RADIATIVE TRANSFER: PROGRESS AND PROBLEMS

V. Ossenkopf

1. Physikalisches Institut der Universität zu Köln, Zülpicher Straße 77, 50937 Köln, Germany

ABSTRACT

Radiative transfer computations are the basic tool to derive the physical parameters of the interstellar medium from astronomical observations. However, many computations are still hindered by severe problems. The uncertainties start with the input data where the dust scattering functions, the line frequencies in the infrared and the collision rates are poorly known. They continue with problems in the treatment of coherent radiation, partial redistribution, and the representation of maser spots.

Nevertheless, numerical simulations of radiative transfer experienced a tremendous progress over the last years. Molecules with hundreds of levels including rotational-vibrational transitions are simulated. The codes start to include the correct treatment of continuum pumping and overlapping lines, and first steps are made to deal with maser beaming, polarisation and partial coherence.

We get a better feeling which approximations are to be applied in certain situations allowing rapid data reduction but acknowledging that there is no standard way. Growing experience from the simulations shows, however, that a large part of the former knowledge on the interpretation of measurements has to be revised. In the analysis of molecular line data ambiguity of the solutions is the rule, not the exception. The results confirm the general wisdom that numerical models can explain many observations but do not necessarily lead to a true physical and intuitive understanding of radiative transfer in the interstellar medium.

Key words: Radiative transfer, Molecular data, Line: formation, Line: profiles

1. INTRODUCTION

Astronomical observations of the interstellar medium can only provide the information carried by the radiation arriving on earth. However there is no direct way to derive the properties of the radiating matter. Radiative transfer (RT) models are required to compute the relation between these properties and the emergent radiation.

In the wavelength range covered by ISO observations the radiation is mixed of contributions from molecular and atomic lines, dust continuum, and emission bands, probably due to PAHs and other large molecules. Hence, all RT computations should include continuum and line radiation in a combined way.

Nevertheless, this review can not cover all aspects but the examples will be biased to the long wavelength range and line radiation simply because this is the field of my main experience. Water lines in the far infrared will provide one of the main examples. Water is an important coolant in several processes and the lines probe an extremely wide range of physical conditions. The lines are strongly affected by radiative pumping from the dust continuum. From the viewpoint of RT computations, water is one of the most challenging problems due to the coexistence of optically very thick, optically thin and masering transitions within the same frequency range for typical cloud conditions.

2. WHAT DO WE NEED ?

2.1. Need for radiative transfer computations

To judge the importance of RT computations we have to consider how information can be retrieved from astrophysical observations. This is schematically shown in Fig. 1. Whereas the space of physical parameters describing a molecular cloud is at least 24 dimensional, measurements can only provide an intensity distribution at the sky as function of frequency[1]. Thus it is principle impossible to extract all information about a cloud from the observations. Ambiguities are unavoidable.

The problem must be tackled by an - often implicit - model for the interstellar object. This model my be given by the assumption of a unique temperature when using rotation diagrams for the line analysis, the assumption of homogeneous conditions when using the escape probability analysis or a more sophisticated model. RT computations are needed to derive observable quantities like intensities distributions, line profiles or polarisation patterns from the models[2]. The comparison of these observable quantities with the real observations constrains the models determining their parameters or detecting violations. One

[1] Often some additional information can be retrieved from polarisation and coherence.

[2] They are also required to make models self-consistent.

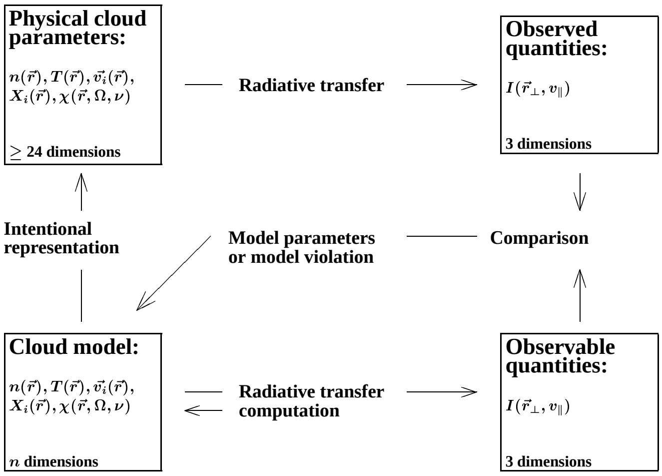

Figure 1. Sketch of the general way astrophysical information can be obtained from the observations. Knowledge comes from the determination of model parameters or the model violation when comparing computed and observed quantities.

has to keep in mind, however, that there is no direct access to the interstellar clouds. We will only learn about the models. The growth of knowledge comes from the iterative process of making "relevant" observations and improving the models. Here, the RT computations play a key role.

2.2. Need for new input data

In the equation system which has to be treated for a self-consistent solution of the radiative transfer problem numerous input parameters are still poorly known.

A first unknown is the frequency of many lines. The line predictions by Schilke et al. (1999) disclose a deficiency of data at frequencies above 1 THz. Here, laboratory data are often lacking. Without them it is impossible to compute the correct mutual pumping and shielding of different molecules and the interpretation of maser emission stimulated by line overlap.

The situation is somewhat better for the transition strengths. The comparison of typical models and experiments (Lynas-Gray et al. 1995, Coudert 1997) reveals errors and uncertainties of up to a factor 2 but typically well below 20 % for many water lines which seems acceptable regarding normal observational uncertainties.

A problem arises, however, when the collisional excitation dominates. The existing data bases for collisional rate coefficients (e.g. Green et al. 1995) are mainly based on computations performed for collisions with He atoms instead of H_2 molecules. A new computation by Phillips (1998) using H_2 collisions for the five lowest water levels showed that the rate coefficients may deviate by up to a factor 1000 from those computed for He collisions (Green et al. 1993). Until now, we have no consistent data set of water collision rates covering the important 30–50 lowest levels.

To estimate the possible effect of this uncertainty, we have used a toy model of a spherical isothermal cloud at 120 K with a power law r^{-2} density decay and computed the emergent line intensities either using the set of He based collision rates or by replacing the collision rates for the lowest five levels in this set by the H_2 based rates. Fig. 2 shows the resulting radial behaviour of the level populations and the profile of the 752 GHz line of para-H_2O.

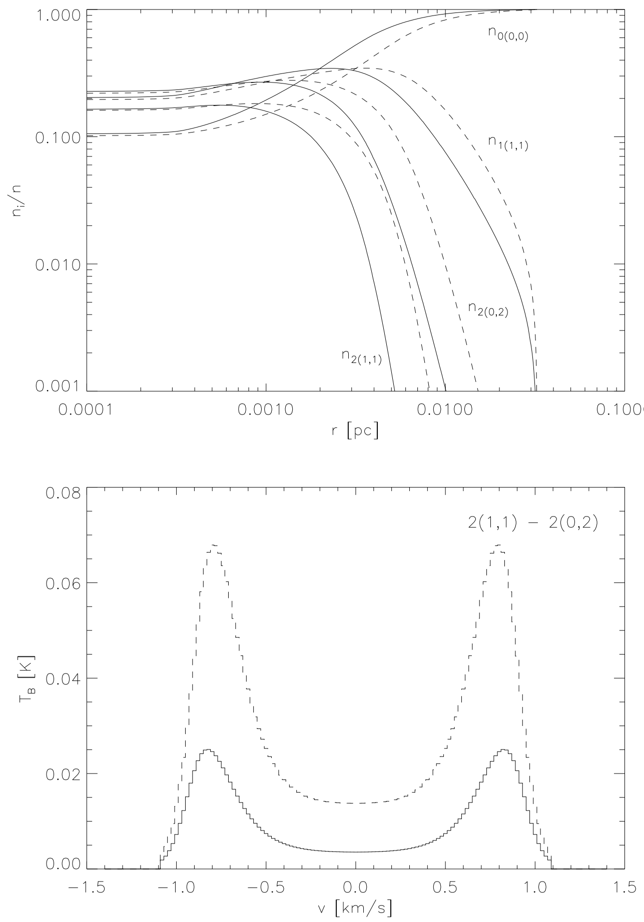

Figure 2. Radial dependence of four level populations of para-H_2O resulting from the toy model (upper panel) and the corresponding 752 GHz line integrated over the whole cloud (lower panel). The solid lines were computed using the He based collision rates and the dashed lines are results using the mixed set of collision rates collision rates with H_2 based rates for the lowest 5 levels.

The modification of the collision rates translates into a change of the line intensity by a factor 3. The effect is lower at smaller temperatures and in case of stronger infrared pumping but may grow up to a factor 10 when taking less realistic cloud parameters. Knowing that the many collision rates that we use today are wrong, we must admit that errors in the order of a factor 3 can occur when computing water abundances, cooling rates or the ortho/para-ratio from far infrared water lines.

In continuum RT the general behaviour of the dust opacity in molecular clouds throughout the infrared is today relatively well known from laboratory data (e.g. Ossenkopf & Henning 1994). However in the details of the band structure, determined by the chemistry and structure of interstellar grains and PAHs, many questions are left which are discussed elsewhere in this volume (see Henning et al., Joblin et al.). One of the big unknowns is the spatial scattering function. Almost all existing RT codes assume isotropic scattering and emission neglecting that the angular scattering function of any dust grain is anisotropic and – even worse - unknown for realistic grain shapes (Bohren & Huffman 1983). Combined effort in the solid-state physics of small particles and the numerical solution of Maxwell's equations under complicated boundary conditions is needed to obtain reliable scattering functions.

These examples demonstrate only a few open questions and many more open questions could be added.

3. WHAT DO WE KNOW ?

Over the last years the numerical simulation of RT experienced a tremendous progress. The development in computer technology pushed codes that are self-consistent in the sense that they compute the local excitation and temperature from the assumed cloud geometry in the same RT frame like the emergent brightness distribution and line profiles. On the other hand new approximations are developed for scenarios that are still beyond the capabilities of the self-consistent codes. They use sophisticated guesses for the local excitation, estimating the radiative interaction of separate cloud regions and shielding effects.

Here, I will introduce only a few aspects of the RT computations and discuss how they are treated in some selected codes. The sketch is far from providing a complete overview and favours codes that are used by me and my collaborators – not because they are more advanced than competitive codes but simply because I know them better.

According to the strong interaction of line and continuum radiation throughout the whole wavelength range observed by ISO a coupled treatment of line and continuum RT is required for many questions. However, a fully self-consistent treatment of both branches, taking the influence of line cooling onto the dust temperature and of the continuum transfer onto the line excitation into account, is only used by Doty & Neufeld (1997) (see also Doty 1999). Most codes still perform a separate treatment or consist only of one of these components. Even if the line and continuum RT problem is not solved simultaneously the interaction can be taken into account in an approximate way. González-Alfonso et al. (1998) showed that the simple coupling of water lines to a continuum radiation field was able to explain the observed near and mid infrared intensities. I will follow the traditional way discussing both branches separately when introducing the different aspects of RT computations.

Dimensionality: In line RT 2-D models became a standard (Hogerheijde et al. 1999, Phillips 1998, Dullemond 1998) and first 3-D simulations are performed (e.g. Juvela 1997). Numerous 1-D codes are available. A collection of references with competitive test cases is provided by van Zadelhoff et al. (1999). In continuum transfer several 3-D codes have been introduced (e.g. Wolf et al. 1999).

Resolution: As the dimensionality of the problem grows, more sophisticated methods are necessary to obtain a reasonable spatial resolution in the simulations. Taking e.g. the 3-D Monte-Carlo line RT code of Juvela (1997), current computer technology constrains the spatial grid to 32^3 points. Applied to the hydrodynamic simulations of Padoan et al. (1998) this is still insufficient to resolve all main structures. A possible solution is to apply nested grids, already implemented in some transfer codes (Wolf et al. 1999, Fabiani Bendicho et al. 1999). They can cover several orders of magnitude in length scale for objects like protoplanetary disks or multiple systems but can hardly help for fractal turbulent clouds.

Another way is the computational separation of the RT at separated scales. A possible approximation is the use of escape probability methods on the small scale embedded in a low-resolution RT code at the large scale (e.g. Ossenkopf et al. 1999). Another approach is the combination of 1-D RT codes at different scales to describe a 3-D problem. Krügel & Siebenmorgen (1994) simulated hot spots in M82 combining 1-D continuum RT computations for the environment of hot stars with a large scale 1-D RT scenario for the inner galaxy. The same kind of separation is used in line RT computations by Zielinsky et al. (2000). They combine a spherical 1-D code with a plane-parallel configuration describing a clumpy photo-dissociation region. This treatment is not self-consistent but provides a reasonable approximation for effects which would be otherwise unaddressable due to resolution limitations.

Turbulence: A physical effect directly related to the resolution problem is the treatment of turbulence stretching over many orders of magnitude in length scale. In most cases a full representation of turbulence is neither required nor possible. Statistical approaches are more appropriate. One can either assume a certain statistics of the coherent structures and simulate the behaviour of a large number ensemble (Martin et al. 1984) or assume certain statistical properties of the fluctuation spectrum (Kegel et al. 1993). Then it is possible to incorporate a representation of turbulence on the microscale into larger RT codes for particular cloud geometries (e.g. Piehler & Kegel 1995, Ossenkopf et al. 2000).

Characteristics: The most time consuming computation in the numerical solution of the RT problem is to follow a large number of rays required to compute the radiative energy density at each point of the cloud. Advanced techniques are required to find optimum spacings and directions to get a good angular coverage at each point with as few rays as possible. Using long characteristics, i.e. rays running through the whole cloud, adaptive ray gridding (e.g. González-Alfonso & Cernicharo 1997) provides an efficient method. In 2-D and 3-D codes, short characteristics, i.e. rays connecting only neighbouring or close points (e.g. Dullemond 1998) are more efficient. An alternative is the use of random characteristics, where the angular coverage is given by Monte-Carlo integration or Monte-Carlo ray-tracing (e.g. Hogerheijde et al. 1999). Juvela (1997) has shown that quasi-random distributions are favourable here.

Saturation: In parts of a cloud which are optically thick with respect to the considered radiation convergence of an ordinary non-local solver is extremely slow because the information propagates in each step of the iteration process only by the free path length of the photons. This can be circumvented applying a local diffusion approximation in continuum transfer (see e.g. Chick et al. 1996) or the more general approximate Λ operator (Olson et al. 1986 or Rybicki & Hummer 1991, 1992). Here, all contributions from the local scattering are implicitly included in a modified system of balance equations, virtually removing them from the RT computation. In line transfer the problem can be addressed in a similar way by a core saturation scheme based on different treatments of photons in the optically thick line center than in the thin line wings (Hartstein & Liseau 1998). This was used e.g. by Juvela (1999) providing a considerable speed-up.

Acceleration of convergence: Even with an approximate Λ operator convergence can be quite slow. It is desirable to guess the converged solution already from a few iteration steps. This kind of acceleration can be performed by extrapolating the changes to minimize residuals (Ng 1974) or by constructing conjugate gradients in the space of residuals (Klein et al. 1989). A comparative overview for both methods was given by Auer (1991). The acceleration methods are numerically simple and improve the convergence drastically.

Line overlap: Whereas traditional line RT codes compute each line individually, energy transfer between different lines can be essential both concerning different transitions of the same molecule and lines from different species. González-Alfonso & Cernicharo (1997) treated the line overlap defining frequency subgroups in a 1-D RT code. Taking the mutual pumping and shielding of different SiO isotopes into account this provided an self-consistent explanation of the observed maser intensities.

Number of species: The example demonstrates that it is necessary to solve the problem simultaneously for all species that may be radiatively coupled. For line RT a detailed analysis of the line positions and Doppler widths is required. The situation is obvious but more difficult for the continuum transfer. All dust materials, sizes, and shapes have to be considered. Taking the possible aggregation of dust into account this opens up a huge parameter space of dust species (Ossenkopf 1993). Some models (e.g. Efstathiou et al. 2000) take the different materials and grain sizes into account, but there exist only first attempts to incorporate a possible shape distribution (e.g. Wolf et al. 2000).

Non-equilibrium effects: For small dust grains the additional effect of temperature fluctuations has to be taken into account. The grains have to be treated as quantum systems like the molecules using a distribution of excitation levels or temperatures. This was addressed by Guhathakurta & Drain (1989) and Siebenmorgen et al. (1992) explaining the observed mid-infrared excess in regions with high ultraviolet flux. Non-equilibrium has to be considered too in systems with high velocities like supernova shells where the material can travel to regions with different excitation conditions within the lifetime of an excited state.

Masering: A moderate inversion of some level populations is typical for many molecules under various interstellar conditions and is taken into account by many RT codes. However, the simulation of saturated masers faces additional problems because coherence becomes a decisive quantity. It is no longer sufficient to compute the intensity of the radiation but phase and polarisation play a role. This results in effects like strong amplification, spot confinement, beaming, partial coherence, and polarisation and requires special techniques in RT. A comprehensive overview on these topics and their treatment in RT was provided by Field (1998).

Polarisation: Beside maser radiation which is often strongly polarised every dust scattering event leads to a partial polarisation. Including scattering polarization in their RT code Wolf & Henning (1999) were abel to derive strong constraints on the geometry of protostellar disks from the observation of polarisation patterns..

Today RT codes have addressed many of the questions and problems mentioned above. Although there is not yet any code dealing with all of them and quite a lot of work has to be added, we are not far from a self-consistent description of RT in the interstellar medium.

4. WHAT DO WE LEARN ?

4.1. How to learn from the numerical computations ?

Taking the possibility to simulate the RT in almost any configuration we are left with the question how to use these codes.

There is a huge space of parameters growing quickly when going from 1-D to 2-D or from 2-D to 3-D. Without additional assumptions, it is impossible to constrain all spatial variations of density, temperature, chemical abundances, velocities etc. from the fit of the observations. Young et al. (1997) derived parameters for a 1-D model from the fit of the observed position velocity diagram in one ammonia line and showed that no more than 7 parameters could be reliably determined. Ossenkopf et al. (2000) found that the simultaneous fit of the profiles and spatial extent of 3 to 4 different CS transitions can provide only 5 or 6 parameters in a 1-D spherical model.

Observable intensities, including the effects of saturation and self absorption, are mainly determined by the size, density, temperature, and gradients in the main exciting region, and less by accurate geometric constraints. A 1-D code is sufficient to deduce these quantities from the observations. Models beyond 1-D are justified, if we have either clear evidence from the observational data or the known physical processes that spherical approaches must fail or if we want to interpret the fine structure of the velocity profiles, e.g. to distinguish infall from rotation. Regarding the number of parameters at higher dimensions it is impossible to retrieve reliable data for all of them without additional constraints from sophisticated theoretical models.

The interpretation of observations needs a balance between the amount of information observed, taking into account the observational limitations like noise and resolution, the assumed or known complexity of the interstellar region used in the modelling, and the complexity of the line formation process. It is for instance useless to construct complicated geometrical models to improve the fit of sub-mm water lines by a few percent as long as we don't know exact collision rates and we do not model the complete radiative pumping from other species. Self-consistent RT computations should be performed to interpret most observational data but there is no general clue which model is to be applied for which data and the balance of information has to be reconsidered for every object.

Because of the natural ambiguity of the observations a numerical fit of observational data never guarantees a true representation of the physical structure. The fit of observational data by an adopted model does not provide much knowledge. We can learn more from an unsuccessful attempt to fit the data since this provides some real constraints. This opens some real understanding of the general relations from the numerical simulations. In general, the results of many runs with different models have to be examined to find the underlying physical laws. On the other hand these results can help us to create new analytic methods and approximations that will be implemented in more efficient numerical codes.

4.2. What about the traditional knowledge ?

Self-consistent numerical RT computations are available today which are at least in 1-D fast and reliable. Is there any way to interpret the results in terms of rules of thumb and how do they influence the traditional rules of thumb used as a kind of standard in the data reduction ? I can address only a few examples.

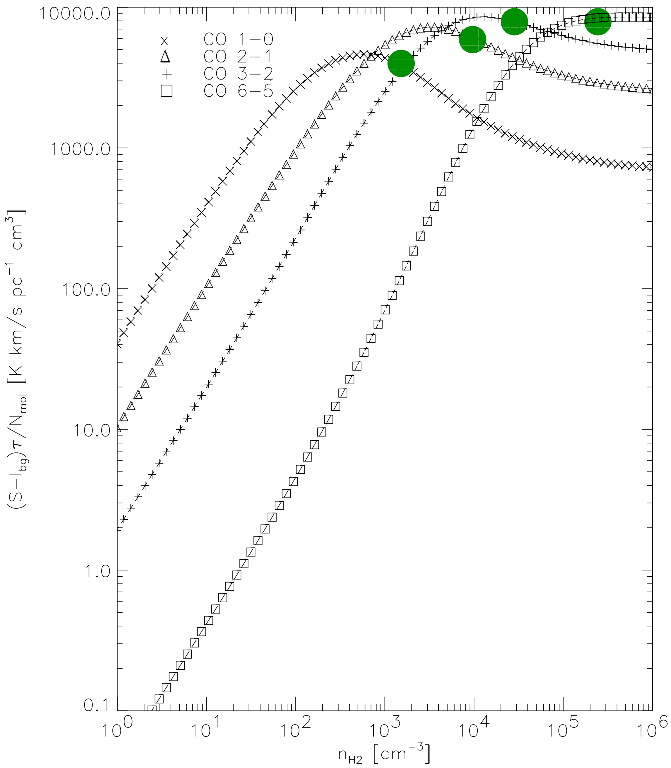

Figure 3. CO molecular emissivity (line intensity per column density in the optically thin case) as a function of the surrounding gas density at 100 K in four frequently observed transitions. The big dots on the curves for the four transitions mark the corresponding critical densities.

The first rules come from the correct treatment of the molecular level structure compared to a two-level system. In Fig. 3 we show the molecular emissivity in four CO lines obtained from the statistical equations without any RT, i.e. taking only the cosmic background radiation into account. The density of the surrounding gas at $T_{\rm kin} = 100$ K is varied. We can critically examine the two traditional concepts that **i)** optically thin lines measure the total column density, **ii)** strong lines indicate material at densities above the critical density of that transition. The plot shows that the emission in each line becomes negligible at very low densities. No line traces the full column density including thin material. Moreover we see that only in the high-J lines the critical density is a good measure for the onset of strong emission. In warm gas, the low-J emission can reach its equilibrium value already at densities a factor 100 below the critical density. The same computation can be done for each atomic system to get a more reliable first hand estimate.

Often escape probability methods, like the LVG approximation, are used to obtain some representative values for the density, column density and temperature of a cloud. Although the LVG approximation is a useful tool in the excitation analysis for many situations (Ossenkopf 1997), the systematic comparison with results from a self-consistent 1-D RT code (Ossenkopf et al. 2000) shows that it is rarely as reliable in data reduction. The derived column densities are accurate within of a factor of a few, but the escape probability provides no reasonable measure for the gas density and is useless for the temperature determination. Regarding the simplicity of 1-D codes they should be always favoured for a first data analysis.

A third rule that is often cited is to interpret asymmetric line profiles with enhanced blue emission in an optically thick transition combined with symmetric profiles in an optically thin transition as proof for infall motion. Computations of Phillips (1998) have shown, however, that the same kind of profiles can be produced by a rotation disk structure. There is no simple way out of this ambiguity and high-resolution observations should help to discriminate.

Most of the frequently applied estimates used for the data reduction should be reinspected in every case to guarantee that the underlying assumptions can be justified. Whenever this is not possible, full simulations are required.

5. CONCLUSIONS

Taking the accuracy of radio- or infrared observations into account the application of approximations in the data reduction and in the RT modelling is often a reasonable approach. Self-consistent models are required to justify the approximations and to test their limits. The common "standard" approximations will fail in many situations.

The complexity of the model used for RT simulations, both concerning the cloud geometry and the combination of species, has to reflect the amount of information from the observational data. Computational "overkill" does not help to understand the physical and chemical processes in the interstellar medium. Simple (e.g. 1-D) RT codes are today fast, easily accessible and appropriate for the analysis of many observations. They are favourable for a first exploration of the parameter space and the error analysis. More complicated codes are required to get an adequate modelling of configurations were geometry effects are obvious and important.

ACKNOWLEDGEMENTS

I want to thank J. Stutzki and C. Kramer for useful discussions improving this paper. The work has been supported by the Deutsche Forschungsgemeinschaft through grants SFB 301C and SFB 494B.

REFERENCES

Auer L.H. 1991, in: Crivellari L. et al. (eds.), Stellar Atmospheres: Beyond Classical Models, Kluwer Acad. Publ., 9
Bohren C.F., Huffman D.R. 1983, Absorption and Scattering of Light by Small Particles, Wiley New York
Chick K.M., Pollack J.B., Cassen P. 1996, ApJ 461, 956
Coudert L.H. 1997, J. Mol. Spectrosc. 181, 246
Doty S.D. 1999, AAS Meeting 195, 128.04
Doty S.D., Neufeld D.A. 1997, ApJ 489, 122
Dullemond C.P., Turolla, R. 1998, Abstracts of the 19th Texas Symposium on Relativistic Astrophysics and Cosmology, Paris
Efstathiou A., Rowan-Robinson M., Siebenmorgen R. 2000, MNRAS in press
Fabiani Bendicho P., Trujillo Bueno J., Auer L. 1997, A&A 324, 161
Field D. 1998, in: Hartquist Th. W, Williams D.A. (eds), The Molecular Astrophysics of Stars and Galaxies, Clarendon Press Oxford, p. 313
González-Alfonso E., Cernicharo J. 1997, A&A 322, 938
González-Alfonso E., Cernicharo J., van Dishoeck E.F., Wright C.M., Heras A. 1998, ApJ 502, L169
Green S., Maluendes S., McLean A.D. 1993, ApJSS 85, 181
Green S. et al. 1995, http://www.giss.nasa.gov/data/mcrates/
Guhathakurta P., Draine B.T. 1989, ApJ 345, 230
Hartstein D., Liseau R. 1998, A&A 332, 702
Hogerheijde M. R., van Dishoeck E. F., Salverda J. M., Blake G. 1999, ApJ 513, 350
Juvela M. 1997, A&A 322, 943
Juvela M. 1999, in: Ossenkopf V. et al. (eds.), The Physics and Chemistry of the Interstellar Medium, GCA-Verlag, p. 220
Kegel W.H., Piehler G., Albrecht M.A. 1993, A&A 270, 407
Klein R.I., Castor J.I., Greenbaum A., Taylor D., Dykema P.G. 1989, J. Quant. Spectrosc. Radiat. Transfer 41, 199
Krügel E., Siebenmorgen R. 1994, A&A 282, 407
Lynas-Gray A.E., Miller S., Tennyson J. 1995, J. Mol. Spectrosc. 169, 458
Martin H.M., Sanders D.B., Hills R.E. 1984, MNRAS 208, 35
Ng K.C. 1974, J. Chem. Phys. 61, 2680
Olson G.L., Auer L.H., Buchler J.R. 1986, J. Quant. Spectrosc. Radiat. Transfer 35, 431
Ossenkopf V. 1993, A&A 280, 617
Ossenkopf V., Henning Th. 1994, A&A 291, 394
Ossenkopf V. 1997, New Ast. 2, 365
Ossenkopf V., Bensch F., Zielinsky M.. 1999, in: Franco J., Carraminaña A. (eds.), Interstellar Turbulence, Cambridge Univ. Press, p. 252
Ossenkopf V., Trojan C., Stutzki J. 2000, A&A submitted
Padoan P., Juvela M., Bally J., Nordlund Å. 1998, ApJ 504, 300
Phillips T.R., Maluendes S., Green S. 1996, ApJSS 107, 467
Phillips R. 1998, in: Ossenkopf V. (ed.), The Physics and Chemistry of the Interstellar Medium - Abstract Book, Shaker-Verlag Aachen, p. 166
Piehler G., Kegel W.H. 1995, A&A 297, 841
Rybicki G.B., Hummer D.G. 1991, A&A 245, 171
Rybicki G.B., Hummer D.G. 1992, A&A 262, 209
Schilke P., Phillips T.G., Mehringer D.M. 1999, in: Ossenkopf V. et al. (eds.), The Physics and Chemistry of the Interstellar Medium, GCA-Verlag, p. 330
Siebenmorgen R., Krügel E., Mathis J.S. 1992, A&A 266, 501
Wolf S., Henning Th., Stecklum B. 1999, A&A 349, 839
Wolf S., Henning Th. 1999, A&A 341, 675
Wolf S., Henning Th., Stecklum B. 2000, High-Resolution Polarization Studies of Disks, in preparation
Young L.M., Keto E., Ho P.T.P. 1998, ApJ 507, 270
van Zadelhoff et al. 1999, http://www.strw.LeidenUniv.nl/~radtrans/
Zielinsky M., Stutzki J., Störzer H. 2000, A&A submitted