Variability of Water Megamasers in AGN

Philip Maloney

CASA, University of Colorado, Boulder CO 80301
maloney@casa.colorado.edu

Abstract: Powerful water masers have been detected in 23 active galactic nuclei (AGN) to date. Most of the sources that have been monitored show evidence for variability, which may be either intrinsic (response of the masers to a time-varying pump or background source) or extrinsic (the effects of scintillation in the interstellar medium of the Galaxy). I briefly review the examples in which these mechanisms may be operating (interstellar scintillation in the Circinus galaxy, a time-varying background source in Mrk 348, and a time-varying AGN luminosity in NGC 1068).

Keywords: galaxies: Seyfert — masers — molecular processes

1 Introduction

Twenty-three extragalactic H_2O megamasers (with isotropic line luminosities reaching up to $L_{H_2O} \sim 6000\,L_\odot$) are now known (see Maloney 2002 for a recent comprehensive review). All are in Sy 2 or LINER galaxies, suggesting an intimate connection between the material responsible for obscuration of the central source and the masing gas. The maser emission typically consists of narrow (a few $km\,s^{-1}$) velocity components spread over $\sim 200\,km\,s^{-1}$ around the systemic velocity. The maser flux from those sources that have been monitored varies on timescales of weeks, with fluctuations typically of the order of tens of per cent. In this paper I briefly review the mechanisms that can produce the observed variability and what can be learned from them.

2 Maser Variability Mechanisms

There are at least four distinct mechanisms that can give rise to variability in the observed maser emission:

- geometry (size, direction, velocity coherence)
- excitation by time-variable pump
- amplification of time-variable background source
- interstellar scintillation

The first of these undoubtedly plays a role in some sources (such as the maser disk of NGC 4258); however, because many of the parameters involved are not known, it is difficult to make detailed inferences from the observations. I will discuss the remaining three mechanisms with specific reference to objects in which they may have been observed.

2.1 Interstellar Scintillation: Circinus

The Circinus galaxy is about 4 Mpc distant, and harbours one of the first H_2O megamasers to be discovered (Gardner & Whiteoak 1982); it possesses a heavily-obscured nuclear hard X-ray source and a compact non-thermal radio source. Line flux variations of $\sim$ tens of per cent have been seen on timescales $\tau_{var} \sim$ minutes

(Greenhill et al. 1997); one velocity feature increased by more than 100% in $\sim$10 min. If the variations are intrinsic, a timescale $\tau_{var} \sim$ minutes implies a size scale $d \sim 1\,AU$. However, even in favourable conditions, the maser gain length $l > 1\,AU$. Hence this variability must be a response to an external process. Of the second and third mechanisms listed above, amplification of a varying background source requires an unsaturated maser, which is not believed to be the case for most of the megamaser sources, and a time-varying radiative pump is unlikely to be the cause since the 22 GHz maser line is believed to be collisionally pumped. This leaves interstellar scintillation (ISS) — scattering in an ionised component of the Galactic interstellar medium — as the default explanation (Greenhill et al. 1997). Lending credence to this possibility is the low Galactic latitude of Circinus, at $b = -3.8°$.

There is not space here to provide a detailed review of ISS; see Rickett (1990, 2002) and Narayan (1993). The variability timescale of weak ISS is approximately the crossing time of the Fresnel scale $l_F \approx \sqrt{D\lambda/2\pi}$, where D is the distance to the scattering screen. For $D \sim 0.1$–1 kpc and a relative velocity (with respect to the observer) $v \approx 50\,km\,s^{-1}$, as expected for a Galactic scattering screen (e.g. Rickett 1990), the weak ISS timescale is too long by about an order of magnitude. Hence the variability must be due to strong diffractive ISS. For strong ISS, the contributing length scales l_{scin} are smaller than l_F; strong diffractive ISS with $\tau_{scin} > 10s$ is seen at 1 GHz towards pulsars near the plane. Since the frequency variation of the scintillation timescale is expected to be $\tau_{scin} \propto \nu^{1.0}$–$\nu^{1.2}$, only the fastest ISS timescales seen at 1 GHz can explain the 22 GHz variations. In addition, the angular size of the masing region must satisfy $\theta_{maser} < \theta_{scin} = l_{scin}/D$ in order not to reduce the amplitude and increase the timescale of the ISS fluctuations; hence we must have

$$d_{maser} < 7\,AU \qquad (1)$$

taking l_{scin} to be the Fresnel scale. This is a not implausible size scale for a maser feature. Interestingly, the

extragalactic water maser source in IC 10, which is also seen to vary with $\tau_{var} < 1$ day, also lies at low Galactic latitude. This lends additional support to the Greenhill et al. (1997) suggestion of interstellar scintillation as the source of the variability in Circinus. Since the maser emission from Circinus has been resolved into multiple spatial/velocity components with VLBI, multiple-epoch imaging observations to study the spatial structure of the variability would be extremely informative, and would provide additional constraints on the nature of the scattering screen.

2.2 Varying Background Source: Mrk 348?

Mrk 348 is a Sy 2 in an S0 galaxy; it possesses a nuclear dust lane and shows broad emission lines in polarised light. There is a luminous ($L_x \sim 10^{43}$ erg s^{-1}), heavily obscured ($N_H \sim 10^{23}$ cm^{-2}) nuclear X-ray source, and a bright, variable radio nucleus. Between the end of 1997 and early 2000 the H$_2$O maser in Mrk 348 flared by about a factor of three in luminosity (Falcke et al. 2000); the line is very broad ($\Delta V \sim 150$ km s^{-1}) compared to most other megamaser sources. Mrk 348 was not detected as a maser source in earlier observations.

The radio continuum emission also flared in this time period, increasing by nearly a factor of five between the beginning of 1997 and late 1998, and increasing by an additional $\sim 40\%$ to early 2000 (Falcke et al. 2000). If the maser flaring results directly from the continuum flare, then the maser emission (or at least a large part of it) must be unsaturated, which is not generally believed to be the case for most extragalactic water masers. In this context, it is of interest to note that the maser emission in the megamaser NGC 1052 (an elliptical galaxy) lies along the radio jet axis, while in the other broad-line megamaser, IRAS 2226-164, the spatial distribution is 'messy' (Moran, Greenhill, & Herrnstein 1999). Hence these sources (in which the emission may be due to shocks) may be quite dissimilar to the majority of the megamasers that have been mapped to date, in which the emission appears to arise in a disk (Moran et al. 1999), and the emission is plausibly powered by the X-ray emission from the central source (Neufeld, Maloney, & Conger 1994; Neufeld & Maloney 1995). The upper limit to the response time lag between the radio flare and the maser flare in Mrk 348 (about two years) sets an upper limit to the distance of the masers from the nucleus of about 0.6 pc. A recent VLBI observation with MERLIN (Xanthopoulos & Richards 2001) of the red half of the line observed by Falcke et al. (2000) confirms that maser emission arises within ~ 0.8 pc of one of the continuum peaks; the peak flux density is about three times higher than when observed by Falcke et al. (2000). Hence there is substantial support for connecting the maser flare causally with the continuum flare. Regular monitoring of the maser and continuum emission from this source should be carried out in order to confirm or refute this model, and would in either case place useful constraints on the maser pumping mechanism (e.g. the level of saturation).

2.3 Time-variable Pump: NGC 1068?

The maser emission in NGC 1068 is rather complex, with 'nuclear' masers associated with the apparent thermal radio source S1 (believed to mark the nucleus) with redshifted, blueshifted, and systemic velocity components, and 'jet' masers associated with radio source C, which appears to arise in a jet/cloud shock (Gallimore et al. 1996). The nuclear masers appear to lie in a disk, a conclusion supported by upper limits to the accelerations of the maser features ($\dot{v} < 1$ km s^{-1} yr^{-1}).

As noted by Claussen & Lo (1986), the entire maser spectrum faded between 1983 and 1984. If the masers are powered by X-rays from the AGN, as argued by Neufeld et al. (1994), we expect to see the masers respond to variations of the central engine. Gallimore et al. (2001, hereafter G01) argue that the nuclear masers vary coherently with $\tau_{var} \gtrsim 2$ yr. (Note that the dynamical timescale $\tau_{dyn} \sim 10^4$ yr.) For the inferred disk geometry, the red and blue features should lag the systemic by $\tau \sim 2$ yr.

G01 compared both the luminosities L_{maser} and the peak flux densities S_ν(peak) of the maser features to one another, searching for coherent temporal variability of the redshifted, blueshifted, and nuclear maser features. (The data are unfortunately not adequate for a cross-correlation analysis.) Correlations of the former would be expected if the timescale of variation of the central power source τ_{pump} is long compared to the timescale for propagation of this signal through the masing region, τ_{prop}, and correlations of the latter if this inequality is reversed. The maser luminosities vary by factors ~ 3 over the 15 yr period for which there are data. G01 find significant ($> 3\sigma$) correlations for the blue and red nuclear masers (in both L_{maser} and S_ν(peak)), although the correlation between the red and systemic features is not much worse. It must be noted, however, that the sampling time was usually coarser than the predicted lag between blue/red and systemic. The jet masers clearly vary independently of the nuclear masers, as expected.

2.3.1 The 1997 Maser Flare

Over 3–4 November 1997, G01 observed the blueshifted $v = 859$ km s^{-1} feature to flare by a factor ~ 10. There was a simultaneous flare of the redshifted 1411 km s^{-1} feature, although it is more difficult to see because it is partly blended with other maser features. The rotating disk model predicts that maser flares in response to variation of the central source should be simultaneous on the blueshifted and redshifted sides (for masing regions symmetric about the disk centre). In fact, since the maximum delay between the 859 and 1411 km s^{-1} features is 22 days (set by the sampling frequency), this implies that any radius difference between the two is $\Delta R/R \lesssim 4\%$.

If the maser emission is powered by the X-ray radiation from the AGN, it can respond to variations in the central source in two ways (Neufeld 2000):

- An increase in the X-ray luminosity L_x will lead to an increase in the maser luminosity L_{maser}, as the maser

emissivity increases with the temperature difference $T_{\mathrm{diff}} = T_{\mathrm{gas}} - T_{\mathrm{dust}}$ (since the X-rays deposit about as much energy into the gas as into the dust, and the gas cools less effectively).

- An increase in the bolometric luminosity L_{bol} will cause L_{maser} to decrease, as T_{diff} declines.

If the former is the case, L_{maser} cannot change faster than the thermal timescale τ_{therm}. If it is the latter, the maser luminosity can change on the radiative diffusion timescale

$$\tau_{\mathrm{rad}} \sim \frac{\tau_{\mathrm{fir}}l}{c} \sim 0.01(N_H/10^{24})(10^9/n_{\mathrm{H_2}})\,\mathrm{yr}, \quad (2)$$

where $\tau_{\mathrm{fir}} = $ FIR optical depth, and $l = $ depth below irradiated surface, since the dust grains have negligible thermal inertia. In application to NGC 1068, Neufeld (2000) shows that either mechanism could explain the flare, provided that the gas density $n_{\mathrm{H_2}} \gtrsim 10^8\,\mathrm{cm}^{-3}$. Note also the opposite sign of correlation with L for the two mechanisms. Unfortunately, in the case of NGC 1068, we do not see the nuclear source directly, so there are considerable uncertainties introduced by the inevitable delays that must occur before we see changes in the intrinsic luminosity. Hence, while this model is certainly plausible, the existing data do not allow us to state unequivocally that the masers are responding to the central source. G01 present evidence that there has been a long-term ($\sim$10 yr) increase in the bolometric luminosity of NGC 1068, as seen in the near infrared emission from warm dust. If this is true, it suggests that the 1997 flare was the result of an increase in the X-ray luminosity. As with the other cases discussed here (interstellar scintillation in Circinus and continuum flaring in Mrk 348), systematic monitoring would greatly improve the situation and allow us to make much more definitive statements regarding the mechanism responsible for the variability.

Acknowledgments

References

Claussen, M. J., & Lo, K.-Y. 1986, ApJ, 308, 592

Falcke, H., Henkel, C., Peck, A. B., Hagiwara, Y., Almudena Prieto, M., & Gallimore, J. F. 2000, A&A, 358, L17

Gallimore, J. F., Baum, S. A., O'Dea, C. P., Brinks, E., & Pedlar, A. 1996, ApJ, 462, 740

Gallimore, J. F., Henkel, C., Baum, S. A., Glass, I. S., Claussen, M. J., Prieto, M. A., & von Kap-Herr, S. 2001, ApJ, 556, 694 (G01)

Gardner, F. F., & Whiteoak, J. B. 1982, MNRAS, 201, 13p

Greenhill, L. J., Ellingsen, S. P., Norris, R. P., Gough, R. G., Sinclair, M. W., Moran, J. M., & Mushotzky, R. 1997, ApJ, 474, L103

Maloney, P. R. 2002, PASA, 19, in press

Moran, J. M., Greenhill, L. J., & Herrnstein, J. R. 1999, JApA, 20, 165

Narayan, R. 1993, in Pulsars as Physics Laboratories, eds R. D. Blandford, A. Hewish, A. G. Lyne, & L. Mestel (London: Oxford), 151

Neufeld, D. A. 2000, ApJ, 542, L99

Neufeld, D. A., & Maloney, P. R. 1995, ApJ, 447, L17

Neufeld, D. A., Maloney, P. R., & Conger, S. 1994, ApJ, 436, L127

Rickett, B. J. 1990, ARA&A, 28, 561

Rickett, B. J. 2002, PASA, 19, in press

Xanthopoulos, E., & Richards, A. M. S. 2001, MNRAS, 326, L37