



2-20-1994

# On the Theory of Astronomical Masers in Three Dimensions

Moshe Elitzur

*University of Kentucky*, [moshe@pa.uky.edu](mailto:moshe@pa.uky.edu)

**[Click here to let us know how access to this document benefits you.](#)**

Follow this and additional works at: [https://uknowledge.uky.edu/physastron\\_facpub](https://uknowledge.uky.edu/physastron_facpub)

 Part of the [Astrophysics and Astronomy Commons](#), and the [Physics Commons](#)

---

## Repository Citation

Elitzur, Moshe, "On the Theory of Astronomical Masers in Three Dimensions" (1994). *Physics and Astronomy Faculty Publications*. 226.  
[https://uknowledge.uky.edu/physastron\\_facpub/226](https://uknowledge.uky.edu/physastron_facpub/226)

This Article is brought to you for free and open access by the Physics and Astronomy at UKnowledge. It has been accepted for inclusion in Physics and Astronomy Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact [UKnowledge@lsv.uky.edu](mailto:UKnowledge@lsv.uky.edu).

---

**On the Theory of Astronomical Masers in Three Dimensions**

**Notes/Citation Information**

Published in *The Astrophysical Journal*, v. 422, no. 2, p. 751-764.

© 1994. The American Astronomical Society. All rights reserved.

The copyright holder has granted permission for posting the article here.

**Digital Object Identifier (DOI)**

<http://dx.doi.org/10.1086/173768>

## ON THE THEORY OF ASTRONOMICAL MASERS IN THREE DIMENSIONS

MOSHE ELITZUR

Department of Physics and Astronomy, University of Kentucky, Lexington, KY 40506

Received 1993 February 8; accepted 1993 August 17

## ABSTRACT

In the standard theory of three-dimensional astronomical masers, the radiation field is described as if the source were comprised of a collection of linear masers, an approximation that has been justified by the highly beamed nature of the radiation. Recently, Neufeld has noted potential difficulties with this description and has supplied the general expressions for the maser problem without assuming beaming at the outset. The consequences of these general expressions, which have been formulated already in 1973 by Litvak and in 1974 by Bettwieser & Kegel, are analyzed here. To leading order, the standard theory is shown to provide the correct description of three-dimensional masers and its results remain intact, but only within a frequency core whose half-width is  $x_s \Delta v_D$ , where  $\Delta v_D$  is the Doppler width and  $x_s$  is a dimensionless parameter. For any given geometry,  $x_s$  is  $\sim 1/\theta_{\text{sat}}$ , where  $\theta_{\text{sat}}$  is the beaming angle of a maser with that geometry that has just saturated. For typical pumping schemes,  $x_s$  is  $\sim 2$  in spherical masers,  $\sim 2.5$ – $3$  in disk masers, and  $\sim 3$ – $5$  in cylindrical masers. For frequencies outside this core region, maser operation corresponds to a mode that will be called *suppressed*, and the standard theory breaks down. In this frequency domain, interaction with core rays that are slightly slanted to the direction of propagation suppresses photon production. In contrast with the core region, in the suppressed regime the rate of maser photon generation never reaches the maximum allowed by the pump processes; this regime effectively corresponds to a maser whose inherent strength is weaker than that of a linear maser whose properties are otherwise identical. Observed maser radiation is effectively confined to the core region since frequencies in a suppressed domain are practically unobservable. In essence,  $x_s$  provides an effective cutoff, defining a width at zero intensity that depends on the geometry but is unaffected by growth at line center. In practice, suppression affects only extreme maser outbursts. Their profiles change in such a way that when fitted with a Gaussian, the linewidth decreases when the line center intensity increases, even for masers that are saturated at the line core—in marked contrast with the predictions of standard analysis of maser linewidths. This behavior could perhaps be related to the inverse relationship between intensity and line width displayed in some intense  $\text{H}_2\text{O}$  maser flares in star forming regions.

*Subject headings:* masers — radiative transfer

## 1. INTRODUCTION

Because the radiation of strong masers is highly beamed, the linear maser has served as the starting point for the formulation of maser theory in three dimensions. At every point in a maser the intensity is highest along the ray that traveled the longest distance through the source, the local dominant ray. Since this ray dominates the emergent intensity, it is the one relevant for observations and the radiation field is envisioned as a collection of linear masers aligned with the dominant rays. Building upon a series of earlier studies, the final formulation of what has become the standard theory of astronomical masers in three dimensions was provided by Goldreich & Kwan (1974). In this formulation, the equation of radiative transfer is

$$\frac{dI_\nu}{dl} = \kappa_\nu I_\nu. \quad (1.1)$$

Here  $\kappa_\nu$  is the maser absorption coefficient, describing the interaction of the maser radiation with the inverted population. The correspondence between frequency and particle velocity, essential for the calculation of  $\kappa_\nu$ , requires a choice of axis for the velocity projection in the Doppler relation. Since radiative transfer is only solved along dominant rays, the frequency  $\nu$  in this equation corresponds to the velocity component along the dominant ray along which it is solved;

transverse velocities are integrated. The standard maser absorption coefficient is

$$\kappa_{\nu,\text{std}} = \frac{\kappa_{0\nu}}{1 + J_\nu/J_s}, \quad (1.2)$$

where  $\kappa_{0\nu}$  is the unsaturated absorption coefficient (determined in the absence of maser radiative interactions),  $J_\nu = \int I_\nu d\Omega/4\pi$  is the angle-averaged intensity and  $J_s$  is the saturation intensity; specifically,  $J_s = \Gamma/2B$ , where  $\Gamma$  is the loss rate of the maser system and  $B$  is the Einstein  $B$ -coefficient of the maser transition. An important property of this result is that different frequencies are completely uncoupled. The absorption coefficient at frequency  $\nu$  depends on quantities associated with that frequency, and that frequency only. As a result, the spatial and frequency properties of the radiative transfer problem are decoupled and can be determined independent of each other. This becomes evident upon inserting the absorption coefficient into the radiative transfer equation. Switching to the variable  $d\tau_{0\nu} = \kappa_{0\nu} dl$ , the frequency becomes merely an index of the resulting equation. A solution in terms of the variable  $\tau_{0\nu}$  provides the maser structure at all frequencies. The spatial structure at every frequency can be obtained from this solution with the aid of a simple transformation in which  $\kappa_{0\nu}$  provides a length scale for frequency  $\nu$ ; in particular, the overall maser size in  $\tau$ -space decreases with frequency shift from line center.

The decoupling of different frequencies reflects the one-to-one correspondence between frequency and velocity that forms the basis for the standard theory expression for the absorption coefficient, equation (1.2). This equation is exact for linear masers and can be expected to provide an adequate approximation along the central ray of a tightly focused beam, the situation in astronomical masers. While this seems like a justified approximation as long as the equation is restricted to dominant rays, a calculation of the intensity of a saturated maser always requires an estimate of the beaming angle, and this requires handling of rays slanted to the dominant ray. The technique developed for this purpose by Goldreich & Keeley (1972; see also Elitzur 1990) involves angular integration of the radiative transfer equation to produce the flux divergence relation

$$\nabla \cdot \mathbf{F}_\nu = \int \kappa_\nu I_\nu d\Omega. \quad (1.3)$$

Although the integration on the right-hand side is effectively confined to the beaming solid angle, it still involves the magnitude of  $\kappa_\nu$  for rays other than the dominant ray. The implicit assumption made in the standard theory is that  $\kappa_\nu$  does not vary appreciably inside the beaming cone and thus can be pulled out of the integral and assigned the value prescribed above for the dominant ray. The remaining angular integration simply becomes  $4\pi J_\nu$ , which for beamed radiation is equal to  $F_\nu$  ( $=|\mathbf{F}_\nu|$ ). The flux divergence relation can thus be solved for  $F_\nu$  independently of the radiative transfer equation, and the beaming angle  $\Omega_\nu$  is determined from the ratio  $F_\nu/I_\nu$ .

The discussion illustrates a fundamental shortcoming of the standard theory: Its basic expressions have not been derived from a more general theory in the limit of narrow beaming angles. Rather, they have been borrowed from the linear maser theory, assuming at the outset infinitely sharp beaming angles. It is impossible to verify the adequacy of the approximations involved and estimate corrections due to small but finite beaming angles because the appropriate expressions are simply not available. For example, in the flux calculation the value of  $\kappa_\nu$ , defined along a dominant ray from the analogy with linear masers is utilized inside the entire beam pattern. Whether that is a justified approximation is a question that cannot even be addressed. Unfortunately, although these shortcomings have been addressed years ago by Litvak (1973) and by Bettwieser & Kegel (1974), who formulated the general equations that describe the maser system without reliance on the analogy with linear masers, they have been largely ignored. Recently, Neufeld (1992) has noticed these problems again and has independently reformulated the general equations. The aim of the present paper is to assess the procedures developed within the standard maser theory, gauge the adequacy of the approximations involved and establish a domain of applicability from this general formalism.

## 2. BASIC FORMALISM

For the sake of completeness and to facilitate comparison with standard procedures, the formalism described by Litvak (1973), Bettwieser & Kegel (1974), and Neufeld (1992) is reproduced here with notations that correspond more closely to those employed in the standard theory. The required equations should describe the coupled behavior of the maser radiation field and the particles that produce it. The studies of Litvak (1970) and Goldreich, Keeley, & Kwan (1974) show that

because of the large dimensions and broad bandwidths of astronomical masers, a full description in terms of field amplitudes is not necessary. The radiation field can be adequately treated in terms of intensities (squared amplitudes) and is properly described by the distribution  $I_\nu(\mathbf{k})$ , the intensity of radiation at frequency  $\nu$  traveling with wavevector  $\mathbf{k} = k\hat{\mathbf{k}} = (\nu/c)\hat{\mathbf{k}}$ . Since the full dependence on  $\mathbf{k}$  is explicitly displayed, the index  $\nu$  is actually redundant; it is kept for the sake of conformity with standard notations.  $I_\nu(\mathbf{k})$  is related to the more commonly used  $\nu$ -distribution  $I(\nu, \Omega_\mathbf{k})$ , where  $\Omega_\mathbf{k}$  characterizes the direction of  $\mathbf{k}$ , through

$$I_\nu(\mathbf{k}) = \frac{c^3}{\nu^2} I(\nu, \Omega_\mathbf{k}). \quad (2.1)$$

The radiation interacts with particles that move with velocities  $\mathbf{v}$  and occupy two levels separated by energy  $h\nu_0$ . The level populations are characterized by  $n_i(\mathbf{v})$  ( $i = 1, 2$ ) and a population difference  $\Delta n(\mathbf{v}) = n_2(\mathbf{v}) - n_1(\mathbf{v})$  ( $> 0$ ). For simplicity, the levels are assumed scalar.

### 2.1. Radiative Transfer

Since radiation at frequency  $\nu$  can interact with particles of any velocity moving in any direction, radiative transfer is described by

$$\frac{dI_\nu(\mathbf{k})}{dl} = \int \kappa(\mathbf{v}) I_\nu(\mathbf{k}) \delta(\nu_0 - \nu + \mathbf{v} \cdot \mathbf{k}) d^3\mathbf{v}, \quad (2.2)$$

where  $l$  measures distances along a path aligned with  $\mathbf{k}$  and where

$$\kappa(\mathbf{v}) = \frac{h\nu_0 B}{4\pi} \Delta n(\mathbf{v}). \quad (2.3)$$

This is the most general form for the transport of maser radiation. It assumes that particle interactions produce radiation with the same frequency and wave vector as the incoming radiation—as appropriate for stimulated emission; maser photons with wavevector  $\mathbf{k}$  are produced only from photons with the same wavevector. The  $\delta$ -function ensures that interaction only occurs for photons in perfect tune with the transition frequency  $\nu_0$  in a particle's rest frame, and its use is justified because the natural line width is so small. Finite width effects can be incorporated by replacing the  $\delta$ -function with a Lorentzian profile.

Because  $I_\nu(\mathbf{k})$  is independent of the integration variable in the radiative transfer equation, it can be pulled out of the integral. The remaining integral suggests that we introduce the quantity

$$n_{i\nu}(\mathbf{k}) = \int n_i(\mathbf{v}) \delta(\nu_0 - \nu + \mathbf{v} \cdot \mathbf{k}) d^3\mathbf{v}, \quad (2.4)$$

the density of particles in level  $i$  that can interact with radiation characterized by  $\nu$  and  $\mathbf{k}$ . It is important to note that for any given  $\mathbf{k}$ ,  $n_{i\nu}(\mathbf{k})$  is not associated with any particular velocity  $\mathbf{v}$  and contains particles that move with all velocities. The operation prescribed by this equation is a projection from  $\mathbf{v}$ -space onto  $\mathbf{k}$ -space. Its meaning is better understood by integrating with the aid of a Cartesian coordinate system whose  $z$ -axis is aligned with  $\mathbf{k}$ , since the  $v_z$ -integration of the  $\delta$ -function is then immediate. Every wavevector  $\mathbf{k}$  defines an associated velocity

subspace through Doppler matching<sup>1</sup>

$$\mathbf{v}_k = \left( v_x, v_y, v_z = c \frac{v - v_0}{v_0} \right) \quad (2.5)$$

so that to leading order in  $v/c$ ,

$$n_{i\nu}(\mathbf{k}) = \frac{c}{v_0} \int n_i(\mathbf{v}_k) dv_x dv_y. \quad (2.6)$$

That is,  $n_{i\nu}(\mathbf{k})$  is the density of particles whose velocity component along the wavevector  $\mathbf{k}$  is properly tuned to  $\nu$ ; velocity components in the transverse directions are irrelevant and are integrated out.

With the aid of this notation, the radiative transfer equation becomes

$$\frac{dI_\nu(\mathbf{k})}{dl} = \kappa_\nu(\mathbf{k})I_\nu(\mathbf{k}), \quad (2.7)$$

where the absorption coefficient  $\kappa_\nu(\mathbf{k})$  is derived from  $\kappa(\mathbf{v})$  through the  $\mathbf{k}$ -projection defined in equation (2.4) (or eq. [2.6]). This is the standard form of the radiative transfer equation (eq. [1.1]) and the definition of the absorption coefficient  $\kappa_\nu(\mathbf{k})$  conforms precisely to the procedure prescribed by Goldreich & Kwan (1974):  $\kappa_\nu(\mathbf{k})$  is obtained by selecting from the population distribution in  $\mathbf{v}$ -space those particles whose velocity component along the ray path is tuned to the frequency  $\nu$ , irrespective of their transverse velocities. This aspect of the standard theory therefore holds rigorously and is independent of any assumptions about beaming or any other property of the maser radiation field. The explicit dependence on  $\mathbf{k}$  displayed in  $\kappa_\nu(\mathbf{k})$  ensures that we can handle an anisotropic absorption coefficient, reflecting possible interaction with anisotropic radiation that could affect the population inversion.

## 2.2. The Absorption Coefficient

A solution of the radiative transfer problem requires an expression for the absorption coefficient  $\kappa_\nu(\mathbf{k})$ , which can be determined from the level population rate equations in steady state. Since particle distributions are specified in  $\mathbf{v}$ -space, they determine directly  $\kappa(\mathbf{v})$ , not  $\kappa_\nu(\mathbf{k})$ . The latter is then derived through the  $\mathbf{k}$ -projection of equation (2.4).

Following standard procedures, we assume that particles are pumped into the maser system from other levels, a process characterized by effective pump rates per unit volume of space and velocity  $P_i(\mathbf{v})$  and a loss rate  $\Gamma$ , taken equal for the two maser levels. Spontaneous decays and elastic collisions are ignored, so the only other process that affects the level populations is interaction with the maser radiation. A particle at velocity  $\mathbf{v}$  can interact with radiation of any frequency when it travels in the right direction; thus, the radiative interaction involves the rate

$$R_i(\mathbf{v}) = B \int n_i(\mathbf{v}) I_\nu(\mathbf{k}) \delta(v_0 - v + \mathbf{v} \cdot \mathbf{k}) d^3\mathbf{k}. \quad (2.8)$$

This expression is similar to the one encountered in the radiative transfer equation (eq. [2.2]), only the integration is performed now over photon wavevector  $\mathbf{k}$  instead of particle

velocity  $\mathbf{v}$ . In analogy with the previous treatment, the density  $n_i(\mathbf{v})$  can be pulled out of the integral. With a factor of  $4\pi$  inserted for convenience, this leads to the quantity

$$I(\mathbf{v}) = \frac{1}{4\pi} \int I_\nu(\mathbf{k}) \delta(v_0 - v + \mathbf{v} \cdot \mathbf{k}) d^3\mathbf{k} \quad (2.9)$$

[designated  $J_k(\mathbf{v})$  by Bettweiser & Kegel, and  $J_*$  by Neufeld; Litvak's  $F(\mathbf{r}, \mathbf{v})$  is  $I(\mathbf{v})/J_s$ ] and  $R_i(\mathbf{v}) = 4\pi B n_i(\mathbf{v}) I(\mathbf{v})$ . Just as  $n_i(\mathbf{k})$  is the density of particles that can interact with radiation propagating with wavevector  $\mathbf{k}$ ,  $I(\mathbf{v})$  is the intensity of radiation that can interact with particles travelling with velocity vector  $\mathbf{v}$ . Similarly,  $n_i(\mathbf{k})$  contains particles with all velocities  $\mathbf{v}$ , and  $I(\mathbf{v})$  contains photons of all wavevectors  $\mathbf{k}$ . This property is best illustrated by expressing the intensity in terms of  $I(\nu, \Omega_k)$  and performing the  $\nu$ -integration. To leading order in  $v/c$  the result is

$$I(\mathbf{v}) = \frac{1}{4\pi} \int I \left[ v_0 \left( 1 + \frac{1}{c} \mathbf{v} \cdot \hat{\mathbf{k}} \right), \Omega_k \right] d\Omega_k. \quad (2.10)$$

For every direction the integrand involves the intensity at a different frequency. Therefore, the interaction rate  $R_i(\mathbf{v})$  for simulated emissions from the upper level produces maser photons of all frequencies and wave vectors, reflecting the original mix contained in  $I(\mathbf{v})$ . Although  $R(\mathbf{v})$  cannot be associated with photon production at any single frequency or wave vector, global conservation holds; from the defining expressions it follows immediately that

$$4\pi \int I(\mathbf{v}) n_i(\mathbf{v}) d^3\mathbf{v} = \int I_\nu(\mathbf{k}) n_{i\nu}(\mathbf{k}) d^3\mathbf{k}. \quad (2.11)$$

Upper and lower limits can be imposed on  $I(\mathbf{v})$  from its definition because at any given direction, the intensity decreases monotonically away from line center. At a given  $\nu$  and any direction  $\Omega_k$ , the integrand is bounded from above by  $I(\nu_0, \Omega_k)$  and from below by  $I[\nu_0(1 \pm v/c), \Omega_k]$ . Therefore,  $I(\mathbf{v})$  is bracketed by

$$J[\nu_0(1 \pm v/c)] \leq I(\mathbf{v}) \leq J(\nu_0) \quad (2.12)$$

for any velocity  $\mathbf{v}$ .

With all the interaction rates identified, the steady state equation for the population difference becomes

$$\Delta n(\mathbf{v}) + \Delta n(\mathbf{v}) I(\mathbf{v}) / J_s = \Delta n_0(\mathbf{v}), \quad (2.13)$$

where  $\Delta n_0(\mathbf{v}) = \Delta P(\mathbf{v}) / \Gamma$  is the population difference in the limit of negligible radiative interaction. Multiplying by  $h\nu_0 B / 4\pi$  (see eq. [2.3]), this becomes an equation for the absorption coefficient in  $\mathbf{v}$ -space

$$\kappa(\mathbf{v}) = \frac{\kappa_0(\mathbf{v})}{1 + I(\mathbf{v}) / J_s}. \quad (2.14)$$

In spite of the formal resemblance to the standard theory expression for the absorption coefficient  $\kappa$ , (eq. [1.2]), there is no direct correspondence between the two; one is a relation in  $\mathbf{k}$ -space and involves radiation with a well defined wave vector, while the other is a relation in  $\mathbf{v}$  space and involves particles with a well-defined velocity. The expression just derived involves distributions in  $\mathbf{v}$ -space and must be  $\mathbf{k}$ -projected to derive the equation for  $\kappa_\nu(\mathbf{k})$ , the desired quantity. With the aid

<sup>1</sup> To leading order in  $v/c$ , the frame independent definition is  $\mathbf{v}_k = \mathbf{v} - c/v_0(v_0 - v + \mathbf{v} \cdot \mathbf{k})\hat{\mathbf{k}}$ .



of equation (2.6), the absorption coefficient in  $k$ -space is

$$\begin{aligned} \kappa_\nu(\mathbf{k}) &= \frac{c}{v_0} \int \frac{\kappa_0(\mathbf{v}_\mathbf{k}) dv_x dv_y}{1 + I(\mathbf{v}_\mathbf{k})/J_s} \\ &= \frac{c}{v_0} \int \frac{\kappa_0(\mathbf{v}_\mathbf{k}) dv_x dv_y}{1 + (1/J_s) \int I_\nu(\mathbf{k}') \delta(v_0 - v' + \mathbf{v}_\mathbf{k} \cdot \mathbf{k}') d^3 \mathbf{k}'} . \end{aligned} \quad (2.15)$$

This equation replaces the standard theory result, equation (1.2), and completes the general formulation of the maser radiative transfer problem. In general, the absorption coefficient at frequency  $\nu$  can be affected by photons at other frequencies, contained in the intensity  $I(\mathbf{v}_\mathbf{k})$ .

### 3. CALCULATION OF $\kappa_\nu(\mathbf{k})$

Consider first the unsaturated limit, in which radiative interaction has a negligible effect on the population inversion. To perform the  $k$ -projection of  $\kappa_0(\mathbf{v})$  or, equivalently,  $\Delta n_0(\mathbf{v})$  we assume from now on that the pump rates  $P_i(\mathbf{v})$  are isotropic; in particular, large velocity gradients are neglected. With the aid of equation (2.6), the  $k$ -projection transforms  $P_i(\mathbf{v})$  into

$$P_{iv}(\mathbf{k}) = \frac{c}{v_0} \int P_i(\mathbf{v}_\mathbf{k}) dv_x dv_y = P_{iv} . \quad (3.1)$$

Because  $P_i(\mathbf{v})$  is isotropic, so is  $P_{iv}(\mathbf{k})$  and the dependence on wavevector  $\mathbf{k}$  can be omitted. This result provides the definition of  $P_{iv}$ , the rate at which particles capable of interaction with photons in the frequency interval  $[\nu, \nu + d\nu]$  are pumped into the maser system. The correspondence between velocity and frequency is established with the choice of an axis whose direction need not be specified, on account of the isotropy. This leads to the expression for the unsaturated absorption coefficient

$$\kappa_{0\nu} = \frac{h\nu_0}{4\pi J_s} \Phi_\nu , \quad (3.2)$$

where  $\Phi_\nu = \frac{1}{2} \Delta P_\nu$  and where  $\Delta P_\nu$  is the difference in pump rates at frequency  $\nu$ , assumed to follow a Gaussian frequency distribution with width  $\Delta\nu_D = (v_0/c)\Delta v_D$ . This is the standard theory expression for the unsaturated absorption coefficient  $\kappa_{0\nu}$ .

With this result and the Gaussian parametrization for  $P_i(\mathbf{v})$ , the absorption coefficient becomes

$$\kappa_\nu(\mathbf{k}) = \kappa_{0\nu} \int \frac{e^{-(v_x^2 + v_y^2)/\Delta v_D^2}}{1 + I(\mathbf{v}_\mathbf{k})/J_s} \frac{dv_x dv_y}{\pi \Delta v_D^2} . \quad (3.3)$$

This is the absorption coefficient of masers of any shape in the absence of large velocity gradients. From equation (2.12) it follows immediately that

$$\kappa_{0\nu} \geq \kappa_\nu(\mathbf{k}) \geq \frac{\kappa_{0\nu}}{1 + J(v_0)/J_s} , \quad (3.4)$$

absolute bounds that hold for all frequencies and propagation directions. The absorption coefficient can be rearranged to the form

$$\kappa_\nu(\mathbf{k}) = \frac{\kappa_{0\nu}}{1 + J_\nu/J_s} [1 + s_\nu(\hat{\mathbf{k}})] , \quad (3.5)$$

so the standard expression emerges with a fractional correction of

$$s_\nu(\hat{\mathbf{k}}) = \frac{1}{2\pi} \int_0^{2\pi} d\psi \int_0^\infty \frac{J_\nu/J_s - I(\mathbf{w}_\nu)/J_s}{1 + I(\mathbf{w}_\nu)/J_s} e^{-w^2} dw^2 , \quad (3.6)$$

where  $\mathbf{w}_\nu = [w\Delta v_D \cos \psi, w\Delta v_D \sin \psi, c(\nu - v_0)/v_0]$  and  $\psi$  is an azimuthal angle around  $\hat{\mathbf{k}}$ . Compared with linear masers with the same pumping scheme, three-dimensional geometries effectively induce a fractional change  $s$  ( $> -1$ ) in the unsaturated absorption coefficient  $\kappa_{0\nu}$ , leaving all other parameters (including  $J_s$ ) unchanged. In borrowing the expression for  $\kappa_\nu(\mathbf{k})$  from the linear maser for all geometries, the standard theory is predicated on the assumption  $|s| \ll 1$  inside the maser beam. The impact of corrections to this assumption on any quantity of interest, such as the emerging intensity, for example, can be gauged by examining the behavior of that quantity when the transformation  $\kappa_{0\nu} \rightarrow \kappa_{0\nu}(1 + s)$  is applied to the solution for the particular geometry, once an estimate for  $s$  has been obtained.

Corrections to the standard theory are significant only when  $I(\mathbf{w}_\nu)$  differs substantially from  $J_\nu$ . From the general bounds of equation (2.12),

$$J(v_0 + \Delta v_D \sqrt{w^2 + x^2}) \leq I(\mathbf{w}_\nu) \leq J(v_0) , \quad (3.7)$$

where  $x = (\nu - v_0)/\Delta v_D$  is the dimensionless frequency shift from line center. Because of the exponential factor, the integration that determines  $s$  is dominated by  $w = 0$  and the lower bound on  $I(\mathbf{w}_\nu)$  at that point is simply  $J_\nu$ . Therefore, in the most significant region of the integration the difference between  $I(\mathbf{w}_\nu)$  and  $J_\nu$  is small around line center, where the upper and lower bounds coincide, and can be expected to become substantial, if ever, only in the line wings; only there can  $I(\mathbf{w}_\nu)$  significantly exceed  $J_\nu$  in the relevant range of  $w$  [note that  $I(\mathbf{w}_\nu)$  is never smaller than  $J_\nu$  at  $w = 0$ ]. To attain the limit  $s \rightarrow -1$ , the maser must be saturated in addition to obeying the relation  $I(\mathbf{w}_\nu) \gg J_\nu$ . In an unsaturated maser, even in the extreme case  $J_\nu = 0$  and  $I(\mathbf{w}_\nu) \neq 0$ , which maximizes the possible magnitude of the correction,  $|s| \ll 1$  as long as  $I(\mathbf{w}_\nu) \ll J_s$ . The reason is simple. Since radiative interactions have a negligible effect on the absorption coefficient in an unsaturated maser, even a large error in the form of the interaction term has little impact on the basic result  $\kappa_\nu(\mathbf{k}) \simeq \kappa_{0\nu}$ . We can conclude that  $s$  is expected to be small and positive ( $0 \lesssim s \ll 1$ ) at line center and that it can become significant and approach the limit  $s \rightarrow -1$  only in the line wings of a maser that has saturated.

It is convenient to introduce the factor

$$\rho_\nu(w, \psi) = 1 - \frac{I(\mathbf{w}_\nu)}{J_\nu} = 1 - \frac{\int I(\omega, \cos \vartheta, \varphi - \psi) d \cos \vartheta d \varphi}{\int I(\nu, \cos \vartheta, \varphi) d \cos \vartheta d \varphi} , \quad (3.8)$$

where

$$\omega = \nu - x\Delta v_D(1 - \cos \vartheta) + w\Delta v_D \sin \vartheta \cos \varphi . \quad (3.9)$$

Then the correction to the standard theory is

$$s_\nu(\hat{\mathbf{k}}) = \frac{1}{2\pi} \int_0^{2\pi} d\psi \int_0^\infty \frac{e_\nu \rho_\nu}{1 - e_\nu \rho_\nu} e^{-w^2} dw^2 , \quad (3.10)$$

where

$$e_\nu = \frac{J_\nu}{J_\nu + J_s} \quad (3.11)$$

is the efficiency with which the system converts pump events into maser photons (e.g. Elitzur, Hollenbach, & McKee 1992); note that  $\kappa_{\nu, \text{std}} = \kappa_{0\nu}(1 - e_\nu)$ . In an unsaturated maser  $e_\nu \simeq J_\nu/J_s \ll 1$  and this is the order of magnitude of  $s$  there. The limit  $s \rightarrow -1$  can be approached only in a saturated maser

( $J_v > J_s$ ), where the efficiency factor approaches its limit of  $e_v \approx 1$ , when  $\rho_v$  is also sufficiently large that  $e_v \rho_v \gg 1$ .

### 3.1. Line Center

Because of the rapid variation of the exponential factor, the integration in the general expression for  $\kappa_v(\mathbf{k})$  (eq. [3.3]) is dominated by  $v_x = v_y = 0$  and to leading order the integral can be approximated by the value of the integrand at that point. Since at  $v = v_0$  the  $z$ -component of  $\mathbf{v}_k$  is also zero, the velocity  $\mathbf{v}_k$  that determines the absorption coefficient at line center is the null vector,  $\mathbf{0}$ . And because  $I(\mathbf{v} = \mathbf{0}) = J(v_0)$ , the standard theory is recovered at line center for all masers.

An estimate for the correction to this leading order result can be obtained from the bounds of equation (3.7). As a distribution sharply peaked around line center,  $J_v$  can be conveniently parameterized as a Gaussian centered on  $v_0$  with a width  $\Delta$ . Then

$$0 \leq \rho_0(w, \psi) \leq 1 - e^{-w^2/\delta^2}, \quad (3.12)$$

where subscript 0 denotes quantities at line center and  $\delta = \Delta/\Delta v_D$  is the dimensionless radiation bandwidth. Standard analysis of line narrowing shows that  $\delta \leq 1$ , approaching the upper limit of unity when the maser saturates and the radiation bandwidth broadens to the full Doppler width (Litvak 1970, 1972; Goldreich & Kwan 1974). Because  $e_0 \rho_0$  is always less than 1, the integrand in equation (3.10) can be expanded as a power series in this variable. This series expansion provides a useful means to estimate bounds on  $s_0$ .

One bound can be obtained by inserting the bounds on  $\rho_0$  from equation (3.12). The series can then be integrated term by term to produce

$$0 \leq s_0(\hat{\mathbf{k}}) \leq \sum_{n=1}^{\infty} \frac{n! e_0^n}{(1 + \delta^2)(2 + \delta^2) \dots (n + \delta^2)}. \quad (3.13)$$

While both inequalities hold for all intensities, the upper bound is mostly useful when  $e_0$  is not too close to unity. In particular, in the unsaturated domain the series is adequately approximated by its first term producing

$$s_0(\hat{\mathbf{k}}) < e_0. \quad (3.14)$$

This absolute bound is general and applies to all unsaturated masers, irrespective of the geometry, up to and including the point of saturation. Indeed, a maser that has just saturated has  $e_0 = \frac{1}{2}$  while the series summation produces  $s_0 \leq 0.386$  if we assume that  $\delta \approx 1$  at this point. The bound on  $s_0$  declines exponentially into the unsaturated domain.

Another bound, more useful in the saturated domain, can be obtained by utilizing the additional fact that maser radiation is tightly beamed. For any  $w$  and  $\psi$ , upper and lower bounds can be placed on the factor  $\rho_0(w, \psi)$  by replacing the frequency  $\omega$  with the extreme values it can take in the integration range. Because of beaming,  $\omega$  is effectively confined to  $\omega \lesssim v_0 + w\theta_B \Delta v_D \cos \varphi$ , where  $\theta_B (\ll 1)$  is the beaming angle, producing the alternative upper limit

$$\rho_0(w, \psi) \lesssim 1 - \frac{1}{4\pi J_0} \int I(v_0 + w\theta_B \Delta v_D \cos \varphi, \cos \vartheta) d \cos \vartheta d \varphi. \quad (3.15)$$

Here the dependence of  $I(v, \vartheta, \varphi)$  on azimuthal angle has been dropped, assuming propagation along the axis of a beam whose pattern is symmetric around that axis. This bound can

now be used in the series expansion instead of the corresponding inequality of equation (3.12). Since  $\theta_B \ll 1$  and, thanks to the exponential factor in the integration that determines  $s_v$ , only  $w \lesssim 1$  is significant, the integrand can be expanded around  $I(v_0, \cos \vartheta)$  in a power series in  $\theta_B$ . To leading order in  $\theta_B$ , this produces the bound

$$s_0(\mathbf{m}) \lesssim \frac{1}{2} \theta_B^2 \frac{e_0}{\delta^2}, \quad (3.16)$$

where  $\mathbf{m}$  is a unit vector along the beam axis. While this result applies to all degrees of saturation, it is mostly useful when the maser is saturated at line center. Then  $e_0 \approx \delta \approx 1$ , and

$$s_0(\mathbf{m}) \lesssim \frac{1}{2} \theta_B^2 \quad (3.17)$$

for all saturated masers. For given pumping scheme and geometry, a maser that has just saturated has a beaming angle  $\theta_{\text{sat}}$  and the correction to the standard theory at line center is bounded by  $\frac{1}{2} \theta_{\text{sat}}^2$  at that point; astronomical masers typically have  $\theta_{\text{sat}}^2 \lesssim 0.1$  (see § 4). As the maser size increases, the beaming angle decreases. Therefore, the upper bound on  $s_0(\mathbf{m})$  becomes tighter with increasing intensity into the saturated regime. This variation typically behaves as a power law.

The general bounds obtained here cover the full range of intensities and all geometries. They show that 40% is an absolute bound on the correction to the absorption coefficient at line-center for all masers irrespective of geometry or degree of saturation. For an arbitrary geometry, a comparison of two solutions to the standard theory, one with  $\kappa_{0v}$ , the other with  $1.4\kappa_{0v}$ , provides an indication of the maximum possible significance corrections to the standard theory can have under any circumstances at line center.

### 3.2. $\delta$ -Function Approximation

We proceed now to evaluate the corrections to the standard theory at all frequencies. Consider first the situation when the radiation angular distribution is approximated with the standard theory assumption, a  $\delta$ -function along the direction of the local dominant ray:

$$I(v, \mu, \varphi) = 4\pi J(v) \delta(\mu - \mu_m) \delta(\varphi - \varphi_m). \quad (3.18)$$

Here  $\mu$  and  $\varphi$  specify  $\Omega_k$ , and  $\mu_m$  and  $\varphi_m$  specify the direction of  $\mathbf{m}$ , the unit vector along the local main ray. It is important to note that this approximation is used only in the angular integration that determines the associated  $v$ -distribution  $I(v)$ . Straightforward integration produces

$$I(v) = J[v_0(1 + \mathbf{v} \cdot \mathbf{m}/c)], \quad (3.19)$$

a result already derived by Bettwieser & Kegel (1974). The velocity dependence of  $I(v)$  involves only  $v_m$ , the component along the beam direction. When  $v_m = 0$ ,  $I(v) = J(v_0)$ , an isotropic distribution. For a fixed magnitude  $v = |\mathbf{v}| \neq 0$ ,  $I(v)$  is a symmetric distribution about the  $\mathbf{v} \perp \mathbf{m}$  plane, where it reaches maximum, and its extreme values are

$$I(v) = \begin{cases} J(v_0) & \mathbf{v} \perp \mathbf{m} \\ J[v_0(1 \pm v/c)] & \mathbf{v} = \pm v\mathbf{m}, \end{cases} \quad (3.20)$$

in agreement with equation (2.12). The top panel of Figure 1 presents the angular distribution of  $I(v)$  for various values of  $v$ .  $I(v)$  is elongated in the direction orthogonal to that of  $I(\mathbf{k})$ , which is an infinitely narrow beam along  $\mathbf{m}$ , and its spatial

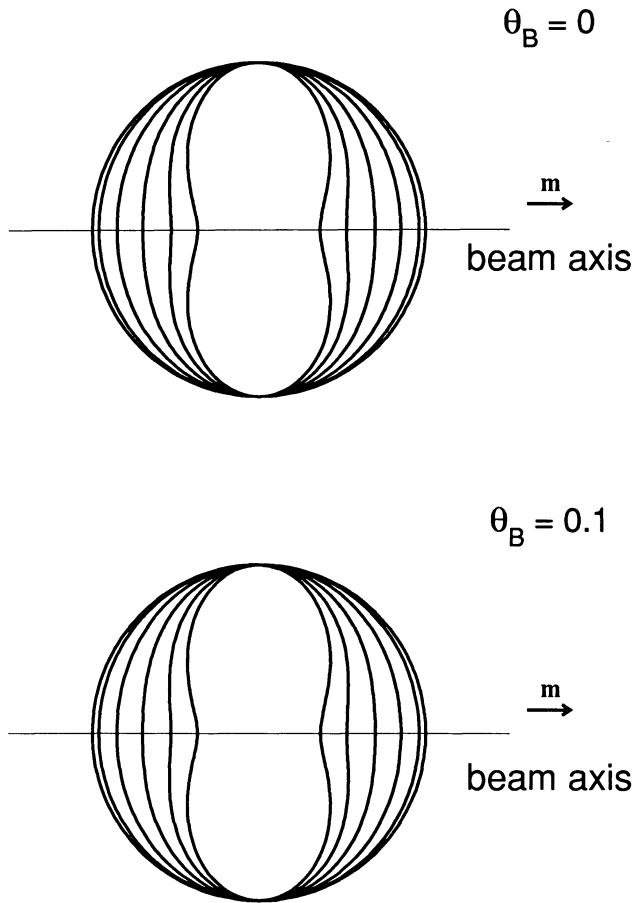


FIG. 1.—Angular distribution of the intensity  $I(\nu)$  around the direction of a beam with rectangular profile (see eq. [3.28]). The angle-averaged intensity  $J(\nu)$  is assumed to have a Gaussian frequency profile with width  $\Delta$ . The different contours, from outside in, correspond to  $\nu_0 v/c\Delta = 0, 0.2, 0.4, 0.6, 0.8, 1$ . The top plot for  $\theta_B = 0$  corresponds to  $\delta$ -function beaming (see eq. [3.19]).

variation reflects the frequency profile of  $J(\nu)$ , assumed Gaussian in this figure. Note that in this case  $I(\nu, \alpha) \propto \exp(-u^2 \cos^2 \alpha)$ , where  $u = \nu_0 v/c\Delta$  and  $\alpha$  is the angle between  $\mathbf{v}$  and  $\mathbf{m}$ .

From its definition in equation (3.8), the factor  $\rho_\nu$  is now

$$\rho_\nu(w, \psi) = 1 - \frac{1}{J(\nu)} \times J[\nu - x\Delta\nu_D(1 - \cos \theta) + w\Delta\nu_D \sin \theta \cos \psi], \quad (3.21)$$

where  $\theta$  is the angle between  $\mathbf{k}$  and  $\mathbf{m}$ . Because only  $\theta \ll 1$  is of interest, this function can be expanded as a power series in  $\theta$ , and to second order

$$\rho_\nu(w, \psi) = -\theta w \cos \psi J'/J + \frac{1}{2}\theta^2(xJ'/J - w^2 \cos^2 \psi J''/J), \quad (3.22)$$

where the prime denotes derivative with respect to  $x$ . With the Gaussian parameterization for  $J_\nu$ , the correction to the standard theory becomes

$$s_\nu(\theta) = \frac{1}{2} \theta^2 \frac{e_\nu}{\delta^2} \left[ 1 - 2x^2 \left( 1 + \frac{1}{\delta^2} \right) + 4x^2 \frac{e_\nu}{\delta^2} \right]. \quad (3.23)$$

The correction vanishes for  $\theta = 0$ , so the standard theory holds rigorously for propagation along the main ray. Terms linear in

$\theta$  have vanished because of the  $\psi$ -integration, and the correction is only second order in  $\theta$ . Note that at line center the correction obeys the bound of equation (3.16) when  $\theta \leq \theta_B$ . In a saturated maser, where the correction  $s$  is maximal, it is simply

$$s_\nu(\theta) \simeq \frac{1}{2} \theta^2, \quad (3.24)$$

assuming  $e_\nu \simeq \delta \simeq 1$ . This result is easy to understand. Radiative interactions reduce the absorption coefficient of a system with an inverted population. At frequency  $\nu$ , the standard expression follows from the assumption that for all particles with velocity  $v_m = c(\nu/\nu_0 - 1)$ , this reduction is controlled by  $J_\nu$ . Instead, the calculation here shows that the relevant intensity is actually  $J\{\nu_0[1 + (1/c)v_k \cdot \mathbf{m}]\}$ , which varies with  $\theta$  (although all motions with the same  $v_m$  are still associated with a single frequency owing to the assumption of  $\delta$ -function beaming). As  $\theta$  increases, the involved frequencies are further removed from line center than  $\nu$ . And because the intensity decreases with distance from line center, the radiative effects are less important and the absorption coefficient is reduced less than the standard theory implies, leading to a correction factor that increases with  $\theta$ . However, the corrections to the standard theory are always negligible ( $s \ll 1$ ) because only rays inclined at  $\theta \ll 1$  are relevant. Note also that an analysis based on  $\delta$ -function beaming produces meaningful results since this assumption was only made in the calculation of  $\rho_\nu$ . In all other applications a finite beaming angle can still be employed. Since maser radiation is highly beamed, this procedure is internally consistent and can be expected to provide an adequate leading order estimate for the absorption coefficient.

### 3.3. Finite Beam; Rectangular Profile

We proceed now to finite beaming patterns and consider first a rectangular profile for the intensity angular distribution

$$I(\nu, \theta, \varphi) = \begin{cases} I_\nu & \theta \leq \theta_B \\ 0 & \theta > \theta_B \end{cases}, \quad (3.25)$$

so that the angle-averaged intensity is  $J_\nu = \frac{1}{4} I_\nu \theta_B^2$ . Here  $\theta_B (\ll 1)$  is the opening angle of a cone whose axis is aligned with  $\mathbf{m}$ , and  $\theta$  and  $\varphi$  are measured with respect to that axis. The  $\delta$ -function profile of the previous section can be obtained from this distribution as the  $\theta_B \rightarrow 0$  limit at fixed  $J_\nu$ . The rectangular profile adequately describes filaments and other matter-bounded masers (see Elitzur et al. 1992). In these sources  $\theta_B$  is frequency independent, the assumption that will be made here.

Since  $\mathbf{m}$  provides an axis of symmetry for the radiation field, the calculation of  $I(\nu)$  is conveniently performed choosing axes such that  $\mathbf{m}$  is along the  $z$ -axis and  $\mathbf{v} = v(\sin \alpha, 0, \cos \alpha)$ , leading to

$$I(\nu) = \frac{4}{\theta_B^2} \int_0^{\theta_B} \theta d\theta \int_0^{2\pi} d\varphi J(\hat{\nu}), \quad (3.26)$$

where

$$\hat{\nu} = \nu_0 [1 + (v/c)(\sin \theta \cos \varphi \sin \alpha + \cos \theta \cos \alpha)]. \quad (3.27)$$

Because  $\theta \leq \theta_B \ll 1$ , the integration can be performed with the aid of a series expansion in  $\theta$  for  $J(\hat{\nu})$ . The linear term in  $\theta$  vanishes in the  $\varphi$  integration and the leading order  $\theta$ -dependence again is quadratic. With the Gaussian param-



eterization previously introduced for  $J_v$ , the result is

$$I(\mathbf{v}) = [1 + (\frac{1}{2})\theta_B^2 b] \times 4\pi J[v_0(1 + \mathbf{v} \cdot \mathbf{m}/c)], \quad (3.28)$$

where the coefficient of the finite beam correction to the  $\delta$ -function result is

$$b(\mathbf{v}) = \left(\frac{v_0 v_m}{c\Delta}\right)^2 - \left(\frac{v_0 v_\perp}{c\Delta}\right)^2 \left[ \frac{1}{2} - \left(\frac{v_0 v_m}{c\Delta}\right)^2 \right] \quad (3.29)$$

and where  $v_\perp$  is the velocity component in the direction perpendicular to  $\mathbf{m}$ . The  $\delta$ -function result is properly recovered in the  $\theta_B \rightarrow 0$  limit. The finite-beam corrections introduce a dependence on  $v_\perp$  on top of the dependence on  $v_m$ , the velocity component along the beam direction. However, these corrections are minute for all values of interest for  $\theta_B$ . As can be seen from Figure 1, the plots for  $\theta_B = 0$  and 0.1 are hardly distinguishable. Note also that according to the bounds listed in equation (2.12),  $b(\mathbf{v})$  must be negative when  $v_m = 0$  and positive when  $v_\perp = 0$ , which it is.

With this result for  $I(\mathbf{v})$ , the calculation proceeds in complete analogy to the case of  $\delta$ -function profile, and to second order in  $\theta$  and  $\theta_B$  the correction factor is

$$s_v(\theta) = \frac{1}{4} \theta_B^2 \frac{e_v}{\delta^2} \left\{ \left( 1 + \frac{2\theta^2}{\theta_B^2} \right) \left[ 1 - 2x^2 \left( 1 + \frac{1}{\delta^2} \right) \right] + 8x^2 \left( \frac{\theta^2}{\theta_B^2} \right) \frac{e_v}{\delta^2} \right\}. \quad (3.30)$$

This result differs from the one for  $\delta$ -function beaming in two significant ways. First, the correction along the main ray does not vanish now. In particular, at line center  $s_0(0) = \frac{1}{4}\theta_B^2 e_0 \delta^{-2}$ , in agreement with the general bound of equation (3.26). And because the bound of equation (3.14) must also always be obeyed by unsaturated masers, we find that

$$\delta^2 > \frac{1}{4}\theta_B^2, \quad (3.31)$$

a new result that places a limit on the amount of line narrowing that can take place in any maser with a rectangular beam pattern.

The top panel of Figure 2 displays the correction at line center along the axis of a cylindrical maser with a fixed radius  $R$  and half-length  $aR$  as a function of the aspect ratio  $a$ . Line-width variation was estimated from the expression derived by Litvak (1970, 1972) for linear masers

$$\delta^2 = \frac{1 + J_0/J_s}{1 + \tau_0} = \frac{1 + J_0/J_s}{1 + J_0/J_s + \ln(J_0/J_{00})}. \quad (3.32)$$

Here  $\tau_0 = \kappa_0 l$  where  $\kappa_0$  is the unsaturated absorption coefficient at line center and  $l$  is the maser half-length, and  $J_{00}$  and  $J_0$  are the line center angle-averaged intensities at the maser's center and edge, respectively; an exact expression for their ratio can be derived from the integral of motion of Alcock & Ross (1985; see also Elitzur 1992). Although this result strictly holds only for linear masers, its final form is expressed purely in terms of angle-averaged intensities so the dependence on geometry is expected to be minimal when the proper value of  $J$  is inserted. The pumping scheme is defined through  $\kappa_0 R = 4$  and  $\gamma = J_s/S_0 = 10^6$ , where  $S_0$  is the source function. With these parameters, which describe typical filaments produced in the  $\text{H}_2\text{O}$  shock models of Elitzur, Hollenbach, & McKee (1989), the maser saturates at an aspect ratio  $a_{\text{sat}} = 2.3$ . As

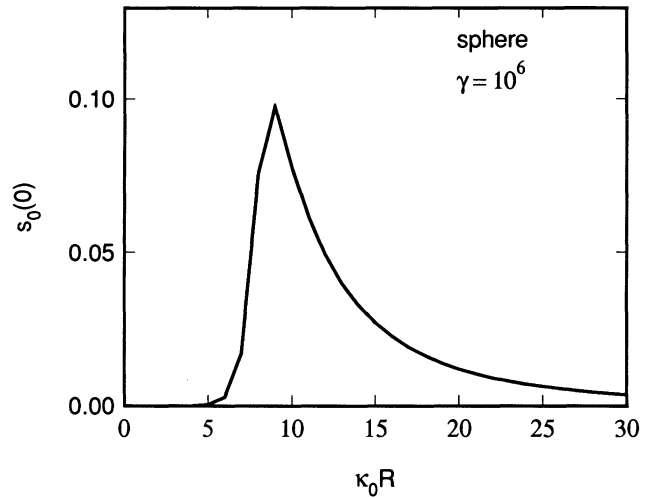
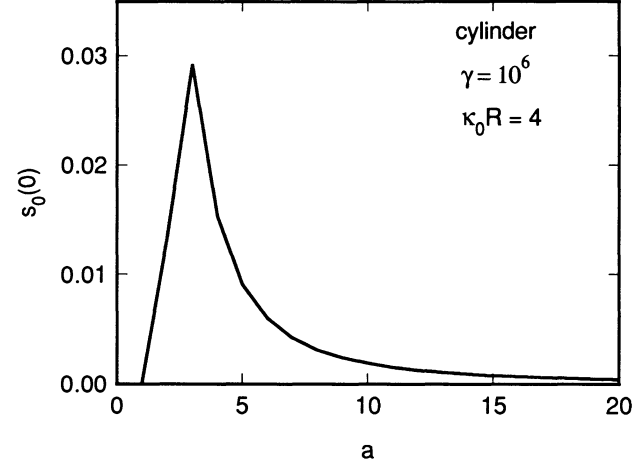


FIG. 2.—Line center corrections to the standard theory along the dominant rays of cylindrical and spherical masers. In both cases the pumping scheme is characterized by  $\gamma = 10^6$ . The cylinder has a radius  $R$  such that  $\kappa_0 R = 4$ , where  $\kappa_0$  is the unsaturated absorption coefficient at line center, and half length  $aR$ . In the case of the sphere,  $R$  is the radius.

evident from the figure,  $s_0(0)$  peaks at  $a = a_{\text{sat}}$ , declining exponentially for  $a < a_{\text{sat}}$  and as a power law for  $a > a_{\text{sat}}$ , the behavior deduced before from general arguments. Since the peak value of  $s_0(0)$  is only about 3%, the standard theory provides an excellent approximation at line center for all cylindrical masers, as expected.

A second, more important consequence of the finite size of the beam is the frequency variation of the correction function. In a saturated maser, where it is maximal, the correction is now

$$s(x, \theta) = \frac{1}{4} \theta_B^2 \left( 1 + 2 \frac{\theta^2}{\theta_B^2} - 4x^2 \right). \quad (3.33)$$

Again, this function displays the behavior deduced before in the general discussion following equation (3.5). At any inclination  $\theta$  the correction starts positive and small at  $x = 0$ , decreasing with distance from line center and approaching the limit  $s \rightarrow -1$  in the wings. With the leading order estimate obtained here,  $s$  actually reaches this limit, first along the axis at  $x = 1/\theta_B$ , signaling a breakdown of the standard theory.

### 3.4. Finite Beam; Gaussian Profile

Consider, finally, a Gaussian angular distribution for the intensity

$$I(\nu, \theta) = I(\nu, 0)e^{-\theta^2/\theta_B^2}, \quad (3.34)$$

so that  $J_\nu = \frac{1}{4}I(\nu, 0)\theta_B^2$ . Again,  $\theta$  is measured from the direction of the local main ray  $\mathbf{m}$  and the  $\delta$ -function profile can be recovered as the  $\theta_B \rightarrow 0$  limit at fixed  $J_\nu$ . The Gaussian profile adequately describes spheres and other amplification bounded masers where  $\theta_B$  is frequency dependent (Elitzur 1990; Elitzur et al. 1992).

Similar to the calculation for a rectangular profile,  $I(\nu)$  is

$$I(\nu) = \int_0^\infty \theta d\theta \int_0^{2\pi} d\varphi I(\hat{\nu}, 0)e^{-\theta^2/\theta_B^2(\hat{\nu})}, \quad (3.35)$$

where the frequency  $\hat{\nu}$  is listed in equation (3.27). Because of the fast decline of the Gaussian profile away from the axis, only rays inclined at  $\theta \lesssim \theta_B \ll 1$  make an appreciable contribution, and a series expansion in  $\theta$  for the  $\hat{\nu}$ -dependence can be utilized as before. Since the argument of the exponential is already proportional to  $\theta^2$ , the function  $1/\theta_B^2(\hat{\nu})$  must be evaluated in lowest order [that is, replaced by  $1/\theta_B^2(\nu)$ ] and only  $I(\hat{\nu}, 0)$  need be expanded. With a Gaussian parameterization with width  $\Delta$  for  $I(\nu, 0)$ , the calculation proceeds in complete analogy to the case of a rectangular profile. The only difference between the two cases is that a rectangular profile gives  $\langle \theta^3 \rangle / \langle \theta \rangle = \frac{1}{2}\theta_B^2$  while a Gaussian profile gives  $\langle \theta^3 \rangle / \langle \theta \rangle = \theta_B^2$ . As a result, the correction factor for a Gaussian profile is simply twice that for a rectangular profile, and the breakdown of the standard theory ( $s = -1$ ) occurs now at  $x = 1/2^{1/2}\theta_B = 0.7/\theta_B$ . The correction at line center along the main ray is now equal to the upper bound of equation (3.16), and from equation (3.14)

$$\delta^2 > \frac{1}{2}\theta_B^2. \quad (3.36)$$

This is the restriction on line narrowing in all masers with a Gaussian beam pattern.

The bottom panel of Figure 2 displays the variation with radius  $R$  of  $s_0(0)$ , the correction at line center along the radial ray of a spherical maser. With  $\gamma = 10^6$ , saturation occurs at  $\kappa_0 R = 8.7$ , and that is where the correction peaks, as expected. The correction factor behaves similarly to the cylindrical case, the main difference is the larger peak value of 10%, reflecting the extra factor of 2 as well as the slightly larger beaming angle at saturation. Still, as expected, the correction is negligible at line center for any radius of the sphere.

The calculation displayed in the figure utilized the standard theory expression for the beaming angle to evaluate  $s$ , the correction to that theory. This procedure is justified when  $|s| \ll 1$ , which is the case for  $x \sim 0$  and  $\theta \leq \theta_B$ , because to leading order the standard expression for the absorption coefficient is applicable, the radiative transfer equation is unmodified, and the standard solution holds. For example, in spherical coordinates the radiative transfer equation is

$$\cos \theta \frac{\partial I_\nu(r, \theta)}{\partial r} - \frac{\sin \theta}{r} \frac{\partial I_\nu(r, \theta)}{\partial \theta} = \kappa_\nu(r, \theta) I_\nu(r, \theta). \quad (3.37)$$

When  $|s| \ll 1$ ,  $\kappa_\nu \simeq \kappa_{\nu, \text{std}}$  and the intensity spatial distribution (both radial and angular) is properly determined by the standard theory; in particular, the standard result for the beaming angle is unchanged to leading order. A number of results rele-

vant to beaming in spherical masers are presented in an appendix.

### 4. SUPPRESSED MASERS

Formal, complete theory for the absorption coefficient  $\kappa_\nu(\mathbf{k})$  shows that, thanks to the tight beaming of maser radiation, to leading order the standard theory always provides the correct description for maser behavior at and near line center and all its results remain intact in that region. However, a previously unrecognized feature of three dimensional masers is the breakdown of the standard theory in the line wings where the absorption coefficient is significantly reduced when  $|x|$  is  $\sim 1/\theta_B$ . This effect will be called *suppression*.

Suppression reflects a fundamental difference between linear and three-dimensional models. Maser radiation with a wavevector  $\mathbf{k}$  (frequency  $\nu$ ) can interact with particles with any velocity  $\mathbf{v}$  when Doppler matching is satisfied by the product  $\mathbf{v} \cdot \mathbf{k}$ . In linear masers  $\mathbf{v}$  and  $\mathbf{k}$  are always parallel, resulting in a unique correspondence between velocity  $v$  and frequency  $\nu$ . As a result, the absorption coefficient  $\kappa_\nu$  is only affected by radiation at that frequency. Three dimensions change this unique correspondence because of two reasons. First, even when the radiative interaction is calculated with the beam pattern of a linear maser, i.e.,  $\delta$ -function beaming,  $\mathbf{v}$  and  $\mathbf{k}$  are only parallel when radiative transfer is solved along the beam. At other directions, Doppler matching introduces the difference  $\theta$  between the directions of  $\mathbf{k}$  and the beam axis. Instead of  $J_\nu$ , the absorption coefficient  $\kappa_\nu(\mathbf{k})$  is now affected by the intensity at a different frequency, uniquely determined by the inclination  $\theta$ . The resulting suppression factor  $s_\nu$  is small, and independent of frequency when the maser saturates.

The second, and much more important, effect results from the finite size of the beam, which can be visualized as a continuum of linear masers inclined at  $\theta \lesssim \theta_B$ . This introduces a spread in the frequencies that can affect the absorption coefficient at frequency  $\nu$ , an effect that can be readily understood from the general expression for the absorption coefficient, equation (3.3). Because of the rapid variation of the exponential function, the integration is dominated by  $v_x = v_y = 0$  and to leading order the integral can be approximated by the value of the integrand at that point. Therefore,

$$\kappa_\nu(\mathbf{k}) \simeq \frac{\kappa_{0\nu}}{1 + \hat{J}_\nu(\hat{\mathbf{k}})/J_s}, \quad (4.1)$$

where  $\hat{J}_\nu(\hat{\mathbf{k}}) = I(\mathbf{u})$  and where  $\mathbf{u} = [0, 0, c(\nu - \nu_0)/\nu_0]$ , so that

$$\hat{J}_\nu(\hat{\mathbf{k}}) = \frac{1}{4\pi} \int I[\nu - (\nu - \nu_0)(1 - \mu), \mu, \varphi] d\mu d\varphi. \quad (4.2)$$

Here  $\theta (= \cos^{-1} \mu)$  and  $\varphi$  are measured with respect to  $\hat{\mathbf{k}}$ . Because  $\hat{J}(\nu_0, \hat{\mathbf{k}}) = J(\nu_0)$  for any  $\hat{\mathbf{k}}$ , the standard theory is recovered at line center for all masers, as before; note again that the beaming assumption is not even required in this case. For all frequencies, consider propagation along the axis of a beam with rectangular profile. A simple change of integration variable yields

$$\hat{J}_\nu(\mathbf{m}) = \frac{1}{x\theta_B^2} \int_{x-x\theta_B^2/2}^x J(x') dx'. \quad (4.3)$$

The effect of radiative interactions on the absorption coefficient at frequency  $\nu$  involves intensities spread over a frequency band whose width increases with the beaming angle and dis-

tance from line center. If the frequency distribution of  $J_\nu$  were flat, the standard theory would hold along the beam for all frequencies since  $\hat{J}_\nu(\mathbf{m}) = J_\nu$  in that case. This is the situation close to line center, where  $\hat{J}_\nu$  can be expected to differ little from  $J_\nu$ . But the difference between these two intensities grows with distance from line center, and to first order

$$\hat{J}_\nu(\mathbf{m}) \simeq J_\nu(1 - \frac{1}{4}x\theta_B^2 J'/J) \quad (4.4)$$

where the prime denotes derivative with respect to  $x$ . Since  $J' < 0$ ,  $\hat{J}_\nu$  exceeds  $J_\nu$ , significantly when  $-\frac{1}{4}x\theta_B^2 J'/J \sim 1$ . For a Gaussian frequency profile this condition becomes  $\frac{1}{2}\theta_B^2 x^2/\delta^2 \sim 1$ . To have an appreciable effect on the absorption coefficient,  $\hat{J}_\nu$  must exceed  $J_s$ , i.e., the maser must be saturated at the corresponding frequency. But in that case it is certainly saturated at line center and  $\delta \simeq 1$ , producing an estimate for the suppression frequency that is similar to but slightly larger than that obtained in the previous section. The reason for this difference is the neglect of transverse motions in the present derivation. At line center, the significance of this neglected contribution can be gauged from  $s_0(0)$  presented in Figure 2 since this function is entirely determined by transverse motions. Moving away from line center, transverse and longitudinal motions contribute with opposite signs, as is evident from equation (3.29). At larger  $x$  (i.e., larger  $v_m$ ) both motions contribute in the same sense, so transverse contributions increase the suppression effect, making it significant closer to line center and bringing the two estimates into agreement. Similar results can be derived for a Gaussian beam pattern, which yields

$$\hat{J}_\nu(\mathbf{m}) = \int_0^\infty e^{-u} J \left[ v - \frac{1}{2}(v - v_0)\theta_B^2 u \right] du \simeq J_\nu \left( 1 - \frac{1}{2} \frac{x\theta_B^2 J'}{J} \right). \quad (4.5)$$

As before, suppression is twice that for a rectangular distribution.

Both saturation and suppression reflect the reduction in  $\kappa_\nu(\mathbf{k})$  due to radiative interactions, a reduction caused because the very process of maser photon production in stimulated emission removes particles from the upper level, reducing the inversion. But the two effects arise from entirely different circumstances and produce completely different results. Saturation is a self-regulating mechanism that involves the effect of radiation at frequency  $\nu$  on the very same particles that produced it. When saturation dominates, the absorption coefficient is slaved to the radiation [ $\kappa_\nu(\mathbf{k}) = \kappa_{0\nu} J_s/J_\nu \propto \Phi_\nu/J_\nu$ ], and its decreasing magnitude simply ensures that the maser produce as many photons as determined by the pump rates. Because of saturation,  $\kappa_\nu(\mathbf{k})$  decreases without bound, in proportion to the intensity. In contrast, maser suppression at frequency  $\nu$  arises because potential interaction partners are lost to that frequency owing to interaction with radiation at other frequencies. The resulting reduction in  $\kappa_\nu(\mathbf{k})$ , the suppression factor  $s_\nu$ , is bounded by a limit which is independent of intensity, and the medium's response at frequency  $\nu$  is reduced *without any direct relation to the intensity at that frequency*. As a result, in the frequency domain where suppression dominates, the maser never reaches the potential photon production rate that the same pumping processes would generate in a linear maser.

From its definition in equation (3.5), suppression can be viewed as a fractional change  $s_\nu$  in  $\kappa_{0\nu}$ , while  $J_s$  remains intact. And from equation (3.2), this is equivalent to a fractional change  $s_\nu$  in the photon production rate  $\Phi_\nu$ . Photon pro-

duction involves competition for the inverted population among all the photons that can interact with particles with a given velocity, and higher intensities always win in such a competition. Photons at a given frequency shift  $x_1$  lose to photons at  $|x| < |x_1|$  and win against photons with  $|x| > |x_1|$ . Consider first the line center. Because this frequency wins against all others, photon production there is enhanced in comparison with the standard theory, which is why  $s_0 > 0$ . This enhancement is negligible, though, because all other intensities are lower than the line center intensity. At a slight shift  $|x| \neq 0$  the maser loses to the line center although it gains from all other frequencies, which is why  $s$  decreases. The ratio of gains to losses continues to decrease with distance from line center until they balance each other at  $x = \frac{1}{2}(1 + 2\theta^2/\theta_B^2)^{1/2}$ , and  $s = 0$  at that point. Farther away from line center, losses outweigh gains and  $s < 0$ . Finally, when  $|x| \gtrsim x_s$  losses become more important than the leading term and suppression dominates.

Because beaming angles decrease when the source size increases and  $x_s$  is  $\sim 1/\theta_B$ , it would appear that the suppression regime should move further into the line wings for larger masers. This is not the case. Each saturated maser solution is approached only through a series of masers with an increasing overall gain. If at a certain frequency the theory breaks down for a certain gain, a structure with a larger gain cannot be reached. Since the largest beaming angles correspond to the smallest masers, the frequency shift determining the suppression region for any given geometry is  $x_s \sim 1/\theta_{\text{sat}}$ , where  $\theta_{\text{sat}}$  is the beaming angle of a maser with the particular shape that has just saturated. In a saturated cylindrical maser with aspect ratio  $a$  the beaming angle is  $0.6/a$  (Goldreich & Keeley 1972; Elitzur, McKee, & Hollenbach 1991) so

$$x_s = \frac{1}{\theta_{\text{sat}}} = 1.7a_{\text{sat}}, \quad (4.6)$$

where for a fixed radius,  $a_{\text{sat}}$  is the aspect ratio of a cylinder that has just saturated. Filaments produced in the  $\text{H}_2\text{O}$  shock models of Elitzur et al. (1989), for example, typically saturate at  $a_{\text{sat}} \sim 2-3$  so  $x_s$  is  $\sim 3-5$ . This estimate should typify most pumping schemes of astronomical masers. In the case of spherical masers the suppression frequency is

$$x_s = \frac{0.7}{\theta_{\text{sat}}} = 0.7(\kappa_{0\nu} R_{\text{sat}})^{1/2}, \quad (4.7)$$

where  $R_{\text{sat}}$  is the radius at which the sphere saturates. For  $\gamma = 10^5-10^7$ ,  $\kappa_{0\nu} R_{\text{sat}} \sim 7.5-9.9$ , so  $x_s \sim 1.9-2.2$ . Disk masers provide the prototype of planar masers, and their properties are intermediate between those of spheres and filaments (Elitzur et al. 1992). The beaming angle of a disk maser is roughly the geometric mean of the beaming angles of a sphere with the same radius and a cylinder whose half-length is the disk radius and whose diameter is its thickness. A similar relation can be expected between the suppression frequencies, so for typical pumping schemes  $x_s$  is  $\sim 2.4-3.3$  for disk masers. As could be expected, since filaments are most similar to linear masers, they have the largest frequency domain where the standard theory is valid while spheres are most dissimilar and have the smallest validity region, with disks in between. In all cases, however, the standard theory applies in a frequency core whose half-width exceeds  $\sim 2\Delta\nu_D$ .

Whatever the geometry, a maser saturates when a characteristic dimension exceeds a length scale that varies in proportion to  $\ln \gamma$ , and typical pumping schemes of astronomical masers



have  $\gamma \sim 10^5\text{--}10^7$  (Elitzur 1992). For a given geometry the onset of suppression is therefore controlled by the pumping scheme through the value of  $\gamma$ ; the larger is  $\gamma$ , the tighter is the beaming angle when the maser saturates, extending the domain of applicability of the standard theory into the line wings. In the limit  $\gamma \rightarrow \infty$  the suppression region moves farther out and the standard theory is fully recovered across the entire line; that is, the standard theory properly provides the  $\gamma \rightarrow \infty$  limit of masers in three dimensions. However, although values of  $\gamma$  of actual pumping schemes are quite large they are finite, and the dependence of the suppression boundary on  $\gamma$  is only logarithmic, resulting in the finite values of  $x_s$  just derived.

#### 4.1. Frequency Profiles

With the general expression for the absorption coefficient of equation (3.5), the flux divergence relation becomes

$$\nabla \cdot \mathbf{F}_\nu = h\nu_0 \Phi_\nu e_\nu \left[ 1 + \int s_\nu(\theta) \frac{I_\nu}{4\pi J_\nu} d\Omega \right]. \quad (4.8)$$

The factor in front of the brackets controls the line width behavior in the standard analysis. Suppression contributes through the second term inside the square brackets, the angle integral of the suppression factor weighted by the corresponding normalized angular profile  $I_\nu/4\pi J_\nu$ . In an unsaturated maser, to leading order in  $e_\nu$  the suppression correction can be neglected and standard line narrowing analysis is recovered. In a saturated maser  $e_\nu \simeq 1$  and the pump rate  $\Phi_\nu$  in front of the square brackets provides the standard Doppler profile for the maser flux. For any of the angular distributions considered above, angular integration of the corresponding suppression profile is straightforward, leading to

$$\frac{F_\nu}{F_0} \simeq e^{-x^2} \left( 1 - \frac{x^2}{x_s^2} \right) = e^{-x^2 + \ln(1 - x^2/x_s^2)}. \quad (4.9)$$

Close to line center, the profile remains a Gaussian with a width  $\Delta\nu_D x_s/(1 + x_s^2)^{1/2}$ , so the line width is reduced when  $x_s < 1$ . However, since  $x_s \gtrsim 2$  for all astronomical masers, this reduction never exceeds 11% in actual sources, an indication of the overall accuracy of the first-order calculation presented here. Suppression was calculated assuming a Gaussian profile, and the resulting profile is again Gaussian only modified by logarithmic terms, which are negligible as long as  $x$  does not approach  $x_s$ .

Figure 3 presents profiles for various values of  $x_s$  typical of different geometries, with the dashed line denoting intensity at half-maximum. A profile with  $x_s = 1.5$  is added for completeness to bracket the possible range of values that may turn out to be important in higher order calculations. Evidently, suppressed profiles are indistinguishable from Gaussian around line center and the width at half-maximum is hardly affected by the suppression effect. The different profiles behave differently only when the intensity varies over a large range, as would be the case during a maser outburst for example. Because the curves are normalized to their peak value, an outburst corresponds to a downward motion of some zero intensity level  $F_{Z1}$ . In the case of a Doppler profile, the width at zero intensity would increase in proportion to  $[\ln(F_0/F_{Z1})]^{1/2}$  without bound. But the width at zero intensity of a suppressed profile would only increase up to a maximum of  $2x_s \Delta\nu_D$ , the effective cutoff imposed by the suppression effect. As a result, the suppressed profile would appear to have narrowed in comparison with the anticipated behavior of an ordinary Doppler profile.

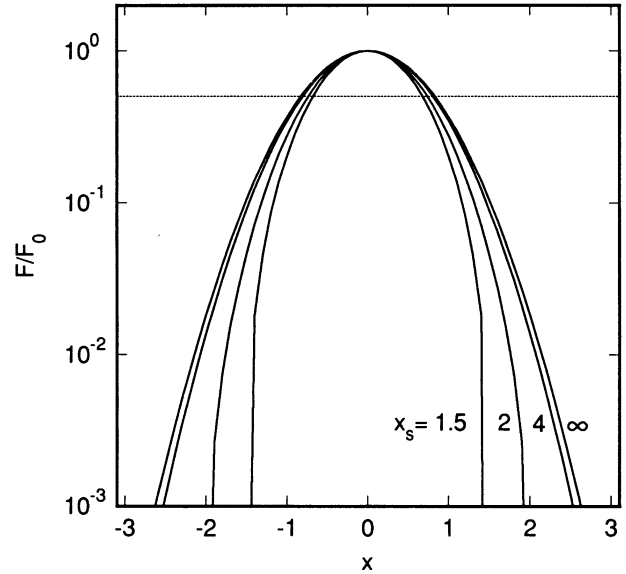


FIG. 3.—Profiles of saturated masers including the suppression effect (see eq. [4.9]) for various values of  $x_s$ , marked on the curves. The limit  $x_s \rightarrow \infty$  corresponds to the Doppler profile, the standard theory result obtained for  $\delta$ -function beaming. The dashed line marks the level of half-peak intensity.

In particular, fitting a suppressed profile with a Gaussian that passes through the peak  $F_0$  and a zero intensity baseline  $F_{Z1}$ , the corresponding width is  $\Delta = x_s \Delta\nu_D [\ln(F_0/F_{Z1})]^{-1/2}$ .

Due to suppression, saturated masers can mimic line narrowing in proportion to  $(\ln F_0)^{-1/2}$ , in contrast with the standard theory where such behavior is confined to unsaturated amplification. Such an inverse correlation between intensity and linewidth has been detected in a number of  $\text{H}_2\text{O}$  maser flares in star-forming regions (e.g., Mattila et al. 1985; Rowland & Cohen 1986; Boboltz et al. 1993; Liljestrom 1993) and was interpreted in terms of possible temperature variations (Strel'nitskij 1982, 1986, 1993). Maser suppression provides an alternative natural explanation, but only in extreme bursts. The effects of suppression on the profile are only noticeable when the dynamic range is such that

$$\frac{F_0}{F_{Z1}} \gtrsim e^{x_s^2} = 10^{0.43x_s^2}. \quad (4.10)$$

The line shape of an outburst in a disk maser, where  $x_s \gtrsim 2.4$ , would be affected by suppression only when  $F_0/F_{Z1} \gtrsim 300$ , while in a filament, where  $x_s \gtrsim 3$ , the effects would only come to play for  $F_0/F_{Z1} \gtrsim 7 \times 10^3$ . It is worth noting, though, that because of the exponential dependence, a slight change in the estimate of  $x_s$  has a significant effect on the threshold dynamic range. For example, with  $x_s = 2$  suppression will have a noticeable effect already at  $F_0/F_{Z1} \gtrsim 55$ .

The failure of saturation rebroadening to the full Doppler profile in violation of the standard theory has been noted already by Bettwieser (1976), albeit only for spherical masers.

#### 5. SUMMARY AND DISCUSSION

Although maser photons of a certain wavevector are only generated from photons of the same wavevector, the generation process involves interaction with particles, and those can interact with photons of all wavevectors. In linear masers all wavevectors point in the same direction and the absorption coefficient at frequency  $\nu$  is only affected by radiation at that



frequency. As a result, the maximal volume production rate of maser photons as determined by the pump processes,  $\Phi_{\nu}$ , is attained at every frequency when the maser saturates and the radiation bandwidth is then equal to the full Doppler width. In three-dimensional masers, on the other hand, wavevectors within the beam are slightly slanted to each other and the absorption coefficient is modified from the standard expression that was carried over from linear masers (eq. [3.5]). The leading order result for the correction term  $s_{\nu}(\theta)$  reveals an effect previously unrecognized—maser suppression. Inside a frequency core  $|\nu - \nu_0| < x_s \Delta\nu_D$ ,  $|s| < 1$ , and suppression can be regarded as a negligible perturbation; to leading order, the standard theory applies and its conclusions remain intact. Outside the core, suppression dominates and the maser intensity is greatly reduced from what the pump processes would prescribe; to leading order, the intensity can be neglected. In the suppressed regime the volume production rate of maser photons never attains the value determined by  $\Phi_{\nu}$ , however large the source may become.

In essence, the maser disappears outside the core and a major modification to the standard theory involves linewidth analysis of three-dimensional masers. Suppression effectively confines maser operation to a full width at zero intensity of  $2x_s \Delta\nu_D$  irrespective of the intensity. Therefore, observed line widths are not directly indicative of the actual Doppler width, signalling a breakdown of the standard line width analysis. When a saturated maser profile is fitted with a Gaussian, it can display line narrowing in proportion to  $(\ln F_0)^{-1/2}$  when the peak intensity is sufficiently high that the observed wings enter the suppressed regime. In contrast, in the standard theory, such behavior is confined to unsaturated amplification. Such an inverse correlation between intensity and line width, which has been detected in a number of H<sub>2</sub>O maser flares in star-forming regions, should be noticeable in strong bursts in disks, which were advocated as a prototype of the H<sub>2</sub>O masers in star-forming regions (Elitzur et al. 1992).

In his paper, Neufeld (1992) suggested that a solution of the maser problem in three dimensions may necessitate a double iteration scheme that couples successive estimates for  $I_{\nu}$  and  $\kappa(\nu)$ . The analysis presented here obviates the need for such a scheme by performing all the velocity integrations, formulating the radiative transfer problem with frequency as the sole parameter. Because maser radiation is tightly beamed, maser theory can be envisioned as a series expansion in the beam width, with the results of each iteration plugged into the expressions for the next order theory. The leading term in such

a series corresponds to zero beam width, i.e., a  $\delta$ -function distribution—the standard theory. This paper provides the general expressions for a complete iteration analysis and calculates the next term in the series, the leading order correction resulting from the finite size of the beam. Although the accuracy of the first-order results derived here is expected to deteriorate as the suppression boundary is approached ( $x \rightarrow x_s$ ), the overall conclusions and specific expressions near line center can be adopted with reasonable confidence. For example, the expansions utilized a Gaussian frequency profile for the intensity with a width obtained from the standard theory, and the resulting profile was shown to closely resemble the same Gaussian near line center, as is evident from Figure 3. More accurate (invariably numerical) estimates of suppression must cope with the fact that it is a correction to the saturation effect. Since solutions to the standard theory are usually performed only to leading order, a calculation of suppression may first require higher order corrections within the standard theory. This problem is evident in the analysis of beaming angles of spherical masers presented in the appendix.

Following the approach developed for maser line width analysis by Litvak (1970, 1972) and Goldreich & Kwan (1974), the only mechanism included in this first study in addition to exchanges with other levels (described by the pump and loss terms) is interaction with the maser radiation. Redistribution in frequency, resulting from elastic collisions for example, has been ignored. Incorporation of this effect is a difficult problem that has not yet been fully accomplished even within the standard theory. Another major, currently unsolved problem in maser theory is the incorporation of large velocity gradients. The general formalism presented here provides the framework for a study of this problem. The only modification necessary in that case is removal of the assumption of isotropic pump rates. With proper expressions for  $P_{\nu}(\nu)$  that reflect the structure of the velocity field, the large velocity gradient problem can be analyzed along the lines developed here.

Discussions with D. J. Hollenbach, C. F. McKee, and D. A. Neufeld were instrumental for formulation of the concepts presented here. Special thanks to D. Neufeld for his persistent questions and for pointing out an error in an earlier version of the paper. I also thank W. H. Kegel who has brought to my attention just as this work was completed the early studies by Bettwieser & Kegel (1974) and by Bettwieser (1976), and M. M. Litvak for pointing out his 1973 paper. The support of NSF grant AST-9016810 is gratefully acknowledged.

## APPENDIX

### BEAMING IN SPECIAL MASERS

This appendix presents some detailed results for beaming angles of spherical masers, including the effect of suppression. Additional results can be found in a paper by Bettwieser (1976) that describes an attempt to construct a detailed, fully self-consistent solution of the spherical maser problem to high orders in  $J_{\nu}/J_s$ . The overall conclusions of both studies are in agreement.

As long as the center of the sphere, where the radiation is isotropic, is avoided, the beam pattern is assumed to follow the Gaussian  $e^{-\theta^2/\theta_B^2}$ , with  $\theta$  the direction from the radius vector. In addition, the radiative transfer equation always admits the formal solution  $I_{\nu} = S_0 e^{\tau_{\nu}}$ , and since  $S_0$  is isotropic, the angular variation of the intensity reflects the  $\theta$  dependence of the optical depth. Expanding  $\tau_{\nu}$  as a power series in  $\theta$ , the Gaussian profile emerges as the first term of this expansion and

$$\frac{1}{\theta_B^2} = - \left. \frac{d\tau_{\nu}(\theta)}{d\theta^2} \right|_{\theta=0}, \quad (\text{A1})$$

where  $\tau_v$  is calculated at a fixed inclination to the diameter (see also Litvak 1973). At the surface of the sphere,

$$\tau_v(\theta) = 2\kappa_{0v} R \int_0^{\sqrt{1-\theta^2}} (1 - e_v)(1 + e_v \sigma_v) dy, \quad (\text{A2})$$

where the suppression factor is written as  $s_v = e_v \sigma_v$  and  $y = (1/R)[r^2 - (R\theta)^2]^{1/2}$  is dimensionless distance along the path. At any point  $(r, \vartheta)$  along the path, the conversion efficiency  $e_v(r)$  and the suppression factor  $\sigma_v(r, \vartheta)$  are expressed here in terms of the variable  $y$  and the parameter  $\theta$  through the replacements  $\vartheta = \theta R/r$  and  $r = R(y^2 + \theta^2)^{1/2}$ .

Saturation ( $e_v \rightarrow 1$ ) diminishes the integrand, effectively confining the integration to unsaturated regimes. To leading order in  $J_v/J_s$ , the saturation factor  $1 - e_v = 1/(1 + J_v/J_s)$  is unity in unsaturated regions and zero in saturated regions. With this approximation in the integral for  $\tau_v$ , the dependence on  $\theta$  is confined to the upper limit of the integration. In an unsaturated sphere [ $J_v(R) < J_s$ ] the integration encompasses the entire path while in a saturated sphere [ $J_v(R) > J_s$ ] it is limited to the unsaturated core of radius  $r_{sv}$ , and the upper limit becomes  $[(r_{sv}/R)^2 - \theta^2]^{1/2}$ . Thus to zeroth order in  $J_v(R)/J_s$ , the beaming angle is

$$\theta_{B0}^2 = \frac{1}{\kappa_{0v} R} \times \begin{cases} 1 & J_v(R) < J_s \\ \frac{r_{sv}}{R} & J_v(R) > J_s \end{cases}, \quad (\text{A3})$$

the familiar standard result (Goldreich & Keeley 1972; Elitzur 1990). In spite of the assumptions used in its derivation, the beaming angle is a continuous function of  $R$ , except that its derivative suffers a discontinuity at the transition to saturation.

Suppression has no effect on the beaming angle to leading order in  $J_v/J_s$ . Its effects enter only at higher orders and thus cannot be evaluated before higher order corrections to the conversion efficiency are introduced. However, if we now attempt to calculate such corrections, we run into an immediate difficulty. Such terms involve not only the value of  $e_v$  at the upper limit of the integration but also its derivative along the path. Therefore, in order to calculate the beaming angle from equation (A1) to higher than leading order,  $e_v$  must be known to the next highest order. In an unsaturated region, expanding the saturation factor in a power series in  $J_v/J_s$  produces in first order  $1 - J_v/J_s$ , and the nonsensical result that the absorption coefficient vanishes at  $J_v = J_s$ . Indeed, an expansion in this variable is inadequate at any finite order. While such an expansion adequately describes the insignificant deviations from a pure exponential of growth away from  $S_0$ , it is meaningless as  $J_v$  approaches  $J_s$  where deviations from exponential must be expanded in a power series in  $1 - J_v/J_s$  instead. The problem is exacerbated by the fact that the optical depth is related to the intensity via  $\tau_v = \ln I_v - \ln S_0$ . During the approach to saturation, the intensity changes its dependence on path length from exponential to power law and  $\tau$  changes its dependence from linear to logarithmic, a dependence prone to meaningless results when expanded in a power series in the wrong variable.

If this problem is ignored, the beaming angle can be calculated to first order in  $J_v/J_s$  utilizing the fact that the variation of  $e_v(r)$  along the path is governed by the exponential growth of intensity along the diameter, which dominates all power law variations of  $\theta_B^2$ ,  $\delta^2$ , and  $\vartheta^2$ . Handling an insignificant logarithmic divergence at  $y \rightarrow 0$  with a cutoff, the result is

$$\theta_B^2 \simeq \frac{\theta_{B0}^2}{1 - 2e_v(R)[1 - \sigma_v(R, 0)]}. \quad (\text{A4})$$

And because  $e_v = \frac{1}{2}$  at saturation, this result produces  $\theta_B^2 \rightarrow \infty$  if extended to  $J_v \rightarrow J_s$  when suppression is neglected, a circuitous way to rediscover the vanishing of the absorption coefficient in this limit. Suppression slightly eases this problem at line center but does not remove the basic difficulty, which, rather than being a real singularity, merely reflects expansion in an inadequate variable. This explains the difficulties encountered by Neufeld (1992) in his attempt to estimate corrections to the beaming angle.

In the saturated regime the leading order correction to  $1 - e_v = 0$  is  $1 - e_v = J_v/J_s$ , which poses no problem at  $J_v = J_s$ ; the reason is that the transition to logarithmic variation of optical depth has already been accounted for [only the next order,  $1 - e_v = (1 - J_s/J_v)J_s/J_v$ , is meaningless]. We can therefore estimate the contributions to  $\tau_v$  from the standard solution in the saturated shell, which provides the leading power-law variation of  $J_v$  with path length. The result produces logarithmic corrections to the beaming angle; therefore, it is meaningless as long as (currently unknown) logarithmic terms in the variation of  $J_v$  are neglected. Another method for estimating beaming angles, presented next, shows that equation (A3) actually provides the correct expression for saturated masers to all orders of the saturation effect when suppression is neglected. Therefore, the logarithmic terms should cancel out in the saturated shell, and an estimate of the corrections induced by the standard solution should provide an indication of the importance of logarithmic corrections to this solution itself in the shell. A straightforward, if somewhat laborious calculation shows that, independent of  $\gamma$ , these corrections increase with  $R$  slowly and smoothly to 60% just before core saturation. The rise of logarithmic corrections as core saturation is approached reflects contributions of the parts of the shell close to the core and is similar to the role of logarithmic terms when an unsaturated sphere approaches saturation. How the transition to core saturation occurs is a problem that has not yet been addressed even for linear masers. This analysis suggests that logarithmic terms will have to be included in any attempt to study this transition.

Obviously, although equation (A1) provides a simple leading order derivation of the beaming angle, it is not the expression one would choose for the study of corrections to this result. A much more useful approach is the method of Goldreich & Keeley (1972) in which the beaming angle is evaluated from the ratio  $F_v/I_v$  of flux to dominant ray intensity. With this method the beaming angle is evaluated from the intensity itself, rather than the variation of its logarithm, at the same order of accuracy. The flux can be determined from the flux divergence relation, equation (4.8). When  $\sigma(\theta)$  contains only zeroth- and second-order terms in  $\theta$ , which is the case at the iteration order presented here, Gaussian angular integration produces  $\sigma(\theta) \rightarrow \sigma(\theta_B)$ . Therefore, for any geometry the

flux divergence relation for Gaussian beaming is

$$\nabla \cdot (fJ_v) = \kappa_v(\theta_B)J_v, \quad (\text{A5})$$

where  $f$  is the unit vector along the direction of the local flux vector. That is, the angle-averaged intensity  $J_v$  is determined by the absorption coefficient at  $\theta = \theta_B$  while the dominant ray intensity  $I_v(0)$  is determined by the absorption coefficient at  $\theta = 0$ . In spherical geometry the equations for  $I_v(r, 0)$  and  $J_v(r)$  can be combined to produce

$$\frac{1}{\theta_B} \frac{\partial \theta_B}{\partial r} = -\frac{1}{r} + \frac{1}{2} [\kappa_v(\theta_B) - \kappa_v(0)]. \quad (\text{A6})$$

Only suppression contributes to the difference between absorption coefficients across the beaming angle. Furthermore, only the terms proportional to  $\theta^2$  in the expression for  $s_v(r, \theta)$  contribute, and produce the same result for  $\delta$ -function and Gaussian angular distributions.

The equation for  $\theta_B$  can be solved treating suppression as a perturbation. The leading order solution (suppression is neglected) is

$$\theta_{B1}(r) = \theta_{B1}(r_i) \frac{r_i}{r}, \quad (\text{A7})$$

where  $r_i$  is an initial radius sufficiently removed from the center of the sphere that the radiation is beamed there. Since our interest is in saturated masers we can select  $r_i = r_{sv}$ , the core radius. When the beaming angle at the core surface is taken from the leading order result for unsaturated masers, the saturated case of equation (A3) is recovered at  $r = R$  as an exact result to all orders in  $J_v/J_s$  in the absence of suppression. Now allow for suppression by inserting a solution  $\theta_B = (1 + q)\theta_{B1}$  into the full equation for  $\theta_B$ . The solution for the suppression correction is

$$q(r) = \exp \left[ \int_1^{r/r_{sv}} \zeta(ur_{sv}) \frac{du}{u^2} \right] - 1, \quad (\text{A8})$$

where

$$\zeta(r) = \frac{1}{2} \left( \frac{e_{v0}}{\delta^2} \right) (1 - e_v) \left[ 1 - 2x^2 \left( 1 + \frac{1}{\delta^2} - \frac{2e_v}{\delta^2} \right) \right]. \quad (\text{A9})$$

This expression follows from the detailed result for the suppression function (eq. [3.30], with an additional factor 2 for Gaussian beaming), and except for the usage of the beaming angle from the first iteration in  $\zeta$  it is exact. Because  $\zeta(r)$  is a decreasing function of  $r$  in the saturated shell ( $r \geq r_{sv}$ ), an upper limit on the integral can be obtained by replacing  $\zeta(ur_{sv})$  with  $\zeta(r_{sv})$ . This produces an absolute upper bound on the suppression correction at the surface of the sphere

$$q(R) < \exp [\zeta(r_{sv})(1 - r_{sv}/R)] - 1, \quad (\text{A10})$$

which at line center becomes

$$q_0(R) < \frac{1}{8\delta^2(r_{sv})} \left( 1 - \frac{r_{sv}}{R} \right). \quad (\text{A11})$$

This correction is negligible as long as the linewidth at the core surface obeys  $\delta^2(r_{sv}) \sim 1$ . When the sphere just saturates,  $r_{sv} = R = R_{\text{sat}}$  and  $q_0(R)$  vanishes even though rebroadening to the Doppler profile is not yet complete. When  $R$  increases, the core shrinks and by the time  $r_{sv}/R \ll 1$ , line narrowing at the core boundary corresponds to  $\delta^2(r_{sv}) \sim 0.67$ . As a result,  $q_0(R)$  never exceeds  $\sim 0.2$ , and the effect of suppression on the beaming angle at the surface of the sphere is always negligible.

Suppression in the shell slightly increases the beaming angle at line center, an effect that can be readily understood. Amplification is the reason that maser radiation is beamed, and suppression effectively increases the unsaturated absorption coefficient at line center. And because the increase at the edge of the beam is larger than along its axis, the beaming angle opens up. The enhanced amplification also affects  $\theta_B(r_{sv})$ , the input to the shell solution, but this effect is even less significant. From equation (A3),  $1/\theta_B^2(r_{sv}) = \tau_s$ , the gain across the core radius. When the maser just saturates, this gain is determined purely by  $\gamma$  and thus is unaffected by suppression. As  $R$  increases,  $\tau_s \simeq \text{const} - 2 \ln(\kappa_{0v} R)$  and suppression causes a relative increase of  $\sim s(R)/\tau_s$  in  $\theta_B(r_{sv})$ , a correction at the  $\lesssim 1\%$  level.

Moving away from line center,  $q$  can be expected to have only a mild dependence on frequency in spite of the explicit dependence of  $\zeta$  on  $x$ . The reason is that this  $x$  dependence disappears when  $e_v \simeq \delta^2 \simeq 1$  and when the line saturates in the wings, it is definitely saturated at its center and re-broadening is complete. It seems safe to conclude that the standard result of eq. A3 can be used as is with reasonable confidence at least within  $|x| \leq 1$ .

## REFERENCES

- Alcock, C., & Ross, R. R. 1985, *ApJ*, 290, 433  
Bettwieser, E. 1976, *A&A*, 50, 231  
Bettwieser, E. V., & Kegel, W. H. 1974, *A&A*, 37, 291  
Boboltz, D., Simonetti, J. H., Diamond, P. J., Uphoff, J. A., & Dennison, B. 1993, in *Astrophysical Masers*, ed. G. Nedoluha & A. W. Clegg (Berlin: Springer-Verlag), 275  
Elitzur, M. 1990, *ApJ*, 363, 638  
———. 1992, *Astronomical Masers* (Dordrecht: Kluwer)  
Elitzur, M., Hollenbach, D. J., & McKee, C. F. 1989, *ApJ*, 346, 983  
———. 1992, *ApJ*, 394, 221  
Elitzur, M., McKee, C. F., & Hollenbach, D. J. 1991, *ApJ*, 367, 333  
Goldreich, P., & Keeley, D. A. 1972, *ApJ*, 174, 517  
Goldreich, P., Keeley, D. A., & Kwan, J. Y. 1973, *ApJ*, 179, 111  
Goldreich, P., & Kwan, J. 1974, *ApJ*, 190, 27  
Liljestrom, T. 1993, in *Astrophysical Masers*, ed. G. Nedoluha & A. W. Clegg (Berlin: Springer-Verlag), 291  
Litvak, M. M. 1970, *Phys. Rev.*, A2, 2107  
———. 1972, in *Atoms and Molecules in Astrophysics*, ed. T. R. Carson & M. J. Roberts (NY: Academic), 201  
———. 1973, *ApJ*, 182, 711  
Mattila, K., Holsti, N., Toriseva, M., Anttila, R., & Malkamaki, L. 1985, *A&A*, 145, 192  
Neufeld, D. 1992, *ApJ*, 393, L37  
Rowland, P. R., & Cohen, R. J. 1986, *MNRAS*, 200, 233  
Strelnitskij, V. S. 1982, *Soviet Astr. Lett.* 8, 86  
———. 1986, *Astr. Circ.*, 1465, 1  
———. 1993, in *Astrophysical Masers*, ed. G. Nedoluha & A. W. Clegg (Berlin: Springer-Verlag), 15