Newell[27] has calculated the spherically symmetrical contribution σ^L to the magnetic shielding constant for H_2. Extrapolating his result to HD through considerations based on the relative internuclear distances for the two molecules combined with an average of Eqs. (82) and (83) yields the following average shielding constant for the HD molecule:

$$\begin{aligned}{}^{HD}{}_0\langle\sigma\rangle_1 &= (e^2/3mc^2)\ {}^{HD}{}_0\langle\textstyle\sum_k r_k^{-1}\rangle_1 - {}^{HD}{}_0\langle\sigma^{hf}\rangle_1 \\ &= (3.216\pm0.010)\times10^{-5} \\ &\quad -(0.594\pm0.030)\times10^{-5} \\ &= (2.622\pm0.032)\times10^{-5}. \end{aligned} \tag{84}$$

PHYSICAL REVIEW VOLUME 112, NUMBER 6 DECEMBER 15, 1958

Infrared and Optical Masers

A. L. SCHAWLOW AND C. H. TOWNES*
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received August 26, 1958)

The extension of maser techniques to the infrared and optical region is considered. It is shown that by using a resonant cavity of centimeter dimensions, having many resonant modes, maser oscillation at these wavelengths can be achieved by pumping with reasonable amounts of incoherent light. For wavelengths much shorter than those of the ultraviolet region, maser-type amplification appears to be quite impractical. Although use of a multimode cavity is suggested, a single mode may be selected by making only the end walls highly reflecting, and defining a suitably small angular aperture. Then extremely monochromatic and coherent light is produced. The design principles are illustrated by reference to a system using potassium vapor.

INTRODUCTION

AMPLIFIERS and oscillators using atomic and molecular processes, as do the various varieties of masers,[1–4] may in principle be extended far beyond the range of frequencies which have been generated electronically, and into the infrared, the optical region, or beyond. Such techniques give the attractive promise of coherent amplification at these high frequencies and of generation of very monochromatic radiation. In the infrared region in particular, the generation of reasonably intense and monochromatic radiation would allow the possibility of spectroscopy at very much higher resolution than is now possible. As one attempts to extend maser operation towards very short wavelengths, a number of new aspects and problems arise, which require a quantitative reorientation of theoretical discussions and considerable modification of the experimental techniques used. Our purpose is to discuss theoretical aspects of maser-like devices for wavelengths considerably shorter than one centimeter, to examine the short-wavelength limit for practical devices of this type, and to outline design considerations for an example of a maser oscillator for producing radiation in the infrared region. In the general discussion, roughly reasonable values of design parameters will be used. They will be justified later by more detailed examination of one particular atomic system.

* Permanent address: Columbia University, New York, New York.
[1] Gordon, Zeiger, and Townes, Phys. Rev. **99**, 1264 (1955).
[2] Combrisson, Honig, and Townes, Compt. rend. **242**, 2451 (1956).
[3] N. Bloembergen, Phys. Rev. **104**, 329 (1956).
[4] E. Allais, Compt. rend. **245**, 157 (1957).

CHARACTERISTICS OF MASERS FOR MICROWAVE FREQUENCIES

For comparison, we shall consider first the characteristics of masers operating in the normal microwave range. Here an unstable ensemble of atomic or molecular systems is introduced into a cavity which would normally have one resonant mode near the frequency which corresponds to radiative transitions of these systems. In some cases, such an ensemble may be located in a wave guide rather than in a cavity but again there would be characteristically one or a very few modes of propagation allowed by the wave guide in the frequency range of interest. The condition of oscillation for n atomic systems excited with random phase and located in a cavity of appropriate frequency may be written (see references 1 and 2)

$$n \geq hV\Delta\nu/(4\pi\mu^2 Q_c), \tag{1}$$

where n is more precisely the difference n_1-n_2 in number of systems in the upper and lower states, V is the volume of the cavity, $\Delta\nu$ is the half-width of the atomic resonance at half-maximum intensity, assuming a Lorentzian line shape, μ is the matrix element involved in the transition, and Q_c is the quality factor of the cavity.

The energy emitted by such a maser oscillator is usually in an extremely monochromatic wave, since the energy produced by stimulated emission is very much larger than that due to spontaneous emission or to the normal background of thermal radiation. The frequency range over which appreciable energy is distributed is given approximately by[1]

$$\delta\nu = 4\pi kT(\Delta\nu)^2/P, \tag{2}$$

where $\Delta\nu$ is the half-width at half-maximum of the resonant response of a single atomic system, P is the total power emitted, k is Boltzmann's constant, and T the absolute temperature of the cavity walls and wave guide. Since in all maser oscillators at microwave frequencies which have so far been considered, $P \gg kT\Delta\nu$, the radiation is largely emitted over a region very much smaller than $\Delta\nu$, or $\delta\nu \ll \Delta\nu$.

As amplifiers of microwave or radio-frequency energy, masers have the capability of very high sensitivity, approaching in the limit the possibility of detecting one or a few quanta. This corresponds to a noise temperature of $h\nu/k$, which for microwave frequencies is of the order of 1°K.

USE OF MULTIMODE CAVITIES AT HIGH FREQUENCIES

Consider now some of the modifications necessary to operate a maser at frequencies as high as that of infrared radiation. To maintain a single isolated mode in a cavity at infrared frequencies, the linear dimension of the cavity would need to be of the order of one wavelength which, at least in the higher frequency part of the infrared spectrum, would be too small to be practical. Hence, one is led to consider cavities which are large compared to a wavelength, and which may support a large number of modes within the frequency range of interest. For very short wavelengths, it is perhaps more usual to consider a plane wave reflected many times from the walls of such a cavity, rather than the field of a standing wave which would correspond to a mode.

The condition for oscillation may be obtained by requiring that the power produced by stimulated emission is as great as that lost to the cavity walls or other types of absorption. That is,

$$\left(\frac{\mu' E}{\hbar}\right)^2 \frac{h\nu n}{4\pi\Delta\nu} \geq \frac{E^2}{8\pi}\frac{V}{t}, \tag{3}$$

where μ' is the matrix element for the emissive transition, E^2 is the mean square of the electric field. (For a multiresonant cavity, E^2 may be considered identical in all parts of the cavity.) n is the excess number of atoms in the upper state over those in the lower state, V is the volume of the cavity, t is the time constant for the rate of decay of the energy, $\Delta\nu$ is the half-width of the resonance at half maximum intensity, if a Lorentzian shape is assumed. The decay time t may be written as $2\pi\nu/Q$, but is perhaps more naturally expressed in terms of the reflection coefficient α of the cavity walls.

$$t = 6V/(1-\alpha)Ac, \tag{4}$$

where A is the wall area and c the velocity of light. For a cube of dimension L, $t = L/(1-\alpha)c$. The condition for oscillation from (3) is then

$$n \geq \frac{3hV}{8\pi^2\mu^2 t}\frac{\Delta\nu}{\nu}, \tag{5}$$

or

$$n \geq \frac{\Delta\nu}{\nu}\frac{h(1-\alpha)Ac}{16\pi^2\mu^2}. \tag{6}$$

Here μ'^2 has been replaced by $\mu^2/3$, since μ'^2 is the square of the matrix element for the transition which, when averaged over all orientations of the system, is just one-third of the quantity μ^2 which is usually taken as the square of the matrix elment.

In a gas at low pressure, most infrared or optical transitions will have a width $\Delta\nu$ determined by Doppler effects. Then the resonance half-width at half-maximum intensity is

$$\Delta\nu = \frac{\nu}{c}\left(\frac{2kT}{m}\ln 2\right)^{\frac{1}{2}}, \tag{7}$$

where m is the molecular mass, k is Boltzmann's constant, and T the temperature. Because of the Gaussian line shape in this case, expression (6) becomes

$$n \geq \frac{\Delta\nu}{\nu}\frac{h(1-\alpha)Ac}{16\pi^2\mu^2(\pi\ln 2)^{\frac{1}{2}}}, \tag{8}$$

or

$$n \geq \frac{h(1-\alpha)A}{16\pi^2\mu^2}\left(\frac{2kT}{\pi m}\right)^{\frac{1}{2}}. \tag{9}$$

It may be noted that expression (9) for the number of excited systems required for oscillation is independent of the frequency. Furthermore, this number n is not impractically large. Assuming the cavity is a cube of 1 cm dimension and that $\alpha = 0.98$, $\mu = 5\times10^{-18}$ esu, $T = 400$°K, and $m = 100$ amu, one obtains $n = 5\times10^8$.

The condition for oscillation, indicated in (5), may be conveniently related to the lifetime τ of the state due to spontaneous emission of radiation by a transition between the two levels in question. This lifetime is given, by well-known theory, as

$$\tau = 3hc^3/(64\pi^4\nu^3\mu^2). \tag{10}$$

Now the rate of stimulated emission due to a single quantum in a single mode is just equal to the rate of spontaneous emission into the same single mode. Hence, $1/\tau$ is this rate multiplied by the number of modes p which are effective in producing spontaneous emission. Assuming a single quantum present in a mode at the resonant frequency, the condition for instability can then be written

$$nh\nu/p\tau \geq h\nu/t_1,$$

or

$$n \geq p\tau/t. \tag{11}$$

This gives a simple expression which may sometimes

be useful, and which is equivalent to (5), since

$$p=\int p(\nu)\frac{(\Delta\nu)^2 d\nu}{(\nu-\nu_0)^2+(\Delta\nu)^2}, \tag{12}$$

where $p(\nu)d\nu$ is the number of modes between ν and $\nu+d\nu$, which is well known to be

$$p(\nu)d\nu=8\pi\nu^2 V d\nu/c^3. \tag{13}$$

From (12) and (13), one obtains for a Lorentzian line shape,

$$p=8\pi^2\nu^2 V\Delta\nu/c^3. \tag{14}$$

Or, for a line broadened by Doppler effects, the corresponding number of effective modes is

$$p=8\pi^2\nu^2 V\Delta\nu/(\pi\ln 2)^{\frac{1}{2}}c^3. \tag{15}$$

If τ and p are inserted into (11) from expressions (10) and (14), respectively, it becomes identical with (5), as one must expect.

The minimum power which must be supplied in order to maintain n systems in excited states is

$$P=nh\nu/\tau=ph\nu/t. \tag{16}$$

This expression is independent of the lifetime or matrix element. However, if there are alternate modes of decay of each system, as by collisions or other transitions, the necessary power may be larger than that given by (16) and dependent on details of the system involved. Furthermore, some quantum of higher frequency than that emitted will normally be required to excite the system, which will increase the power somewhat above the value given by (16). Assuming the case considered above, i.e., a cube of 1-cm dimension with $\alpha=0.98$, $\lambda=10^4$ A, and broadening due to Doppler effect, (16) gives $P=0.8\times10^{-3}$ watt. Supply of this much power in a spectral line does not seem to be extremely difficult.

The power generated in the coherent oscillation of the maser may be extremely small, if the condition of instability is fulfilled in a very marginal way, and hence can be much less than the total power, which would be of the order of 10^{-3} watt, radiated spontaneously. However, if the number of excited systems exceeds the critical number appreciably, e.g., by a factor of two, then the power of stimulated radiation is given roughly by $h\nu$ times the rate at which excited systems are supplied, assuming the excitation is not lost by some process not yet considered, such as by collisions. The electromagnetic field then builds up so that the stimulated emission may be appreciably greater than the total spontaneous emission. For values even slightly above the critical number, the stimulated power is of the order of the power $nh\nu/\tau$ supplied, or hence of the order of one milliwatt under the conditions assumed above.

The most obvious and apparently most convenient method for supplying excited atoms is excitation at a higher frequency, as in optical pumping or a three-level maser system. The power supplied must, of course, be appreciably greater than the emitted power in expression (9). There is no requirement that the pumping frequency be much higher than the frequency emitted, as long as the difference in frequency is much greater than kT/h, which can assure the possibility of negative temperatures. Since, for the high frequencies required, an incoherent source of pumping power must be used, a desirable operating frequency would be near the point where the maximum number of quanta are emitted by a given transition from a discharge or some other source of high effective temperature. This maximum will occur somewhere near the maximum of the blackbody radiation at the effective temperature of such a source, or hence in the visible or ultraviolet region. The number of quanta required per second would probably be about one order of magnitude greater than the number emitted at the oscillating frequency, so that the input power required would be about ten times the output given by (16), or 10 milliwatts. This amount of energy in an individual spectroscopic line is, fortunately, obtainable in electrical discharges.

Very desirable features of a maser oscillator at infrared or optical frequencies would be a high order of monochromaticity and tunability. In the microwave range, a maser oscillator is almost inherently a very monochromatic device. However, a solid state maser can also normally be tuned over a rather large fractional variation in frequency. Both of these features are much more difficult to obtain in the infrared or optical regions. Frequencies of atomic or molecular resonances can in principle be tuned by Stark or Zeeman effects, as they would be in the radio-frequency or microwave range. However, such tuning is usually limited to a few wave numbers (or a few times 10 000 Mc/sec), which represents a large fractional change in the microwave range and only a small fractional change in the optical region. Certain optical and infrared transitions of atoms in solids are strongly affected by neighboring atoms. This may be the result of Stark effects due to internal electric fields or, as in the case of antiferromagnetic resonances, internal magnetic fields may vary enough with temperature to provide tuning over a few tens of wave numbers. Hence variation of temperature or pressure can produce some tuning. However, it appears unreasonable to expect more than a small fractional amount of tuning in an infrared or optical maser using discrete levels.

SPECTRUM OF A MASER OSCILLATOR

Monochromaticity of a maser oscillator is very closely connected with noise properties of the device as an amplifier. Consider first a maser cavity for optical or infrared frequencies which supports a single isolated mode. As in the microwave case, it is capable of

detecting, in the limit, one or a few quanta, corresponding to a noise temperature of $h\nu/k$. However, at a wavelength of 10 000 A, this noise temperature is about 14 000°K, and hence not remarkably low. Furthermore, other well-known photon detectors, such as a photoelectric tube, are capable of detecting a single quantum. At such frequencies, a maser has no great advantage over well-known techniques in detecting small numbers of quanta. It does offer the new possibility of coherent amplification. However, if many modes rather than a single one are present in the cavity, a rather large background of noise can occur, the noise temperature being proportional to the number of modes which are confused within the resonance width of the atomic or molecular system. A method for isolation of an individual mode which avoids this severe difficulty will be discussed below.

Let us examine now the extent to which the normal line width of the emission spectrum of an atomic system will be narrowed by maser action, or hence how monochromatic the emission from an infrared or optical maser would be. Considerations were given above concerning the number of excited systems required to produce stimulated power which would be as large as spontaneous emission due to all modes of a multimode cavity which lie within the resonance width of the system. Assume for the moment that a single mode can be isolated. Spontaneous emission into this mode adds waves of random phase to the electromagnetic oscillations, and hence produces a finite frequency width which may be obtained by analogy with expression (2) as

$$\Delta\nu_{\text{osc}}=(4\pi h\nu/P)(\Delta\nu)^2, \tag{17}$$

where $\Delta\nu$ is the half-width of the resonance at half-maximum intensity, and P the power in the oscillating field. Note that kT, the energy due to thermal agitation, has been replaced in expression (15) by $h\nu$, the energy in one quantum. Usually at these high frequencies, $h\nu\gg kT$, and there is essentially no "thermal" noise. There remains, however, "zero-point fluctuations" which produce random noise through spontaneous emission, or an effective temperature of $h\nu/k$.

For the case considered numerically above, $4\pi h\nu\Delta\nu/P$ is near 10^{-6} when P is given by expression (16), so that $\Delta\nu_{\text{osc}}\sim 10^{-6}\Delta\nu$. This corresponds to a remarkably monochromatic emission. However, for a multimode cavity, this very monochromatic emission is superimposed on a background of stimulated emission which has width $\Delta\nu$, and which, for the power P assumed, is of intensity equal to that of the stimulated emission. Only if the power is increased by some additional factor of about ten, or if the desired mode is separated from the large number of undesired ones, would the rather monochromatic radiation stand out clearly against the much wider frequency distribution of spontaneous emission.

Another problem of masers using multimode cavities which is perhaps not fundamental, but may involve considerable practical difficulty, is the possibility of oscillations being set up first in one mode, then in another—or perhaps of continual change of modes which would represent many sudden jumps in frequency. If the cavity dimensions, density distribution of gas and distribution of excited states remains precisely constant, it seems unlikely that oscillations will build up on more than one mode because of the usual nonlinearities which would allow the most favored mode to suppress oscillations in those which are less favored. However, if many nearby modes are present, a very small change in cavity dimensions or other characteristics may produce a shift of the oscillations from one mode to another, with a concomitant variation in frequency.

SELECTION OF MODES FOR AMPLIFICATION

We shall consider now methods which deviate from those which are obvious extensions of the microwave or radio-frequency techniques for obtaining maser action. The large number of modes at infrared or optical frequencies which are present in any cavity of reasonable size poses problems because of the large amount of spontaneous emission which they imply. The number of modes per frequency interval per unit volume cannot very well be reduced for a cavity with dimensions which are very large compared to a wavelength. However, radiation from these various modes can be almost completely isolated by using the directional properties of wave propagation when the wavelength is short compared with important dimensions of the region in which the wave is propagated.

Consider first a rectangular cavity of length D and with two square end walls of dimension L which are slightly transparent, its other surfaces being perfectly reflecting. Transparency of the end walls provides coupling to external space by a continuously distributed excitation which corresponds to the distribution of field strength at these walls. The resulting radiation produces a diffraction pattern which can be easily calculated at a large distance from the cavity, and which is effectively separated from the diffraction pattern due to any other mode of the cavity at essentially the same frequency.

The field distribution along the end wall, taken as the xy plane, may be proportional, for example, to $\sin(\pi rx/L)\cos(\pi sy/L)$. The resonant wavelength is of the form

$$\lambda=\frac{2}{[(q/D)^2+(r/L)^2+(s/L)^2]^{\frac{1}{2}}}, \tag{18}$$

where q is the number of half-wavelengths along the z direction. If L is not much smaller than D, and if $q\gg r$ or s, the resonant wavelength is approximately

$$\lambda=\frac{2D}{q}\left[1-\tfrac{1}{2}\left(\frac{Dr}{Lq}\right)^2-\tfrac{1}{2}\left(\frac{Ds}{Lq}\right)^2\right], \tag{19}$$

which is primarily dependent on q and insensitive to r or s. The direction of radiation from the end walls, however, is critically dependent on r and s. The Fraunhofer diffraction pattern of the radiation has an intensity variation in the x direction given by

$$I \propto (2\pi r)^2 \sin^2\left(\frac{\pi L \sin\theta}{\lambda}+\frac{\pi r}{2}\right)\Big/ \left(\pi r+\frac{2\pi L \sin\theta}{\lambda}\right)^2\left(\pi r-\frac{2\pi L \sin\theta}{\lambda}\right)^2, \quad (20)$$

where θ is the angle between the direction of observation and the perpendicular to the end walls. For a given value of r, the strongest diffraction maxima occur at

$$\sin\theta=\pm r\lambda/2L,$$

and the first minima on either side of the maxima at

$$\sin\theta=\pm r\lambda/2L\pm\lambda/L.$$

Thus the maximum of the radiation from a mode designated by $r+1$ falls approximately at the half-intensity point of the diffraction pattern from the mode designated by r, which is just sufficient for significant resolution of their individual beams of radiation. This provides a method for separately coupling into or out of one or a few individual modes in the multimode cavity. A practical experimental technique for selecting one or a few modes is to focus radiation from the end walls by means of a lens onto a black screen in the focal plane. A suitable small hole in the screen will accept only radiation from the desired mode or modes.

There may, of course, be more than one mode which has similar values of r and s but different values of q, and which radiate in essentially identical directions. However, the frequencies of such modes are appreciably different, and may be sufficiently separated from each other by an appropriate choice of the distance D between plates. Thus if only one mode with a particular value of r and s is wanted within the range of response $2\Delta\nu$ of the material used to produce oscillations, D should be less than $c/4\Delta\nu$. Or, if it is undesirable to adjust D precisely for a particular mode, and approximately one mode of this type is wanted in the range $2\Delta\nu$, one may choose

$$D\approx c/4\Delta\nu. \quad (21)$$

For the conditions assumed above, the value of D given by (21) has the very practical magnitude of about 10 cm.

It is desirable not only to be able to select radiation from a single mode, but also to make all but one or a few modes of the multimode cavity lossy in order to suppress oscillations in unwanted modes. This again can be done by making use of directional properties. Loss may be introduced perhaps most simply by removing the perfectly reflecting walls of the cavity. The "cavity" is then reduced to partially transparent end plates and nonexistent (or lossy) perfectly-matched side walls.

Suppose now that one of the modes of such a cavity is excited by suddenly introducing the appropriate field distribution on one of the end walls. This will radiate a wave into the cavity having directional properties such as those indicated by the diffraction pattern (20). If r and s have their minimum values, the maximum energy occurs near $\theta=0$, and the wave travels more or less straight back and forth between the two plates, except for a gradual spreading due to diffraction. If r or s are larger, the maximum energy occurs at an appreciable angle θ, and the wave packet will wander off the reflecting plates and be lost, perhaps after a number of reflections. Those modes for which θ is large are highly damped and merge into a continuum, since energy radiated into them travels immediately to the walls and is lost from the cavity. However, modes for which θ is quite small may have relatively high Q and hence be essentially discrete.

For estimates of damping, consider first two end plates of infinite extent, but excited only over a square area of dimension L by a distribution which corresponds to one of our original modes. The radiated wave will be reflected back and forth many times, gradually spreading out in the diffraction pattern indicated by (20). If a mode with small values of r and s is used, the wave undergoes reflection every time it travels a distance D, and the rate of loss of energy W is given by the equations

$$dW/dt=-c(1-\alpha)W/D, \qquad W=W_0e^{-c(1-\alpha)t/D}. \quad (22)$$

The decay time t is then $D/c(1-\alpha)$ rather than that given by (4) for the multimode case, or the effective distance traveled is $D/(1-\alpha)$.

Since the wavelength for modes with small r and s is given by (19), the frequency separation between modes with successive values of q is given by the usual Fabry-Perot condition

$$\delta\nu=c/2D. \quad (23)$$

Thus $\delta\nu\gg 1/t$ and the modes with successive values of q are discrete if $1-\alpha$, the loss on reflection, is much less than unity. On the other hand, the various modes given by small values of r or s and the same value of q are nearly degenerate, since according to (19) their frequency difference is less than $\delta\nu$ given in (23) by the factor r/q, which is of the order 10^{-4} for a typical case. These modes must be separated purely by their directional properties, rather than by their differences in frequency.

After traveling a distance $D/(1-\alpha)$, the radiation resulting from the excitation discussed above will have moved sideways in the x direction along the infinite parallel plates a distance of approximately $D\theta/(1-\alpha)$,

where θ is the angle of one of the two large diffraction maxima given by (20). This distance is then

$$x=D\lambda r/[2(1-\alpha)L]. \tag{24}$$

Consider now the case of finite end plates of dimension L without their infinite extension which was assumed immediately above. After a number of reflections, the diffraction pattern would no longer be precisely that given by (20). However, expression (24) would still give a reasonable approximation to the distance of sideways motion, and if this distance is larger than the end-wall dimension L, the radiation will have been lost to the cavity, and the decay time for the mode in question is appreciably shorter than that indicated by (22). This condition occurs when

$$D\lambda r/[2(1-\alpha)L]\gtrsim L, \quad \text{or} \quad r\gtrsim 2(1-\alpha)L^2/D\lambda. \tag{25}$$

Thus to damp out modes with $r\gtrsim 10$, when $L=\frac{1}{2}$ cm, $\alpha=0.98$, and $\lambda=10^{-4}$ cm, the separation D between plates needs to be as large as a few centimeters. By choosing L sufficiently small, it is possible to discriminate by such losses between the lowest mode ($r=1$), and any higher modes. Too small a ratio $L^2/D\lambda$ will, however, begin to appreciably add to the losses from the lowest mode, and hence is undesirable if the longest possible delay times are needed.

The precise distribution of radiation intensity in the plane of the end walls which will give minimum loss, or which will occur during maser oscillation, cannot be very easily evaluated. It must, however, be somewhat like the lowest mode, $r=1$, $s=0$. A normal and straightforward method for exciting a Fabry-Perot interferometer is to use a plane wave moving perpendicular to the reflecting plates, and screened so that it illuminates uniformly all but the edge of the plates. Such a distribution may be expressed in terms of the nearly degenerate modes of the "cavity" with various r and s, and the considerable majority of its energy will be found in the lowest mode $r=1$, $s=0$, if it is polarized in the y direction. There is, of course, an exactly degenerate mode of the same type which is polarized in the x direction. Any much more complicated distribution than some approximation to uniform illumination or to the lowest mode $r=1$, $s=0$ of our rectangular cavity will produce a wider diffraction pattern which would be lost to a detector arranged to accept a very small angle θ near zero, and which would also be subject to greater losses when $L^2/D\lambda$ is small. However, nonuniform distribution of excited atoms between the reflecting plates could compensate for the larger diffraction losses, and in some cases induce oscillations with rather complex distributions of energy.

The above discussion in terms of modes of a rectangular cavity illustrates relations between the arrangement using a Fabry-Perot interferometer and the usual microwave resonant cavity.† An alternative approach which uses the approximation of geometrical optics more directly may also be helpful and clarifying. An atom radiating spontaneously in any direction has a decay time τ given by expression (10). The probability per unit time of emission of a quantum within a given solid angle $\Delta\Omega$ is then

$$\frac{1}{t'}=\frac{16\pi^3\nu^3}{3hc^3}\mu^2\Delta\Omega. \tag{26}$$

Hence if a sufficiently small solid angle is selected from the radiation, the amount of spontaneous emission can be made arbitrarily small. However, if essentially all the stimulated emission emitted from the end-wall of the interferometer is to be collected in a receiver or detector, allowance must be made for diffraction and a solid angle as large as about $(\lambda/L)^2$ must be used, so that the rate of spontaneous emission into the detector is

$$\frac{1}{t'}=\frac{16\pi^3\mu^2\nu}{3hcL^2}. \tag{27}$$

The rate of spontaneous emission (27) within the diffraction angle may be compared with the rate of induced transitions produced by one photon reflected back and forth in the volume L^2D. This rate is, as in (3), $(\mu' E/\hbar)^2(4\pi\Delta\nu)^{-1}$, where $E^2L^2D/8\pi=h\nu$. That is, since $\mu'^2=\mu^2/3$,

$$\frac{1}{t''}=\frac{8\pi^2\mu^2\nu}{3\Delta\nu L^2Dh}. \tag{28}$$

If D is $c/4\Delta\nu$ as in expression (21), so that there is approximately one and only one interference maximum of the interferometer in a particular direction within the range $2\Delta\nu$ of emission, then (28) becomes

$$\frac{1}{t''}=\frac{32\pi^2\mu^2\nu}{3hcL^2}. \tag{29}$$

Except for a small numerical factor of the order of the accuracy of the approximations used here, $1/t''$ given by (29) may be seen to equal $1/t'$. That is, use of the limiting amount of directional selection reduces the background of spontaneous emission to the same rate as that of stimulated emission due to a single photon. This is similar to the situation of a single mode in a cavity at microwave frequencies. It affords the limit of sensitivity which can be obtained by the usual maser amplifier, and the smallest possible noise for such a system as an oscillator.

† *Note added in proof.*—Use of two parallel plates for a maser operating at short wavelengths has also recently been suggested by A. M. Prokhorov [JETP 34, 1658 (1958)] and by R. H. Dicke [U. S. Patent 2,851,652 (September 9, 1958)]. These sources do not, however, discuss the reduction of excess modes or spontaneous emission.

Consider now the rate of loss of energy from a beam being reflected back and forth between the two end plates in the approximation of geometric optics. If the angle of deviation from the direction perpendicular to the plates is θ, then the additional rate of energy loss from a plane wave due to its spilling off the edges of the reflecting surfaces is

$$dW'/dt = -c\theta W/L. \tag{30}$$

Expression (30) assumes, to be precise, that the deviation is parallel to one edge of the end plate. Thus, when $\theta=(1-\alpha)L/D$, the decay time is one-half that for $\theta=0$. Because of nonlinearities when oscillations set in, it may be seen from expression (3) that only those modes with the largest decay times will fulfill the condition for oscillation. The fraction ϵ of all modes of the "cavity" which have decay times greater than one-half that of the maximum decay time is approximately $(2\theta)^2/2\pi$, or

$$\epsilon=2(1-\alpha)^2L^2/\pi D^2. \tag{31}$$

Letting $(1-\alpha)=1/20$, $L=1$ cm, and $D=10$ cm, one obtains $\epsilon=1.6\times10^{-5}$. This enormously reduces the number of modes which are likely to produce oscillations. Since the total number of modes is, from (14), $(8\pi^2\nu^2\Delta\nu/c^3)L^2D$, this number which may produce oscillations is

$$p'=\frac{16\pi(1-\alpha)^2\nu^2\Delta\nu L^4}{c^3D}. \tag{32}$$

Under the assumptions used above, p' may be found to be approximately 10^5, which is very much smaller than the total number of modes in the multimode cavity, but still may be an inconveniently large number. By using limiting values of the solid angle $(2\theta)^2$ set by diffraction, the number of modes can be further reduced to approximately unity, as was seen above.

FURTHER DISCUSSION OF PROPERTIES OF MASERS USING LARGE DIMENSIONS

It is important to notice that in the parallel plate case a very large amount of spontaneous emission may be radiating in all directions, even though only the very small amount indicated above is accepted in the detector, or is confused with the amplified wave. This property is quite different from the normal case in the microwave or radio-frequency range, and requires a rather rapid rate of supply of excited systems in order to maintain enough for maser action. Furthermore, great care must be taken to avoid scattering of light from undesired modes into the one which is desired. The fraction of spontaneous light which is scattered into the detector must be typically as low as about 10^{-6} or 10^{-7} in order to approach genuine isolation of a single mode.

Admission of a signal into the region between the two parallel plates is very similar to the process involved in a microwave cavity. The partially reflecting surfaces are analogous to coupling holes. If a monochromatic plane wave strikes the outside surface of one of the partially reflecting planes, energy will build up with the region between the planes, and the relations between input wave, energy in the "cavity", and output waves are just analogous to those for a microwave impinging on an appropriate cavity with input and output coupling holes.

Another interesting property of optical or infrared maser action which is associated with directional selection is that a beam of light may be passed through an ensemble of excited states with resulting amplification, but no important change in the wave front or phase. This amplification is just the inverse of an absorption, where it is well known that the wave front and phase are not distorted. Suppose, for example, that parallel light is focused by a lens. If an amplifying medium of excited gas is interposed between the lens and its focal point, the image will be intensified, but not otherwise changed except for some more or less normal effects which may be attributed to the dielectric constant of the gas. The same situation can, in principle, occur for maser amplification of microwaves. However, at these lower frequencies the amplification per unit length is usually so small that an impractically large volume of excited material would be required for amplification of a wave in free space to be evident. There may be a considerable amount of spontaneous emission in all directions, but only a very small fraction of the total spontaneous emission will fall at the focal point of the lens and be superimposed as noise on the intensified image. Noise from spontaneous emission decreases, for example, with the inverse square of distance from the emitting material, whereas the intensity of the focused beam increases as one approaches the focal point.

A SPECIFIC EXAMPLE

As an example of a particular system for an infrared maser, let us consider potassium. Atomic potassium is easily vaporized and has a simple spectrum as indicated by the energy levels shown in Fig. 1. Absorption transitions can occur from the $4s\ ^2S_{\frac{1}{2}}$ ground state only to the various p levels. In particular, the atoms can be excited to the $5p\ ^2P_{\frac{1}{2},\frac{3}{2}}$ by radiation of wavelength 4047 A. Just the right exciting frequency can be obtained from another potassium lamp, whose light is filtered to remove the red radiation at 7700 A. These excited atoms will decay to the $5s$ or $3d$ states in about 2×10^{-7} sec, or more slowly to the $4s$ ground state. However, if excited atoms are supplied fast enough, a sizable population can be maintained in the $5p$ state.

The minimum number of excited atoms required for maser-type oscillation may be found from (6), if the dipole matrix element were known, or from (11) if the

lifetime were known. Although the wave functions necessary for obtaining the matrix elements have been calculated,[5,6] only estimates of the matrix element or lifetime can be made at present. The rate at which atoms must be supplied may, however, be obtained without detailed knowledge of the matrix elements by a small modification of expression (11). If τ is the mean life for spontaneous radiation of the desired wavelength, and φ is the fraction of the all decay processes from the upper level which occur in this manner, then the actual mean life in the excited state is $\tau'=\varphi\tau$. The number of atoms needed per second can be obtained from (11) as

$$\frac{n}{\tau'}=\frac{p}{\varphi t}=\frac{8\pi^2V}{(\pi\ln 2)^{\frac{1}{2}}\varphi t\lambda^3}\frac{\Delta\nu}{\nu}=8\pi(2\pi)^{\frac{1}{2}}\frac{V}{\varphi\lambda^3 ct}\left(\frac{kT}{m}\right)^{\frac{1}{2}}, \quad (33)$$

where λ is the wavelength. Thus, if the fraction φ is known, no other detailed properties of the atomic transition are required to evaluate the rate at which excited atoms must be supplied.

For the particular case of a gas (such as potassium vapor) at sufficiently low pressure that collisions are not too frequent, we can obtain φ from the relative intensities of the various radiative transitions out of the excited state. Observed relative intensities[7] show that for potassium the $5p\rightarrow 3d$ transitions are about 4 times more intense than the $5p\rightarrow 5s$ transitions. Within the $5p\rightarrow 3d$ transitions it follows from elementary angular momentum theory that the $5p_{\frac{3}{2}}\rightarrow 3d_{\frac{5}{2}}$ is the most intense, accounting for 9/15 of the radiation

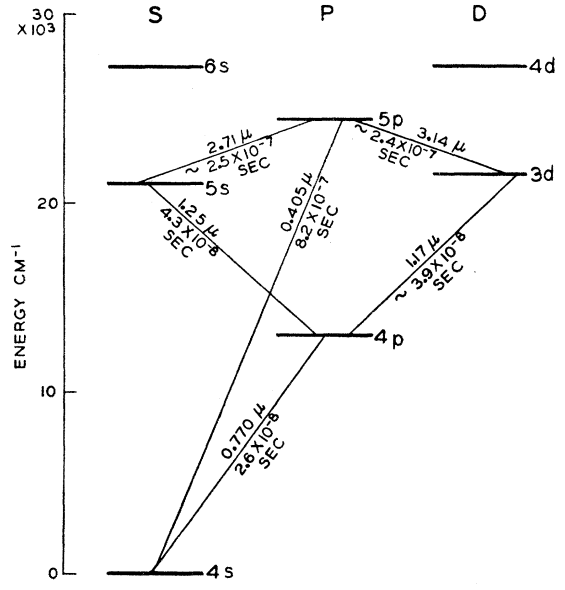

FIG. 1. Low-lying energy levels and transitions of atomic potassium.

emitted. Using the observed intensity ratio to allow for transitions to the $5s$ level, we conclude that about $9/18=\frac{1}{2}$ of those atoms excited to the $5p_{\frac{3}{2}}$ level decay to the $3d_{\frac{5}{2}}$ level. Decay to the $4s$ ground state is almost certainly less likely since we do know that this matrix element is not very large ($f=0.010$,[5] so that $\mu=0.65\times10^{-18}$ esu). Thus $\varphi=\frac{1}{2}$ for the transition $5p_{\frac{3}{2}}-3d_{\frac{5}{2}}$ at 31 391 A.

Assume now two parallel plates of area 1 cm^2 and 10 cm apart, having a reflectivity α of 0.98. The decay time t for radiation in the space between the plates is $(10/3\times10^{10})\times50$ sec and $V=10$ cm^3. For potassium vapor of suitable pressure, $T=435$°K and $m=39$ amu. Hence, from (29), the number of excited atoms needed per second is $dn/dt\geq2.5\times10^{15}$.

The energy needed per second is $d/dt(nh\nu)$, where ν is the frequency of the exciting radiation. Its value is 1.2×10^{-3} watt. This energy requirement is quite attainable. Incomplete absorption of the exciting radiation, reflection losses and multiplicity of the atomic states might raise this requirement somewhat. The absorption of the existing radiation is easily calculable and can be adjusted by controlling the density of the vapor:

$$k_0=\frac{1}{\Delta\nu_D}\left(\frac{\ln 2}{\pi}\right)^{\frac{1}{2}}\frac{\pi e^2}{mc}Nf, \quad (34)$$

where k_0 is the absorption coefficient at the peak of the line $\Delta\nu_D$ is the (Doppler) line half-width, e is the electron charge, m is the electron mass, c is the velocity of light, N is the number of initial state atoms per cc, and f is the oscillator strength of the transition; i.e.,

$$k_0=1.25\times10^{-2}(Nf/\Delta\nu_D).$$

For the exciting transition, 4046 A, in potassium, $\nu_0=7.42\times10^{14}$ cycles/sec and at 435°K, $\Delta\nu_D=0.84\times10^9$ cycles/sec. At this temperature the vapor pressure is 10^{-3} mm of mercury, so that in saturated vapor $N=2.5\times10^{13}$/cc. Since $f=0.10$ for the $4s_{\frac{1}{2}}\rightarrow5p_{\frac{3}{2}}$ transition,[6] $k_0=3.72$. This is high enough that the exciting radiation would be absorbed in a thin layer; if necessary it can be reduced by changing the pressure or temperature.

The light power for excitation is proportional to

$$\frac{V}{t}=\frac{AD}{(1-\alpha D/c)}=\frac{Ac}{1-\alpha}, \quad (35)$$

where V is the volume of the cavity, t is the decay time for light in the cavity, D is the length of the cavity, A is the cross-section area of the cavity, α is the reflectivity of the end plates, and c is the velocity of light. This is independent of length, so that for a given cross-sectional area the light density needed can be reduced by increasing the length.

[5] L. Biermann and K. Lubeck, Z. Astrophys. **25**, 325 (1948).
[6] D. S. Villars, J. Opt. Soc. Am. **42**, 552 (1952).
[7] *Viz.*, Tabulation in the *Handbook of Chemistry and Physics*, edited by D. Hodgman (Chemical Rubber Publishing Company, Cleveland, 1957), thirty-ninth edition.

LIGHT SOURCES FOR EXCITATION

A small commercial potassium lamp (Osram) was operated with an input of 15 watts, 60 cycles, and its output was measured. In the red lines (7664–7699 A), the total light output was 28 mw from about 5 cc volume. At the same time, the total output in the violet lines (4044–4047 A) was 0.12 mw,[8] so that the output in $4s-5p_{\frac{3}{2}}$ was 0.08 mw. By increasing the current from 1.5 to 6 amp, with forced air cooling (the outer jacket being removed), the total violet output was increased to 0.6 mw. These outputs are somewhat short of the power level needed, but they may be considerably increased by adjusting discharge conditions to favor production of the violet line, and by using microwave excitation. With a long maser cell, the lamp area can be greatly increased. If necessary, very high peak light powers could be obtained in pulsed operation, although one would have to be careful not to broaden the line excessively.

Another possibility for excitation is to find an accidental coincidence with a strong line of some other element. The $8p$ level of cesium is an example of this type, since it can be excited very well by a helium line. The 4047 A line of mercury is 5 cm^{-1} from the potassium line, and is probably too far away to be useful even when pressure broadened and shifted.

Different modes correspond to different directions of propagation, and we only want to produce one or a few modes. Thus the cavity need only have two good reflecting walls opposite each other. The side walls need not reflect at all, nor do they need to transmit infrared radiation.

Unfortunately, most elements which have simple spectra, are quite reactive. Sapphire has good chemical inertness and excellent infrared transmission, being almost completely transparent as far as about 4 microns wavelength.[9] With such good transmission, the principal reflecting surfaces can be put outside the cell, and hence chosen for good reflectivity without regard to chemical inertness. Thus, one could use gold which has less than 2% absorption in this region, and attain a reflectivity of ~97% with 1% transmission. Even better reflectivity might be obtained with multiple dielectric layers of alternately high- and low-dielectric constant. The inner walls of the sapphire cell would reflect about 5% of the infrared light, and the thickness should be chosen so that the reflections from the two surfaces are in phase. The phase angle between reflections from the two surfaces depends on the thickness and the refractive index. Since sapphire is crystalline and the index is different for ordinary and extraordinary rays, the thickness could be chosen to give constructive interference for one polarization, and destructive interference for the perpendicular polarization. Thus, one could discriminate, if desired, between modes traveling in the same direction with different polarization.

To select just one from among the very many modes possible within the line width, the stimulated emission of radiation with one chosen direction of propagation must be favored. Thus the cell should be made long in the desired direction and fitted with highly reflecting end plates. The desired wave then has a long path as it travels back and forth, and so has a good chance to pick up energy from the excited atoms. A large width decreases the angular discrimination, and increases the pumping power needed.

For the potassium radiation at 3.14×10^{-4} cm wavelength, and $\Delta\nu$ being the Doppler width at 435°K, i.e., $\Delta\nu/\nu_0=1.2\times10^{-6}$, the number of modes is $2.0\times10^6\,V$ from expression (15). If we consider a cavity 1 cm square by 10 cm long, this number is 2.0×10^7, or 3.2×10^6 modes per steradian (forward and backward directions are taken as equivalent for standing waves). The angular separation between modes is then $(32\times10^6/2)^{-\frac{1}{2}}=2.5\times10^{-4}$ radian, where the 2 in the denominator removes the polarization degeneracy. The angular aperture accepted by this cavity is 1/10, but if the end plates had 98% reflectivity, the effective length would be increased by a factor of 50, and the angular aperture reduced to 2×10^{-3} radian. Thus there would be only 8 modes of each polarization within the effective aperture of the cell. Obviously this type of mode selection could be pushed further by making the cavity longer or narrower or more reflecting but this should not be necessary. Furthermore, the emission line does not have constant intensity over the width $\Delta\nu$, and the mode nearest the center frequency would be the first to oscillate at the threshold of emission.

SOLID-STATE DEVICES

There are a good many crystals, notably rare earth salts, which have spectra with sharp absorption lines, some of them having appeared also in fluorescence. In a solid, a concentration of atoms as large as 10^{19} per cc may be obtained. The oscillator strengths of the sharp lines are characteristically low, perhaps 10^{-6}. If the f value is low, radiative lifetimes are long, and in some cases lifetimes are as long as 10^{-3} sec or even more.

If the lifetime is primarily governed by radiation in the desired line, the pumping power required for the onset of stimulated oscillation is independent of the f value, as was shown above. For the atomic potassium level considered earlier, there are several alternative radiative decay paths (to the $4s$ and $3d$ states). In a solid there may also be rapid decay by nonradiative processes. If the storage time is long, because of a small f value, there is more time for competing processes to occur. Even lines which are sharp for solids are likely to be broader than those obtainable in gases. This larger width makes the attainment of maser oscillation more difficult, and it adds greatly to the difficulty of

[8] We are indebted to R. J. Collins for making these measurements.

[9] R. W. Kebler, *Optical Properties of Synthetic Sapphire* (Linde Company, New York).

selecting a single mode. However, there may very well be suitable transitions among the very many compounds.

The problem of populating the upper state does not have as obvious a solution in the solid case as in the gas. Lamps do not exist which give just the right radiation for pumping. However, there may be even more elegant solutions. Thus it may be feasible to pump to a state above one which is metastable. Atoms will then decay to the metastable state (possibly by nonradiative processes involving the crystal lattice) and accumulate until there are enough for maser action. This kind of accumulation is most likely to occur when there is a substantial empty gap below the excited level.

SUMMARY AND HIGH-FREQUENCY LIMITS

The prospect is favorable for masers which produce oscillations in the infrared or optical regions. However, operation of this type of device at frequencies which are still very much higher seems difficult. It does not appear practical to surround an atomic system with cavity walls which would very much affect its rate of spontaneous emission at very short wavelengths. Hence any ensemble of excited systems which is capable of producing coherent amplification at very high frequencies must also be expected to emit the usual amount of spontaneous emission. The power in this spontaneous emission, from expressions (14) and (16), increases very rapidly with frequency—as ν^4 if the width $\Delta\nu$ is due to Doppler effects, or as ν^6 if the width is produced by spontaneous emission. By choice of small matrix elements, $\Delta\nu$ can, in principle, be limited to that associated with Doppler effects, but the increase in spontaneously emitted power as fast as ν^4 is unavoidable.

For a wavelength $\lambda = 10^4$ A, it was seen above that spontaneous emission produced a few milliwatts of power in a maser system of dimensions near one centimeter, assuming reflectivities which seem attainable at this wavelength. Thus in the ultraviolet region at $\lambda = 1000$ A, one may expect spontaneous emissions of intensities near ten watts. This is so large that supply of this much power by excitation in some other spectral line becomes very difficult. Another decrease of a factor of 10 in λ would bring the spontaneous emission to the clearly prohibitive value of 100 kilowatts. These figures show that maser systems can be expected to operate successfully in the infrared, optical, and perhaps in the ultraviolet regions, but that, unless some radically new approach is found, they cannot be pushed to wavelengths much shorter than those in the ultraviolet region.

For reasonably favorable maser design in the short wavelength regions, highly reflecting surfaces and means of efficient focusing of radiation must be used. If good reflecting surfaces are not available, the number of excited systems used must, from (6), be very much increased with a resulting increase in spontaneous emission and difficulty in supply of excited systems. If focusing is not possible, the directional selection of radiation can in principle be achieved by detection at a sufficiently large distance from the parallel plates. However, without focusing the directional selection is much more difficult, and the background of spontaneous emission may give serious interference as noise superimposed on the desired radiation.

Finally, it must be emphasized that, as masers are pushed to higher frequencies, the fractional range of tunability must be expected to decrease more or less inversely as the frequency. The absolute range of variation can be at least as large as the width of an individual spectral line, or as the few wave numbers shift which can be obtained by Zeeman effects. However, continuous tuning over larger ranges of frequency will require materials with very special properties.

ACKNOWLEDGMENTS

The authors wish to thank W. S. Boyle, M. Peter, A. M. Clogston, and R. J. Collins for several stimulating discussions.