1977 Summer Student Lecture Interstellar Masers

Some History

Shklovskii (1949) predicted that absorption by interstellar OH in the ground-state Λ-doublet ought to be observable against background continuum sources. Unfortunately, his frequencies were wrong by almost a factor of two. Weinreb et al. (1963), using the laboratory measurements of Ehrenstein et al. (1959) [4 transitions, at 1612, 1665, 1667, and 1720 MHz, with relative line strengths 1:5:9:1], made the first astronomical detection of OH towards Cas A. It was found almost immediately that the relative intensities of the different OH transitions, particularly towards Sgr A, did not match the theoretical relative line strengths. While the data could be fit assuming clumps of moderate optical depth, Gardner et al. (1964) noted that "an alternate explanation, such as perturbations of the populations of the [energy] levels, cannot be excluded."

Gundermann (1965) (working with Goldstein) and Weaver et al. (1965) independently, and almost simultaneously, found strong emission lines at the OH frequencies towards the HII region W49. Because they only saw the strong emission at 1665 MHz, Weaver et al. attributed it to an unknown emitter, "mysterium", that was probably not OH. They noted with surprise that the narrow line widths implied kinetic temperatures much smaller than those of the HII region with which they were associated. Gundermann (1965) and Weinreb et al. (1965), however, detected the strong emission at other OH frequencies, and correctly identified it as "anomalously excited OH". Weinreb et al. noted that some of the 1665 emission was as much as 37% linearly polarized. Barrett and Rogers (1966) and Davies et al. (1966) found that the lines were circularly polarized. Invoking the Zeeman effect, Davies et al. deduced magnetic fields $\sim$ milligauss. But, Barrett and Rogers noted that the

velocity structure of the observed emission was so complex that the interpretation of specific features as Zeeman patterns would be hazardous. [This is still considered true.] More significantly, Barrett and Rogers determined an upper limit to the source size in W3(OH). This implied line brightness temperatures >2000°K, while the linewidths (if interpreted as thermal broadening) implied kinetic temperatures <50°K. From the fact that the lines were intense but narrow, Barrett and Rogers inferred that they were in maser emission. They considered that the anomalous hyperfine intensity ratios and polarization supported this interpretation, because it is easy to get such effects in laboratory masers. Later interferometric observations (Rogers et al. 1966; Cudaback et al. 1966) implied brightness temperatures $>2 \times 10^6$°K; at such large intensities, maser emission is essentially required.

Some Motivation

We now know that many interstellar molecules are masing, in the sense that the observed line intensities are enhanced because the radiation has propagated through regions of negative optical depth and negative excitation temperature. Three molecules produce fairly strong emission lines in many sources: OH, H_2O, and SiO. These masers are particularly interesting because: 1) They are fairly common astronomical objects. The number of type I OH masers in Turner's (as yet unpublished) survey is >400. By comparison, there are <200 known pulsars. Also, Wilson and Barrett (1972) estimate that ∿6% of all Mira variables are type IIb OH masers. For ∿4600 Miras catalogued (see Allen 1973), that implies >270 type IIb masers. (These classifications will be defined later.) 2) They are found in a wide range of astrophysical conditions: dark dust clouds, dense molecular clouds, protostellar condensations, circumstellar regions of young stars, circumstellar envelopes of evolved stars, stellar remnants.

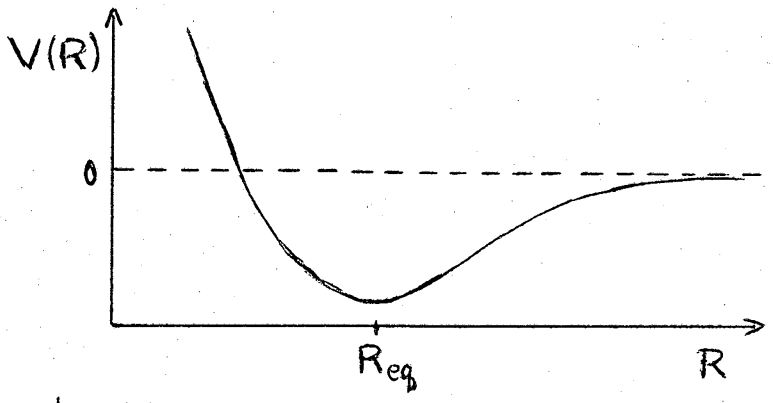
3) They seem to involve rather sophisticated physics. In particular, people who study the origin of OH maser polarization find themselves worrying about fundamental details of the interaction of strong radiation with matter.

4) They provide unique information about interesting regions. The excitation of many maser transtions (H_2O, excited OH, SiO) requires high densities and temperatures. (One would have expected to be able to study such regions only with presently unobservable far-infrared transitions.) The high intensities enable the study of regions of small angular size: e. g., a 1 a.u. source at 1 kpc suffers a beam-dilution factor of 4×10^{-10} when observed at 1.35 cm with a 100-m telescope; in order to be detectable at the 0.1°K level, it must have a brightness temperature of 2.5×10^{8}° K. (Which is to say that the only 1 a.u. sources that we could expect to detect are masers.)

Spectroscopic Structure of the Important Masers

In order to understand pumping schemes, the excitations and de-excitations that eventually produce an excess of molecules in the upper energy level of a masing transition ('population inversion'), a brief descent into molecular spectroscopy is de rigeur. Only the three strongest masers are discussed here (ordered by decreasing excitation of the masing transitions); for further details, see Townes and Schawlow (1955) or Gordy and Cook (1970).

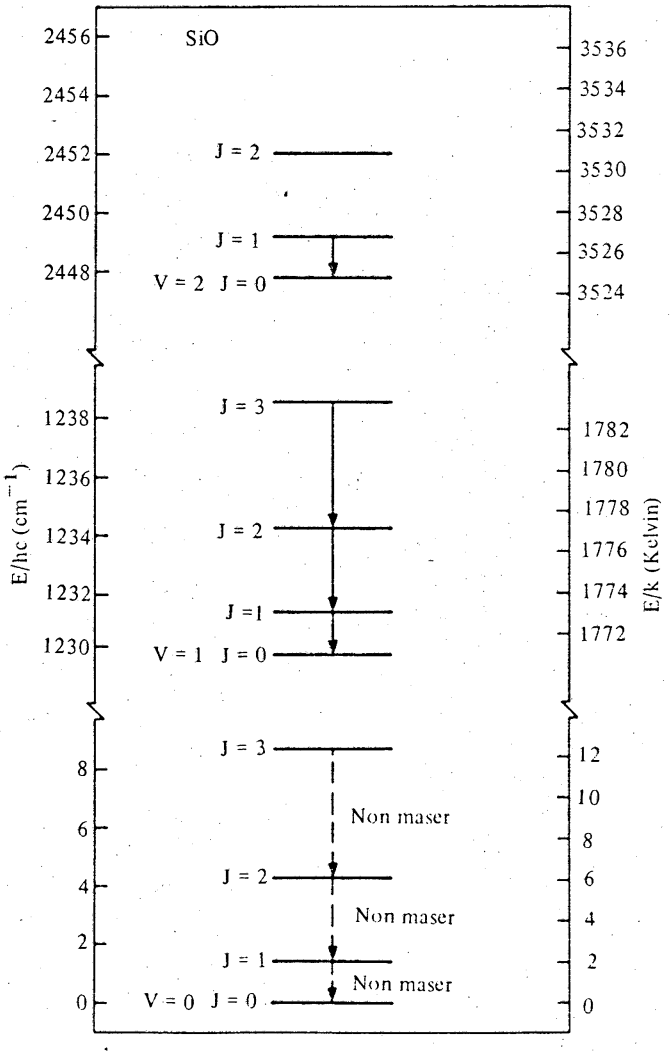
SiO is a simple diatomic molecule. It can be pictured as two nuclei embedded in a cloud of electrons. Each nucleus feels the electrostatic attraction of the electron cloud and the electrostatic repulsions of the other nucleus. The resulting potential of the ground electronic state, as a function of the distance between the nuclei, is:



For small $|R-R_{eq}|$, V(R) looks like a harmonic oscillator potential, so there is a sequence of vibrational energy levels. In addition, there is end-over-end rotation, the energy of which is small compared to that of vibrational motion. The rotation can thus be considered to be superimposed on each vibrational state. The resulting energy level scheme is shown in figure 1. The masing transitions are v = 1, J = 3 → 2, 2 → 1, 1 → 0 and v = 2, J = 1 → 0.

Figure 9-8. Some of the low rotational and vibrational energy levels of SiO showing the maser transitions which have been detected. There may be one case of maser emission in the $V = 0$, $J = 2 \rightarrow 1$ transition (Buhl et al., 1975).

Fig. 1



(from Moran 1976)

H_2O is an asymmetric molecule, which can rotate about three different axes with three unequal moments of inertia, so its rotational energy level structure is much more complex. Fortunately, it is only observed in the v = 0 state. Some of the energy levels are shown in figure 2. The masing transition is $6_{16} \rightarrow 5_{23}$.

Figure 9-10. Part of the rotational energy levels of H_2O, an asymmetrical rotator. The microwave maser transition is due to a chance proximity of the 6_{16} and 5_{23} levels. Data taken from de Jong (1973).

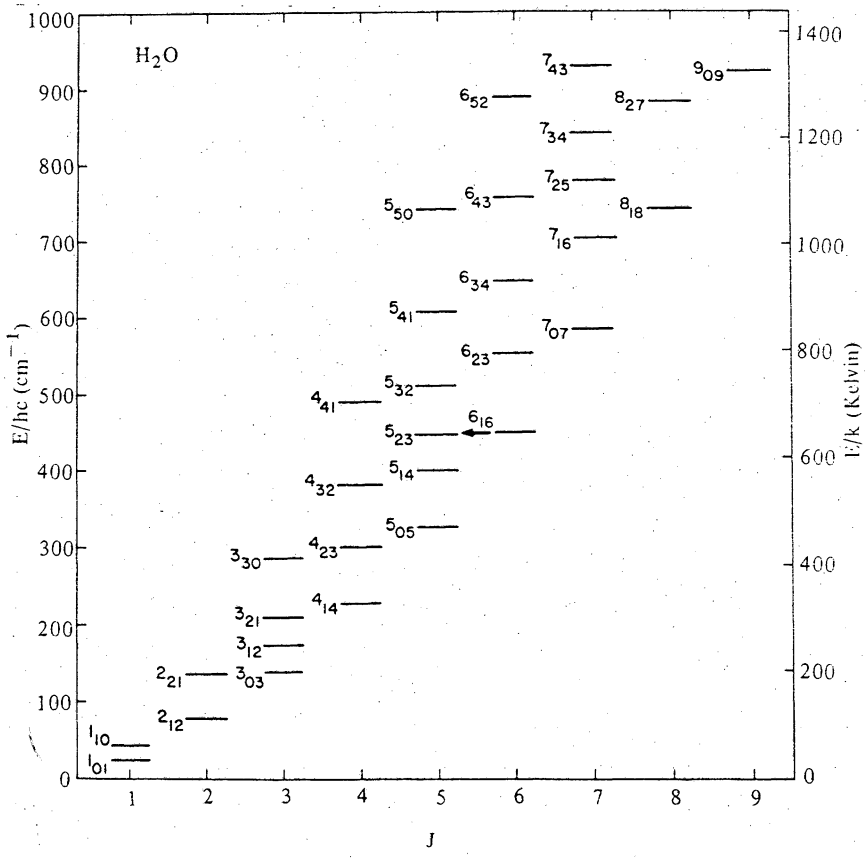


Fig. 2

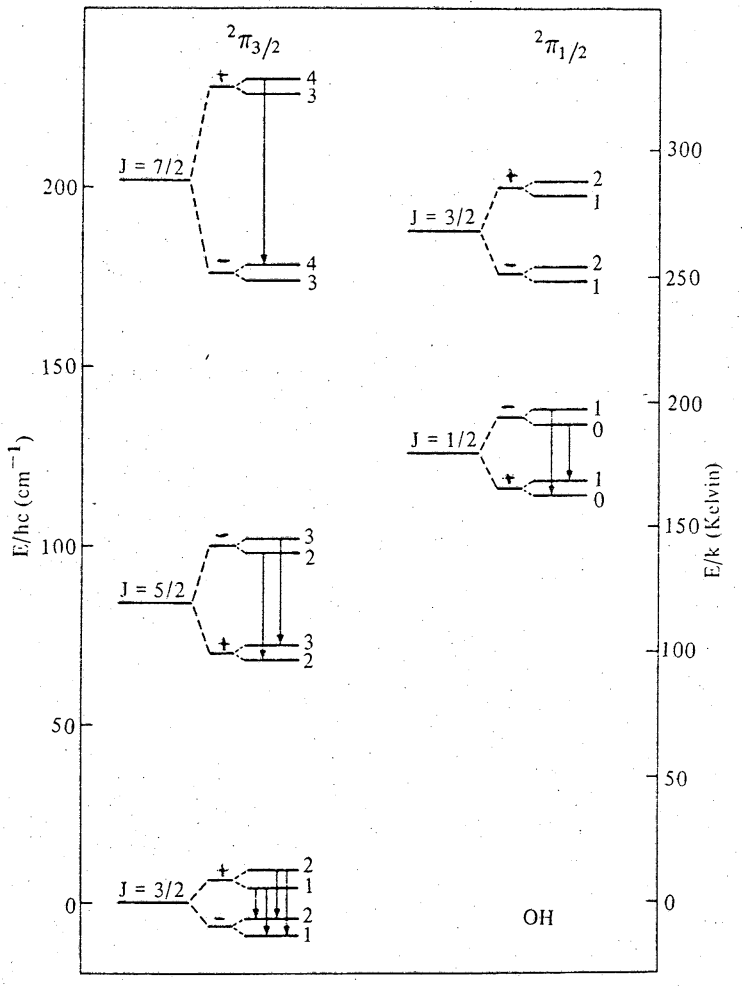
(from Moran 1976)

OH is a diatomic molecule, so it has a simple rotational spectrum. But the rotational transitions are in the far infrared (the first is at 120 μ). The radio-frequency transitions arise from more subtle interactions: (i) spin-orbit interaction splits the energy level structure into two rotational ladders, $^2\Pi_{3/2}$ and $^2\Pi_{1/2}$; (ii) Λ-doubling splits each rotational state into states of opposite parity; (iii) hyperfine interaction splits each Λ-doublet state

into two states of different total angular momentum, F. The transitions observed in maser emission, marked on figure 3, are: $^2\Pi_{3/2}$, J = 3/2, 2 → 2, 1 → 1, 2 → 1, 1 → 2; $^2\Pi_{3/2}$, J = 5/2, 3 → 3, 2 → 2; $^2\Pi_{3/2}$, J = 7/2, 4 → 4; and $^2\Pi_{1/2}$, J = 1/2, 1 → 0, 0 → 1.

Figure 9-9. Part of the rotational spectrum of OH. The rotational ladder has two branches due to spin splitting. The Λ doubling and hyperfine splitting which split each rotational level into four sublevels are not shown to scale. The number on the right side of each energy level is the total angular-momentum quantum number *F*. The known maser transitions are indicated.

Fig. 3



(from Moran 1976)

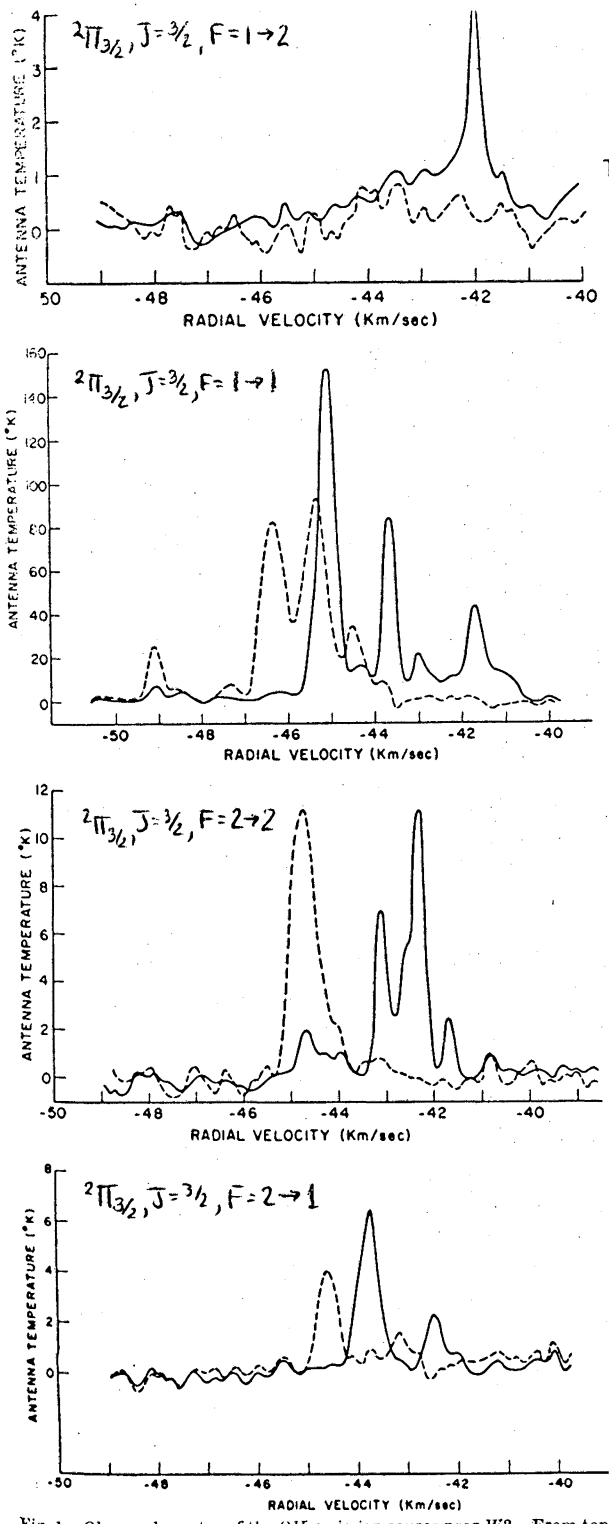
Two Main Types of Maser Sources

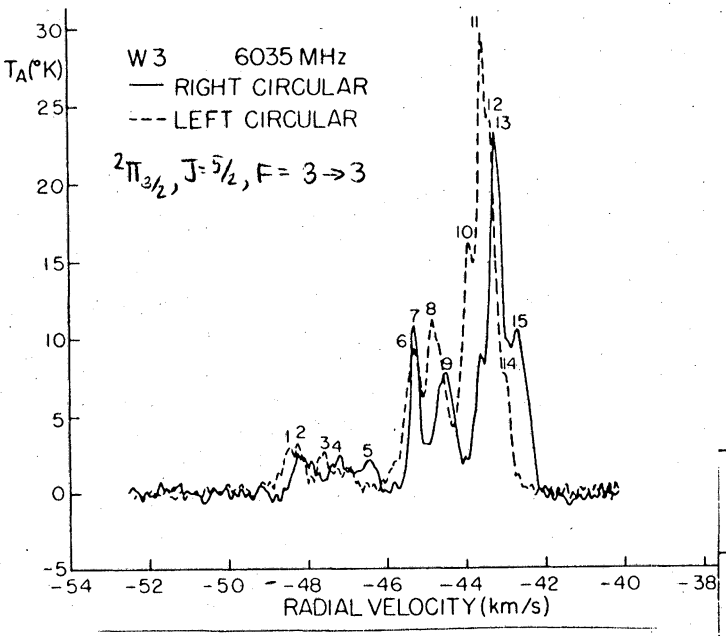
There is a general classification scheme for OH masers, based on their ground-state emission characteristics (Turner 1970). It turns out that two

categories correspond to the two types of astronomical source associated with most strong OH, H_2O, and SiO masers.

A) HII Region masers: Type I OH masers have strong emission at 1665 and 1667 MHz, with the 1665 lines almost always stronger than the 1667 lines. Emission at 1612 or 1720 MHz is usually weak or nonexistent. The OH spectra generally show many features with widths $\sim$0.1 to 1.0 km/s spread over a range $\lesssim$ 20 km/s. There are almost never any obvious velocity correspondences between transitions. The features are usually circularly polarized, perhaps as much as 100%; linear polarization is also seen, but much less often. Some features vary on timescales $\gtrsim$ months. A typical spectrum, for W3(OH), is shown in figure 4.

Figure 4 (next page): Maser emission from W3(OH). The ground-state ($^2\Pi_{3/2}$,J=3/2) spectra are from Barrett and Rogers (1966). The excited-state ($^2\Pi_{3/2}$,J=5/2) spectra are from Zuckerman et al. (1972, Ap. J., 177, 59.) The H_2O spectra are from Sullivan, W. T. III (1973, Ap. J. Suppl., 25, 393.)





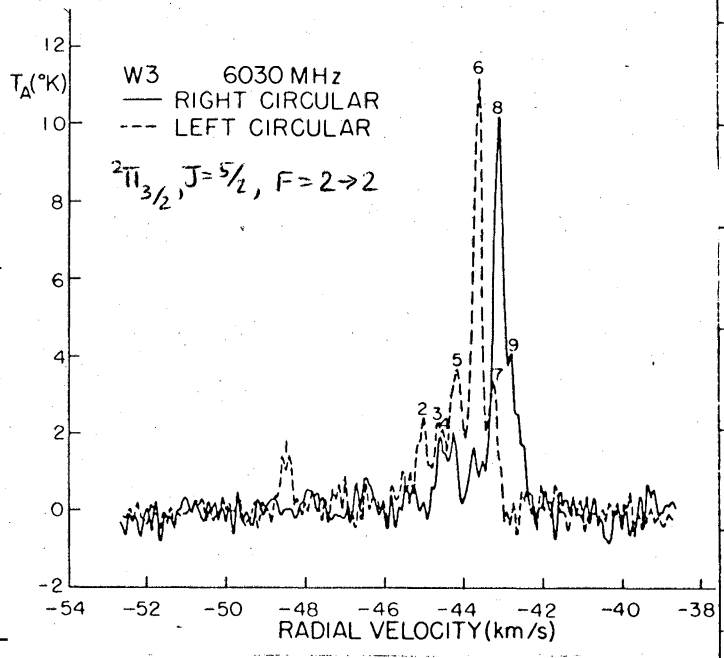
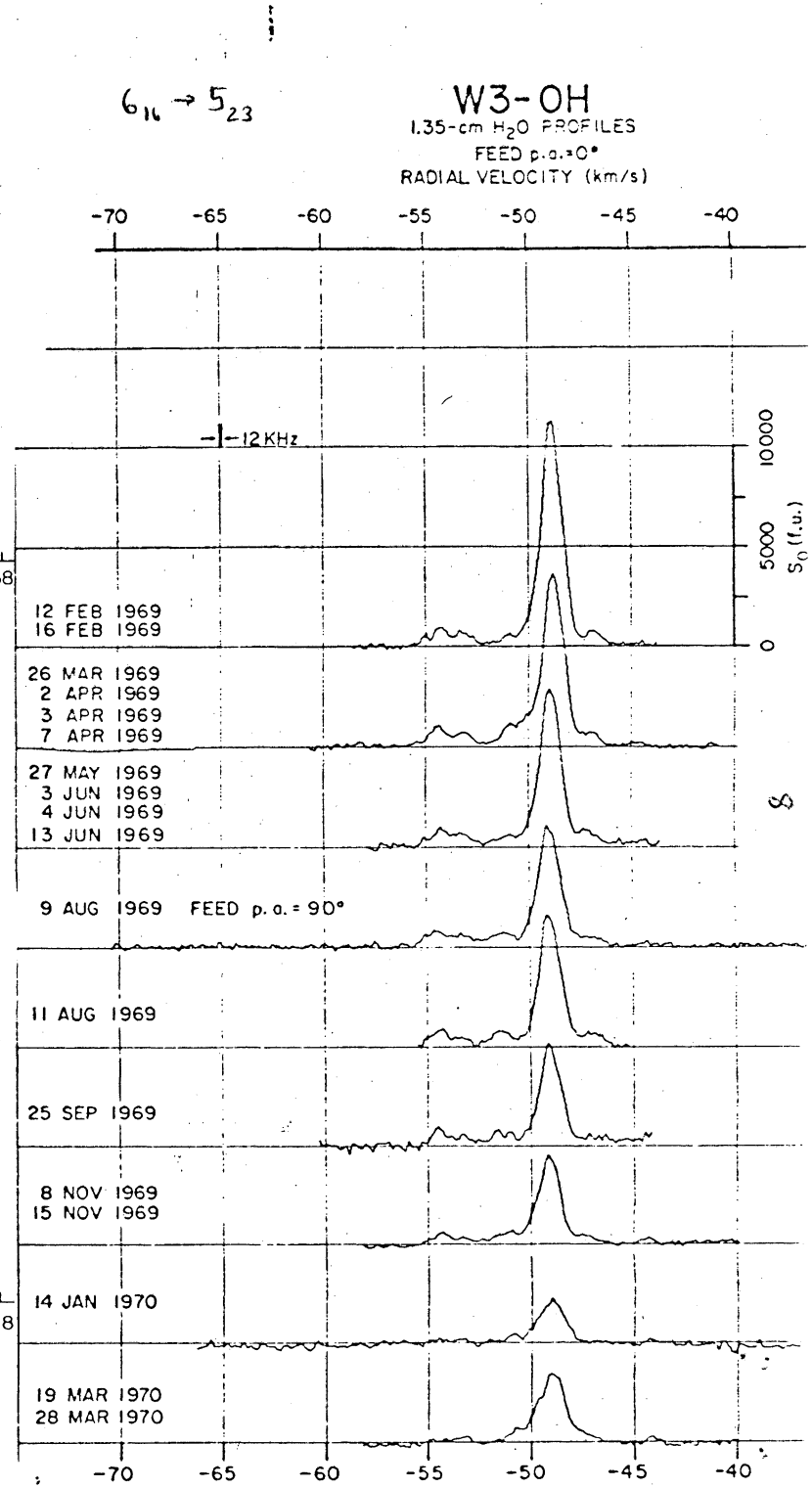


Fig. 1. Observed spectra of the OH-emission source near W3. From top to bottom, the frequencies are 1,612, 1,665, 1,667, and 1,720 Mc/s. The solid line shows the right-hand circularly polarized component, and the dashed line shows the left-hand. Note that the velocity scales are slightly different. The frequency resolution is 1·25 kc/s



The strongest type I sources also have rotationally-excited OH masers, generally showing fewer velocity features. Circular and linear polarization are seen as are slow time variations. A few rapid variations, with timescales $\sim$ 1 day, have also been observed.

Most type I OH masers, when checked to a sufficiently low level, are also found to be H_2O masers. The H_2O spectra are often very complex, occasionally showing features over a wide velocity range (> 300 km/s). An example, W49 is shown in figure 5. The H_2O features are never circularly polarized, but are often linearly polarized by as much as 40%. Individual features are often variable, with timescales as short as days. There is some evidence that the percent of polarization of particular features can also vary.

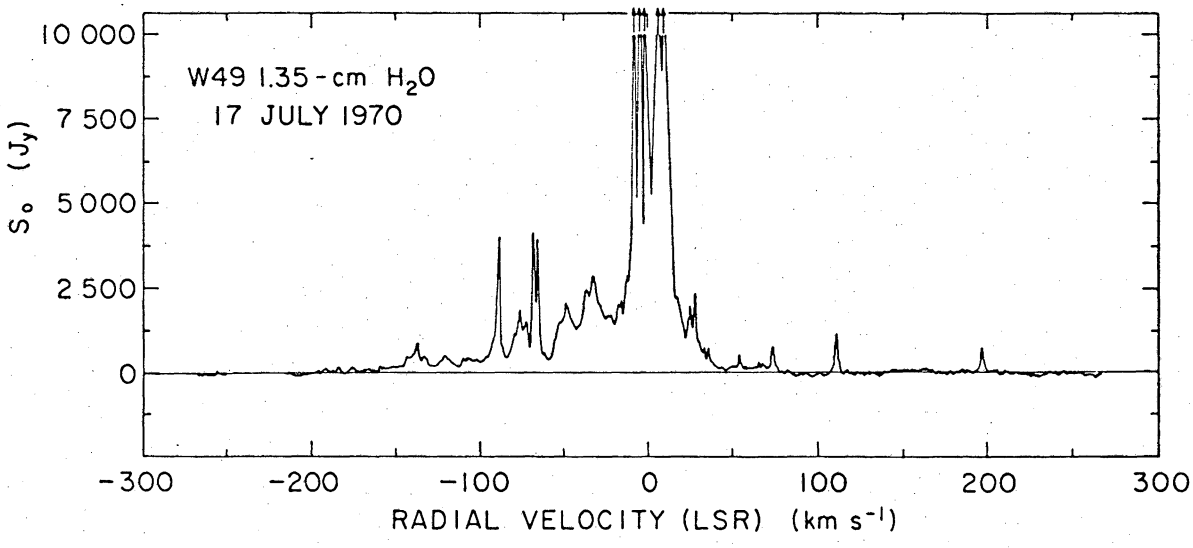


FIGURE 1 Typical 1.35 cm H_2O profile of W49 covering a wide velocity range and with a resolution of 0.7 km s^{-1}. The most intense low velocity features have been truncated in order to show better the high velocity features. The rms noise in the profile is $\sim$25 Jy. These data were obtained with the 85 ft radio telescope of the Naval Research Laboratory at Maryland Point, Md.

Fig. 5 : taken from Heckman and Sullivan (1976, Ap. Lett., 17, 105.)

When observed with an interferometer, both OH and H_2O emission features are found to arise from small spots distributed through larger region (see figure 6).

Figure 9-2. A map of the spectral components, labeled by velocity, of the water vapor maser in W49. The error bars show the relative measurement accuracy with respect to the −1.8 km/sec component. The position of the map is uncertain by about 1 arcsec. The sizes of the components vary between 0.0003 and 0.003 arcsec. This map was made from an analysis of the relative rates of change of fringe phase (equation 9-24) in a VLBI experiment with an antenna spacing of 845 km (Moran et al., 1973).

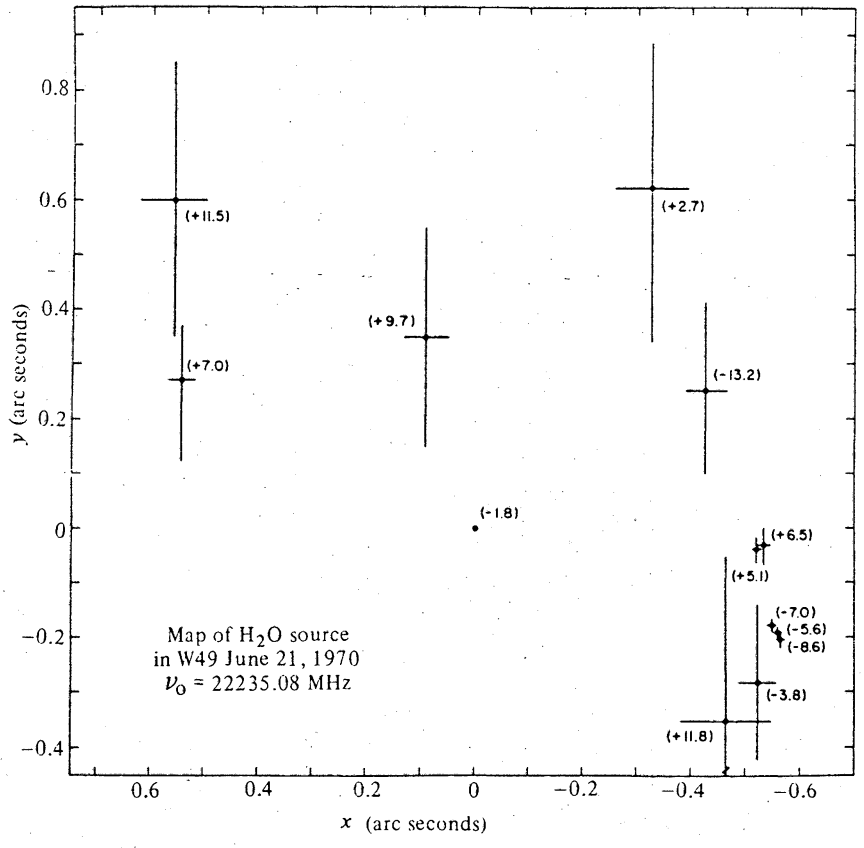
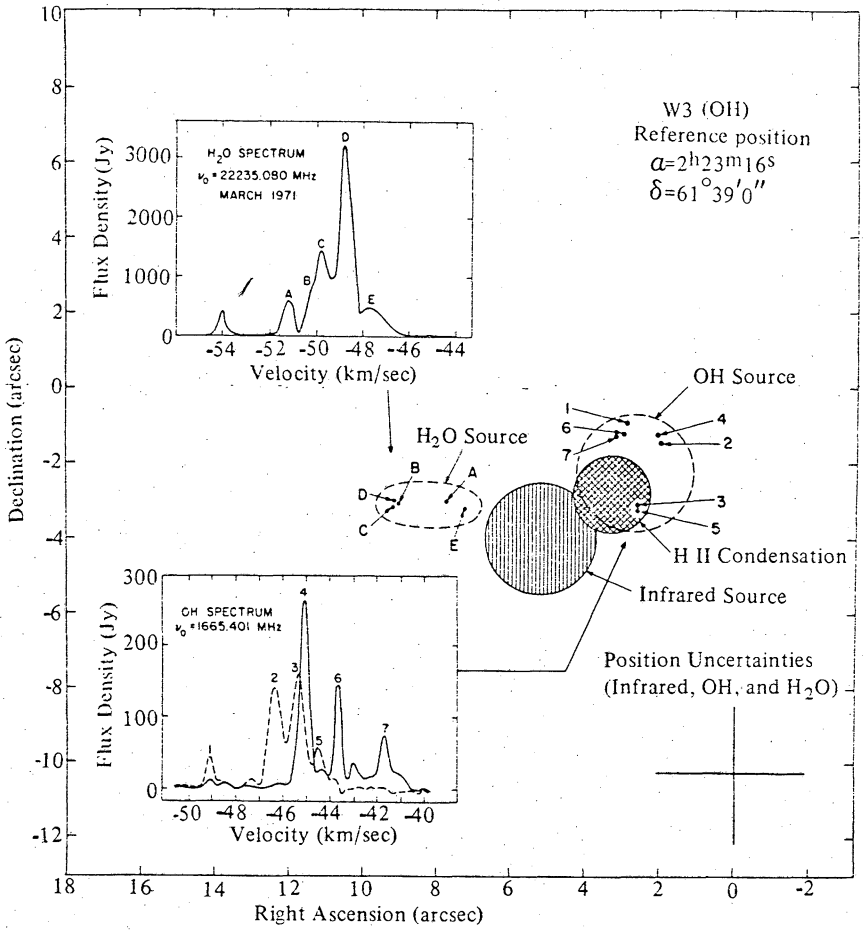


Fig. 6 : taken from Moran(1976)

Each spot corresponds to a different velocity feature. The OH spots have diameters $\sim 10^{14}$ cm, the H_2O spots $\sim 10^{13}$ cm. Both are distributed over regions of diameter $\sim 10^{16}$ cm but, in the few cases where the observations give accurate absolute positions, the OH and H_2O regions don't overlap. About half the time, the masers are found to be near (but not coincident with) compact HII regions (continuum knots ∿0.1 to 0.01 pc in size, with $n_e \sim 10^{5-6}$ cm^{-3}). (For W3(OH), it is possible that the OH spots are in a shell around the compact HII region.) Again about half the time, the masers are also associated with compact infrared sources (see figure 7).

Figure 9-17. The region near W3(OH) showing the infrared source IRS9 (Wynn Williams, Becklin, and Neugebaurer, 1972), the compact H II region (Baldwin et al., 1973), the H_2O maser (Hills et al., 1972), and OH maser (Mader et al., 1975). The errors on the position of the H II condensation are about 0.1 arcsec. The uncertainties in position of the infrared, OH, and H_2O sources are about 2 arcsec. The origin of the coordinates is located at right ascension (1950) $2^h23^m16^s$ and declination (1950) 61°39′0″.

Fig. 7



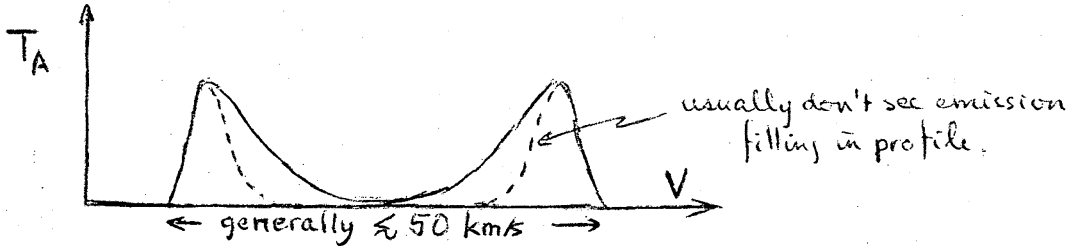
taken from Moran (1976)

If the source sizes are as measured by the interferometers (which is not necessarily ture, as will be discussed later), they imply brightness temperatures $T_B(OH) = 10^{12-13}$° K, $T_B(H_2O) = 10^{13-15}$° K, and luminosities $L(OH) = 10^{25-29}$ erg/s, $L(H_2O) = 10^{28-32}$ erg/s. Note that these figures refer to individual velocity features.

Arguing from the statistics of type I OH/H_2O masers (association with HII, IR; sizes and luminosities; pump models [discussed later]; etc.) a general picture of the evolution of HII region masers has been sketched by some authors (e.g., Lo 1974, Habing et al. 1974): A massive star forms within a dense dust cocoon. The heated dust pumps maser emission in the surrounding H_2O. As the star's UV radiation ionizes the cocoon (to form a detectable

compact HII region), it dissociates the H_2O, producing OH which is also pumped into maser emission. The HII region grows and eventually (after $\sim 10^4$ yr) destroys the OH as well. This picture leaves many questions unanswered; e.g., the excited-state OH masers cannot be fit into this scheme in any systematic way.

B) Circumstellar Masers: Type IIb OH masers have strong emission at 1612 MHz, occasionally weaker emission at 1665 and 1667 MHz, and no 1720 MHz emission. The 1612 MHz emission usually has a distinctive line shape,



and is usually unpolarized. Most type IIb OH masers also show H_2O maser emission , and ∿20% have SiO maser emission. The H_2O and SiO features generally fall at velocities within the OH pattern.

These masers are associated with late-type stars that have strong near-IR excesses. These are generally Mira variables or irregular supergiant variables with oxygen-rich atmospheres, typical effective temperatures ∿1800-2800°K and total luminosities $\sim 10^4 L_\odot$. The strong emission over 3-20 μ implies the presence of a dust shell with a temperature ∿600-800° K. The stellar velocities are difficult to determine, but are probably always within the OH pattern.

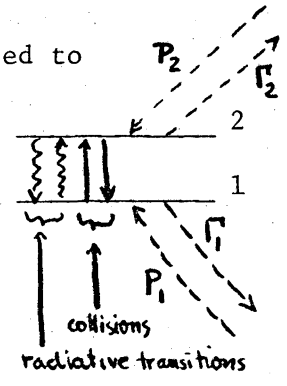
Interferometer maps have been made for a few sources; they again show small spots. The OH spots have diameters $\sim 10^{15}$ cm (luminosities $\sim 10^{24-28}$ erg/s) and are distributed over a region of size $\sim 3 \times 10^{16}$ cm. The H_2O spots have diameters $\sim 10^{14}$ cm (luminosities $\sim 10^{24-28}$ erg/s) and are distributed over $\sim 10^{15}$ cm.

The type IIb masers are often variable, but in a manner different from the HII region masers. The OH and H_2O flux varies in a rough sinousoidal fashion as does the infrared continuum. The periods are 10^{2-3} days long, and the maser maxima lag the IR maxima by $\sim$0.1 - 0.2 periods. This is interpreted as radiative coupling between the stellar flux and the maser, presumably through the pumping mechanism.

The Elitzur et al. (1974) model is fairly typical among those proposed for type IIb masers. The stellar mass loss includes dust and H_2O. The dust is heated by the star, and radiates at 6μ, exciting the H_2O, $v = 0 \rightarrow 1$ transition and pumping the H_2O maser. As the circumstellar cloud flows outward, the ambient interstellar UV field dissociates the H_2O, producing OH, which is also pumped by the near-IR emission of the dust. In other models (e.g., Harvey et al. 1974; Schwartz et al. 1974), there is a similar configuration, but the molecules are pumped directly by stellar radiation.

Basic Radiative Transfer for Masers

Consider the energy levels of some maser transition, connected to other (unspecified) energy levels (including continuum states) in unspecified ways.



Statistical equilibrium:

$$\frac{dn_1}{dt} = An_2 + B_{21}In_2 - B_{12}In_1 + C_{21}n_2 - C_{12}n_1 + P_1N - \Gamma_1 n_1$$

$$\frac{dn_2}{dt} = \underbrace{B_{12}In_1 - An_2 - B_{21}In_2}_{\text{Radiative transitions between levels 1 + 2}} + \underbrace{C_{12}n_1 - C_{21}n_2}_{\text{Collisional transitions between 1 + 2}} + \underbrace{P_2N}_{\text{All input from other levels}} - \underbrace{\Gamma_2 n_2}_{\text{All outflow to other levels}}$$

Here $N = n_1 + n_2$ is proportional to the total number of masing molecules.

Assuming steady state: $\frac{dn_1}{dt} = \frac{dn_2}{dt} = 0$, and after tedious algebra,

$$\Delta n \equiv n_2 - n_1 \equiv N \frac{(P_2-P_1) + 1/2(\Gamma_1-\Gamma_2) - [A + B_{21}I - B_{12}I] - [C_{21}-C_{12}]}{(A + B_{21}I + B_{12}I) + (C_{21} + C_{12}) + 1/2\ (\Gamma_1 + \Gamma_2)}$$

For $\Delta n > 0$, the energy level populations are inverted, and we have a maser.

Let $\Delta W_r = A + B_{21}I - B_{12}I$, $\Delta W_c = C_{21} - C_{12}$ $\quad (\Delta W_r,\ \Delta W_c \geq 0)$

$$W_{rt} = A + B_{21}I + B_{12}I, \quad W_{ct} = C_{21} + C_{12}$$

$$\Delta n = N \frac{(P_2-P_1) + 1/2(\Gamma_1-\Gamma_2) - \Delta W_r - \Delta W_c}{W_{rt} + W_{ct} + 1/2(\Gamma_1 + \Gamma_2)}$$

Note that if levels 1 and 2 are the only energy levels of the molecule ($P_1=P_2=\Gamma_1=\Gamma_2=0$), then Δn is always ≤ 0, i.e. we can never have a maser. This is just a consequence of the 2nd Law of Thermodynamics. The two levels have a unique temperature (the excitation temperature T_{ex}) that you can change by connecting up with various reservoirs (radiation fields, colliding molecules, etc.) at various temperatures. But you can never cool Tex below the temperature of the coolest reservoir which, for the physical situations available, will be > 0.

For convenience, let $P_{eff} \equiv (P_2 - P_1) + 1/2(\Gamma_1 - \Gamma_2) - \Delta W_r - \Delta W_c$ (the effective pump rate). Furthermore, assume equal statistical weights for the 2 levels ($B_{12} = B_{21}$) and a rectangular line of width $\Delta\nu$. The one dimensional equation of transfer becomes:

$$\frac{dI}{dx} = \frac{h\nu}{4\pi\Delta\nu} [\Delta n B I + n_2 A]$$

$$= I \frac{h\nu B}{4\pi\Delta\nu} \frac{NP_{eff}}{W_{ct}+A+1/2(\Gamma_1+\Gamma_2)} \frac{1}{1 + \frac{2BI}{W_{ct}+A+1/2(\Gamma_1+\Gamma_2)}} + \frac{h\nu n_2 A}{4\pi\Delta\nu}$$

$$= \frac{\alpha_o I}{1 + I/I_s} + \varepsilon$$

Note that $\alpha_o I = \frac{h\nu}{8\pi\Delta\nu} NP_{eff} \frac{I}{I_s}$ and $\Delta n = \frac{(NP_{eff}/2BI_s)}{1 + I/I_s}$

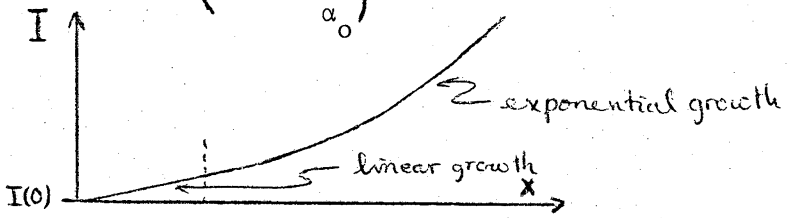
We can isolate two regimes:

A) Unsaturated amplification: $I/I_s << 1$

Here $\frac{dI}{dx} \simeq \alpha_o I + \varepsilon$, so $I(x) = I(0)e^{\alpha_o x} + \frac{\varepsilon}{\alpha_o}(e^{\alpha_o x} - 1)$

For $\alpha_o x$ small, $I(x) \simeq I(0) + \varepsilon x$ (linear growth)

As $\alpha_o x$ becomes larger, $I(x) \simeq \left(I(0) + \frac{\varepsilon}{\alpha_o}\right) e^{\alpha_o x}$ (exponential growth)

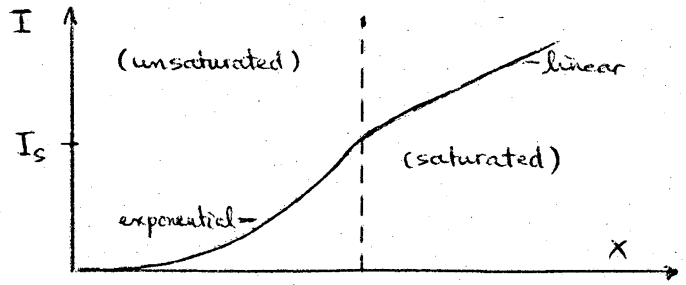


So, when the initial radiation field - consisting both of emission from an external source (like a nearly HII region), $I(0)$, and of spontaneous emission within the maser, ε - tranverses a pathlength L, it is amplified by a factor $\sim e^G$, where $G = \alpha_o L$ = the gain. Note that $\Delta n \simeq \frac{NP_{eff}}{2BI_s}$, so that the inversion is completely controlled by the pumping mechanism.

B) Saturated amplification: $I/I_s >> 1$

Here $\frac{dI}{dx} \simeq \alpha_o I_s + \varepsilon$, so $I(x) = I(0) + \alpha_o I_s x + \varepsilon x$ (linear growth)

Note that $\Delta n \simeq \frac{NP_{eff}}{2BI}$. The radiation field is beginning to compete with the pump, trying to destroy the inversion by stimulating the $2 \rightarrow 1$ transition.



Consider this another way: at $I = I_s$,

$$\frac{dI}{dx} = \frac{1}{2}\alpha_o I + \epsilon = \frac{h\nu}{16\pi\Delta\nu} NP_{eff} + \varepsilon$$

Neglecting ε (which is small compared to I_s), this means that as the pump energy is put in, it is immediately converted into the energy of the radiation field (modulo a factor of 1/4 at $I = I_s$), and can't be used to build up the inversion any more.

The distinction between these two regimes has observable effects

i) Line narrowing

Suppose the initial radiation had a Doppler profile:

$$N = N(\nu) = N_o \exp\left[-4 \ln 2 \frac{(\nu-\nu_o)^2}{\Delta\nu^2}\right]$$

If you plug this into the unsaturated maser equation, the resulting profile is roughly Gaussian, but with width

$$\Delta\nu' \simeq \frac{\Delta\nu}{\sqrt{G_o}}, \quad G_o = \alpha_o x \text{ evaluated at } \nu = \nu_o.$$

The stronger parts of the line (the center) are amplified more ($\frac{dI}{dx} \propto I$), so the line narrows. For $G_o \sim 25$, you can make a line with a Doppler width corresponding to 10^3 °K look like a line of Doppler width corresponding to <50 ° K.

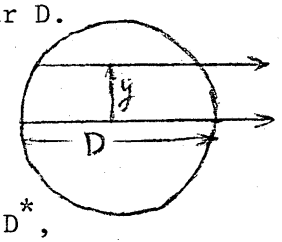
But if the maser is saturated, the line core grows linearly while the line wings (still unsaturated) grow exponentially and catch up. For a strongly saturated maser, the line is braodened out to its original width. [However, Goldreich and Kwan (1974a) have a mechanism that will keep the line from broadening in saturated masers.]

ii) Time variations

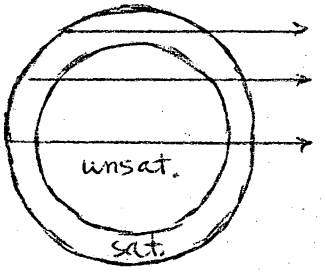
Neglecting ε, $I = I_o e^G$ for an unsaturated maser, so $\delta I = I_o e^G \delta G$, i.e. $\frac{\delta I}{I} = \delta G$; a 1% change in G produces a G% change in I. For a saturated maser, though, $I = I_o + GI_s$, $\delta I = I_s \delta G$, and $\frac{\delta I}{I} = \frac{I_s \delta G}{I_o + GI_s} = \frac{\delta G}{G} \frac{1}{1 + I_o/I_s G} \leq \frac{\delta G}{G}$; a 1% change in G produces <u>at most</u> a 1% change in I.

iii) Apparent source size (a rough description)

Consider two path lengths through a spherical maser of diameter D. The path through the diameter has length D, the chord has length $D\sqrt{1-4y^2/D^2}$, so $\frac{G(y)}{G(0)} = \sqrt{1 - 4y^2/D^2}$, and $\frac{I(y)}{I(0)} \propto \exp\,[G(0)(\sqrt{1-4y^2/D^2} - 1)]$. Thus the apparent source size, D^*, will be smaller than the true source size D: $\frac{D^*}{D} \simeq \sqrt{\frac{2\ln 2}{G(0)}}$.



As G(0) increases, the apparent source size decreases. Eventually G(0) is sufficiently high so that the intensity (at the surface) of radiation moving along the diameters (not necessarily toward the observer) saturates the edge. A line of sight through the saturated edge shows only linearly amplified background radiation and spontaneous emission. The saturated edge is essentially invisible in comparison with the unsaturated core. So D* is further reduced. When the maser becomes saturated throughout, the apparent size increases again, because $\frac{I(y)}{I(0)} = \sqrt{1 - 4y^2/D^2}$. In this limit, $\frac{D^*}{D} = \frac{\sqrt{3}}{2}$. Also note: because the gains decrease for directions away from the diameter, the radiation in any particular direction becomes strongly beamed [hence, the source size effect]. So at the cloud surface, the radiation field is very anisotropic. This may be a very important affect in more realistic analyses of the radiative transfer.



iv) Statistics.

For simple spontaneous emission, the observed radiation arrives with random phases, and the total electric field amplitude observed at a given instant of time is an average over these different phases. Because of the large number of sources, the probability of observing some amplitude $|\vec{E}|$ is Gaussian distributed:

$$P\,(|\vec{E}|) = \frac{1}{\sqrt{2\pi}} \exp\left\{ - \frac{|\vec{E}|^2}{2\sigma^2} \right\}.$$

In a laboratory maser with feedback, the phases are all aligned, because the radiation is essentially all stimulated emission, having the same phase and direction as the stimulating radiation.

Astrophysical masers are intermediate. For an unsaturated maser, you expect Gaussian statistics because the maser is amplifying the random phase distribution of the incident radiation field or the internal spontaneous emission. But with the onset of saturation, there are nonlinear effects: the gain is smaller for signals with stronger $|\vec{E}|$. So the probability of observing the larger values of $|\vec{E}|$ is decreased below the Gaussian value.

So are interstellar masers saturated or unsaturated? There are some tests:

i) Measure I (i.e., T_B) and compare it with I_S. Doing so, it turns out that the smallest sources (small H_2O spots) are just barely saturated. Since we expect that the physical sizes are even larger, it seems likely that all OH and H_2O masers are saturated.

ii) Sources often show several velocity components with comparable intensities. If the sources were unsaturated, small differences in G would mean large differences in I, and we wouldn't expect so many sources to have similar intensities.

iii) Look at the statistics. Evans et al. (1972) did so, and found no detectable deviation from Gaussian statistics. This would suggest unsaturated masers, but:

(a) scattering of signals by the interveneing interstellar medium may wash out expected deviations, and

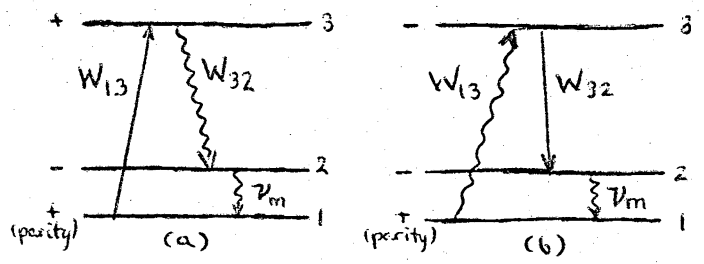
(b) even with interferometers, we may be averaging over many regions of (different) constant phase.

[We can define a correlation length $\frac{4\pi\Delta\nu}{h\nu Bn_{OH}} = \frac{8\pi\Delta\nu}{\lambda^2 An_{OH}} \sim 2 \times 10^{11}$ cm, for $A \sim 10^{-10} s^{-1}$, $n_{OH} \sim 10$ cm^{-3}. This implies a constant phase area $\sim 4 \times 10^{22}$ cm^2, vs. a spot area $\sim 10^{26}$ cm^2. So there's room for 2500 phase patches, quite enough to produce Gaussian statistics via the central limit theorem.]

Pump Models

There are many OH, H_2O, and SiO pump models; were we to discuss the advantages and disadvantages of each, these notes would rapidly become tedious. So we shall concentrate on general pump theory, especially thermodynamics.

Consider these simple three-level pumps, W_{13} and W_{32} correspond to energy reservoirs at different temperatures.



a) W_{13}: collisions with H_2 molecules at high $T_{kinetic}$

W_{32}: radiation into field with $T_{rad} << T_{kin}$ (e.g. 3°K background)

b) W_{13}: radiation excitation by field at high T_R (e.g. IR from hot dust)

W_{32}: collisional deexcitation by H_2 molecules at low T_{kin}

Using the Boltzmann formula and writing $\frac{n_2}{n_1} = \frac{n_3/n_1}{n_3/n_2}$, we derive:

i) an approximate relation: $\frac{1}{T_{32}} - \frac{1}{T_{31}} \simeq \frac{1}{T_{21}} \frac{\nu_m}{\nu_p}$,

where T_{32}, T_{31}, and T_{21} are the excitation temperatures of the 3 pairs of levels and $\nu_{32} \simeq \nu_{13} = \nu_p$. So to get inversion ($T_{21} < 0$), we need $T_{31} > T_{32}$, as written for examples (a) and (b), The efficiency of energy transfer between pump and maser, $\eta_m = \nu_m/\nu_p$.

ii) more carefully: $\frac{n_2}{n_1} = \exp\left[\frac{h\nu_m}{kT_{32}}\left(\underbrace{\frac{\nu_p}{\nu_m}}_{\eta_m^{-1}} \underbrace{\frac{T_{31} - T_{32}}{T_{31}}}_{\eta_c} - 1\right)\right]$

(cf, Scovil et al. 1959).

η_c is just the efficiency of the Carnot cycle describing the pump process. For inversion ($\frac{n_2}{n_1} \geq 1$) we must have $\eta_m \leq \eta_c$. So the maser is almost never as efficient as the pump. Furthermore, the most efficient maser is one that is very saturated ($\frac{n_2}{n_1} \to 1$).

The presence of a radiative transition in the pump cycle leads to a constraint on the process. If the optical depth in the radiative transition becomes large, the radiation field in the line is thermalized to the kinetic temperature; i.e. $T_{32} \simeq T_{31}$, and we can't sustain an inversion. The resulting constraint is that the rate of photon leakage in the line be sufficiently fast (at least as fast as the maser transition rate). The maximum leakage rate is $\Lambda_{max} = \rho \frac{c\Delta\nu}{4\pi h\nu} 4\pi r^2_{cloud}$, $\rho = \frac{8\pi h\nu^3/c^3}{e^{(h\nu/kT_B)}-1}$, T_B = line brightness temperature. This often turns out to be a crucial factor in the ability of a hypothetical pump to work in a real cloud. For example, the -3.5 km/s H_2O feature in W49 (very famous because $L_{H_2O} \sim 1 L_\odot$) radiates $\sim 5 \times 10^{48}$ photons/s, and has an apparent size $\sim 3 \times 10^{13}$ cm. If the radiative part of the pump works at $\sim 100\mu$, with $T_B \sim 500°K$ and $\Delta\nu/\nu \sim 7 \times 10^{-6}$ (typical values from a far-IR pump model), $\Lambda_{max} \sim 5 \times 10^{40}$ photons/s. This misses having maser rate $\leq$ leakage rate by a factor of $\sim 10^8$. We can try to get around this by invoking the fact that the physical cloud size is larger than the apparent one, but we'd need $(D^*/D) \sim 10^{-4}$ (quite unlikely).

The usual way that people get around this is by distributing the radiation source (or sink) throughout the masing material. Consider, e.g., the Goldreich and Kwan (1974b) H_2O pump. Dust is heated by a continuum source (UV from an HII region, near-IR from a star) and emits strongly in the near infrared. It emits strongly at 6μ, exciting nearby H_2O to the v = 1 state. The excited H_2O is collisionally deexcited by cool H_2 molecules ($T_{gas} < T_{dust}$), and the

$6_{16} - 5_{23}$ transition ends up being inverted. Both the pump (hot dust radiation) and the sink (cool gas collisions) are mixed with the masing H_2O, so the optical depths can be kept low. Providing that the timescale for collisional equilibration of T_{gas} and T_{dust} is long enough, Goldreich and Kwan can get a maser.

The Elitzur et al. (1976) 1612 MHz maser pump is similar. Warm dust grains excite the $^2\Pi_{3/2}$, J = 3/2 → $^2\Pi_{1/2}$, J = 5/2 transition at 35μ. The excited states radiatively decay, with small optical depths for the most part. The details of the decay from $^2\Pi_{1/2}$, J = 1/2 to $^2\Pi_{3/2}$, J = 3/2 lead to an inversion of the hyperfine levels within each Λ-doublet state, sufficient to bring about 1612 MHz inversion. The pump is again distributed through the maser. We do have to worry about the leakage rate in the (optically-thick) $^2\Pi_{1/2}$, J = 1/2→ $^2\Pi_{3/2}$, J = 3/2 transitions, but it turns out to be sufficiently fast to explain even the strongest IR star masers.

The Gwinn et al. (1973) pump for type I OH masers is similar to example (a). OH is produced in highly-excited states by collisions with H or H_2 or by the collisional dissociation of H_2O. The excited-OH decays down the rotational ladder and, because of subtle quantum-mechanical details of the collision process, preferentially populates the upper Λ-doublet levels in the $^2\Pi_{3/2}$ ladder. Again, the pump is distributed through the maser, but the sink (radiative deexcitation of rotational transitions) involves the leakage of photons out of the cloud in optically-thick lines. It turns out that the leakage rate is just fast enough to account for the strongest OH line known in W49 (the strongest source). Maybe we don't see stronger lines because of these thermodynamic considerations.

Other Masers

The masers discussed so far are the most abundant, the most intense, and the most astrophysically interesting. There is a wide variety of other circumstances, though, in which maser action is important. From considerations of their pumping mechanisms and typical environments, these masers can be grouped into two broad categories

A) Infrared - pumped masers in dust clouds.

Type IIa OH masers - in which 1720 MHz emission is stronger, and 1612 MHz emission weaker, than expected in LTE - were originally thought to be associated with supernova remnants. Subsequent work has shown that the emission sources are actually dark dust clouds, only coincidentally associated with SNRs (Turner 1969); and that, while most dust cloud masers show enhanced 1720 and suppressed 1612, some have reversed satellite-line patterns, and a few have main-line anomalies (Turner 1973). Several sources have unusually broad lines ($\Delta v \sim 2$ km s^{-1}) and angular sizes as large as 30' (Goss et al. 1973), while others are clearly small sources (Hardebeck 1971).

The typical explanation for satellite-line inversion without main-line inversion is the far-IR pump of Litvak (1969). The propagation of far-IR radiation through a cloud induces transitions from the ground state to the first two excited rotational states, $^2\Pi_{3/2}$, J = 5/2 and $^2\Pi_{1/2}$, J = 1/2. The infrared transitions between different hyperfine levels have different line strengths and, depending on the details of the optical depths in various lines, population can be shuttled between the F = 2 and F = 1 levels of the ground state. The result is inversion of either the 1720 or the 1612 MHz transition, but not of the main lines.

CH has an energy-level structure like that of OH, except that the ground-state is in the $^2\Pi_{1/2}$ ladder. When CH was first discovered (Turner and Zuckerman

1974), it was seen in weak emission ($T_A \sim 0.3$K) towards the dark clouds in front of the strong continuum source Cas A ($T_C \sim 320$ K). (The relative positions of line source and continuum were deduced from the velocity structure of the emission.) Since $T_A \simeq (T_{ex} - T_c)\tau$, normal excitation would require $T_{ex} > T_c$, rather unlikely for the physical conditions in dust clouds. The alternative interpretation - weak masing - has since been supported by observation of anomalous hyperfine ratios. The far-infrared pump works for CH in the same way as for OH and is in fact the presently preferred mechanism (Zuckerman and Turner 1975).

B) Collisionally - pumped masers in dense clouds.

It often turns out that in the complicated energy-level structure of a large molecule, the details of the radiative deexcitation rates can lead to population inversions - provided that there is a suitable amount of initial excitation by a process that ignores radiative selection rules (e.g., 'hard' collisions). Goldsmith (1972) showed how this could occur even in linear molecules: Hard collisions (large ΔJ) take molecules from $J = 0$ to high J states, from which they cascade downwards, primarily by spontaneous emission. Since the Einstein A-coefficients decrease roughly as J^3 in linear molecules, population tends to pile up in the lower states, and can lead to inversion of the $J = 1 \rightarrow 0$ transition. This specific effect appears to be operating in HC_3N (Morris et al. 1976) and HC_5N (Broten et al.1976).

Litvak (1973) has argued that the overall energy-level structure of asymmetric-top molecules (combined with rather general excitation conditions) virtually guarantees population inversions in the lowest K-doublets. Indeed, the asymmetric tops NH_2CHO, HCOOH, CH_3CHO, CH_3OH, CH_2NH, CH_2CHCN, and $HCOOCH_3$ all show K-doublet emission, even though they are seen against the strong

continuum sources Sgr B2 and Sgr A. (H_2CO and H_2CS absorption in the same direction, at the same velocity, implies that the molecular cloud is in front of the continuum source.) By the same argument used for CH, it is unlikely that $T_{ex} - T_c > 0$ (especially with nearly the same $T_{ex}-T_c$ for so many different molecules). Detailed statistical equilibrium calculations (e.g., Gottlieb et al. 1973) support the conclusion that these are masers. (Figure 8 is the 'spider web' diagram of acetaldehyde. It shows the masing K-doublets, $1_{10} \rightarrow 1_{11}$ and $2_{11} \rightarrow 2_{12}$, and illustrates the complexity of these problems.)

Figure 8 (next page): Energy level diagram of acetaldehyde, taken from Gilmore et al. (1976, Ap. J., 204, 43.) The K-doublets are the states $J_{K_-\ K_+}$ with equal K_-, closely spaced in energy.

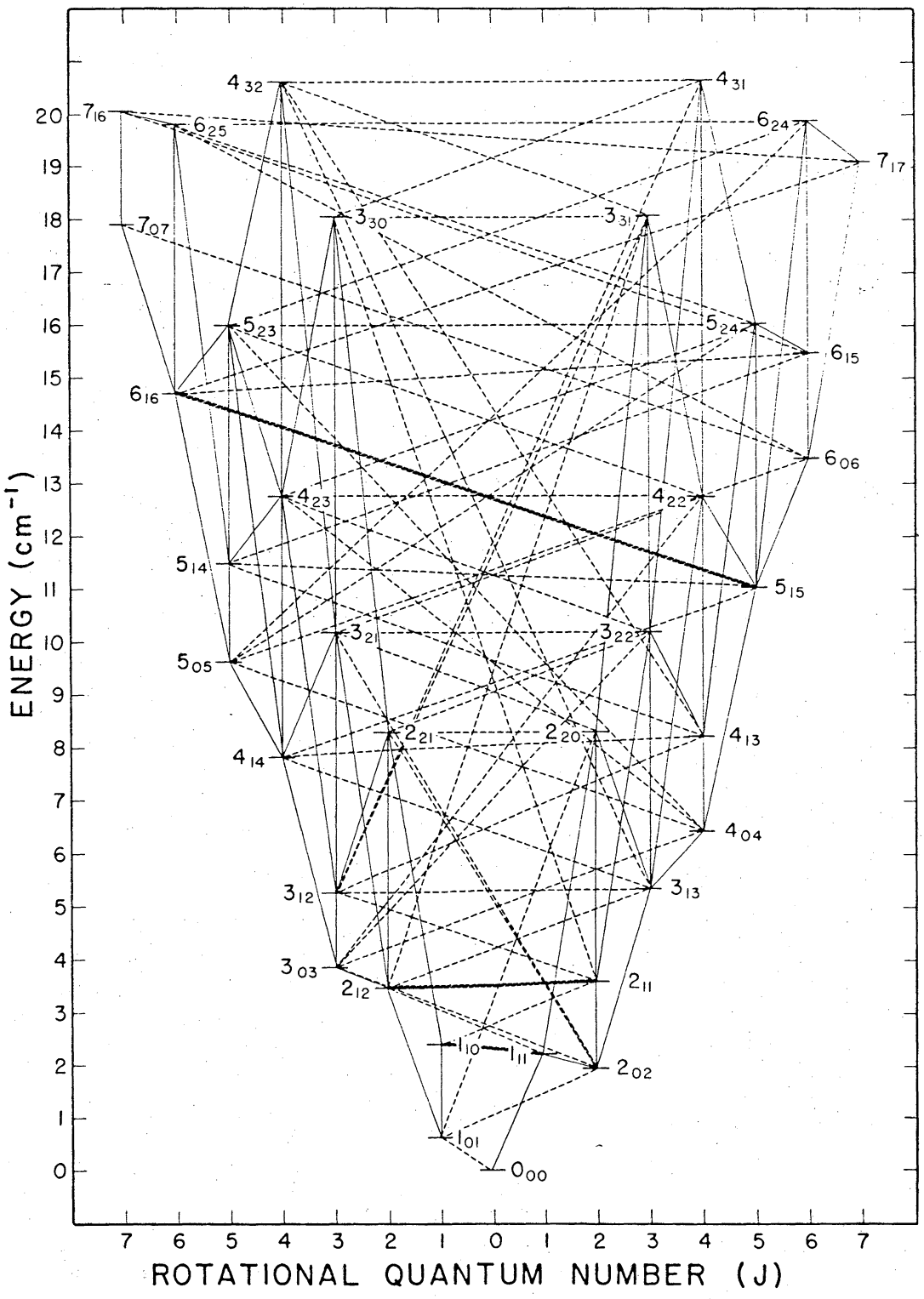


FIG. 2.—The lowest ground-state rotational energy levels of acetaldehyde (CH_3CHO). Each level is labeled with its $J_{K_-K_+}$ asymmetric rotor quantum numbers; levels with even $(K_- + K_+)$ parity are on the right, and those with odd parity are on the left. Allowed transitions active through the μ_b dipole moment are indicated by thin solid lines, and those active through μ_a are indicated by thin dashed lines. The observed interstellar transitions are indicated by heavy dashed lines connecting the appropriate levels. Essentially parallel and independent sets of energy levels such as these exist for the A and E states of the molecule with the E state 0_{00} level offset 0.069 cm^{-1} to higher energy.

For other complex molecules, there are no general rules; identification of maser behavior requires much work. Consider the case of the millimeter-wavelength methanol maser. Barrett et al. (1971) observed a set of CH_3OH lines in Orion (the 4_2-4_1, 5_2-5_1, 6_2-6_1, 7_2-7_1, and 8_2-8_1, E transitions) for which the source size was <1' (implying moderate T_B). From the relative intensities of the lines, they concluded that departures from LTE were not likely to be important. But Zuckerman et al. (1972) and Turner et al. (1972) argued that the intensities of those lines were anomalous relative to other methanol lines. From statistical equilibrium calculations, they inferred that the lines were weakly masing. Chiu et al. (1974) reobserved the $7_2 \rightarrow 7_1$ and $6_2 \rightarrow 6_1$, lines, finding them to be stronger than before. But, because the $(6_2 - 6_1)/(7_2 - 7_1)$ ratio hadn't changed, and because there were no variations during the 3 months of their observations (3^{mo} corresponds to the light-travel time of a 1' diameter source at Orion), they decided that they were including another source within their larger beam. (Still, they noted that the only time variations in spectral lines in radio-astronomical sources were seen in masers.)

However, Hills et al. (1974) derived upper limits to the CH_3OH source sizes $\sim$10" - 30", and also resolved the lines into several narrow (0.4 km/s) features. Since T_B (implied by source size) >> T_K (implied by line width), they concluded that the lines were masing. They noted, though, that because the time variations were weak, T_B is probably not very large and, because of the similarity of the $6_2 - 6_1$, and $7_2 - 7_1$ line shapes, the overall gain is probably not too high.

Collisional pumping of the variety discussed by Zuckerman et al. and Turner et al. has also been invoked in the case of the H_2O maser. de Jong (1973) has argued that, given sufficient collisional excitation of high-J

levels (as might be expected in a collapsing protostar), the strongest radiative decay routes naturally funnel population into the upper level of the masing transition. The maser must occur at the surface of the source, though, where the IR transitions feeding the inversion will be optically thin.

REFERENCES

Review Articles

Goldreich, P. 1975, in Atomic and Molecular Physics and the Interstellar Medium, eds. R. Balian, P. Encrenaz, and J. Lequeux (Amsterdam: North-Holland), p. 409.

Kegel, W. H. 1975, in Problems in Stellar Atmospheres and Envelopes, eds. B. Baschek, W. H. Kegel, and G. Traving (Heidelberg: Springer-Verlag), p. 257.

Litvak, M. M. 1974, Ann. Rev. Astr. and Ap., 12, 97.

Moran, J. M. 1976, in Frontiers of Astrophysics, ed. E. H. Avrett (Cambridge: Harvard University Press), p. 385.

ter Haar, D., and Pelling, M.A. 1974, Repts. on Prog. in Phys., 37, 481.

Turner, B. E. 1970, R.A.S.C. Journal, 64, 221, 282.

Cited Articles

Allen, C. W. 1973. Astrophysical Quantities.

Barrett, A. H., and Rogers, A. E. E. 1966, Nature, 210, 188.

Barrett, A. H., Schwartz, P. R., and Waters, J. W. 1971, Ap. J. (Letters), 168, L101.

Broten, N.W., MacLeod, J.M., Oka, T., Avery, L.W., Brooks, J.W., McGee, R.X., and Newton, L. M. 1976, Ap. J. (Letters), 209, L143.

Chui, M. F., Cheung, A.C., Matasakis, D., Townes, C.H., and Cardiasmenos, A. G. 1974, Ap. J. (Letters), 187, L19.

Cudaback, D. D., Read, R. B., and Rougoor, G. W. 1966, Phys. Rev. Lett., 17, 452.

Davies, R. D., de Jager, G., and Verschuur, G. 1966, Nature, 209, 974.

Ehrenstein, G., Townes, C. H., and Stevenson, M.J. 1959, Phys. Rev. Lett., 3, 40.

Elitzur, M., Goldreich, P., and Scoville, N.Z. 1976, Ap. J., 205, 384.

Evans, N. J., Hills, R. E., Rydbeck. O. E. H., and Kollberg, E. 1972, Phys. Rev. A. 6, 1643.

Gardner, F. F., Robinson, B.J., Bolton, J.G., and van Damme, K. J. 1964, Phys. Rev. Lett., 13, 3.

Goldreich, P., and Kwan, J. 1974a, Ap. J., 190, 27.

_____. 1974b, Ap. J., 191, 93.

Goldsmith, P. F. 1972, Ap. J., 176, 597.

Goss, W. M., Johansson, L.E.B., Elldér, J., Höglund, B., N-Q-Rieu, and Winnberg, A. 1973, Astr. and Ap., 28, 89.

Gottlieb, C.A., Palmer, P., Rickard, L. J , and Zuckerman, B. 1973, Ap. J., 182, 699.

Gundermann, E. J. 1965. Ph.D. Thesis, Harvard University.

Gwinn, W.D., Turner, B. E., Goss, W. M., and Blackman, G.L. 1973, Ap. J., 179, 789.

Habing, H. J., Goss, W. M., Matthews, H.E., and Winnberg, A. 1974, Astr. and Ap. 35, 1.

Hardebeck, E. 1971, Ap. J., 170, 281.

Harvey, P. M., Bechis, K. P., Wilson, W. J., and Ball, J. A. 1974, Ap. J. Suppl., 27, 331.

Hills, R., Pankonin, V., and Landecker, T. L. 1975, Astr. and Ap., 39, 149.

Jong, T. de 1973, Astr. and Ap., 26, 297.

Litvak, M. M. 1969, Ap. J., 156, 471.

Litvak, M. M. 1973, in Atoms and Molecules in Astrophysics, eds. T. R. Carson and M. J. Roberts (New York: Academic), p. 201.

Lo, K. Y., 1974, Ph.D. thesis, M. I. T.

Morris, M., Turner, B. E., Palmer, P., and Zuckerman, B. 1976, Ap. J., 205, 82.

Rogers, A. E. E., Moran, J. M., Crowther, P. P., Burke, B. F., Meeks, M. L., Ball, J. A., and Hyde, G. M. 1966, Phys. Rev. Lett., 17, 450.

Schwartz, P. R., Harvey, P. M., and Barrett, A. H., 1974, Ap. J., 187, 491.

Scovil, H. E. D., and Schulz-DuBois, E. O. 1959, Phys. Rev. Lett., 2, 262.

Shklovskii, I. S. 1949, Astr. Zhur., 26, 10.

Turner, B. E. 1969, Ap. J., 157, 103.

_____. 1973, Ap. J., 186, 357.

Turner, B. E., and Zuckerman, B. 1974, Ap. J. (Letters), 187, L59.

Turner, B. E., Gordon, M.A., and Wrixon, G. T., 1972, Ap. J., 177, 609.

Weinreb, S., Barrett, A. H., Meeks, M. L., and Henry, J. C. 1963, Nature, 200, 829.

Weinreb, S., Meeks, M. L., Carter, J. C., Barrett, A. H., and Rogers, A. E. E. 1965, Nature, 208, 440.

Weaver, H., Williams, D. R. W., Dieter, N. H., and Lum, W. T. 1965, Nature, 208, 29.

Wilson, W. J., and Barrett, A. H. 1972, Astr. and Ap., 17, 385.

Zuckerman, B., and Turner, B.E. 1975, Ap. J., 197, 123.

Zuckerman, B., Turner, B. E., Johnson, D. R., Palmer, P., and Morris, M. 1972, Ap. J., 177, 601.

Molecular Spectroscopy

Gordy, W., and Cook, R. L. 1970, Microwave Molecular Spectra (New York: Interscience).

Townes, C. H., and Schawlow, A. L. 1955, Micorwave Spectroscopy (New York: McGraw-Hill).