THEORY OF ASTRONOMICAL MASERS

Nikolaos D. Kylafis
University of Crete
Physics Department
P. O. Box 470
714 09 Iraklion, Crete
GREECE

The theory of astronomical masers is reviewed. As with laboratory masers, masing occurs when a transition between two energy levels of a molecule exhibits inverted populations. In order to present the basic concepts about masers, an idealized two-level system is used. The exact energy level structure is taken into account later on when the pumping of specific molecules is discussed. Unlike laboratory masers, where the radiation must be bounced between two mirrors to accumulate gain, the propagation of radiation in astronomical masers is a lot simpler. This is because astronomical masers are single-pass and broadband. Thus, the main theoretical effort has concentrated on inventing efficient mechanisms that produce population inversion. Specific pumping mechanisms for the three molecules (H_2O, SiO, and OH) that exhibit strong masing are presented and their ability to explain the observations is discussed.

1. BASIC CONCEPTS

I would like to start with a qualitative description of masers and then proceed to a discussion of the basic concepts regarding astronomical masers (for a more extensive discussion, see Elitzur 1982a). I will briefly discuss the differences between laboratory masers and astronomical masers, and I will devote the major part of this review to the specific pumping mechanisms proposed for the different molecules in various astronomical environments.

The name MASER comes from the capitalized initials of the following words: Microwave Amplification through Stimulated Emission of Radiation. Microwave radiation is emitted when a radiative transition occurs between two energy levels of a molecule. For the purposes of a qualitative discussion, we will idealize the energy level diagram of a molecule as follows. We will isolate the two energy levels, between which the maser transition occurs, and treat all the other levels as a reservoir. As a result of collisional and radiative transitions, the molecules circulate between the reservoir and the masing levels. If, as a consequence of this circulation, and when steady state is achieved, the population per magnetic sublevel of the upper masing

level is larger than that of the lower masing level, then the maser phenomenon takes place. This is because a photon that was emitted by a spontaneous emission from the upper to the lower masing level is more likely as it propagates to cause a stimulated emission than to be absorbed. Under normal conditions, the optical depth of the medium, which is proportional to the population per magnetic sublevel of the lower level minus that of the upper level, is positive and causes attenuation of the radiation. However, under conditions of inverted populations, the optical depth is negative and causes amplification of the radiation. This amplification is exponential because each maser photon causes the emission in the same direction of an identical photon.

The question that naturally arises at this point is: Can the intensity of a maser grow exponentially forever, assuming that we have a large enough region? The answer is clearly no, because with each stimulated emission a molecule goes from the upper masing level to the lower one. Thus, stimulated emission tends to reduce the population inversion. We will say that a maser is *saturated* when the stimulated emission rate is comparable to or greater than the intrinsic pumping rate. After saturation sets in, the growth of the intensity is no longer exponential. In the case of a one-dimensional maser, the growth after saturation is linear.

It is often argued that astronomical masers are strongly beamed. In order to understand how strong beaming can occur, we consider a maser region that is not spherically symmetric but is elongated. The maser photons travel, of course, in all directions. Nevertheless, for simplicity we approximate the maser radiation field as consisting of two beams, one in the direction of elongation and the other perpendicular to it. From geometrical considerations alone, the beam propagating in the direction of elongation is stronger than the other. This anisotropy is further enhanced because at every point in the maser, the two beams compete for the available pump photons and the more intense beam wins. Thus, an anisotropy in the shape of the masing region results in strong beaming. This statement has been quantified by Alcock and Ross (1985a,b; 1986a). Observers usually quote *isotropic* luminosities for the various sources. That is, they make the assumption that the observed flux is the same in all directions. It should be kept in mind that this assumption may result in significant overestimates of the true luminosity.

Many maser sources exhibit significant amounts of polarization. In order for a masing region to emit polarized radiation, the molecules must not only exhibit *inverted* populations, but also *unequal* populations in the magnetic sublevels of a maser level. Magnetic fields in the masing region play a central role in determining the polarization of emitted radiation, as demonstrated in a classic paper by Goldreich, Keeley, and Kwan (1973). This work has been generalized and extended in recent years (Western and Watson 1983a,b,c; 1984; Deguchi and Watson 1986a,b; Deguchi, Watson and Western 1986). These authors have shown that anisotropic radiation fields in the maser region can also cause polarization. Thus, it is not always clear what is the cause of the observed polarization. Notable exceptions, however, do exist. For example, Moran *et al.* (1978) identified 12 Zeeman pairs of left and right circular polarization in the spectrum of W3(OH) and determined the magnetic field strength in the region to be 2–9 mG.

In closing the qualitative discussion about astronomical masers, I would like to

mention the major differences between laboratory masers and astronomical masers. In laboratory masers, the maser region is typically ~ 1 m long. In order to have significant gain, the radiation is bounced between two mirrors, one of which is semi-transparent to allow the amplified radiation to escape. In contrast, astronomical masers are single-pass, broadband, and have linear dimensions of a few AU or larger. Because the radiation transfer in astronomical masers is relatively simple, the main theoretical effort has concentrated on inventing efficient pumping mechanisms. In what follows, I will discuss in detail the various pumping mechanisms that have been proposed. For each of the three molecules that exhibit the strongest maser emission, namely H_2O, SiO, and OH, I will discuss both the environments that host them and the specific pumping mechanism that is believed to be operating there.

2. PUMPING MECHANISMS

2.1. H_2O Masers

The H_2O molecule exhibits one known maser transition between levels 6_{16} and 5_{23}. An inspection of the energy level diagram of H_2O (see, e.g., de Jong 1973, his Fig. 1) shows that the above two levels accidently lie close to each other, thus giving rise to a transition of 22 GHz. It is not an exaggeration to say that if you abuse this molecule in some way, it will most likely mase. A common method of abuse in astronomical environments is to expose the H_2O molecules to collisions and/or radiation. Thus, as collisions and/or radiation move the H_2O molecules among the energy levels, the upper masing level always gets more than the lower one, and population inversion is established. This is only a fluke of the transition rates. As it will be stressed later on, this pumping mechanism has its limitations.

2.1.1. *H_2O masers in the envelopes of late-type stars.* This class of H_2O masers presents no problem in its theoretical understanding because the masers in the envelopes of late-type stars are not intense. Cooke and Elitzur (1985), using plausible parameters to describe the environment, have shown that collisional pumping more than accounts for the observed luminosities.

2.1.2. *H_2O masers in star-forming regions.* These masers are the most intense of all galactic masers. Contrary to previous beliefs, Strelnitskij (1984) demonstrated that all models, suggested up to that time, suffer difficulties in explaining the extremely strong sources such as W49, W51, and others. The basic problem with these models is that the maximum power they can produce is lower than what is observed, and, if the models are pushed to their limits, the H_2O molecules thermalize.

Strelnitskij (1980; 1984) proposed the elegant idea that collisional pumping of H_2O masers could avoid the difficulty of thermalization if the H_2O molecules collide with two kinds of particles (e.g., H_2 molecules and electrons) at different kinetic temperatures, but with comparable collision rates. In other words, if the H_2O molecules find themselves in contact with two heat baths, they do not know which of the two Maxwellian distributions to obey and in fact they obey neither. With the help of analytical and numerical model calculations, it has been demonstrated (Kylafis and Norman 1986; 1987) that collisional pumping of H_2O, in environments where the temperature of the neutral particles T_H is significantly less than that

of the electrons T_e, can sustain inverted populations. Such environments exist in the magnetic precursors of MHD shocks (Draine 1980; 1981; Draine, Roberge and Dalgarno 1983).

To understand physically why inversion occurs when $T_H \ll T_e$, consider the artificial situation where the electron collisions are turned off. Then, the H_2O molecules are in thermal equilibrium with the neutral particles. This means that for low enough T_H, practically all H_2O molecules are in the lowest rotational state (level 1_{01}). If the electron collisions are now turned on, there is a flow of H_2O molecules upward in energy space. The electron collision rates for these upward transitions are such that there is a wide "pathway" leading to the upper masing level but a narrow one to the lower masing level. Since the downward collision rates by the neutral particles do not destroy this asymmetry, population inversion is established. For more details see Kylafis and Norman (1987). The two-temperature model appears to explain the strong galactic H_2O maser sources, but it needs testing.

2.1.3. *Extragalactic H_2O masers.* Extragalactic H_2O megamasers with luminosities roughly in the range of 100 to 1000 solar luminosities have been observed. For reviews, see Baan (1985) and Diamond (1987). As discussed at the beginning of this section, infrared radiation *can* cause inversion in the H_2O molecules. However, the observations suggest that the infrared luminosity is not enough to pump the masing regions. It is possible that two-temperature collisional pumping is operating there (Claussen and Lo 1986). Because of this, I would like to point out that hot electrons and cold neutrals can also coexist in ionization fronts (Kylafis and Norman 1987). It is therefore interesting to note the similarity between galactic star-forming regions and the region somewhere between the broad line (BLR) and narrow line regions (NLR) in active galactic nuclei. Both have densities $\sim 10^9 - 10^{11}$ cm^{-3} and characteristic lengths $\sim 10^{11} - 10^{13}$ cm. If the extragalactic H_2O masers are indeed located in the BLR/NLR, then VLBI observations can probe these regions and provide invaluable information.

2.2. SiO Masers

The SiO molecule exhibits the most masing transitions of all other molecules. Maser emission has been observed in rotational transitions as high as $J = 5$ to $J = 4$ and in vibrational transitions as high as $v = 3$. Therefore, the maser emission in SiO is a general feature of the molecule and does not depend on the properties of a particular transition and the associated energy levels. For a review of the observations and a diagram of the energy levels of SiO, see Lane (1982, her Fig. 1).

The pumping mechanism that is believed to cause population inversion in SiO molecules was first suggested by Kwan and Scoville (1974). The basic idea is that if the radiative transition from the vibrational state v to $v-1$ is optically thick, then the transition rate is inversely proportional to $2J+1$, where J is the rotational level in the vibrational level v. Thus, the smaller the value of J is, the faster the level is drained and inversion is established. The upward transitions can be caused either by collisions (Elitzur 1980) or infrared radiation (Deguchi and Iguchi 1976). In the case of infrared pumping, the upward transitions must be optically thin while the downward ones must be optically thick. In other words, the optical depth must be anisotropic and larger than unity when it is averaged over all directions.

2.2.1. *SiO masers in the envelopes of late-type stars.* Detailed model calculations, taking into account both collisions and radiation, of SiO maser emission in a smooth stellar wind around a late-type giant have been performed by Langer and Watson (1984). They have found that, contrary to previous beliefs, for a variety of system parameters the model fails to produce the observed power. VLBI observations (Lane 1982) indicate that the maser regions are blobs or shocks. No model calculations have as yet been performed for such maser regions (see, however, Alcock and Ross 1986b).

I would like to point out here that strong magnetic fields (of order 10–80 G) have been inferred in the maser regions of late-type giants and supergiants (McIntosh 1987). One arrives at a similar conclusion by scaling a magnetic field strength of ~ 1 mG at $R \sim 3 \times 10^{16}$ cm for supergiant OH–IR sources (Cohen *et al.* 1987) down to a few stellar radii. If that is the case, then the magnetic field dominates the dynamics in the SiO maser region. Detailed studies of maser polarization with VLBI would therefore be of great help to the theorists.

2.2.2. *SiO masers in star-forming regions.* Up to 1985, it was thought that only IRc2 in Orion exhibited SiO maser emission. A model by Elitzur (1982b) could be made to work if stringent requirements on the parameters were imposed. This led Elitzur to suggest that the model automatically explained the lack of SiO maser emission in other star-forming regions.

Hasegawa *et al.* (1985) reported observations of SiO masers in the star-forming regions W51 IRS2 and Sgr B2 MD5. These SiO masers coincide in position with strong H_2O masers within a positional uncertainty of 5 arcseconds, and their isotropic luminosities are comparable to that in Orion. Further theoretical work is therefore needed for this class of SiO masers.

SiO masers in other galaxies have not been observed yet.

2.3. OH Masers

The OH molecule exhibits maser emission in all four allowed transitions in the ground state and also in several higher rotational levels (Reid and Moran 1981, see their Table 1). For a diagram of the energy levels of OH, see, for example, Elitzur (1982a, his Fig. 1). Of all the observed maser transitions in OH, the best studied ones are the four transitions in the ground state. The F-conserving transitions $F = 2$ to $F = 2$ (1667 MHz) and $F = 1$ to $F = 1$ (1665 MHz) are called *main* lines. The F-changing ones $F = 1$ to $F = 2$ (1612 MHz) and $F = 2$ to $F = 1$ (1720 MHz) are called *satellite* lines. As with the other two molecules discussed above, the OH molecules are pumped by radiation and collisions from the ground state to the higher rotational levels and then cascade back. The route they follow on their cascade back to the ground state determines which of the two satellite lines gets inverted (for details, see Elitzur 1982a). Unlike the satellite lines, the inversion of the main lines requires the pump process to have a parity selectivity built in. Radiative pumping can have this selectivity if the intensity of the radiation field that causes the excitation is a monotonically increasing function of frequency (Elitzur 1982a).

2.3.1. *OH masers in the envelopes of late-type stars.*

This class of OH masers is fairly well understood theoretically. The pumping is done radiatively (Elitzur 1982a), and the stellar radiation is sufficient to do it (Dickinson 1987). The masers lie more or less in a spherical shell (as an example, see Chapman and Cohen 1986), but observations and theory suggest that they are located in blobs in the shell (Alcock and Ross 1986c).

2.3.2. *OH in star-forming regions.*

The OH masers in star-forming regions are associated with compact HII regions. There is currently a controversy as to where exactly the OH masers lie, but I will not enter into it. The interested reader can consult a review of the observations by Genzel (1986) and an assessment of the theoretical understanding by Elitzur (1986). I will restrict myself to pumping considerations. Unlike the OH masers in the envelopes of late-type stars, radiative pumping here can be excluded because the number of exciting photons in the UV, the near infrared, and the far infrared are smaller than the number of maser photons observed. Contrary to previous beliefs, collisions with neutral particles also fail to establish inversion (Andresen *et al.* 1984). Johnston (1967) showed that if the OH molecules are exposed to collisions with a stream of electrons, then the excitation rate is larger than the de-excitation rate for $\Delta m = 0$ transitions and population inversion is established. The problem with this mechanism is how to get these streams of electrons. In order to speak of a stream, the ordered velocity of the electrons must be larger than their thermal speed, which is difficult to achieve. Because of this, Elitzur (1979) suggested ion streams. Ions stream when their ordered velocity exceeds ~ 1 km/s, a condition which is easily satisfied in MHD shocks. However, the collision rates of ions with OH are not known, and quantitative calculations cannot be performed. It would be very interesting to carry out two-temperature calculations in MHD shocks, similar to the ones for H_2O, to see whether this pumping mechanism works also for OH.

2.3.3. *Extragalactic OH masers.*

Extragalactic OH megamasers have been observed in active galaxies with high far infrared emission. Baan (1985) and Staveley-Smith *et al.* (1987) give a detailed account of the observations and report the results of systematic searches for such masers. Their luminosity is comparable to that of H_2O masers, i.e., roughly in the range of 100 to 1000 solar luminosities, and it comes from amplification of the background radio source. The pumping is done by infrared radiation, which is abundant in the host galaxies (Baan 1985).

One point that I would like to stress before closing this review is the fact that high velocity wings have been observed (Staveley-Smith *et al.* 1987) in the OH maser line profiles of the Seyfert galaxies Mrk 231 (~ 800 km/s) and Mrk 273 (~ 1100 km/s). These velocities are comparable to those in the forbidden line regions of Seyfert galaxies. VLBI observations are therefore recommended in order to locate the emission regions relative to the nuclei of the galaxies.

3. REFERENCES

Alcock, C. and Ross, R. R. 1985a, *Ap. J.*, **290**, 433.
Alcock, C. and Ross, R. R. 1985b, *Ap. J.*, **299**, 763.
Alcock, C. and Ross, R. R. 1986a, *Ap. J.*, **306**, 649.
Alcock, C. and Ross, R. R. 1986b, *Ap. J.*, **310**, 838.
Alcock, C. and Ross, R. R. 1986c, *Ap. J.*, **305**, 837.
Andresen, P., Hausler, D., Lulf, H. W., and Kegel, W. H. 1984, *Astr. Ap.*, **138**, 617.
Baan, W. A. 1985, *Nature*, **315**, 26.
Chapman, J. M., and Cohen, R. J. 1986, *M.N.R.A.S.*, **220**, 513.
Claussen, M. J., and Lo, K. -Y. 1986, *Ap. J.*, **308**, 592.
Cohen, R. J., Downs, G., Emerson, R., Grimm, M., Gulkis, S., Stevens, G., and Tarter, J. 1987, *M.N.R.A.S.*, **225**, 491.
Cooke, B., and Elitzur, M. 1985, *Ap. J.*, **295**, 175.
Deguchi, S., and Iguchi, T. 1976, *Pub. Astr. Soc. Japan*, **28**, 307.
Deguchi, S., and Watson, W. D. 1986a, *Ap. J. (Letters)*, **300**, L15.
Deguchi, S., and Watson, W. D. 1986b, *Ap. J.*, **302**, 750.
Deguchi, S., Watson, W. D., and Western, L. R. 1986, *Ap. J.*, **302**, 108.
de Jong, T. 1973, *Astr. Ap.*, **26**, 297.
Diamond, P. J. 1987, these proceedings.
Dickinson, D. F. 1987, *Ap. J.*, **313**, 408.
Draine, B. T. 1980, *Ap. J.*, **241**, 1021.
Draine, B. T. 1981, *Ap. J.*, **246**, 1045.
Draine, B. T., Roberge, W. G., and Dalgarno, A. 1983, *Ap. J.*, **264**, 485.
Elitzur, M. 1979, *Astr. Ap.*, **73**, 322.
Elitzur, M. 1980, *Ap. J.*, **240**, 553.
Elitzur, M. 1982a, *Rev. Mod. Phys.*, **54**, 1225.
Elitzur, M. 1982b, *Ap. J.*, **262**, 189.
Elitzur, M. 1986, in *Masers, Molecules, and Mass Outflows in Star Forming Regions*, ed. A. D. Haschick (Haystack Observatory), p. 299.
Genzel, R. 1986, in *Masers, Molecules, and Mass Outflows in Star Forming Regions*, ed. A. D. Haschick (Haystack Observatory), p. 233.
Goldreich, P., Keeley, D. A., and Kwan, J. Y. 1973, *Ap. J.*, **179**, 111.
Hasegawa, T., Morita, K.-I., Okumura, S., Kaifu, N., Suzuki, H., Ohishi, M., Hayashi, M., and Ukita, N. 1986, in *Masers, Molecules, and Mass Outflows in Star Forming Regions*, ed. A. D. Haschick (Haystack Observatory), p. 275.
Johnston, I. D. 1967, *Ap. J.*, **150**, 33.
Kwan, J., and Scoville, N. 1974, *Ap. J. (Letters)*, **194**, L97.
Kylafis, N. D., and Norman, C. 1986, *Ap. J. (Letters)*, **300**, L73.
Kylafis, N. D., and Norman, C. 1987, *Ap. J.*, in press.
Lane, A. 1982, Ph.D. Thesis, University of Massachusetts.
Langer, S. H., and Watson, W. D. 1984, *Ap. J.*, **284**, 751.
McIntosh, G. 1987, Ph.D. Thesis, University of Massachusetts.
Moran, J. M., Reid, M. J., Lada, C. J., Yen, J. L., Johnston, K. J., and Spencer, J. H. 1978, *Ap. J. (Letters)*, **224**, L67.
Reid, M. J., and Moran, J. M. 1981, *Ann. Rev. Astr. Ap.*, **19**, 231.
Staveley-Smith, L., Cohen, R. J., Chapman, J. M., Pointon, L., and Unger, S. W. 1987, *M.N.R.A.S.*, in press.
Strelnitskij, V. S. 1980, in *IAU Symposium 87, Interstellar Molecules*, ed. B. H.

230

Andrew (Boston: Reidel) p. 591.

Strelnitskij, V. S. 1984, *M.N.R.A.S.*, **207**, 339.

Western, L. R., and Watson, W. D. 1983a, *Ap. J.*, **268**, 849.

Western, L. R., and Watson, W. D. 1983b, *Ap. J.*, **274**, 195.

Western, L. R., and Watson, W. D. 1983c, *Ap. J.*, **275**, 195.

Western, L. R., and Watson, W. D. 1984, *Ap. J.*, **285**, 158.