## VINOD KRISHAN

Indian Institute of Astrophysics, Koramangala, Bangalore 560 034, India

AND

## PAUL J. WIITA<sup>1</sup>

Department of Physics and Astronomy, Georgia State University, Atlanta, GA 30303-3083 Received 1993 February 2; accepted 1993 September 13

## **ABSTRACT**

We consider several physical processes, dominated by plasma effects, that might play roles in producing variability in AGN. One class of processes involves oscillations in the plasma itself. If accretion disks around supermassive black holes have coronae not too different from those of the Sun, then magnetoacoustic oscillations of coronal loops can yield fluctuations on timescales from days to years. Such coronal loops could also produce flares with substantial emission variations over the short times comparable to those associated with the recently observed broad-band microvariability. A second class of processes depends upon electromagnetic fields acting to modulate themselves; these modulational instabilities can produce filamentation and thence rapid variability. We also consider how a shift in emission and scattering mechanism (e.g., from Compton scattering to Raman scattering) can yield variability in the observed emission.

Subject headings: BL Lacertae objects: general — galaxies: active — galaxies: nuclei — MHD radiation mechanisms: nonthermal

#### 1. INTRODUCTION

Variability has long been considered an important property of active galactic nuclei (AGNs); recent conferences on this subject include Miller & Wiita (1991) and Valtaoja & Valtonen (1992). Observations over the past few years have demonstrated the reality of microvariability, i.e., rapid variability of a few percent over less than 1 day, in the radio (e.g., Quirrenbach et al. 1989; Wagner et al. 1990), optical (e.g., Miller, Carini, & Goodrich 1989; Carini et al. 1991; Wagner et al. 1990), ultraviolet (Edelson et al. 1991), and X-ray bands (e.g., Turner & Pounds 1988; Kunieda et al. 1990). Thus, the range of timescales for AGN variability is quite broad, ranging from minutes to decades.

Theoretical explanations for variability in AGNs fall into two main classes: extrinsic and intrinsic. While extrinsic processes such as refractive interstellar scintillation (e.g., Heeschen et al. 1987) or gravitational microlensing (e.g., Gopal-Krishna & Subramanian 1991) can certainly contribute to the observed changes in some frequency bands or for some sources, they cannot be the explanation for the majority of the observed variability. Within the intrinsic category of variability explanations there have also been two main divisions: models that are based upon relativistic jets and those placing their emphasis on accretion disks.

tivistic speeds have long been dominant in attempts to understand BL Lacertae objects (e.g., Blandford & Königl 1979). Detailed models for flux and polarization variations in the radio on timescales of years have been successfully developed within this framework (Marscher & Gear 1985; Hughes, Aller, with turbulent regions within the jet whose varying densities

Models involving shocks propagating down jets at rela-

& Aller 1985, 1991). However, there have only been very preliminary attempts to explain microvariability using this paradigm. One way to do so is to assume that the shock interacts and velocities produce noticeable flux variations (Marscher, Gear, & Travis 1992). Another crude jet-based picture relies upon modest swings in the direction at which the shock is observed (Gopal-Krishna & Wiita 1992). A related model involves nonaxisymmetric bubbles carried outward in relativistic magnetized jets (Camenzind & Krockenberger 1992).

Variability arising from accretion disks is known to occur for binary stellar systems, and it is certainly logical to expect that disk fluctuations could account for at least some of the observed variability in AGNs. Relatively long-term variations could arise from disk pulsations (e.g., Vila 1979). The possibility that a large number of hot spots or flares on accretion disks could provide the bulk of the observed rapid changes in the X-ray (Abramowicz et al. 1991; Zhang & Bao 1991), and optical and UV (Wiita et al. 1991, 1992; Mangalam & Wiita 1993) has been recently proposed. Purely phenomenological models of this type are able to match the amplitude and power spectrum of the observations, and unstable spiral shocks seem to provide one physical basis for this picture (Wiita et al. 1992; Chakrabarti & Wiita 1993). Still, the underlying physics of this class of models is even less developed than that based upon shocks in jets. So far, all of this microvariability has been clearly demonstrated only in radio-loud AGNs (mainly BL Lacertae objects), and thus a search for microvariability in radio-quiet QSOs is most desirable (Gopal-Krishna, Wiita, & Altieri 1993b; Gopal-Krishna, Sagar, & Wiita 1993a); if radioquiet AGNs do exhibit rapid fluctuations, disk-based models would be preferred, but if they do not, then the jet-based models would be favored.

We note that if there is a corona around accretion disks (e.g., Galeev, Rosner, & Vaiana 1979; Liang & Thompson 1979) then processes similar to those occurring in the solar corona may be very important. Note that variations from milliseconds to minutes and from microwaves through X-rays are seen in solar flares (Kiplinger et al. 1983). We also note that while the standard models for AGNs assume that the bulk of the spectrum is a combination of synchrotron and synchrotron self-

<sup>&</sup>lt;sup>1</sup> Also, Indian Institute of Astrophysics, Bangalore.

Compton emission (mainly from the jet) and quasi-thermal emission (mainly from the disk and reradiating dust), several models relying on coherent emission processes within plasmas for substantial portions of the AGN continuum (Krishan 1983; Baker et al. 1988; Krishan & Wiita 1990, hereafter KW) and variability (Benford 1992; Gangadhara, Krishan, & Shukla 1993; Lesch & Pohl 1992) have been proposed over the past few years.

In this paper we briefly discuss some additional variability mechanisms, based on plasma processes that might be relevant within AGNs, and we estimate reasonable timescales for them. In no way do we claim that our treatments are complete; to a large extent this paper should be taken as providing a catalog of processes that might play a role in AGN variability and that ought to be examined more carefully in future work. There are two key concepts upon which we will expand: (1) oscillations in plasma density can affect the observed radiation; (2) electromagnetic fields can act to modulate themselves. One can logically divide (1) into (a) slow changes in the medium producing magnetohydrodynamic (MHD) waves, which are discussed in § 2; and (b) impulsive changes in the medium yielding flares and substantial additional radiation (§ 3). Modulation induced by the radiation itself (2) can yield filamentation and other effects outlined in § 4 and can also alter the basic type of emission process.

## 2. MHD FLUCTUATIONS

We first consider magnetoacoustic oscillations in looplike structures and base our treatment on the work of Roberts, Edwin, & Benz (1984, hereafter REB). Assuming that the accretion disk's corona, like the Sun, is permeated with structures that have transverse dimensions small compared with their lengths, significant density and temperature inhomogeneities are expected. Even if the magnetic field strength is close to uniform, the density variations will provide large differences in the Alfvén speed, which determines the nature of the oscillations (REB). Both standing modes and propagating modes are possible, which can lead to both periodic and nonperiodic fluctuations in emissivity; of course, periodic signals would only be detected if the energy released was dominated by a single loop at a given time, which is quite unlikely for an AGN.

For simplicity, the coronal loop is treated as a uniform straight cylinder of radius a and length L, and gravitational and curvature effects are neglected. The field strength is  $B_0$  inside the cylinder and is taken as uniform and equal to  $B_x$  outside; the respective densities, pressures, and temperatures are:  $\rho_0$ ,  $\rho_x$ ;  $p_0$ ,  $p_x$ ;  $T_0$ ,  $T_x$ . These lead to the following characteristic speeds (REB): sound speeds,  $c_0 = (\gamma p_0/\rho_0)^{1/2}$ ,  $c_x = (\gamma p_x/\rho_x)^{1/2}$ ; Alfvén speeds,  $v_A = (B_0^2/4\pi\rho_0)^{1/2}$ ,  $v_{Ax} = (B_x^2/4\pi\rho_x)^{1/2}$ ; and tube speeds  $c_T = c_0 v_A/(c_0^2 + v_A^2)^{1/2}$ ,  $c_{Tx} = c_x v_{Ax}/(c_x^2 + v_{Ax}^2)^{1/2}$ .

When adiabatic oscillations are considered using the usual MHD equations, a fairly complex dispersion relation is obtained (Spruit 1982; Edwin & Roberts 1983) for waves with wavenumber k along the magnetic field direction (z). Using cylindrical coordinates  $(r, \phi, z)$ , the velocity perturbations have the form

$$\mathbf{v} = \mathbf{v}(r) \exp \left[i(\omega t + m\phi - kz)\right]$$
.

As long as the conditions are reasonable for a corona, i.e.,  $v_{\rm Ax}>v_{\rm A}>c_0>c_T>c_x>c_{\rm Tx}$ , this dispersion relation has two classes of modes whose radial motions are essentially con-

fined to the loop and which correspond to the usual fast and slow magnetoacoustic modes (REB).

Both the sausage (m=0) and the kink (m=1) slow modes have very weak dependences of their phase speed on the wavenumber, and for the low  $\beta$  plasma we assume here,  $c_T \approx c_0$ . Thus, as long as ka is not too large, the phase speeds of both the sausage and kink modes are nearly given by  $\omega/k = c_T$ . Hence, for a standing wave in such a coronal loop, whose footpoints are assumed to be anchored in a higher density disk atmosphere, the disturbances can be assumed to be zero at the ends of the loop (z=0,L), and one can take  $k=j\pi/L$  for positive integers j. Then the periods of these slow modes, defined as  $(\tau_s=2\pi/\omega)$  are, for solar abundance and  $\gamma=5/3$ ,

$$\tau_s = \frac{2L}{jc_T} = \frac{1.2 \times 10^{-4}}{j} L T_0^{-1/2} \left( 1 + \frac{c_0^2}{v_A^2} \right)^{1/2}$$
 (1)

(REB, eq. 3a).

The fast modes have strong dependences of  $\omega/k$  on k. For this case, both the sausage and kink modes have phase velocities between  $v_A$  and  $v_{Ax}$  and only exist as free (real k and  $\omega$ ) modes if  $v_{Ax} > v_A$ ; i.e., the regions with lower Alfvén speed (basically those with higher density) trap these fast magneto-acoustic modes. The fast sausage modes are also cut off for too low wavenumbers (roughly k < 1/a) as are the higher order kink modes, but the principle kink oscillation exists for all k, and its phase speed has the following long wavelength limit:  $c_k = [(\rho_0 v_A^2 + \rho_x v_{Ax}^2)/(\rho_0 + \rho_x)]^{1/2}$ . The periods of the fast kink and fast sausage modes, respectively, turn out to be

$$\tau_f = \frac{2L}{jc_k} = \frac{4\pi^{1/2}L}{j} \left(\frac{\rho_0 + \rho_x}{B_0^2 + B_x^2}\right)^{1/2},$$
 (2)

$$\tau_f' = \frac{2\pi a}{c_k} = 4\pi^{3/2} a \left( \frac{\rho_0 + \rho_x}{B_0^2 + B_x^2} \right)^{1/2} \tag{3}$$

(REB, eq. 3b, c). In substituting typical values of the parameters expected in the coronal loops ( $T_0 = 2 \times 10^6$  K,  $L = 10^{10}$  cm) REB find equation (1) yields  $\tau_s = 850$  s for j = 1. The period of the fast kink mode can be estimated if one assumes the loop is much denser than its surroundings but has a similar field strength. REB take  $B_0 = 40$  G and  $n_0 = 10^9$  cm<sup>-3</sup> as reasonable for the solar corona, yielding  $\tau_f = 50$  s for j = 1. While no periodicities in solar loops corresponding to the slow mode seem to have been observed, REB argue that the many observed cases of quasi-periodic pulsations in solar flares with periods of the order of minutes could be attributed to the fast kink mode.

When we turn to parameters possibly relevant for AGNs, the uncertainties become much greater. But we recall that typical models invoke supermassive black holes with masses,  $M_{\rm BH}$ , between  $10^7-10^{10}~M_{\odot}$ , yielding Schwarzschild radii  $(R_{\rm S})$  of  $3\times10^{12}-3\times10^{15}$  cm, respectively. The accretion disks usually posited within these models (e.g., Rees 1984) have radii extending to hundreds of  $R_{\rm S}$ ; standard thin disks (e.g., Shakura & Sunyaev 1973) have thicknesses of a few percent of their radii. The disk half-thicknesses should provide approximate physical scales for the lengths of coronal loops that might extend above and below these accretion disks. The surface temperatures of the accretion disks are mainly expected to lie in the range  $5\times10^3-5\times10^4~{\rm K}$  (scaling roughly as  $M_{\rm BH}^{-1/4}r^{-3/4}$ ), which are not too far from stellar surface temperatures. Hence, in the absence of any firm theoretical or observational evidence we assume that coronal temperatures

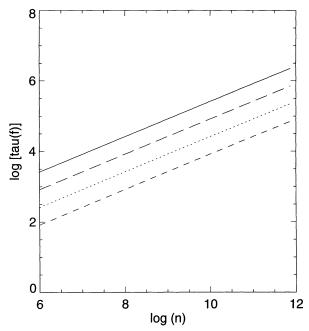


Fig. 1.—The fast mode timescales (in s) plotted against ambient density (in cm<sup>-3</sup>); in all cases shown the magnetic field strength is taken as 100 G. The solid line corresponds to the kink mode for j=1 and a tube length of  $L=10^{14}$  cm, while the dotted line has  $L=10^{13}$  cm but is otherwise the same. The long dashed line is the fast sausage mode for a tube diameter of  $a=10^{13}$  cm, while the short dashed line corresponds to  $a=10^{12}$  cm.

are  $> 10^6$  K and disk thicknesses are  $\sim 10^{12}$  cm. Number densities and magnetic field strengths within these coronal loops are also quite uncertain, but reasonable crude estimates (see KW) would give  $n_0 \sim 10^9$  cm<sup>-3</sup> and  $B \sim 100$  G.

Thus, to estimate the periods of the magnetoacoustic waves,

Thus, to estimate the periods of the magnetoacoustic waves, we assume  $n_0 \gg n_x$  and  $B_0 > B_x$ . Then  $v_A = 6.9 \times 10^8 B_{0,2} \, n_{0,9}^{-1/2}$  cm s<sup>-1</sup>, where  $B_0 = 10^2 B_{0,2}$  G and  $n_0 = 10^9 n_{0,9}$  cm<sup>-3</sup>. Using equation (1) and defining  $L = 10^{15} \, L_{15}$  cm and  $T = 10^6 T_{0,6}$  K (in general we will use the notation  $X_n = X/10^n$  and express quantities in cgs units), we find

$$\tau_s = \frac{2L}{jc_T} = \frac{1.7 \times 10^8 L_{15}}{jT_{0,6}^{1/2}} \left( 1 + \frac{2.9 \times 10^{-4} T_{0,6} n_{0,9}}{B_{0,2}^2} \right)^{1/2} \text{ s} . \quad (4)$$

As it is likely that  $L_{15} < 1$  while  $T_{0,6} > 1$ , these slow oscillations have timescales of roughly a few years. Allowing for the plausible range in disk coronal parameters, this nonthermal emission could, in principle, range from the millimeter to UV bands and would show very weak polarization. Many observed variations in the optical (e.g., Smith, Nair, & Clements 1991) and radio (e.g., Aller, Hughes, & Aller 1991) occur on these multiyear timescales. But it must be recalled that shock-in-relativistic-jet models for the radio variations seem to provide excellent matches to the data (Marscher & Gear 1985; Hughes et al. 1985, 1991).

If the loop magnetic field is substantially stronger than the field outside the loop, then equations (2) and (3) for the fast kink and fast sausage modes reduce to

$$\tau_f \approx \frac{L}{\pi^{1/2} i v_A} = 8.2 \times 10^5 \, \frac{L_{15} \, n_{0.9}^{1/2}}{i B_{0.2}} \, \text{s} \,,$$
 (5)

$$\tau'_f \approx \frac{\pi^{1/2}a}{v_A} = 2.6 \times 10^4 \frac{a_{13} n_{0,9}^{1/2}}{B_{0,2}} \text{ s} .$$
 (6)

These timescales are illustrated in Figure 1; as these values are typically from hours to months, they certainly do overlap with the observed microvariability timescales.

#### 3. FLARES

Our treatment in this section is based closely upon the work of Spicer (1981, hereafter S81), who considers a generalization of the Alfvén-Carlqvist (1967) current interruption model for solar flares as well as an improvement of Spicer's (1977) tearing mode model for instabilities occurring within magnetic loops. In the picture of S81 a current channel of thickness  $\delta r$  exists within the loop as a skin current or sheath, which is expected because of the small resistive skin depth for coronal conditions. One has  $\delta r = c(v_{ei} \Delta t)^{1/2} \omega_{pe}$ , where  $v_{ei}$  is the electron-ion collision frequency,  $\Delta t$  is the timescale over which time dependent induction fields are driving the current, and the electron plasma frequency  $\omega_{pe} = (4\pi e^2 n_e/m_e)^{1/2} = 5.64 \times 10^4 n_e^{1/2}$  rad s<sup>-1</sup>.

## 3.1. Anomalous Resistivity

The physical assumption behind the Alfvén-Carlqvist flare model is that a current flows parallel to the magnetic field within a loop. In the original picture, a density rarefaction in such a loop leads to a relative increase in the drift velocity,  $v_D$ , between the electrons and ions; if great enough, this increased  $v_D$  should drive a microinstability, producing an anomalous resistivity much higher than the classical value (Krishan 1978). This anomalous resistivity implies a great increase in Joule heating, thence a big voltage drop across the unstable region, which is then expected to accelerate particles.

For plasmas with  $T_e \approx T_i$ , the relevant instability is that of Buneman (1959), which requires

$$v_D \ge c_{T_e} \equiv (k_b T_e/m_e)^{1/2} ,$$
 (7)

so that the resulting electron heating quenches the instability. If this is the situation, the growth rate of the instability is given by

$$\Gamma = \frac{3^{1/2}}{2^{4/3}} \left( \frac{m_e}{m_i} \right)^{1/3} \omega_{p_e} \,. \tag{8}$$

(S81, Table 1).

But if  $\xi \equiv T_e/T_i \gg 1$ , then the ion-acoustic instability will be excited; this instability is quenched due to the heating of both electrons and ions. To instigate the ion-acoustic instability one requires

$$v_D \ge c_s \alpha = (k_b T_e/m_i)\alpha \tag{9}$$

(S81, eq. 3), where  $c_s$  is the ion sound speed and

$$\alpha = 1 + (m_i/m_e)^{1/2} \xi^{3/2} \exp(-1.5 - 0.5\xi)$$
.

In this case, the growth rate is

$$\Gamma = \left(\frac{m_e}{m_i}\right)^{1/2} \omega_{pi} \tag{10}$$

(S81, Table 1). We continue to follow Spicer (1981) in making the fundamental assumption, which he justifies, that the current is being driven; i.e., it is maintained by an external mechanism at a steady state value limited by nonlinear effects so that the microinstability is excited always at marginal stability. This is reasonable despite the very high values of the growth rate given by equation (10), because nonlinear analyses demonstrate the saturation of these instabilities. If one assumes

that an electric field, E, is applied to the plasma so that a current forms with a drift velocity roughly equal to the ion sound speed,  $c_s$ , and  $T_e/T_i \gg 1$ , then the ion-acoustic instability will be excited. Obtaining the power dissipated from the current driving the marginal instability  $(J_{MS})$  and equating it with the power emitted in a flare,  $P_f$ , leads to the characteristic time for dissipation of the current, which is a substantial fraction of the total flare lifetime (S81, eq. 16).

We now evaluate the temperature, characteristic length, and dissipation timescale using numbers that might be expected in AGNs. Doing so, we find that after correcting typographical errors in equations (28) and (29) of S81 and scaling to the larger sizes discussed in § 2.

$$T_{e,6} = 8.7 \times 10^5 \left( \frac{P_{f,41}}{a_{14} \delta B_{p,2} L_{15} \sqrt{n_{e,9} \alpha}} \right)^2;$$
 (11)

$$\delta r \approx \frac{4 \times 10^4 \delta B_{p,2}}{n_{e,9} T_{e,6}^{1/2} \alpha} \text{ cm} ;$$
 (12)

$$\tau_d = \frac{4.6 \times 10^8 a_{14} \, \delta B_{p,2}}{\left(n_{e,9} \, T_{e,6}\right)^{1/2} \alpha} \, \mathrm{s} \,\,, \tag{13}$$

where  $\delta B_p$  is the change in magnetic field induced by  $J_{MS}$  across the characteristic width of the sheath,  $\delta r$ , required to excite the instability.

A moderate fluctuation, arising from a single flare in a corona around an accretion disk in a Seyfert galaxy of total luminosity of  $10^{44}$  ergs s<sup>-1</sup> might have  $P_{f,41} = 1$ . While all these quantities are uncertain, it would be fair to take  $a_{14} \approx$  $\delta B_{p,2} \approx n_{e,9} \approx L_{15} \approx 1$  as well. For  $\xi = 3$ ,  $\alpha \approx 12$ , and then the flare could yield high temperatures,  $T_{e,6} \approx 5.9 \times 10^3$ . The decay timescale would then be  $\tau_d \approx 5.0 \times 10^5$  s, or around a week. For a fluctuation in a more powerful quasar,  $P_{f,41} = 10^3$ might be expected. But if the central black hole is also bigger, as would be likely in this case, then higher values of a, L, and possibly  $\delta B_p$  (by roughly a factor of 10 each) would also be expected, so that the value of  $T_{e,6}$  would probably be comparable to that mentioned above. While the value of  $n_{e,9}$  might also be somewhat higher, the net effect would be to increase the timescale by roughly one order of magnitude (to a few months) for this substantial flare. Most of the emission generated by these flares would usually emerge in the X-ray band.

# 3.2. Tearing Modes

The main difference between the modified Alfvén-Carlqvist model discussed above and the tearing mode model of Spicer (1977) is that the former invokes microinstabilities to produce the anomalous resistivity while the latter requires a macroinstability—the resistive kink or tearing mode that can occur in any sheared magnetic configuration. For this case, assume that a current density,  $J_z(r)$ , flows along an axial magnetic field,  $B_z(r)$ , producing self-consistently a component  $B_\phi(r)[\ < B_z(r)]$ . This combination yields a sheared field, and can generate multiple tearing modes which can interact with each other, substantially increasing the rate at which energy can be released in a flare through reconnection (S81). The following conditions must be satisfied for tearing modes to exist:

$$\mathbf{k} \cdot \mathbf{B}_0 = 0 , \quad |\mathbf{k}| \, \delta a < 1 , \quad \Delta' > 0$$
 (14)

(Furth, Killeen, & Rosenbluth 1963), where  $B_0$  is the equilibrium magnetic field,  $\delta a$  is the field gradient scale length, and  $\Delta'$  measures the free magnetic energy available to drive the

tearing mode and depends upon the steepness of the current density gradient.

For a double tearing mode (equivalent to the reconnection of the field between two neutral sheets) the growth rate,  $\Gamma \propto S^{2/3}/\tau_R$ , where  $S = \tau_R/\tau_A$ ,  $\tau_R = 4\pi(\delta a)^2/\eta c^2$ , and  $\tau_A = \delta a/v_A$ . The normal resistivity,  $\eta = (m/ne^2)v_{ei}$ , yields extraordinarily long timescales for any plausible values of  $B_0$ ,  $n_e$ , and  $\delta a$  applicable to AGNs, though they can be quick enough for the circumstances expected for solar coronal loops. Defining  $\tau_T = 1/\Gamma$ , we note

$$\tau_T \propto \frac{\tau_R}{S^{2/3}} \propto \tau_R^{1/3} \tau_A^{2/3} = \left[ \frac{(4\pi)^{1/3} (\delta a)^{2/3}}{(nc^2)^{1/3}} \right] \left( \frac{\delta a}{v_A} \right)^{2/3} \propto (\delta a)^{4/3} \ . \tag{15}$$

If the ion-acoustic effective collision rate is used (S81, Table 1) along with equation (15), we eventually find

$$\tau_T = 1.2 \times 10^{10} (\delta a_{12})^{4/3} n_{e,9}^{-1/6} B_2^{2/3} \text{ s} . \tag{16}$$

Unfortunately, all estimates of  $\delta a$  are very uncertain, the only clear point being that  $\delta r \ll \delta a \ll L$ , motivating the numerical scaling chosen in the previous equation. Despite this uncertainty, it is likely that tearing modes produce timescales longer than those arising directly from Alfvénic waves and are thus unlikely to be important for *rapid* variability.

## 4. MODULATIONS OF ELECTROMAGNETIC WAVES

An entirely separate family of instabilities arises from the interaction of electromagnetic waves with the plasma through which they propagate. These waves modify the plasma properties and, in turn, the plasma acts back upon the "pump" wave to produce a nonlinear active medium and additional EM waves which induce temporal variability in the emerging radiation. Well known plasma instabilities such as the Langmuir modulational instability (e.g., Zakharov 1972), the oscillating two-stream instability (e.g., Nishikawa 1978), and the self-modulation of EM waves in a plasma (e.g., Kaw, Schmidt, & Wilcox 1973) are all examples of modulational instabilities. These modulational instabilities can all be described in terms of four-wave interactions, as was nicely shown by Bingham (1990, hereafter B90).

Physically, these four waves consist of two pump "quanta" and two sidebands (e.g., Bingham & Lashmore-Davies 1979). The modulational instability can lead to filamentation of the wave if the modulation wavenumber is perpendicular to the direction of propagation, but it can lead to the breakup of the beam, yielding a train of short pulses if the modulational wavenumber is parallel to the direction of propagation (B90). For simplicity, in this section we follow B90 in examining only unmagnetized plasmas, although similar processes occur in magnetized plasmas (Shapiro & Shevchenko 1984).

# 4.1. A Simple Model

Consider a uniform, infinite plasma through which a high-frequency pump wave is propagating. This initial pump wave's electric field is (B90, eq. 1)

$$E_0(x, t) = A_0(x, t) \exp [i(k_0 \cdot x - \omega_0 t)],$$
 (17)

where the linear dispersion relation

$$\omega_0^2 = \omega_{pe}^2 + v^2 k_0^2 \tag{18}$$

176

KRISHAN & WIITA Vol. 423

holds, with v=c for a transverse wave, while  $v=\sqrt{3}v_{\,{\rm th},e}$  for a longitudinal (Langmuir) wave. The other key component of this four-wave interaction model is a low-frequency density perturbation with frequency and wavenumber  $(\Omega,k_s)$ . This perturbation, which affects both ions and electrons, beats with the initial pump wave to yield high-frequency sidebands with wavenumbers  $k_0-k_s$  (Stokes) and  $k_0+k_s$  (anti-Stokes). Those sidebands can couple to produce another perturbation at  $2k_s$ , which can in turn lead to additional Stokes and anti-Stokes waves. But there will always be one fastest growing mode  $(k_{ms})$  which will dominate this process, and we will focus on it.

The slow amplitude variation of the EM wave is determined by linear dissipation and nonlinear interactions. A perturbation technique, described in detail in Bingham & Lashmore-Davies (1979), combines equations (17) and (18) with similar ones for these waves and with Maxwell's equations (including relativistic motion). A key assumption is that the wave-vectors match perfectly, i.e.  $k_0 = k_{1,2} \pm k_s$ , but a slight mismatch between the frequencies is allowed, where they are given by  $\omega_j^2 = \omega_{pe}^2 + v^2 k_j^2$  (j = 0, 1, 2). If the amplitudes of the sidebands (with  $k_{1,2}$ ) are  $A_{1,2}$ , and the amplitude of the density perturbation is  $N(\mathbf{x}, t)$ , then solutions can be found analytically.

Using the static approximation (e.g., Zakharov 1972),  $\gamma(\partial/\partial t)$ ,  $\partial/\partial t^2 \ll k_s^2 c_s^2$  (where  $\gamma$  is the Lorentz factor), one finds

$$\left(\frac{\partial}{\partial t} + \boldsymbol{v}_0 \cdot \frac{\partial}{\partial \boldsymbol{x}} + \gamma_0\right) A_0(\boldsymbol{x}, t) = iG_0 \left[\omega_1^3 A_0 |A_1|^2 + \omega_2^3 A_0 |A_2|^2 + \omega_1 \omega_2 (\omega_1 + \omega_2) \times A_0^* A_1 A_2 e^{i(\delta_1 + \delta_2)t}\right]$$
(19)

(B90, eq. 17), where  $c_{s0}$  is a coupling constant that depends upon the specific interaction process,  $\delta_{1,2} = \omega_0 - \omega_{1,2}$  are the frequency mismatch terms, and  $v_0$  is the group velocity. The equivalent equation for  $A_1$  is (that for  $A_2$  arises from the interchange of the subscripts 1 and 2):

$$\left(\frac{\partial}{\partial t} + \boldsymbol{v}_1 \cdot \frac{\partial}{\partial \boldsymbol{x}} + \gamma_1\right) A_1(\boldsymbol{x}, t) = iG_1$$

$$\times \left[ \omega_0^3 |A_0|^2 A_1 + \omega_0^3 \frac{\omega_2}{\omega_1} A_0^2 A_2^* e^{-i(\delta_1 + \delta_2)t} + \alpha_1 \omega_2^3 A_1 |A_2|^2 \right],$$
(20)

where the coupling coefficients for filamentation and selfmodulation of EM waves are

$$G_0 \equiv \frac{e^2 \omega_{pe}^2}{8m_e m_i \omega_0^2 \omega_1^4 c_s^2}, \quad G_1 \equiv G_0 \frac{\omega_1^2}{\omega_0^2}, \quad \alpha_1 \equiv \frac{\omega_0^4}{\omega_1^4}$$
 (21)

(B90), while those for the oscillating two-stream instability are

$$G_{0} \equiv \frac{e^{2}}{8m_{e}m_{i}\omega_{0}\omega_{pe}^{2}\omega_{1}c_{s}^{2}}, \quad G_{1} \equiv G_{0}\frac{k_{s}^{2}}{k_{LI}^{2}}\frac{\omega_{pe}^{2}\omega_{1}^{2}}{\omega_{0}^{4}},$$

$$\alpha_{1} \equiv \frac{\omega_{0}^{5}}{\omega_{pe}^{4}}\frac{k_{LI}^{2}}{\omega_{1}k_{s}^{2}}, \qquad (22)$$

and  $k_{LI} \equiv k_s^2 + k_0^2$ . For both of these cases,  $\omega_1 = \omega_2$ , so  $G_2 = G_1$ .

# 4.2. Instability Growth Rates

The dominant initial terms in equation (20) and its twin for  $A_2(x, t)$  can be obtained by assuming that the pump wave amplitude  $A_0$  stays constant, so that these equations are a set of linear differential equations with constant coefficients. By assuming the variation is of the form  $\exp(i\Omega t)$ , B90 obtains this simple form for the dispersion relation for  $\Omega$ :

$$(\Omega - \delta_1 + i\gamma_1)(\Omega + \delta_2 + i\gamma_2) - \frac{1}{2}(\delta_1 + \delta_2)\omega_0 K = 0$$
, (23)

where  $K = 2G_1 |A_0|^2/\omega^2$ . Defining  $\Delta = (\delta_1 + \delta_2)/2$ , this equation gives the following threshold for instability:

$$K = -\frac{\gamma^2 + \Delta^2}{\omega_0 \Delta} \tag{24}$$

(B90, eq. 21), so that the instability grows for  $\Delta < 0$ , which can be shown to be usually the case for the conditions we are considering.

For the limit of an infinite wavelength pump wave or a transverse perturbation  $(v_s \perp k_0)$ , the density perturbation is purely growing. For the perpendicular case the frequency of the density perturbations is zero, and the frequencies of the sidebands are the same, locked to  $\omega_0$ . But for a longitudinal perturbation  $(k_s \parallel k_0)$  the instability excites a low-frequency wave and the two high-frequency sideband waves whose frequencies are shifted from their unperturbed values  $(\omega_{1,2})$  to  $\omega_0 \mp (\omega_1 - \omega_2)$ .

The growth rate arising from equation (23) is

$$\frac{\Gamma}{\omega_0} = -\frac{\gamma_1}{\omega_0} + \kappa_s (K - \kappa_s^2)^{1/2} \tag{25}$$

(correcting eq. 23 of B90), where  $\kappa_s = k_s v/\sqrt{2}\omega_0$ , and it is assumed  $\gamma_1 = \gamma_2$ . The maximum growth rate occurs at the wavenumber,  $k_{sm}$  for which  $\kappa_{sm} = (K/2)^{1/2}$ , and is roughly  $\Gamma_m \approx K\omega_0/2$ .

In order to evaluate these timescales in the framework of AGN conditions, we use the formula from Bingham & Lashmore-Davies (1979):

$$K = \frac{\omega_{pe} v_0^2}{4\omega_0^2 v_{T_e}^2},\tag{26}$$

where  $v_0 = eE_0/m_e \omega_0$  and  $v_{T_e}^2 = k_b T_e/m_e = c_s^2 m_i/m_e$ . Therefore,  $K = \omega_{pe}^2 e^2 E_0^2/4\omega_0^4 c_s^2 m_i m_e$ . We wish to estimate the strength of the pump wave in terms of the properties of the AGN. Assuming the bulk of the AGN continuum in the band around  $\omega_0$  is carried by this beam, then a rough approximation gives

$$E_0^2 \approx \frac{2P}{cR^2} \tag{27}$$

(see KW), where P is the luminosity in this beam and R is the size of the region from which it emerges. Utilizing equations (25)-(27) we find

$$\Gamma_m \approx \frac{e^2 \omega_{pe}^2 P}{4c \omega_0^3 k_b T_a m_a R^2} \,. \tag{28}$$

In terms of dimensionless variables using numbers expected to be of O(1) for optical emission from a quasar, we see from equation (28) that the timescale for the growth of these modu-

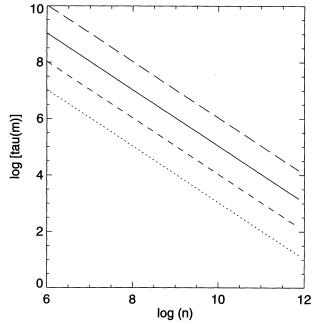


Fig. 2.—The modulational timescales (in s) plotted against ambient density (in cm $^{-3}$ ) at an observed frequency of 6  $\times$   $10^{14}$  Hz for a beam of  $P=10^{45}$  ergs s $^{-1}$ . The solid line is for  $R=10^{15}$  cm and  $T_e=10^6$  K; the dotted line corresponds to  $R=10^{14}$  cm and  $T_e=10^6$  K; the long dashed line has  $R=10^{15}$  cm and  $T_e=10^7$  K; and the short dashed line is for  $R=10^{14}$  cm and  $T_e=10^7$  K.

lational instabilities is

$$\tau_m \approx 2 \times 10^4 \frac{\omega_{0,15}^3 T_{e,6} R_{15}^2}{n_{e,9} P_{45}} \text{ s} .$$
(29)

Thus, as shown in Figure 2, fluctuations on subday timescales can certainly arise through this mechanism. In principle, this mechanism can work in any band of the electromagnetic spec-

trum if a strong pump wave can be established; the nonthermal emission produced in this fashion would be mildly polarized.

### 5. CHANGE OF EMISSION AND SCATTERING PROCESS

As mentioned in KW, coherence can yield very quick changes. More specifically, if the radiation generation process changes from the stimulated Raman scattering (SRS) to the stimulated Compton scattering (SCS) or vice versa owing to changes in the plasma parameters, the corresponding *e*-folding times of the radiation differ significantly. The inverses of the *e*-folding times for the energy loss of a relativistic electron beam through the processes of SRS and SCS are plotted as functions of the frequency of the emitted radiation for several sets of model parameters in Figure 3 (reproduced from Fig. 4 of Gangadhara & Krishan 1992, hereafter GK92).

It is seen in Figure 3 (also see Fig. 5 of GK92) that the change of emission process from SCS to SRS could manifest itself through a much quicker rate of energy release with a characteristic time of a few microseconds to a few milliseconds. The position of the change of processes could be shifted to other frequencies by appropriately changing the parameters. But because most of the sources show a "bump" in the blue region, we have chosen parameters that allow us to interpret at least part of this blue excess as arising from the change of emission process. This type of microvariability could arise if the electron beam suddenly encounters a plasma region with unfavorable parameters which thus disrupt the emission process completely or change it from SRS to SCS (or vice versa).

The lower frequency scattered emission occurs over much wider ranges of angles than does the higher frequency scattered emission (see Fig. 11 of GK92). Hence, a small change in the emission angle could either bring the radiation into our line of sight or completely remove this radiation from it. Thus, another reason for variability could be the scattering of the radiation in and out of the line of sight (to the receiver) during its propagation through the plasma surrounding its emission

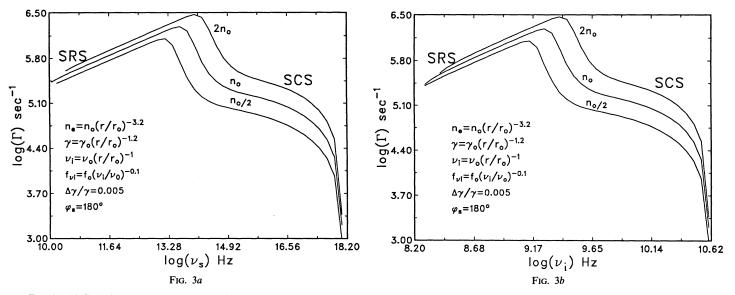


Fig. 3.—(a) Growth rate  $\Gamma$  vs. scattered wave frequency  $v_s$  at three different values of the electron density. At the maxima, the scattering process changes from SRS to SCS. The fixed parameters are:  $\gamma_0 = 3 \times 10^3$ ,  $n_0 = 9.24 \times 10^{17}$  cm<sup>-3</sup>,  $v_0 = 4 \times 10^{10}$  Hz, and  $f_0 = 7.1 \times 10^{-25}$  ergs cm<sup>-2</sup> s<sup>-1</sup> Hz<sup>-1</sup>. (b) As in (a), but for  $\Gamma$  plotted against incident wave frequency  $v_i$  at three different values of the electron density.

.994ApJ...423..172k

region. The scattering of radiation in an underdense plasma  $(\omega_0 \gg \omega_p)$  could be through SRS or SCS. The shortest characteristic timescales in which the wave vector  $k_0$  of the incident radiation swings into  $k_s$ , the wavevector of the scattered radiation in a thermal plasma of density n and temperature T for the two processes of SRS and SCS, are given by

$$t_{\text{SRS}}^{\min} = \left[\frac{v_0}{c} \omega_0 \left(\frac{\omega_{pe}}{\omega_0}\right)^{1/2} \sqrt{\frac{1 - \cos \phi_s}{2}}\right]^{-1} ; \qquad (30)$$

$$t_{\text{SCS}}^{\min} = \left[\frac{1}{2}\sqrt{\frac{\pi}{2}}\exp\left(-\frac{1}{2}\right)\left(\frac{v_0}{c}\right)^2 \frac{\omega_0}{(k_0 l_{\text{D}})^2}\sin^2\phi_s\right]^{-1}$$
(31)

(Drake et al. 1974), where  $\phi_s$  is the angle between  $k_0$  and  $k_s$ . Now, substituting in the expressions for the quiver velocity  $v_0$ , the plasma frequency  $\omega_{pe}$  and the Debye length  $l_D$ , we get:

$$t_{\rm SRS}^{\rm min} = \frac{cm_e}{e} \sqrt{\frac{c}{2}} \left(\frac{m_e}{4\pi e^2}\right)^{1/4} \left(\frac{P}{\omega_0 R^2}\right)^{-1/2} n_e^{-1/4} \left(\frac{1-\cos\phi_s}{2}\right)^{-1/2};$$
(32)

$$t_{\text{SCS}}^{\min} = \frac{m_0^2 c k_b}{4\pi e^4} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{2}\right) \frac{T_e}{n_e} \left(\frac{P}{\omega_0^3 R^2}\right)^{-1} \sin^{-2} \phi_s. \quad (33)$$

We compute the values of  $t_{\rm SRS}^{\rm min}$  and  $t_{\rm SCS}^{\rm min}$  for this nonthermal and weakly circularly polarized radiation using the typical plasma and radiation parameters in the emission region (Krishan & Wiita 1985). Choosing the temperature  $T_e = 10^4$  K and using our usual notation,

$$t_{\rm SRS}^{\rm min} = 2.931 \times 10^{-7} \left(\frac{P_{45}}{\omega_{10} R_{16}^2}\right)^{-1/2} n_2^{-1/4} \left(\frac{1 - \cos \phi_s}{2}\right)^{-1/2} s;$$
(34)

$$t_{\rm SCS}^{\rm min} = \frac{2.485 \times 10^{-5}}{n_2} \left( \frac{P_{45}}{\omega_{10}^3 R_{16}^2} \right)^{-1} \sin^{-2} \phi_s \, \text{s} \, . \tag{35}$$

In the radio emission region, if we take  $P_{45}=1$ ,  $\omega_{10}=1$ ,  $R_{16}=50$ , and  $n_2=6$ , we get  $t_{\rm SRS}^{\rm min}=9.36\times 10^{-6}([1-\cos\phi_s]/2)^{-1/2}$  s. For forward scattering  $(\phi_s=0)$ ,  $t_{\rm SRS}^{\rm min}\to\infty$  and for backward scattering  $(\phi_s=\pi)$ ,  $t_{\rm SRS}^{\rm min}\to 9.36\times 10^{-6}$  s. Thus, through SRS the radio wave vector  $k_0$  can change its direction into  $k_s$  in microseconds to years depending upon the scattering angle  $\phi_s$ . Again for SCS,  $t_{\rm SCS}^{\rm min}=0.01$  sin<sup>-2</sup>  $\phi_s$  s. If  $\phi_s\to 0$  or  $\pi$ , then  $t_{\rm SCS}^{\rm min}\to\infty$ , and when  $\phi_s\to \pi/2$  then  $t_{\rm SCS}^{\rm min}\to 0.01$  s, which is the shortest time in which  $k_0$  changes into  $k_s$  through SCS.

into  $\kappa_s$  in microseconds to years depending upon the scattering angle  $\phi_s$ . Again for SCS,  $t_{\rm SCS}^{\rm min} = 0.01 \sin^{-2} \phi_s$  s. If  $\phi_s \to 0$  or  $\pi$ , then  $t_{\rm SCS}^{\rm min} \to \infty$ , and when  $\phi_s \to \pi/2$  then  $t_{\rm SCS}^{\rm min} \to 0.01$  s, which is the shortest time in which  $k_0$  changes into  $k_s$  through SCS. In the optical region we can take  $R_{16} = 22$ ,  $P_{45} = 10^2$ ,  $\omega_{10} = 10^5$ , and  $n_2 = 10^8$ , and get  $t_{\rm SRS}^{\rm min} = 2.04 \times 10^{-6} ([1 - \cos \phi_s]/2)^{-1/2}$  s and  $t_{\rm SCS}^{\rm min} = 1.20 \times 10^3 \sin^{-2} \phi_s$  s. Finally, in the X-ray band we take  $R_{16} = 1$ ,  $P_{45} = 1$ ,  $\omega_{10} = 10^7$ ,  $n_2 = 10^{10}$ , to find  $t_{\rm SRS}^{\rm min} = 2.93 \times 10^{-6} ([1 - \cos \phi_s]/2)^{-1/2}$  s and  $t_{\rm SCS}^{\rm min} = 2.49 \times 10^6 \sin^{-2} \phi_s$  s.

Coherent emission processes have also been invoked by Weatherall & Benford (1991) and, particularly for explaining variability, by Lesch & Pohl (1992). The scattering of energetic electrons off strong Langmuir plasma waves is the basic radiation process proposed in these papers. Invoking relativistic beaming for enhancing the luminosity, Lesch & Pohl (1992) then go on to ascribe the intraday variability to the light travel time through a region of plasma with velocity distribution function f(v) such that  $(\partial f/\partial v) > 0$ , a necessary condition for the excitation of these Langmuir waves. In contrast, in SRS and

SCS, a low-frequency electromagnetic wave scatters off, respectively, a weak and a strongly damped Langmuir wave into a high-frequency electromagnetic wave in a relativistic plasma. The SRS changes into SCS when the Langmuir wavevector becomes larger than or equal to the Debye wavevector. As discussed above, this transition may cause variable emission.

Let us also mention that in addition to the processes discussed here, there are other nonlinear wave-wave interaction processes through which the radiation could be modulated. The modulation may take the form of spikes, coherent nonlinear structures called solitons, quasi-periodic intensity oscillations (e.g., Sastry, Krishan, & Subramanian 1981), and explosively growing waves. A clear description of some of these processes can be found in Weiland & Wilhelmsson (1977). We plan to study these processes in the context of AGNs in the near future.

## 6. CONCLUSIONS

Continued studies of the rapid variability of AGN in all wave bands should prove to be extremely fruitful, as they have the potential to probe regions very close to the central engine. An assortment of models, most of which were mentioned in § 1, have already been proposed for the physical source of such fluctuations. In this paper we have outlined a series of additional mechanisms that can act within the plasmas that almost certainly must exist in the vicinity of these powerhouses. Several of these plasma processes are, in principle, capable of producing significant flux variations on timescales of less than 1 day, and they can act over a very wide range of emission frequencies.

We have shown that both MHD fluctuations (§ 2) and modulations arising from the interaction of electromagnetic waves with the plasma through which they propagate (§ 4) are suitable candidates for these processes. It is also possible that flares arising from anomalous resistivity-driven instabilities can occur quickly enough to acount for much of the observed microvariability, although tearing modes probably operate too slowly (§ 3). Furthermore, coherent radiation mechanisms, one of which we have presented here (§ 5), may also play an important role. Coherent processes and some of the flare processes described here may even be relevant to the variability of AGNs in  $\gamma$ -rays, as recently discovered by the EGRET telescope (e.g., Hartman et al. 1992).

It does not seem as if the data collected to date can clearly eliminate any of the above ideas. But surely this is primarily because none of the proposed scenarios have been worked out in sufficient detail to allow strict confrontations with the data at hand. We plan on developing the details of several of these plasma-based mechanisms so that such careful comparisons can be performed, and we hope that all alternative models will also be developed and tested in more detail. It will be extremely important to determine the extent to which emission at different frequencies is correlated within both the models and in the observations. If the remarkable apparent correlations between optical and radio variations reported by Quirrenbach et al. (1991) for 0716 + 714 are seen again, either for that object or for other AGNs, then they will place strong constraints on most models. As discussed in Gopal-Krishna et al. (1993a, b), searches for microvariability in radio quiet QSOs can provide very important information for choosing between different classes of theoretical models.

Aside from flux monitoring at different wavelengths, it will be most important to examine the variation in polarization fraction and angle in as many bands as possible. Once the models are improved to the extent that they give specific predictions about polarization behavior, it is highly likely that many will be excluded, and a few supported, by current and future polarization measurements.

We thank the referee, Phillip Hughes, for his useful suggestions on presentation. V. K. is grateful to R. T. Gangadhara for his assistance. P. J. W.'s research in India is supported by the Smithsonian Institution through grant FR 10263600. This work was supported in part by NSF grant AST 91-02106 and by the Chancellor's Initiative Fund at Georgia State University.

#### REFERENCES

Kunieda, H., Turner, T. J., Awaki, H., Koyama, K., Mushotzky, R., & Tsusaka, Y. 1990, Nature, 345, 786

Lesch, H., & Pohl, M. 1992, A&A, 254, 29