

PHYSICS OF COSMIC MASERS

V. S. Strel'nitskij
Astronomical Council, U.S.S.R. Academy of Sciences,
Moscow, U.S.S.R.

1. INTRODUCTION

Natural masers, which emit powerful narrow-band radiation in molecular microwave lines, were discovered about 20 years ago. They draw the astrophysicist's attention in two ways: 1) as a unique physical phenomenon and 2) as very informative probes of physical conditions in those astrophysical objects (very young and very old stars) with which they are intimately connected. In this paper, the first aspect of masers will be stressed, though some consideration will be given to the astrophysical aspect as well. The limited space here will not allow a complete discussion. For a more systematic account, refer to reviews on astrophysical masers by Litvak 1974; Strel'nitskij 1974; Kegel 1975; Cook 1977; Moran 1976; Reid and Moran 1981; Elitzur 1982. I shall limit myself only to several topics of general character, which, to me personally, seem to be particularly important for understanding the fundamental principles of cosmic maser physics. I shall then try to illustrate these general principles by the example of H₂O masers in star formation regions. These masers, in my opinion, are the most intriguing of all the cosmic masers known at present. And (to make the subjectivity of the report complete!) I shall lay special stress on the results obtained by our group in the last few years.

2. SOME HISTORICAL REMARKS

Cosmic maser radiation was first observed in 1965 – in OH radio lines near 18-cm wavelength (Weaver *et al.* 1965). It would be wrong to consider this discovery as a *complete* surprise. For example, long before the discovery of not only astrophysical, but also laboratory, masers the astrophysicist Menzel (1937) had pointed out the possibility, in principle, of the “negative absorption” (i.e., amplification) of radiation in spectral lines in the absence of local thermodynamic equilibrium (LTE). A really unexpected result was the *strength* of the maser effect detected in the interstellar medium: amplification in many sources attains $10^{10} - 10^{12}$, so that radiation brightness temperatures reach 10^{16} K, and the luminosity in a narrow radio line is sometimes comparable to the total (bolometric) luminosity of a solar class star (assuming isotropic emission from the maser).

Soon after the discovery of the OH masers Shklovskij (1966) suggested that the maser radiation was generated in the upper layers of protostars. It then became clear that a unique observational method for probing the star-formation regions became available to astrophysicists.

In 1969 the Berkeley group (Cheung *et al.* 1969) discovered the first interstellar H₂O maser at 1.35-cm wavelength ($6_{16} - 5_{23}$ rotational transition). It was soon found that some H₂O masers exhibit very broad spectra, up to several hundred km/s in radial velocity. Strel'nitskij and Sunajev (1972) pointed out that the simplest, kinematical, interpretation of such wide spectra automatically implies dynamical instability of the corresponding system of maser condensations. Thus, they predicted the proper motions of the maser spots of the order of milliarcseconds per year, a prediction that was confirmed a decade later (Genzel *et al.* 1981). They also suggested that the systems of maser condensations are systems of *light* (planetary masses) gas condensations blown away from very young stars by strong stellar winds or other mechanisms. Many subsequent observations by different techniques confirmed extremely strong gas outflows (up to $\sim 10^{-2} M_{\odot}/\text{year}$) from very young stars. As a rule, these flows have a well-defined "bipolar" structure (e.g., Bally and Lada 1983).

The first extragalactic masers, probably also related to the star-formation activity in other galaxies, were discovered in the late seventies (Churchwell *et al.* 1977). More than a dozen "masing" galaxies (in OH, H₂O, or H₂CO lines) are known now.

In 1968, the OH maser emission was first observed from the *old red giant* stars (Wilson and Barrett 1968). Later on, many of these stars were also found to have H₂O and SiO maser radiation.

Weak masers in CH₃OH (methanol) lines at 1.2 cm wavelength have been discovered in Orion by Hills, Pankonin, and Landecker (1975). Stronger methanol masers at 2.5 cm wavelength appear to be common (Batrla, Matthews, and Menten 1987). Weak maser effects have also been observed in the radio lines of some other molecules such as H₂CO and CH.

In parallel with the observations of cosmic masers was the development of maser theory. The major relations for the simplest, one-dimensional, model of a cosmic maser amplifier were formulated by Litvak *et al.* (1966) soon after the discovery of the OH masers. In 1972–74, a series of papers by Goldreich and others (e.g., Goldreich and Keeley 1972; Goldreich and Kwan 1974a,b) appeared, which considerably extended our understanding of the physical processes in astrophysical masers.

The study of cosmic masers is closely connected with the development of VLBI. The first interferometric measurements of masers (Rogers *et al.* 1966; Cudaback, Read, and Rougoor 1966) showed that their brightness temperatures were in excess of 10^6 K. Shortly thereafter, VLBI observations (Moran *et al.* 1968) revealed the extremely high brightness temperatures of the OH sources (as high as 10^{13} K), making the maser interpretation of these sources inevitable.

3. KINETICS AND THERMODYNAMICS OF COSMIC MASERS

The relaxation time scales in strong cosmic masers ($\sim 10^{-1} - 10^4$ s) are much smaller than the time scales of their variability, and this condition is believed to be sufficient for the assumption of *stationarity*. The level populations of a stationary

non-LTE system are described by the stationarity equations:

$$\begin{aligned} \sum_{k \neq i} n_k W_{ki} - n_i \sum_{k \neq i} W_{ik} &= 0, \\ \sum_k n_k &= N \quad (k = 1, 2, \dots, m), \end{aligned} \quad (1)$$

where n_k is the population of level k , W_{ik} is the total probability of the ik transition, m is the total number of the levels involved, and N is the total population. Transitions are produced by collisions of the “working” molecules with other particles, by their interactions with the radiation fields, and as a result of chemical processes. These external agents can be regarded as “thermostats” that tend to bring the level populations into equilibrium with their own temperatures. These temperatures need not be strict equilibrium temperatures; they can be defined as the inverse derivatives of the entropy with respect to the energy in nonequilibrium situations, or simply as convenient temperature-like parameters, which, together with another well-known temperature parameter, the excitation temperature of the transition, satisfy the usual rule of the direction of energy transfer between energy reservoirs (Strel'nitskij 1983).

Assuming some physical model of the maser, one can, in principle, fix all the rate coefficients in the system (1), or indicate the procedure by which they can be found. Solution of this system allows us to say whether the signal transition will or will not be *inverted* in such a physical situation, a condition for a maser to function. In practice, however, it is impossible to take into consideration all the possible processes, and one tries to *estimate* the most important nonequilibrium process leading to inversion: the *pump mechanism*. A clear understanding of the thermodynamical properties of nonequilibrium quantum-mechanical systems can help substantially in this analysis.

In a stationary system, the transfer of populations can only be *cyclic*. It can be shown (e.g., Strel'nitskij 1983) that each pump cycle leading to inversion of the signal transition works as a “heat engine.” As such, the pump cycle must have at least one step corresponding to the receipt of energy from a high-temperature thermostat and at least one corresponding to the sink of energy to a low-temperature thermostat. The net population transfer realizing these steps can be of radiative (R), collisional (C) or chemical (X) nature. Thus, all the pump mechanisms can be divided into nine types, according to the combination of the kinds of energy source and sink:

$$\begin{array}{l} RR \quad RC \quad RX \\ CR \quad CC \quad CX \\ XR \quad XC \quad XX \end{array} \quad (2)$$

To describe precisely the type of the pump one must also indicate the kind of transitions by which the pump cycles are constructed: electronic (e), vibrational (v), or rotational (r). The matrix (2) is a more informative and heuristic classification than the previous division of all pump mechanisms into radiative, collisional, or chemical, solely based on the kind of energy source. The scheme described here serves as a relatively good guide for the search of the adequate pump mechanism by the preliminary comparative analysis of possible pump schemes and crude estimates of their pump rates (the “equivalent transition approximation,” Strel'nitskij 1983,

1984). When a potentially effective mechanism is chosen in this manner, it can be investigated in more detail by solving the system of equations (1). The effective procedures to identify the real pump cycles and to describe the type of the pump after solving equation (1) were worked out by Sobolev (1987) and references therein.

4. MASER SATURATION AND LINE WIDTH

Neglecting spontaneous emission, which is not substantial for strong masers, one can give the one-dimensional transfer equation of maser radiation the following form (Goldreich and Kwan 1974a):

$$\frac{dI_\nu}{dz} = k_\nu (I_\nu^{-1} + I_s^{-1})^{-1} , \quad (3)$$

where

$$k_\nu = \frac{h\nu B}{4\pi} \frac{\Delta\lambda_\nu}{\Gamma} \quad (4)$$

is the amplification coefficient at frequency ν within the line, B is the Einstein B -coefficient, $\Delta\lambda_\nu$ is the difference of pump rates for the upper and lower maser levels (per unit volume, per unit frequency interval), Γ is the decay constant, describing the relaxation (collisional and/or radiative), which is assumed to be the same for both levels, h is Planck's constant, and

$$I_s = \frac{2\pi\Gamma}{B\Omega} \quad (5)$$

is the intensity at which saturation effects become important. What does saturation mean? The solution of equation (3) is:

$$\ln \frac{I_\nu}{I_\nu(0)} + \frac{I_\nu - I_\nu(0)}{I_s} = \int_0^z k_\nu dz \equiv \tau_\nu , \quad (6)$$

where z is the distance along the line of sight through the maser, and τ_ν is the optical depth. It is readily seen that when $I_\nu \ll I_s$, the intensity depends exponentially on the optical depth:

$$I_\nu \propto e^{\tau_\nu} . \quad (7)$$

However, when $I_\nu \gg I_s$ (i.e., when $2R \gg \Gamma$, $R \equiv I_\nu B\Omega/4\pi$, the stimulated emission rate, and Ω is the solid angle of the maser beam), I_ν depends only linearly on τ_ν :

$$I_\nu \propto \tau_\nu . \quad (8)$$

It must be stressed that the decay constant Γ , which determines the saturation condition, accounts for *all* the relaxation processes, both collisional and radiative. Ignoring this fact leads to misunderstandings.

One potential misunderstanding concerns line broadening. It is well known that the exponential dependence of I_ν on τ_ν in the case of unsaturated maser amplification leads to line narrowing relative to the Doppler width $\Delta\nu_D$:

$$\frac{\Delta\nu}{\Delta\nu_D} \approx \tau_0^{-0.5} , \quad (9)$$

where $\Delta\nu$ is the linewidth, and τ_0 is the optical depth at the line center (the proof can be found, for example, in Strel'nitskij 1982a). Comparing equations (7) and (9), we find that

$$\Delta\nu \propto (\ln I_0)^{-0.5}, \quad (10)$$

where I_0 is the intensity at line center. Hence, there is only a very weak inverse dependence of $\Delta\nu$ on I_0 .

In the last few years many cases of much stronger $\Delta\nu(I_0)$ dependence have been reported for H₂O sources:

$$\Delta\nu \propto I_0^{-\alpha}, \quad (11)$$

where α varies from ≈ 0.3 (Lada *et al.* 1981) to ≈ 1 (Rowland and Cohen 1986), with $\alpha \approx 0.5$ as the most typical value (Mattila *et al.* 1985; Rowland and Cohen 1986). In the last two papers, the $\Delta\nu \propto I_0^{-0.5}$ dependence is explained by the mechanism of radiative relaxation proposed by Goldreich and Kwan (1974a), but this explanation seems to be incorrect since it does not include the radiative relaxation in Γ (Strel'nitskij 1986). Another explanation for the observed line width dependence is described in the next section.

5. PHYSICS OF H₂O MASERS IN REGIONS OF STAR FORMATION

H₂O maser sources are particularly interesting because: 1) they are the strongest of all known masers; 2) they are probably connected with the earliest stages of stellar evolution observed in maser emission; 3) their kinematics promise new and very attractive possibilities of the distances measurements in the Galaxy and beyond (Reid, Moran, and Gwinn 1986, 1987).

The radiative-radiative and the collisional-radiative pump schemes (*RRv*, *CRr*, *RCv*, and *CRv*) that have been proposed (Litvak 1969; Strel'nitskij 1971; de Jong 1973; Goldreich and Kwan 1974b; Shmeld, Strel'nitskij, and Musilev 1976; Norman and Silk 1979) are barely capable of explaining the extremely strong sources such as W49, Ori A or W51. The basic reason is the existence of an upper limit for the number density in the maser condensation (above which the radiative population transfer is hindered by collisions). With this limitation and the limitation on the dimensions of the source determined by VLBI measurements, these pump schemes cannot account for the observed maser emission power (Strel'nitskij 1984).

The analysis of matrix (2) suggested to the author a new type of possible pump, *CC*, which is able to work, in principle, at any density (Strel'nitskij 1980, 1984). For the *CCr* pump, two kinds of colliding particles with different kinetic temperatures are required, e.g., H₂ molecules and electrons. The hot particles provide the source, and the cool particles provide the sink of the quantum heat engine. Solutions of the stationarity equations obtained by Bolgova (1981) and by Bolgova, Makarov, and Sobolev (1987) have shown a strong tendency of the *CCr* pumping with $T_H > T_e$ to invert the signal transition 6₁₆ – 5₂₃ as well as other transitions. This inversion takes place over a large range of fractional ionization ($10^{-5} - 10^{-1}$) and kinetic temperature ($\sim 10^2 - 10^3$ K). In these calculations, the poorly known collision probabilities were varied over a wide range.

In the astrophysical model of the H₂O maser with the *CCr* pump proposed by the author (Strel'nitskij 1984), the required condition $T_H > T_e$ arises behind the shock front in a circumstellar gas condensation impinged upon by a strong stellar

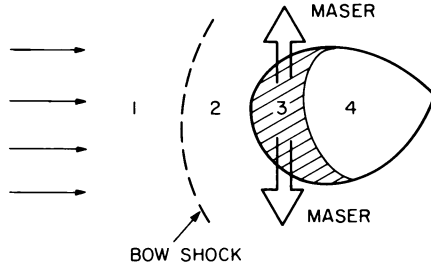


Figure 1. A stellar wind impinging on a maser condensation. 1—the unperturbed stellar wind; 2—the compressed gas behind the bow shock; 3—the gas of the condensation compressed by the internal shock; 4—the unperturbed gas of the condensation.

wind (regions 3 and 4 in Fig. 1). The required ionization can be produced by high-energy (*MeV*) protons accelerated by plasma instabilities behind the shock front (regions 1 and 2). The main advantage of this model, described in detail by Strel'nitskij (1984), is the possibility to explain the observed extremely high flux densities from the H_2O sources. It can also explain many other properties of the sources. Let us see here how the model can explain the observed strong anticorrelation of $\Delta\nu$ and I_0 mentioned in the preceding section. The optical depth in the line center is inversely proportional to $\Delta\nu_D$ if the pump rate, as is commonly assumed, does not depend on the velocities of the molecules:

$$\tau_0 \propto k_0 \propto \frac{\Delta\Lambda}{\Delta\nu_D} \quad , \quad (12)$$

where k_0 is the amplification coefficient at line center, and $\Delta\Lambda$ is the full difference of the pump rates for the upper and lower maser levels. The previous unsuccessful (see §4) explanations of the observed variations of I_0 and $\Delta\nu$ were based on the assumption that these variations are due to changes of $\Delta\Lambda$, leaving $\Delta\nu_D$ constant. Our hypothesis is that I_0 and $\Delta\nu$ change because $\Delta\nu_D$ changes, e.g., due to changes of the kinetic temperature in the amplification region.

We note now that in *all* stages of maser saturation, the observed $\Delta\nu$ will be proportional to $\Delta\nu_D$, whereas τ_0 is inversely proportional to $\Delta\nu_D$. Suppose that the maser is partly saturated. Then we can write $I_0 \propto \tau_0^\beta$ (see Fig. 2). For a large part of the $I_0(\tau_0)$ curve, β is between ~ 1 and ~ 3 , and one will have $\Delta\nu \propto I_0^{0.3-1}$, as observed.

In the model shown in Figure 1, the rise of kinetic temperature necessary to explain the rise of $\Delta\nu_D$ and the decrease of I_0 can be connected with perturbations in the stellar wind coming to the shock front 1–2. Any such perturbation will change the conditions behind the shock and in particular will change the number and energies of protons accelerated here and producing the ionization in the pump region 3.

At first, the wind condensation coming to the front 1–2 will increase the maser intensity due to these extra protons, which reach the maser region in less than 1 day due to their high velocities. The gas-dynamical perturbation, on the other hand, takes ~ 10 days to cross into region 2 and reach the amplification layer 3. Thus, within this period, I_0 rises, but $\Delta\nu$ remains almost unchanged. Then the perturbation reaches regions 2–3, and the kinetic temperature within the

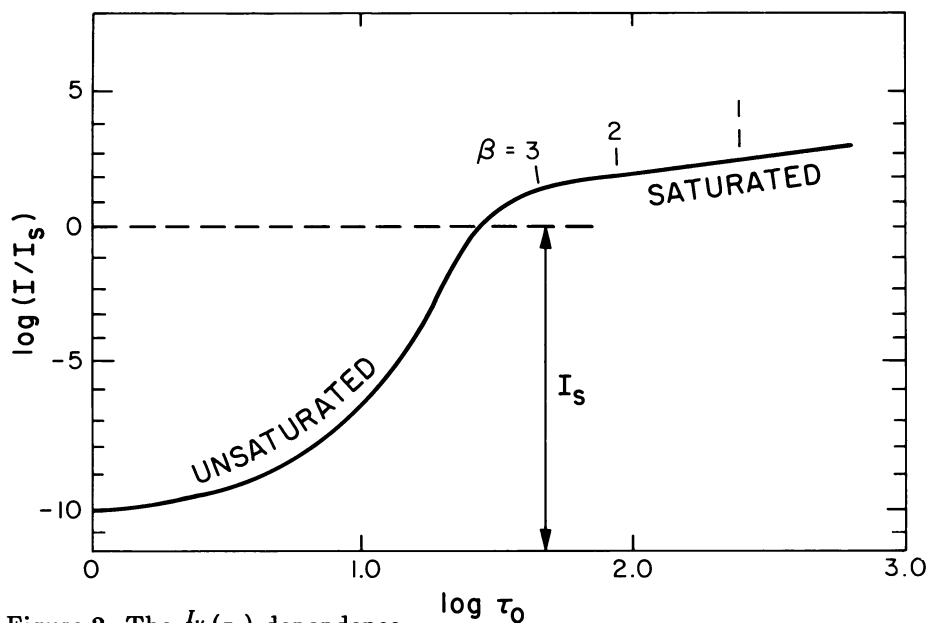


Figure 2. The $\frac{I}{I_s}(\tau_0)$ dependence.

amplification region begins to rise, increasing $\Delta\nu_D$ and decreasing I_0 . The time scale for this phase is of the order of sound propagation time through region 3, which is approximately several months. This simple model satisfactorily explains the typical time scales in H_2O maser outbursts observed by Rowland and Cohen (1986). It also explains why, for a well-studied, flare-up of the -11 km/s feature in Cepheus A (Rowland and Cohen 1986; Mattila *et al.* 1985), the increase of I_0 was not accompanied by changes of $\Delta\nu$, whereas the decrease of I_0 was. Recently Kylafis and Norman (1986) proposed a variant of the CCr -pump model in which $T_H \ll T_e$. According to these authors, such a situation can arise in the magnetic precursor of a MHD J-type wave.

6. CONCLUDING REMARKS

I would like to conclude with three suggested observations for VLBI specialists (without regard to their technical difficulty):

1. Try to measure the spatial evolution of outbursts described in §§4–5.
2. Try to measure the proper motion of the flared-up 8 km/s feature in Orion in order to discriminate between the “stationary” (Matveenko 1981) and “outflowing” (Strel'nitskij 1982a) models of this source.
3. Try to measure the time evolution of the H_2O maser distribution in Mira variables in order to discriminate between the “expanding shock front” model and “stationary base of the stellar wind” model of these sources (Strel'nitskij 1982b).

Good luck!

REFERENCES

- Bally, J., and Lada, C. J. 1983, *Ap. J.*, **265**, 824.
- Batrla, W., Matthews, H. E., and Menten, K. M. 1987, *Nature*, **326**, 49.
- Bolgova, G. T. 1981, *Nauchnye Inform. Astron. Sov. AN SSSR*, **47**, 9.
- Bolgova, G. T., Makarov, S. V., and Sobolev, A. M., 1987, *Astrofisika*, submitted.
- Cheung, A. C., Rank, D. M., Townes, C. H., Thornton, D. D., and Welch, W. J. 1969, *Nature*, **221**, 626.
- Churchwell, E., Witzel, A., Huchtmeier, W., Pauliny-Toth, I., Roland, J., and Sieber, W. 1977, *Astr. Ap.*, **54**, 969.
- Cook, H. N. 1977, *Celestial Masers* (Cambridge: Cambridge University Press).
- Cudaback, D. D., Read, R. B., and Rougoor, G. W. 1966, *Phys. Rev. Lett.*, **17**, 452.
- de Jong, T. 1983, *Astr. Ap.*, **26**, 297.
- Eltizur, M. 1982, *Rev. Mod. Phys.*, **54**, 1225.
- Genzel, R., Reid, M., Moran, J., and Downes, D. 1981, *Ap. J.*, **244**, 884.
- Goldreich, P., and Keeley, D. A. 1972, *Ap. J.*, **174**, 517.
- Goldreich, P., and Kwan, J. 1974a, *Ap. J.*, **190**, 27.
- Goldreich, P., and Kwan, J. 1974b, *Ap. J.*, **191**, 93.
- Hills, R., Pankonin, V., and Landecker, T. L. 1975, *Astr. Ap.*, **39**, 149.
- Kegel, W. H. 1975, in *Problems in Stellar Atmospheres and Envelopes*, ed. B. Baschek et al. (Berlin: Springer), p. 257.
- Kylafis, N. D., and Norman, C. 1986, *Ap. J. (Letters)*, **300**, L73.
- Lada, C. J., Blitz, L., Reid, M. J., Moran, and J. M. 1981, *Ap. J.*, **243**, 769.
- Litvak, M. 1969, *Science*, **165**, 855.
- Litvak, M. M., McWhorter, A. L., Meeks, M. L., and Zeiger, H. J. 1966, *Phys. Rev. Lett.*, **17**, 821.
- Litvak, M. M. 1974, *Ann. Rev. Astr. Ap.*, **12**, 97.
- Mattila, K., Holsti, N., Toriseva, M., Anttila, R., and Malkamäki, L. 1985, *Astr. Ap.*, **145**, 192.
- Matveenko, L. I. 1981, *Sov. Ast. Lett.*, **7(1)**, 54.
- Menzel, D. 1937, *Ap. J.*, **85**, 330.
- Moran, J. M. 1976, in *Frontiers of Astrophysics*, ed. E. Avrett (Cambridge: Harvard University Press), p. 385.
- Moran, J. M., Burke, B. F., Rogers, A. E. E., Carter, J. C., Ball, J. A., and Cudaback, D. D. 1968, *Ap. J. (Letters)*, **152**, L97.
- Norman, C., and Silk, J. 1979, *Ap. J.*, **228**, 197.
- Reid, M. J., and Moran, J. M. 1981, *Ann. Rev. Astr. Ap.*, **19**, 231.
- Reid, M. J., Moran, J. M., and Gwinn, C. R. 1986, in *Radio Astronomy in Space*, Proceedings of Green Bank Workshop, ed. K. W. Weiler (Green Bank, WV: NRAO), p. 145.
- Reid, M. J., Moran, J. M., and Gwinn, C. R. 1987, this volume.
- Rogers, A. E. E., Moran, J. M., Crowther, P. P., Burke, B. F., Ball, J. A., Meeks, M. L., and Hyde, G. M. 1966, *Phys. Rev. Lett.*, **17**, 450.
- Rowland, P. R., and Cohen, R. J. 1986, *M. N. R. A. S.*, **220**, 233.
- Schmeld, I. K., Strel'nitskij, V. S., and Musilev, V. V. 1976, *Astron. Zh.*, **53**, 728.
- Shklovskij, I. S. 1966, *Astron. Tsirc.*, **372**, 1.
- Sobolev, A. M. 1987, *Astron. Tsirc.*, submitted.
- Strel'nitskij, V. S. 1971, *Astron. Tsirc.*, **609**, 1.
- _____. 1974, *Usp. Fis. Nauk*, **113**, 463.

- _____. 1980, *Pis'ma v Astron. Zh.*, **6**, 354.
- _____. 1982a, *Pis'ma v Astron. Zh.*, **8**, 165.
- _____. 1982b, *Publ. Astrophys. Obs. Potsdam*, **110**, 121.
- _____. 1983, *Nauchnye Inform. Astron. Sov. AN SSSR*, **2**, 75.
- _____. 1984, *M. N. R. A. S.*, **207**, 339.
- _____. 1986, *Astron. Tsirc.*, **1465**, 1.
- Strel'nitskij, V. S., and Sunjaev, R. A. 1972, *Astron. Zh.*, **49**, 704 [*Sov. Ast. A. J.*, **16**, 579, 1973].
- Weaver, H., Williams, D. R. W., Dieter, W. H., and Lum, W. T., 1965, *Nature*, **208**, 19.
- Wilson, W. J., and Barrett, A. H. 1968, *Science*, **161**, 728.