Inversion mechanism and maser emission in the interstellar medium

Riccardo Cesaroni

INAF–Osservatorio Astrofisico di Arcetri, Firenze, Italy
E-mail: cesa@arcetri.astro.it

MASER emission in the astronomical environment is the subject of this lecture. In order to explain the origin of this phenomenon, the problem of molecular excitation in a galactic molecular cloud is treated formulating the equations of statistical equilibrium in a stationary situation. Special attention is devoted to the idealized cases of 2-level and 3-level systems, with the scope to illustrate the concept of thermalization and population inversion – the latter being at the origin of maser lines. A phenomenological description of astronomical masers is then given, with reference to the criteria used to recognize maser lines. Finally, pros and contras of maser observations are discussed.
1. Introduction

The acronym MASER stands for Microwave Amplification by Stimulated Emission of Radiation, one of the most impressive phenomena in astrophysics. Maser lines have been observed in a variety of environments, from star forming molecular clouds to late-type stars, from normal galaxies to active galactic nuclei. In this contribution we wish to summarize our knowledge of the maser characteristics from an observational point of view. However, for a better understanding of the maser phenomenon it is necessary to analyse the mechanism which leads to inversion of the level population. We will thus first discuss the effects of collisions and radiation on the level populations in an atomic or molecular species and then describe the main properties of maser emission in the interstellar medium.

2. Statistical equilibrium

Let us consider a cloud of gas containing particles of at least one atomic or molecular species. The population of the different energy levels is determined by two types of interactions: collisions among the particles themselves and, possibly, with particles of other species; and radiative interactions with both the photons produced by radiative decays of the levels and those emitted by other sources. The problem of the determination of the level populations can be significantly simplified if one considers that molecular clouds are made of molecular hydrogen (carbon monoxide, the most common molecule after H$_2$ is $10^4$ times less abundant). For this reason the only relevant collisions are those between the given particles and the H$_2$ molecules.

The problem we want to treat can be formulated as follows. We consider a cloud of H$_2$ molecules, whose number density $n_{H_2}$ and temperature $T_{H_2}$ are given, containing a small number of particles of the species X, with number density $n_X$. The cloud is also experiencing a known radiation field $I_\nu$. In a stationary situation, the number density $n_i$ in a generic energy level $i$ of the species X can be obtained by imposing that the rate of the transitions out of that level equals that of the transitions into it. This is expressed by the so-called “statistical equilibrium equations”:

$$n_i \left[ \sum_{j \neq i} (B_{ij}J_\nu + C_{ij}) + \sum_{j < i} A_{ij} \right] = \sum_{j \neq i} (n_jB_{ji}J_\nu + n_jC_{ji}) + \sum_{j > i} n_jA_{ji}$$  \hspace{1cm} (2.1)

In addition, one must satisfy the constraint

$$\sum_i n_i = n_X.$$ \hspace{1cm} (2.2)

Here, $A_{ij}$ and $B_{ij}$ are the Einstein coefficients, $J_\nu = \frac{1}{4\pi} \int_{4\pi} I_\nu \, d\Omega$ is the mean brightness over 4$\pi$ steradians, and $C_{ij} = n_{H_2} \langle v \sigma_{ij} \rangle$ are the collisional rate coefficients for X–H$_2$ interactions. The average $\langle v \sigma_{ij} \rangle$ is calculated over the distribution (e.g. a Maxwellian) of the relative velocity between H$_2$ and X, with $\sigma_{ij}$ the cross section for a collision from level $i$ to $j$.

Despite appearances, the previous equations are non-linear in the unknown variables $n_i$, because these are hidden in $J_\nu$. In fact, $I_\nu$ is the solution of the radiative transfer equation

$$\frac{dI_\nu}{d\tau} = S_\nu - I_\nu$$ \hspace{1cm} (2.3)
where the optical depth along a given line of sight (l.o.s.), \( \tau(\nu) \equiv \int_{\text{l.o.s.}} \kappa \, ds \), and the source function, \( S_\nu \equiv \frac{\nu}{\kappa(\nu)} \), depend on \( n_i \) through the absorption and emission coefficients

\[
 j_\nu = \frac{h\nu}{4\pi} n_i A_{ij} \phi(\nu - \nu_{ij}) \tag{2.4}
\]

\[
 \kappa(\nu) = \frac{h\nu}{4\pi} \left( n_j B_{ji} - n_i B_{ij} \right) \phi(\nu - \nu_{ij}) \tag{2.5}
\]

where \( h\nu_{ij} \) is the energy difference between levels \( i \) and \( j \) and \( \phi \) the normalised line profile \( \int_0^{+\infty} \phi(\nu - \nu_{ij}) \, d\nu = 1 \). The previous expression can be simplified using the well known relations

\[
 g_i B_{ij} = g_j B_{ji}, \quad A_{ij} = \frac{2h\nu^3}{c^2} B_{ij}, \quad \text{and} \quad C_{ij} = \frac{g_i g_j}{g_j} e^{-\frac{\nu_{ij}}{T_{\text{ex}}}}.
\]

In conclusion, the statistical equilibrium equations in general can only be solved numerically and under suitable approximations. However, in order to explain the effect of collisions and radiation on the level populations it is very instructive to consider a couple of simple, idealized cases where Eqs. (2.1) and (2.2) can be solved analytically. This will be done in Sects. 2.2.1 and 2.2.2.

### 2.1 Radiative transfer equation: the line case

In order to understand the origin of population inversion and maser transitions, it is instructive to solve the radiative transfer equation in the simple case of a homogeneous, isothermal cloud. For this purpose it is also convenient to define the quantity \( T_{\text{ex}} \) named “excitation temperature”, through the relation:

\[
 \frac{n_i}{n_j} \equiv \frac{g_i}{g_j} e^{\frac{\nu_{ij}}{kT_{\text{ex}}}}. \tag{2.6}
\]

Clearly, in thermodynamic equilibrium Eq. (2.6) is the Boltzmann equation and \( T_{\text{ex}} \) corresponds to the physical temperature of the gas; otherwise \( T_{\text{ex}} \) depends on the pair of levels \( i \) and \( j \). Roughly speaking, one can say that the closer \( T_{\text{ex}} \) is to e.g. the kinetic temperature of the gas, the closer the system will be to local thermodynamic equilibrium (LTE). We will illustrate this point in Sect. 2.2.1.

With this definition, \( \tau \) and \( S_\nu \) take the form:

\[
 S_\nu = B_\nu(T_{\text{ex}}) \tag{2.7}
\]

\[
 \tau = \frac{h\nu}{4\pi} \int_{\text{l.o.s.}} n_i B_{ij} \left( e^{\frac{\nu}{T_{\text{ex}}} - 1} \right) \phi(\nu - \nu_{ij}) \, ds \tag{2.8}
\]

The solution of the radiative transfer equation for line \( i \rightarrow j \) in the homogeneous, isothermal case considered here can now be written as

\[
 I^L_\nu \equiv I_\nu - B_\nu(T_{\text{BG}}) = (B_\nu(T_{\text{ex}}) - B_\nu(T_{\text{BG}})) \left( 1 - e^{-\tau} \right) \tag{2.9}
\]

where the background continuum emission \( B_\nu(T_{\text{BG}}) \) has been subtracted to calculate the “pure” line brightness \( I^L_\nu \). It is also convenient to define the line brightness temperature \( T_B \) by writing

\[
 I^L_\nu \equiv B_\nu(T_B) \tag{2.10}
\]

which in the Rayleigh-Jeans limit, \( h\nu \ll kT_B \), becomes

\[
 I^L_\nu \simeq \frac{2kT_B \nu^2}{c^2}. \tag{2.11}
\]
In conclusion, we obtain

\[ T_B = (T_{\text{ex}} - T_{\text{BG}}) \left(1 - e^{-\tau}\right) \]  

(2.12)

From this result one sees that if \( T_{\text{ex}} > T_{\text{BG}} \) the line is in emission, whereas if \( 0 < T_{\text{ex}} < T_{\text{BG}} \) the line is in absorption. The interesting case for us is that of \( T_{\text{ex}} < 0 \). According to Eq. (2.6) this occurs when the upper level of the transition is statistically overpopulated with respect to the lower level. Equation (2.8) demonstrates that in this case one has \( \tau < 0 \) and Eq. (2.12) can be written as

\[ T_B = |T_{\text{ex}} + T_{\text{BG}}| \left(e^{\mid\tau\mid} - 1\right) \gg T_{\text{BG}}. \]  

(2.13)

Negative excitation temperature and optical depth correspond to population inversion between the upper and lower levels and cause dramatic amplification of the corresponding line, which is said to be “masing”. In fact, large values of \( \tau \) can explain the observed brightness temperatures of maser lines, which may be as high as \( 10^{10} \) K or even more.

2.2 LTE conditions and the inversion mechanism

Under what conditions a molecule or atom can be considered in local thermodynamic equilibrium? The most effective way to illustrate this concept is to consider the simplest case, namely a particle with only 2 levels. As we will see later, this cannot explain the inversion mechanism, for which at least an additional level is needed. In the next two sections we will thus discuss the effects of the kinetic temperature, density, and radiation field on the 2-level and 3-level systems.

2.2.1 The 2-level system

Here, we want to examine the excitation of the species \( \text{X} \) with only 2 energy levels depending on the relevant physical parameters that describe the gas and radiation field. For this purpose, we consider collisions with \( \text{H}_2 \) molecules and interaction with a black-body radiation field \( B_\nu(T_R) \). We assume that the \( \text{H}_2 \) molecules have constant density, \( n_{\text{H}_2} \), and are in thermodynamic equilibrium at a temperature \( T_K \). In this case, Eq. (2.1) reduces to

\[ n_2(A_{21} + B_{21}J_\nu + C_{21}) = n_1(B_{12}J_\nu + C_{12}) \]  

(2.14)

from which one can obtain the ratio \( n_2/n_1 \) and thus the excitation temperature \( T_{\text{ex}} \) from Eq. (2.6):

\[ \frac{n_2}{n_1} = \frac{g_2}{g_1} e^{-\frac{\nu}{kT_{\text{ex}}}}. \]  

(2.15)

Besides, the following relationships hold:

\[ \frac{C_{12}}{C_{21}} = \frac{g_2}{g_1} e^{-\frac{\nu}{kT_K}} \]  

(2.16)

\[ C_{21} = n_{\text{H}_2}\langle \nu \sigma_{21}\rangle \]  

(2.17)

\[ J_\nu = S_\nu(1 - e^{-\tau}) + B_\nu(T_R)e^{-\tau} \]  

(2.18)

\[ S_\nu \equiv \frac{j_\nu}{\kappa'(\nu)} = \frac{n_2A_{21}}{n_1B_{12} - n_2B_{21}} = B_\nu(T_{\text{ex}}) \]  

(2.19)

\[ \tau = \frac{h\nu}{4\pi\Delta\nu} B_{21}X N_{\text{H}_2} \frac{e^{\frac{\nu}{kT_{\text{ex}}}} - 1}{\frac{g_1}{g_2} e^{\frac{\nu}{kT_{\text{ex}}}} + 1} \]  

(2.20)
where \( X \) is the abundance of particle \( X \) with respect to \( \text{H}_2 \), \( N_{\text{H}_2} \) the column density of \( \text{H}_2 \) molecules, and \( \tau \) the mean opacity over the line profile.

In principle, one can obtain \( T_{\text{ex}} \) as a function of the relevant physical parameters, namely \( n_{\text{H}_2} \), \( T_K \), and \( T_R \). However, analytical solutions are possible only in the two limits \( \tau \gg 1 \) and \( \tau \ll 1 \). These are:

\[
T_{\text{ex}} \approx T_K
\]

for \( \tau \gg 1 \), and

\[
T_{\text{ex}} \approx \frac{h \nu}{k} \ln \left[ 1 + \left( \frac{e^{h \nu / k T_K} - 1}{e^{h \nu / k T_K} - 1} \right)^{-1} n_{\text{H}_2} / n_{\text{cr}} \right]
\]

for \( \tau \ll 1 \). \( n_{\text{cr}} \equiv A_{21} / \langle \nu \sigma_{21} \rangle \) is named critical density. A plot of \( T_{\text{ex}} \) as a function of \( n_{\text{H}_2} \) for given \( T_K \) and \( T_R \) is shown in Fig. 1. As expected (see Eq. 2.22), one has \( T_{\text{ex}} \approx T_R \) for \( n_{\text{H}_2} \ll n_{\text{cr}} \) and \( T_{\text{ex}} \approx T_K \) for \( n_{\text{H}_2} \gg n_{\text{cr}} \). This means that at low densities the particles attain equilibrium at the radiation temperature, because collisions become too rare to be relevant. Vice versa, at sufficiently high densities the radiation field cannot affect the level populations, which obey the Boltzmann equation at the gas temperature. LTE at temperature \( T_K \) occurs also when \( \tau \gg 1 \), since the radiation field cannot penetrate the cloud and the particles are exposed only to collisions with \( \text{H}_2 \). This explains why the analytical solution (solid line in Fig. 1), derived in the optically thin approximation, approaches \( T_K \) at higher densities than the numerical solution taking properly into account the optical depth (dashed line in Fig. 1).

### 2.2.2 The 3-level system

Let us now consider a 3-level system. Albeit more complicated, the determination of the level population is basically the same as in the 2-level case. Instead of one \( T_{\text{ex}} \), here two independent excitation temperatures can be defined, one for levels 1 and 2, \( T_{21} \), and another for levels 2 and 3, \( T_{32} \). The excitation temperature between levels 1 and 3 can be expressed as a combination of \( T_{21} \) and \( T_{32} \) and will be ignored in our analysis.

The statistical equilibrium equations can be written as

\[
n_1(B_{12}J_\nu + C_{12} + B_{13}J_\nu + C_{13}) = n_2(A_{21} + B_{21}J_\nu + C_{21}) + n_3(A_{31} + B_{31}J_\nu + C_{31}) \quad (2.23)
n_2(A_{21} + B_{21}J_\nu + C_{21} + B_{23}J_\nu + C_{23}) = n_1(B_{12}J_\nu + C_{12}) + n_3(A_{32} + B_{32}J_\nu + C_{32}) \quad (2.24)
\]

and the following relationships hold:

\[
\frac{n_i}{n_j} = \frac{g_i}{g_j} e^{-\frac{h \nu}{k T_{ij}}} \quad (2.25)
\]

\[
\frac{C_{ji}}{C_{ij}} = \frac{g_i}{g_j} e^{-\frac{h \nu}{k T_K}} \quad (2.26)
\]

\[
C_{ij} = n_{\text{H}_2} \langle \nu \sigma_{ij} \rangle \quad (2.27)
\]

\[
J_\nu = S_\nu (1 - e^{-\tau_{ij}}) + B_\nu(T_{ij}) e^{-\tau_{ij}} \quad (2.28)
\]

\[
S_\nu \equiv \frac{j_\nu}{\kappa} = \frac{n_i A_{ij}}{n_j B_{ji} - n_i B_{ij}} = B_\nu(T_{ij}) \quad (2.29)
\]
Figure 1: Plot of the excitation temperature of a 2-level system subject to collisions with \( \text{H}_2 \) molecules with temperature \( T_K = 50 \text{ K} \) and density \( n_{\text{H}_2} \), and interacting with a black-body radiation field with temperature \( T_R = 2.7 \text{ K} \). The solid curve corresponds to an approximate analytical solution obtained in the optically thin limit, while the dashed curve is a numerical solution taking into account the gas optical depth, given by the dotted curve. The shadowed area marks the values of \( n_{\text{H}_2} \) for which \( \tau > 1 \). Note how \( T_{ex} \) rapidly approaches \( T_K \) when \( n_{\text{H}_2} > n_{cr} \), in the optically thin case, and when \( \tau > 1 \) in the general case.

As before, one can obtain an approximate analytical solution in the optically thin and thick limits:

\[
\tau_{ij} = \frac{\hbar \nu}{4\pi \Delta \nu B_{ij} X N_{\text{H}_2}} \frac{e^{\nu T_{ij}} - 1}{\nu e^{\nu T_{ij}} + 1} \tag{2.30}
\]

As before, one can obtain an approximate analytical solution in the optically thin and thick limits:

\[
T_{ij} \simeq T_K \tag{2.31}
\]

for \( \tau \gg 1 \), and

\[
T_{21} \simeq \frac{h \nu_{21}}{k} \ln \left[ \frac{g_1}{g_2} \langle A_{31} + B_{31} J_0 + C_{31} \rangle \langle B_{12} J_0 + C_{12} \rangle \langle B_{12} J_0 + C_{12} + B_{12} J_0 + C_{13} \rangle \right] \tag{2.32}
\]

\[
T_{32} \simeq \frac{h \nu_{32}}{k} \ln \left[ \frac{g_1}{g_2} \langle A_{31} + B_{31} J_0 + C_{31} \rangle \langle B_{12} J_0 + C_{12} \rangle \langle B_{12} J_0 + C_{12} + B_{12} J_0 + C_{13} \rangle \right] \tag{2.33}
\]

for \( \tau \ll 1 \). Note that this time two critical densities are defined: \( n_{21} \equiv A_{21} / \langle \nu \sigma_{21} \rangle \) and \( n_{32} \equiv A_{32} / \langle \nu \sigma_{32} \rangle \). A plot of \( T_{ij} \) as a function of \( n_{\text{H}_2} \) is shown in Fig. 2 (solid curves). For the same reasons
Figure 2: Plot of the excitation temperature of a 3-level system subject to collisions with H$_2$ molecules with temperature $T_K = 10$ K and density $n_{H_2}$, and interacting with a black-body radiation field with temperature $T_R = 2.7$ K. The two curves correspond to the excitation temperatures obtained from an analytical solution in the optically thin limit.

discussed in the case of the 2-level system, the excitation temperatures tend to the radiation and gas temperatures, respectively at low and high densities. The new feature with respect to Fig. 1 is the fact that for intermediate densities $T_{21}$ becomes even greater than the gas temperature. Naïvely, one can explain this effect because in our example $A_{32} \simeq 10A_{21}$, which makes transitions to level 2 favoured with respect to those away from it. As a result level 2 becomes overpopulated with respect to what is expected in thermodynamic equilibrium at the gas temperature, $T_K=10$ K. Such an effect becomes dramatic for $T_K = 50$ K (see Fig. 3), because the population of level 2 with respect to 1 becomes greater than what can be achieved for any physical temperature in thermodynamic equilibrium, so that the population ratio $n_2/n_1$ can be described by the Boltzmann equation only for $T_{21} < 0$. This phenomenon is called “population inversion” and – as explained in Sect. 2.1 – is at the basis of the maser phenomenon.

3. Astronomical Masers

In the light of the discussion in Sects. 2.1 and 2.2, one can define a maser line as the one originating from a transition between two levels undergoing population inversion. According to
In practice, how can one establish that $T_{\text{ex}} < 0$? An example is given by Rieu et al. [5], who observed the 1720 MHz line of the hydroxil radical (OH) emitted from a molecular cloud intercepting the line of sight to a background quasar. They noticed that, while the 1665 and 1667 MHz transitions of OH were seen in absorption against the quasar, the 1720 MHz line was in emission. They could determine the values of $T_{\text{ex}}$ and $\tau$ from the solutions of the radiative transfer equations for the line of sight to the quasar ($T_{\text{B on}}$) and another one passing very close to it ($T_{\text{B off}}$):

$$T_{\text{B on}} = \left( T_{\text{ex}} - T_{\text{bg}} \right) \left( 1 - e^{-\tau} \right)$$

$$T_{\text{B off}} = T_{\text{ex}} \left( 1 - e^{-\tau} \right).$$

Here $T_{\text{bg}}$ is the quasar continuum brightness temperature. These hold under the (reasonable) assumption that the cloud properties are basically the same along the two lines of sight. Rieu and collaborators derived $T_{\text{ex}} = -10$ K and $\tau = -0.1$. This result provides us with a typical example of “weak” maser, as opposed to “amplifying” masers for which $|\tau| \gg 1$, the latter being characterised by their intense brightness. The best example of an amplifying maser is the 22 GHz line of H$_2$O.

To date a large number of molecular species with one or more maser lines have been discovered. In the following we will summarize the major maser properties and phenomenology. The
interested reader can refer to a number of excellent books, reviews, and articles to improve his knowledge on this intriguing phenomenon [1, 2, 3].

### 3.1 How to recognize a maser line

Clearly, the method adopted by Rieu et al. to identify a maser line may be applied to a handful of cases. However, there are several criteria that are quite effective to establish if a transition is masering, although most of them are not iron clad proofs. Also, it must be said that most of these work only with amplifying masers. The most important ones are the following:

**Large** $T_B$. As demonstrated by Eq. (2.13), the maser amplification mechanism can produce large brightness temperatures, in excess of any sensible physical temperature of the molecular gas. This is a clear indication of the fact that one is observing a maser line.

**Point-like source.** The exponential amplification mechanism not only produces large $T_B$, but also privileges a few lines of sight where $\tau$ is much larger and hence $T_B$ is much brighter than along adjacent lines of sight corresponding to even slightly smaller values of $\tau$. This makes the apparent angular size of the source extremely small, basically unresolved in many cases.

**Narrow line width.** For the same reason causing small source sizes, also the width of a maser line is quite small, as the brightness of the central part of the line (where $\tau$ is maximum) is boosted with respect to the line wings.

**Multiple lines.** Since the maser emission is coming from point-like sources corresponding to optimal amplification paths where velocity coherence is attained, one commonly observes a few of these privileged lines of sight across a the cloud, each of these “spots” emitting a distinct spectral feature. This causes the presence of multiple lines at different velocities in the maser spectra.

**Time variability.** The maser intensity depends exponentially on $\tau$ and is hence extremely sensitive to even small changes of the amplification path (due to changes of the gas physical parameters, break of velocity coherence, etc.). In the most extreme cases, maser emission can undergo changes by orders of magnitude in intensity and several km s$^{-1}$ in velocity over a few weeks.

### 3.2 Usefulness of maser lines

Maser emission is observed from a variety of species and in many astronomical environments. In our Galaxy, maser emission has been detected in both molecular clouds hosting star formation (from H$_2$O, OH, CH$_3$OH, SiO, H$_2$CO, NH$_3$, Hn$\alpha$, etc.) and the atmospheres of old late-type stars (from H$_2$O, OH, SiO, HCN, etc.). Powerful maser lines have been also found in star-burst galaxies (H$_2$O, OH, H$_2$CO) and active galactic nuclei (H$_2$O, OH).

Given such a widespread presence of maser emission, it is worth finding out what kind of information can be obtained from the observation of maser lines. Unfortunately, it is difficult to derive the gas physical parameters, because, for instance, $T_{\text{ex}}$ has nothing to do with the gas kinetic temperature and $\tau$ is not trivially related to the gas column density. Nonetheless, maser emission has two advantages which make it ideal for high-angular resolution measurements: it can be very
bright and point like. These makes astronomical masers an ideal target for Very Long Baseline Interferometry (VLBI) observations, which allow us to derive a number of important quantities:

**3-D velocity.** Since maser spots are basically unresolved even with VLBI observations achieving angular resolutions of 1 milli-arcsec, they can be used as “target particles” to study the motion of the gas from which the maser emission originates. Multi-epoch imaging allows to derive the spot proper motion and thus obtain the velocity in the plane of the sky. This, combined with the velocity along the line of sight obtained from the Doppler shift of the line, makes it possible to derive the 3-D velocity of the gas.

**Distance.** Measurements of the apparent proper motion of the maser spots (parallax) is used to estimate the distance to the associated source up to several kpc.

**Magnetic field.** Zeeman splitting of the maser lines has been used for some species (OH and H$_2$O) to estimate the magnetic field on scales as small as a few AU.

The previous list is certainly not exhaustive of the many possible applications of maser studies. These have shed light on a variety of issues, from the study of jets in the early stages of star formation to that of circumnuclear disks in AGN. In particular, the long observing wavelengths of masers such as H$_2$O or CH$_3$OH make it possible to reveal the existence and determine the exact location of forming stars inside heavily obscured molecular condensations, otherwise inaccessible even at IR wavelengths. We conclude mentioning what is probably the most impressive result obtained so far: the determination of the mass of the black hole at the center of the galaxy NGC4258 [4]. This could be achieved by reconstructing the velocity field of the Keplerian disk rotating about the nucleus, by means of VLBI observations of the H$_2$O masers associated with it.

**References**


