (15), (15a) (cf. Section 4).

The F_m's are determined in such a way that also

$$\text{div } F_m = 0. \tag{66}$$

Moreover, the F_m satisfy the wave equation

$$\nabla^2 F_m + k^2 F_m = 0 \tag{66a}$$

with

$$k_m = \omega_m / c. \tag{66b}$$

Because of (15), (66), (66a) the second Maxwell equation

$$\text{curl } E = -\frac{1}{c}\frac{\partial H}{\partial t} \tag{66c}$$

is also fulfilled by the particular solutions F_m.

Now assume a distribution of magnetic currents of current density J^* per unit volume. Then the Maxwell equation (66c) is replaced by (11a). Inserting Eqs. (14), (14a) we obtain

$$\text{curl curl } F + \frac{1}{c^2}\frac{\partial^2 F}{\partial t^2} = -4\pi J^*. \tag{66d}$$

Using the expansion (65), the condition (66), and the wave equation (66a), and expanding the right-hand side of (66d) in terms of eigenfunctions F_m, we obtain

$$\frac{d^2 q_m}{dt^2} + \frac{\omega_m}{Q_m}\frac{dq_m}{dt} + \omega_m{}^2 q_m = -\frac{4\pi c^2}{V}\int dV J^* \cdot F_m. \tag{67}$$

The second term has been added to take into account the energy dissipation in the walls of the cavity. If we have magnetic surface currents K instead of volume currents J^*, the right-hand side of (67) is replaced by

$$-\frac{4\pi c^2}{V}\int d\sigma K \cdot F_m. \tag{67a}$$

Equations (67), (67a) are obviously the exact analog of Condon's Eq. (64).

It will be noted that (67a) depends only on K, not on η. Indeed, as we have shown in Section 5b, the magnetic charge η serves only to eliminate the longitudinal component of the magnetic field and does not affect the transverse component at all. Now the normal modes of the cavity all have transverse magnetic fields. Therefore to determine their excitation we need only the magnetic current distribution K, not η. Moreover, our expansion in transverse modes will give us the complete field because the longitudinal field produced by K is canceled by the effect of η.

We may now evaluate (87a) in terms of $H_{0\,\text{tan}}$ and E_{0n}. Inserting (33), we find

$$\int K_H \cdot F_m d\sigma = \frac{ik}{\pi^2} H_0 \cdot F_m(0) \int (a^2 - r'^2)^{\frac{1}{2}} 2\pi r' dr'$$

$$= \frac{2i}{3\pi} k a^3 H_0 \cdot F_m(0), \tag{68}$$

where $F_m(0)$ is the value of the "eigenfunction" F_m at the center of the hole. To evaluate the contribution of K_E, Eq. (33), we must expand F_m in powers of r' and keep the linear term:

$$F_{my} = F_{my}(0) + \frac{\partial F_{my}}{\partial y'} y' + \frac{\partial F_{my}}{\partial z'} z'. \tag{68a}$$

The constant term in (68a) gives no contribution, and we obtain

$$\int K_E \cdot F_m d\sigma = \frac{E_{0x}}{2\pi^2}\int_0^{2\pi} d\beta \int_0^a \frac{r' dr'}{(a^2 - r'^2)^{\frac{1}{2}}}$$

$$\cdot\left[\left(\frac{\partial F_{my}}{\partial y'} y' + \frac{\partial F_{my}}{\partial z'} z'\right) z' - \left(\frac{\partial F_{mz}}{\partial y'} y' + \frac{\partial F_{mz}}{\partial z'} z'\right) y'\right]$$

$$= -\frac{1}{3\pi} a^3 E_{0x} \,\text{curl}_x\, F_m$$

$$= -\frac{1}{3\pi} k_m a^3 E_{0x} A_{mx}(0). \tag{69}$$

In the last transformation, we have used relation (65c). Inserting (68), (69) into (67) we obtain

$$\frac{d^2 q_m}{dt^2} + \frac{\omega_m}{Q_m}\frac{dq_m}{dt} + \omega_m{}^2 q_m$$

$$= \frac{4}{3}\frac{c^2 a^3}{V}[k_m E_{0x} A_{mx}(0) - 2ik H_0 \cdot F_m(0)]. \tag{70}$$

We must now determine H_0 and E_0. The simplest case, perhaps, is that of a cavity communicating with outside space through a small hole. If radiation comes from the outside and if there is no radiation in the cavity, then H_0, E_0 are simply the incident magnetic and electric field at the hole. If there is also a field in the cavity, it is easy to see that in order to satisfy the boundary conditions in the hole, we must choose H_0 and E_0 equal to the *difference* between the outside and the inside field.

A case of some interest is that of a cavity with a field inside, radiating into space through a small hole. Then H_0, E_0 must be replaced by the inside field at the hole, with negative sign. Some caution is necessary since it has been assumed in the derivation in Section 4 that the fields depend on time as $e^{-i\omega t}$. Therefore, if the inside field H is to be obtained from (65a), we must put[6]

$$(1/c)(dq_m/dt) = -ikq_m. \tag{70a}$$

Now let us suppose that the frequency is close to the natural frequency ω_1 of a certain mode, 1, of the cavity. Then we may put in sufficient approximation

$$E_{0x} = -q_1k_1A_{1x}, \tag{71}$$

$$H_0 = -\frac{1}{c}\frac{dq_1}{dt}F_1 = ik_1q_1F_1 \tag{71a}$$

and obtain

$$\frac{d^2q_m}{dt^2}+\frac{\omega_m}{Q_m}\frac{dq_m}{dt}+\omega_m{}^2q_m = -\frac{4}{3}\frac{a^3}{V}\omega_1q_1[2\omega F_1(0)\cdot F_m(0) - \omega_mA_{1x}(0)A_{mx}(0)], \tag{72}$$

where the argument 0 denotes the value of the respective function at the center of the hole.

Equation (72) describes the excitation of all normal modes m by one normal mode, 1, through the action of the hole. In the stationary state, i.e., when

$$q_1 = A\sin\omega_1 t \tag{73}$$

the excitation of the mth mode will be given by

$$q_m = \frac{4}{3}A\frac{a^3}{V}\sin\omega_1 t\frac{\omega_1}{\omega_m{}^2-\omega_1{}^2} \times[2\omega_1F_1(0)\cdot F_m(0) - \omega_mA_{1x}(0)A_{mx}(0)]. \tag{73a}$$

Thus quite similar to the perturbation theory of wave mechanics, the other modes m will be "mixed in" with the main mode on account of the "perturbation" by the small hole. However, it is in general not necessary to use the coefficients q_m because the field distribution near the hole is given more directly by the integrals (17), (18) and is only very little modified by the presence of the rest of the cavity (for given H_0, E_0).

The most interesting result is obtained by applying Eq. (72) to the main mode $m=1$ which gives

$$\frac{d^2q_1}{dt^2}+\frac{\omega_1}{Q_1}\frac{dq_1}{dt}+\omega_1{}^2q_1 = 2\omega_1{}^2\gamma q_1, \tag{74}$$

where the emission coefficient γ is

$$\gamma = \frac{2a^3}{V}[2F_1{}^2(0) - A_{1x}^2(0)]. \tag{74a}$$

The right-hand side of (74) will cause a shift of frequency from ω_1 to

$$\omega_1' = \omega_1(1-\gamma). \tag{75}$$

The amount and sign of this shift depends on the value of F_1 and A_{1x} at the hole. If A_{1x} is small, the frequency will be decreased; if $F_1 \ll A_{1x}$, there will be an increase of frequency.

Both alternatives may occur in practice: E.g., in a rectangular cavity with electric field in the x direction, the field will be given by (52) with $m_1=0$. Then the electric field will be zero on the XY and the XZ plane; a hole in either of these planes will therefore *lower* the natural frequency. On the YZ plane, at the points determined by (52c), $F_1=0$ and A_{1x} is a maximum; therefore a hole at these positions will increase the frequency. If $m_2=m_3=1$, this is the case in the middle of the YZ plane ($y=\frac{1}{2}L_2$, $z=\frac{1}{2}L_3$). For the same mode, a hole at $y=\frac{1}{4}L_2$, $z=\frac{1}{4}L_3$ will leave the frequency unchanged.

At least the first of these results can be understood qualitatively from the pattern of

[6] It can be easily verified that the correct final result is obtained by this method even if $q_m \sim \sin\omega t$ or $\cos\omega t$. It is only necessary to go through the derivations from the beginning with a different time dependence.

electric field lines. If the hole is in a plane parallel to the undisturbed electric field (XY or XZ plane), then in the hole the boundary condition $E_x=0$ is relaxed. This is equivalent to a slight extension of the cavity, and therefore the characteristic frequency will be lowered. The electric field near the hole will behave about as indicated in Fig. 3 (exaggerated).

The case of a hole in the YZ plane is not quite so clear. A possible interpretation is that the hole reduces the value of E_x to one-half and that therefore the cavity behaves as if it were smaller, giving a higher frequency. However, a closer investigation of the fields near the hole seems necessary to understand the effects properly.

It will be noted that γ is purely real. This means that only the frequency is changed but not the energy dissipation. This is in agreement with the result in Section 5c that the intensity of the emitted radiation is of order a^6 rather than a^3. The radiation would therefore contribute a dissipative term of order γ^2 only. However, it must be remembered that our theory does not properly take into account the finite conductivity of the cavity walls. The dissipation in the walls will be increased by the presence of the hole (cf. end of Section 7), the increase being of order $a^2\sim\gamma^{\frac{2}{3}}$ [Eq. (62b)].

10. COUPLED CAVITIES

Let us assume two cavities with one common wall with a hole in it (Fig. 1). The fields in the

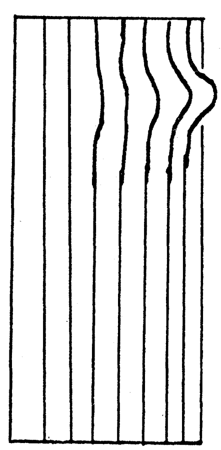

FIG. 3. Electric field lines in a cavity with a small hole in one of the sides parallel to the field. The density of the lines is not proportional to the field strength, the bulge near the hole is exaggerated.

two cavities will be denoted by subscripts α and β, respectively. E.g., $q_{\alpha m}$ will be the amplitude of the mth normal mode of cavity α. Cavity α shall be on the "left" of the boundary $x=0$, while β extends to positive x. We shall assume that mode 1 of α, and mode 1 of β have nearly equal frequencies, and that the cavities are excited with a frequency

$$\omega\approx\omega_{\alpha 1}\approx\omega_{\beta 1}. \tag{76}$$

Considering the excitation of cavity β, we have to insert for H_0 and E_{0x} the difference between the fields in cavities α and β. We need only take into account the first mode of each cavity so that, similar to (71):

$$E_{0x}=k_{\alpha 1}q_{\alpha 1}A_{\alpha 1x}-k_{\beta 1}q_{\beta 1}A_{\beta 1x}, \tag{77}$$

$$H_0=-ik_{\alpha 1}q_{\alpha 1}F_{\alpha 1}+ik_{\beta 1}q_{\beta 1}F_{\beta 1}. \tag{77a}$$

We may drop the subscript 1 and refer in the following only to the first mode. Moreover, because of (76), we may put $k_{\alpha 1}=k_{\beta 1}$ in E_{0x} and H_0. Then we obtain the differential equations

$$\begin{aligned}\frac{d^2q_\alpha}{dt^2}+\frac{\omega_0}{Q_\alpha}\frac{dq_\alpha}{dt}+\omega_\alpha^2q_\alpha&=\frac{4}{3}\omega_0^2\frac{a^3}{V_\alpha}(c_{\alpha\alpha}q_\alpha-c_{\alpha\beta}q_\beta),\\ \frac{d^2q_\beta}{dt^2}+\frac{\omega_0}{Q_\beta}\frac{dq_\beta}{dt}+\omega_\beta^2q_\beta&=\frac{4}{3}\omega_0^2\frac{a^3}{V_\beta}(c_{\beta\beta}q_\beta-c_{\alpha\beta}q_\alpha),\end{aligned} \tag{78}$$

with

$$\begin{aligned}c_{\alpha\alpha}&=2F_\alpha^2-A_{\alpha x}^2,\quad c_{\beta\beta}=2F_\beta^2-A_{\beta x}^2,\\ c_{\alpha\beta}&=c_{\beta\alpha}=2F_\alpha\cdot F_\beta-A_{\alpha x}A_{\beta x}.\end{aligned} \tag{78a}$$

The values of F_α, etc., have to be taken in the hole. In all small terms in (78), ω_α and ω_β have been replaced by ω_0. Equation (78) has exactly the same form as for two conventional coupled circuits and can be treated in the same way.

Let us consider the case of two nearly equal cavities, and let us choose the eigenfunctions so that the normalized magnetic fields F_α and F_β are identical at the common boundary. Then, according to (65c) we have also $A_{\alpha x}=A_{\beta x}$ and therefore

$$c_{\alpha\alpha}=c_{\beta\beta}=c_{\alpha\beta}. \tag{79}$$

We put

$$\gamma=\frac{4}{3}\frac{a^3}{V}(2F_\alpha^2(0)-A_{\alpha x}^2(0)); \tag{79a}$$

then γ is a dimensionless quantity which measures the coupling between the two cavities. Then (78) becomes

$$\frac{d^2q_\alpha}{dt^2}+\frac{\omega_0}{Q}\frac{dq_\alpha}{dt}+\omega_\alpha^2q_\alpha=\omega_0^2\gamma(q_\alpha-q_\beta),$$
$$\frac{d^2q_\beta}{dt^2}+\frac{\omega_0}{Q}\frac{dq_\beta}{dt}+\omega_\beta^2q_\beta=\omega_0^2\gamma(q_\beta-q_\alpha). \quad (80)$$

In the last term on the left hand we have kept ω_α and ω_β to allow for a possible lack of tuning. In the other (correction) terms we have put $\omega_\alpha=\omega_\beta=\omega_0$. Neglecting the damping, we find the frequencies of the coupled cavities are given by

$$\omega^2=\tfrac{1}{2}(\omega_\alpha^2+\omega_\beta^2)-\omega_0^2\gamma \pm(\tfrac{1}{4}(\omega_\alpha^2-\omega_\beta^2)^2+\omega_0^4\gamma^2)^{\frac{1}{2}}. \quad (81)$$

If the two cavities are considerably out of tune, i.e., if

$$|\omega_\alpha-\omega_\beta|\gg\omega\gamma, \quad (81a)$$

they will oscillate almost independently, i.e., there will be one proper frequency close to ω_α at which mostly the first cavity is excited while the second has very little excitation, another proper frequency close to ω_β for which the reverse is the case (cf. Section 11).

The cavities are well tuned if

$$|\omega_\alpha-\omega_\beta|\ll\omega\gamma. \quad (81b)$$

The tuning requirement is thus determined by the size of the hole, cf. (79). If $\omega_\alpha=\omega_\beta=\omega_0$, (81) gives

$$\omega^2=\omega_0^2(1-\gamma\pm\gamma). \quad (82)$$

Thus we get one mode whose frequency is the same as for the uncoupled cavities,

$$\omega_1=\omega_0, \quad (82a)$$

and another mode of frequency

$$\omega_2=\omega_0(1-\gamma) \quad (82b)$$

when we neglect terms of order γ^2.

In the first mode we have [cf. (80)]

$$q_\alpha=q_\beta. \quad (83)$$

This means that the tangential magnetic field at the common boundary is continuous, $H_\alpha=H_\beta$, and the same holds for the normal component of the electric field. The hole, then, does not perturb the field at all and the two cavities oscillate as if there were no coupling. This explains immediately why the frequency is not disturbed. Nevertheless, the coupling is important because it determines the phase relations between the two cavities. For the simple case of two rectangular cavities oscillating in their lowest mode we have plotted the electric and magnetic field in Fig. 4. The electric field is assumed to be parallel to the common boundary; it is zero inside the hole just as on the conducting wall.

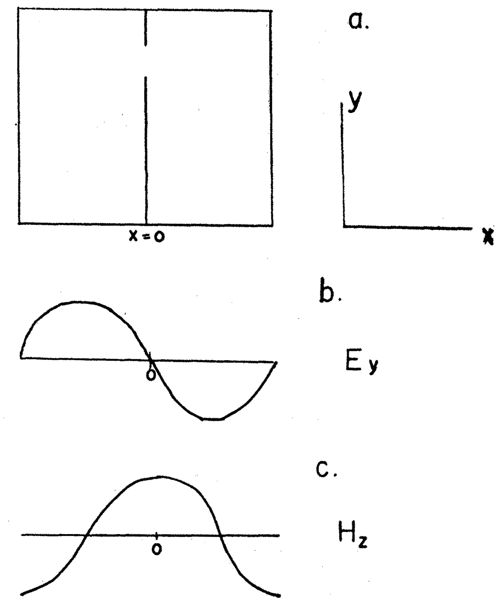

FIG. 4. Field distribution in the first mode ω_1 of two coupled tuned cavities. (a) Cross section of cavities. (b) Electric field E_y, as a function of x (same curve for any value of y). (c) Magnetic field H_z.

The second mode has

$$q_\alpha=-q_\beta. \quad (83a)$$

Then, the magnetic field changes its sign when the boundary between the cavities is crossed (cf. Fig. 5). In the hole, the tangential component of the magnetic field is zero. Around the hole, there is a large perturbation of the field. The tangential electric field has a maximum in the hole. In Fig. 5 we have plotted the electric and magnetic field for the same cavity as Fig. 4, but this time for the second mode, $\omega_2=\omega_0(1-\gamma)$. Since in our case the unperturbed electric field has no component normal to the common boundary, γ is positive, and the frequency is reduced by the coupling.

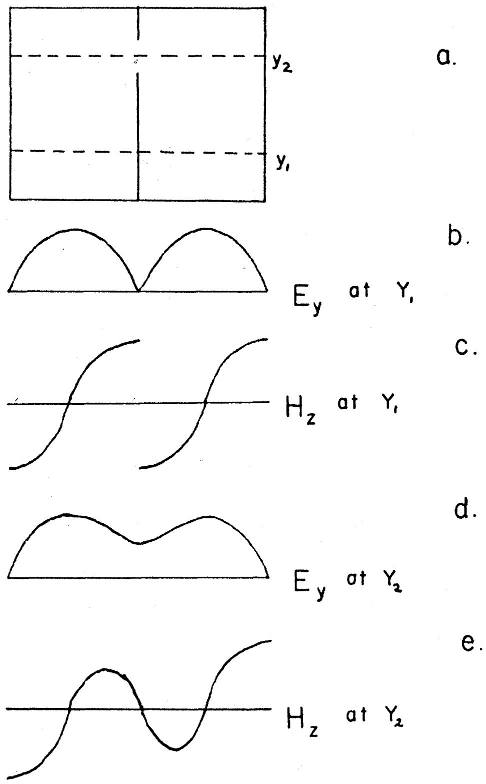

FIG. 5. Field distribution in the second mode ω_2 of the cavities of Fig. 4. (a) Cross section [identical with 4(a)]. (b) Electric field E_y along a line not going through the hole ($y=y_1$). (c) Magnetic field along the same line. (d) Electric field E_y along a line $y=y_2$ going through the hole. (e) Magnetic field along the same line.

This corresponds to the fact that now, so to speak, the two cavities oscillate as a whole; the electric field E_y no longer goes to zero in the hole but has a secondary maximum there, thus beginning to develop towards the field corresponding to the lowest mode in the larger cavity $\alpha+\beta$ with the central boundary absent (Fig. 6). This mode is entirely different from that of the separate cavities (Fig. 4); therefore in mode 2 (which is the lowest harmonic for the larger cavity of Fig. 6), a large perturbation field exists around the hole which varies rapidly in space. The energy dissipation is likely to be considerably greater than for mode 1 (cf. Section 7, end).

The coupling coefficient γ depends sensitively on the position of the hole, and may be zero for certain positions. For example, in our case of two rectangular cavities with an electric field in the y direction, the only field component which is not zero on the common boundary ($x=0$) is H_z. If we do not restrict our attention to the lowest mode, we have for the normalized field:

$$F_z=\frac{m_1/L_1}{[(m_1/L_1)^2+(m_3/L_3)^2]^{\frac{1}{2}}}\times 2\cos\frac{\pi m_1 x}{L_1}\sin\frac{\pi m_3 z}{L_3}, \quad (84)$$

where m_1 and m_3 are positive integers not zero and $L_1L_2L_3$ are the lengths of the three edges of a single cavity. Then we have for the coupling coefficient [cf. (79)]

$$\gamma=\frac{8}{3}\frac{a^3}{L_1L_2L_3}\times\frac{(m_1/L_1)^2}{(m_1/L_1)^2+(m_3/L_3)^2}4\sin^2\frac{\pi m_3 z}{L_3}. \quad (84a)$$

The last factor possesses nodes on the common boundary, $x=0$, for any value of m_3 greater than one. In general, large coupling is desired; then the hole should be put at or near the maximum of γ.

If definite phase relations between the oscillations of the two cavities are required, the correct frequency must be selected. It seems possible to predict, in many cases at least, whether the desired phase relation corresponds to the higher or the lower frequency. In any case, two coupled

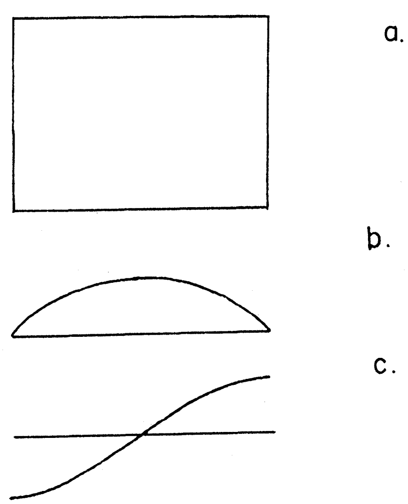

FIG. 6. (a) The cavity of Figs. 4 and 5 without the wall in the middle. (b) Electric field distribution E_y in the lowest mode of the cavity a. (c) Magnetic field in the same mode.

cavities which are exactly tuned will give two modes with a relative frequency difference of the order $\gamma \sim a^3/V$, the phase relations between the fields in the two cavities being opposite for the two modes.

11. STEP-UP COUPLING

It is often desirable to couple two circuits in such a way that the field in β is stronger than in α. This can be achieved for two exactly tuned cavities if they are coupled asymmetrically, *viz.*, by a hole which is in a position corresponding to a high value of the normalized field F_α and to a low one of F_β (Fig. 7). Then the excitation may be expected to adjust itself so that the magnetic field is continuous in the hole, *viz.:*

$$H_\alpha = q_\alpha F_\alpha \text{ (hole)} = H_\beta = q_\beta F_\beta \text{ (hole)}, \quad (85)$$

where F_α, F_β are the normalized fields. Then, since F_β is made small compared with F_α, we get $q_\beta \gg q_\alpha$ as desired.

We shall show in this section that (85) is actually fulfilled for *one* of the two modes, *viz.*, that having the same frequency as the uncoupled cavities. There is, however, another mode in which the field in β is smaller than in α. Furthermore, we are going to show that two slightly mistuned cavities coupled symmetrically will give the same results as two exactly tuned cavities coupled asymmetrically.

The problem of cavities coupled asymmetrically in the way indicated in Fig. 7 can be easily treated by Eqs. (78) to (78c). It is convenient to define the "displaced frequencies."

$$\begin{aligned} \omega_\alpha'^2 &= \omega_\alpha^2 - (4/3)\omega_0^2 c_{\alpha\alpha}(a^3/V), \\ \omega_\beta'^2 &= \omega_\beta^2 - (4/3)\omega_0^2 c_{\beta\beta}(a^3/V), \end{aligned} \quad (86)$$

and the coupling coefficient

$$\gamma = (4/3)\omega_0^2 c_{\alpha\beta}(a^3/V), \quad (86a)$$

the volumes of the two cavities having been assumed equal. Then, if ω denotes the actual frequency, (78) becomes

$$\begin{aligned} (\omega_\alpha'^2 - \omega^2)q_\alpha + \omega_0^2\gamma q_\beta &= 0, \\ \omega_0^2\gamma q_\alpha + (\omega_\beta'^2 - \omega^2)q_\beta &= 0, \end{aligned} \quad (87)$$

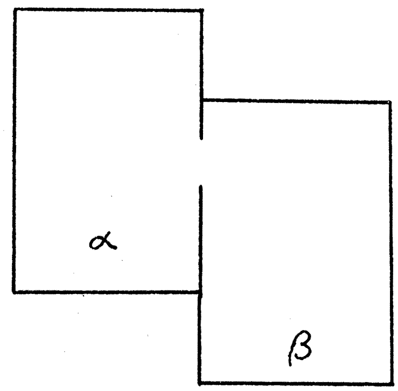

FIG. 7. Asymmetrical coupling of the two tuned cavities.

which has the two solutions

$$\omega_{1,2}^2 = \tfrac{1}{2}(\omega_\alpha'^2 + \omega_\beta'^2) \pm [\tfrac{1}{4}(\omega_\alpha'^2 - \omega_\beta'^2)^2 + \omega_0^4\gamma^2]^{\frac{1}{2}}. \quad (87a)$$

The ratio of the excitation amplitudes q_α, q_β is:

$$\frac{q_\beta}{q_\alpha} = \lambda \pm (1+\lambda^2)^{\frac{1}{2}}, \quad (88)$$

the upper sign corresponding to the higher frequency ω_1, the lower sign to ω_2. λ is given by

$$\lambda = \frac{\omega_\beta'^2 - \omega_\alpha'^2}{2\gamma\omega_0^2} \approx \frac{\omega_\beta' - \omega_\alpha'}{\gamma\omega_0}. \quad (88a)$$

We shall now make the simplifying assumptions [cf. (78a), (78b)] that A_x is zero in the hole, and that F_α and F_β have the same direction. These assumptions are fulfilled by the cavities treated in Section 10 where $A = A_y$ and F is in the z direction everywhere on the common boundary of the cavities. Then we may write [cf. (78a), (78b)]

$$c_{\alpha\alpha} = c_{\alpha\beta}\mu, \quad (89)$$

$$c_{\beta\beta} = c_{\alpha\beta}/\mu, \quad (89a)$$

with

$$\mu = F_\alpha/F_\beta. \quad (89b)$$

Then (86) reduced to

$$\begin{aligned} \omega_\alpha'^2 &= \omega_\alpha^2 - \gamma\omega_0^2\mu, \\ \omega_\beta'^2 &= \omega_\beta^2 - \gamma\omega_0^2/\mu, \end{aligned} \quad (90)$$

and (88a) becomes

$$\lambda = (\mu^2 - 1)/2\mu. \quad (90a)$$

If we now assume that the cavities were originally tuned, i.e., $\omega_\alpha = \omega_\beta = \omega_0$, we obtain, by inserting (90), (90a) into (87a) and (88)

$$\omega_1^2 = \omega_0^2, \tag{91}$$

$$\omega_2^2 = \omega_0^2[1 - \gamma(\mu^2+1)/\mu], \tag{91a}$$

$$(q_\beta/q_\alpha)_1 = \mu, \tag{92}$$

$$(q_\beta/q_\alpha)_2 = -1/\mu. \tag{92a}$$

We find thus one mode ω_1 of frequency equal to the unperturbed frequencies, and another mode ω_2 of lower frequency. In the first mode the ratio of the amplitudes of excitation q_β/q_α is such that the magnetic fields in the hole are matched, in agreement with Eq. (85). This mode, then, gives the desired stepping up of the amplitude from α to β. However, in the second mode, the ratio of the q's is reversed (in addition, the phase relation is opposite). Caution is therefore necessary to excite the coupled cavities with the correct frequency in order to obtain the desired results. (For frequencies other than ω_1 and ω_2, beats will occur.)

Moreover, we see that it is not at all necessary to couple the cavities unsymmetrically in order to step up the excitation from α to β. It is only important [cf. (88)] that λ be different from zero, i.e. [cf. (88a)], *that* ω_α' *and* ω_β' *be different.* This can be achieved either by making $\omega_\alpha = \omega_\beta$ but $c_{\alpha\alpha} \neq c_{\beta\beta}$; then we have two tuned but unsymmetrically coupled cavities as treated in (89) to (92a). But it can equally well be achieved by two slightly mistuned and symmetrically (or also unsymmetrically) coupled cavities; then $\omega_\alpha \neq \omega_\beta$ and $c_{\alpha\alpha} = c_{\beta\beta}$ so that

$$\lambda = \frac{\omega_\beta - \omega_\alpha}{\gamma\omega_0}. \tag{93}$$

In either case, the more strongly excited cavity is the one whose ω' is closer to the actual frequency.