Neutron Stars Observations as Astrophysical Probes

Thesis by

Roger William Romani

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1987

(Submitted May 6, 1987)

Acknowledgements

It is a great pleasure to thank Professor Roger Blandford who, in acting as my research advisor, has been a very inspiring source of ideas, encouragement, and insight into the business of doing astrophysics. It has been exciting to participate, in some measure, in his pursuit of this business; his excellent approach to research provides an example most worthy of study and emulation. I have particularly valued Roger's willingness to discuss a wide range of exciting topics. Perhaps the most interesting ideas were those for which the research is not yet complete, or even those which were merely subject of certain enthusiastic speculations.

The contribution of Professor Ramesh Narayan to this thesis is also most gratefully acknowledged. Ramesh served very much as a second advisor, particularly during the initial part of the research. His patient tutelage and his admirably lucid and didactic explanations were a welcome resource to an often rather dense student. I thank Professor Joe Taylor, as well, who generously provided an early opportunity to participate in astrophysical research.

I also wish to thank Professors Goldreich, Phinney and Thorne for providing a most pleasant, open and productive research environment. An early course from Peter was most helpful in showing the value of dimensional analysis and of seeking the important part of a problem. The several other members of the TAPIR research group have been very helpful at numerous junctures in this research, as well. In particular, I and the whole group owe Sam Finn a debt of gratitude for his unflagging help with the computing facilities and his willingness to tutor group members the intricacies of their use. John Hawley was also most helpful in discussing various issues of physics and numerics. Our administrative aide, Pat Lyon, has been a perpetual wonder as she not only succeeds in keeping the group running, but manages to do so smoothly, efficiently and with remarkable good humor.

I am happy to acknowledge the generous support of a fellowship from the Fannie and John Hertz Foundation which allowed me to embark on this research program much earlier than would otherwise have been possible. Additional support for the work described was provided by the National Science Foundation.

Support, non-financial, but no less important, has been extended by a number of persons. In this context, I am very happy to acknowledge the ongoing enthusiasm of my parents, Roger and Ann

Romani, and the rest of my family. The support of numerous friends, teammates and colleagues, and particularly Ruth, was important in making this period at Caltech both pleasant and productive.

Finally, as a brief apologia, it should be mentioned that the stated 'theme' for the research described herein should be not interpreted as the underlying reason for the program of investigation, but rather as something somewhat closer to an excuse. The fundamental goal of these studies was, of course, the further elucidation of certain astrophysical phenomena as well as the acquisition of some understanding of the process of astrophysical research. The investigation of neutron stars, dealing with a wide range of physical parameters and processes, provides an admirable vehicle for such studies, and it is in this spirit that the following thesis is assembled.

Abstract

Several aspects of observations of neutron stars and compact extragalactic radio sources are considered, with particular regard to their use in constraining certain astrophysical phenomena.

A theoretical treatment of the pulse arrival time analysis of millisecond pulsars is made; we consider how a detailed timing analysis can be used to quantitatively probe noise processes affecting the pulsar period and the propagation of the radiation. The intrinsic noise may be used to study the neutron star interior, while propagation effects due to gravitational perturbations of the neutron star and interstellar refraction of the emitted radio waves provide probes of the pulsar environment and the intervening plasma. In addition, important constraints on the background of cosmological gravitational radiation can be derived from timing such pulsars.

We consider the thermal X-rays emitted from a warm ($10^5\,\mathrm{K} \lesssim T_{eff} \lesssim 3 \times 10^6\,\mathrm{K}$) neutron star, either cooling from its initial formation or heated by internal dissipation, accretion, *etc*. Constructing model atmospheres appropriate to such stars with various effective temperatures and elemental abundances, we calculate their emergent spectra and the bolometric correction for observation bands of various X-ray satellites. We conclude that the present limits on neutron star surface flux are even more constraining than those derived assuming that the spectra are blackbody and examine how this effects models of neutron star interiors, formation and cooling. We also examine the consequences of similar X-ray observations for neutron star models of various gamma-ray stars.

The refraction of radio waves from pulsars and other compact sources by interstellar plasma is also studied. We show how pulsar observations, in particular, can be used to characterise the large scale inhomogeneities in the ionized ISM and compute a number of observable effects for various electron density perturbation spectra. It is shown how similar refraction can account for the low and intermediate frequency variation of compact extragalactic radio sources. We argue that the observations indicate that more power is present in the large scale fluctuations than previously believed. In addition, single $\sim 10^{14}$cm scale clouds in a previously unrecognized dense, ionised phase of the ISM can apparently dominate the refractive scintillation for some lines of sight.

Contents

Introduction

To preface a discussion involving neutron stars, it is appropriate to open by recalling the prescient suggestions of Landau (Rosenfeld 1974), Baade and Zwicky (1934); and, later, Wheeler (1966) and Pacini (1967); the dramatic and unexpected discovery of pulsars by Bell working in the research team of Hewish (Hewish *et al.* (1968)); and the rapid and clear interpretation of these observations by Gold (1968). The subsequent elaboration of our description of pulsars (*eg.* Goldreich and Julian 1969, Ostriker and Gunn 1969) and the realization that neutron stars are the central powerhouses of a number of astrophysical sources (particularly the X-ray binaries) has led to a rapid growth in the depth and breadth of theoretical studies of these compact objects, so that today neutron stars reign supreme as the *deus ex machina* of most galactic high energy astrophysical sources, whether modeled or as yet unexplained.

Observational advance has been rapid on a number of fronts, as well. The early observations of the Crab pulsar, PSR0531+21, in particular, were of critical importance in consolidating support for the basic neutron star model. Firstly, since the Crab and, nearly simultaneously, the Vela pulsars were found towards the centers of young supernova remnants, the compact objects were strongly associated with the collapse of massive stars. Indeed the discovery that Crab's short (33 ms) period was growing, with a $\dot{P} \sim 4\times 10^{-13}$s s^{-1}, both substantiated the belief that the object was powered by tapping the rotational energy, and, via the maximum characteristic age $P/2\dot{P} \sim 2000$yr, established the association with the historical supernova of 1054. The discovery of the Crab pulsar also solved the long-standing puzzle of the energization of the surrounding synchrotron nebula; conversely, the observed power in the nebula gives a measure of the pulsar moment of inertia, since the rotational spindown supplies $I\Omega\dot{\Omega}$ in power. The resulting value, $I \sim 10^{45-46}$gm cm^2, is in agreement with that estimated for a ~ 1M$_\odot$, 10km radius neutron star. The rate of change of the period derivative has also been measured for the Crab pulsar and the results are in rough accord with the hypothesis of magnetic dipole braking. Finally the Crab, and especially the Vela pulsar are notable for showing occasional large discontinuities in the period and period derivative; these glitches yield important clues to the internal structure.

More recently, study of the ~ 500 presently known rotation powered (radio) pulsars (*cf.* Manchester and Taylor 1977) has allowed a substantial investigation of the number and origin of this observed sample of neutron stars. Detailed consideration of the observational biases in pulsar searches

(Lyne, Manchester and Taylor 1985; Narayan 1987) have shown that the number of potentially observable radio pulsars in the galaxy is $\sim 10^5$. The galactic distribution is strongly concentrated to the plane, giving a disk population of scale height $\sim$ 400pc, with some increase in the pulsar number density towards the galactic center. An important tool for understanding the pulsar population is the $P - \dot{P}$ diagram (see Fig. 1, from Dewey, *et al.* 1986). The quantity $P\dot{P}$ is related to the pulsar magnetic field in the dipole braking model, and hence to the luminosity. The spin down evolution of a pulsar can be thought of as driving a "current" in the $P - \dot{P}$ plane (Phinney and Blandford 1981). A pulsar is born in the upper left (*cf.* Crab and Vela) of the diagram, moves down and to the right and disappears when the luminosity (related to some combination of P and $\dot{P}$) decreases below a critical value (the "Death Line"). The pulsar statistics give evidence for an additional luminosity evolution due to decay of the magnetic field, with a timescale $\lesssim 10^7$ years. The birthrate inferred is about 1 pulsar per 30 to 120 years, which is in rough agreement with the deathrate for the extreme population I OB stars, believed to be the principal progenitors of pulsars. Although the scale height for this progenitor population is very much less than that of the pulsars, the observed pulsar velocity dispersion normal to the plane is $\sim$ 100km/s, so it is clear that they can travel substantial distances from their birthsites; the observed scale height is consistent with the field decay lifetimes.

The association of neutron star births with extreme population I stellar deaths and supernovae is, however, somewhat problematic. Although theoretical arguments and the situation of certain young pulsars in supernova remnants suggest that neutron stars are born in the demise of massive stars, it is not certain if the estimated galactic supernova rate is adequate to explain the pulsar birthrate. If the estimates are to agree, it seems necessary that the majority of supernovae give rise to radio pulsars. However, observations with the *Einstein* satellite (Helfand and Becker 1984) have failed to find evidence for neutron stars in a number of young supernova remnants. The inference of an additional, more quiet, arena for neutron star birth has sparked a number of theoretical and observational investigations. In this context observations of the recent supernova SN1987a in the LMC will be of particular interest. The apparent detection of a supernova neutrino burst in water Cerenkov experiments indicates that core collapse to a compact remnant has probably occurred; if a neutron star has formed future X-ray and radio observations should detect the thermal flux from the cooling remnant and determine if a young pulsar is present.

Although certain anomalous objects (*eg.* the binary pulsar PSR1913+16, Hulse and Taylor 1975) provided an indication of the presence of a second pulsar population, the discovery of the

millisecond pulsar (Backer, *et al.* 1982) highlighted the importance of a subgroup of rapidly spinning radio pulsars with low magnetic fields. These objects, lying in the lower-left hand region of the $P - \dot{P}$ diagram, are believed to be "reborn", low field pulsars brought above the death line by accretion induced spin-up, which should shift the "dead" field-decayed pulsars to the left until they reach the "spin-up" line (a limiting period for a given field and accretion rate). While it has been suggested that these objects are associated with the low mass X-rays binaries (Alpar, *et al.* 1982) and, in particular, the recently discovered X-ray binaries showing Quasi-Periodic Oscillations (van der Klis 1985), the evolutionary connection is still far from clear.

The monitoring of departures from the very stable rotation rates of radio pulsars has led to a description of the interior of a neutron star as composed of several superfluid and normal components, interacting in complex ways. In particular, the observations of pulsar glitches (*cf.* review by Pines and Alpar) and timing noise (*eg.* Cordes and Helfand 1980) have given important insights into the structure of neutron star interiors. The current picture is summarized in Figure 2 (from Shapiro and Teukolsky 1983). While there are variations in the internal structure as a function of the poorly known high-density equation of state, most models have in common a number of distinct regions. In the outermost few meters the matter is normal, although it can be significantly affected if strong magnetic fields are present. In the low density crust, neutron-rich high-Z nuclei exist along with a degenerate electron gas. At higher densities, neutrons can exist freely (outside of nuclei) and a superfluid neutron gas forms along with exotic, high-Z nuclei. At still higher densities, nuclei lose their identities and the bulk of the matter is superfluid neutrons, with a small admixture of superfluid protons. Finally, at the core, densities of several times nuclear density can be reached and formation of a hyperon-, pion- or quark-condensate may be energetically favored. In a glitch, the steady spin-down mediated by the outward diffusion of quantized superfluid vortices through the various regions is perturbed. The dynamical response of these regions as the star relaxes to a steady state gives information on the couplings, moments of inertia, *etc.* of the various zones. The Crab, Vela and a number of older pulsars have shown glitches with dynamic relaxations broadly consistent with the above description, although the detailed behaviour is not fully explained.

At higher energies there is much to be learned about basic neutron star physics as well. Pulse timing and eclipse modeling has yielded mass estimates for a number of neutron stars in binary X-ray pulsar systems; along with the mass estimate for PSR1913+16, derived from general relativistic effects, these are all consistent with a value $\sim 1.2 - 1.4 M_{\odot}$, *ie.* close to the Chandrasekhar limit

(Joss and Rappaport 1984). Recent advances in modeling the time dependence of X-ray bursts, believed to arise from thermonuclear explosions on neutron star surfaces, are also beginning to provide significant constraints on neutron star masses and radii. Finally, as a more speculative possibility, some evidence suggests that the enigmatic gamma-ray bursters may be associated with an old population of slowly rotating neutron stars. If so, such behaviour may be the only means of detecting these low period, low field neutron stars, aside from an occasional thermal pulse from rare spindown glitches in very nearby examples.

While many observed aspects of neutron stars await adequate explanation (most notably the question of how pulsars emit coherent radio waves, in the first instance) in this discussion we will take the basic properties of radio pulsars and neutron stars as given and consider the interpretation of the observations in somewhat different light; *ie.* we study how certain aspects of radio and X-ray observations of neutron stars may be used to probe extreme and otherwise unobservable regimes of astrophysics. More specifically, the investigations conducted during the course of this thesis and reported herein relate to three different uses of neutron stars as tools for the study of other physical phenomena; timing of fast, quiet pulsars and their use in searching for ultralow-frequency gravitational radiation, X-ray flux limits for neutron stars in young supernova remnants and the possible need for exotic cooling agents due to phase transitions at supranuclear densities, and the use of pulsar arrival time and intensity modulations in the study of the ionized component of the intervening interstellar medium (ISM). This division allows a rough grouping of the research undertaken during this thesis into three sections. In the remainder of this introduction, we consider these three sections in turn, providing a brief description of previous studies using neutron stars as probes of the specified phenomena, describing how the investigation undertaken relates to these earlier findings, and summarizing the principal results of the papers contained.

I. Pulsar Timing and Limits on Ultralow-Frequency Gravitational Radiation and Other Sources of Timing Noise.

Insofar as quiet pulsars do not depart measurably from a simple, steady spindown, one may use the limits on their irregularity to constrain a number of processes which would result in timing noise. In particular, Sazhin (1978) realized that gravitational waves emitted by a line-of-sight binary could produce a periodic perturbation in a pulsar's arrival time (*ie.* a residual to the timing model) and Detweiler (1979) showed how a stochastic background of ultralow-frequency gravitational radiation,

a relic of the early universe, could produce a measurable rms timing residual in an otherwise quiet pulsar. Romani and Taylor (1983), Hellings and Downs (1983) and Bertotti, Carr and Rees (1983) made use of this idea to show that one could limit the energy density in a background of gravitational waves with periods ∼1-10 years, to less than the density required to close the universe with such radiation.

The discovery of the millisecond pulsar (Backer, *et al.* 1982), with its short period, slow spindown and remarkably small timing noise has proved most valuable for this method; monitoring this object over the past several years has yielded limits on the relic gravity wave background many orders of magnitude stronger than those mentioned above (Taylor 1987). During the same period there has been increasing interest in cosmic strings, singularities frozen out from phase transitions in the very early universe, and scenarios in which these objects might produce the density perturbation spectrum that seed the formation of the galaxies seen today (Zel'dovich 1980, Vilenkin 1981). However, these massive, relativistic objects must produce copious gravitational radiation as they oscillate and collide, indeed such radiation is the principal sink for their energy. Current limits from timing the millisecond pulsar show that the relic background from these strings must be

$$\Omega_{GW}(f) \leq 10^{-6}(f/10^{-8}Hz)^4, \qquad f \gtrsim f_{min}$$

in units of the closure density, where the wave frequency, f, is greater than some minimum wavelength of order 10^{-8}Hz. Cosmic string scenarios for galaxy formation require $\Omega_{GW} \gtrsim 10^{-7}$ (Hogan and Rees 1984, Vachaspati and Vilenkin 1985) and so pulsar timing is already providing important constraints on these theories. Since this limit depends quite sensitively on f_{min}, only a detailed treatment of the timing model allows one to be quantitative in placing these bounds. Moreover, for long periods, the timing accuracy of PSR1937+214 currently exceeds that of terrestrial clocks, accordingly the avenue for progress in pulsar timing gravity wave searches lies in finding other fast, quiet pulsars and establishing an array of well-timed pulsars on the sky. Such a timing network would allow the separation of several timing noise processes with different angular dependencies; in particular, a gravitational wave perturbing the Earth will produce timing residuals (in effect a Doppler ranging to the various pulsars) which will have a quadrupolar angular signature. The recently discovered PSR1855+09 (Segelstein *et al.* 1986) shows some promise for such use.

In paper I (Blandford, Narayan and Romani 1984) the effect of various types of timing noise on a pulsar arrival time analysis were considered. In particular, red noise processes dominated by low

frequencies, such as those produced by a cosmic string gravity wave background or pulsar seismicity, were shown to induce variances in the fitted pulsar parameters that reduce more slowly with time than those caused by random, white-noise processes. Conversely, since intrinsic pulsar parameters (such as the period, period derivative and position) are refined in the analysis, a certain fraction of the residual due to the noise process is absorbed. An evaluation of this effect for various noise spectra was performed; in the context of a limit on the gravity wave background, this means that the effective f_{min} above is somewhat larger than the reciprocal of the observation period. A more precise limit on the background energy density is obtained and it is shown that, if PSR1937+214 continues to show predictable timing behavior, string scenarios of galaxy formation will soon face severe constraints.

In paper II (Blandford, Romani and Applegate 1987), we consider briefly a most promising candidate for a new fast pulsar located in the direction of the globular cluster M28. Certain evolutionary arguments suggest that pulsars may be spun-up in the cores of globular clusters, and the present limit on the pulse period for this source suggests that if it is a pulsar, its period must be on the order of a few msec. We point out that the gravitational perturbations of the other cluster members should produce a measurable variation in the arrival times and that this will degrade such pulsar's utility in gravity wave search timing programs.

II. Model Atmospheres for Cooling Neutron Stars and Limits on the Surface Flux of Isolated Neutron Stars.

The association of neutron stars with supernovae suggests that immediately after birth the collapsed core will be very hot, $\sim 1 - 10 \text{MeV}$, and should emit a large flux of X-rays. Early analyses (Chiu and Salpeter 1964, Tsuruta 1964) showed that once the photon cooling dominates that of neutrinos, the surface temperature of the young neutron star will decrease relatively slowly, leading to a potentially detectable X-ray flux persisting for $\sim 10^4 - 10^5$ years. However, it was pointed out by Bahcall and Wolf (1965) that if a substantial portion of the dense core underwent a phase transition to a pion condensate, the initial neutrino cooling would be dramatically enhanced and the thermal flux of photons would decrease by a factor $\sim 10^3 - 10^4$ at early times. With the advent of the *Einstein* satellite, detection of the thermal flux of young neutron stars became a realistic observational goal and a number of groups made detailed recalculations of neutron star cooling, including substantially more realistic interior physics. Superfluidity, magnetic fields and departures

from isothermality in the core were shown to have important, but relatively modest effect on the cooling (*eg*. Hernquist 1984); exotic phases at supranuclear densities such as quark, kaon, and pion condensates were found to produce dramatic departures from the standard cooling theory (Tsuruta 1985). In addition, it was suggested that older neutron stars might maintain modest $\gtrsim 10^5$ K surface temperatures for extended times via accretion from the general ISM, heating of the polar caps by magnetospheric particles, or dissipative heating from internal seismicity (Helfand, Chanan and Novick 1980).

The *Einstein* survey of nearby pulsars and young supernova remnants (*eg*. Helfand, Chanan and Novick 1980, Helfand 1983) did not, however, provide any convincing detections of thermal flux from the initial cooling. Moreover, rather severe limits were placed on the flux from any X-ray point source in several young supernova supernova remnants and important upper bounds were obtained for the thermal component of the X-ray emission of the Crab and Vela pulsars. These observations are particularly important since, although one may miss radio pulsars in young supernova remnants because of beaming or the late turn-on of the pulsar phenomenon (*eg*. Blanford, Applegate and Hernquist 1984), the thermal flux cannot be hidden in this manner. In the case of several well studied remnants, one may infer that either no neutron star is present or that accelerated cooling from an exotic interior phase was important. If the former supposition applies and pulsars are produced only in a minority of supernovae, then the already strained agreement of the supernova and pulsar birthrates would be violated and an additional, quiet arena for pulsar formation must be found. If the alternative of accelerated cooling is applicable, then these observations give important information on the equation of state at densities and not presently accessible in terrestrial experiments. In the case of the Vela pulsar the limits on surface flux are at best only marginally compatible with standard cooling theory and it may again be necessary to invoke an exotic internal phase. However, before such significant conclusions may be reached, it is important to quantify the connection between the surface heat fluxes predicted by the cooling calculations and the upper limits to counting rates in specific energy bands provided by the observations. In particular, previous authors have assumed that the flux emerging from the neutron star surface had a simple blackbody spectrum, an assumption which might not, in general, be valid. Further, *Einstein* observations have provided important limits on the flux from older radio pulsars, as well as some possible detections of various other astronomical sources suspected to involve neutron stars, *eg*. gamma-ray burst sources and Geminga. To compare these observations with the source models it is clearly important to characterise the thermal flux from a neutron star surface.

In paper III (Romani 1987), we have computed model atmospheres for cool neutron stars, considering the effect of the radiative transfer through the final few centimeters, where atomic processes will have important effects on the emergent spectrum. Realistic opacity data for densities and temperatures appropriate to non-magnetised cooling neutron stars were obtained from the Los Alamos Astrophysical Opacity Library and LTE model atmospheres were constructed using a temperature correction scheme employing detailed radiative equilibrium. We found that the largest effect of the varying surface gravity was due to a simple red-shift to infinity and atmospheres were computed for a range of effective temperatures and surface compositions. The effect of accretion and gravitational settling on the surface layers is discussed. The results show that, for atmospheres dominated by low-Z elements such as hydrogen and helium the emergent spectra would be substantially harder than those of blackbodies at the corresponding effective temperatures. Conversely, atmospheres dominated by heavy elements would have spectra closer to the blackbody form, but with strong absorption edges. The application of these results to the limits obtained for young SN remnants and radio pulsars with *Einstein* and *EXOSAT* were considered and we show that the upper bounds on the effective temperatures are substantially lower, in many cases, than those found assuming blackbody spectra. This strengthens the conclusion that if neutron stars are present in a number of well-studied SN remnants, then phase transitions to exotic matter should occur at supranuclear densities.

In paper IV, we apply the results of Romani (1987) to *Einstein* observations of the fields of several gamma-ray sources. In the case of the gamma ray burster GRB781119 and the enigmatic source Geminga, candidates for soft X-ray counterparts have been found. This X-ray flux has been interpreted as thermal radiation from the neutron star surface. We note that the presence of accreted light elements can alter the spectrum of the emergent flux as described above, discuss how the limits on the thermal flux will be affected and briefly consider the implications for astrophysical modeling of such sources.

III. Low Frequency Variability of Pulsars and Compact Extragalactic Radio Sources, Refractive Interstellar Scintillation and Ionised Plasma in the Interstellar Medium

Pulsars are remarkably good probes of conditions in the galactic interstellar medium (ISM). Since pulsar signals are broad-band, impulsive and polarized, one can study the interstellar hydrogen and free-free electron absorption, measure dispersion by the ISM plasma, and, using the observed

Faraday rotation, estimate the galactic magnetic field along various lines of sight. Moreover, since pulsars are effectively point sources, interstellar scintillation allows one to study inhomogeneities in the ISM electron density distribution. Since the earliest days of pulsar observation, the narrow band, $\sim$100s intensity fluctuations and the pulse broadening seen at low frequencies have been attributed to scattering by these electron density fluctuations (Scheuer 1968). This diffractive scintillation causes a mean delay in the pulse arrival time and a decorrelation bandwidth for the diffraction pattern seen at the Earth, whose transverse scale, b, at an observing wavelength λ gives an estimate of the typical scattering angle, $\theta \sim \lambda/2\pi b$. A pulsar at a distance D subtends an angle very much smaller than b/D (which is $\sim 10^{-12}$ for typical parameters) and so rays arriving from this spread of angles can interfere incoherently. However, this condition on the intrinsic source size, $\theta_{int} < 10^{-3}$mas, is so restrictive that pulsars are the only known sources compact enough to show strong diffractive scintillation.

From low frequency VLBI observations of the apparent size of pulsars (Mutel, *et al.* 1974) and measurements of the pulse broadening (Slee, *et al.* 1980) it is determined that the mean bending angle experienced by a meter wavelength ray propagating from a $\sim$1kpc distant pulsars is of order 1-10 mas. In the standard picture, this angular broadening is effected by the random superposition of the ensemble of small-scale, weak scatterers along the line of sight. The density fluctuation power spectrum of these scatterers has been most commonly characterised as a extended three-dimensional power law in spatial frequency, $\Phi_k \propto k^{-\beta}$; in particular it has been argued that the index of this spectrum has the Kolmogorov value, $\beta = 11/3$ (Rickett 1977, Armstrong, Cordes and Rickett 1981) and that the scattering is dominated by fluctuations on scales $\sim 10^9$cm. However, this value is close to the critical value, 4, above which large-scale fluctuations dominate, so the low spatial frequency fluctuations can be important. Indeed, the observation that diffraction maxima in the frequency-time plane (dynamic scintillation spectra) of pulsars often show shearing and large-scale organization suggested that large scale prisms (*ie.* the edges of $\sim 10^{13-14}$cm perturbations) induce chromatic aberration into the scintillation (Shishov 1974, Hewish 1980).

Moreover, it was shown by Sieber (1982) that the modulation indices describing the long period (months-years) variation of pulsar fluxes correlated with the dispersion measure, suggesting a propagation effect. Following this observation, Rickett, Coles and Bourgois (1984) (also *cf.* Shapirovskaya 1978) argued that the intensity fluctuations were *refractive* in nature, caused by focussing and defocussing of the pulsar flux by electron density perturbations on the scale of the scatter-broadened

image. It was also noted that a number of other astrophysical radio sources were sufficiently compact, $\lesssim$ a few mas, to show such effects. In particular, it was suggested that this effect could explain the problematic low-frequency variability of compact extragalactic radio sources. The fluctuations of these sources on month to year timescales at meter wavelengths, if intrinsic, contrast sharply with the light travel time across the minimum source size inferred from brightness temperature arguments, often suggesting apparent velocities of several hundred c. Although VLBI observations have directly demonstrated the presence of bulk relativistic motions in several extragalactic radio sources (Cohen 1986), the prevalence of low frequency variability creates difficulties for the relativistic beaming models (invoked to explain the apparent superluminal motions in the VLBI observations), which would be mitigated if most low frequency variability was a propagation effect. Rickett, Coles and Bourgois (1984) also suggested that a related phenomenon, the few percent, $\sim$10d timescale fluctuation of compact sources at $\sim$10cm observing wavelengths discovered by Heeschen (1984) could be similarly explained as an effect of refractive scintillation.

It was pointed out by Blandford and Narayan (1984,1985) that when pulsars are subjected to this refractive scintillation, parameters other than the flux would be modulated, as well. They established a simple model for the effect of refraction caused by a thin, phase-changing, $\beta \lesssim 4$ screen and showed how various observables such as flux, apparent angular size, and pulse width would show correlated fluctuations with amplitudes depending on the slope of the underlying perturbation spectrum. These calculations indicated that the observed flux variations required an ISM electron density fluctuation spectrum with $\beta > 11/3$. Goodman and Narayan (1985) performed a more exact calculation of the flux variation spectrum from a thin power law phase screen, including the case $4 < \beta < 6$, where large scale fluctuations dominate the scattering. Further, this work showed that the wavelength dependence of the decorrelation band width, scatter broadening, *etc.* used previously to infer a value of 11/3 for β were equally compatible with $\beta \sim 4.3$. For this spectrum the refractive effects are greatly enhanced and large, frequency independent values are expected for the modulation indices of pulsars and compact extragalactic radio sources, in better agreement with observation. In addition, observations of dynamic scintillation spectra by Roberts and Ables (1982) and Hewish, Wolszczan and Graham (1985) also suggested that the effective β for large scales is > 4.

In paper V (Romani, Narayan and Blandford 1986), the scattering model of Blandford and Narayan (1985) was extended to include diffractive effects, such as the scintillation timescale, the decorrelation bandwidth and the drifting bands seen in pulsar dynamic scintillation spectra. Auto-

and cross-correlations for the fluctuations of these and several other observables were computed and shown to be useful probes of refractive scintillation theory. In addition, it was shown how the scattering model could be extended to spectra described by power laws with $\beta > 4$ and these calculations were shown to agree with the more precise results derived for the flux variations by Goodman and Narayan(1985). The validity of the thin screen approximation was also tested, by developing a formalism to treat scattering in an extended medium. It was found that computing in terms of an equivalent thin scattering screen sometimes underestimates the refractive effects by as much as a factor of ~ 2, especially for shallow, $\beta < 4$ spectra. The comparison of these results with observations of pulsars and other compact radio sources was discussed and it was suggested that a more complicated spectrum than a simple power law might be indicated.

In paper VI (Blandford, Narayan and Romani 1986) it was shown how flicker of compact extragalactic radio sources, *ie.* few percent fluctuations at ~ 10 cm observing wavelengths, could be explained as an effect of refractive scintillation. In particular, it was shown how an extended scattering medium and a variation of source size with observing frequency, consonant with that expected for self absorbed synchrotron sources, could bring the theory into agreement with the observations of Simonetti, Cordes and Heeschen (1985). We also examined the interesting possibility that these fluctuations, if sufficiently understood, could provide a means of resolving very compact sources.

Paper VII (Goodman, Romani, Blandford and Narayan 1987) examines the refractive scintillation from a power law spectrum of ISM electron density fluctuations with a significant inner scale. It is shown that suppressing the small scale fluctuations causes a relative enhancement of the slow, "refractive" portion of the flux variation. An evaluation of the exact flux variation spectrum demonstrates that there is additional power between the refractive and diffractive regimes due to caustic focussing events and that, for inner scales somewhat larger than the Fresnel scale, this can be the dominant component of the modulation. These caustics are examined in the context of catastrophe theory, which allows them to be classified into certain generic forms. In the case of interstellar scintillation, the plasma dispersion introduces important frequency dependencies into the scaling laws for the various catastrophes; the most important cases, the fold and the cusp, are given explicitly. It is shown that when a substantial inner scale obtains and caustics are important, there are specific predictions for the time and frequency behavior of many pulsar observables, such as the dynamic scintillation spectra. We demonstrate how features such those observed can be manifestations of

particular catastrophes. It is not clear if such large inner scales are indicated for the ISM; however, we note that observations of pulsars seem to require an enhancement of the refractive component of interstellar scintillation over that expected from a simple $\beta < 4$ power law and caustics are likely to be found, at least on occasion, for any scintillation spectrum with such enhanced large scale focussing.

In paper VIII (Romani, Blandford and Cordes 1987) we show that the dramatic focussing events found in the light curves of certain compact extragalactic sources at 2.7 and 8.1GHz by Fiedler, *et al.* (1987) can be understood as the signature of the caustics from single large scale, high-density plasma lenses in the ISM. It is argued that these observations provide evidence for a new component of the ionized ISM (termed the DIM for dense, ionized medium) with single, $\sim 10^{14}$cm scale regions of very high density ($\sim 100-1000$cm^{-3}) plasma. It is suggested that this plasma might plausibly be confined in old supernova remnants or in magnetic ropes. Since scintillation by such single strong scatterers is in marked contrast to the standard interpretation in terms of multiple weak scatterings, this picture is shown to have wide ranging for observations of pulsars and high-resolution, intermediate frequency VLBI of compact sources.

References

Alpar, A., Cheng, A., Ruderman, M. and Shaham, J. 1982, *Nature*, **300**, 728.

Armstrong, J.W., Cordes, J.M. and Rickett, B.J., 1981, *Nature*, **291**, 561.

Baade, W. and Zwicky, F. 1934. *Phys. Rev*, **45**, 138.

Backer, D., Kulkarni, S., Heiles, C., Davis, M., and Goss, M. 1982, *Nature*, **300**, 615.

Bahcall, J.N. and Wolf, R.A. 1965, *Phys. Rev.*, **140B**, 1452.

Bertotti, B., Carr, B.J. and Rees, M.J. 1983 *Mon. Not. R. ast. Soc.*, **203**, 945.

Blandford, R., Applegate, J. and Hernquist, L. 1983, *Mon. Not. R. ast. Soc.*, **204**, 1025.

Blandford, R. and Narayan, R. 1984,Proceedings of the Workshop on Millisecond Pulsars, Reynolds, S. and Steinbring, D. eds., Grren Bank, Va.

Blandford, R. and Narayan, R. 1985, *Mon. Not. R. astr. Soc.*, **213**, 590.

Blandford, R., Narayan, R. and Romani, R.W. 1984, *J. Astrophys. Astron.*, **5**, 369.

Blandford, R., Narayan, R. and Romani, R.W. 1986, *Astrophys. J.*, **301**, L57.

Blandford, R., Romani, R.W. and Applegate, J. 1987, *Mon. Not. R. ast. Soc.*, in press.

Chiu, H.-Y. and Salpeter, E. 1964, *Phys. Rev. Lett.*, **12**, 413.

Cohen M. 1986, in *Highlights of Modern Astrophysics*, Shapiro, S. and Teukolsky, S., eds., Wiley, New York.

Cordes, J. and Helfand, D. 1980, *Astrophys. J.* , **239**, 640.

Detweiler, S. 1979, *Astrophys. J.*, **234**, 1100.

Dewey, R.J., Maguire, C.M., Rawley, L.A., Stokes, G.H. and Taylor, J.H. 1986, *Nature*, 322, 712.

Fiedler, R.L., Dennison, B., Johnston, K.J. and Hewish, A. 1987, *Nature*, submitted.

Gold, T. 1968, *Nature*, **218**, 731.

Goldreich, P. and Julian, W.H. 1969, *Astrophys. J.*, **157**, 869.

Goodman, J. and Narayan, R., 1985, *Mon. Not. R. astr. Soc.*, **214**, 519.

Goodman, J., Romani, R.W., Blandford, R. and Narayan, R. 1987, *Mon. Not. R. ast. Soc*, submitted.

Heeschen, D.S. 1984. *Astron J.*, **82**, 701.

Helfand, D. 1983, in IAU 110, *Supernova Remnants and their X-Ray Emission*, eds. Danziger, J. and Gorenstein, P., Riedel, Dordrecht, 471.

Helfand, D. and Becker, R. 1984, *Nature*, **307**, 215.

Helfand, D., Chanan, G. and Novick, R. 1980, *Nature*, **283**, 337.

Hellings, R.W. and Downs, G.S., *Astrophys. J. Lett.*, **265**, L39.

Hewish, A., 1980, *Mon. Not. R. astr. Soc.*, **192**, 799.

Hewish, A., Bell, J.S., Pilkington, J., Scott, P. and Collins, R. 1968, *Nature*, **217**, 709.

Hewish, A., Wolszczan, A. and Graham, D.A., 1985, *Mon. Not. R. astr. Soc.*, **213**, 167.

Hogan, C.J. and Rees, M.J. 1984, *Nature*, **311**, 109.

Hulse, R. and Taylor, J. 1975, *Astrophys J. Lett.*, **195**, L51.

Joss, P.C. and Rappaport, S.A. 1984, *Ann. Rev. Ast. Astrophys.*, **22**, 537.

Lovelace, R.V.E., 1970, Ph.D. Thesis, Cornell University.

Lyne, A. Manchester, R. and Taylor, J. 1985, *Mon. Not. R. ast. Soc.*, **213**, 613.

Manchester, R. and Taylor, J. 1977, *Pulsars*, Freeman, San Francisco.

Mutel, R.L., Broderick, J.J. Carr, T.D., Lynch, M., Desch, M., Warnock, W.W. and Klemperer, W.K. 1974, *Astrophys J.*, **193**, 279.

Narayan, 1987, preprint.

Ostriker, J.P. and Gunn, J.E. 1969, *Astrophys. J.*, **157**, 1395.

Pacini, F. 1967, *Nature*, **216**, 567.

Phinney, E.S. and Blandford, R.D. 1981, *Mon. Not. R. ast. Soc.*, **194**, 137.

Pines, D. and Alpar, A. 1985, *Nature*, **316**, 27.

Rickett, B.J., 1977, *Ann. Rev. Astron. Astrophys.*, **15**, 479.

Rickett, B.J., Coles, W.A. and Bourgois, G., 1984, *Astron. Astrophys.*, **134**, 390.

Roberts, J.A. and Ables, J.G., 1982, *Mon. Not. R. astr. Soc.*, **201**, 1119.

Romani, R.W. 1987, *Astrophys. J.*, **313**, 718.

Romani, R.W., Narayan, R. and Blandford, R. 1986, *Mon. Not. R. astr. Soc.*, **220**, 19.

Romani, R.W., Blandford, R. and Cordes, J. 1987, *Nature*, submitted.

Romani, R.W. and Taylor, J.H. 1983, *Astrophys. J. Lett.* , **265**, L35.

Rosenfeld, L. 1974. in *Astrophysics and Gravitation*, Proc. 16th Solvay Conference on Physics, Editions de l'Universite de Bruxelles, Brussels.

Sazhin, M.V. 1978, *Soviet Astronomy-AJ*, **22**, 36.

Scheuer, P.A.G. 1968. *Nature*, **218**, 920.

Segelstein, D.J., Rawley, L. Steinbring, D. Fruchter, A. and Taylor, J. 1986, *Nature*, **322**, 714.

Shapiro, S.L. and Teukolsky, S.A. 1983, *Black Holes, White Dwarfs and Neutron Stars*, Weiley, New York.

Shapirovskaya, N. Ya., 1978, *Soviet Astron.*, **22**, 544.

Shishov, V.I., 1974, *Sov. Astron.*, **17**, 598.

Sieber, W., 1982, *Astr. Astrophys.*, **113**, 311.

Simonetti, J.H., Cordes, J.M. and Heeschen, D.S. 1985. *Astrophs. J.*, **296**, 46.

Slee, O.B., Dulk, G.A. and Otrpcek, R.E. 1980. *Proc. Aust. Ast. Soc.*, 4(1), 100.

Taylor, J. 1987, in *Proc. 11th Int. Conf. on Gen. Rel. and Grav.*, Cambridge University Press, Cambridge.

Tsuruta, S. 1964, Ph.D. Thesis, Columbia University.

Tsuruta, S. 1985, Max-Planck Institut fuer Astrophysik preprint.

Vachaspati, T. and Vilenkin, A. 1985, *Phys. Rev. D*, **31**, 3052.

van der Klis, M., Jansen, F., van Paradijs, J., Lewin, W.H.G., Truemper, J. and Sztajno, M. 1985, *Nature*, **316**, 22.

Vilenkin, A. 1981, *Phys. Rev. D*, **24**, 208

Wheeler, J. 1966, *Ann. Rev. Astrn. Astrophys.*, **4**, 393.

Zel'dovich, Ya. B. 1980, *Mon. Not. R. ast. Soc.*, **192**, 663.
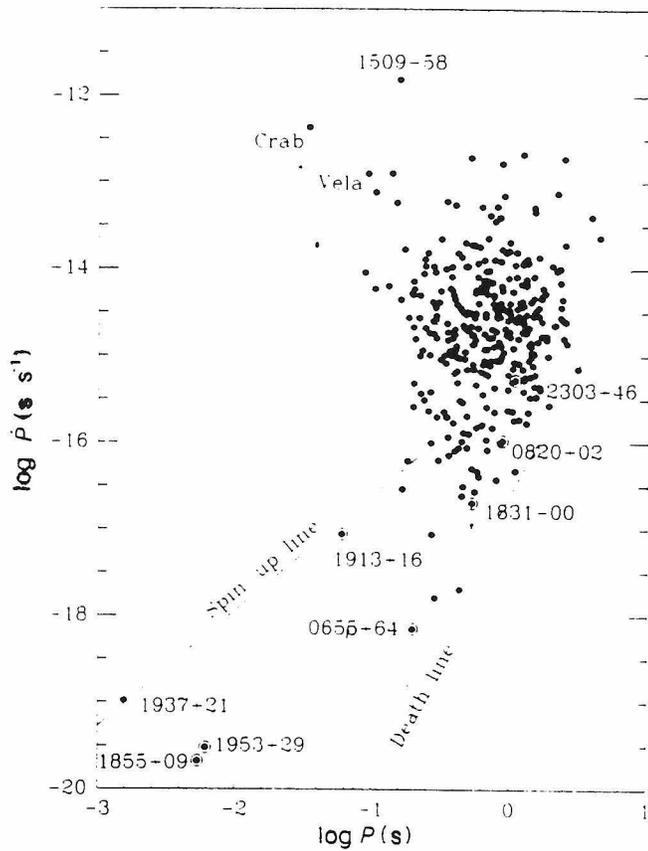
Figure 1. $P - \dot{P}$ diagram for 353 pulsars (from Dewey, *et al.* 1986). Note that young pulsars are found in the upper left, while binary (circled) pulsars and the single 1.5ms pulsar are found between the death line and the spin-up line, extending to the lower left.
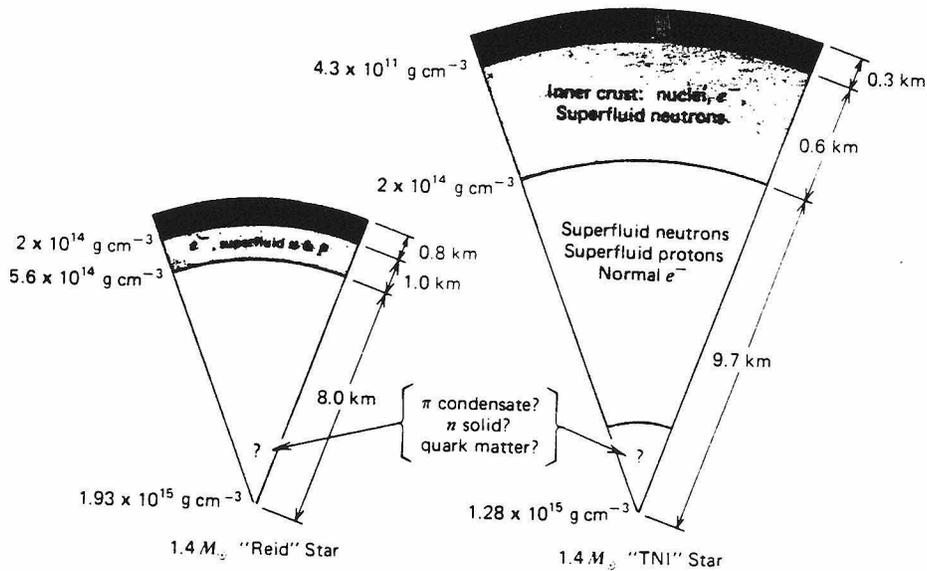
Figure 2. Schematic depiction of the internal structure of neutron stars with soft, "Reid", and relatively stiff, "TNI", equations of state (from Shapiro and Teukosky 1983).

Chapter 1:

Arrival Time Analysis for a Millisecond Pulsar

Roger D. Blandford, Ramesh Narayan, and Roger W. Romani

Theoretical Astrophysics

California Institute of Technology

Pasadena, CA 91124, U.S.A.

Arrival-Time Analysis for a Millisecond Pulsar

Roger Blandford, Ramesh Narayan* & Roger W. Romani
Theoretical Astrophysics, California Institute of Technology, Pasadena CA 91125 USA

(Invited article)

Abstract. Arrival times from a fast, quiet pulsar can be used to obtain accurate determinations of pulsar parameters. In the case of the millisecond pulsar, PSR 1937 + 214, the remarkably small rms residual to the timing fit indicates that precise measurements of position, proper motion and perhaps even trigonometric parallax will be possible (Backer 1984). The variances in these parameters, however, will depend strongly on the nature of the underlying noise spectrum. We demonstrate that for very red spectra *i.e.* those dominated by low-frequency noise, the uncertainties can be larger than the present estimates (based on a white-noise model) and can even grow with the observation period. The possibility of improved parameter estimation through 'pre-whitening' the data and the application of these results to other pulsar observations are briefly discussed. The post-fit rms residual of PSR 1937 + 214 may be used to limit the energy density of a gravitational radiation background at periods of a few months to years. However, fitting the pulsar position and pulse-emission times filters out significant amounts of residual power, especially for observation periods of less than three years. Consequently the present upper bound on the energy density of gravitational waves $\Omega_g \lesssim 3 \times 10^{-4} R_{\mu s}^2$, though already more stringent than any other available, is not as restrictive as had been previously estimated. The present limit is insufficient to exclude scenarios which use primordial cosmic strings for galaxy formation, but should improve rapidly with time.

Key words: millisecond pulsar—arrival times—gravitational background radiation

1. Introduction

The discovery of the millisecond pulsar, PSR 1937 + 214 (Backer *et al.* 1982), has opened up several new possibilities in the study of pulsar timing. The high-spin frequency (642 Hz) and the apparently small intrinsic timing noise combine to make this object an excellent clock. Arrival times have been monitored with an accuracy exceeding 1 μs over periods of two years (Backer, Kulkarni & Taylor 1983; Backer 1984; Davis *et al.* 1984) and it appears that we are already limited by the accuracy of planetary ephemerides and the stability of atomic clocks. As has been pointed out by several authors, PSR 1937 + 214 can be used as a sensitive detector of low-frequency gravitational radiation (*e.g.* Hogan & Rees 1984), as a probe of electron-density

* On leave from: Raman Research Institute, Bangalore 560080, India.

R. Blandford, R. Narayan & R. W. Romani

fluctuations in the interstellar medium (Armstrong 1984, Cordes & Stinebring 1984, Blandford & Narayan 1984a,b) and perhaps for the study of neutron-star seismology (*e.g.* Cordes & Greenstein 1981). Our purpose in the present paper is twofold. Firstly, we wish to develop the analysis of pulsar arrival times so as to estimate the sensitivity of fast pulsars as detectors of gravitational radiation and dispersion-measure fluctuations under the assumption that they remain as good clocks as is indicated by present observations. Secondly, we explore the limits to the use of accurate arrival times to measure pulsar spin-down, position, proper motion and parallax distance, in the presence of a particular noise spectrum.

In Section 2, we give a general analysis of the fitting of residuals in the measured pulse arrival times with an assumed timing model that includes the pulsar phase, period and period derivative, together with its position, proper motion and parallax. We specialize to the case of a stationary noise source and consider in Section 3 the particular case of a power-law power spectrum. We give estimates of the accuracy with which the pulsar parameters and the noise strength can be determined with standard least squares and suggest that 'pre-whitening' could lead to improvement if the noise spectrum is very 'red' (*i.e.* noise-power increasing strongly towards low frequencies). In Section 4, we apply our results to PSR 1937 + 214 and give quantitative estimates of its sensitivity to three potential sources of noise—gravitational waves, interstellar electron-density fluctuations and intrinsic pulsar noise. Applications to other pulsars are discussed in Section 5.

2. Analysis of timing residuals

Measured sequences of pulsar arrival times are conventionally fitted to a linear expression, whose parameters (essentially the corrections to various unknown quantities) are determined by the method of least squares. Unfortunately, contributions to the residuals that have quite different physical origins—for example the response to a gravitational wave of period exceeding several years and the slowing down of the pulsar's spin—can have very large covariances and are therefore not easily separated. In this section we describe a method for estimating the true sensitivity of a rapid pulsar to gravitational radiation and interstellar effects. We do this by analysing a simple timing model that includes all of the essential sources of covariance, omitting some inessential terms that would otherwise lengthen the analysis. We emphasize that the timing model has been chosen purely for analytical convenience and is not to be used in fitting real data, which should be fitted to a model based on a complete ephemeris, including general-relativistic corrections (*e.g.* Romani & Taylor 1983; Backer 1984).

In our model we assume that a point earth describes a circular orbit of known radius *a* about the solar system barycentre and so the transverse Doppler shift and gravitational redshift terms represent constant offsets (*e.g.* Manchester & Taylor 1977). This is equivalent to assuming that we possess a sufficiently accurate planetary ephemeris determined by independent means so that errors in the telescope position relative to the barycentre do not contribute to the timing noise. We discuss this approximation further in Section 4. We also assume that the pulsar position on the sky is known well enough that a linear fit to its true position, proper motion and distance is adequate.

We restrict our attention to stationary sources of noise that can be completely described by a power spectrum. In order to keep the algebra manageable, we further

idealize the observations by assuming that they are uniformly spaced and extend over an integral number of years starting at a particular epoch which we shall specify. This restriction greatly simplifies the theory and will slightly overestimate the sensitivity of the timing data if our results are applied to non-uniform observations taken over a non-integral number of years.

For a pulsar with parallax $p = a/d$ (with d the pulsar's distance from the barycentre), whose heliocentric latitude and longitude measured from the vernal equinox are respectively β and λ, the distance a pulse travels to earth is given by

$$D = [d^2 \sin^2 \beta + \{d \cos \beta + a \cos (\phi - \lambda)\}^2 + a^2 \sin^2 (\phi - \lambda)]^{1/2}$$

$$= a\left[\cos \beta \cos (\phi - \lambda) - \frac{1}{4} p \cos^2 \beta \cos 2 (\phi - \lambda)\right], \tag{2.1}$$

where ϕ is the earth's mean anomaly, and we have dropped some constant terms. Let the small errors in the pulsar latitude and longitude be

$$\delta\beta = \delta\beta_0 + \delta\mu_\beta t \tag{2.2}$$

$$\delta\lambda = \delta\lambda_0 + \delta\mu_\lambda t \tag{2.3}$$

where $\delta\mu_{\beta,\lambda}$ are the two components of the proper motion and t is the time of observation, which we measure in years from the mid-point of the observation, fixed to occur at an anomaly $\phi = \lambda + \pi/4$.

As usual, we fit the time of emission of the pulses to a quadratic function parametrized by the unknown phase, frequency and frequency derivative. Ignoring constant additive and multiplicative factors, the pulse arrival time is given by the emission time plus the variable part of the propagation time to earth, D/c. We define the timing residual $R(t)$ to be the difference between the observed arrival time of a pulse and the arrival time predicted on the basis of our best guesses to the unknown parameters. These residuals are fitted to an expression that is linear in the corrections to the unknown parameters, *i.e.*,

$$R(t) = \sum_{a=1}^{8} \alpha_a \psi_a(t) \tag{2.4}$$

where

$$\alpha_1 = -K, \qquad \psi_1 = 1,$$

$$\alpha_2 = -\frac{\delta\nu}{\nu}, \qquad \psi_2 = t,$$

$$\alpha_3 = -\frac{\delta\dot{\nu}}{2\nu}, \qquad \psi_3 = t^2,$$

$$\alpha_4 = \frac{a}{c\sqrt{2}} (\delta\beta_0 \sin \beta + \delta\lambda_0 \cos \beta), \qquad \psi_4 = \sin 2\pi t,$$

$$\alpha_5 = \frac{a}{c\sqrt{2}} (-\delta\beta_0 \sin \beta + \delta\lambda_0 \cos \beta), \qquad \psi_5 = \cos 2\pi t,$$

$$\alpha_6 = \frac{a}{c\sqrt{2}} (-\delta\mu_\beta \sin \beta + \delta\mu_\lambda \cos \beta), \qquad \psi_6 = t \cos 2\pi t,$$

$$\alpha_7 = \frac{a}{c\sqrt{2}} (\delta\mu_\beta \sin \beta + \delta\mu_\lambda \cos \beta), \qquad \psi_7 = t \sin 2\pi t,$$

$$\alpha_8 = \frac{a}{4c} p \cos^2 \beta, \qquad \psi_8 = \sin 4\pi t. \tag{2.5}$$

R. Blandford, R. Narayan & R. W. Romani

Timing parallax has not so far been measured in any pulsar. Therefore, we have repeated our calculations for a linear combination of 7 parameters, leaving out ψ_8.

Now suppose that we measure n equally spaced and comparably accurate arrival times each year for a total of N years, *i.e.*, we have Nn residuals $R_i = R(t_i)$, $i = 1, Nn$. We wish to obtain least squares estimates of the parameters α_a. As there are 8 independent parameters to fit, it turns out to be algebraically easier to diagonalize the normal equations by introducing a set of orthonormal fitting functions, $\psi'_{ai} = \psi'_a(t_i)$, which are linear combinations of the original ψ_i; *i.e.*,

$$R(t) = \sum_{a=1}^{8} \alpha'_a \psi'_a(t), \qquad \psi'_{ai} = L_{ab}\psi_{bi}, \qquad (2.6)$$

where

$$\sum_{i=1}^{Nn} \psi'_{ai}\psi'_{bi} = \delta_{ab}. \qquad (2.7)$$

In fact, the number of observations is usually so large that the sum in Equation (2.7) can be approximated by an integral over the observing period; *i.e.*, $\sum_{i=1}^{Nn} \sim n \int_{-N/2}^{N/2} dt$. A convenient choice of orthonormal functions for the case in point is defined uniquely through the Gram-Schmidt orthogonalization procedure:

$$\psi'_1 = L_{11},$$
$$\psi'_2 = L_{22}t,$$
$$\psi'_3 = L_{31} + L_{33}t^2,$$
$$\psi'_4 = L_{42}t + L_{44}\sin 2\pi t,$$
$$\psi'_5 = L_{51} + L_{53}t^2 + L_{55}\cos 2\pi t,$$
$$\psi'_6 = L_{62}t + L_{64}\sin 2\pi t + L_{66}t\cos 2\pi t,$$
$$\psi'_7 = L_{71} + L_{73}t^2 + L_{75}\cos 2\pi t + L_{77}t\sin 2\pi t,$$
$$\psi'_8 = L_{82}t + L_{84}\sin 2\pi t + L_{86}\cos 2\pi t + L_{88}\sin 4\pi t. \qquad (2.8)$$

where the L_{ab} are constants that depend upon N. The best-fitting primed parameters, α'_a, are given by the solution of the normal equations

$$\alpha'_a = \sum_{i=1}^{Nn} R_i\psi'_{ai}. \qquad (2.9)$$

Now suppose that the residuals are entirely due to timing noise generated by a stationary power spectrum $P(f)$ so that

$$\langle R_i R_j \rangle = \int_0^{\infty} df P(f)\cos 2\pi f(t_i - t_j) \qquad (2.10)$$

where $\langle \ \rangle$ signifies an ensemble average over many realizations of the fitting procedure. We obtain an expression for the mean-square residual after subtracting the best-fitting solution to Equation (2.6)

$$\overline{R^2(t)} = \frac{1}{Nn} \sum_i \sum_j \langle R_i R_j \rangle [\delta_{ij} - \sum_a \psi'_{ai}\psi'_{aj}]$$
$$= \int_0^{\infty} df P(f)T(f) \qquad (2.11)$$

Arrival-time analysis for a millisecond pulsar

where the transmission or filter function, $T(f)$, is given by

$$T(f) = 1 - \frac{1}{N} \sum_a \tilde{\psi}'_a(f)\tilde{\psi}'^*_a(f), \qquad (2.12)$$

and

$$\tilde{\psi}'_a(f) = \int_{-N/2}^{N/2} dt\, \psi'_a(t) \exp(2\pi i f t) \qquad (2.13)$$

are the Fourier transforms of the orthonormal fitting functions.

Equation (2.11) is an expression of the fact that when we try to detect background timing noise, much of this noise will be filtered out by the fit for the pulsar period, position and other parameters. We can think of the factor $T(f)$ as being a transmission coefficient for the noise and the individual factors $|\tilde{\psi}'_a(f)|^2$ as being absorption coefficients associated with the individual fitting functions. The latter are presented for $N = 3$ in Figs 1 and 2 and the transmission function $T(f)$ is presented for $N = 1, 3, 10$ in Fig. 3. The pulsar will thus be a less sensitive detector of the noise than if we had prior knowledge of the exact phase, period, position, *etc.* (in which case the filter function is $T(f) = 1$).
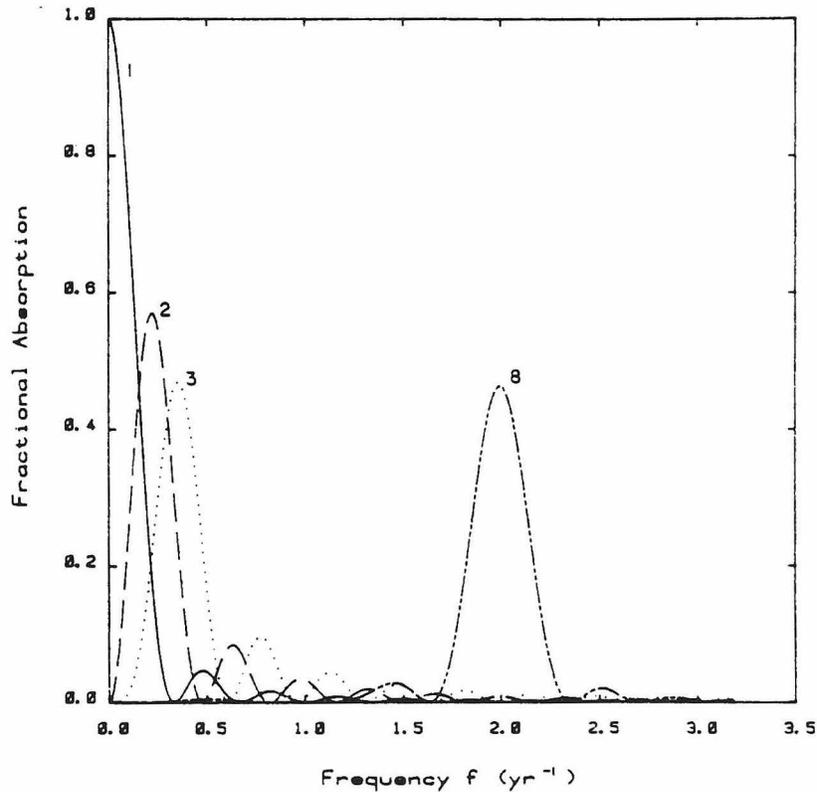
Figure 1. Absorption coefficients $|\tilde{\psi}'_a|^2$ for $a = 1, 2, 3$ and 8 at $N = 3$ years. The first three functions generate the dip near the origin in Fig. 3, and the last function generates the feature at $f = 2\,\mathrm{yr}^{-1}$.

23

R. Blandford, R. Narayan & R. W. Romani
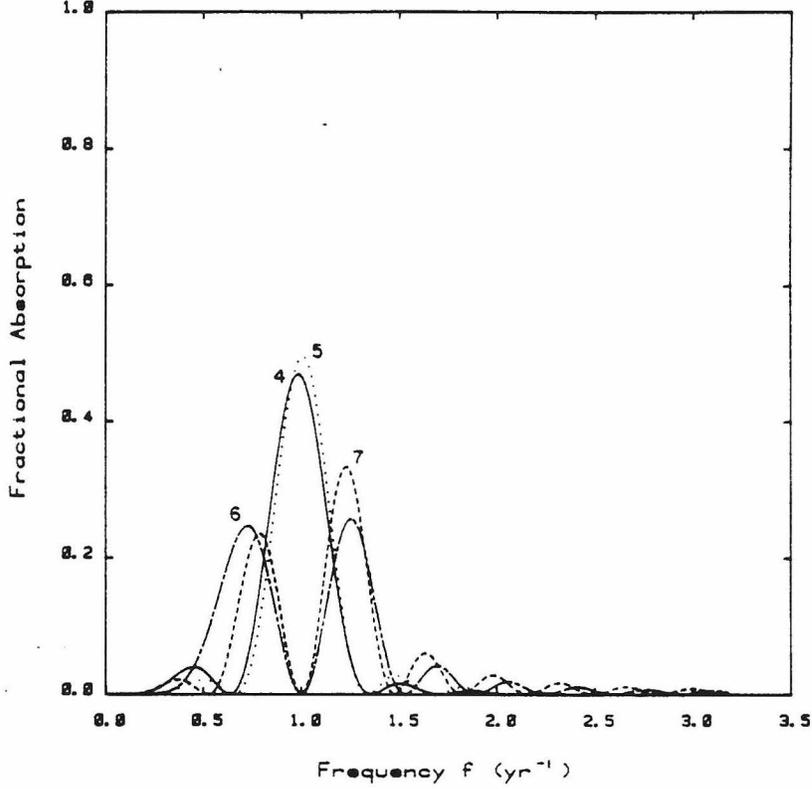
Figure 2. Absorption coefficients $|\tilde{\psi}'_a|^2$ for $a = 4$–7 at $N = 3$ years. The functions 4 and 5 are largely due to position errors while 6 and 7 are dominated by the proper motion terms. These generate the minimum at $f = 1 \text{ yr}^{-1}$ in Fig. 3.

We can also use Equation (2.6) to estimate the covariance matrix of the parameters α'_a after performing a least-squares fit to the measured arrival times

$$\langle \delta\alpha'_a \delta\alpha'_b \rangle = \int_0^\infty df\, P(f)\tilde{\psi}'_a(f)\tilde{\psi}'^*_b(f) \qquad (2.14a)$$

or

$$\langle \delta\alpha'_a \delta\alpha'_b \rangle = \left[\frac{\int_0^\infty df\, P(f)\tilde{\psi}'_a(f)\tilde{\psi}'^*_b(f)}{\int_0^\infty df\, P(f)T(f)} \right] \bar{R}^2. \qquad (2.14b)$$

Note that the quantity in square brackets is independent of the strength of the noise and depends only on the shape of its spectrum. Finally, the covariance matrix of the original fitting parameters is given by

$$\langle \delta\alpha_a \delta\alpha_b \rangle = \sum_a \sum_b L_{ca} L_{db} \langle \delta\alpha'_c \delta\alpha'_d \rangle. \qquad (2.15)$$

Equations (2.14) and (2.15) allow us to make an unbiased estimate of the expected error in the various fitting parameters in terms of either the noise strength or the residual. However, as we discuss further in Section 3.4 below, we may be able to filter out much of the noise so as to obtain a much smaller variance for the unknown parameters. The

Arrival-time analysis for a millisecond pulsar
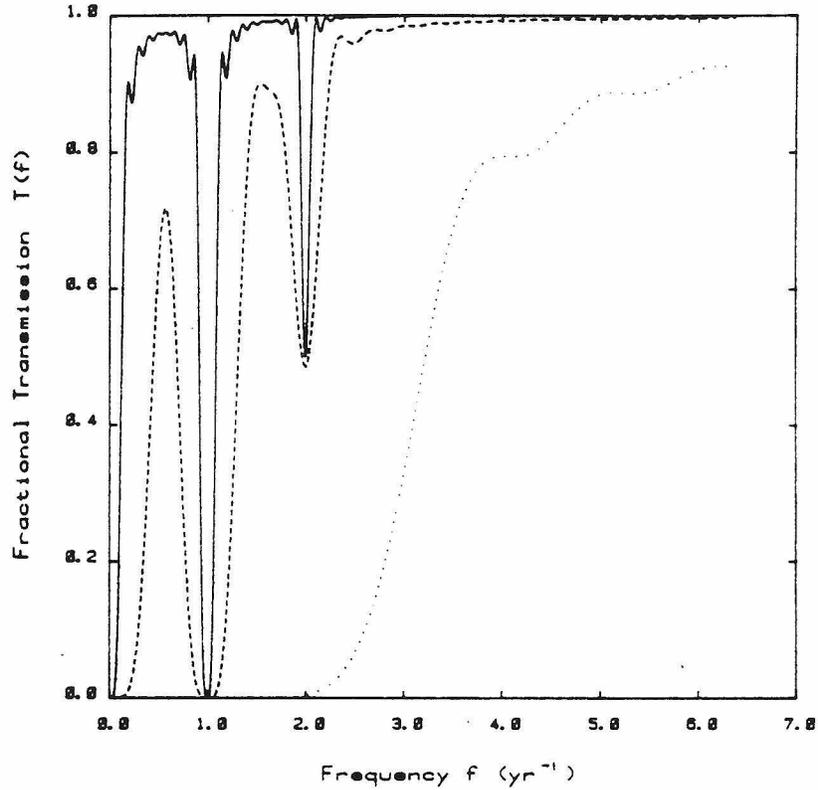
Figure 3. Transmission coefficient $T(f)$ defined in Equation (2.12) for an 8-parameter fit for $N = 1$ year (dotted line), 3 years (dashed line), 10 years (solid line). The dip near the origin corresponds to power removed by the polynomial fit, the dip at $1\ yr^{-1}$ is from fitting position and proper motion and that at $2\ yr^{-1}$ is due to parallax. As N increases, the three features become narrower (width $\propto 1/N$) showing that the corresponding sets of functions become more nearly orthogonal to one another.

usual variance estimated by standard least squares corresponds to the case of white noise, *i.e.*, $P(f) = $ constant.

3. Power-law noise spectra

3.1 *General Considerations*

We now assume that the noise spectrum has a power law form

$$P(f) = P_0 f^{-s}, \qquad f \geqslant 0. \tag{3.1}$$

P_0 is the noise power in waves with a period of around one year. We confine our attention to the exponents $s = 0, 2, 3, 4, 5, 6$ and data spans of $N = 1, 2, 3, 5, 10\,yr$. The exponent $s = 0$ corresponds to white noise, which is the spectrum usually assumed (at least implicitly) when analysing the arrival times by least-squares fitting. It is appropriate when individual independent measurement errors dominate other sources

R. Blandford, R. Narayan & R. W. Romani

of noise. 'Red' spectra with slopes $s = 2, 4, 6$ correspond to random walks in phase, frequency and torque respectively. Spectra with slopes $s = 3, 5$ may be produced, respectively, by interstellar density fluctuations and a hypothetical background of primordial gravitational waves.

Our procedure is to compute the elements L_{ab} of the transformation matrix defined by (2.6) for each value of N and then to calculate the Fourier transforms of the orthonormal functions, $\tilde{\psi}'_a(f)$, by taking suitable linear combinations of the analytical Fourier transforms of the $\psi_a(t)$. Next, we evaluate the filter function $T(f)$ (Equation 2.12), and then compute the mean expected residual through Equation (2.11). In order to make contact with earlier work we express our results in terms of an equivalent filter which is 0 for $f < \beta/N$ and 1 for $f > \beta/N$. In other words, we determine β so that the calculated mean square residual $\bar{R}^2$ satisfies the relation

$$\bar{R}^2 = \int_0^\infty df P(f) T(f) = \int_{\beta/N}^{n/2} df P(f) \tag{3.2}$$

The upper cut-off in the frequency arises from the sampling theorem and is not important for red noise. The lower cut-off takes account of the fact that lower frequencies are fitted away by the polynomial fit and periods around 1 yr and 6 months are fitted by position/proper motion and parallax respectively. In the past β has been assumed to be ~ 1 (Detweiler 1979; Bertotti, Carr & Rees 1983; Romani & Taylor 1983), but no quantitative estimates have been reported to date.

We also compute the uncertainties in the various parameters α_1–α_8 and present each as the ratio, (variance)$^{1/2}$ per μs of post-fit rms residual. These can be converted to variance per unit power at 1 yr period, P_0, through Equations (3.1) and (3.2).

3.2 White Noise

To bring out the salient features of our formalism we first consider white noise, corresponding to $s = 0$. Calculations show that, for white noise with $n \gg 1$, $\beta = 4$ when all 8 parameters are fitted and $\beta = 3.5$ when parallax is not refined.

Consider next the variance in the position estimate of the pulsar. We can make the following approximate estimate. If N is sufficiently large, $\psi_4(t)$ and $\psi_5(t)$ are almost orthogonal to the other $\psi_i(t)$. Then, the variance v_4 in the estimate of α_4 is approximately given by simplifying Equation (2.14a) to

$$v_4 = \langle \delta\alpha_4 \delta\alpha_4 \rangle = \frac{\int_0^{n/2} P(f) |\tilde{\psi}_4(f)|^2 df}{\left[2\int_0^{n/2} |\tilde{\psi}_4(f)|^2 df \right]^2} . \tag{3.3}$$

The denominator is necessary because $\psi_4(t)$ is not normalized and the factor of 2 is because the integral has been restricted to positive f. There is a similar expression for v_5. Taking $P(f) = P_0$ for white noise and substituting

$$\tilde{\psi}_4(f) = \frac{iN}{2}\left[\frac{\sin \pi N(f-1)}{\pi N(f-1)} - \frac{\sin \pi N(f+1)}{\pi N(f+1)} \right], \tag{3.4}$$

$$\tilde{\psi}_5(f) = \frac{N}{2}\left[\frac{\sin \pi N(f-1)}{\pi N(f-1)} + \frac{\sin \pi N(f+1)}{\pi N(f+1)} \right], \tag{3.5}$$

Arrival-time analysis for a millisecond pulsar

we obtain

$$v_4 \sim v_5 \sim \frac{2\bar{R}^2}{nN}. \tag{3.6}$$

We thus recover the well-known result that the variance decreases inversely with the number of independent measurements. Substituting $a = 1.5 \times 10^{13}$ cm in (2.5) we thus have

$$\sin \beta \, (\delta\beta_0)_{\rm rms} \sim \cos \beta \, (\delta\lambda_0)_{\rm rms} \sim \frac{0.6 \, (\bar{R}^2)^{1/2}_{\mu s}}{(nN)^{1/2}} \, {\rm mas}. \tag{3.7}$$

More detailed calculations through the Gram-Schmidt orthogonalization procedure described in Section 2 confirm the coefficient as well as the scaling with n and N. The rms error in the proper motion is given by

$$\sin \beta \, (\delta\mu_\beta)_{\rm rms} \sim \cos \beta \, (\delta\mu_\lambda)_{\rm rms} \sim \frac{2 \, (\bar{R}^2)^{1/2}_{\mu s}}{(nN^3)^{1/2}} \, {\rm mas \ yr}^{-1}. \tag{3.8}$$

3.3 Red Noise

Red noise spectra have $s > 0$, *i.e.* the residuals are dominated by low-frequency noise. In the cases of interest, all the integrals converge rapidly at high f and so none of the results are sensitive to n so long as $n \gtrsim 10$. This is an important qualitative feature of red noise, showing that one cannot improve the precision of the refined parameters by increasing the number of measurements. As we demonstrate below, one does not gain by increasing the number of years of data either since the variances often increase as N increases.

Red noise has a divergent spectrum at low f. However, since the filter function $T(f)$ $\propto f^6$ at low f (for the present problem), the post-fit mean-square residual $\bar{R}^2$ converges so long as $s < 7$. Equation (3.2) can now be written in the form

$$\bar{R}^2 = \frac{P_0}{(s-1)} \left(\frac{N}{\beta}\right)^{(s-1)} \tag{3.9}$$

where the upper limit in the integral should ideally be $n/2$ but has been set to ∞ (continuous sampling) because the integral converges rapidly. β has been evaluated for various values of N and s; the results are presented in Table 1. We give β for a 7-parameter fit (without parallax) for $N = 1, 2, 3, 5$ and also for an 8-parameter fit for $N = 5, 10$. Note that β is large, $\gtrsim 2$ for $N < 3$, showing that the parameter fit removes a substantial part of the noise. Our values of β are somewhat larger than those assumed by Bertotti, Carr & Rees (1983) and Hogan & Rees (1984).

Press (1975) and Lamb & Lamb (1976) have developed a least-squares analysis of pulsar timing noise in terms of a complete set of orthogonal polynomials, but considered only a white-noise spectrum. Our approach, which involves an ortho-gonalization of the functions relevant to physical parameters, can be extended to accommodate red-noise processes. Groth (1975a) and Cordes (1980) have analysed red-noise spectra as well, but employ a model in the time domain. This time series approach, in principle, has more information than is contained in the power spectrum alone; we

R. Blandford, R. Narayan & R. W. Romani

Table 1. Values of the effective spectral cutoff β (*cf.* Equation 3.2) corresponding to a 7-parameter fit (no parallax) for $N = 1, 2, 3, 5$ and an 8-parameter fit (including parallax) for $N = 5, 10$.

N/s	2	3	4	5	6
7-parameter fit:					
1	2.91	2.80	2.73	2.67	2.59
2	3.12	2.95	2.61	2.05	1.53
3	1.75	1.34	1.15	1.01	0.88
5	1.35	1.15	1.05	0.95	0.84
8-parameter fit:					
5	1.36	1.16	1.05	0.96	0.84
10	1.22	1.11	1.03	0.94	0.83

make a comparison between the time domain and power-spectrum methods in the Appendix.

Equation (3.9) shows that the post-fit residuals grow rapidly as data are collected over longer spans of time. Physically, large-amplitude low-frequency noise becomes increasingly important over longer data spans. The rate of growth of $\bar{R}^2$ with N can be used to estimate the spectral index s, as Groth (1975b) and Cordes (1980) have emphasized. Deeter & Boynton (1982) and Deeter (1984) describe another interesting technique (based on a formalism that has some similarity to our methods) for estimating the shape of the noise spectrum. Their analysis treated finite samples of unevenly spaced data, but considered only even integral values of s, and did not include the refinement of intrinsic pulsar parameters. Odd s can, however, be of physical interest. In principle, since $T(f)$ is known, it should always be possible to recover $P(f)$ from the power spectrum of the residuals. With the complexities of a finite time series of data, however, a discrete method such as that developed by Deeter (1984) may be more accurate.

As can be seen from Equation (2.14), the variances of the parameters involve integrals over the power spectrum $P(f)$ weighted by the appropriate absorption coefficient. All the integrals converge in the limit as $f \to \infty$ but their properties vary as $f \to 0$. It can be shown that the weighting functions vary as f^0, f^2 and f^4 for α_1, α_2 and α_3 and as f^6 for the rest of the parameters. Consequently, depending on the value of s, one or more of the parameters could have a divergent variance. Physically, this means that the error in the estimated parameter is dominated by noise of very long period and so the variance is essentially determined by the lower cut-off in the spectrum. Uncertainties in α_1 and α_2 are of no consequence. The variance in α_3, however, is of interest. Results are given in Table 2 for various values of N and s. For $s = 5, 6$, the answer depends on $N_{\max}$, the longest-period wave present in the spectrum. If the source of noise is gravitational radiation, $N_{\max}$ is the light travel time to the pulsar (since beyond $N_{\max}$ the effective spectral slope reduces by 2 and so the integral converges), while if it is intrinsic pulsar noise (say a random walk in the rate of spin-down), $N_{\max}$ will probably be of the order of the characteristic spin-down age of the pulsar, $P/2\dot{P}$.

The uncertainty in $\dot{P}$ also affects the accuracy with which $\ddot{P}$ can be measured. The error in $\ddot{P}$ is approximately the difference between the errors in $\dot{P}$ at the beginning and end of the observations, divided by N years. Clearly, errors in $\dot{P}$ caused by very-long-

Arrival-time analysis for a millisecond pulsar

Table 2. · Root-mean-square error $\delta\dot{P}/P$ per µs post-fit residual R_{rms} in units of $10^{-20}\,\text{s}^{-1}$. For $s = 5$ and 6 the results depend on the cut-off frequency $f_{\text{min}} = 1/N_{\text{max}}$ and hence two numbers, A and N^*, are given. For $s = 5$, $\delta\dot{P}/P = A[\ln(N_{\text{max}}/N^*)]^{1/2}$ and for $s = 6$, $\delta\dot{P}/P = A(N_{\text{max}} - N^*)^{1/2}$.

$N \diagdown s$	2	3	4	5	6
7-parameter fit:					
1	46.6	67.6	86.7	43.9, 0.010	74.4, −0.231
2	0.792	2.20	5.78	7.79, 2.83	4.14, 3.07
3	0.293	0.492	0.793	0.841, 3.94	0.389, 4.28
5	0.088	0.148	0.247	0.269, 6.67	0.096, 7.25
8-parameter fit:					
5	0.089	0.148	0.248	0.269, 6.67	0.096, 7.25
10	0.021	0.035	0.059	0.065, 13.6	0.016, 14.7

period waves are not relevant since they coherently affect $\dot{P}$ over the whole range of observations. Therefore, for this calculation, we have used the rms error in $\dot{P}$ contributed by waves with periods less than πN. We then find that the rms error in the braking index, $n_b = P\ddot{P}/\dot{P}^2$, contributed by a red noise process is

$$\delta n_b \sim 8 \left(\frac{\tau}{10^7\,\text{yr}}\right)^2 \left(\frac{N}{3}\right)^{(s-7)/2} \left(\frac{R_{\text{rms}}}{1\,\mu\text{s}}\right) \tag{3.11}$$

where τ is the pulsar timing age $P/2\dot{P}$ and s is the index of the noise spectrum. This error is to be compared with $n_b = 3$ predicted by magnetic dipole braking.

Fig. 4 shows the rms uncertainties per µs post-fit residual of pulsar position, proper motion and parallax for $s = 4$ and various values of N. The results are relatively insensitive to s, particularly at large N. This can be understood on the basis of approximate analytical estimates of the variances similar to those made in Section 3.1. Noting that for large N and sufficiently steep spectra the respective variances are dominated by the integrals near $f \sim 1/N$ (below which the integrands fall off as f^{6-s}), it can be shown that the position and parallax variances $\propto \bar{R}^2/N^2$ and the proper motion variances $\propto \bar{R}^2/N^4$, with no dependence on s. These scalings are consistent with the more accurate calculations of Fig. 2. Combining with Equation (3.9), the surprising result is that for a given power spectrum, the position and parallax variances $\propto N^{s-3}$ and the proper motion variances $\propto N^{s-5}$, *i.e.*, for a sufficiently steep spectrum *the variance increases with increasing N*. This is quite contrary to the normal wisdom on parameter uncertainties in least squares, which is based on white noise. A comparison of the above scaling laws with those presented in Equations (3.7) and (3.8) shows that the true variance in the presence of red noise can be significantly greater than that estimated on the basis of standard least squares whenever $n \gg N$.

3.4 *Variance Reduction*

We now discuss how prior knowledge of the spectrum can, in principle, be used to reduce the variances in the estimated parameters. For simplicity consider a model consisting of only one parameter, *i.e.*

$$R(t) = \alpha\psi(t). \tag{3.12}$$

R. Blandford, R. Narayan & R. W. Romani
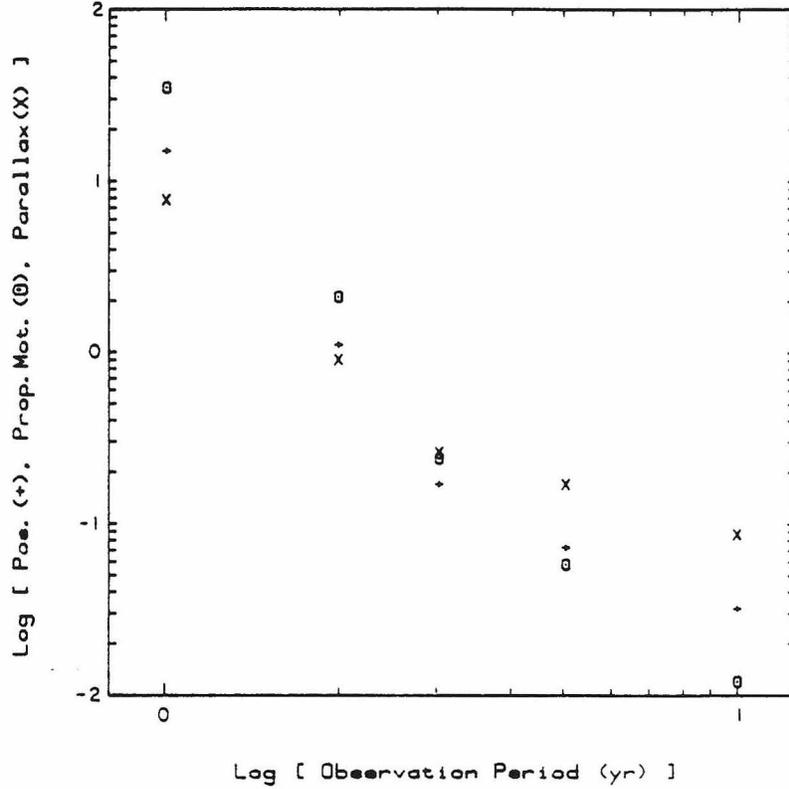
Figure 4. Root-mean-square error in pulsar parameters per μs post-fit residual R_{rms} as a function of the number of years of observation. The results are for $s = 4$, but do not vary a great deal for other values of s. The symbol + shows position errors, $\sin\beta(\delta\beta_0)_{rms}$ and $\cos\beta(\delta\lambda_0)_{rms}$, in mas (milli-arcsec). For large N the error scales as $1/N$. The symbol $\odot$ shows proper-motion errors, $\sin\beta(\delta\mu_\beta)_{rms}$ and $\cos\beta(\delta\mu_\lambda)_{rms}$, in mas yr^{-1}; scaling as $1/N^2$. The symbol $\times$ shows $(\sin^2\beta/d_{kpc}) \times$ parallax error $\delta d/d$; scaling as $1/N$.

As before we take $\tilde{\psi}(f)$ to be the Fourier transform of $\psi(t)$. Now let us suppose that we convolve the measured residuals $R(t_i)$ with an arbitrary function $K(t)$, which is equivalent to multiplying $P(f)$ by $|\tilde{K}(f)|^2$. Correspondingly, the new model that is to be fitted is $\tilde{\psi}(f)\tilde{K}(f)$. Proceeding as in Section 2, the variance of x is given by

$$v = \frac{\int_0^\infty P(f)|\tilde{\psi}(f)|^2|\tilde{K}(f)|^4\,df}{\left[2\int_0^\infty |\tilde{\psi}(f)|^2|\tilde{K}(f)|^2\,df\right]^2}. \tag{3.13}$$

We now optimize v with respect to the function $|\tilde{K}(f)|^2$. This gives

$$|\tilde{K}(f)|^2 = K_0/P(f) \tag{3.14}$$

where K_0 is an arbitrary constant. Thus the uncertainty in the parameter is minimum when the noise is 'pre-whitened' before the least squares is performed, with the fitting model being suitably modified.

When there are several parameters the analysis becomes a little more complicated because the variances in (2.14) depend on the orthogonal functions $\tilde{\psi}'_a(f)$ which change

Arrival-time analysis for a millisecond pulsar

as $\tilde{K}(f)$ is varied. However, a proof can be devised, based on a variational technique where one constantly rotates into a local orthogonal set of functions, to show that (3.14) continues to be optimal even for this case.

Simple estimates indicate that the 'pre-whitened' variances in pulsar position and parallax will be $\propto \bar{R}^2/n^{s-1} N^{s-1}$ while the variances in proper motion will be $\propto \bar{R}^2/n^{s-1} N^{s+1}$. The coefficients in these relations, however, are quite large and therefore significant gains are probably possible only for large s, n and N. A practical matter is that at high frequencies measurement errors, which behave like white noise, will dominate. Hence the appropriate n to use in the above estimates is not the actual sampling rate but some $n' < n$ where the spectrum changes from red to white noise. We are currently exploring the practicality of implementing this pre-whitening procedure.

4. Application to PSR 1937 + 214

4.1 *General Considerations*

For the particular case of PSR 1937 + 214, $\nu = 642$ Hz and $\dot{\nu} = -4.3 \times 10^{-14}$ Hz s^{-1} (Backer 1984). If we assume that the braking index is 3, then $\ddot{\nu} = 8.6 \times 10^{-30}$ Hz s^{-2}. If we were to include a cubic term in the fitting formula, then the contribution to the residual would be $7 \times 10^{-5} N^3$ μs. This may possibly be detectable after ~ 10 yr but will be significantly harder to measure than the parallax term. We have therefore omitted it from the fitting formula.

The heliocentric latitude and longitude of the pulsar are respectively $\beta = 42.3°$ and $\lambda = 301.3°$. The distance, determined from hydrogen absorption measurements (Heiles *et al.* 1983) is $d \sim 5$ kpc which is consistent with the dispersion measure of $DM = 71$ cm^{-3} pc. Scintillation studies suggest that the speed of the pulsar transverse to the line of sight is ~ 80 km s^{-1} (J. M. Cordes, personal communication) which translates into a proper motion of ~ 3.4 mas yr^{-1}. However, the pulsar is unusually close to the galactic plane for its apparent age and so we expect that the velocity lies within the plane. The parameters α_4–α_8 are expected to have the following magnitudes

$$\alpha_4 \sim \alpha_5 \sim 1.7 \left[\frac{\delta\beta_0, \delta\lambda_0}{1 \text{ mas}} \right] \mu\text{s}, \tag{4.1}$$

$$\alpha_6 \sim \alpha_7 \sim 4.2 \left[\frac{\mu_\beta, \mu_\lambda}{3.4 \text{ mas yr}^{-1}} \right] \mu\text{s}, \tag{4.2}$$

$$\alpha_8 \sim 0.066 \ \mu\text{s}. \tag{4.3}$$

It is clear that the signal given by Equation (4.3) will be very hard to measure; for this reason we have not included parallax within the fitting formula for observing periods $N < 5$. In fact, from the results of Fig. 2, we see that a ~ 30 per cent measurement of the parallax will require that the rms residual over 5 years from red noise should be less than 0.2 μs. Unfortunately, however, dispersion measure fluctuations alone introduce a residual of $\sim 2(N/10)^{1/2} \mu$s (*c.f.* Section 4.3).

We should also consider the accuracy of solar system ephemerides over ~ 10 yr timescales. The internal agreement over periods of ~ 10 yr for the best ephemerides is about 3000 metres, *i.e.* 10 μs in arrival time. There is some prospect that improvements

R. Blandford, R. Narayan & R. W. Romani

in our knowledge of the position of the telescope relative to the solar system barycentre, which must be known to better than 10 m to exploit the timing fully, will occur over the same period, especially if plans to land a ranger on Phobos in the early 1990's are realized (R. Hellings, personal communication). A related requirement is that local time as measured by atomic clocks be able to avoid drifts in excess of a few μs over ten year periods. Trapped ion clocks may achieve the necessary stability. Of course, the discovery of another quiet millisecond pulsar (or preferably several others) would allow the separation of intrinsic pulsar noise and ephemeris errors to a large extent.

4.2 *Gravitational Radiation*

Several authors (*e.g.* Detweiler 1979; Mashhoon 1982; Bertotti, Carr & Rees 1983) have suggested that an upper bound can be placed on the energy density of primordial gravitational radiation with periods ~ 1 yr using the pulsar timing residuals. In particular, a substantial energy density in gravitational radiation may be produced by primordial cosmic strings and indeed pulsar timing is probably the best way to set limits on the density of these strings (*e.g.* Hogan & Rees 1984). If the energy density in the gravitational radiation between frequency f and $f + df$ is $\rho_g(f)$ then the expected power spectrum for the timing noise is

$$P(f) = \frac{G\rho_g(f)}{3\pi^3 f^4},$$ (4.4)

i.e. $P_0 = 1.3 \times 10^4 \, \Omega_g(f)\mu s^2$ where $\Omega_g(f) = [8\pi G\rho_g(f)f]/(3H_0^2)$ is the ratio of the wave energy density per natural-logarithm frequency interval at frequency f to the critical cosmological density (setting the Hubble constant $H_0 = 100 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$).

If a fixed fraction of the energy within a horizon during the radiation-dominated era is channelled in some self-similar way into gravitational radiation of comoving wavelength equal to a fraction of the horizon size, then we expect Ω_g to be constant, *i.e.* $P(f) \propto f^{-5}$. Under other circumstances, as discussed by Vilenkin (1981) and Hogan & Rees (1984), structure may be imprinted on the spectrum at the epoch when the universe becomes matter-dominated. Spectral slopes of 5.5 and 7 in the frequency range $0.1 \gtrsim f \gtrsim 10^{-4}$ have also been proposed. Existing observations of the millisecond pulsar can only place a rather modest limit on the energy density of gravitational radiation at frequencies on the order of a few cycles per year. Setting $N = 2$, we see that

$$\Omega_g < 3.4 \times 10^{-4} \, R_{\mu s}^2.$$ (4.5)

The difference between this estimate and that given by Hogan & Rees (1984) is due mainly to their assumed value of β. After observations have been carried out for more than 5 years, however, a limit

$$\Omega_g < 4.0 \times 10^{-7} \, R_{\mu s}^2$$ (4.6)

may be set, which would certainly be more interesting. For instance, cosmological models in which primordial strings are created during the earliest epochs of the expanding universe and re-enter the horizon during the radiation era require the string parameter ε to be $\gtrsim 10^{-6}$ if the strings are to have a significant effect on galaxy formation. Since $(\varepsilon/10^{-6}) \sim (\Omega_g/2 \times 10^{-7})^2$ (Hogan & Rees 1984), 5 years of sub-μs residuals on PSR 1937 + 214 would be sufficient to exclude such scenarios.

Arrival-time analysis for a millisecond pulsar

4.3 *Interstellar Density Fluctuations*

Arrival-time fluctuations can also be caused by a variable dispersion measure along the line of sight to the pulsar (Armstrong 1984; Blandford & Narayan 1984a,b). Essentially what happens is that as the observations proceed, larger and larger interstellar clouds can cross the line of sight, causing progressively greater changes in the dispersion measure. The importance of this effect depends upon the spectrum of interstellar density fluctuations in the length-scale range 10^4–10^{16} cm. It has been argued that the spectrum of density fluctuations has a power law form,

$$\Phi_k = C_N^2 k^{-\eta} \tag{4.7}$$

where Φ_k is the three-dimensional power spectrum of the density fluctuations at spatial frequency k. The exponent η has been estimated to be close to the 'Kolmogorov' value of 11/3 (*e.g.* Armstrong, Cordes & Rickett 1981) although there are some indications that it may be somewhat larger (Blandford & Narayan 1984b). Here we adopt a value $\eta = 4$, *i.e.*, $s = 3$. For PSR 1937 + 214 we take C_N^2 to be 10^{-4}, compatible with the measured decorrelation bandwidth (Cordes & Stinebring 1984), together with a measured speed of the scintillation pattern relative to earth of 80 km s^{-1} (J. M. Cordes, personal communication). At an observing wavelength of 1400 MHz, we then find that

$$P(f) = 0.3 f^{-3} \, \mu s^2 \tag{4.8}$$

(*cf.* Armstrong 1984). If most of the measurement error is removed, leaving (4.8) as the dominant noise component in the spectrum, then after three years the timing position can be determined with an uncertainty of ~ 0.23 mas, and the proper motion can be measured to an accuracy of ~ 0.33 mas yr^{-1}. The scaling laws of Section 3.3 indicate that these uncertainties will remain constant for the first parameter and scale as $1/N$ for the second. The uncertainty in the braking index, $\delta \eta_b$, induced by DM fluctuations will be $\sim 2 \times 10^4/N$ (for $N \gtrsim 3$). After three years, the fractional uncertainty in the parallax distance, $\delta d/d$, will be ~ 2.6, and will not improve with time. Therefore, unless dispersion measure fluctuations are monitored and corrected for, parallax distance cannot be determined.

4.4 *Intrinsic Noise*

It has long been known that many pulsars exhibit intrinsic timing noise. The best-analysed case is the Crab pulsar for which successive studies have found that the noise is principally describable as a random walk in frequency (called frequency noise, FN) with $s = 4$ (*e.g.* Groth, 1975b; Cordes 1980). This also appears to be true for a variety of other pulsars, although there are indications that admixtures of random walks in phase and torque must also be included (*e.g.* Cordes & Helfand 1980). We can relate the expected mean squared residual to the diffusion coefficient expressed as the strength of the random walk in frequency P_0/P^2, through

$$\frac{P_0}{P^2} = 1.5 \times 10^{-25} \left[\frac{\beta}{N} \right]^3 \left[\frac{\overline{R^2}}{1 \, ms^2} \right] \left[\frac{1s}{P} \right]^2 \, Hz^2 \, s^{-1}. \tag{4.9}$$

If we assume that FN contributes the bulk of the residual (currently $\sim 0.7 \, \mu s$) in PSR 1937 + 214, then the present data imply that $P_0/P^2 \lesssim 1.4 \times 10^{-25} \, Hz^2 \, s^{-1}$. For

R. Blandford, R. Narayan & R. W. Romani

comparison the measured strength of FN in the Crab pulsar is $5.3 \times 10^{-23}\,\text{Hz}^2\,\text{s}^{-1}$ (Groth 1975b) and the upper limit on FN for a quiet pulsar, PSR $1237 + 25$, is $P_0/P^2 \lesssim 7 \times 10^{-30}\,\text{Hz}^2\,\text{s}^{-1}$. To measure $\ddot{P}$ in the millisecond pulsar the rms residual must be less than $10^{-3}\,\mu\text{s}$ over a period of 10 years. This limits the strength of any FN random walk to $P_0/P^2 \lesssim 6 \times 10^{-32}\,\text{Hz}^2\,\text{s}^{-1}$. We thus require the millisecond pulsar to be less restless (by this measure) than any other pulsar we know if the timing is to be exploited fully.

5. Application to other pulsars

Although other pulsars do not have the remarkably small timing residuals of PSR 1937 $+ 214$, the time baselines of the observations are considerably longer ($\gtrsim 10$ yr) and so the results of Section 3 for low-frequency noise can still be of interest. Following Bertotti, Carr & Rees (1983), we consider the orbit decay of the binary pulsar, PSR 1913 $+ 16$. The secular decrease in the binary period has been measured to an accuracy of 4 per cent (Weisberg & Taylor 1984) and agrees to this accuracy with the result $\dot{P}/P = 3 \times 10^8$ yr predicted by general relativity. We can therefore take the error in $\dot{P}/P$ to be $< 0.04/3 \times 10^8$ yr $= 4.2 \times 10^{-18}\,\text{s}^{-1}$. As we have demonstrated, gravitational waves with periods longer than the duration of the observations (but shorter than the light travel time to the pulsar) can cause unusually large variances in period derivatives. PSR 1913 $+ 16$ can be used to set a limit on the energy density in such waves. A background with equal energy density in logarithmic intervals has a spectrum $\propto f^5$ with $P_0 = 1.3 \times 10^4\,\Omega_g^2\,\mu\text{s}^2$. The resulting rms timing residual is given by Equation (3.9) with $s = 5$. Therefore, taking $N = 10$ yr, $\beta = 0.94$, and $N_{\text{max}} = 10^4$ yr and using Table 2 for $s = 5$, we see that the variance in the measured orbit decay time is

$$\frac{\delta\dot{P}}{P} = 0.17 \times 10^{-20} R_{\mu s} = 1.1 \times 10^{-17}\Omega_g^{1/2}\,\text{s}^{-1}. \tag{5.1}$$

Thus, the measured limit $\delta\dot{P}/P < 4.2 \times 10^{-18}\,\text{s}^{-1}$ yields the upper bound $\Omega_g < 0.15$. The limit on the integrated Ω between $N = 10$ and $N_{\text{max}} = 10^4$ is $\Omega_{\text{tot}} < 1.0$.

A similar bound can be obtained from PSR $1952 + 29$, which has the largest known timing age. We can consider its observed $\dot{P}/P = 4.7 \times 10^{-18}$ to be a statistical upper bound on the rms error in the estimate of its age. Using $N_{\text{max}} = 10^3$ yr and $N = 10$ yr, one obtains, as above, the limits $\Omega_g < 0.26$ and $\Omega_{\text{tot}} < 1.2$. Other noise spectra are also strongly limited. The expected variance for spindown noise (SN, $s = 6$) is

$$\frac{\delta\dot{P}}{P} = 9 \times 10^{-15}\left[\frac{R_{\text{rms}}}{1\,\text{ms}}\right]\text{s}^{-1} \tag{5.2}$$

so that SN processes are unlikely to contribute more than $\sim 10^{-4}$ of the measured timing residual.

Cordes & Helfand (1980) have determined the dominant noise process for a number of pulsars; the timing noise of PSR $0823 + 26$, for example, is apparently described by SN. If the observed 12.6 ms residual is in fact SN dominated, then for ~ 10 years of observation, our model predicts the rms error in $\delta\dot{P}/P$ to be $4.4 \times 10^{-15}\,\text{s}^{-1}$. The measured timing age, $\tau = 4.9 \times 10^6$ yr could then be in error by as much as a factor of two or three. This suggests the interesting possibility that such pulsars with a

Arrival-time analysis for a millisecond pulsar

sufficiently small spindown rate could actually have an observed *spinup* because of strong noise with a steep red spectrum.

As has been previously noted, timing noise makes $\ddot{P}$ measurements and braking index determinations very uncertain. The nominal braking indices reported by Gullahorn & Rankin (1982), ranging up to 10^5 and of both signs, are evidently spurious and can be largely accounted for by the variance expressed by Equation (3.11). Both SN and FN processes as well as a gravitational radiation background can produce δn_b's of the appropriate magnitude.

There are three independent methods for estimating the proper motions of pulsars. Direct interferometry appears to be the most accurate and gives reproducible results (Lyne, Anderson & Salter 1982). Measuring the speeds of scintillation diffraction patterns at the Earth is less accurate and does not provide a direction for the motion but the results here appear to be in agreement with the interferometric determination. The third method, however, which relies on fitting arrival times has only produced a credible result in the case of PSR 1133 + 16 (Manchester, Taylor & Van 1974). Furthermore, the timing *positions* do not agree with those determined interferometrically (Fomalont *et al.* 1984). The discussion of Section 3.3 shows that, in the presence of red noise, uncertainties in the pulsar parameters are often much larger than the reported experimental errors which are calculated assuming white noise alone. The variances in position and proper motion determinations can, in fact, grow with increased observation time. It seems worthwhile to try to pre-whiten the timing noise in these pulsars to see if their timing positions and proper motions can be brought into agreement with the interferometrically determined values.

Acknowledgements

Appendix

In this paper we have described the timing noise exclusively in terms of power spectra in the arrival residuals. This approach differs from that followed by earlier authors and we now relate the two methods.

Following Boynton *et al.* (1972), Groth (1975) and Cordes (1980), consider three distinct forms of noise, which they describe as random walks in phase (PN), in frequency (FN) and in the time derivative of the frequency (SN). We have corresponding noise spectra with associated exponents $s = 2, 4$ and 6. However, we make an essential simplification in that we assume the noise to be completely described by its power spectrum. This restricts us to random walk steps that are sufficiently small and frequent to be unresolved by the observations. The formalisms of Groth and Cordes are developed to enable them to detect finite step sizes as well. In practice this has not yet been possible as these effects appear to be masked by measurement errors. (In fact, it

R. Blandford, R. Narayan & R. W. Romani

should also be possible to develop the power spectrum approach along these lines by considering bispectra and three-point correlation functions. We shall not pursue this.)

A second important difference is in the treatment of transients associated with the start of the observations. Cordes artificially assumes that the noise commences at the same instant as the observations. The influence of all prior noise can then be absorbed in the fitted values for the phase, the period and its derivative. A Monte Carlo method is used to relate the ensemble-averaged rms phase residual after a least-squares polynomial fit to the rms phase residual that would have resulted from the same noise adopting the phase, the period and its derivative at the start of the observations. The ratio of these two rms residuals is the quantity $C_R(m, T_{obs})$ where m denotes the order of the polynomial and T_{obs} the duration of the observations. $C_R(m, T_{obs})$ is independent of T_{obs} provided the rate of occurrence r of random walk steps satisfies $rT_{obs} \gg 1$. Groth deals with the transients in a related manner but instead makes an orthogonal polynomial fit to the observations and compares the coefficients of these polynomials with their expectation values. Both approaches accommodate the non-stationary nature of the phase residuals through a memory of the start of the observations, although the underlying noise process is white in the relevant parameter (*e.g.* frequency), is stationary and possesses a well-defined correlation function.

In our approach, we deal with the transients by assuming that the noise process has been switched on adiabatically in the distant past and that the phase noise (or equivalently arrival time noise) has a power spectrum which is simply related to the frequency noise spectrum. If the Wiener-Khintchine theorem for the frequency is written

$$\frac{\langle \delta v(t)\delta v(t+\tau)\rangle}{v^2} = \int_0^\infty df Q(f) e^{2\pi i f \tau};$$ (A1)

then the true underlying arrival time power spectrum is simply

$$P(f) = Q(f)/4\pi^2 f^2$$ (A2)

and so on for other types of power spectra. These power spectra as defined here are all stationary.

In fact, we can compute the correction factors, $C_R(m, T_{obs})$, introduced by Cordes and evaluated by him through a Monte Carlo procedure directly from these power spectra. Consider phase noise first. The quantity that Cordes considers is

$$\overline{R_{PN}^2(T_{obs})} = \left\langle \frac{1}{T_{obs}} \int_0^{T_{obs}} [R(t) - R(0)]^2 dt \right\rangle.$$ (A3)

Taking Fourier transforms and expressing the result in terms of the power spectrum of the residuals yields

$$\overline{R_{PN}^2} = \int_0^\infty P(f) T_{PN}(f) df$$ (A4)

where the filter function, $T_{PN}(f)$, is

$$T_{PN}(f) = 2\left[1 - \frac{\sin 2\pi f T_{obs}}{2\pi f T_{obs}}\right]$$ (A5)

Arrival-time analysis for a millisecond pulsar

and T_{obs} is in years. In comparison, the formalism of Section 2 gives the filter function for a quadratic fit (3 parameters, α_1, α_2, α_3 only) to be

$$T_3(f) = \left[1 - \frac{9}{2x^2} - \frac{9}{x^4} - \frac{45}{2x^6} \right] + \cos 2x \left[\frac{3}{2x^2} - \frac{36}{x^4} + \frac{45}{2x^6} \right]$$
$$+ \sin 2x \left[-\frac{12}{x^3} + \frac{45}{x^5} \right] \tag{A6}$$

where $x = \pi f T_{obs}$. Substituting $P(f) = P_0 f^{-2}$ for phase noise one can calculate $\overline{R_{PN}^2}$ from (A4), and $\overline{R_3^2}$ by substituting $T_3(f)$ instead of $T_{PN}(f)$. Their ratio is the correction factor $C_R(2, T_{obs})$ of Cordes; we obtain the same numerical value. In the case of frequency noise, $s = 4$, and Cordes considers

$$\overline{R_{FN}^2(T_{obs})} = \left\langle \frac{1}{T_{obs}} \int_0^{T_{obs}} \left[R(t) - R(0) - t\,(dR/dt)_{t=0} \right]^2 dt \right\rangle. \tag{A7}$$

The appropriate filter function in this case is

$$T_{FN}(f) = 2 \left[1 + \frac{y^2}{6} - \frac{\sin y}{y} + \frac{y \cos y - \sin y}{y} \right] \tag{A8}$$

where $y = 2\pi f T_{obs}$. Finally, for spindown noise we have

$$T_{SN}(f) = 2 \left[1 + \frac{y^4}{40} + 2\frac{y \cos y - \sin y}{y} + \frac{1}{2}(y^2 - 2)\frac{\sin y}{y} \right]. \tag{A9}$$

We verify the numerical results of Cordes in each case.

References

Armstrong, J. W. 1984, *Nature*, **307**, 527.
Armstrong, J. W., Cordes, J. M., Rickett, B. J. 1981, *Nature*, **291**, 561.
Backer, D. C. 1984, *J. Astrophys. Astr.*, **5**, 187.
Backer, D. C., Kulkarni, S. R., Heiles, C., Davis, M. M., Goss, W. M. 1982, *Nature*, **300**, 615.
Backer, D. C., Kulkarni, S. R., Taylor, J. H. 1983, *Nature*, **301**, 314.
Bertotti, B., Carr, B. J., Rees, M. J. 1983, *Mon. Not. R. astr. Soc.*, **203**, 945.
Blandford, R., Narayan, R. 1984a, *Proc. Workshop Millisecond Pulsars*, (in press).
Blandford, R., Narayan, R. 1984b, *Mon. Not. R. astr. Soc.*, (in press).
Boynton, P. E., Groth, E. J., Hutchinson, D. P., Nanos, G. P., Jr., Patridge, R. B., Wilkinson, D. T. 1972, *Astrophys. J.*, **175**, 217.
Cordes, J. M. 1980, *Astrophys. J.*, **237**, 216.
Cordes, J. M., Greenstein, G. 1981, *Astrophys. J.*, **245**, 1060.
Cordes, J. M., Helfand, D. J. 1980, *Astrophys. J.*, **239**, 640.
Cordes, J. M., Stinebring, D. R. 1984, *Astrophys. J.*, **277**, L53.
Davis, M. M., Taylor, J. H., Weisberg, J., Backer, D. C. 1984, in preparation.
Deeter, J. E. 1984, *Astrophys. J.*, **281**, 482.
Deeter, J. E., Boynton, P. E. 1982, *Astrophys. J.*, **261**, 337.
Detweiler, S. 1979, *Astrophys. J.*, **234**, 1100.
Fomalont, E. B., Goss, W. M., Lyne, A. G., Manchester, R. N. 1984, *Mon. Not. R. astr. Soc.*, (in press).
Groth, E. J. 1975a, *Astrophys. J. Suppl. Ser.*, **29**, 443.
Groth, E. J. 1975b, *Astrophys. J., Suppl. Ser.*, **29**, 453.

Gullahorn, G. E., Rankin, J. M. 1982, *Astrophys. J.*, **260**, 520.
Heiles, C., Kulkarni, S. R., Stevens, M. A., Backer, D. C., Davis, M. M., Goss, W. M. 1983, *Astrophys. J.*, **273**, L75.
Hogan, C. J., Rees, M. J. 1984, *Nature*, **311**, 109.
Lamb, D. Q., Lamb, F. K. 1976, *Astrophys. J.*, **204**, 168.
Lyne, A. G., Anderson, B., Salter, M. J. 1982, *Mon. Not. R. astr. Soc.*, **201**, 503.
Manchester, R. N., Taylor, J. H. 1977, *Pulsars*. Freeman, San Francisco.
Manchester, R. N., Taylor, J. H., Van, Y. Y. 1974, *Astrophys. J.*, **189**, L119.
Mashhoon, B. 1982, *Mon. Not. R. astr. Soc.*, **199**, 659.
Press, W. H. 1975, *Astrophys. J.*, **200**, 182.
Romani, R. W., Taylor, J. H. 1983, *Astrophys. J.*, **265**, L35.
Vilenkin, A. 1982, *Phys. Lett.*, **107B**, 47.
Weisberg, J. M., Taylor, J. H. 1984, *Phys. Rev. Lett.*, **52**, 1348.

Chapter 2:

TIMING A MILLISECOND PULSAR IN A GLOBULAR CLUSTER

Roger D. Blandford, Roger W. Romani

Theoretical Astrophysics

California Institute of Technology

Pasadena, CA 91124, U.S.A.

and James H. Applegate

Department of Astronomy

Columbia University

New York, N.Y. 10027, U.S.A.

Summary. Certain evolutionary scenarios predict that millisecond pulsars should form in the cores of globular clusters; such a source may be present in M28. Gravitational perturbations of the cluster stars should subject the pulsar to a time-varying acceleration. This may be detectable in the arrival time analysis as an anomalous $\ddot{P}$, much larger than that expected from magnetic dipole braking. This effect can be used as a probe of the cluster dynamics and will degrade the usefulness of such pulsars in searches for long wavelength cosmological gravitational radiation.

1. Introduction

Certain evolutionary scenarios associate millisecond pulsars with the low mass X-ray binaries, and consequentially predict that they may be found in the cores of globular clusters (Alpar, *et al.* , 1982). Condensed clusters have core radii $r_c \sim 1$pc and velocity dispersions of $\sigma \sim 10$ km s^{-1}, giving central densities of $\rho_c \sim 10^4 - 10^5 \mathrm{M}_\odot pc^{-3}$. A pulsar in this environment will move through the stippled gravitational field of the cluster members and experience varying accelerations due to close encounters with nearby stars. The resultant perturbations in position, velocity and acceleration will not be measurable using a pulse arrival time analysis, as they will be absorbed into the fundamental parameters of the timing model: ϕ_0 (fiducial phase), P (spin period), and $\dot{P}$ (period derivative). The rate of change of the pulsar's acceleration, $\dot{a}$ will, however, produce an anomalous $\ddot{P}$, which is observable. In particular, although we have no *a priori* knowledge of P or $\dot{P}$ (which depends on the pulsar dipole moment), in the context of magnetic dipole spindown we expect the "braking index", $n = 2 + P\ddot{P}/(\dot{P})^2$ to have the value 3. Measured n's for young, fast pulsars are close to this value (Crab, n=2.515±0.005, Groth 1975; PSR1509-58, n=2.8±0.2, Manchester and Durdin 1984; PSR0540-69, n=3.6±0.8, Middleditch, *et al.* , 1987). If the measured $\ddot{P}$ differs substantially from the spindown prediction, then we can ascribe this difference to the varying acceleration of the pulsar.

Recent studies (*eg.* Erickson, *et al.* , 1987) suggest that the faint, highly polarized source in the globular cluster M28 (NGC 6626) has the steep spectrum $\alpha \sim -2.5$ characteristic of a fast pulsar. Present limits from period searches suggest that, if the source is a pulsar, $P \le 5$ ms. From Peterson and King (1975) M28 has $r_c \sim 0.4$pc and $\sigma \sim 8$ km s^{-1}; the source position is $\sim 0.6 r_c$ from the optical cluster center.

2. Estimation of $\ddot{P}$

The typical interstellar distance in the core is $b \sim (\mathrm{M}_\odot/4\pi\rho_c)^{1/3}$, for cluster members of $\sim 1\mathrm{M}_\odot$. Thus the pulsar acceleration contributed by the nearest cluster member, $\sim G\mathrm{M}_\odot/b^2$, changes on the timescale $t \sim b/\sigma$

$\sim 2 \times 10^3 (10^5 \mathrm{M}_\odot pc^{-3}/\rho_c)^{1/3}(10$ km s$^{-1}/\sigma)$ years. Accordingly, the rate of change of a is

$$\dot{a} \sim \frac{G\mathrm{M}_\odot \sigma}{b^3} \sim 3/4\pi G\rho_c\sigma \sim \frac{3\sigma^3}{r_c^2} \tag{1}$$

since $\sigma^2 \simeq (4\pi/9)G\rho_c r_c^2$. The varying Doppler shift gives $\ddot{P}/P \sim \dot{a}/c$ so that the $\ddot{P}$ induced by dynamics is

$$\frac{\ddot{P}}{P} \sim 10^{-29}(\sigma/10 \text{ km s}^{-1})^3 (\text{pc}/r_c)^2 \text{s}^{-2} \tag{2}$$

Note that there will also be a contribution to $\dot{a}$ from the mean cluster potential of the same order as (1). In addition, since $\dot{a}$ grows as b^{-3}, rare very close encounters become important in estimating the average value of $\ddot{P}$. As the rms value diverges for $b \to 0$ we use the median absolute value of $\dot{a}$ for estimates. A numerical simulation of 1.8×10^4 stars in a Plummer potential with a core radius and isotropic velocity dispersion appropriate to M28 confirms the estimate above, giving a median $\dot{a}$ about twice (1), $ie.$ $\ddot{P}/P = 7 \times 10^{-29}$ for a pulsar at $0.6r_c$, and we adopt this value. In comparison, for a normal field pulsar with a velocity ~ 100 km s^{-1} placed in the galactic disk of mean density $\sim 0.1 M_\odot \text{pc}^{-3}$, the induced $\ddot{P}/P$ from (1) will be a factor of $\sim 10^5$ smaller. The problem of calculating the rate of change of acceleration for a star in a uniform background distribution of stars has been solved by Chandrasekhar (1943,1960). He evaluates the rms value of $\dot{a}$ given a stellar density ρ and a *fixed* star speed v. Substituting $v = \sqrt{3}\sigma$ gives the estimate $\dot{a}_{rms} \approx 13.5 G\rho\sigma$, significantly larger than the median and attributable to the skewness of the distribution.

The anomalous braking index corresponding to a $\ddot{P}$ other than that due to spindown is $n^* \sim 4(\ddot{P}/P)\tau^2$ where $\tau = P/2\dot{P}$ is the characteristic age. This is

$$n^* \sim 4 \times 10^3 (\ddot{P}/P)_{-28}(\tau_8)^2 \tag{3}$$

with τ_8 in units of 10^8 years, typical of other millisecond pulsars, and $(\ddot{P}/P)_{-28}$ in units of 10^{-28}s^{-2}. Clearly, the cluster dynamics can dramatically affect the observed braking index. To determine if an n as large as (3) will be observable, we must estimate other possible error sources in the timing analysis. If the dominant source of an rms timing residual of $R_{rms}\mu$s is white measurement error then the resultant uncertainty in $\ddot{P}$ will be $(\ddot{P}/P)_{-28} \sim 4N^{-3.5}n_{10}^{-1/2}R_{rms}$, where N is the observation period in years and roughly $10n_{10}$ observations are made yearly (Blandford, Narayan and Romani 1984). Thus ~ 3 years of 10 measurements per year with 1μs rms residuals will give a ~ 10 percent measurement of the braking index in (3). If pulsar seismicity in the form of low-frequency "red" noise dominates the timing residual, then the estimate of n will improve more slowly; in fact, for spin frequency (FN) noise the accuracy will remain constant. If we assume FN at the level of the quietest pulsar for which timing noise has been measured, PSR 1237+25 (millisecond pulsars may well be even quieter), then a few percent measurement of n^* should be possible. Note that this red

timing noise contribution causes the very large ($10^4 - 10^5$) spurious n's found for ordinary pulsars (Gullahorn and Rankin 1982) which, unlike the n^* from cluster dynamics estimated above, do not converge to a fixed value as the timing proceeds.

3. Conclusions

If the timing accuracy achievable for a fast globular cluster pulsar is comparable to that attained for other millisecond pulsars and if internal red noise processes are not too strong, then ~ 3 years of timing should give a measurement of a braking index much larger than that expected for dipole spindown. The various other effects of the gravitational perturbations will probably be too small to be detectable. For example if, as appears likely on evolutionary grounds, millisecond pulsars are reborn in tight binaries, then the low escape velocity of the cluster core suggests that the pulsar should still have a close companion. If the binary separation is r_b, then tidal effects on the orbit from the nearest cluster star should scale as $(r_b/b)^3$, *ie.* a part in $\sim 10^9$ for orbits small enough to be detectable. The binary center of mass would, of course, experience the $\dot{a}$ estimated above. We note that if a family of planet sized companions could be acquired during the binary evolution, the additional gravitational perturbations could introduce significant timing noise. Finally, since for sources such as that in M28 the cluster-induced $\ddot{P}$ is detectable, it must be fitted for in any arrival time analysis. This extra fitting term in the timing model will absorb power from the residual at periods comparable to the observation span and weaken bounds that can be placed on unknown perturbations at those frequencies, such as arrival-time delays caused by low frequency gravitational waves. Accordingly, the utility of globular cluster millisecond pulsars as probes of relic background gravitational radiation will be degraded (Blandford, Narayan and Romani 1984). The timing estimate of the cluster potential should, however, provide an interesting and potentially useful comparison with the optically derived values.

References

Alpar, A., Cheng, A., Ruderman, M. and Shaham, J. 1982, *Nature*, **300**, 728.

Blandford, R., Narayan, R. and Romani, R.W. 1984, *J. Astrophys. Astron.*, **5**, 369.

Chandrasekhar, S. 1943. *Annals of the New York Academy of Sciences*, Vol. **XLV**, article 3.

Chandrasekhar, S. 1960. in *Principles of Stellar Dynamics*, (New York:Dover), 277.

Erickson, W.C., Mahoney, M.J., Becker, R.H. and Helfand, D.J. 1987, *Astrophys. J.* **314**, L45.

Groth, E.J. 1975, *Astrophys. J. Suppl.*, **29**, 453.

Gullahorn, G.E. and Rankin, J.M. 1982, *Astrophys. J.*, **260**, 520.

Manchester, R.N. and Durdin, J.M. 1984, *Bull. AAS*, **16**, 542.

Middleditch, J., Pennypacker, C.R. and Burns, M.S. 1987, *Bull. AAS*, **18**, 1049.

Peterson, C.J. and King, I.R. 1975, *Astr. J.*, **80**, 427.

Chapter 3:

MODEL ATMOSPHERES FOR COOLING NEUTRON STARS

Roger W. Romani

Theoretical Astrophysics

California Institute of Technology

Pasadena, CA 91125, U.S.A.

ABSTRACT

A number of investigators have computed the surface temperature of a cooling neutron star as a function of time after its birth and the physics of the high density interior. *Einstein* observations of supernova remnants and radio pulsars in the soft X-ray band have confronted this cooling theory with rather low upper limits to the surface temperature and a few possible detections, based on the assumption that the surface emits as a blackbody. We examine the effect of various surface compositions on the blackbody assumption, calculating model atmospheres for the physical conditions typical of neutron star surfaces with realistic opacities. We find that, for hydrogen or helium dominated surfaces or neutron stars with very low effective temperature, the soft X-ray flux can be much greater than the blackbody value. If high Z elements dominate the surface the number of counts expected is comparable to the blackbody value. In this case, however, we show that absorption edges will be prominent in the spectrum and discuss the possibilities for future spectroscopy. We also discuss potential surface compositions and the importance of settling of heavy elements. Comparing our results with *Einstein* observations, we examine the consequences for searches for young neutron stars--in particular, our results tend to strengthen the conclusion that any neutron star present in several well-studied young remnants must have cooled faster than permitted in standard scenarios.

I INTRODUCTION

Prompted by the discovery in rocket experiments of galactic X-ray sources, a number of authors, starting with Tsuruta (1964) and Bahcall and Wolf (1965), have estimated the X-ray flux of a neutron star cooling after its initial hot ($\sim 10^{11}$ K) birth in a supernova explosion. Successive refinements to these calculations have considered the effect of a wide range of physical processes on the cooling of the star and have resulted in a number of investigations of the properties of matter at high densities. While these calculations do not agree in detail, the general consensus is that although surface temperatures will fall to a few times 10^6K in $\sim 10^3$ years due to neutrino energy losses, neutron stars should remain potentially detectable as soft X-ray sources ($T \gtrsim 10^5$K) for on the order of 10^5 years. The neutrino-dominated phase is sensitive to the poorly understood physics of the dense interior where the presence of exotic phases such as quark matter or pion condensates can dramatically affect the cooling curves. At later times the thermal evolution of the star is dominated by the transport of energy through a thin envelope where most of the temperature gradient occurs. Here there is less theoretical uncertainty, and much progress has been made recently in the detailed treatment of the transport, notably in the work of Yakovlev and Urpin (1981), Gudmundsson (1981) and Hernquist (1984). This body of work has been reviewed in detail by Tsuruta (1985).

While it is now apparent that the bright galactic X-ray sources are powered by accretion and the fluxes predicted for typical cooling neutron stars have decreased greatly, the hope of constraining the properties of matter at high densities and the problematic association of neutron stars with known supernova remnants ensures that detection of their thermal flux has remained an important observational goal. The advent of *Einstein* and similar imaging X-ray satellites made this goal a realistic possibility. Helfand, Chanan and Novick (1980) and Helfand (1983) have described the results of the *Einstein* program, which included studies of 72 galactic supernova remnants and nearly two dozen known radio pulsars. In addition to the thermal flux arising from the initial explosion, they considered heating of older neutron stars caused by internal release of energy in vortex creep and glitch events and surface heating by magnetospheric phenomena.

The observations, using either the Imaging Proportional Counter (IPC) or the High Resolution Imager (HRI) (*cf*. Giacconi, *et al*. 1979), have placed rather strong limits on the flux in the soft ($\sim 0.3-3$keV) band. The corresponding upper limits to the surface temperature were roughly $2-5 \times 10^5$K for radio pulsars and ranged from 0.7 to 2×10^6K for

young supernova remnants. For some remnants, notably SN1006 and Tycho, the limits may be significantly lower than the temperatures given by standard cooling theory. One remnant, RCW103, has yielded a possible point source detection, but here spectral information is lacking-- assuming the spectrum to be thermal gives a temperature of $\sim 2 \times 10^6$K. Similarly, only the nearest old radio pulsar, PSR1929+10, was detected as a point source. Its surface temperature was estimated to be $\sim 2 \times 10^5$K, although the spectrum may be inconsistent with a blackbody law. These temperatures are derived assuming that the surface emits as a blackbody. If, however, there is a substantial atmosphere present, the spectrum and hence the inferred effective temperatures may be quite different.

For the early cooling theory calculations, Morton (1964) calculated the spectra of hot grey atmospheres and Orszag (1965) used approximate multilevel ions to estimate the effect of K-edges on neutron star spectra. These results were however primarily for the higher temperature ($\sim 10^7$K) models considered as explanations of bright galactic X-ray sources. Recently, the problem of the apparent super-Eddington flux of X-ray bursting neutron stars has led to several calculations of helium-rich, high-temperature neutron star atmospheres (Ebisuzaki and Nomoto, 1986; London, Taam and Howard 1984, 1986) at temperatures of $\sim 1-3 \times 10^7$K. These atmospheres are dominated by scattering and Comptonization and display significant flux excesses above the thermal peak, explaining the high (super-Eddington) color temperatures found in the 1-10keV energy band. London, Tamm and Howard (1986, hereafter LTH) also considered somewhat lower temperatures and include numerical spectra for models as cool as 2.9×10^6K.

In this paper we calculate model atmospheres to evaluate the accuracy of the blackbody assumption for surface temperatures and compositions more typical of cooling neutron stars and isolated pulsars. The temperatures considered are somewhat lower, ranging from 10^5K to $10^{6.5}$K where the effects of the atomic (bound-bound, bound-free, and free-free) processes dominate the opacity. Opacities are obtained from the Los Alamos Astrophysical Opacity Library (Huebner, *et al.* 1977). In section II we describe our procedure for calculating the emergent spectra of LTE model atmospheres. In section III we present some representative spectra and quantify the departures from a blackbody, computing the ratio of the flux in a soft X-ray energy band to that expected from a blackbody and fitting power laws to the emergent flux. In section IV we fold spectra through *Einstein* detector response curves and estimate the consequences for the central temperature bounds determined from the satellite data. With an eye toward future spectroscopy, we

also suggest certain absorption features which should be detectable in AXAF caliber experiments and could provide useful estimates of the surface conditions.

II MODEL ATMOSPHERE CALCULATIONS

Our computation is based on the classic temperature correction approach to the construction of radiative equilibrium, LTE model atmospheres (e.g. Mihalas 1978). The surface of a neutron star, however, represents a rather different set of physical conditions than is commonly considered. The surface gravity will be $g_s \sim 10^{14}$g cm/s^2 and the temperature is $\sim 10^6 K$ so that an optical depth of unity in soft X-rays occurs at a depth of ~ 1 cm and densities of $\rho \sim 1 - 10$g/cm^3. Since the stellar radius is $\sim 10^6$cm, the photosphere is plane parallel to a high accuracy. At these temperatures and densities the dominant sources of absorption are free-free and bound-free processes so the opacity can be very non-grey, being roughly Kramers's law, $\kappa \sim E^{-3}$, between absorption edges.

The calculation starts from an approximate grey atmosphere on a Rosseland mean depth scale at an effective temperature T_{eff}

$$T(\tau_R) = T_{eff}(3/4(\tau_R + q))^{1/4} \tag{1}$$

where $q = .71044$ ensures correspondence with the exact grey solution at large optical depth. A typical model has a grid of ~ 100 depth levels, logarithmically spaced in τ_R from 10^{-3} to $10^{2.5}$, and $\sim 100 - 300$ energy levels, logarithmically spaced from ~ 0.1Kev to the highest energies of interest, ~ 10Kev. We interpolate and smooth the $T(\tau_{Ri})$ with a fitted spline function to provide a continuous temperature run. We impose hydrostatic equilibrium

$$\frac{dP}{d\tau_R} = \frac{g_s}{\kappa_R(\tau_R)} \tag{2}$$

by iteratively solving for $P(\tau)$

$$P'(\tau) = \left[\frac{3}{2} g_s \int_0^\tau \frac{P^{1/2}(\tau')}{\kappa_R(\rho(\tau'), T(\tau'))} d\tau' \right]^{2/3} \tag{3}$$

where κ_R is the Rosseland mean opacity and the equation of state is used to obtained the pressure from the density. For most of the atmospheres studied $P^{1/2}dP$ was nearly linear in τ in the outer layers, so that integration of the equation of hydrostatic equilibrium in

the form of Eq. 3 was found to ensure rapid numerical convergence to the solution for $P(\tau)$. At the relatively low T_{eff} considered here, radiation pressure is negligible.

To determine the flux at each energy and depth $\tau_{R\,i}$ we use the Milne relation

$$F_E(\tau_E) = 2\pi \left[\int\limits_{\tau_E}^{\infty} S_E(\tau_E')E_2(\tau_E'-\tau_E)d\tau_E' - \int\limits_{0}^{\tau_E} S_E(\tau_E')E_2(\tau_E-\tau_E')d\tau_E' \right] \tag{4}$$

where E_2 is the second exponential integral function, $S_E(\tau)$ is the source function at energy E and level τ, and τ_E is the optical depth at energy E corresponding to $\tau_{R\,i}$ on our mean depth scale, determined from

$$\tau_E = \int\limits_{0}^{\tau_{Ri}} \frac{\chi_E}{\kappa_R}\, d\tau_R'. \tag{5}$$

The source function, S_E, contains, in general, terms for both scattering and thermally emitted flux. The electron scattering opacity is $\sigma = \sigma_T n_e/\rho \sim 0.1-0.2 \mathrm{cm}^2/\mathrm{g}$ in term of the electron density and the Thompson cross section. At the mean energy, however, the absorptive opacity will be $\sim 10^3-10^5 \mathrm{cm}^2/\mathrm{g}$, so scattering should be unimportant except for the highest temperature models at the highest frequencies. In this regime, however, we can approximate the effect of scattering quite well by using the Planck function for $S_E(T(\tau_R))$ and by taking the opacity to be

$$\chi_E = \kappa_{abs} + \left[\frac{\kappa_{abs}}{\kappa_{abs}+\sigma} \right]^{1/2} \sigma \tag{6}$$

where the small additive term accounts for the fraction of the random walk scatterings that result in absorption. Varying σ by a factor of two had a negligible effect on the flux at detectable energies.

Using the smoothed temperature run and (5), we now know both S_E and E_2 as (continuous) functions of τ_E'. We may therefore evaluate the integrals in (4) as simple quadratures, using a Romberg integration scheme with an adaptive stepsize and taking the τ_E derived from (5) for the appropriate limits of integration. At each depth we sum the flux over the energies to find $F(\tau_{R\,i})$, the net flux, which in radiative equilibrium should be constant at $\overline{F} = \sigma_R T_{eff}^4$. The departures from constancy, ΔF, are used to estimate corrections to the temperature run using the Lucy-Unsöld procedure (*eg.* Mihalas 1978)

$$\Delta T(\tau) \simeq \frac{T(\tau)^{-3}}{16\sigma_R} \left\{ \frac{\kappa_J}{\kappa_P} \left[3\int_0^\tau \frac{\chi_F(\tau')}{\kappa_R(\tau')} \Delta F(\tau')d\tau' + 2\Delta F(0) \right] - \frac{\kappa_R}{\kappa_P}\frac{d\Delta F(\tau)}{d\tau_R} \right\} \qquad (7)$$

where κ_R, χ_F, κ_P and κ_J are the Rosseland, flux, Planck and absorption mean opacities, respectively. The correction was found to be sufficiently accurate using $\kappa_J = \kappa_P$. The corrections to the temperature are applied at the τ_{Ri}, giving an improved approximation to the desired temperature run. These new $T(\tau_{Ri})$ are then smoothed as above, hydrostatic equilibrium is reinstated using (3), the energy fluxes are recalculated and the procedure is iterated to convergence. We find that departures from flux constancy of less than 1% throughout the atmosphere are generally achieved after three to four iterations.

The code was tested against the analytic solution for the grey atmosphere by using a grey opacity and against the lowest temperature ($T_{eff} = 2.9\times10^6$) helium atmospheres of LTH using a Kramers's law opacity and an ideal gas equation of state. In each case the agreement was to better than 25% at all frequencies of interest. The LTH models are quite hot for our purposes and the "scattering" term in the opacity (6) was found to have a small, but significant, effect at high frequencies.

The absorptive opacities, electron density and equation of state for several pure elements and two astrophysically interesting mixtures were obtained from the Los Alamos Astrophysical Opacity Library data (Huebner, et al. 1977) with the latest improvements to the equation of state. These data covered the range of density $10^{-3}<\rho(g/cm^3)<10^2$, temperatures from $\sim10^4$ to $\sim5\times10^7$, and energies from 1eV to 10keV. Occasionally points at lower density were used, in which case the EOS was extrapolated with the ideal gas law and the opacity with a Kramers's law run.

III EMERGENT SPECTRA

Models were calculated for surface effective temperatures of 10^5, $10^{5.5}$, 10^6, and $10^{6.5}$K and surface compositions of pure iron (as appropriate for BPS matter; Baym, Pethick, and Sutherland 1971), pure helium, and a cosmic abundance (X=0.73; Y=0.25; Z=0.02, in cosmic proportions). To investigate the effect of various absorption edges upon the emergent spectra, carbon, oxygen and helium rich (Y=0.98, Z=0.02) models were calculated as well. The effect of varying the surface gravity is largely given by the variation in the redshift of the spectrum to infinity, as is shown in figure 1. This shows the surface

(unredshifted) flux for two iron surface composition, 10^6K models with surface gravities of 10^{15} cm/s^2 and 10^{14} cm/s^2. These surface gravities bracket those appropriate for realistic neutron star equations of state. The variations in the spectrum are small and are largely due to an increased pressure ionization at one optical depth in the high gravity model for energies below the absorption edges. The variation in redshift (z=0.13 for $g_s = 10^{14}$ cm/s^2 *vs.* z=0.84 for $g_s = 10^{15}$ cm/s^2) is quite large and will be the dominant effect upon the spectrum measured at infinity. We therefore restrict our calculations to $g_s = 2.43 \times 10^{14}$ cm/s^2, appropriate for a typical neutron star mass of $1.4M_\odot$ and radius of 10 Km. We present fluxes that would be measured at the stellar surface. In order to convert to fluxes measured at a distance D, one must renormalize

$$F_{E,\infty}(E) = \left[\frac{R_{NS}}{D}\right]^2 \frac{F_{E,NS}(E(1+z))}{(1+z)} \qquad (8)$$

where $F_{E,\infty}$ is the spectrum at infinity, $F_{E,NS}$ is the surface spectrum and R_{NS} is the locally measured radius of the star. This allows us to restrict our attention to a two parameter (T_{eff}, composition) family of models. Note that the effective temperatures are surface values, as well; the T_{eff} of the equivalent blackbody spectrum at infinity will also be lowered by (1+z). When comparing against *Einstein* observations we have used the mass and radius given above to find the redshifted values at the observer.

Figures 2 and 3 show the effect of surface composition for two surface temperatures, $T_{eff} = 10^{5.5}$ and $T_{eff} = 10^{6.5}$, and two extremes of composition, pure iron and pure helium. For comparison, blackbody spectra and *Einstein* IPC response curves are also shown, with and without the effects of interstellar absorption. The on-axis IPC response taken from Harnden, et al. (1984) shows significant sensitivity in both a low energy (LE, ~.15-.5 keV) range and a somewhat higher energy (HE, ~.5-4.5 keV) channel while parameters for a typical line of sight ($D = 1$kpc, $n_H = 1$ cm^{-3}) are used to illustrate the effect of interstellar absorption, showing that the sensitivity to low energy photons is severely reduced. It is clear that the counts measured in a soft X-ray experiment can be greatly affected by varying the composition of the surface. We quantify this by considering a typical detector response window (eg. IPC or HRI) blueshifted to the surface of the star, 0.5-5.0 keV. In this energy band we calculate the ratio of the expected number of counts to those expected from a blackbody, R_{BB}. In Table 1 we give values of R_{BB} for various effective temperatures and for iron, helium and cosmic abundance surface compositions. We also give the local

(unredshifted) effective temperature that would be assigned to the total flux, assuming the spectrum to follow the Planck Law. Since absorption edge features can affect the spectrum significantly, a blackbody fit will not, in general, be appropriate. A simple description appropriate to a limited spectral resolution is obtained by fitting a power law, $F(E) \propto E^{-\alpha}$, to the first decade in flux in the 0.5-5.0 keV window. We give α for the various models and blackbody spectra in the last columns of the table. Although our spectra for 10^5K atmospheres have significant flux in the specified window, the corresponding blackbody flux is negligible; we therefore omit the entries for R_{BB} at this temperature.

To demonstrate the effects of the composition on the opacity and the emergent spectrum we display spectra for $T_{eff} = 10^6$K and surface compositions of C, O, and Fe in figure 4 and He, cosmic abundances and He with metals (in cosmic proportion) in figure 5 along with blackbody curves. The iron spectrum has a rather sharp set of L edges at $\sim$0.7−1.0keV. The K edge at $\sim$7 keV is prominent at somewhat higher effective temperatures (*cf.* fig. 3). Carbon and oxygen also show significant K edges at $\sim$0.3 keV and $\sim$0.5 keV, respectively. The redistribution of the flux in the C model gives rise to a high energy excess at $\sim$2 keV; at somewhat higher T_{eff} an excess in the oxygen spectrum becomes prominent at $\sim$5 keV. For pure helium the extensive Kramers's law run in the opacity causes a similar shift to a higher color temperature. Note, however, that the addition of even 2% metals, as in the helium-rich (He+Z) and cosmic abundance models, makes the opacity sufficiently grey to bring the spectra close to the blackbody curve for energies below a keV. Despite this agreement, some of the L edge structure of iron is still apparent near the blackbody peak. As the Los Alamos data are not intended for spectroscopy and our LTE approximations to the radiative transfer are clearly oversimplifications, the finer details of the calculated spectra are likely to be suspect. Nonetheless, it is clear that there are significant features in the bound-free opacity in the energy range considered if heavy elements are present on the surface.

IV DISCUSSION

Although our choice of surface compositions is largely illustrative, we can provide some motivation for the models considered. The neutron star surface is commonly assumed to be iron, the zero pressure equilibrium state of BPS matter. Further, supernova explosion calculations leaving neutron stars (*e.g.* Weaver, Woosley and Fuller 1985) indicate that the "mass cut" occurs within the iron core, leading to ejection of the

envelope of lighter elements. It is possible, however, for other nuclei to be present in, and even dominate the thin surface layer. If, for example, the explosion is asymmetric or the ejecta are perturbed by the presence of a companion it may be possible to accrete some fraction of the lighter nuclei. Alternatively, material could be transferred from a companion or, in the case of an isolated older neutron star, accreted from the general interstellar medium. A further possibility is the creation of light nuclei as spallation products of the high energy collisions induced by magnetospheric particles and radiation thought to impinge on the surface of an active radio pulsar. To cover the surface to an X-ray optical depth of unity it suffices to deposit only $\sim 10^{14}$g $\sim 10^{-19} M_\odot$ of material. Moreover, this material is likely to suffer significant fractionation on short timescales. The density and temperature at the neutron star surface are similar to the conditions found at the base of low-luminosity white dwarf convection envelopes. The very much higher surface gravity means that the separation will be quite rapid. Using the results of Alcock and Illarionov (1980), we estimate the separation timescale for a helium photosphere of ~ 10cm depth to be $\sim 1 - 100$s for typical temperatures and densities. It is thus possible for the surface be chemically very pure and dominated by the lightest element present. In this scenario a pure hydrogen photosphere is also quite likely. The emergent spectrum would then be very similar to our results for helium. A full treatment of separation and surface abundances is important but somewhat complicated as it involves the computation of diffusion coefficients in high density, high Z plasmas; we defer this to a future communication.

As the present observations constrain only the total flux and not the spectrum we bracket the hypothetical surface compositions by considering pure helium and pure iron models. To compare with the measured flux limits we should also consider the low energy cut-off due to interstellar absorption. From the IPC sensitivity curves in figures 2 and 3 and the steepness of the calculated and blackbody spectra in the 0.2-5. keV energy range, it is clear that interstellar absorption will have a dramatic effect on the number of counts expected from a given emergent spectrum. For temperatures lower than several times 10^6K, this difficulty will dominate the uncertainty in determining an effective temperature. Further as the absorption increases, moving the low energy cut-off of the sensitivity higher, the departures from the counts expected for a blackbody spectrum will, in general, be greater. This means that useful estimates of the surface temperature depend rather critically on the absorption cut-off. In practice this is often estimated from associated column densities, especially when X-ray spectral information is limited. The free

electron column density or that of neutral hydrogen is commonly used.

While such estimates are generally good approximations, in some cases they may not be appropriate. In a young supernova remnant the 1-3$M_{\odot}$ of heavy elements released into the circumstellar environment will produce an excess absorption (not reflected in the hydrogen column density or the absorption cut-off as determined from the remnant spectrum) equivalent to

$$\Delta N_H \sim 10^{21} M_{ej,A} \ r_{ej}^{-2} \left[\frac{20}{A} \right] \left[\frac{\rho_{A,ej}}{500 \ \rho_{A,cos}} \right] cm^{-2} \tag{9}$$

where $M_{ej,A}$ is the mass of the ejecta of atomic weight A in $M_{\odot}$ thrown off to a typical distance of r_{ej} parsecs and $\rho_{A,cos}$ and $\rho_{A,ej}$ are the cosmic and ejecta abundances for species A. Since a typical column density at a kpc is $\sim 3 \times 10^{21}$ cm^{-2}, this can be substantial. The Vela pulsar, where the best fit equivalent N_H from X-ray absorption is ~ 1.5 times the measured value (F. Harnden, private communication), may be an indication of this. Indeed, in some cases, clumping of the ejecta combined with rather steep spectra in the soft X-ray window may be sufficient to obscure an otherwise detectable neutron star. Thus IPC determinations of the equivalent N_H should be used, whenever available.

Keeping in mind that the absorption dominates the uncertainty, we give estimates of adjustments to temperature limits reported for certain soft X-ray observations of supernovae and radio pulsars. We assume a 10km, 1.4$M_{\odot}$ star and compare the blackbody results with spectra for helium and pure iron surface compositions. The spectra are redshifted to infinity and convolved with a low-energy cut-off $T(E)$ of the form

$$T(E_{\infty}) \sim e^{-\tau}, \quad \tau = (N_H/10^{21} cm^{-2}) \ (E_{\infty}/.5 keV)^{-8/3} \tag{10}$$

where N_H is the hydrogen column density, E_{∞} is the energy measured at infinity and the functional form of the optical depth approximates the results of Zombeck (1980).

For a point source in the remnant of SN1006, the temperature limits quoted by Tsuruta (1985) correspond to a value $T_{\infty} \lesssim 7.0 \times 10^5$K (*i.e.* a local surface temperature $T_s \lesssim 9.2 \times 10^5$K) for our assumed stellar mass and radius, a distance of 1.2kpc and the estimated column density $N_H = 1 \times 10^{21}$ (Pye, *et al.* 1981). For a helium surface, with all other assumptions held fixed, the effective surface temperature giving the same flux is 5.7×10^5K, corresponding to $T_{\infty} \lesssim 4.4 \times 10^5$K. For a surface composed of iron the revised temperature is closer to the blackbody value, $T_{\infty} \lesssim 6.2 \times 10^5$K. Similarly, for a neutron star

in Tycho's supernova remnant ($D = 3.0$kpc, $N_H = 3 \times 10^{21}$cm^{-2}, Seward, *et a.l* 1982) we find that the blackbody limit $T_\infty < 1.1 \times 10^6$K is lowered to $T_\infty < 9.2 \times 10^5$K for a helium surface. For a surface composed of iron, the effect of an L edge just below the observing window decreases the flux slightly and the revised temperature limit is $T_\infty < 1.2 \times 10^6$K. Using the computations of Hernquist (1985) we can find the corresponding limits to the neutron star core temperature, the value most useful in constraining the high density physics of the interior. Using the zero field results, we find that for SN1006 the blackbody value $T_c < 1.2 \times 10^8$K is lowered to $T_c < 5.2 \times 10^7$K by a helium surface. For Tycho the limit in the presence of helium is reduced from $T_c < 2.8 \times 10^8$K to $T_c < 1.9 \times 10^8$K. For a surface composed of iron the differences are small. Our result, then, strengthens the conclusion that any neutron star present in these two remnants must have cooled faster than conventional interior physics allows. As these supernovae are believed to be type I, however, the conventional view asserts the absence of a neutron star rather than the presence of exotic stellar interiors.

In the supernova remnant RCW103 Tuohy and Garmire (1980) report the detection of a HRI point source at $(3.8 \pm 0.7) \times 10^{-3}$ counts per second. At an estimated distance of 3.3 kpc and hydrogen column density of 1.5×10^{22} cm^{-2} this corresponds to a blackbody surface temperature of 2.7×10^6K ($T_\infty = 2.1 \times 10^6$K). For helium and iron surfaces the temperature as measured at infinity is revised to 1.7×10^6K and 2.2×10^6K, respectively. Both of these values are in reasonable accord with standard cooling theory.

The strongest candidate for detection of surface flux from an old neutron star, PSR 1929+10, is also the closest known radio pulsar at 0.08 kpc. With an estimated N_H of 5×10^{19}, its IPC detection, at 6×10^{-3} counts s^{-1} (Helfand 1983), corresponds to a blackbody surface temperature of $T_\infty = 2.0 \times 10^5$K. The corresponding temperature for a helium surface is $T_\infty = 1.2 \times 10^5$K; the iron surface value is 1.9×10^5K. Subsequent HRI observations detected a point source coincident with the radio position at $\sim 7 \times 10^{-4}$ counts s^{-1}. For a blackbody spectrum at the IPC determined temperature the expected counting ratio is IPC/HRI ~ 3. Folding through the two detector response curves, we estimate IPC/HRI count ratios of 4.1 for iron and 5.0 for helium. These are in somewhat better agreement with the observed value, IPC/HRI=8 ± 3, as well as the limited IPC spectral data. Thus if PSR 1929+10 has been detected, this crude spectral information suggests that the surface supports a substantial atmosphere. Moreover, recent observations of PSR 1929+10 with the EXOSAT low-energy (LE) telescope and CMA detector (Alpar, *et al.* 1986) found no

significant flux, placing limit of $T_{BB} \lesssim 1.9 \times 10^5$K on a blackbody surface temperature. Since the LE response ($\sim$ 0.04-2.4 keV) is even softer than that of the HRI, this limit is consistent both with the non-blackbody interpretation of the *Einstein* counts and the substantially harder spectrum expected for an absorbing atmosphere.

In the case of PSR1929+10 and other known radio pulsars the surface is believed to be subject to magnetic fields of order 10^{12}G. While it has been speculated that such a field would cause the surface to solidify or form molecular chains, recent computations (Neuhauser, Langanke, and Koonin 1986) indicate that for all Z greater than 2 independent atoms are energetically favored over solid chain configurations. Even for helium the binding energy is only of order 25 eV so the chains will be substantially disrupted at typical surface temperatures. The effect of the magnetic field on the opacity should then largely be a rescaling due to suppression of absorption normal to the field lines. If the overall form of the opacity is preserved our qualitative results will still be relevant. It is still important to calculate the detailed modification of the absorption at high fields and to account for the anisotropic opacity in the radiative transfer. The effect of strong magnetic fields on the emission from a solid neutron star surface has been considered by several authors (Brinkmann 1980, Cheng and Helfand 1983, Greenstein and Hartke 1983) who conclude that while the spectrum amplitude is modulated at the stellar rotation frequency it remains largely blackbody in nature. We note that Cheng and Helfand's conclusion that atmospheric effects are not important was based on a very high binding energy for a high-field iron surface lattice. The presence of an atmosphere will clearly modify these discussions; for example, while Greenstein and Hartke emphasized that the flux modulation could be visible and could explain the X-ray pulse profile of PSR 1509-58, their model requires a rather large temperature difference between the polar cap and the rest of the star which might be difficult to sustain if there is a substantial absorbing atmosphere. Consideration of high magnetic fields is probably less relevant for the limits placed on cooling neutron stars in young supernova remnants as one would expect a high field neutron star to manifest itself in more visible ways than *via* its thermal surface flux-- *i.e.* as a radio pulsar or a synchrotron nebula (Helfand 1983).

In this investigation we have concentrated on a broadband description of the effect of surface composition on neutron star X-ray spectra and have shown that some absorption edges can be prominent at soft X-ray energies. For detailed comparison of such calculated features with future observations, more refined treatments of the opacity and

radiative transfer will clearly be desirable. In particular, the LTE approximation is of limited value when the spectra contain excess high energy fluxes and strong absorption edges, since our species are generally very highly ionized. LTH for example, find that in X-ray bursting neutron stars iron K edges are significantly suppressed due to non-LTE effects. A detailed treatment of scattering may also be important at higher effective temperatures. Since the Los Alamos opacity data are based on the assumption that LTE is valid, such studies necessitate extensions to that code.

Our approximate model atmospheres suffice, however, to show that the surface composition of a cooling neutron star can have a significant effect on the emergent X-ray spectrum. The calculations demonstrate that for a surface composed of iron, assuming the spectrum to be blackbody gives a fairly accurate estimate of the flux in the soft X-ray band for moderate effective temperatures. For a low Z surface composition or a surface temperature less than $\sim 5 \times 10^5$K, the flux expected is much greater than that appropriate for a blackbody curve. Indeed, for relatively low effective temperatures, this difference can have a significant effect on the surface temperature bounds derived from recent satellite observations. Conversely, when heavy elements are dominant, absorption edges can be prominent, although the total flux measured will be close to the black body value. It is thus important to delineate the likely surface compositions, considering the effects of accretion, internal processes and settling in the photosphere. Future studies of the thermal emission from neutron star surfaces should take account of these compositional and spectral uncertainties, uncertainties which the next generation of X-ray imaging satellites should be able to resolve.

TABLE CAPTION

Spectral characteristics in the .5–5.keV energy range at the neutron star surface. Columns are 1) the effective temperature, 2)the surface composition, 3) the ratio of counts to those from a blackbody at T_{eff}, 4) the blackbody temperature giving the expected number of counts, 5) the energy index α, from a fit to to the first decade in flux, 6) the equivalent α for a blackbody.

T_{eff}	Comp.	R_{BB}	T_{BB}	α	α_{bb}
$10^{5.0}$	He	——	2.87×10^5	-7.2	
	Fe	——	2.04×10^5	-14.	-56.3
	Cosmic	——	2.11×10^5	-24.	
$10^{5.5}$	He	6852	5.65×10^5	-3.6	
	Fe	225	4.35×10^5	-5.5	-16.5
	Cosmic	338	4.46×10^5	-8.1	
$10^{6.0}$	He	2.83	1.17×10^6	-1.59	
	Fe	1.84	1.09×10^6	-4.02	-4.5
	Cosmic	1.15	1.01×10^6	-5.22	
$10^{6.5}$	He	.714	2.89×10^6	-.69	
	Fe	1.02	3.16×10^6	-.86	-1.2
	Cosmic	.953	3.12×10^6	-1.1	

REFERENCES

Alcock, C. and Illarionov, A. 1980. *Ap.J.*, **235**, 534.

Alpar, M.A., Brinkmann, W., Kiziloglu, U., Ogelman, H. and Pines, D. 1986. preprint.

Bahcall, J.N. and Wolf, R.A. 1965. *Phys. Rev.*, **140B**, 1452.

Baym, G., Pethick, C.J., and Sutherland, P. 1971, *Ap.J.*, **170**, 299.

Brinkmann, W. 1980. *Astron. Astrophys.*, **82**, 352.

Cheng, A. and Helfand D.J. 1983. *Ap. J.*, **271**, 271.

Ebisuzaki, T. and Nomoto, K. 1986. *Ap. J. (Letters)* **305**, L67.

Giaconni, R. *et al.* 1979. *Ap. J.*, **230**, 540.

Greenstein, G. and Hartke, G. 1983. *Ap. J.*, **271**, 283.

Gudmundsson, E.H. 1981. licentiate thesis, University of Copenhagen.

Harnden, F.R., Fabricant D. G., Harris, D. E. and Schwarz, J. 1984. **Smithsonian Astrophysical Observatory Special Report 393**.

Helfand, D.J. 1983. **Supernova Remnants and Their X-Ray Emission**, eds. J. Danziger and P. Gorenstein, D. Reidel: Dordrecht, 471.

Helfand, D.J., Chanan, G.A. and Novick, R.A. 1980. *Nature*, **283**, 337.

Hernquist, L. 1984. Ph.D. Thesis, Caltech.

Hernquist, L. 1985. *Mon. Not. R. Ast. Soc.*, **213**, 313.

Huebner, W.F., Mertz, A.L., Magee, N.H., Jr., and Argo, M.F. 1977. Astrophysical Opacity Library, Los Alamos Report LA-6760-M.

London, R.A., Taam, R.E., and Howard, W.M. 1984. *Ap. J. (Letters)*, **287**, L27.

------------------ 1986. *Ap. J.* **306**, **170 (LTH)**.

Mihalas, D. 1978. **Stellar Atmospheres**, San Francisco: Freeman.

Morton, D.C. 1964, *Ap.J.*, **140**, 460.

Neuhauser, D., Langanke, K. and Koonin, S.E. 1986. *Phys. Rev. A.*, **33**, 2084.

Orszag, S.A. 1965, *Ap.J.*, **142**, 473.

Pye, J., Pounds, K., Rolf, D., Seward, F., Smith, A., and Willingale, R. 1981. *Mon. Not. R. Ast. Soc.*, **194**, 569.

Seward, F., Gorenstein, P., and Tucker, W. 1982, *Ap.J.*, **266**, 287.

Tuohy, I., and Garmire, G. 1980, *Ap.J.(Letters)*, **239**, L107.

Tsuruta, S. 1964. Ph.D. thesis, Columbia University.

------------------ 1985. Max-Planck Institut Für Astrophysik preprint.

Weaver, T., Woosley, S., and Fuller G. 1985. in **Numerical Astrophysics**, ed.
    J. Centrella, J. LeBlanc, and R. Bowers, Jones and Bartlett:Boston, 374.

Yakovlev, D., and Urpin, V. 1981, *Soviet Astr.(Letters)*, **7**, 88.

Zombeck, M.V. 1980. High Energy Astrophysics Handbook,
    Smithsonian Astrophysical Observatory Special Report 386.

FIGURE CAPTIONS

Figure 1. Effect of surface gravity on unredshifted spectra. Pure iron models at $T_{eff} = 10^6$K and $g_s = 10^{14}, 10^{15}$g cm s^{-2}.

Figure 2. Flux at the stellar surface against energy for $T_{eff} = 10^{5.5}$K models with iron (Fe) and helium (He) surfaces, compared with the corresponding blackbody spectrum (BB). The IPC response function (arbitrary normalization) shows sensitivity in both LE ($\sim$.15 -.5 keV) and HE ($\sim$.5 -4.5 keV) channels; with typical interstellar absorption (IPC w/ abs), $N_H \sim 3 \times 10^{21}$cm^{-2}, the LE response is suppressed.

Figure 3. Flux at the stellar surface against energy for $T_{eff} = 10^{6.5}$K models with iron (Fe) and helium (He) surfaces, compared with the corresponding blackbody spectrum (BB). The IPC response functions are as in Figure 2.

Figure 4. Flux at the stellar surface against energy for $T_{eff} = 10^6$K models for carbon (C), oxygen (O) and iron (Fe) surfaces compared with a blackbody spectrum (BB).

Figure 5. Flux at the stellar surface against energy for $T_{eff} = 10^6$K models for helium (He), helium with 2% metals (He+Z) and cosmic abundance (Cosmic) surfaces compared with a blackbody spectrum (BB).
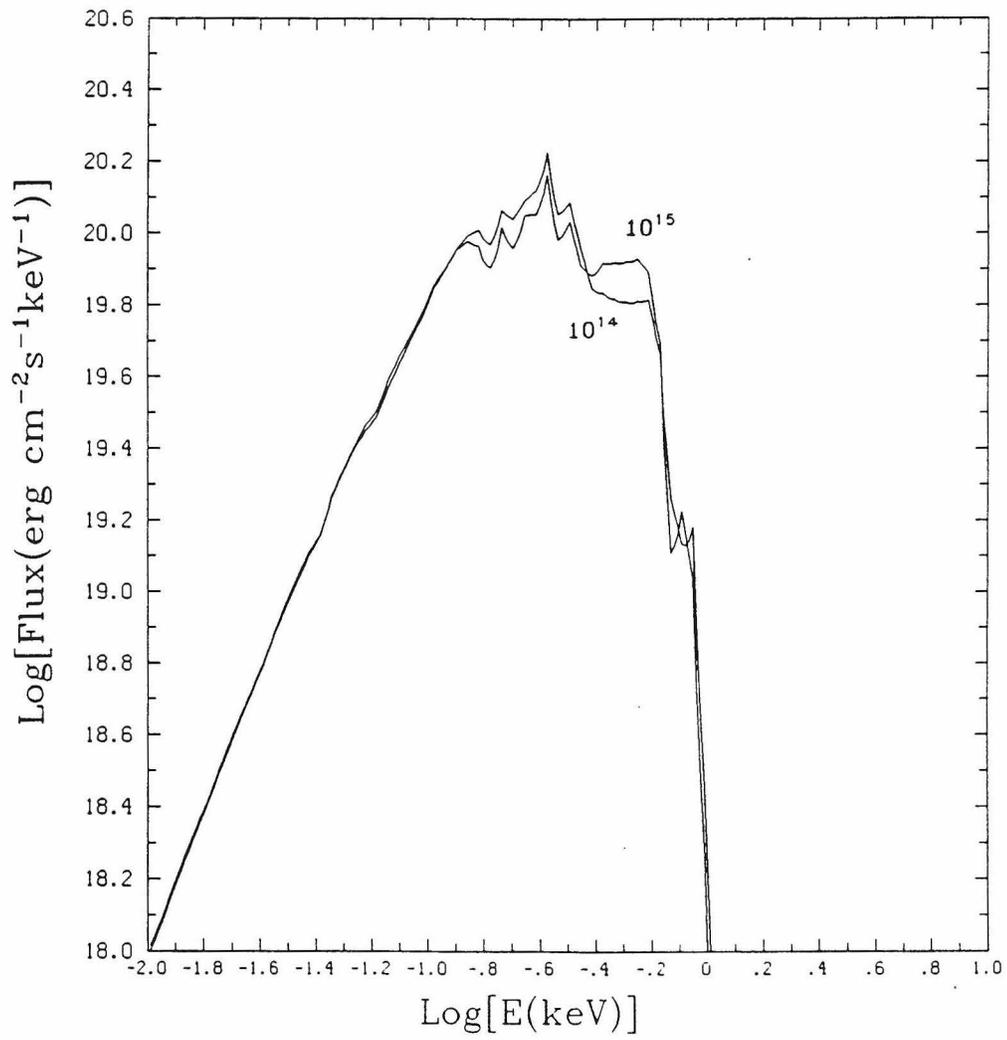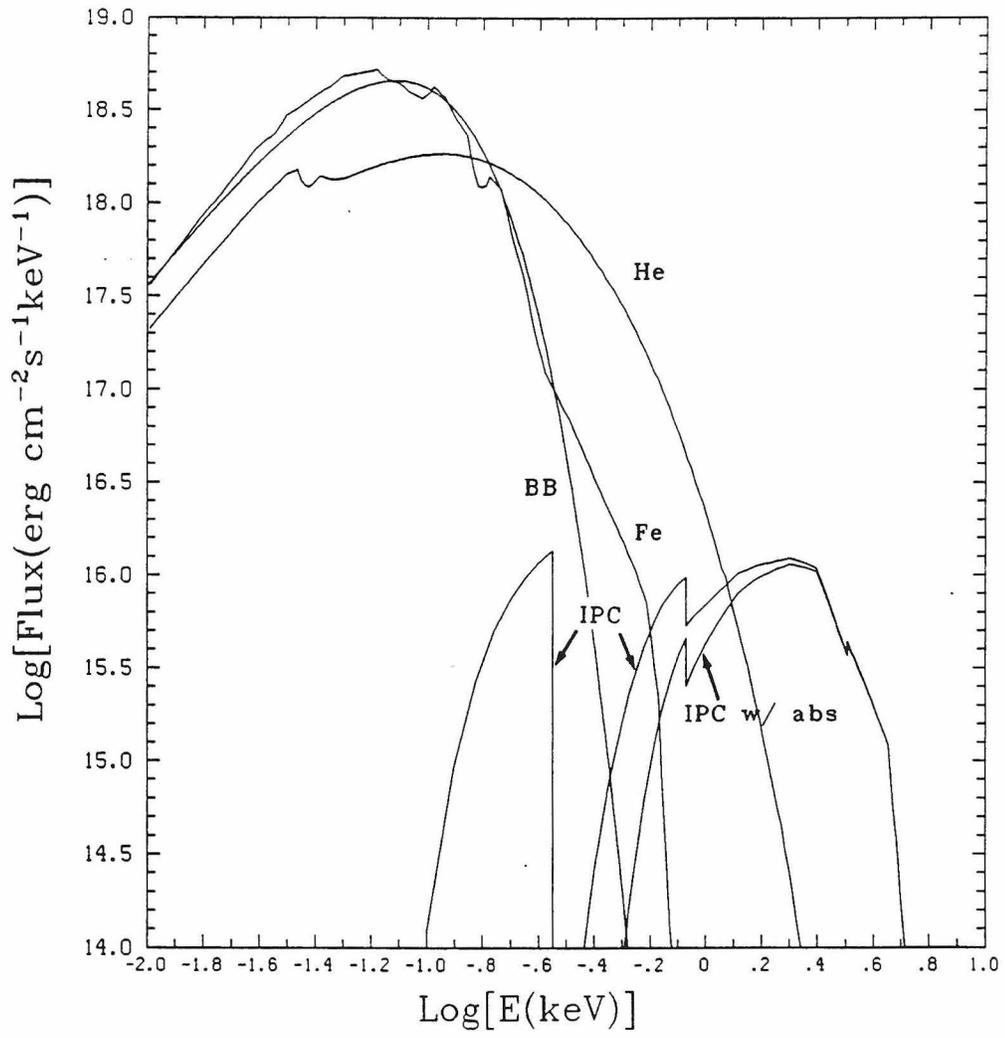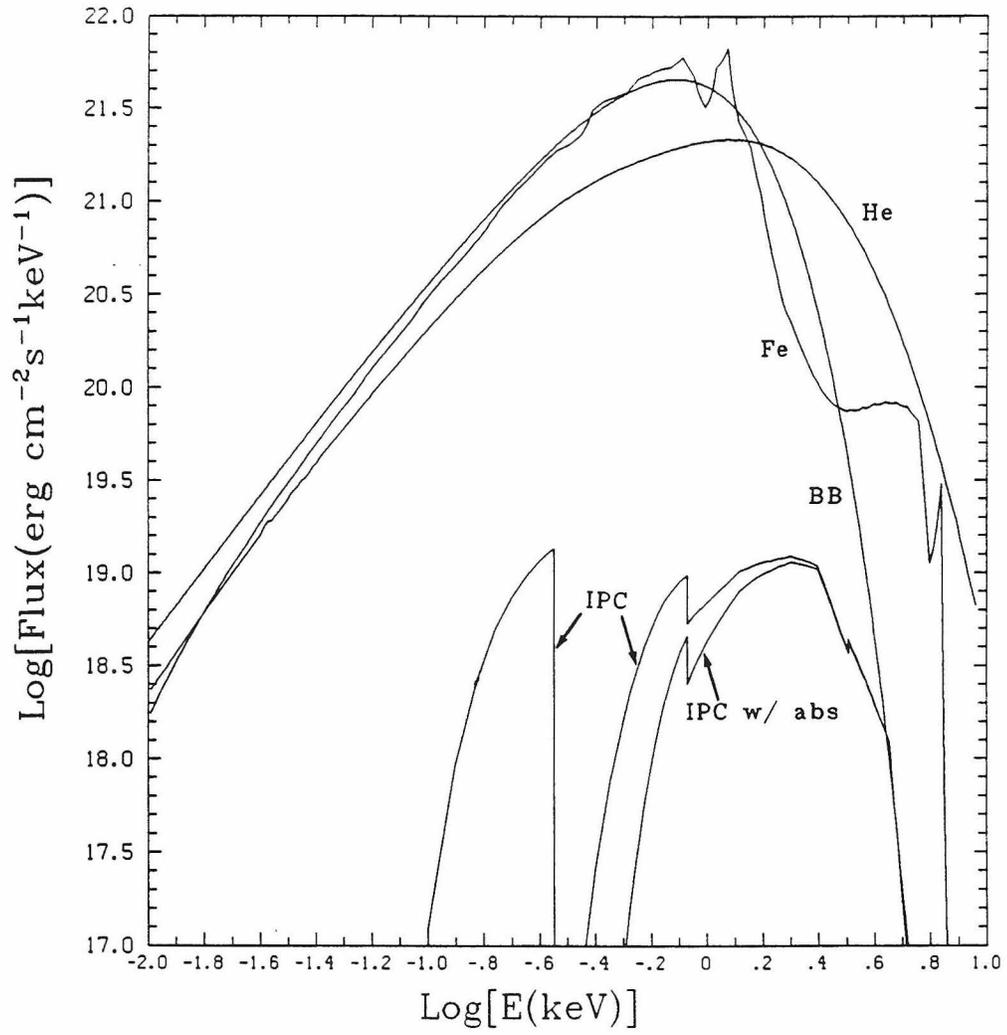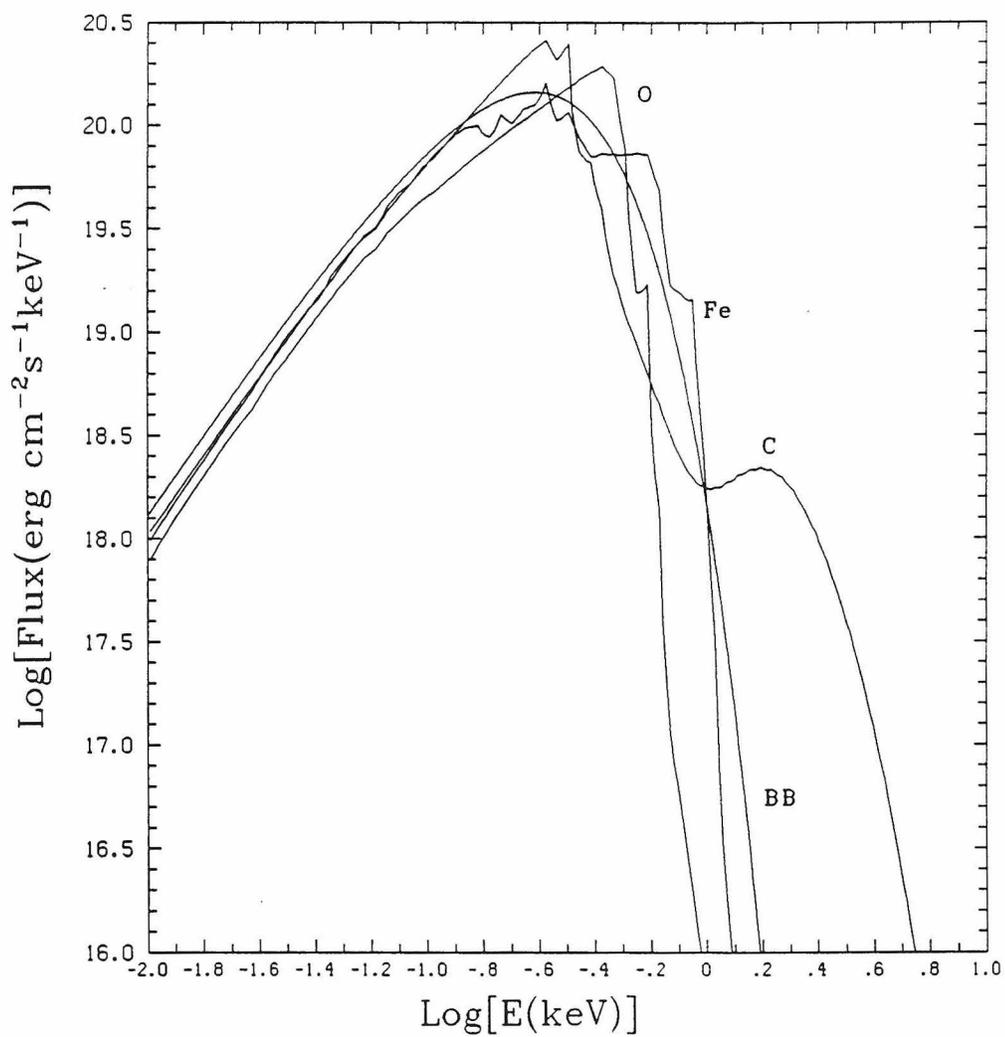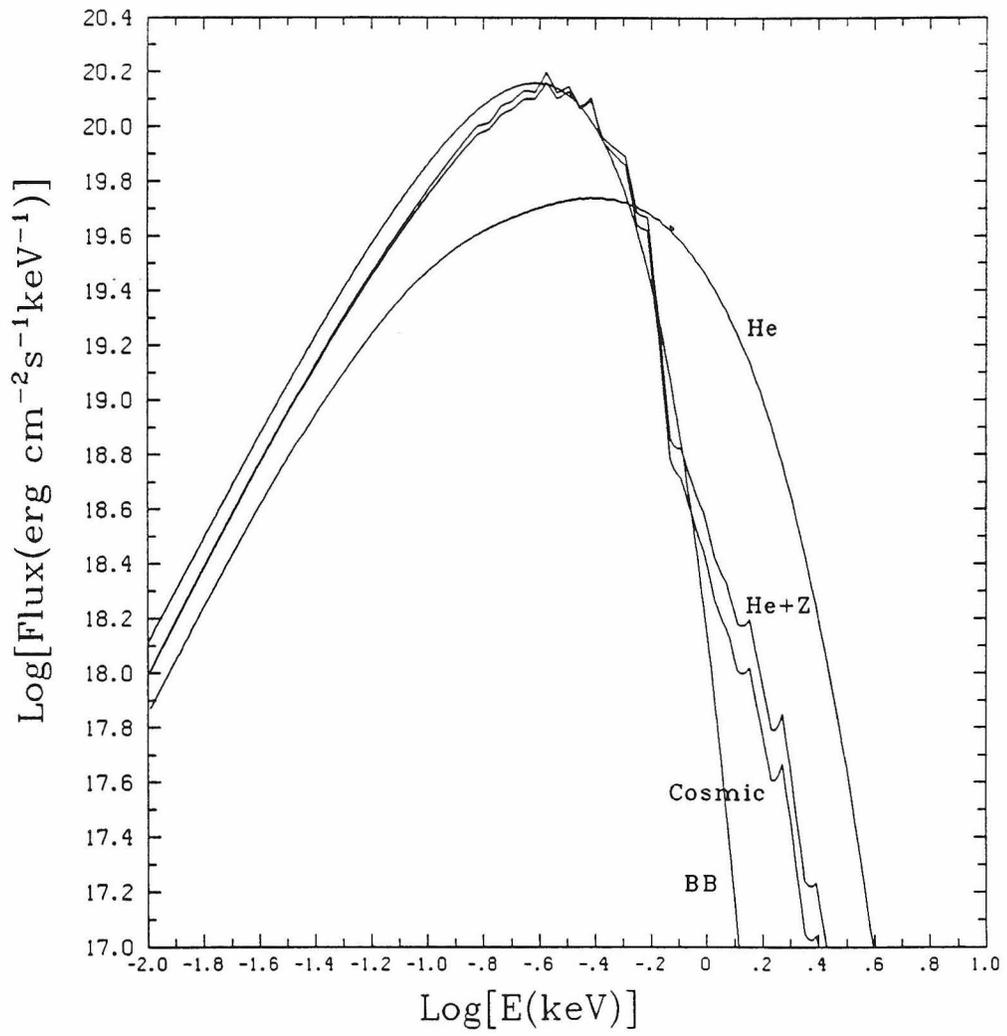
Chapter 4:

X RAY OBSERVATIONS AND CONSTRAINTS ON MODELS OF GAMMA RAY STARS

Roger W. Romani

Theoretical Astrophysics

California Institute of Technology

Pasadena, CA 91125, U.S.A.

ABSTRACT

Neutron star models of gamma ray sources are subject to limits on thermal emission obtained from *Einstein* observations in the soft X-ray band. Examining the effect of light elements accreted onto the stellar surface, we find departures from a blackbody spectrum leading to substantially stronger bounds on the thermal flux. We discuss the effect of such modified temperature limits on thermonuclear burst models. The application of these limits to other emission models and the interpretation of some possible detections are treated briefly.

THERMAL EMISSION FROM NEUTRON STAR GAMMA SOURCES

One of the greatest difficulties in the study of gamma burst sources and other gamma ray stars lies in finding their counterparts in other energy bands. Correspondingly, observational limits on counterpart flux at lower energies can be used to place important constraints on source models. For example, archival searches for optical transients in burster fields can be used to estimate the burst recurrence rate (Schaefer 1981), while Pizzichini, *et al.* (1986) have shown how observations in the soft X-ray band made with the *Einstein Observatory* (Giacconi, *et al.* 1979) can be used to constrain the surface temperatures of quiescent gamma burst sources. In a few cases possible X-ray counterparts have been found, indicating that future observations should allow the study of these sources in detail.

On quite general grounds, we expect gamma ray bursters to be soft X-ray sources. Theoretical consensus is that bursters are associated with neutron stars with surface areas of $\sim 10^{13} A_{13} \text{cm}^2$. The presence of associated optical transients implies recurrence times, in years, of order $\tau_{yr} \sim 0.1-10$. Isotropy of the source distribution suggests a local galactic plane population (although halo and cosmic populations have also been proposed) at distances $\sim 100\, d_{100}$pc and corresponding energies of order $E_{38} \sim 10^{38} d_{100}^2 \text{erg sec}^{-1}$. If the γ-ray emission is isotropic, one can expect the neutron star to absorb an energy of order that observed in the burst and re-radiate it on a recurrence timescale, giving a minimum effective surface temperature

$$T_{eff} \gtrsim 2.8 \times 10^5 (d_{100})^{1/2} (A_{13}\tau_{yr}/E_{38})^{-1/4} \text{K}$$

For specific models the T_{eff} should in general be higher. For example, in thermonuclear explosion models, the energy released due to accretion is $E_{grav} \gtrsim 100\, E_\gamma$.

Such effective temperatures are comparable to the limits achieved in *Einstein* observations of isolated radio pulsars (Helfand, Chanan and Novick 1980, Helfand 1983). In Romani (1987) it was shown that neutron stars experiencing modest accretion rates should have surface atmospheres dominated by low Z elements and the surface temperature bounds derived from *Einstein* flux limits would be lower than those based on the assumption of blackbody surface emission. The model atmosphere calculations are based on realistic opacities and equations of state from the Los Alamos Opacity Library (Huebner, *et al.* 1977). Since convergence to the radiative zero solution is rapid in the

atmosphere and magnetic field will not affect the shape of the opacity law law strongly, these results can be applied to observations of gamma burst fields, as well.

In analogy with white dwarf convection envelopes (Alcock and Illarionov 1980), material accreted onto a neutron star should experience substantial settling, leading to a photosphere of pure light elements (H, He) for accretion rates $\lesssim 10^{-12} M_\odot yr^{-1}$. The Kramers' law opacity of these elements at soft X-ray energies leads to a hardened emergent spectrum with a color temperature higher than the effective temperature. The resultant upper limits on T_{eff} and the associated accretion rate

$$\dot{M} \simeq \sigma T_{eff}^4 A_{13}(R_*/GM_*)$$

are correspondingly strengthened.

COMPARISON WITH OBSERVATIONS

Following Pizzichini, *et al.* (1986), we assume $R_* = 16$ Km, $M_* = 1.3 M_\odot$ and determine the most stringent limit on the effective temperature from the *Einstein* IPC low energy (.15-.5 keV), high energy (.5-4.5 keV) and total (.15-4.5 keV) counting rates using the values in their Table 1. We find that for $T_{eff} \lesssim 3 \times 10^5$K our limits on the thermal flux and $\dot{M}$ are about a factor of 10 more stringent than values based on blackbody emission, while for higher T_{eff} the limits are comparable. Figure 1 shows the surface temperature limits as a function of distance for the April 6, 1979 burster assuming the entire surface is heated. Dotted lines represent limits based on blackbody spectra, solid lines are based on our calculated spectra. To bound the possible effect of absorption, the upper and lower line of each type represent limits based on hydrogen column densities of $N_H = 1.5 \times 10^{20}$ cm^{-2} (the total along the line of sight) and zero, respectively.

The minimum accretion rate compatible with current thermonuclear burst models, $\dot{M} \gtrsim 10^{-15} M_\odot yr^{-1}$, corresponds to $T_{eff} \gtrsim 3 \times 10^5$K. Accordingly, for GRB790406 distances of less than .7 kpc are excluded for such models. In general, the harder spectra expected from a low Z atmosphere strengthens Pizzichini *et al.*'s conclusion that the observed X-ray flux limits are compatible with thermonuclear flash models only if the accretion is episodic or the γ emission is beamed. Distances of less than 0.5 kpc are generally excluded for flash models. For comparison, Epstein (1985) finds that observations of high energy tails in a number of bursts allow maximum distances of $\sim$ 0.2-0.5 kpc. Limits on other

models are similar; models relying on internal energy supplies (*eg.* starquakes) require the release of $10^{40}-10^{41}$erg into pulsational modes. The X-ray flux limits require very high efficiencies for conversion to gamma rays since substantial energy will be dissipated in the surface layers during transmission of flux to the optically thin emission region.

When X-ray counterparts to gamma sources are detected, estimations of their spectra are clearly crucial for interpretation as flux from a neutron star surface. In the case of the November 19, 1978 burster, a marginal (3.5σ) source was detected in the IPC high energy channel at the gamma ray position (Grindlay, *et al.* 1982). In the lower energy channel only an upper limit to the flux is reported. In Figure 2 we follow Pizzichini, *et. al.* in displaying effective temperatures determined from the three sigma limits to the high energy data (shaded regions) as well as the upper bound from the low energy channel. As in figure 1, dotted lines represent results for blackbody spectra, solid lines correspond to calculated spectra and accretion over the entire surface is assumed. The hydrogen column density along the line of sight 2.5×10^{20}cm^{-2} has been assumed in estimating absorption. Note that, for blackbody spectra only distances $\gtrsim 2$ kpc are compatible with the 0.15-5. keV flux limits, while for our harder calculated spectra distances as small as 50 pc are acceptable. If the source is fueled by steady accretion, however, the distance to the source is probably $\gtrsim 1$ kpc.

The *Einstein* counterpart to the elusive gamma ray source Geminga (2CG 195+04), identified by Bignami, Caraveo and Lamb (1983), also has spectral properties useful in constraining the source model. This source, 1E 0630+178, is very soft and shows little absorption ($N_H \lesssim 2\times10^{20}$cm^{-2}), leading to a probable distance ~100 pc. At this distance the effective surface temperature would be $2-3\times10^5$K ($0.017 < $ kT < 0.026). Bignami, *et al.*, however, report that the spectrum is harder, being adequately fit by powerlaw spectra with $3 < \alpha < 3.5$ or thermal spectra with $0.08 < $ kT < 0.1 with column densities of a few times 10^{19}cm^{-2}. A possible interpretation is that $\sim10^{-3}$ of the surface in the form of a hot polar cap is emitting at this temperature and supplying the observed thermal flux, although in this case one might expect modulation of the X-ray flux at the neutron star's rotation period. If however the source has accreted the $\sim10^{-19}M_\odot$ of matter sufficient to form a low Z atmosphere, then we find that the reported number of counts can be produced from emission over the full surface of a neutron star of $R_* = 16$km, $M_* = 1.3M_\odot$ with $T_{eff} \simeq 3.1\times10^5$K. The color temperature in the IPC energy range is substantially higher, giving a power law exponent $\alpha \sim 3.1$ and a thermal fit at $T_{color} \sim9\times10^5$ (kT~0.08), in good

agreement with the observations.

REFERENCES

Alcock, C. and Illarionov, A. 1980. *Ap.J.*, **235**, 534.

Bignami, G.F., Caraveo, P.A. and Lamb, R.C. 1983. *Ap. J.*, **272**, L9.

Epstein, R.I. 1985. *Ap. J.*, **297**, 555.

Giaconni, R. *et al.* 1979. *Ap. J.*, **230**, 540.

Grindlay, J.E., *et al.* 1982. *Nature*, **300**, 730.

Helfand, D.J. 1983. **Supernova Remnants and Their X-Ray Emission**, eds.
    J. Danziger and P. Gorenstein, D. Reidel: Dordrecht, 471.

Helfand, D.J., Chanan, G.A. and Novick, R.A. 1980. *Nature*, **283**, 337.

Huebner, W.F., Mertz, A.L., Magee, N.H. Jr., and Argo, M.F. 1977.
    Astrophysical Opacity Library, Los Alamos Report LA-6760-M.

Pizzichini, G., *et al.* 1986. *Ap. J.*, **301**, 641.

Romani, R.W. 1987. *Ap. J.***313**, 718.

Shaefer, B.E. 1981. *Nature*, **294**, 722.

FIGURE CAPTIONS

Figure 1. Effective surface temperature limits versus source distance from *Einstein* observations of GRB790406. Dotted lines represent limits derived from blackbody spectra, solid lines from computed spectra. The upper and lower curves represent limits for absorption from the full column density along the line of sight and no absorption, respectively.

Figure 2. Limits on effective surface temperature using blackbody and computed spectra for GRB781119, as in figure 1. The full absorption along the line of sight has been assumed; shaded regions represent the three sigma range allowed by the IPC high energy detection, while the single lines give the upper limits from the low energy channel.

EINSTEIN FLUX LIMITS−−GRB790406

EINSTEIN FLUX LIMITS--GRB781119
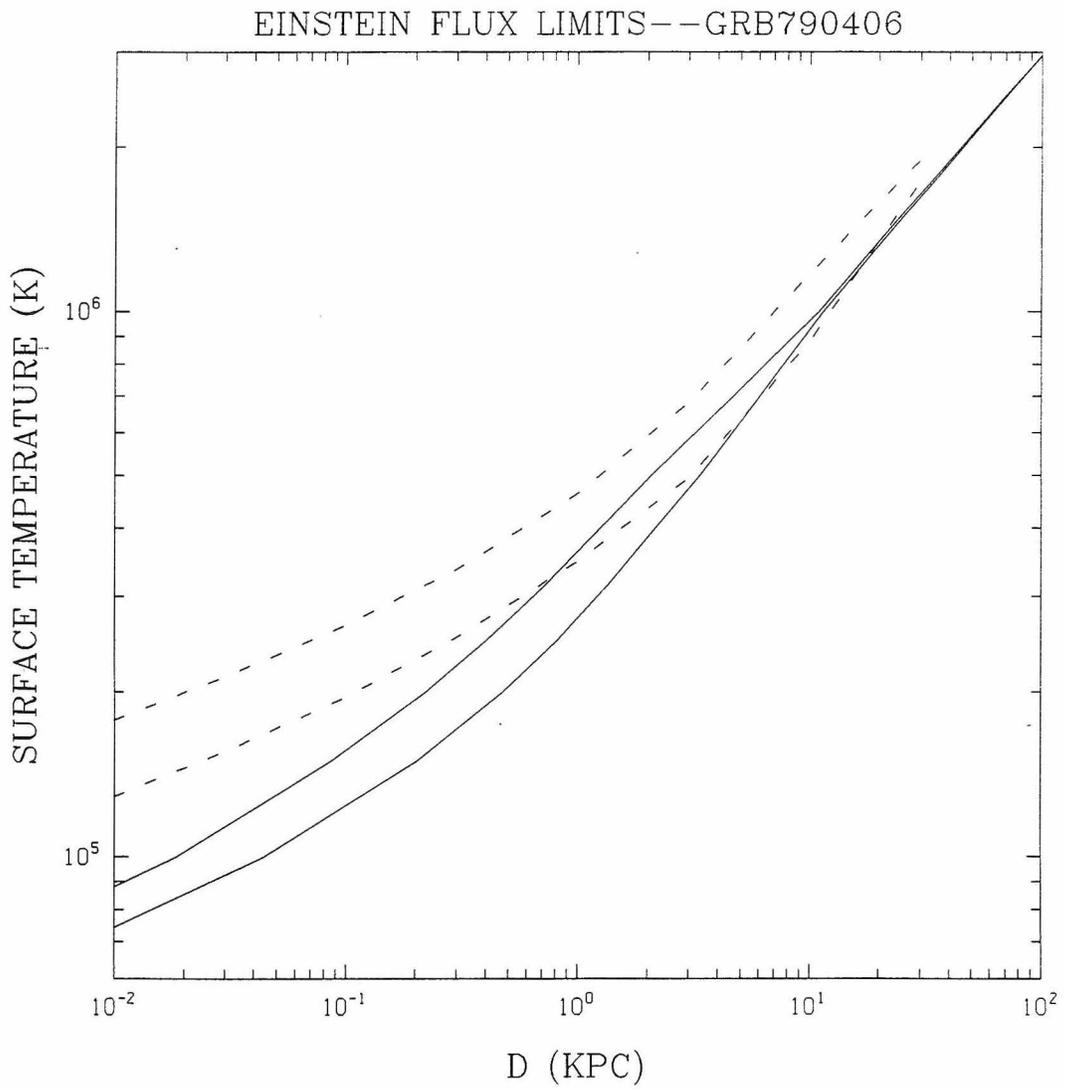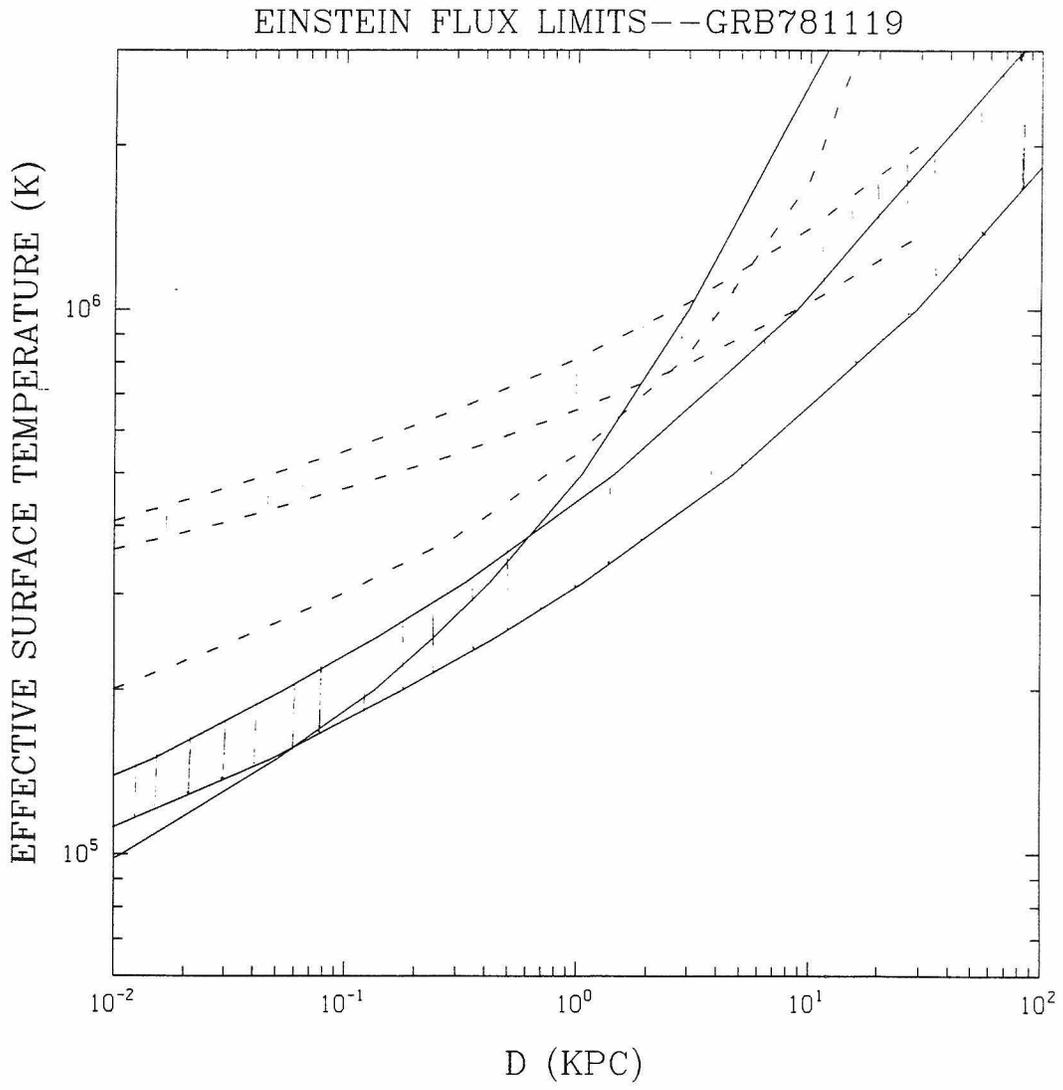
Chapter 5:

REFRACTIVE EFFECTS IN PULSAR SCINTILLATION

Roger W. Romani, Ramesh Narayan and Roger Blandford

Theoretical Astrophysics

California Institute of Technology

Pasadena, CA 91125, U.S.A.

ABSTRACT

Recent studies have focused attention on the refractive effects of long wavelength ($\lesssim 10^{14}$ cm) electron density fluctuations in the interstellar medium upon radio observations of pulsars and compact extragalactic radio sources. In earlier work, a simple scattering model was introduced which allowed us to compute fluctuations in mean intensity, image size, pulse width and pulse arrival time, along with their cross-correlations and fluctuation timescales when there is a power-law spectrum of density perturbations in a thin "equivalent screen" of scattering material. In this work, we extend the analysis to include refraction-induced fluctuations in intrinsically diffractive quantities such as the scintillation timescale, t_s, and the decorrelation bandwidth, ν_{dc}. We then use the theory to study the drifting bands in dynamic scintillation spectra caused by the dispersive steering of the diffraction pattern. We also estimate the fluctuations in the position of the image on the sky, rates of variation of intensity and position, and the root mean square elongation of the scatter-broadened image. We make two further extensions of the theory. First we show that, despite certain formal divergences, the theory can be extended to accommodate steeper density fluctuation spectra (power law indices $\beta > 4$) than the conventionally assumed Kolmogorov spectrum ($\beta = 11/3$). Second, we test the validity of the thin screen approximation, developing a formalism to treat scattering in an extended medium. We find that the thin screen theory sometimes underestimates the refractive fluctuations by a factor ~ 2. The auto- and cross-correlations of the various observables are calculated and comparison is made with the known scintillation properties of pulsars to select those effects most suited to observational verification. The predicted cross-correlation between decorrelation bandwidth and flux fluctuations seems particularly well-suited to measurement. These measurements should, in turn, provide insights into the density fluctuation spectrum and the distribution of the scattering along the line of sight.

1 INTRODUCTION

The effect of the small scale electron density perturbations in the interstellar medium upon the propagation of pulsar radio signals has been recognized since the earliest observations (Scheuer 1968, Rickett 1977, Manchester & Taylor 1977). These inhomogeneities scatter the rays by a root mean square scattering angle θ in propagating a distance D to Earth. The extra pathlength traversed by a given ray leads to a mean *geometric time delay* $t \sim D\theta^2/2c$ and the dispersion in this value among a large number of received rays causes a *pulse broadening* of the same magnitude. Since the phases of the rays are uncorrelated, their interference produces deep scintillation and creates a diffraction pattern at Earth with a lateral coherence length $b \sim \lambda/2\pi\theta$ and a *decorrelation bandwidth*, $\nu_{dc} \sim c/\pi D\theta^2$. The motion of an observer relative to this pattern at a speed v (dominated by the pulsar's peculiar velocity) leads to a diffractive *scintillation timescale*, $t_s \sim b/v$.

It has been argued on observational grounds (*e.g.* Lee & Jokipii 1975, Rickett 1977) that the electron density perturbations have a three-dimensional power spectrum, $\Phi_k \propto k^{-\beta}, 2 < \beta < 4$, *ie.* the density fluctuations on a scale a vary as $\delta n \propto a^{(\beta-3)/2}$. For such a spectrum, the scattering angle induced by a fluctuation on scale a satisfies $\delta\eta \sim \delta n(a) r_e\lambda^2/2\pi$ with $r_e = e^2/mc^2$. If there are D/a such regions along the pulsar-Earth path then the scatterings will add incoherently to give an rms scattering angle $\theta(a) \sim (D/a)^{1/2}\delta\eta \sim a^{(\beta-4)/2}\lambda^2 D^{1/2}$. Hence, for $\beta<4$, the scattering will be dominated by the smallest scale, a_{min}, permitted by the strong scattering condition: *viz.*, phase fluctuations on a scale a_{min} satisfy $\phi(a_{min}) \sim a_{min}\theta/\lambda > \pi$. Thus, the angular size of the image has the following scaling, $\theta \propto \lambda^{\beta/(\beta-2)}D^{1/(\beta-2)}$. Observations reveal that $\partial(\log\theta)/\partial(\log\lambda) \gtrsim 2$ (*e.g.* Mutel *et al.* 1974), implying that $\beta \sim 4$. More specifically, it has been argued that the exponent β has the Kolmogorov value of $11/3$ (*e.g.* Armstrong, Cordes & Rickett 1981).

Recently, there has been increasing evidence that perturbations much larger than the diffractive scale a_{min} are also important in the propagation of radio frequency radiation in the interstellar medium. In particular, it seems that electron density perturbations on a scale comparable to the size of the scatter-broadened spot on the sky, $\sigma \sim \theta D$, can cause refractive focusing and defocusing of the pulsar image (see fig. 1). Rickett, Coles & Bourgois (1984), following Sieber (1982) (*cf.* also Shapirovskaya 1978), proposed that these effects may account for long period ($\sim$ days-years) variations in pulsar

intensity as well as the low frequency variability of compact extragalactic radio sources. In addition, the sloping drift patterns of pulsar dynamic scintillation spectra could also arise from the influence of these large scales (Shishov 1974, Hewish 1980, Roberts & Ables 1982, Hewish *et al.* 1985).

Blandford & Narayan (1984), pointed out that, in addition to the received flux, many other observed parameters can also vary due to refractive perturbations. In a later paper, (Blandford & Narayan 1985, hereafter BN), they developed a simple model, reviewed in section 2, to calculate the magnitudes of the various fluctuations and the correlations among them. A simple extension of their analysis allows us to estimate the fluctuations in the observed source position, which should be correlated with $\dot{F}$, the rate of variation of flux F, as shown in Figure 1. We also consider the random elongation of the scatter-broadened image.

The technique employed in BN and in this paper involves *intensity*-weighted averages of quantities of interest. At first sight this approach would appear to be unsuitable for a description of diffraction-related phenomena such as the decorrelation bandwidth ν_{dc} or the scintillation timescale t_s. In fact, the theory is also capable of treating such phenomena as we show in section 3. Since ν_{dc} and t_s are related to the angular spread in the rays received, their fluctuations are correlated with variations in the flux, angular size Ω, pulse broadening τ, etc. Another interesting diffractive phenomenon that we study is the drifting bands seen in dynamic scintillation spectra. The sloping patterns are believed to be produced by gradients or "prisms" in the scattering medium (Shishov 1974, Hewish 1980); they should thus be correlated with position shifts and $\dot{F}$. In section 4 we evaluate the various correlations for power-law spectra, $k^{-\beta}$, of density perturbations.

The analysis presented in BN required that the scattering due to refractive scales be smaller than the small scale scattering, θ. This is valid in the strong scintillation regime provided the spectral exponent $\beta \leq 4$. Recently, however, there has been some suggestion that β may exceed 4 (BN, Goodman & Narayan 1985, hereafter GN, Hewish *et al.* 1985). We show in section 5 that, although the refractive scattering angle can in principle diverge for $\beta > 4$, a simple "renormalization" of the theory can be developed to handle this regime. Another important question concerns the validity of the thin screen approximation when the scattering really occurs in an extended medium. We consider this issue in section 6 using an extension of our formalism and find that a single screen underestimates the magnitude of refractive flux fluctuations by a factor of ~ 2.

·In section 7 we give numerical estimates of the various fluctuations for power-law spectra with $\beta = 11/3$, 4 and 4.3. The magnitudes of the observable quantities as well as their scalings as a function of λ, D, velocity v, and the strength of the density perturbations C_N^2 are collected together in Table 1. The magnitudes of the cross-correlations are presented in Table 2 and the relevant formulae are given in Appendix A. Using these, we discuss in section 8 the feasibility of detecting the various refractive effects. Refractive fluctuations in diffractive quantities such as t_s and ν_{dc} seem to be quite suitable for experimental verification, particularly since the measurements can be made at high frequencies where the refractive timescale $(T_{ref} \sim D\theta/v)$ is short. Further, the slopes of drifting bands in dynamic scintillation spectra have the unique property of being sensitive to density fluctuations on scales much larger than the refractive scale $\sim D\theta$. Hence they are particularly well-suited to placing limits on the outer scale (*ie.* large length scale cut-off) of the fluctuation spectrum. The data on drift slopes presently available already seem to suggest that an outer scale must be present and that a simple power-law estimate of the fluctuation spectrum is insufficient to explain all results. An outer scale is also suggested in several cases by the physical requirement that the electron density fluctuation $\delta n(a)$ on a scale a must not exceed the mean density n. This places stringent limits on regions of high C_N^2, such as the line of sight to the Vela pulsar and the galactic center. Apart from an outer scale, there are distinctive phenomena, related to caustics, that can occur in the presence of an inner scale. We hope to treat these in a later communication.

2 THE SCATTERING MODEL AND REFRACTIVE FLUCTUATIONS

BN treated the effects of long wavelength ("refractive") fluctuations in the ISM as weak perturbations of an underlying bundle of rays scatter-broadened by the diffractive scale inhomogeneities. When averaged over a time much greater than the scintillation timescale t_s, the image of a point source, such as a pulsar, will be essentially gaussian with a characteristic angular radius θ. This gaussian bundle will be focused, defocused, steered, etc. (figure 1) by density fluctuations on the scale of the "spot" or image size, σ. For simplicity, the refractive effect of the scattering medium is estimated in terms of an equivalent thin screen with large-scale phase variations, $\phi(\mathbf{r})$; the scattering strength of the screen and its distance from the observer, L, are adjusted so that the observed angular size θ of a point source as well as the mean geometrical time delay $(t = L\theta^2/2c)$ are the same as in the model of the medium (see Appendix A in BN). The spot size on the

screen will then be $\sigma \sim \theta L$. The extra refractive bending angle $\eta(\mathbf{r})$ of a ray at transverse location $\mathbf{r}$ on the screen is given by

$$\eta(\mathbf{r}) = -\lambdabar \left(\partial \phi(\mathbf{r}) / \partial \mathbf{r} \right) \tag{2.1}$$

where $\lambdabar = \lambda / 2\pi$.

Since the diffraction pattern is moving relative to the observer (due to the motions of the pulsar, Earth and the medium), the time dependence of the various observable quantities will be given by their spatial dependence in the observer plane (see Fig. 1). Thus, if $\overline{F}$ is the mean flux from the source, then the intensity received at a general point $\mathbf{x}$ from unit area around the point $\mathbf{x+r}$ on the screen is

$$I(\mathbf{r},\mathbf{x}) = \frac{\overline{F}}{\pi\sigma^2} \exp-\left[\frac{L\eta+\mathbf{r}}{\sigma}\right]^2 \tag{2.2}$$

where η is evaluated at the point $(\mathbf{x}+\mathbf{r})$ and we have assumed a gaussian spot shape. A more formal justification of this approximation is given in Appendix B. As the deflection η is, by assumption, small compared with σ/L, we can expand the argument of the exponential to first order in η and integrate over $\mathbf{r}$ to calculate the fluctuations in the observed flux. Substituting for η from (2.1), integrating by parts and normalizing to the mean flux, $\overline{F}$, we obtain the *fractional* intensity fluctuation

$$\delta F(\mathbf{x}) \equiv \frac{\Delta F(\mathbf{x})}{\overline{F}} = \frac{4\lambdabar L}{\pi\sigma^6} \int d^2r \; \phi \; (r^2 - \sigma^2) \exp(-\frac{r^2}{\sigma^2}) \tag{2.3}$$

with $\phi = \phi(\mathbf{x}+\mathbf{r})$. Throughout the paper, we use the symbol Δ to denote the fluctuation in some quantity and δ to describe its *fractional* fluctuation, as above.

BN calculate similar expressions for fractional fluctuations in the angular size of the image, Ω, the mean time delay of the pulse, t, and the pulse width, τ (the last two are normalised with respect to the mean pulse broadening $L\theta^2/2c$). Each of these quantities is of the form (*cf*. Appendix A)

$$\delta A_i(\mathbf{x},\lambdabar) = \int d^2r \, \phi(\mathbf{x}+\mathbf{r},\lambdabar) f_i(r,\lambdabar) \tag{2.4}$$

As the observer moves through the diffraction pattern at a speed $\mathbf{v}$, these fluctuations will vary. The random internal velocities of the phase screen will generally be much smaller than $\mathbf{v}$, and so a time lag T is equivalent to a spatial lag $\mathbf{s} = \mathbf{v}T$. The cross-correlations at

lag s between the fluctuations in two quantities A_1 and A_2 is given by

$$\delta A_1(\mathbf{x},\lambda_1)\delta A_2(\mathbf{x}+\mathbf{s},\lambda_2)=\int d^2r_1 \, d^2r_2 \, \phi(\mathbf{x}+\mathbf{r}_1,\lambda_1)\phi(\mathbf{x}+\mathbf{s}+\mathbf{r}_2,\lambda_2) \, f_1(r_1,\lambda_1)f_2(r_2,\lambda_2)$$

where f_1, f_2 denote any of the f_i. Taking two-dimensional Fourier transforms and averaging over all $\mathbf{x}$ keeping $\mathbf{s}$ constant gives the mean correlation (BN)

$$\langle \delta A_1(\mathbf{x},\lambda_1)\delta A_2(\mathbf{x}+\mathbf{s},\lambda_2)\rangle = \lambda_1\lambda_2 \int \frac{d^2q}{(2\pi)^2} \, \tilde{f}_1(\mathbf{q},\lambda_1) \, \tilde{f}_2^{*}(\mathbf{q},\lambda_2) \, Q(\mathbf{q}) \, \exp(i\mathbf{q}\cdot\mathbf{s}) \qquad (2.5)$$

where

$$\tilde{f}_i(\mathbf{q},\lambda) = \int d^2r \, f_i(r,\lambda) \, \exp(-i\mathbf{q}\cdot\mathbf{r}) \qquad (2.6)$$

$$Q(\mathbf{q}) = <\frac{1}{\lambda_1}\tilde{\phi}(\mathbf{q},\lambda_1) \, \frac{1}{\lambda_2}\tilde{\phi}^{*}(\mathbf{q},\lambda_2) >/A \qquad (2.7)$$

$$\tilde{\phi}(\mathbf{q},\lambda) = \int d^2r \, \phi(\mathbf{x}+\mathbf{r},\lambda) \, \exp(-i\mathbf{q}\cdot\mathbf{r}) \qquad (2.8)$$

and the random phase approximation has been used for $\tilde{\phi}(\mathbf{q})$. Note that since $\phi \propto \lambda$, $Q(q)$ is wavelength-independent and is given by the power spectrum of density fluctuations. The $\tilde{f}_i$ for the various parameters are listed in Appendix A. Each of these is of the form

$$\tilde{f}_i \propto P_i(q) \, \cos^{\alpha}(\psi) exp(-\frac{1}{4} \, q^2\sigma^2) \qquad (2.9)$$

where $P_i(q)$ is a polynomial in q and $\mathbf{q}\cdot\mathbf{s} = q \, s \, cos(\psi)$. We compute the time averaged correlations (2.5) with $\lambda_1 = \lambda_2$. For an isotropic power law spectrum

$$Q(q) = Q_0 q^{-\beta} \qquad (2.10)$$

the angular integrals in equation (2.5) generate Bessel functions and the wave number integrals give functions of the form

$$h_n^{\alpha}(s) \equiv \int_0^{\infty}(q\,\sigma)^{(2n+3-\beta)} \exp(\frac{-q^2\sigma^2}{2}) \, J_{\alpha}(sq)d(q\,\sigma) \qquad (2.11)$$

$$= \left[\frac{s^2}{2\sigma^2}\right]^{\alpha/2} \frac{2^{(n+1-\beta/2)}}{\Gamma(\alpha+1)} \, \Gamma\left[n+2+\frac{(\alpha-\beta)}{2}\right] M\left[n+2+\frac{(\alpha-\beta)}{2},\alpha+1,-\frac{s^2}{2\sigma^2}\right]$$

where $M(a,b,x)$ is the confluent hypergeometric function (Abramowitz and Stegun 1970). The different angular factors of the Fourier transforms (2.9) will cause various

combinations of the $h_n^i(s)$ to appear, so we define the following linear combinations, g_n^i:

$$g_n^2 = \frac{1}{2} h_n^0 - \frac{1}{2} h_n^2$$

$$g_n^4 = \frac{3}{8} h_n^0 - \frac{1}{2} h_n^2 + \frac{1}{8} h_n^4 \tag{2.12}$$

and $g_n^0 = h_n^0$. At zero lag ($s = 0$) the g_n^i are proportional to h_n^0.

The mean auto- and cross-correlations of the fractional fluctuations are given by a dimensionless constant, K, characterising the strength of the scattering medium, multiplied by some linear combination of the $g_n^\alpha(s)$. The autocorrelation of the flux fluctuations, for instance, is

$$\langle \delta F(\mathbf{x}) \delta F(\mathbf{x+s}) \rangle = K g_1^0(s) \tag{2.13}$$

where

$$K = \frac{Q_0 \lambda^4 L^2}{2\pi \sigma^{(6-\beta)}} \qquad \beta \leq 4 \tag{2.14}$$

Similar expressions for other auto- and cross-correlations of interest are tabulated in Appendix A.

Correlations in wavelength can be derived by substituting the $\tilde{f}_i$ in equation (2.5) and evaluating with $\lambda_1 \neq \lambda_2$ and $\mathbf{s} = 0$. Several of the wavelength autocorrelations are also listed in Appendix A, where it is noted that the correlations are generally quite broad band.

The spot size on the screen, σ, can be calculated in terms of the assumed power spectrum, $Q_0 q^{-\beta}$. A simple order of magnitude estimate was given by BN who argued that the smallest scale phase fluctuation that scatters a ray is one that contributes a total phase change $\Delta\phi \gtrsim \pi$ (see also Gapper & Hewish 1981). The spectrum is thus truncated at an appropriate $q = q_{max}$ and the mean square angular size of the image is estimated to be

$$\frac{\sigma^2}{L^2} = \lambda^4 \int_{q_{min}}^{q_{max}} \frac{d^2q}{(2\pi)^2} q^2 Q(q) \tag{2.15}$$

The integral is cut off below $q_{min} = \sigma^{-1}$ because the corresponding spatial wavelengths are larger than σ and do not contribute to the image size. (For $\beta < 4$ q_{min} can be taken to be 0).

In this paper, we use the following more exact evaluation of the angular size derived in Appendix B.

$$\sigma = L \lambda \left[\frac{\Gamma((6-\beta)/2) \lambda^2 Q_0}{\pi(4-\beta)(\beta-2)\Gamma(\beta/2)} \right]^{1/(\beta-2)} \tag{2.16}$$

Thus, given a power spectrum, one can solve (2.16) for σ and substitute into (2.14) to obtain the normalization of the fluctuation magnitudes.

We now consider the wander in the position of the image on the sky. If we take the vector defined by the velocity $\mathbf{v}$ to denote the x direction, then, in a given realization of the screen, the angular displacement in the x direction of the image from its "true" time-averaged position is

$$\delta\theta_x \equiv \frac{\Delta\theta_x}{\sigma/L} = \frac{1}{\bar{F}} \frac{L}{\sigma} \int d^2r \ r_x \ I(\mathbf{r}, \mathbf{x}) \tag{2.17}$$

where r_x is the distance from the point $\mathbf{x}$ on the phase screen in the x direction and $\Delta\theta_x$ is normalized by the image half width, σ/L. Substituting from equation (2.1) and integrating by parts gives

$$\delta\theta_x = \left[\frac{-2L\lambda}{\pi\sigma^7} \right] \int d^2r \ \phi(\mathbf{r}+\mathbf{x}) \ r_x (3\sigma^2 - 2r^2) e^{-r^2/\sigma^2} \tag{2.18}$$

There will also be fluctuations in the direction transverse to $\mathbf{v}$. This displacement $\delta\theta_y$ is obtained by substituting r_y for r_x in equation (2.18).

An examination of figure 1 reveals that one can expect some correlation between the angular displacement of the spot, θ_x, and the *rate of change* of the received flux. In particular, when the spot is shifted farthest from its mean position, the flux will be varying most rapidly. Further, there will be a similar correlation between F and the rate of change of θ_x. To compute $\dot{F}$ and $\dot{\theta}$ we take the derivatives of (2.3) and (2.18) with respect to $\mathbf{x}$ and normalize by the refractive time scale $T_{ref} \ \sigma/v$. Thus

$$\dot{F} \equiv \frac{\sigma}{\bar{F}} \frac{\partial F}{\partial x} = \frac{4L\lambda}{\pi\sigma^5} \int d^2r \ \frac{d\phi}{dr_x} \ (r^2 - \sigma^2) e^{-r^2/\sigma^2}$$

$$i.e. \qquad f_{\dot{F}} = \frac{8L\lambda}{\pi\sigma^7} \ r_x (r^2 - 2\sigma^2) e^{-r^2/\sigma^2} \tag{2.19}$$

Similarly, we write the normalized rate of wander of θ_x as

$$\dot{\theta} \equiv \sigma \frac{\partial \theta}{\partial x} = \frac{-2L\lambda}{\pi\sigma^6} \int d^2r \frac{d}{dr_x}(\phi \, r_x(3\sigma^2 - 2r^2)e^{-r^2/\sigma^2})$$

$$i.e. \quad f_. = \frac{2L\lambda}{\pi\sigma^6}[3\sigma^2 - 2r^2 - r_x^2(10 - 4r^2/\sigma^2)]e^{-r^2/\sigma^2} \qquad (2.20)$$

Next we consider the expected elongation of the gaussian image. If one averages over a time long compared to σ/v the mean shape of a point source scatter-broadened by an *isotropic* ISM will be circular. Since the fluctuations of the spot's diameter in two orthogonal directions are independent, the rms elongation of the scattering disk can be non-zero for an instantaneous "snapshot" (*i.e.* single realization of $\phi(\mathbf{r})$). The spot will have some major axis with gaussian width $2\sigma_1$ and a minor axis of width $2\sigma_2$. The orientation of these axes will be random, but we can relate σ_1, σ_2 to the measured widths along fixed axes x and y via the relations

$$\sigma_1^2 + \sigma_2^2 = 2(\int r_x^2 + \int r_y^2) = 2\int r^2 \qquad (2.21)$$

$$\sigma_1^2\sigma_2^2 = 4[\int r_x^2 \cdot \int r_y^2 - (\int r_x r_y)^2] \qquad (2.22)$$

where we have introduced the following shorthand for intensity weighted averages

$$\int f(r_x, r_y) \equiv \int d^2r f(r_x, r_y)I(\mathbf{r}, \mathbf{x})$$

Then defining the elongation of the spot, e_s, as follows

$$e_s \equiv \frac{\Delta\sigma}{\sigma} = \frac{\sigma_1 - \sigma_2}{\sigma} \simeq \frac{(\sigma_1^2 - \sigma_2^2)}{\sigma_1^2 + \sigma_2^2} \qquad (2.23)$$

and substituting from (2.21) and (2.22) we have

$$e_s^2 = \left[\frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right]^2 = \frac{(\int(r_x^2 - r_y^2))^2 + (2\int r_x r_y)^2}{(\int r^2)^2}$$

$$= \frac{16L^2\lambda^2}{\pi^2\sigma^{16}}\left[\left[\int d^2r\,\phi(\mathbf{x},\mathbf{r})r^2(2\sigma^2 - r^2)e^{-r^2/\sigma^2}\right]^2\right.$$

$$- 4\left[\int d^2r\,\phi(\mathbf{x},\mathbf{r})r_x^2(2\sigma^2 - r^2)e^{-r^2/\sigma^2}\int d^2r\,\phi(\mathbf{x},\mathbf{r})r_y^2(2\sigma^2 - r^2)e^{-r^2/\sigma^2}\right]$$

$$+ 4\left.\left[\int d^2r\,\phi(\mathbf{x},\mathbf{r})r_x r_y(2\sigma^2 - r^2)e^{-r^2/\sigma^2}\right]^2\right] \qquad (2.24)$$

3 FLUCTUATIONS IN DIFFRACTIVE PHENOMENA

The flux received from a typical pulsar is found to be correlated at any instant over both a frequency interval, the decorrelation bandwidth ν_{dc}, and a range of time lags, the scintillation timescale t_s. Moreover, when plotted as dynamic scintillation spectra in the frequency-time plane, the patterns of enhanced flux do not display uncorrelated modulation in the two coordinates but instead often show an organized drifting behaviour with a typical drift slope of order a few kHz/sec (fig. 2). These frequency drifts are believed to be caused by refractive density perturbations on large scales, $\gtrsim \sigma$, pictured either as large prisms (Shishov 1974, Hewish 1980) or as gradients causing the interference of a few distinct bundles of rays (Roberts & Ables 1982).

To include diffractive effects in our scattering model we must explicitly deal with the phases of the individual rays received by the observer. We conceptually decompose the image into a large number of point scatterers, each located at a position of "stationary phase". The mean separation of these scatterers on the screen is a_{min}, giving $N \sim (\sigma/a_{min})^2$ scatterers. Since $N >> 1$ in the strong scattering limit, we may use statistical methods in the analysis of the interference of rays from these scattering centers. Consider a scatterer on the screen at a transverse distance r_j from the observer located at x. The phase advance of a wave propagating a distance L to the earth after being scattered at r_j consists of two parts: a free rotation over the distance propagated, $\theta_j = -(L + r_j^2/2L)/\lambda$, and an advance intrinsic to the phase screen, $\theta_j = \phi(x + r_j) \equiv \phi_j$. The *difference* in the total phase of a wave scattered from r_j on the phase screen when the observer moves a distance Δx and changes his observation frequency to $\lambda + \Delta\lambda$ is given by

$$\Delta\theta_j = \frac{r_{xj}}{L\lambda}\Delta x + \frac{(L + \frac{r_{xj}^2}{2L})}{\lambda^2}\Delta\lambda + \frac{\phi(x + r_j)}{\lambda}\Delta\lambda \tag{3.1}$$

where we have taken the observer velocity to be in the x direction and have used the fact that $\Delta\phi = (\phi/\lambda)\Delta\lambda$. The electric vector received by the observer from the point j is described by $E_j \propto e^{i\theta_j}$ and the total flux is $|\sum_j E_j|^2$. In the strong scintillation regime, we can assume that the θ_j are uniformly distributed and independent of one another. This is reasonable as the scales that dominate diffractive scintillation are much smaller than σ. We then have $\langle e^{i\theta_j} \rangle = 0$ and $\langle e^{i(\theta_j - \theta_k)} \rangle = \delta_{jk}$ where θ_j, θ_k denote the phases of two electric vectors received at x, λ from two point r_j, r_k. Let $\theta_j + \Delta\theta_j$, $\theta_k + \Delta\theta_k$ denote the phases from

the same points as seen at $x + \Delta x$, $\lambda + \Delta \lambda$, where $\Delta \theta_j$ is given by (3.1). The auto correlation of the flux is then

$$\langle F(x,\lambda) F(x + \Delta x, \lambda + \Delta \lambda) \rangle = F^2 \langle \frac{1}{N^2} \sum_{j,k,l,m} e^{i(\theta_j - \theta_k + \theta_l + \Delta \theta_l - \theta_m - \Delta \theta_m)} \rangle \qquad (3.2)$$

where for convenience we have assumed that the magnitudes of the electric vectors from the various scatters are the same. The $\Delta \theta_j$ of (3.1) have both a random part $\propto \phi$ and a deterministic part. As we are only interested in the $1/e$ width and not the details of the distribution it is reasonable to assume that the $\Delta \theta_j$ are distributed in a gaussian manner. Then, summing (3.2) in pairs and using $\langle \exp[i(\Delta \theta_j - \Delta \theta_k)] \rangle = \exp[-\langle (\Delta \theta_j - \Delta \theta_k)^2/2 \rangle]$, we can write the flux autocorrelation as

$$\langle F(x,\lambda) F(x + \Delta x, \lambda + \Delta \lambda) \rangle - F^2 = F^2 e^{-\langle (\Delta \theta_j - \Delta \theta_k)^2 \rangle/2} \qquad (3.3)$$

When the relative phases from the various scatters within the image fluctuate by ~ 1 radian the net intensity becomes decorrelated.

Thus, we need to estimate $\langle (\Delta \theta_j - \Delta \theta_k)^2 \rangle$ as a function of Δx and $\Delta \lambda$ in a given realisation of the phase screen. Let $\mathbf{r}_1, \mathbf{r}_2$ denote the points of origin at the screen of two rays received at $\mathbf{x}$ and let $(\mathbf{r}_1 - \mathbf{r}_2) \cdot \Delta \mathbf{x} = (x_1 - x_2) \Delta x$. We now define via equation (3.1)

$$f(r_1, r_2) \equiv (\Delta \theta_1 - \Delta \theta_2)^2 = \left[\frac{x_1^2 + x_2^2 - 2 x_1 x_2}{L^2 \lambda^2} \right] (\Delta x)^2 \qquad (3.4)$$

$$+ \left[\frac{r_1^4 + r_2^4 - 2 r_1^2 r_2^2}{4 L^2 \lambda^4} + \frac{\phi_1 (r_1^2 - r_2^2) + \phi_2 (r_2^2 - r_1^2)}{L \lambda^3} \right] (\Delta \lambda)^2$$

$$+ \left[\frac{x_1 (r_1^2 - r_2^2) + x_2 (r_2^2 - r_1^2)}{L^2 \lambda^3} + \frac{2 \phi_1 (x_1 - x_2) + 2 \phi_2 (x_2 - x_1)}{L \lambda^2} \right] \Delta x \, \Delta \lambda$$

where, as before, we have kept terms to linear order in ϕ. The number density of scatterers within the image is clearly proportional to $I(\mathbf{r}, \mathbf{x})$ defined in (2.2). Thus, to find the flux autocorrelation, (3.3), we must compute the intensity weighted average of $f(r_1, r_2)$ over all r_1, r_2 on the spot

$$\langle f(r_1, r_2) \rangle = \frac{\int d^2 r_1 \int d^2 r_2 \, I(\mathbf{r}_1, \mathbf{x}) I(\mathbf{r}_2, \mathbf{x}) \, f(r_1, r_2)}{\int d^2 r_1 \int d^2 r_2 \, I(\mathbf{r}_1, \mathbf{x}) I(\mathbf{r}_2, \mathbf{x})} \qquad (3.5)$$

We evaluate the integrals to lowest order in ϕ and use the fact that f is symmetric with respect to r_1, r_2 to write (3.5) in the form

$$\langle f \rangle = \langle (\Delta\theta_1 - \Delta\theta_2)^2 \rangle \equiv A\,(\Delta x)^2 + B\,(\Delta\lambda)^2 + C\,\Delta x\,\Delta\lambda \tag{3.6}$$

$$= \left[\frac{\sigma^2}{L^2\lambda^2} - \frac{4}{\pi L\,\sigma^4\lambda} \int d^2r\ \phi[2x^2(2-\frac{r^2}{\sigma^2}) - \sigma^2 + r^2]e^{-r^2/\sigma^2} \right](\Delta x)^2$$

$$+ \left[\frac{\sigma^4}{2L^2\lambda^4} + \frac{2}{\pi L\,\sigma^2\lambda^3} \int d^2r\ \phi[\frac{r^6}{\sigma^4} - \frac{5r^4}{\sigma^2} + 5r^2 - \sigma^2]e^{-r^2/\sigma^2} \right](\Delta\lambda)^2$$

$$+ \left[\frac{4}{\pi\sigma^2 L\,\lambda^2} \int d^2r\ \phi\,x\,[4 - \frac{7r^2}{\sigma^2} + \frac{2r^4}{\sigma^4}]e^{-r^2/\sigma^2} \right]\Delta x\,\Delta\lambda$$

From (3.3) we see that the autocorrelation of the intensity will fall to $1/e$ of its maximum value when $\langle f \rangle = 2$. Let us define the scintillation timescale t_s to correspond to the $1/e$ width along Δx of the flux auto-correlation. Thus, since t_s is the spatial coherence length of the diffraction pattern at Earth divided by $\mathbf{v}$ we have

$$t_s \equiv t_{s0}(1+\delta t_s) = \frac{\sqrt{2}L\lambda}{\sigma\upsilon}(1 + \frac{2L\lambda}{\pi\sigma^6} \int d^2r\ \phi[2x^2(2-r^2/\sigma^2) - \sigma^2 + r^2]e^{-r^2/\sigma^2}) \tag{3.7}$$

where t_{s0} is the mean scintillation timescale and δt_s is its fractional fluctuation. Similarly, the $1/e$ half width of the diffraction pattern in λ can be converted to units of frequency to find the instantaneous decorrelation bandwidth ν_{dc}

$$\nu_{dc} = \nu_{dc0}(1+\delta\nu_{dc}) = \frac{Lc}{\pi\sigma^2}(1 - \frac{2L\lambda}{\pi\sigma^{10}} \int d^2r\ \phi[r^6 - 5r^4\sigma^2 + 5r^2\sigma^4 - \sigma^6]e^{-r^2/\sigma^2}) \tag{3.8}$$

Since the mean pulse broadening is described by the time constant $\tau_0 = \sigma^2/2LC$ (cf. BN), we see that

$$2\pi\,\nu_{dc}\,\tau_0 = 1 \tag{3.9}$$

This "uncertainty relation" has been verified observationally in the case of the Vela pulsar (Slee et al. 1980). Furthermore, the fluctuation $\delta\nu_{dc}$ is the exact negative of the fractional fluctuation in the pulse broadening $\delta\tau$ (cf. (A.4) and (A.9)). So we find that the uncertainty relation holds even for the fluctuations about the mean. This encourages us in believing that our simple scattering model can indeed be applied to diffractive phenomena.

· The rotated ellipse described by (3.6) represents an "average" frequency drift pattern for the given realization of the phase screen ϕ or, equivalently, the shape of the 2-d autocorrelation function in the (ν,t) plane. It is of interest to calculate the expected tilt or drift slope that an observer would measure. The angle of tilt will clearly depend on our scaling of the Δx and $\Delta \lambda$ axes; accordingly, we normalise by σ and λ, respectively. Then, taking ω to be the rotation from the orientation in the absence of refractive effects we write the drift slope, m_d, as

$$m_d \equiv \tan\omega = \frac{\Delta x}{\Delta \lambda} \frac{\lambda}{\sigma} = \left[\frac{d\nu}{dt} \cdot \frac{\sigma}{\nu \mathbf{v}} \right]^{-1}$$

In terms of the coefficients A, B, C in (3.6)

$$m_d = \left[\frac{\sigma^2 A - \lambda^2 B}{\sigma \lambda C} \right] \left[-1 + \left[1 + (\frac{\sigma \lambda C}{\sigma^2 A - \lambda^2 B})^2 \right]^{1/2} \right]$$

$$\simeq \frac{4L\lambda}{\pi \sigma^5} \int d^2r \; \phi \; x \; [4 - 7r^2/\sigma^2 + 2r^4/\sigma^4] e^{-r^2/\sigma^2} \tag{3.10}$$

The appearance of a dynamic scintillation spectrum is also characterized by the average elongation of its drift bands, e_d. This is a measure of the prominence of the frequency drift phenomenon, since a circular pattern in the (ν,t) plane has no well defined slope. As e_d again depends on the normalization chosen, we make an unambiguous definition by calculating the elongation at a fixed drift slope of $\pi/4$. Since observers tend to record the drift patterns with the most conspicuous drift, this will facilitate comparison with the data. We use the same definition as for e_s in section 2; if the gaussian drifting band has a semi-minor axis a and a semi-major axis b, then

$$e_d \simeq \frac{b^2 - a^2}{b^2 + a^2} \tag{3.11}$$

We can eliminate a and b in favour of A, B and C to obtain

$$e_d = \frac{[(A - B)^2 + C^2]^{\frac{1}{2}}}{A + B}$$

Fixing the drift slope at $\pi/4$ with our normalization corresponds to setting $\sigma A = \lambda B$ and letting C go to $C \cdot (\sigma A / \lambda B)^{1/2}$. Thus to first order in ϕ, we find that

$$e_d = 2^{3/2} \frac{L\lambda}{\pi \sigma^5} \int d^2r \; \phi \; x \; (4 - \frac{7r^2}{\sigma^2} + \frac{2r^4}{\sigma^4}) e^{-\frac{r^2}{\sigma^2}} \tag{3.12}$$

4 CORRELATIONS

To compute the auto- and cross-correlations of the above quantities, we take the Fourier transforms $\tilde{f}$ of the expressions for the fluctuations ((2.18), (2.19), *etc.*) and substitute into (2.5). These transforms, tabulated in Appendix A, can be grouped into two classes:

I) Curvature-induced fluctuations (those $\tilde{f}$ that are real and proportional to even powers of q)

i.e. δF, $\delta \Omega$, δt, $\delta \tau$, $\delta \nu_{dc}$, θ, δt_s

II) Gradient-induced fluctuations (those $\tilde{f}$ that are imaginary and proportional to odd powers of q)

i.e. $\dot{F}$, $\delta \theta_x$, m_d, e_d

In addition there are $\delta \theta_y$ and e_s^2 which do not belong to either class. Class I quantities correspond to those effects which are caused by focusing or defocusing *lenses*. Class II effects, on the other hand, are caused by *prisms* which steer the wavefronts. In general, one can expect Class I and Class II quantities to co-vary among themselves, but not with each other. The strength of these cross-correlations is discussed in section 7 and Table 2. The autocorrelations of the refractive fluctuations of the various observables that we have studied are shown in figures 3 and 4 for a power-law spectrum with $\beta = 11/3$ (expressions are given in Appendix A). We note that the variations with lag can be quite different from one another. However, most of the curves have half-widths that are characteristically of order σ, as would be expected for a refractive effect. For a combined Earth-pulsar velocity of v this corresponds to the refractive timescale

$$T_{ref} = \frac{\sigma}{v} \tag{4.1}$$

However, $\delta \theta_x$, $\delta \theta_y$ and m_d decorrelate over a time $\gtrsim 10\sigma/v$, a consequence of the incipient divergence in these quantities as $\beta \to 4$.

The Fourier transform of $f_{e_s^2}$, (A.8) has, as expected, no angular dependence. It is, in fact, identical to the square of $\tilde{f}_\Omega$, which means that the rms elongation of the spot is equal to the normalised rms fluctuation in the angular size. This is because the diameter fluctuations in orthogonal directions in the image are independent. Formally, e_s^2 will

correlate with the parameters of group I, but as this is a higher order effect, we do not calculate it. θ_y will be uncorrelated with any of the other quantities.

We also note that $\tilde{f}_{\dot{F}}\tilde{f}_{\theta_x} = -\tilde{f}_{\dot{\theta}}\tilde{f}_F$ so that the cross-correlations of these two pairs will be equal but opposite. This implies that in the $(\delta F, \delta\theta_x)$ plane a set of pulsar observations will statistically follow elliptical trajectories with a fixed sense of rotation (clockwise). Thus, an observer can gain $\sqrt{2}$ in signal-to-noise by measuring the *curl* of this field: $\Delta\theta\cdot\dot{F} - \dot{\theta}\cdot\Delta F$.

Certain other pairs of Fourier transforms are related, as well. We have already noted that $\tilde{f}_\tau = -\tilde{f}_{\nu_{dc}}$ and so the correlation functions of the latter follow immediately from those of the former. For frequency drifts, the rms elongation at fixed slope is seen to be half the rms value of the slope, m_d, so the average stretching of the drift slope bands is directly related to their inclinations. This is, however, a simple geometrical effect.

5 DENSITY FLUCTUATION SPECTRA WITH $\beta > 4$

So far, we have considered spectra with $\beta \leqslant 4$ for which the arguments of section 1 show that the scattering is dominated by the small scale density perturbations. Recent work (*e.g.* BN, GN, Hewish, *et al.* 1985) has indicated that spectral indices of greater than the critical value $\beta=4$ may also be relevant to electron density perturbations in the ISM. In this regime, the scattering is dominated by the large spatial scales and, in the absence of an outer scale, the rms value of the bending angle $\eta(\mathbf{r})$ of (2.1) diverges. Accordingly, the linearization of the exponential in (2.2), which is central to the earlier development, is no longer valid. However, most of the correlations computed in the previous section are finite even for $\beta>4$. This suggests that the divergence of $\eta(r)$ may be removable by a suitable modification of the theory.

For concreteness, we consider the variations in the flux, F. The autocorrelation function $C_{FF}(s)$ has a zero-lag magnitude $\sim K$ and has a half-width $s_{1/2}\sim\sigma$. This means that we are rarely interested in correlating observers separated by more than $\sim\sigma$. The mean bending angle seen by two such observers will be large but will be dominated by large scale perturbations ($q^{-1}\gg\sigma$) that contribute the same bending angle for both observers. This common steering will be indistinguishable from a shift in the image position. The physically interesting quantity, the *difference* in the bending between the two observers (equivalently the phase curvature or focusing of the screen) is, however, finite

and reasonably small as shown below. Therefore we should obviously measure the bending angle with respect to some mean bending, η_0, common to the two observers. (We can take η_0 to be the refractive bending by the screen at a point half-way between the observers.) Equation (2.2) can now be written

$$I(\mathbf{r},\mathbf{x}) = \frac{\overline{F}}{\pi\sigma^2} \exp(-[\frac{L(\eta-\eta_o)+r}{\sigma}]^2) \tag{5.1}$$

$$\simeq \frac{\overline{F}}{\pi\sigma^2} \left[1 - \frac{2L\,\mathbf{r}\cdot(\eta-\eta_o)}{\sigma^2} + \varepsilon \right] e^{-r^2/\sigma^2}$$

where $\varepsilon \propto (\eta-\eta_0)^2(L/\sigma)^2$. If we assume that ε is small compared to 1, we then have as before

$$\delta F = \frac{-2L}{\pi\sigma^4} \int d^2r \; (\eta(\mathbf{r}+\mathbf{x})-\eta_o)\cdot\mathbf{r} \; e^{-r^2/\sigma^2}$$

Since η_0 is a constant, terms proportional to it vanish by symmetry. Substituting (2.1) and integrating by parts, we recover (2.3) as before.

We must now show that ε in (5.1) can, in fact, be neglected. To do this we compute $\langle(\eta(\mathbf{s})-\eta_o)^2\rangle$ for $s \sim \sigma/2$ (half the observer separation). From (2.1), (2.9) we have

$$\eta(\mathbf{s}) = \frac{-\lambda}{(2\pi)^2} \int d^2q \, (iq)\tilde{\phi}(q)e^{iq\cdot s} \tag{5.2}$$

In computing $\langle(\eta(\mathbf{s})-\eta_o)^2\rangle$, we should include only the effect of the refractive scales since the small scales $< \sigma$ have been counted in determining the spot size (Appendix B). (One can equivalently note that all the $\tilde{f}_i$ have a weighting factor $\exp(-1/4\,q^2\sigma^2)$ which effectively damps out the small scales). We thus have

$$\frac{L^2}{\sigma^2}\langle(\eta(s)-\eta(0))^2\rangle = 2K(\frac{s}{\sigma})^{(\beta-4)}\int_0^\infty x^{3-\beta}(1-J_0(x))e^{-\frac{x^2}{2(s/\sigma)^2}}dx$$

$$= K\, 2^{(4-\beta)/2}\Gamma\left[\frac{4-\beta}{2}\right] M\left[\frac{4-\beta}{2},1,-\frac{s^2}{2\sigma^2}\right] \tag{5.3}$$

For $\beta>4$ we evaluate (2.15) between $q=\sigma^{-1}$ and ∞ to obtain $K = \beta-4$. We then find that (5.3) is small compared to 1 for $s\lesssim\sigma/2$ so long as $\beta\lesssim5$. Since the range of β of interest to us is $3.5\lesssim\beta\lesssim4.5$ (c.f. GN) we are justified in neglecting ε in the expansion of (5.1). Similar arguments show that for each of the correlations computed above that remains finite for $\beta>4$

(*i.e.* $C_{\Omega\Omega}$, $C_{\Omega t_s}$, *etc.*) we may continue to use the expressions derived previously for $\beta < 4$ as long as $s \lesssim \sigma$.

Certain correlations, however, are formally divergent for $\beta \leq 4$. For example the position shift, $\delta\theta$, depends directly on the bending angle and thus diverges at $\beta = 4$. A steep spectrum (in the absence of some outer scale) will cause the image to wander arbitrarily far from its true position. An observer, however, must estimate the true position by the mean over his observation period, $T_{obs} = s/v$. The relevant measure of the amplitude of refractive position fluctuations is then the fluctuation across the duration of the observation ,*i.e.*

$$\langle [\delta\theta(s) - \delta\theta(0)]^2 \rangle = 2[C_{\theta\theta}(0) - C_{\theta\theta}(s)] \tag{5.4}$$

where $C_{\theta\theta}$ is the autocorrelation function of $\delta\theta$ (A.18). Although $C_{\theta\theta}$ diverges, the difference (5.4), equivalent to a first order structure function in η, is finite for $\beta < 6$. When $s \gg \sigma$, we can use (2.5) and (A.6) to obtain the following approximate estimate

$$\langle [\delta\theta_x(s) - \delta\theta_x(0)]^2 \rangle \sim \frac{\lambda^4 L^2 Q_0}{2\pi\sigma^2} \int\limits_{2\pi/s}^{1/\sigma} dq \; q^{3-\beta}$$

$$= K \ln(s/2\pi\sigma), \qquad \beta = 4 \tag{5.5}$$

$$= \frac{K}{(\beta-4)} \left[\frac{s}{2\pi\sigma} \right]^{\beta-4}, \qquad \beta > 4$$

It is seen explicitly that, when $\beta > 4$, the image wander diverges in the limit of large baselines s. In practice, of course, the power-law spectrum (2.11) that we have considered will have a physical cut-off at some q_{min} and so the image wander will saturate for $s \gtrsim 2\pi/q_{min}$.

The arrival time of pulses from a pulsar has a random delay whose dominant component is proportional to the mean phase fluctuation of ϕ averaged over the spot size σ on the scattering screen. This formally diverges for $\beta \geq 2$. We note, however, that observers measure pulse arrival time residuals only after fitting a low-order polynomial model of intrinsic pulsar behaviour, $\Delta t = \alpha_1 + \alpha_2 T + \alpha_3 T^2$, as well as sinusoidal components of period 1 year to refine the position and proper motion of the pulsar. Consequentially, the *post-fit* arrival time residual Δt_{pf} is finite for all $\beta < 8$. Blandford, Narayan & Romani (1984) have considered the effect of parameter-fitting on post-fit residuals. For the

timing noise contributed by phase fluctuations in the ISM, we can apply their results in combination with (A.3) to obtain

$$\langle \delta t^2 \rangle = \frac{4 \lambda^4 L^2 Q_0}{2\pi\sigma^4} \int_0^\infty dq \ q^{1-\beta}(1-\frac{1}{2}q^2\sigma^2+\frac{1}{8}\dot{q}^4\sigma^4)^2 \exp(-\frac{1}{2}q^2\sigma^2)T(q) \qquad (5.6)$$

where the transmission or filter function $T(q)$ is defined in the above-mentioned paper. If the observations extend over a time $T_{obs} \gg 1$ year and if $vT_{obs} \gg \sigma$, then we can use Table 1 of Blandford *et al.* (1984) to simplify (5.6)

$$\langle \delta t^2 \rangle = \frac{4 \lambda^4 L^2 Q_0}{2\pi\sigma^4} \int_{6.9/s}^\infty dq \ q^{1-\beta} = \frac{4K}{\beta-2} \left[\frac{s}{6.9 \ \sigma} \right]^{\beta-2}, \qquad \beta \sim 4 \qquad (5.7)$$

where the lower limit $q_{\min} = 6.9/s$ in the integral is appropriate for $\beta \sim 4$ (*i.e.* spectral index ~ 3 in Blandford *et al.*). The divergences associated with arbitrarily large electron density perturbations are thus absorbed into the timing model. A similar treatment will excise the divergences in the cross-correlations of δt and other parameters.

The frequency drift slope $d\nu/dt$ has, however, a divergence which cannot be removed by the above techniques. The drift slope is directly proportional to the phase gradient on the screen just as the position shift $\delta\theta$. However, unlike $\delta\theta$, the true mean value of $d\nu/dt$ is known *a priori* to be 0. Thus, in the absence of an outer scale, a $\beta>4$ spectrum will cause frequency drifts with arbitrarily large slope and this would be observed even in a single epoch of observations. The finite observed drift slopes therefore place limits on the perturbation spectrum as discussed in section 8.

6 SCATTERING BY A THICK SCREEN

In the development so far we have assumed that the source is distant and that all the scattering is localised within a single thin screen at a distance L from the observer. In many circumstances (*e.g.* interplanetary scintillation) this will be a good approximation. It is, however, important to understand the changes that are introduced if the scattering is shared between several screens or indeed distributed uniformly along the line of sight to the source. Fortunately, the present formalism allows us to treat these cases as well.

Suppose that there are n phase screens between the source and the observer. Let the strength of the fluctuations on these screens be Q^i and the associated scattering angles be ρ_i (equal to σ/L in the 1 screen case). The separation between screen i and

screen j is denoted L_{ij} with $i=0$ signifying the source and $i=n+1$ the observer. The distant source case is recovered by taking the limit $L_{0i} \to \infty$.

Now consider a ray propagating from the source to the observer and undergoing angular deflections ξ_i at each of the n screens (Fig. 5). We can relate the transverse position vectors $\mathbf{r}_i$ of the ray's intersection with the screens to the ξ_i through the recursion relation

$$\xi_i = \frac{\mathbf{r}_{i+1}-\mathbf{r}_i}{L_{i\ i+1}} - \frac{\mathbf{r}_i-\mathbf{r}_{i-1}}{L_{i-1\ i}}$$

We set $\mathbf{r}_0=0$ and denote the observer position by $\mathbf{r}_{n+1}=\mathbf{x}$ to obtain

$$\mathbf{r}_i = \frac{L_{0i}}{L_{0n+1}} \mathbf{x} - \sum_{j=1}^{n} M_{ij}\ \xi_j \tag{6.1}$$

where
$$M_{ij} = \begin{cases} \dfrac{L_{0j}\ L_{i\ n+1}}{L_{0\ n+1}} & 1 \leq j \leq i \leq n \\[2ex] \dfrac{L_{0i}\ L_{j\ n+1}}{L_{0\ n+1}} & 1 \leq i \leq j \leq n \end{cases} \tag{6.2}$$

One can change from the variables $\mathbf{r}_i$ to the variables ξ_i via the Jacobian

$$\frac{\partial(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n)}{\partial(\xi_1, \xi_2, \cdots, \xi_n)} = \left[\frac{L_{01}}{L_{0\ n+1}}\right]^2 \prod_{i=1}^{n} L_{i\ i+1}^2 \tag{6.3}$$

Let $I_i(\mathbf{r}_i, \theta_i)$ be the intensity leaving the i'th screen at an angle θ_i to the reference direction. The intensity leaving the $(i+1)$'th screen can then be written

$$I_{i+1}(\mathbf{r}_{i+1}, \theta_{i+1}) = \int \frac{d^2 r_i}{L_{i\ i+1}^2}\ G_{i+1}\left[\theta_{i+1} - (\frac{\mathbf{r}_{i+1}-\mathbf{r}_i}{L_{i\ i+1}});\mathbf{r}_{i+1}\right] I_i\left[\mathbf{r}_i, (\frac{\mathbf{r}_{i+1}-\mathbf{r}_i}{L_{i\ i+1}})\right] \tag{6.4}$$

where $G_i[\xi_i,\mathbf{r}_i]$ is the scattering kernel. Using the linear approximation outlined in section 2, we write

$$G_i[\xi_i,\mathbf{r}_i] = \left[1 + \lambda\frac{\partial\phi_i}{\partial\mathbf{r}_i}\cdot\frac{\partial}{\partial\xi_i}\right]\frac{e^{-\xi_i^2/\rho_i^2}}{\pi\rho_i^2} \tag{6.5}$$

If the mean flux at the position of the first screen from a point source is denoted $\overline{F}_1$, the intensity at the observer plane can be written formally

$$I_n = \int \frac{d^2 r_{n-1}}{L_{n-1\ n}^2}\ G_n \int \cdots \int \frac{d^2 r_1}{L_{1\ 2}^2}\ G_2 \cdot G_1 \overline{F}_1 \tag{6.6}$$

where the integration must be carried out along a ray connecting the observer to the source.

Let us initially ignore refractive effects. The mean flux received by the observer is

$$\overline{F} \equiv \overline{F}_{n+1} = \int \frac{d^2 r_n}{L_{n\,n+1}^2} I_n$$

$$= \int \prod_{i=1}^{n} \left\{ \frac{d^2 r_i}{\pi L_{i\,i+1}^2 \rho_i^2} \right\} \overline{F}_1 e^{-\sum_{j=1}^{n} \xi_j^2 / \rho_j^2} = \overline{F}_1 \left[\frac{L_{0\,1}}{L_{0\,n+1}} \right]^2 \tag{6.7}$$

where we have incorporated the Jacobian from equation (6.3). This is just the inverse square law. Similarly, the mean angular size of the observed image is

$$\langle \theta^2 \rangle = \frac{1}{\overline{F}} \int \frac{d^2 r_n}{L_{n\,n+1}^2} \theta_n^2 I_n \tag{6.8}$$

where θ_n, the angle of the ray incident on the observer plane at $\mathbf{x}$, is given by

$$\theta_n = \frac{\mathbf{r}_1}{L_{0\,1}} + \xi_1 + \cdots + \xi_n - \frac{\mathbf{x}}{L_{0\,n+1}} = \sum_{i=1}^{n} \frac{L_{0\,i}}{L_{0\,n+1}} \xi_i \tag{6.9}$$

So, using equations (6.3) and (6.7) we write the mean angular size of the image as

$$\langle \theta^2 \rangle = \int \prod_{i=1}^{n} \left[\frac{d^2 \xi_i e^{-\xi_i^2 / \rho_i^2}}{\pi \rho_i^2} \right] \cdot \left[\sum_{k=1}^{n} \frac{L_{0\,k}}{L_{0\,n+1}} \xi_k \right]^2 = \sum_{i=1}^{n} \left[\frac{L_{0\,i}}{L_{0\,n+1}} \right]^2 \rho_i^2 \tag{6.10}$$

This is, as expected, the weighted sum of the individual scattering angular widths. Equation (6.10) agrees in the continuum limit with equation (A2) of BN.

Now we introduce refractive effects by including the phase fluctuations on the screens. The perturbation to the intensity can be written by combining equations (6.5) and (6.6) as

$$\delta I_n = \lambda \overline{F}_1 \int \prod_{i=1}^{n-1} \left\{ \frac{d^2 r_i}{\pi L_{i\,i+1}^2 \rho_i^2} \right\} \sum_{j=1}^{n} \left[\frac{\partial \phi_j}{\partial \mathbf{r}_j} \frac{\partial}{\partial \xi_j} \right] \frac{e^{-\sum_{k=1}^{n} \xi_k^2 / \rho_k^2}}{\pi \rho_n^2} \tag{6.11}$$

In the spirit of the earlier development, we express the phases ϕ_i as Fourier transforms and integrate by parts using equation (6.2) to obtain an expression for the normalised flux fluctuation

$$\delta F(x) = \frac{1}{\overline{F}} \int \frac{d^2 r_n}{L_{n\,n+1}^2} \, \delta I_n \;=\; \frac{\lambda \overline{F}_1}{\overline{F}} \int \prod_{i=1}^{n} \left[\frac{d^2 r_i \, e^{-\xi_i^2/\rho_i^2}}{\pi L_{i\,i+1}^2 \, \rho_i^2} \right] \sum_{j=1}^{n} \left[M_{jj} \frac{\partial^2 \phi_j}{\partial \mathbf{r}_j^2} \right]$$

$$= - \frac{\lambda \overline{F}_1}{\overline{F}} \int \prod_{i=1}^{n} \left\{ \frac{d^2 r_i \, e^{-\xi_i^2/\rho_i^2}}{\pi L_{i\,i+1}^2 \, \rho_i^2} \right\} \sum_{j=1}^{n} \int \frac{d^2 q_j}{(2\pi)^2} \, e^{i\mathbf{q}_j \cdot \mathbf{r}_j} \, M_{jj} q_j^2 \tilde{\phi}_j \qquad (6.12)$$

Equation (6.12) is the multi-screen generalization of equation (2.4). We obtain the generalization of the flux correlation function (eq. (2.5) with $\lambda_1 = \lambda_2$) by averaging over $\mathbf{x}$

$$\left\langle \delta F(\mathbf{x}) \delta F(\mathbf{x}+\mathbf{s}) \right\rangle = \lambda^2 \sum_{i=1}^{n} \int \frac{d^2 q_i}{(2\pi)^2} \, \tilde{f}_F^i(\mathbf{q}_i) \tilde{f}_F^{i*}(\mathbf{q}_i) Q^i(\mathbf{q}_i) e^{i\mathbf{q}_i \cdot \mathbf{s}(L_{0i}/L_{0\,n+1})} \qquad (6.13)$$

where

$$\tilde{f}_F^i(\mathbf{q}_i) = - \left[\frac{L_{0\,n+1}}{L_{0\,1}} \right]^2 \lambda M_{ii} q_i^2 \int \prod_{j=1}^{n} \left[\frac{d^2 r_j}{\pi L_{j\,j+1}^2 \rho_j^2} e^{-\xi_j^2/\rho_j^2 - i M_{ij} \mathbf{q}_i \cdot \xi_j} \right] \qquad (6.14)$$

We can change from the variables $\mathbf{r}_j$ to the variables ξ_j using the Jacobian (6.3) and carry out the integrations to obtain

$$\tilde{f}_F^i(\mathbf{q}_i) = - \lambda \frac{L_{0\,i} L_{i\,n+1}}{L_{0\,n+1}} \, q_i^2 \exp\left(- \frac{1}{4} q_i^2 \sigma_i\right) \qquad (6.15)$$

where σ_i, the effective size of the scattering disk at screen i, is defined by

$$\sigma_i^2 = \sum_{j=1}^{n} M_{ij}^2 \rho_j^2 \qquad (6.16)$$

If we have n evenly spaced similar screens and denote the distance from the source to the observer by D, then we can use equation (6.10) to obtain

$$\sigma_i^2 = \frac{i(n+1-i)[2i(n+1-i)+1]D^2 \langle \theta^2 \rangle}{n(2n+1)(n+1)^2} \qquad (6.17)$$

Of greater interest is the continuum limit. We denote the distance along the line of sight from the source by z and the mean scattering rate by $\Psi(z) = \langle \Delta \rho^2 / \Delta z \rangle$. Then letting $n \to \infty$ we obtain

$$\sigma^2(z) = \left[\frac{D-z}{D} \right]^2 \int_0^z dz' z'^2 \Psi(z') + \frac{z^2}{D^2} \int_z^D dz' (D-z')^2 \Psi(z') \qquad (6.18)$$

and

$$\langle \delta F(\mathbf{x}) \delta F(\mathbf{x+s}) \rangle = \frac{\lambda^4}{D^2} \int_0^D dz \, (D-z)^2 z^2 \int \frac{d^2q}{(2\pi)^2} \, q^4 \frac{dQ}{dz} \, e^{i\mathbf{q \cdot s}(z/D)} e^{\frac{-1}{2}q^2\sigma^2} \qquad (6.19)$$

We have therefore expressed the autocorrelation function for the flux fluctuations as a sum (or an integral) over the scattering screens. To proceed further, we must substitute an expression for the spectrum of density fluctuations. For the spectrum in (2.10)

$$\frac{dQ}{dz} = q^{-\beta} \frac{dQ_0}{dz} \qquad (6.20)$$

Substituting in (6.13) the n-screen case gives

$$\langle \delta F(0) \delta F(s) \rangle = \frac{\lambda^4 \, \Gamma(3-\beta/2)}{\pi 2^{(\beta/2-1)}} \sum_{i=1}^n \frac{M_{ii}^2 \, Q_0^i}{\sigma_i^{6-\beta}} \, M\left[\frac{6-\beta}{2}, 1, \frac{-s^2}{2(1-x)^2\sigma_i^2} \right] \qquad (6.21)$$

while for a uniform medium we obtain

$$\langle \delta F(0) \delta F(s) \rangle = \frac{3K \, \Gamma(3-\beta/2)}{2^{2-\beta/2}} \int_0^1 dx \, [x(1-x)]^{\beta-4} M\left[\frac{6-\beta}{2}, 1, \frac{-s^2}{2(1-x)^2\sigma_1^2} \right] \qquad (6.22)$$

where K is given by the single screen value (2.14) and $\sigma_1 = \theta_{rms}D/2$ as for a single screen.

In Figure 6 we show the flux auto correlation function for a uniform scattering medium with $\beta=4$ and compare this with the results for 1, 3, and 5 equally spaced screens. The 1 screen case corresponds to the equivalent screen approximation introduced in BN and used in the earlier part of the present paper. We see that the rms refractive flux fluctuation at zero lag from a uniform medium is larger by a factor $\sqrt{3}$ than was predicted by the equivalent screen. In the case of a Kolmogorov spectrum, $\beta=11/3$, the corresponding factor is 2.3, while for $\beta=4.3$ the fluctuations are 1.4 times the result for the equivalent thin screen. For these three spectra the flux autocorrelation functions for an extended medium and single thin screen are shown in Figure 7.

We have also calculated the angular size fluctuation autocorrelation and the cross-correlation with flux fluctuations. For a uniform medium with $\beta=4$ their expectation values at zero lag are 7/8 and 3/4, respectively, of the single screen correlations.

7 NUMERICAL RESULTS

We now calculate the normalization, K, and numerical estimates of the fluctuations in various observables for three power law models of the interstellar medium; $\beta=11/3$, $\beta=4$ and $\beta=4.3$. In the following we specialize to the *single screen* equivalent of a uniform distribution of inhomogeneities between the source and the observer. (The theory of section 6 for an extended medium could in principle be used to obtain more accurate estimates, but the present status of the observations does not warrant such calculations). For an extragalactic point source well out of the galactic plane one should replace $C_{-4}D$ by $3C_{-4}H \, csc(b)$ and D by $2H \, csc(b)$ in the expressions below and in Table 1, where b is the galactic latitude of the source and H is the scale height of the inhomogeneities in kpc (see Appendix A of BN for details).

We first assume that the power spectrum of phase fluctuations on the scattering screen has a Kolmogorov power law form

$$Q(q) = Q_0 \, q^{-11/3}, \qquad Q_0 = 3.7 \times 10^{-18} C_{-4}D \; cm^{-11/3} \qquad (7.1)$$

which corresponds to $C_N^2 = 10^{-4}C_{-4}m^{-20/3}$ in the notation of Armstrong *et al.* (1981)(*cf.* BN). D is the distance to the pulsar measured in kpc. The amplitude Q_0 differs from that used by BN because they used an approximate estimate of the image angular size and needed to adjust Q_0 suitably to fit the observations. We avoid this by using the improved angular size estimate given in (2.16). The scaling of C_N^2 has been selected such that the parameter C_{-4} has a value ~ 1 for nearby pulsars. However, C_{-4} can be as large as $\sim 10^4$ for distant pulsars in the plane of the galaxy and for the particular case of the Vela pulsar (Manchester & Taylor 1977, Cordes, Weisberg & Boriakoff 1984) and in fact is $>10^5$ for the radio source at the galactic center (Lo *et al.* 1985). We substitute (7.1) and (2.16) into equation (2.14) to obtain

$$K = 1.27 \times 10^{-2} C_{-4}^{-2/5} \lambda^{-17/15} D^{-11/15} \qquad (7.2)$$

where we measure the wavelength λ in meters.

We next consider the "critical" or "gaussian" spectrum with $\beta=4$. Here we fix the normalization constant Q_0 by requiring that the calculated angular size be the same as that for a Kolmogorov spectrum when $C_{-4} = \lambda = D = 1$:

$$Q(q) = Q_0 q^{-4}, \quad Q_0 = 1.6 \times 10^{-21} C_{-4}D \; cm^{-4} \qquad (7.3)$$

Substituting in equations (2.14) and (2.15), where we retain the lower limit, we obtain

$$K = \frac{1}{\ln(430C_{-4}\lambda^3 D^2/K^{1/2})} = \frac{1}{7\gamma} \qquad (7.4)$$

where the correction factor γ is unity for $C_{-4}=\lambda=D=1$ and has only a weak logarithmic dependence on C_{-4}, λ and D.

Finally, we consider a spectrum with $\beta=4.3$. For this value the scalings of various observables with λ and D are quite close to those of the Kolmogorov spectrum so that a $\beta=4.3$ spectrum is equally compatible with scintillation observations as $\beta=11/3$ (GN). The refractive effects, however, will be much larger for $\beta=4.3$. For $\beta>4$, the lower cut-off at $q_{min}\sim\sigma^{-1}$ will dominate the integral (2.15). Taking $q_{max}=\infty$, we can solve for σ

$$\sigma = \left[\frac{Q_0}{2\pi(\beta-4)}\right]^{1/(6-\beta)} \chi^{4/(6-\beta)} L^{2/(6-\beta)} \qquad (7.5)$$

If we again require that the calculated σ equal that for the Kolmogorov spectrum when $C_{-4} = \lambda = D = 1$, we find that

$$Q(q) = Q_0 q^{-4.3}, \quad Q_0 = 3.6\times10^{-25}C_{-4}D \text{ cm}^{-4.3} \qquad (7.6)$$

Substituting in (2.14) we find

$$K = (\beta-4) = .3 \qquad (7.7)$$

We use these three normalizations and the autocorrelation functions listed in Appendix A to calculate the magnitudes of the zero-order quantities and their rms fluctuations for the three spectra. These values, along with their scalings as a function of C_{-4}, λ, and D are listed in Table I. As discussed in section 5, certain quantities are formally divergent for large values of β. For these, we therefore list approximate magnitudes of the fluctuations about the *observed* mean as a function of the observation period in years, T_y.

The *normalized* cross-correlation of two parameters A and B is defined to be

$$C_{AB} = \frac{\langle A\,B\rangle}{\langle A^2\rangle^{1/2}\langle B^2\rangle^{1/2}} \qquad (7.8)$$

Class I (curvature-induced) fluctuations will correlate with one another to varying degrees, but will not be correlated with Class II (gradient-induced) fluctuations and vice-

versa. We list the non-divergent Class I and II cross-correlations in Table 2.

8 DISCUSSION

In the preceding sections, we have extended the computations of BN to include several more potentially observable effects arising from long wavelength density fluctuations in the interstellar medium. We now have the theoretical machinery to estimate the magnitudes and timescales of refractive effects for virtually any observable parameter in terms of any power-law spectrum of density fluctuations, including those with $\beta > 4$. We have also outlined the extension from a single screen to an extended medium and have shown how this can introduce significant changes.

Since the theory depends on a simple linearised model of the scattering (equation 2.2), we should address the question of the reliability of the theoretical predictions. Fortunately, Goodman and Narayan (1985) have presented exact results for the flux fluctuations produced by a single screen for both $\beta < 4$ and $\beta > 4$. A comparison between their results and those of our approximate theory is made in Appendix C. We find that the agreement is extremely good for $3.5 \lesssim \beta \lesssim 4.5$; this encourages us to believe that the other computed correlation functions are also quite accurate.

The detection of any of the fluctuations predicted by our theory, particularly the cross-correlations, would confirm the importance of propagation effects for the long timescale variability of pulsars and compact extragalactic radio sources. The predicted magnitudes of the fluctuations are relatively small in the case of the Kolmogorov spectrum ($\beta = 11/3$), increasing with the observation frequency. On the other hand, if $\beta \geq 4$, the fluctuations are relatively large but independent of λ. The shapes of the auto- and cross-correlations also depend on the value of β and upon whether the density fluctuations are restricted to a thin screen or are distributed throughout an extended scattering medium (see figure 7). Thus, observations of refractive effects promise to be a sensitive probe of the spectrum of ISM density perturbations as well as the distribution of the scattering irregularities along the line of sight.

Perhaps the easiest observations to make will be those that include fluctuations in the scintillation timescale, δt_s, the decorrelation bandwidth, $\delta \nu_{dc}$, and the flux ,δF. Since the fluctuation timescale T_{ref} decreases as one moves to shorter wavelengths (Table 1), observations for a relatively short period at the highest frequency allowed by the

multipath propagation condition, *viz.* $\sigma > a_{min}$, should be sufficient to detect the predicted correlations. We note that the thin screen theory predicts normalised cross-correlation coefficients of between 50% and 75% (*cf.* Table 2), so the effects are large.

It is difficult to resolve pulsars with conventional VLBI. At meter wavelengths there are some candidates for observation of angular size fluctuations, but the timescales will be rather long. Another possible VLBI experiment involves measuring the relative separation of pairs of pulsars close enough in the sky to be contained within the same primary beam of a radio telescope ($\sim 1°$), *e.g.* PSR 2016+28 and PSR 2020+28. This observation should be carried out at two or more low frequencies and it seems possible to achieve positional accuracies $\sim 0.1\lambda/b \sim 1$mas. The detection of a refraction-induced shift may be possible in a steep spectrum, though the variation timescale will again be quite long. In fact, it may be easier to monitor the rate of position fluctuation $\dot{\theta}$, which increases with observing frequency. Correlations with $\delta\nu_{dc}$ and δt_s are quite strong. To detect $C_{\dot{\theta}\cdot t_s}$, for example, one desires a moderately dispersed, nearby pulsar with a high space velocity. PSR1818-04 seems to be a good candidate with sufficient flux for an intermediate frequency VLBI experiment. Other possible pulsars are tabulated in BN. The correlation between $\dot{\theta}$ and F is however too small to be observationally interesting for β close to 4.

Frequency drifts in dynamic scintillation spectra provide one of the best probes of large scale density fluctuations (Hewish 1980, Roberts & Ables 1982, Smith & Wright 1985, Hewish *et al.* 1985). The correlations of the drift slope with $\dot{F}$ and $\delta\theta$ are probably too small to be measured. However, the *magnitudes* of the drifts and their scaling with β, λ, D, and C_{-4} can be compared with our theory. Smith & Wright (1985) have measured the drift slopes of 32 pulsars. They present their data in the form

$$\frac{d\nu}{dt} = \frac{\nu\upsilon}{\theta_r D} \, sec\,\phi \qquad (8.1)$$

where θ_r is the rms refractive bending angle and ϕ is the angle between the plane of maximum dispersion (*i.e.* the orientation of the ISM prism) and the pulsar velocity υ. For 24 pulsars in which an independent measurement of υ is available, they define $\theta_r = m\,\theta_s$ where θ_s is the rms scattering angle (*i.e.* σ/L), and estimate the value of $|mcos(\phi)|$, a measure of the relative importance of long and short wavelength perturbations in the ISM. Noting that $\theta_s D = 2\sigma$, we can compute $|mcos(\phi)|$ directly in our model, using the

measured distances and scattering strengths for these 24 pulsars. For $\beta = 11/3$ we obtain $|m cos(\phi)| = .40$, somewhat larger than the observed value of .24. Thus, the observed magnitude of refractive fluctuations in frequency drifts is *smaller* than that predicted by the Kolmogorov spectrum. On the contrary, the observed elongations of the drift patterns seem to be significantly larger than the value $e_d \lesssim .1$ expected for a Kolmolgorov spectrum, although observational bias towards the most prominent examples may be reflected in the published spectra. Further, the observed flux variations are also larger than predicted for a Kolmogorov spectrum (BN,GN). These conflicting indications could mean that the continuous Kolmogorov power-law spectrum commonly assumed is too simplistic. They may also reflect deficiencies in the thin screen model.

As further evidence that a single extended power-law spectrum with $\beta \lesssim 4$ is insufficient for the explanation of all scintillation phenomena, we consider the observation of periodicities in the spacing of the drift bands. Striking examples of quasi-periodic frequency drifts in dynamic scintillation spectra have been presented by Hewish, Wolszczan & Graham (1985). In these instances the patterns are interpreted as arising from the interference of a few, well-separated bundles of rays which have passed through an image-scale dispersive wedge on their way from the pulsar to the observer plane (*cf.* also Ewing *et al.* 1970, Roberts & Ables 1982). Hewish *et al.* go on to argue that the effective value of β can exceed 4. This attractive physical picture may, however, be difficult to realize in an extended power-law spectrum as the inhomogeneities intermediate between the diffractive scale and the spot size will generally break the image into too many beams to give the observed patterns. An alternative possibility is that the small scale irregularities are absent and the spectrum has an inner scale somewhat smaller than the size of the image. This would create a few caustics which could give the observed periodic modulation. A possible prediction of this idea is that the periodicities of dynamic scintillation spectra should only be found at frequencies where the spot size is comparable to this inner scale and that this frequency should be larger for the more distant and more highly scattered pulsars.

Another implication of the observed quasi-periodicities is that a snapshot image of the pulsar would reveal a few bright blobs within the time-averaged spot (figure 2). GN have argued that such a "fractal" geometry for the image is expected in theories with $\beta > 4$. This may be testable with VLBI on selected pulsars. The extended periodicities seen would probably still be rather rare unless the effective β were close to 6. However, a

spectrum with $\beta>4$ predicts extremely large values for the average drift slope (in fact, m_d will technically diverge) unless the spectrum cuts off at lengths not much longer than the refractive scale σ. To maintain the large amplitude refractive fluctuations indicated by other observations(*e.g.* flux), it might be necessary to impose an inner scale as well and invoke the focusing effects of caustics. (We note that there may already be evidence for caustics in the cusplike peaks in pulsar intensity fluctuation records *e.g.* Cole *et al.* 1970, Helfand *et al.* 1977). The resulting spectrum of the ISM density perturbations would thus be severely truncated at both ends, containing only a limited power-law regime. GN showed that the λ and D scalings of observed quantities are relatively unaffected by the absence of short scales for $\beta>4$. If, however, $\beta<4$, then the absence of short wavelength fluctuations drives the scaling laws towards the "gaussian spectrum" case, *i.e.* $\beta=4$.

It is possible, for a given value of the spectral index β and a given distribution of the scattering inhomogeneities, to estimate an upper bound on the outer scale for the power-law spectrum from the observed angular broadening and the physical constraint that the amplitude of the electron density fluctuations on this scale still be linear (*i.e.* $\delta n < n$). For spectra with $\beta \leq 4$ the scattering is dominated by the smallest scale consistent with the strong scintillation condition, $a_{min} \sim \lambda/\theta(a)$. To allow for the possibility that the spectrum cuts off at an inner scale somewhat larger than this we define $a_{min} \sim \alpha\lambda/\theta(a_{min})$ with $\alpha \gtrsim 1$. If one has an independent estimate of the total number of scattering electrons, for example from the dispersion measure DM in the case of pulsars, one can use the scaling $\delta n(a) \propto a^{(\beta-3)/2}$ (section 1) to estimate the scale at which the perturbation spectrum must become nonlinear, *i.e.* $\delta n \sim n$. If we consider a source at distance z, a scattering screen of thickness L at z_0 and use the typical pulsar observables $\nu_{50} \equiv \nu_{dc}$(in units of 50 kHz) and DM_{30} (in units of 30 pc cm^{-3}), we find for a Kolmogorov spectrum

$$a_{nl} \simeq 4.2 \times 10^{20} \text{cm} \left[DM_{30}^3 L^{-1.5}(\frac{z_0}{z})^{1.75}[\nu_{50}(z-z_0)]^{1.25}\alpha^{-.5}\lambda^{5.5} \right] \qquad (8.2)$$

where λ is in meters and all other lengths are in kpc. This is an upper bound for the outer scale of the power-law spectrum. For spectra with $\beta>4$ the scattering is dominated by fluctuation scales on the order of the spot size $\sigma = \theta_{rms}(z-z_0)$. We can again scale the fluctuation strength with the scale size to find an upper bound on the outer scale for a $\beta=4.3$ spectrum

$$a_{nl} \simeq 1.5 \times 10^{16} \mathrm{cm} \left[DM_{30}^{1.54} L^{-.77} (\frac{z_0(z-z_0)}{z})^{.88} \nu_{50}^{.65} \lambda^{3.1} \right] \qquad (8.3)$$

If a_{nl} is less than the spot size σ, then refractive fluctuations cannot be important for that pulsar. For Vela and a few other pulsars, (8.3) is actually a significant constraint, as the bulk of the scattering is believed to be provided by the local effect of the Gum nebula. For Vela, $a_{nl} \sim \sigma$ when β=4.3, so refractive effects such as frequency drifts are likely to be quite restricted in such a steep spectrum.

Propagation-induced fluctuations can also be significant for sources other than pulsars. Rickett, *et al.* (1984) suggested that the phenomenon of low frequency variability of extragalactic radio sources can, in many cases, be explained as a propagation effect. The timescales inferred from a spectrum of refractive fluctuations are compatible with those observed. Our theory predicts that the flux variations will be correlated with position shifts and angular size variations. Although the predicted magnitudes will be small, detection of this covariance using VLBI would allow a critical test of the ISM modulation hypothesis. There is the further possibility that the flicker of extragalactic radio sources (Heeschen 1984) could again be an effect of the ISM (Rickett *et al.* 1984). These variations may, however, also be intrinsic to the sources. In this context, it is worth noting that interstellar refraction by density irregularities should not affect the direction of linear polarization observed from pulsars and the compact components of extragalactic radio sources.

Observations by Lo *et al.* (1985) show that the VLBI resolved core of the galactic center has a diameter$\sim 2.1 mas$ at λ=1.35cm and scales with the observation wavelength as $\sim \lambda^2$, suggesting scatter-broadening. The source appears to be elongated at 3.6 cm with an axial ratio of 1.8:1 which corresponds to an elongation parameter $e_s \simeq .54$. From the source broadening we can estimate the scattering strength to be $C_{-4} \sim 5 \times 10^5$(for β=4). Then, using this value, we estimate the expectation value of e_s at λ=3.6cm for an isotropic scattering medium and find it to be $\lesssim .1$ unless β is somewhat greater than 4. We note, however, that observations of other sources near the galactic center indicate significantly lower scatter-broadening. Hence, the bulk of the scattering medium is probably within $\sim$100pc of the galactic center (Backer, private communication). Estimating the total number of electrons along the line of sight via the observed total extinction, one can use (8.2) to show that refractive effects in the galactic center must be very small. As further confirmation for the unimportance of refractive effects for this source, we note that the

data of Backer & Sramek (1982) place a limit of < 10 mas on the wander of the source over a 5 year baseline. Moreover, they find the centroids of the images at 3.6 cm and 11 cm to agree within 10mas, indicating that there are no large scale "prisms" in the line of sight. A possibility one should consider is that the scattering medium in the vicinity of the galactic center could be strongly anisotropic, as in the model of Higdon (1984), in which case the image spot would be elongated in the ratio of the scattering strengths along the two principal axes. For magnetic fields stretched in the plane of the Galaxy by differential rotation, the long axis of the image should be perpendicular to the galactic plane, as observed. A distinction between this picture and the random elongation we have considered (in section 2) is that the position angle of the elongation will not change as a function of time for the anisotropic medium, whereas in our theory it is expected to do so on a timescale $\sim T_{ref} \sim \sigma/\mathbf{v}$. A second epoch of observations separated by $\gtrsim T_{ref}$ would be helpful in clarifying this question.

In conclusion, we urge that future single dish observations of radio pulsars include accurate measurements of the mean flux and the parameters $\tau, \delta\nu_{dc}, \delta t_s, m_d$, and e_d which characterise the scintillation properties. In addition we advocate a modest simultaneous VLBI program of observation of pulsars such as PSR 1818-04 designed to resolve the scatter-broadened image and detect position wander. Successful detections of the predicted correlations would, in addition to determining which variations in pulsars and extragalactic radio sources are intrinsic, also yield valuable data on the interstellar turbulence spectrum of particular relevance to theories of cosmic ray propagation. They would also motivate further calculations using the techniques outlined in this paper.

ACKNOWLEDGEMENTS

REFERENCES

Abramowitz, M. & Stegun, J.A., 1970. *Handbook of Mathematical Functions*, Dover Publications, New York.

Armstrong, J.W., Cordes, J.M. & Rickett, B.J., 1981. *Nature*, **291** , 561.

Backer, D.C. & Sramek, R.A., 1982.*Astrophys. J.*, **260**, 512.

Blandford, R. & Narayan, R., 1984. *Proceedings of the Workshop on Millisecond Pulsars*, Reynolds, S.P. and Stinebring, D.R. , eds., Green Bank, Va.

Blandford, R. & Narayan, R., 1985. *Mn. Not. R. astr. Soc.* **213**, 591 (BN).

Blandford, R., Narayan, R. & Romani R.W., 1984. *J. Astrophys. Astr.*, **5**, 369.

Cordes, J.M., Weisberg, J. & Boriakoff, V., 1984. *Astrophys. J.*, **288**, 221.

Cole, T.W. , Hesse, H.K. & Page C.G., 1970. *Nature*, **225**, 712.

Ewing, M.S., Batchelor, R.A., Friefeld, R.D., Price R.M. & Staelin, D.H., 1970. *Astrophys. J. Lett.* **162**, L169.

Gapper, G.R. & Hewish, A., 1981. *Mon. Not. R. astr. Soc.* **197**, 209.

Goodman, J. & Narayan, R., 1985. *Mon. Not. R. astr. Soc.*, in press (GN).

Heeschen, D.S. 1984. *Astron. J.*, **89**, 1111.

Helfand, D.J, Fowler, L.A. & Kuhlman, J.V., 1977. *Astron. J.*, **82**, 701.

Hewish, A., 1980. *Mon. Not. R. astr. Soc.*, **192**, 799.

Hewish, A., Wolszczan, A. & Graham, D.A., 1985. *Mon. Not. R. astr. Soc.*, **213**, 167.

Higdon, J.C., 1984. *Ap.J.*, **285**, 109.

Lee ,L.C. & Jokipii, J.R. 1975. *Astrophys. J.*, **201**, 532.

Lo, K.Y., Backer, D.C., Ekers R.D., Kellermann, K.I., Reid M. & Moran, J.M., 1985. *Nature*, in press.

Manchester, R.N. & Taylor, J.H., 1977. *Pulsars*, Freeman, San Francisco.

Mutel, R.L., Broderick, J.J., Carr, T.D., Lynch, M., Desch, M., Warnock, W.W. & Klemperer, W.K., 1974. *Astrophys. J.*, **193**, 279.

Rickett, B.J., 1977. *Ann. Rev. Astron. Astrophys.*, **15**, 479.

Rickett, B.J., Coles, W.A. & Bourgois, G., 1984. *Astron. Astrophys.*, **134**, 390.

Roberts, J.A. & Ables, J.G., 1982. *Mon. Not. R. astr. Soc.*, **201**, 1119.

Scheuer, P.A.G. 1968. *Nature*, **218**, 920.

Shapirovskaya, N. Ya., 1978. *Soviet Astron*, **22**, 544.

Shishov, V.I., 1974. *Sov. Astron.*, **17**, 598.

Sieber, W., 1982. *Astr. Astrophys.*, **113**, 311.

Slee, O.B., Dulk, G.A. & Otrupcek R.E., 1980. *Proc. Aust. Ast. Soc.*, **4**(1), 100.

Smith, F.G. & Wright, N.C., 1985. *Mon. Not. R. astr. Soc.*, *in press*.

APPENDIX A

The $\tilde{f}_i$ described by (2.7) are listed below for the various observables that we have considered ,$viz.$ flux F, angular size Ω, pulse arrival time t, pulse width τ, time derivative of flux $\dot{F}$, position shift $\delta\theta_x$, time derivative of position shift $\dot{\theta}$, spot elongation e_s^2, decorrelation bandwidth ν_{dc}, scintillation timescale t_s, drift slope m_d and elongation of drift pattern e_d.

$$\tilde{f}_F = -\frac{\lambda L}{\sigma^2}q^2\sigma^2\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.1}$$

$$\tilde{f}_\Omega = -\frac{\lambda L}{\sigma^2}(q^2\sigma^2 - \frac{1}{4}q^4\sigma^4)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.2}$$

$$\tilde{f}_t = 2\frac{\lambda L}{\sigma^2}(1 - \frac{1}{2}q^2\sigma^2 + \frac{1}{8}q^4\sigma^4)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.3}$$

$$\tilde{f}_\tau = -3/2\frac{\lambda L}{\sigma^2}(q^2\sigma^2 - \frac{1}{3}q^4\sigma^4 + \frac{1}{48}q^6\sigma^6)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.4}$$

$$\tilde{f}_{\dot{F}} = \frac{\lambda L}{\sigma^2}(iq_x\sigma)q^2\sigma^2\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.5}$$

$$\tilde{f}_{\delta\theta_x} = -\frac{\lambda L}{\sigma^2}(iq_x\sigma)(1 - \frac{1}{2}q^2\sigma^2)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.6}$$

$$\tilde{f}_{\dot{\theta}} = -\frac{\lambda L}{\sigma^2}(q_x^2\sigma^2)(1 - \frac{1}{2}q^2\sigma^2)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.7}$$

$$|\tilde{f}_{e_s^2}|^2 = \frac{\lambda^2 L^2}{\sigma^4}(q^4\sigma^4 - \frac{1}{2}q^6\sigma^6 + \frac{1}{16}q^8\sigma^8)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.8}$$

$$\tilde{f}_{\nu_{dc}} = 3/2\frac{\lambda L}{\sigma^2}(q^2\sigma^2 - \frac{1}{3}q^4\sigma^4 + \frac{1}{48}q^6\sigma^6)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.9}$$

$$\tilde{f}_{t_s} = \frac{\lambda L}{\sigma^2}q_x^2\sigma^2(1 - \frac{1}{4}q^2\sigma^2)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.10}$$

$$\tilde{f}_{m_d} = -4\frac{\lambda L}{\sigma^2}(iq_x\sigma)(1 - \frac{5}{8}q^2\sigma^2 + \frac{1}{16}q^4\sigma^4)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.11}$$

$$\tilde{f}_{e_d} = 1/\sqrt{2}\,\tilde{f}_{m_d} = -2\sqrt{2}\frac{\lambda L}{\sigma^2}(iq_x\sigma)(1 - \frac{5}{8}q^2\sigma^2 + \frac{1}{16}q^4\sigma^4)\,e^{-\frac{1}{4}q^2\sigma^2} \tag{A.12}$$

We substitute these expressions into (2.5) and evaluate according to (2.11) and (2.12) to obtain the various correlations. The autocorrelation functions are

$$\langle \delta F(\mathbf{x})\delta F(\mathbf{x+s})\rangle = K\, g_1^{\,0} \tag{A.13}$$

$$\langle \delta\Omega(\mathbf{x})\delta\Omega(\mathbf{x+s})\rangle = K\left(g_1^{\,0} - \frac{1}{2}g_2^{\,0} + \frac{1}{16}g_3^{\,0}\right) \tag{A.14}$$

$$\langle \delta t(\mathbf{x})\delta t(\mathbf{x+s})\rangle = K\left(g_{-1}^{\,0} - g_0^{\,0} + \frac{1}{2}g_1^{\,0} - \frac{1}{8}g_2^{\,0} + \frac{1}{64}g_3^{\,0}\right) \tag{A.15}$$

$$\langle \delta\tau(\mathbf{x})\delta\tau(\mathbf{x+s})\rangle = \frac{9}{4}K\left(g_1^{\,0} - \frac{2}{3}g_2^{\,0} + \frac{11}{72}g_3^{\,0} - \frac{1}{72}g_4^{\,0} + \frac{1}{2304}g_5^{\,0}\right) \tag{A.16}$$

$$\langle \dot F(\mathbf{x})\dot F(\mathbf{x+s})\rangle = K\, g_2^{\,2} \tag{A.17}$$

$$\langle \delta\theta_x(\mathbf{x})\delta\theta_x(\mathbf{x+s})\rangle = K\left(g_0^{\,2} - g_1^{\,2} + \frac{1}{4}g_2^{\,2}\right) \tag{A.18}$$

$$\langle \delta\theta_y(\mathbf{x})\delta\theta_y(\mathbf{x+s})\rangle = \frac{1}{2}K\left(h_0^{\,0} + h_0^{\,2} - h_1^{\,0} - h_1^{\,2} + \frac{1}{4}h_2^{\,0} + \frac{1}{4}h_2^{\,2}\right) \tag{A.19}$$

$$\langle \dot\theta(\mathbf{x})\dot\theta(\mathbf{x+s})\rangle = K\left(g_1^{\,4} - g_2^{\,4} + \frac{1}{4}g_3^{\,4}\right) \tag{A.20}$$

$$\langle \delta(\nu_{dc})(\mathbf{x})\delta(\nu_{dc})(\mathbf{x+s})\rangle = \frac{9}{4}K\left(g_1^{\,0} - \frac{2}{3}g_2^{\,0} + \frac{11}{72}g_3^{\,0} - \frac{1}{72}g_4^{\,0} + \frac{1}{2304}g_5^{\,0}\right) \tag{A.21}$$

$$\langle \delta t_s(\mathbf{x})\delta t_s(\mathbf{x+s})\rangle = K\left(g_1^{\,4} - \frac{1}{2}g_2^{\,4} + \frac{1}{16}g_3^{\,4}\right) \tag{A.22}$$

$$\langle \delta\tan\omega(\mathbf{x})\delta\tan\omega(\mathbf{x+s})\rangle = 2\langle e_d(\mathbf{x})e_d(\mathbf{x+s})\rangle$$

$$= 16K\left(g_0^{\,2} - \frac{5}{4}g_1^{\,2} + \frac{33}{64}g_2^{\,2} - \frac{5}{64}g_3^{\,2} + \frac{1}{256}g_4^{\,2}\right) \tag{A.23}$$

Equation (A.15) diverges for $\beta \geq 2$ and (A.18), (A.19) and (A.23) for $\beta \geq 4$. These are discussed in section 5. The others are convergent for $\beta < 6$, but are reliable only for $\beta \lesssim 5$.

Autocorrelations with respect to changes in the observation frequency may also be of interest and we give these for certain parameters in (A.24) to (A.27). The frequency behaviour of position shifts are of interest since VLBI measurements can be performed at several frequencies. Fluctuations in ν_{dc} and t_s should also be accessible over a moderate frequency range.

$$\langle \delta\dot F(\lambda_1)\delta\dot F(\lambda_2)\rangle \propto \frac{\lambda_1^2\lambda_2^2\,\sigma_1\sigma_2}{(\sigma_1^2+\sigma_2^2)^{(6-\beta)/2}} \tag{A.24}$$

$$\langle\delta\theta(\lambda_1)\delta\theta(\lambda_2)\rangle \propto \frac{\lambda_1^2\lambda_2^2}{\sigma_1^2\sigma_2^2} \frac{1}{(\sigma_1^2+\sigma_2^2)^{(4-\beta)/2}} \left[\beta - 3 + (6-\beta)(4-\beta)\frac{\sigma_1^2\sigma_2^2}{(\sigma_1^2+\sigma_2^2)^2}\right] \tag{A.25}$$

$$\langle\delta\nu_{dc}(\lambda_1)\delta\nu_{dc}(\lambda_2)\rangle \propto \frac{\lambda_1^2\lambda_2^2}{(\sigma_1^2+\sigma_2^2)^{(6-\beta)/2}} \left[\frac{2\beta-9}{3}\right.$$

$$+ \frac{(6-\beta)(8-\beta)}{6(\sigma_1^2+\sigma_2^2)^2} \left\{\frac{(\sigma_1^4+\sigma_2^4)}{2} + \frac{(\beta-2)}{3}\sigma_1^2\sigma_2^2\right.$$

$$\left.\left.+ \frac{1}{24}\frac{\sigma_1^4\sigma_2^4}{(\sigma_1^2+\sigma_2^2)^2}(10-\beta)(12-\beta)\right\}\right] \tag{A.26}$$

$$\langle\delta t_s(\lambda_1)\delta t_s(\lambda_2)\rangle \propto \frac{\lambda_1^2\lambda_2^2}{(\sigma_1^2+\sigma_2^2)^{(6-\beta)/2}} \left[2\beta - 4 + (8-\beta)(6-\beta)\frac{\sigma_1^2\sigma_2^2}{(\sigma_1^2+\sigma_2^2)^2}\right] \tag{A.27}$$

where $\sigma_{1,2}=\sigma(\lambda_{1,2})$ and $\sigma\propto\lambda^{\beta/(\beta-2)}$ for $\beta\leqq4$ and $\sigma\propto\lambda^{4/(6-\beta)}$ for $\beta\geqq4$. The constants of proportionality are given by the corresponding spatial autocorrelations evaluated at zero lag, ($s = 0$). Numerically, we find that for $\beta=11/3$ the correlations reach their half power points at the following values of λ_1/λ_2: $\dot{F}$, (.6,1.4); $\delta\theta$, (.6,1.8); ν_{dc}, (.6,1.4); t_s, (.6,1.4). For $\beta=4.3$ the half power points are: $\dot{F}$, (.7,1.8); ν_{dc}, (.6,2.5); t_s, (.6, 2.3). The fluctuation $\delta\theta$ grows arbitrarily large for $\beta>4$. In general, the correlations are seen to be quite broad band.

Normalised cross-correlations are given in Table 2.

APPENDIX B

We wish to formalise the separation of the perturbation spectrum into diffractive and refractive regimes, leading respectively to angular broadening and refractive steering of the image of a point source. Let us imagine that we image the source with a gaussian aperture of full-width W. We assume that W is intermediate between the diffractive scale, a_{min}, and the Fresnel scale, $r_F=\sqrt{\lambda L}$. (Note that, in the strong scintillation regime, $a_{min} \ll r_F \ll \sigma$.) The angular amplitude of the signal received at the aperture is

$$\Phi(\alpha) = \frac{1}{\lambda}\int\int d^2x \exp[-i\frac{\alpha x}{\lambda} - \frac{2x^2}{W^2}]\,\mathbf{E}(x) \tag{B.1}$$

where $\mathbf{E(x)}$ is the instantaneous electric vector measured at the point x on the observer

plane. If we represent the source by a plane wave incident on the phase screen at a distance L and account for the phase rotations of the electric vectors received from different positions r on the phase screen, we can write this as

$$\Phi(\alpha) \propto \int\int d^2x\, d^2r\, \exp[-i\frac{\alpha x}{\lambda} - \frac{2x^2}{W^2} + i(\phi(r) - \frac{(r-x)^2}{2L\lambda})]$$ (B.2)

Since $W << r_F$, we neglect the term in $x^2/L\lambda$. Integrating over x we then obtain

$$\Phi(\alpha) \propto \int d^2r\, \exp[\,i\phi(r) - \frac{ir^2}{2L\lambda} - \frac{(r-\alpha L)^2 W^2}{8L^2\lambda^2}\,]$$ (B.3)

We see that $\Phi(\alpha)$ is dominated by a region on the screen around the point $r = \alpha L$ of width $r_{tr} \sim \sqrt{8}L\lambda/W$. We now handle separately the phase fluctuations $\phi_<(r)$ due to scales smaller than r_{tr} and those $\phi_>(r)$ due to scales larger than r_{tr}. By Taylor expanding $\phi_>$ about $r = \alpha L$, we can write the argument of the exponent in (B.3) as

$$\chi(r) = i[\phi_<(r) + \phi_>(\alpha L) + \phi'_>(\alpha L)(r - \alpha L)$$

$$+ \frac{1}{2}\phi''_>(\alpha L)(r-\alpha L)^2 - r^2/(2L\lambda)\,] - \frac{W^2}{8L^2\lambda^2}(r-\alpha L)^2$$ (B.4)

The angular intensity is then given by

$$I(\alpha) = |\Phi(\alpha)|^2 \propto \int\int d^2r\, d^2r'\, e^{(\chi(r) + \chi^*(r'))}$$

We now introduce new coordinates $u = r + r' - 2\alpha L$, $v = r - r'$ and perform the integral over u. Since there are many diffractive scales α_{min} within the range of integration, we can ensemble average over $\phi_<$. Thus

$$I(\alpha) \propto \int d^2v\, e^{-\frac{1}{2}D_{\phi<}(v)}\, e^{-\frac{i}{\lambda}(\alpha - \lambda\phi'_>)v}\, e^{-[\frac{W^2}{16L^2\lambda^2} + \frac{1}{W^2}(1 - \lambda L\phi''_>)^2]v^2}$$ (B.5)

where $D_{\phi<}(v) \equiv <[\phi_<(0) - \phi_<(v)]^2>$ is the phase structure function at lag v due to scales smaller than r_{tr}.

Noting that the first exponential in (B.5) cuts off at $v \sim \alpha_{min}$, we now show that, over this range of v, the third exponential can be set to 1 with negligible error. For $W < r_F$, the first term in the argument of the third exponential is $\sim W^2v^2/r_F^4 < \alpha_{min}^2/r_F^2 \ll 1$. Next, we have $v^2/W^2 \sim \alpha_{min}^2/W^2 \ll 1$, provided $W \gg \alpha_{min}$ as already assumed. Finally, noting that $\alpha_{min} \sim (\lambda^2 Q_0)^{-1/(\beta-2)}$ and $\phi''_>(r_{tr}) \sim \lambda^2 Q_0 r_{tr}^{\beta-6}$, we have

$$\frac{\hbar^2 L^2 \phi_>''^2 \alpha_{\min}^2}{W^2} \sim \left[\frac{\alpha_{\min}W}{r_F^2}\right]^{4-\beta} \ll 1$$

for the assumed range of W. Thus, (B.5) simplifies to

$$I(\alpha) \propto \int d^2v \; e^{-\frac{1}{2}D_{\phi<}(v)} \; e^{-\frac{i}{\lambda}(\alpha - \lambda\phi_>')v} \tag{B.6}$$

Let us first neglect the contribution from $\phi_>'$ in the second exponential. $I(\alpha)$ is then given by the Fourier transform of $exp\left[-\frac{1}{2}D_{\phi<}(v)\right]$. Now, $D_{\phi<}(v)$ varies as $v^{\beta-2}$ which, for $\beta \sim 4$, is not very different from v^2. We can therefore conveniently approximate $I(\alpha)$ by a gaussian

$$I(\alpha) = \frac{\bar{F}}{\pi\theta_0^2} \exp(-\frac{\alpha^2}{\theta_0^2}) \tag{B.7}$$

where the beam-width θ_0 can be determined in terms of the $1/e$ width of $exp[-\frac{1}{2}D_{\phi<}(v)]$ to be

$$\theta_0 = \lambda \left[\frac{\Gamma((6-\beta)/2)\hbar^2 Q_0}{\pi(4-\beta)(\beta-2)\Gamma(\beta/2)}\right]^{1/(\beta-2)} \tag{B.8}$$

This leads to equation (2.16) for $\sigma = \theta_0 L$. When the contribution from $\phi_>'$ is included in (B.6), the angular intensity received is modified to

$$I(\alpha) = \frac{\bar{F}}{\pi\theta_0^2} \exp\left[-\frac{(\alpha - \lambda\phi_>'(\alpha L))^2}{\theta_0^2}\right] \tag{B.9}$$

which is equivalent to equation (2.2).

APPENDIX C

To test the accuracy of the approximations in our formalism, we compare our results with the exact results obtained by Goodman & Narayan (1985) for the flux fluctuations produced by a thin scattering screen for power-law spectra with $2<\beta<6$. From (2.6) we see that the spectrum of the correlation $\langle \delta A_1(x)\delta A_2(x+s)\rangle$ is given by $\hbar^2 \tilde{f}_1(q)\tilde{f}_2^*(q)Q(q)$. For flux fluctuations, we substitute from (A.1) to obtain for a power-law spectrum of index β

$$\tilde{C}_{FF}(q) = Q_0\hbar^4 L^2 q^{4-\beta}\exp(-\frac{1}{2}q^2\sigma^2) \tag{C.1}$$

In comparison, GN give for $\beta < 4$,

$$\check{C}_{FF}(q) \equiv \check{W}(q) = Q_0 \chi^4 L^2 q^{4-\beta} \exp\left[-\left[\frac{q}{q_{ref}}\right]^{\beta-2}\right] \qquad (C.2)$$

We note that the normalization as well as the power-law index below the cut-off is identical. Moreover, the cut-off scales in the two formulae, *viz.* $\sqrt{2}/\sigma$ and q_{ref}, are also exactly equal showing that our approximate theory is extremely accurate for $\beta < 4$. The form of the cut-off is gaussian in our theory because we have made the simplifying assumption of a gaussian image whereas the exact result has the true spot shape. For $\beta \lesssim 4$, however, the difference is small.

When $\beta > 4$, our theory again predicts the form (C.1) and GN still give

$$\check{C}_{FF}(q) = Q_0 \chi^4 L^2 q^{4-\beta}, \qquad q < q_{ref} \qquad (C.3)$$

so that the form and normalization of the spectrum below the cut-off continue to be in perfect agreement. However, whereas in (C.1) we have a gaussian cut-off at q_{ref}, the true spectrum has a second power-law regime $\check{C}_{FF} \propto q^{-4/(\beta-4)}$, $q_{ref} < q < q_{int}$ out to an intermediate scale q_{int}. This region of the spectrum arises from the patchy "fractal" nature of the image and is filtered out in our gaussian approximation of the spot shape. Its contribution to the flux variance is, however, quite small. Thus the rms fluctuation of flux predicted by the present theory is

$$\delta F_{rms} = [Kh_1(0)]^{1/2} = (\beta-4)^{1/2} \, 2^{(2-\beta/2)} \Gamma(3-\beta/2) \qquad (C.4)$$

whereas the exact result given by Goodman and Narayan is

$$\delta F_{rms} = [2\sqrt{\beta-3}/(6-\beta) - 1]^{1/2} \qquad (C.5)$$

A numerical comparison of the two formulae confirms that the agreement is quite good up to $\beta \sim 4.5$.

We thus find that the approximate theory that we have developed is in very good agreement with more exact calculations in the regimes of interest. The advantage in our approach is that it can be extended to calculate a variety of effects that would be very difficult to compute using the more rigorous theory.

TABLE CAPTIONS

Table 1

Numerical estimates and scalings of scintillation parameters for power-law spectra of interstellar electron density fluctuations with $\beta = 11/3, 4$ and 4.3. D is in kpc, λ in m, $C_{-4} = 10^4 C_N^2$ (C_N^2 as in Armstrong et $al.$ 1984) and γ is defined in equation (7.4). The increase in amplitude of the refractive fluctuations with increasing β is explicitly seen. An asterisk (*) denotes an estimate for the fluctuations about the measured mean ($eg.$ 5.5) over an observation period T_y in years. v_7 is the combined Earth-pulsar velocity in 10^7cm/sec; the equivalent spatial lag of the observation period is written $s = v_7 T_y$. $\mu = \ln[\dfrac{v_7 T_y}{\gamma^{.5} C_{-4}^{.5} \lambda^2 D^{1.5}}]$ is a logarithmic correction factor for θ_x and $\beta = 4$. For $\beta \geqq 4$, e_d and m_d depend strongly on the outer scale and are therefore omitted.

Table 2

Normalised cross-correlations as defined in equation (7.8) for power-law spectra of interstellar electron density fluctuations. Each entry consists of three values corresponding from top to bottom to $\beta = 11/3$, 4 and 4.3, respectively. The flux F, angular size Ω, decorrelation bandwidth ν_{dc}, scintillation timescale t_s and the position shift derivative $\dot{\theta}$ correlate with one another. Cross-correlations with the pulse broadening τ are identical to those with ν_{dc}, but with the opposite sign. The rate of flux variation $\dot{F}$ correlates with position shift θ_x and drift slope m_d. For $\beta \geqq 4$, m_d diverges in the absence of an outer scale and therefore the corresponding correlations have been omitted.

Table 1

	$\beta = 11/3$	$\beta = 4$	$\beta = 4.3$
θ(mas)	$2.2C_{-4}^{0.6}\lambda^{2.2}D^{0.6}$	$2.2\gamma^{5}C_{-4}^{.5}\lambda^{2}D^{.5}$	$2.2C_{-4}^{.59}\lambda^{2.4}D^{.76}$
τ(μs)	$2.9C_{-4}^{1.2}\lambda^{4.4}D^{2.2}$	$2.9\gamma C_{-4}^{1.0}\lambda^{4}D^{2}$	$2.9C_{-4}^{1.2}\lambda^{4.8}D^{2.5}$
ν_{dc}(kHz)	$54C_{-4}^{1.2}\lambda^{-4.4}D^{-2.2}$	$54\gamma^{-1}C_{-4}^{-1}\lambda^{-4}D^{-2}$	$54C_{-4}^{-1.2}\lambda^{-4.7}D^{-2.5}$
t_{s}(s)	$149C_{-4}^{-0.6}\lambda^{-1.2}D^{-0.6}v^{-1}$	$149\gamma^{-.5}C_{-4}^{-.5}\lambda^{-1}D^{-.5}v_{7}^{-1}$	$149C_{-4}^{-.59}\lambda^{-1.4}D^{-.76}v_{7}^{-1}$
T_{ref}(d)	$19C_{-4}^{0.6}\lambda^{2.2}D^{1.6}v_{7}^{-1}$	$19\gamma^{-.5}C_{-4}^{.5}\lambda^{2}D^{1.5}v_{7}^{-1}$	$19C_{-4}^{.59}\lambda^{2.4}D^{1.76}v_{7}^{-1}$
δF	$.12C_{-4}^{-.2}\lambda^{-.57}D^{-.37}$	$.38\gamma^{-.5}$	$.55$
$2\Delta\theta$(mas)	$.17C_{-4}^{.4}\lambda^{1.6}D^{.23}$	$.58C_{-4}^{.5}\lambda^{2}D^{.5}$	$.89C_{-4}^{.59}\lambda^{2.4}D^{.76}$
Δt(μs)	$1.2C_{-4}^{.91}\lambda^{2}D^{.5}s^{.83}$ *	$4.3C_{-4}^{.5}\lambda^{2}D^{-.5}s$ *	$6.7C_{-4}^{.52}\lambda^{2}D^{.5}s^{1.2}$ *
$\Delta\tau$(μs)	$.30C_{-4}^{1.0}\lambda^{3.8}D^{1.8}$	$1.0\gamma^{5}C_{-4}^{1}\lambda^{4}D^{2}$	$1.6C_{-4}^{1.2}\lambda^{4.7}D^{2.5}$
$\Delta\dot{F}$(d^{-1})	$(\frac{\cdot}{150})C_{-4}^{.20}\lambda^{-2.8}D^{-2.0}v_{7}$	$(\frac{\cdot}{50})\gamma^{-1}C_{-4}^{-.5}\lambda^{-2}D^{-1.5}v_{7}$	$(\frac{\cdot}{37})C_{-4}^{-.59}\lambda^{-2.4}D^{-1.8}v_{7}$
ΔS_{z}(mas)	$.32C_{-4}^{.40}\lambda^{1.6}D^{.23}$	$.62\gamma^{-.5}C_{-4}^{.5}\lambda^{2}D^{.5}\mu$ *	$1.5C_{-4}^{.5}\lambda^{2.0}D^{.5}s^{.15}$ *
$\Delta\theta$(mas/d)	$(\frac{1}{109})C_{-4}^{-.20}\lambda^{-.57}D^{-1.4}v_{7}$	$(\frac{1}{45})\gamma^{-.5}D^{-1}v_{7}$	$(\frac{1}{28})D^{-1}v_{7}$
$\Delta\nu_{dc}$(kHz)	$5.3C_{-4}^{-1.4}\lambda^{-5.0}D^{-2.6}$	$19\gamma^{-1.5}C_{-4}^{-1}\lambda^{-4}D^{-2}$	$31C_{-4}^{-1.2}\lambda^{-4.7}D^{-2.5}$
Δt_{s}(s)	$7.1C_{-4}^{-.80}\lambda^{-1.8}D^{-.97}v_{7}^{-1}$	$25\gamma^{-1}C_{-4}^{-.5}\lambda^{-1}D^{-.5}v_{7}^{-1}$	$37C_{-4}^{-.59}\lambda^{-1.4}D^{-.76}v_{7}^{-1}$
e_{s}	$.08C_{-4}^{-.20}\lambda^{-.57}D^{-.37}$	$.27\gamma^{-.5}$	$.40$
e_{d}	$.36C_{-4}^{-.20}\lambda^{-.57}D^{-.37}$	$-$	$-$
m_{d}	$.50C_{-4}^{-.20}\lambda^{-.57}D^{-.37}$	$-$	$-$

Table 2

	Ω	ν_{dc}	t_s	$\dot{\theta}$	
	.61	-.76	-.50	-.125	
	.71	-.80	-.58	.00	F
	.78	-.84	-.64	.13	
		-.91	-.82	.56	
		-.94	-.82	.58	Ω
		-.96	-.82	.61	
	.06		.75	-.37	
$\dot{F}$	.00		.77	-.44	ν_{dc}
	—		.79	-.51	
	.10	.07		-.69	
m_d	—	—		-.71	t_s
	—	—		-.74	
	$\delta\theta_x$	$\dot{F}$			

FIGURE CAPTIONS

Figure 1. Schematic representation of the focusing and steering effects of refractive perturbations. The upper section shows the effect on an incident plane wave of a thin screen with a single long-wavelength sine-wave perturbation. As the observer moves relative to the resulting intensity distribution, the flux F and position of the source θ are modulated as shown below. F is maximal at A, varies most rapidly at B and is minimal at C. Thus F correlates with the rate of change of θ and vice-versa.

Figure 2. Refractive and diffractive contributions to the phenomenon of frequency drifts in dynamic scintillation spectra. For pulsars observed through the interstellar medium (left panel), the effect is produced by the combined action of small and large scale inhomogeneities. The scintillation timescale t_s and the instantaneous decorrelation bandwidth ν_{dc} are indicated on the resulting frequency drift patterns. On the right we depict an equivalent optical system consisting of a diffracting mask and a dispersing prism. An observer moving past the pattern with velocity v detects intensity maxima at the different frequencies ν_4 to ν_1 ($\nu_4 > \nu_1$) at successively later times, leading to sloping bands in the frequency-time plane. In the example given the diffracting mask has only a small number of slits, thus producing a *periodic* diffraction pattern as shown. Similar periodicities are sometimes seen in dynamic scintillation spectra of pulsars, suggesting that at these times the pulsar image is dominated by a few bright well-separated spots.

Figure 3. Normalised autocorrelations as a function of normalised spatial lag (s/σ), *i.e.* normalised time lag T/T_{ref} for a Kolmogorov spectrum $(\beta=11/3)$. The flux F, its derivative $\dot{F}$, the rate of position wander $\dot{\theta}$, the decorrelation bandwidth ν_{dc} and the scintillation timescale t_s are shown. The amplitudes are expressed in terms of the dimensionless constant K (equation 2.14) using the expressions given in Appendix A.

Figure 4. Normalised autocorrelations for $\beta=11/3$ of θ_x, the fluctuation in image position parallel to the direction of relative motion (dotted line), and θ_y (dashed line). The solid line corresponds to a tenth of the autocorrelation of the drift slope m_d. Amplitudes are in units of K and expressions are given in Appendix A. Note that the decorrelation timescale of the fluctuations shown here is much greater than for those shown in figure 3. This is because θ_x, θ_y and m_d all have a divergent variance for $\beta \geqq 4$ in the absence of an outer scale.

Figure 5. Scattering geometry for an extended medium represented by n thin screens. The angles θ_i and ξ_i described in the text are shown for screen 1 and a ray path that connects the source to the observer is denoted by the bold line.

Figure 6. Normalised flux autocorrelation functions for a $\beta=4$ spectrum are shown for 1(dots), 3(short dashes), 5(long dashes) equally spaced thin screens and for a continuous scattering medium (solid line). Normalisation is in terms of the single screen dimensionless constant K defined in equation (2.14).

Figure 7. Comparison of the normalised flux autocorrelation functions for a single equivalent screen (dashed lines) and a continuous medium (solid lines) for three values of β (11/3, 4, 4.3). Normalisation is in terms of the single screen K in equation (2.14).

Figure 1

-123-
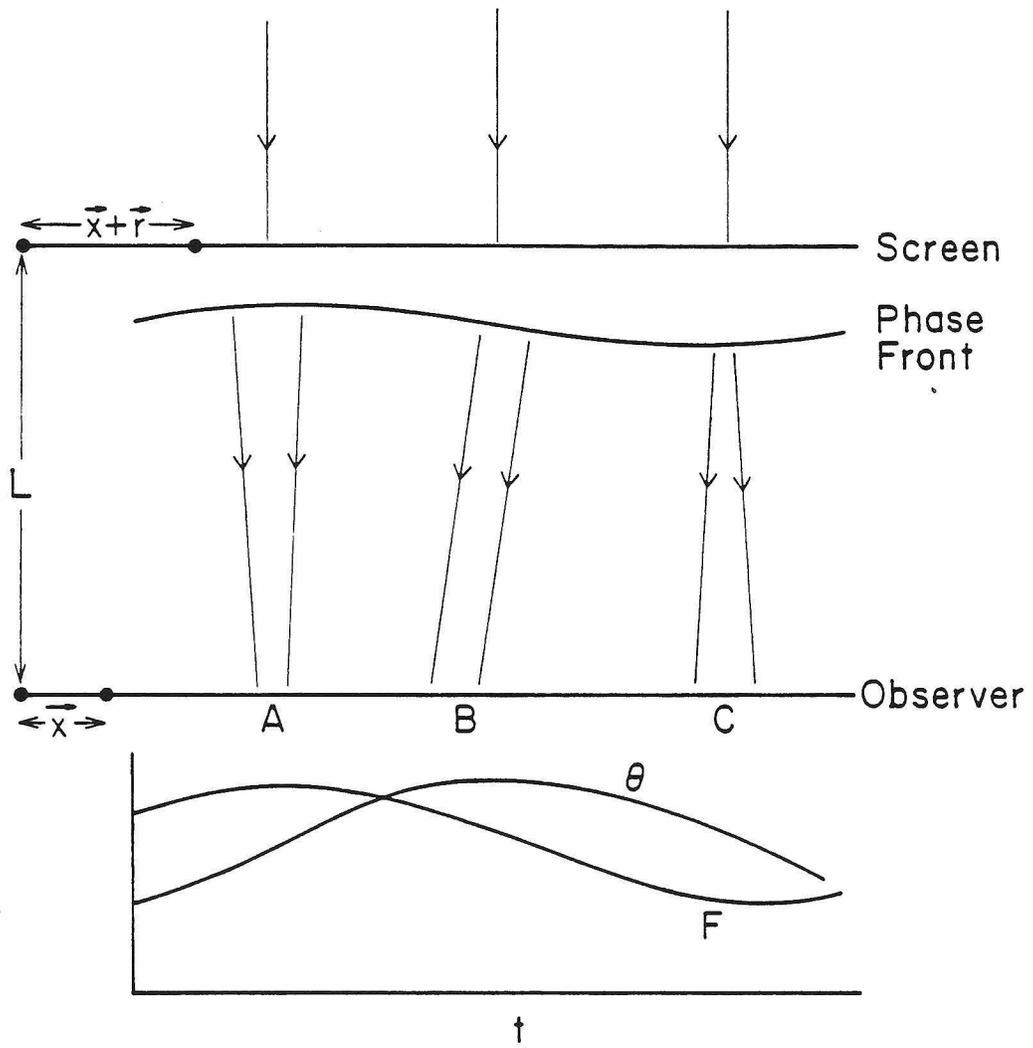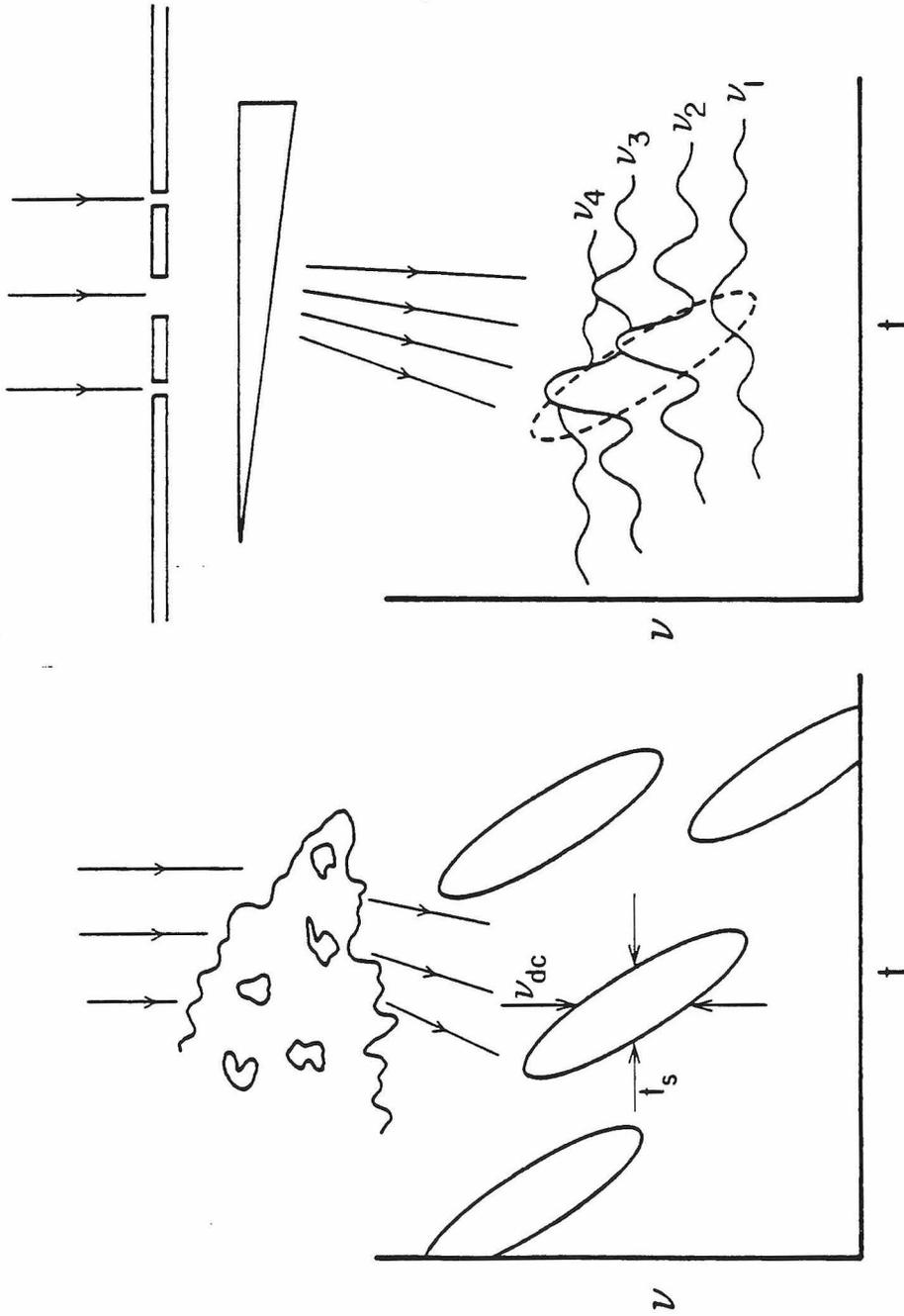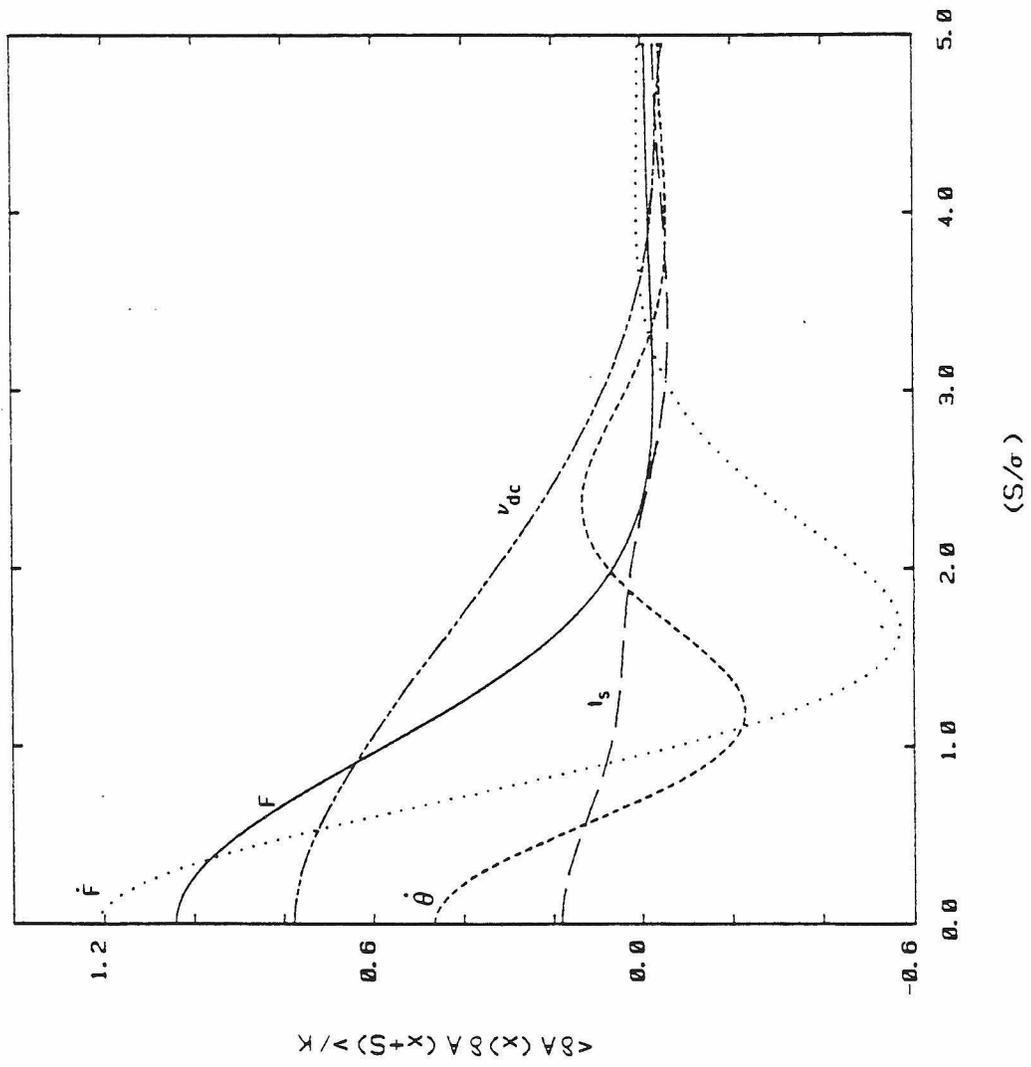
Figure 2

Figure 3

Figure 4
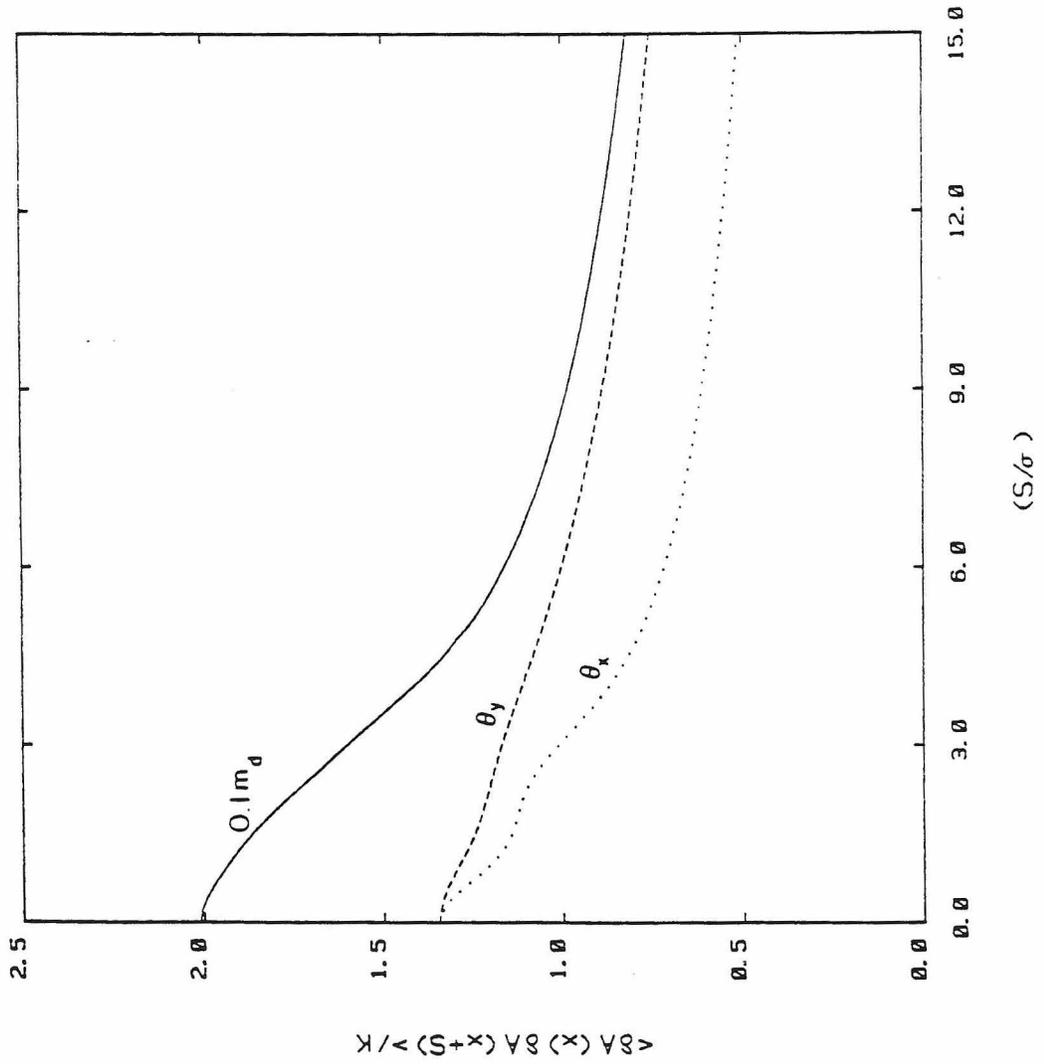
Figure 5
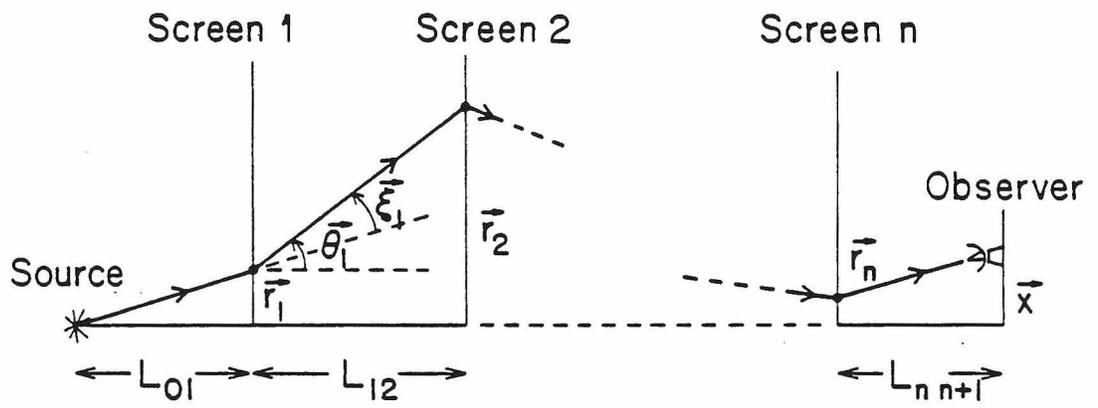
Figure 6
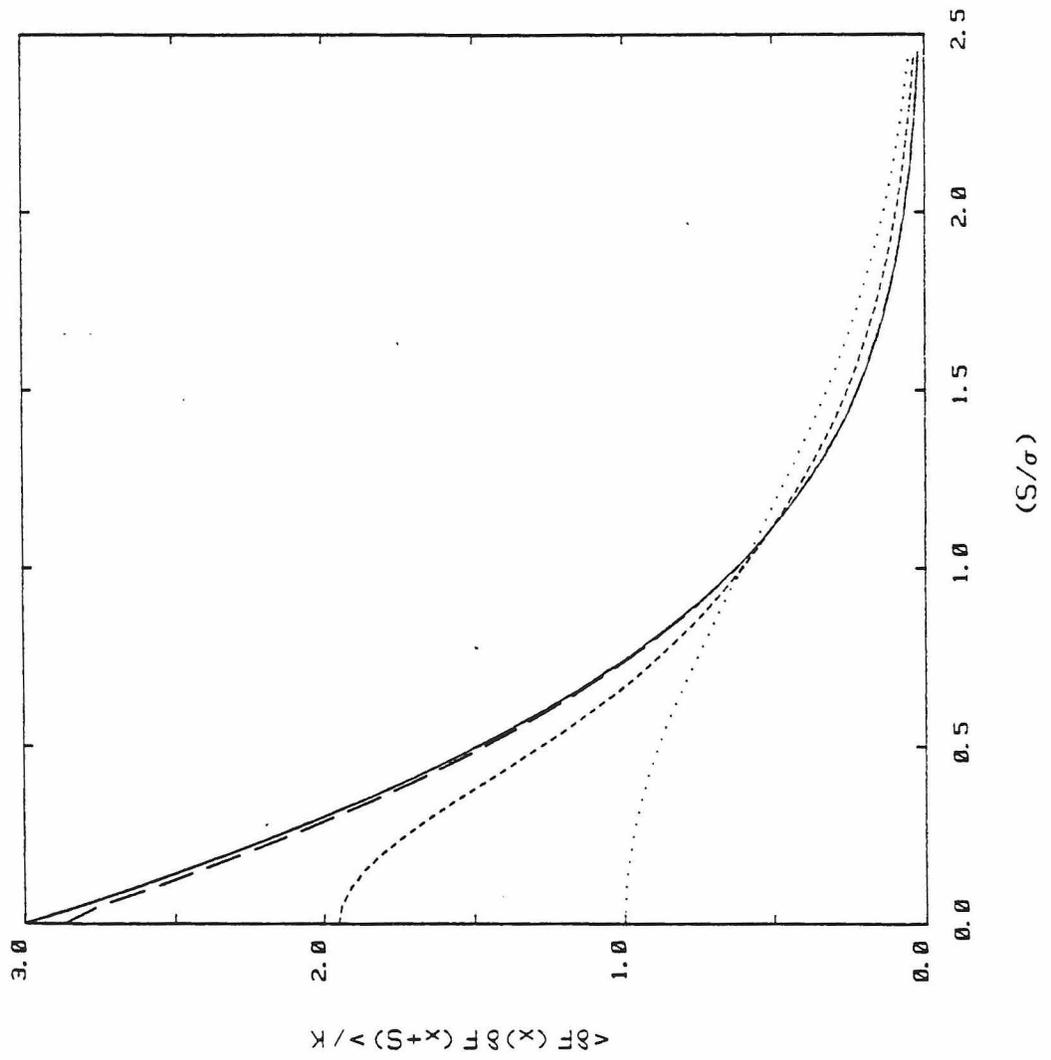
Figure 7
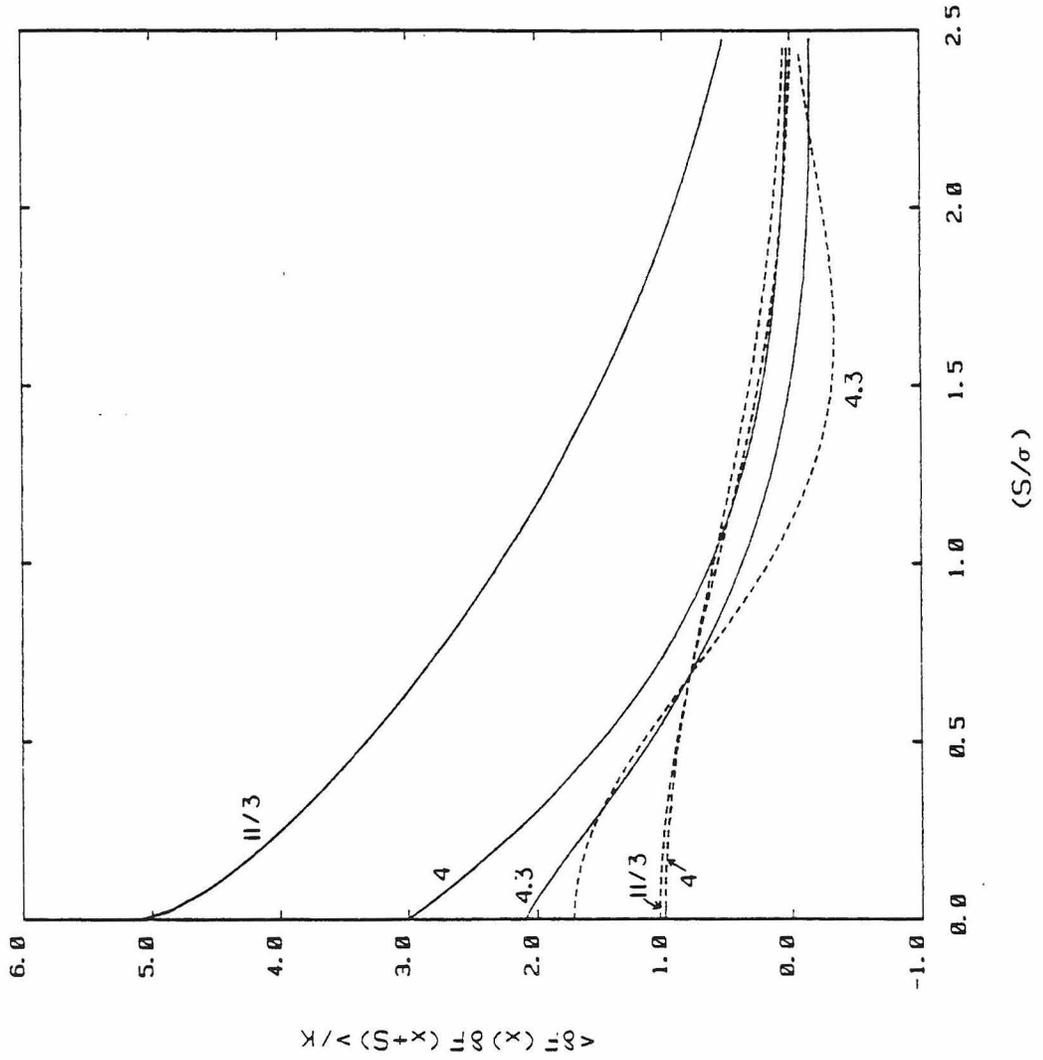
Chapter 6:

FLICKER OF EXTRAGALACTIC RADIO SOURCES
AND REFRACTIVE INTERSTELLAR SCINTILLATION

Roger Blandford, Ramesh Narayan and Roger W. Romani

Theoretical Astrophysics

California Institute of Technology

Pasadena, CA 91125, U.S.A.

ABSTRACT

Recent work has identified variability of flat-spectrum extragalactic radio sources at $\lambda \approx 10$ cm with rms amplitude of ~2-3% and time scale of days. We show that this "flicker" is consistent with intensity fluctuations caused by refractive scintillation in an extended interstellar medium in our Galaxy. Further observation of flicker may allow the structure of suitable sources to be partially resolved on angular scales smaller than those probed by VLBI.

INTRODUCTION

Variability of compact extragalactic radio sources has been found over a range of frequencies. Large amplitude intensity fluctuations at centimeter wavelengths on time scales of approximately months are consistent with (possibly relativistic) variations of synchrotron sources. Variation found at meter wavelengths on time scales of approximately years was originally discovered by Hunstead (1972) has since been demonstrated to be widespread (*eg.* Condon, *et al.* 1979). If this variation is also intrinsic to the sources, then highly relativistic internal motions and prodigious energy requirements are indicated for a large fraction of the objects.

As an alternative, Rickett, Coles and Bourgois (1984) proposed that the low frequency variability of compact extragalactic radio sources might be an effect of propagation through large-scale electron density inhomogeneities in the interstellar medium (ISM) of our Galaxy. With a sufficiently steep power spectrum for these ISM density perturbations (*cf.* Blandford and Narayan 1985; Goodman and Narayan 1985), intensity fluctuations of large amplitude can be produced in pulsars and extragalactic sources by refractive focusing and defocusing of the radio waves.

There is a third type of variability that has been detected in careful observations by Heeschen (1982,1984). This is a low-amplitude (~2%) short time scale (~few days) "flicker" at wavelengths λ~10 cm. This variability was also associated by Rickett, *et al.* with refractive interstellar scintillation, the small amplitude being attributable to the finite size of the source. However, recent observations by Simonetti, Cordes and Heeschen (1985) of flicker at two radio frequencies seem at first sight to be in conflict with an interpretation in terms of refractive scintillation. In this *Letter*, we show that the apparent discrepancies can be resolved by including the effect of an extended galactic scattering medium and allowing the source size to vary with the observation wavelength in a manner consistent with synchrotron models of the radio emission.

REFRACTIVE SCINTILLATION AND FLICKER

Simonetti, *et al.* monitored 14 flat- and 20 steep-spectrum extragalactic sources for 20 days at 1410 MHz and 2380 MHz. The flat spectrum sources showed flicker in their normalized flux fluctuations $\delta F(t)$ at both frequencies with average rms amplitudes μ of ~2.3% and ~2.8%, respectively. Steep spectrum sources, as expected, do not flicker. First-order structure functions of the fluctuations, defined to be

$D^{(1)}(\tau) = <\delta F(t)\delta F(t)> - <\delta F(t)\delta F(t+\tau)>$, of the flat-spectrum sources were found to vary with time lag τ approximately as $D^{(1)}(\tau) \propto \tau$ throughout the observation period. Significant power was also found in the cross-frequency structure function, indicating correlation in the fluctuations at the two frequencies.

Blandford and Narayan (1985) have introduced a simple formalism that gives an approximate description of refractive interstellar scintillation (ISS) and used it to compute the autocorrelation function of the normalized flux fluctuations for the case of an equivalent thin scattering screen at a distance L from the observer. In this *Letter*, we assume a power spectrum of interstellar refractive density perturbations $Q(\beta) = Q_0 q^{-\beta}$, where $Q(q)$ is dimensionless. (It is widely believed that $\beta = 3.7$ [*eg.*, Lee and Jokipii 1975; Armstrong, Cordes and Rickett 1981], though there are suggestions that β may be as large as 4.3 [Blandford and Narayan 1985; Goodman and Narayan 1985].) In this *Letter*, for simplicity, we set $\beta = 4$. The autocorrelation is then

$$<\delta F(x)\delta F(x+v\tau)> = \frac{Q_0 \lambda^4}{2\pi\theta^2} e^{-v^2\tau^2/(2\theta^2 L^2)} \tag{1}$$

where $\lambda = 2\pi\lambdabar$ is the observation wavelength, θ is the apparent angular size of the source, and v is the velocity at which the scattering screen moves past the line of sight joining the source and the observer. The angle brackets denote a spatial average in the observer's plane or equivalently a time average in the observations as used above. In the case of flicker, the broadening of the source image due to scattering is negligible and θ is almost equal to the intrinsic angular size θ_I of the source. As noted by Simonetti, *et al.* if θ is constant, the rms fluctuation amplitude μ should then scale approximately as λ^2, implying $\mu(1410)/\mu(2380) \approx 3$. The mean of the observed ratios of $\mu(1410)$ to $\mu(2380)$ is 1.0 ±0.6, arguing against an explanation of flicker in terms of ISS in a thin scattering screen. Further, from equation (1), the average first order structure function, $<D^{(1)}(\tau)>$ is proportional to τ^2, instead of the observed linear scaling.

Since the density perturbations on scales $\sim \theta L$ contribute the bulk of the focusing that leads to flux variations, refractive ISS in an equivalent screen is dominated by a single length scale. In an extended scattering medium, however, the diameter subtended by the image will be a function of distance from the observer and so a range of scales can contribute to the fluctuations. This changes the functional form of $D^{(1)}(\tau)$. For simplicity, we take the galactic scattering medium to be a Gaussian disk of scale height H (an

exponential distribution gives similar results). The scattering strength Q_0 in equation (1) is then replaced by a scattering density

$$\frac{d\,Q_0(L)}{d\,L} = \frac{2Q_0}{\sqrt{\pi}H}\,e^{-L^2/H^2} \tag{2}$$

We also assume that the velocity v is dominated by the Earth's orbital motion and so is the same for all screens. Since $\theta \approx \theta_I$ is the same for all the screens, we can integrate over the scattering medium to obtain

$$<\delta F(t)\delta F(t+\tau)> = \int_0^\infty \frac{\lambda^4 Q_0}{\pi^{3/2}\theta^2}\,e^{-y^2}e^{-v^2\tau^2/(2H^2\theta^2 y^2)}\,dy \tag{3}$$

$$= \frac{\lambda^4 Q_0}{2\pi\theta^2}\,e^{-\sqrt{2}v\tau/\theta H} \tag{4}$$

Expanding for small τ we see that

$$D^{(1)}(\tau) = \frac{\lambda^4 Q_0}{2\pi\theta^2}\left(1 - e^{-\sqrt{2}v\tau/\theta H}\right) \sim \frac{\lambda^4 Q_0 v}{\sqrt{2}\pi\theta^3 H}\left(\tau - \frac{\sqrt{2}v\tau^2}{\theta H}\right) \tag{5}$$

Thus, for a total observing time $T \lesssim T_{sat} = \theta H/\sqrt{2}v$ the structure function is linear in τ, as observed.

The observed wavelength independence of the rms flicker magnitude can also be understood if the intrinsic source angular size increases with wavelength and we observe for times $T \lesssim T_{sat}$. For example $\theta_I \propto \lambda$ for a flat spectrum, constant brightness temperature core (eg., Kellermann and Pauliny-Toth 1981) and thus for $T \lesssim T_{sat}$

$$\mu(T) = [D^{(1)}(\tau)]^{1/2} \propto \lambda^{1/2}, \tag{6}$$

which gives an estimate of the effective scintillation index as a function of the observation period. This is in reasonable accord with the observations, given the substantial measurement error.

In addition to being able to produce the observed scaling relations we should also check that the amplitude of the fluctuations is in the expected range. Ideally this should be done on a case-by-case basis using VLBI observations to determined the unresolved flux and low frequency variability to constrain the scattering properties of the interstellar medium. At low frequencies, the scatter-broadened angular size of a point source θ_s, which varies as λ^2, will be comparable to or even exceed the intrinsic angular size θ_I, and

equations (3)-(5) are not valid. Romani, Narayan and Blandford (1986) have developed an approximate theory to compute the flux variability amplitude and time scale for a scatter-broadened point source. With a simple extension, this theory can be used for an extended source. For a circular Gaussian source of intrinsic angular radius θ_I and a scattering density described by equation (2), we find that the angular size θ in the integral of equation (3) must be replaced by

$$\theta^2(y) = \theta_s^2(y) + \theta_I^2$$

$$\theta_s^2(y) = \theta_s^2(0)\left[1 - \operatorname{erf}(y)(1 - y^2/2) - e^{-y^2}/\sqrt{\pi}y\right] \tag{7}$$

Numerically, we adopt a typical flux and brightness temperature of $S \approx 1$ Jy and $T_B \approx 5 \times 10^{11}$K to obtain $\theta_I \approx 3.6 \lambda_m$ mas and take the scattering angular size to be $\theta_s(0) \approx 2.2 C_{-4}^{1/2} \lambda_m^2 H_{kpc}^{1/2}(\operatorname{cscb})^{1/2}$ mas (Cordes, Weisberg and Boriakoff 1984, Romani, et al.), where the observation wavelength is in meters, C_{-4} is a measure of the strength of the scattering and b is the galactic latitude of the source. Thus for an average line of sight with $H_{kpc} \approx 1$, $C_{-4} \approx 1$, and $b \approx 30°$, this gives $\theta(0) \approx 4\lambda_m(1 + 0.7\lambda_m^2)^{1/2}$ mas. We assume that the velocity v is dominated by the Earth's component of motion perpendicular to the line of sight, on average $v \approx (\sqrt{2/3})30$Km/s. Using these estimates, we compute the first-order structure functions for flicker at 1410 MHz and 2380 MHz and low frequency variability at 327 MHz and display the results as a function of the observation period in Figure 1.

The saturations time scales for the growth of the intensity fluctuations, are 38 days and 23 days at 1410 MHz and 2380 MHz with the above parameters, so after 20 days the observations are still within the linear range. The expected rms flicker magnitudes are 4.4% at 1410 MHz and 3.1% at 2380 MHz, somewhat higher than the observed values, but not unreasonably so in light of the difficulty of estimating the contribution of measurement error. (We note that the presence of source components on angular scales larger than θ_I can quench the flicker.) The ratio of the flicker amplitudes, $\mu(1410)$ and $\mu(2380)$, is 1.4, in reasonable agreement with the average flicker ratio computed by Simonetti, et al.. The variability time scale at 327 MHz is $T_{var} \approx 1$ yr and the amplitude is ~26%, again in reasonable accord with the observations. (Note that this refers to the rms fluctuations in a large sample of similar sources and does not describe the peak-to-peak variations in the currently most variable sources.)

We can also demonstrate the broadband nature of the intensity fluctuations caused by refractive scintillation by considering the cross-correlation between intensity fluctuations measured at two wavelengths λ_1, λ_2. Using the simple model of equations (3)-(5), we obtain

$$<\delta F(t,\lambda_1)\delta F(t+\tau,\lambda_2)> = \frac{\lambda_1^2 \lambda_2^2 Q_0}{\pi(\theta_1^2+\theta_2^2)} e^{-\sqrt{2}\upsilon\tau/H(\theta_1^2+\theta_2^2)^{1/2}} \tag{9}$$

where $\theta_{1,2}$ are the apparent angular sizes at $\lambda_{1,2}$. Simonetti, *et al.* have estimated from their data the first order cross-frequency structure function, normalized by the autofrequency results and averaged over all flat spectrum sources and find a significant correlation. Computing this object with our theory, we obtain

$$\frac{D_{12}^{(1)}(\tau)}{[D_{11}^{(1)}D_{22}^{(1)}]^{1/2}} = \frac{2\theta_1\theta_2}{\theta_1^2+\theta_2^2} \frac{1-\exp[-2\upsilon\tau/H(\theta_1^2+\theta_2^2)^{1/2}]}{(1-\exp[-\sqrt{2}\upsilon\tau/H\theta_1])^{1/2}(1-\exp[-\sqrt{2}\upsilon\tau/H\theta_2])^{1/2}} \tag{10}$$

Inserting our estimates for $\theta(0)$, we find that for τ less than T_{sat} the value is ~0.8, becoming ~0.9 after saturation. The measured cross-frequency structure function between 1410 and 2380 MHz has large errors, but clearly reaches values greater than 0.5.

Note that, except for changes in the intrinsic source structure, the percentage polarization of the source should not fluctuate. This contrasts with the (presumably) intrinsic variation of BL Lac objects where the polarization varies more rapidly than the total flux.

DISCUSSION

We have shown how existing observations of flickering in compact radio sources can be interpreted as an interstellar propagation effect if we allow the source size to vary with wavelength and abandon the thin screen approximation. Although some sources probably do vary intrinsically at low frequencies (*eg.* AO 0235+164) we endorse the proposal of Rickett *et al.* that low frequency variability and flicker are mainly due to refractive scintillation. Multi-frequency observations of selected sources over time scales long enough to measure the saturated structure function should confirm the approximate scaling $\mu_{sat} \propto \lambda^2/\theta_I(\lambda)$ for the flicker amplitude and $T_{var} \propto \lambda^2 H/\upsilon$ for the low frequency variability time scale.

Detailed observations of flicker might, in principle, be used via a sort of "Earth orbit aperture synthesis" to obtain a superior angular resolution to that attainable using conventional Earth rotation VLBI. Suppose that the scattering medium moves slowly compared with the Earth. We can define a power spectrum of the flux fluctuations $P(k)$ by taking the Fourier transform of the autocorrelation function in equation (1). It is then straightforward to show that for a k^{-4} spectrum

$$P(k) = \lambda^4 \int dL \ L^2 \frac{d \ Q_0}{d \ L} |V(kL)|^2 \tag{11}$$

where $V(b)$ is the complex visibility function, the Fourier transform of the normalized intensity of the source. A similar relation between the visibility function and the intensity autocorrelation has been derived by Salpeter (1967) (and also M.H. Cohen, reference therein). The shape of the effective "beam" and the sampling in the observer plane will, of course, be determined by the Earth's motion and the motion of the solar system barycenter. Obviously equation (11) is easiest to invert when the scattering is confined to a thin screen [*ie.* when $D^{(1)}(\tau) \propto \tau^2$], and when the wavelength is long enough to give easily measurable fluctuations yet short enough for the interstellar scattering not to obscure the source structure.

In practice, intrinsic source variability and motions of the interstellar plasma will probably prevent this method from ever being implemented. Nevertheless, this analysis does serve to highlight a complementary consequence of refractive fluctuations, namely that daily variation in the angular structure of compact radio sources at the level $\sim 10^{-2}$ limits the accuracy of VLBI mapping. Dynamic ranges of 1000:1 claimed in some contemporary experiments will not be attainable in strongly flickering sources. It remains to be seen how severe this restriction turns out to be in practice.

REFERENCES

Armstrong, J.W., Cordes, J.M. & Rickett, B.J. 1981. *Nature*, **291**, 561.

Blandford, R. & Narayan, R., 1985. *Mon. Not. R. astr. Soc.*, **213**, 591.

Condon, J.J., Ledden, J.E., O'Dell, S.L., and Dennison, B. 1979, *A.J.*, **84**, 1.

Cordes, J.M., Weisberg J. & Boriakoff, V. 1984. *Ap. J.* **288**, 221.

Goodman, J. & Narayan, R., 1985. *Mon. Not. R. astr. Soc.* **214**, 519.

Heeschen, D.S., 1982. in *IAU Symposium* 97, *Extragalactic Radio Sources*,
    eds. D.S. Heeschen, and C.M. Wade (Dordrecht: Reidel), p. 327.

Heeschen, D.S., 1984. *Astron. J.*, **89**, 1111.

Hunstead, R.W. 1972. *Ap. Letters*, **12**, 193.

Kellermann, K.I., & Pauliny-Toth, I.I.K., 1981. *Ann. Rev. Astr. Ap.*, **19**, 373.

Lee, L.C. and Jokipii, J.R. 1975. *Ap. J.* **201**, 532.

Rickett, B.J., Coles, W.A. & Bourgois, G., 1984. *Astron. Astrophys.*, **134**, 390.

Romani, R.W., Narayan, R. & Blandford, R., 1986. *Mon. Not. R. astr. Soc.*, **220**, 19.

Salpeter, E.E., 1967. *Ap.J.*, **147**, 433.

Simonetti, J.H., Cordes, J.M. & Heeschen, D.S., 1985. *Ap. J.*, **296**, 46.

Figure Caption

Figure 1. Variation of the average first-order structure function $<D^{(1)}(\tau)>$ with the observation period τ for flicker at 1410 MHz and 2380 MHz and for low-frequency variability at 327 MHz. For times less than the saturation time scale $T_{sat} = \theta H / \sqrt{2} v$ the growth rate is approximately linear with τ, and the amplitudes are similar. The saturation time becomes longer at lower radio frequencies, and the saturated amplitude increases. The typical rms fluctuation in the flux is given by the square root of the saturated value of $<D^{(1)}(\tau)>$.
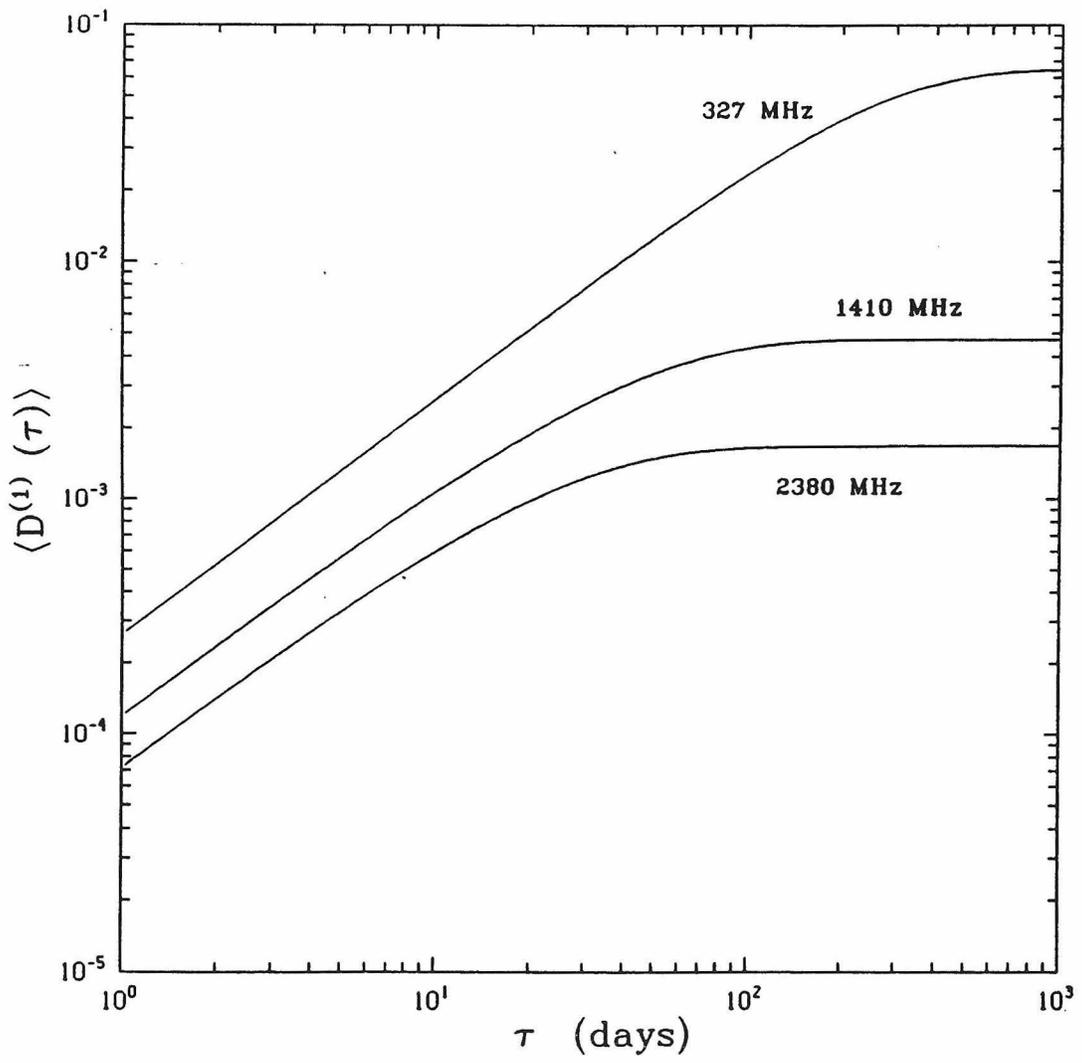
Chapter 7:

THE EFFECTS OF CAUSTICS ON SCINTILLATING RADIO SOURCES

Jeremy J. Goodman,

Institute for Advanced Study

Princeton, N.J. 08540, U.S.A.

Roger W. Romani, Roger D. Blandford,

Theoretical Astrophysics, California Institute of Technology

Pasadena, CA 91124, U.S.A.

and *Ramesh Narayan*

Steward Observatory, University of Arizona

Tucson, AZ, 85721, U.S.A.

Summary. We consider the scintillation properties of compact radio sources for a spectrum of interstellar electron density fluctuations which is a power-law over a finite range of spatial frequencies. In particular, if the power-law is truncated at an inner scale intermediate between the diffractive scale and the refractive scale, we find that there is additional power in the spectrum of the intensity variations at these scales. This power is associated with strong focusing events, or *caustics*. These events are best described as simple diffraction catastrophes, which are classified and analysed on the basis of geometrical and wave optics, taking account of the strong dispersion of the ISM scattering, which introduces important frequency dependencies. It is shown that such caustics can enhance the flux modulation and may produce patterns in dynamic spectra similar to those observed. Caustics are best observed by monitoring pulsar fluxes through a range of frequencies and specific observations are suggested which should either detect their effects or delimit their importance. We conclude that caustic effects are only likely to dominate the flux variations when the inner scale is ~ 10 percent of the refractive scale. For somewhat larger inner scales individual caustics could be resolved. Whether such conditions obtain in the ISM is still uncertain; if they do not, a power-spectrum description may not be adequate and focusing events from single scatterers (*ie.* rare high-density clouds instead of an incoherent superposition of small-amplitude density fluctuations along the line of sight) may dominate the refractive variations.

I. Introduction

While the possibility of a slow component of the flux scintillation of radio pulsars by the interstellar medium (ISM) has been discussed for some time (e.g. Salpeter (1967), Lovelace (1970), Shishov (1974), Shapirovskaya (1978), Sieber (1982)), more recent work (Rickett, Coles and Bourgois (1984); Blandford and Narayan (1985), hereafter BN) has emphasized the importance of such refractive (i.e. geometrical optics) fluctuations and shown that they may account for the low frequency variability of pulsars and extragalactic radio sources. In these studies the spectrum of electron density fluctuations in the ISM giving rise to the scintillations has been most frequently characterised as an extended power-law in spatial wavenumber, $Q(q) \sim q^{-\beta}$ where β is close to the "Kolmogorov" value, $\beta = 11/3$ Rickett (1977). It has been shown (BN; Romani, Narayan, and Blandford (1986), hereafter RNB) that refractive scintillation should cause a number of observables other than the flux to fluctuate, as well, and that there should be correlations among these fluctuations. Comparison of the variations computed for a $\beta = 11/3$ spectrum with the observations has, however, met with several difficulties—(i) the observed fluctuations (i.e. modulation indices) exceed those computed for the assumed spectrum of density perturbations, (ii) strongly non-gaussian "spikes", inconsistent with the simple theory, are sometimes seen in long-term studies of pulsar flux (e.g. Cole, Hesse and Page (1970), Helfand, Fowler and Kuhlman (1977)), (iii) quasi-periodic bands are sometimes observed in the "dynamic scintillation spectra" (i.e. contour plots of flux density in the frequency-time plane) of many pulsars (e.g. Roberts and Ables (1982); Cordes, Pidwerbetsky and Lovelace (1986)), again not explained by the standard theory.

These three observations argue for increased power in the low frequency refractive fluctuations relative to the normal "diffractive" scintillation. There are three ways in which the spectrum of electron density perturbations could be modified to achieve this: (i) the power-law index of the density perturbations could be increased to $\beta \gtrsim 4$, (ii) the power spectrum could be truncated at an inner scale, (iii) an extra component may be added to the perturbation spectrum at low frequencies. The first approach has been suggested by Roberts and Ables (1982) and Hewish, Wolszczan, and Graham (1985) based on observations of dynamic scintillation spectra. BN, Goodman and Narayan (1985; henceforth GN) and RNB have calculated fluctuations based on $\beta > 4$ spectra, finding that the modulation indices, etc. do increase although some difficulties remain in interpreting the observations. Truncation of the power spectrum has been considered by Coles et al. (1987) and is

the subject of further investigation in this work. We find that, in addition to the enhancement of the refractive low frequency fluctuations discussed by these authors, there should be an additional contribution to the flux modulation from intermediate scales in the form of strong focusing events or *caustics*. The third approach, superposition of extra fluctuation power at a preferred scale has been advocated by Cordes and Wolszczan (1986). Although this third possibility generates clear predictions in terms of preferred timescales for variability and organization of dynamic spectra, we do not discuss it in detail. However, we note that this extra fluctuation power may be characterised by a random modulation at large scales or, as a more radical alternative, the distributed line-of-sight scattering may be dominated by single large fluctuations (Romani, Blandford and Cordes 1987). In either case, strong focussing can arise from scattering on these large scales and many of our qualitative results are applicable to the resulting modulations.

In section 2 we compute the autocorrelation function of flux variations for a power-law spectrum of phase fluctuations with an inner scale and show that there is excess power on intermediate scales. We identify conditions under which this contribution to the flux variations is dominant. In section 3, we review the general appearance of simple diffraction catastrophes, (i.e. caustics) under conditions appropriate to interstellar scintillation and show that they are the cause of the new component in the intensity spectrum in the presence of an inner scale. In section 4 we apply our results to the case of radio scintillation in the interstellar medium. We describe the variation of the diffraction catastrophes as a function of observation frequency and discuss strategies for the detection of individual caustics. In section 5 the predictions of the caustic picture for the appearance of dynamic scintillation spectra are described and compared with other scenarios. We conclude in section 6 by suggesting specific observational tests which could help us to gauge the importance of caustic events.

2. Flux Variations due to Power-Law Phase Fluctuations
with an Inner Scale

In this section we present a formal treatment of the power spectrum of flux variations arising from a power-law spectrum of phase fluctuations with an inner scale cut-off. We find that the cut-off introduces extra power at scales between those normally associated with diffraction and refraction.

We show in section 3.3 that this power is due to the effect of caustics. For readers wishing to avoid technical complications, the principal results are displayed in Figures 2 and 3.

2.1 The Thin Screen

To simplify the analysis, we adopt the thin-screen model throughout this paper, i.e. we replace the electron-density fluctuations of the interstellar medium by a thin scattering layer ("screen") at distances z_{os} and z_{sp} respectively from the observer's plane and from the pulsar (Fig. 1). We use the notations $\vec{R} = (X, Y)$ and $\vec{r} = (x, y)$ for positions in the observer's plane and in the thin screen, respectively, and take the origins for these two coordinate systems at the intersections with the optical axis OSP. To make matters even simpler, we replace the spherical wavefront emanating from the pulsar and impinging on the far side of the screen by a plane wave of the same flux, thus in effect letting $z_{sp} \to \infty$. In order to preserve the pattern of fluxes on the observer's plane, we change the distance between the screen and the observer to

$$z' \equiv \left(\frac{1}{z_{os}} + \frac{1}{z_{sp}} \right)^{-1}, \tag{2.1.1}$$

and use the rescaled coordinates

$$\vec{R}' \equiv \left(\frac{z_{sp}}{z_{os} + z_{sp}} \right) \vec{R} \tag{2.1.2}$$

in the observer's plane. These scalings ensure that the variation of path length with $\vec{r}, \vec{R}$ is preserved. Henceforth we drop the primes. We further assume that $r \ll z$ and $R \ll z$ (the "paraxial approximation"), which is reasonable since the scattering angles in our problem are measured in milliarcseconds, at most.

The advantage of the thin-screen model is that it permits us to describe a particular ray path between the pulsar and the observer by $\vec{r}$ and $\vec{R}$ alone (for a given z), whereas with an extended scattering medium an infinite number of parameters would be required. Predictions of the thin-screen and extended-medium models have been compared by Tatarskii and Zavorotnyi (1980). In general the agreement between the two models is very good (but see RNB, Frehlich 1986). In particular, Coles, et $al.$ have evaluated an approximate computation of the refractive fluctuations due to an extended scattering medium with an inner-scale cut-off and find an effect comparable to the thin screen case. It should be noted that insofar as the extended medium is assumed to be statistically

homogeneous, it is as much an idealization as the thin-screen model. Indeed, the thin-screen model is to be preferred when the scattering is dominated by a relatively small segment of the line of sight, e.g. an HII region.

2.2 The Diffraction Integral

If we ignore the very small depolarising effects of the scattering, then the electric field on the observer's plane can be represented by a scalar $\psi(\vec{R})$ whose absolute square is proportional to the flux $F(\vec{R})$, i.e.

$$F(\vec{R}) = F_0 |\psi(\vec{R})|^2, \tag{2.2.1}$$

where $F_0 \propto (z_{os} + z_{sp})^{-2}$ is the average flux. If the flux is measured in a very narrow band centered on wavelength $\lambda \equiv 2\pi/k$, and the Fresnel scale r_F is defined by

$$r_F = \sqrt{z/k}. \tag{2.2.2}$$

then in the thin-screen model described above,

$$\psi(\vec{R}) = \int \frac{d^2\vec{r}}{2\pi r_F^2} \exp[i\zeta(\vec{r}, \vec{R})], \tag{2.2.3}$$

where

$$\zeta(\vec{r}, \vec{R}) = \phi(\vec{r}) + \frac{(\vec{r}^{\,2} - 2\vec{r} \cdot \vec{R})}{2r_F^2} \tag{2.2.4}$$

is the total phase along the path that passes through $\vec{r}$ and $\vec{R}$, apart from an irrelevant term that depends on $\vec{R}$ only. The quantity $\phi(\vec{r})$ is the excess phase at $\vec{r}$ introduced by the scattering material, while the second term is a geometrical contribution due to path-length variations.

2.3 Phase Statistics

We are interested in evaluating certain moments of ψ. We regard the excess phase $\phi(\vec{r})$ to be a random function of $\vec{r}$, and we assume that it has zero mean, which is allowed since the mean can be subtracted from ζ without affecting the flux. Because the relation between ψ and ϕ is nonlinear, the problem is tractable only if we make the usual assumption that ϕ is a stationary gaussian process with correlation function

$$C(\vec{r}) \equiv \langle \phi(\vec{r}' + \vec{r}) \phi(\vec{r}') \rangle. \tag{2.3.1}$$

Since the electron column density must be positive, the gaussian approximation requires at the very least that the fluctuations in column density be small compared to the mean. In relating our thin-screen results to the extended medium, $C(\vec{r})$ should be interpreted as the mean-square difference in integrated phase along two *parallel* lines of sight with linear separation $\vec{r}$ and common length z.

$C(\vec{r})$ is uniquely determined by its Fourier transform $k^{-2}Q(\vec{q})$, the power spectrum of ϕ:

$$Q(\vec{q}) = k^2 \int d^2\vec{r}\, e^{i\vec{q}\cdot\vec{r}} C(\vec{r})$$
$$= k^2 \langle \tilde{\phi}(\vec{q})\tilde{\phi}^*(\vec{q})\rangle/A, \tag{2.3.2}$$

where $\tilde{\phi}(\vec{q})$ is the Fourier transform of $\phi(x,y)$, calculated over an area A on the screen. The extra factor of k^2 in the definition makes Q dimensionless. For simplicity we assume that the scattering is isotropic, so that C and Q depend only on $r \equiv |\vec{r}|$ and $q \equiv |\vec{q}|$, respectively. We can interpret $q^2 Q(q)/k^2$ as the mean-square amplitude of the phase fluctuations on the scale $r = 2\pi/q$.

We are particularly interested in spectra that are well approximated by power laws in some range:

$$Q(q) = Q_0 q^{-\beta}, \quad 2\pi/r_{out} \equiv q_{out} < q < q_c \equiv 2\pi/r_c. \tag{2.3.3}$$

We assume that Q is negligibly small outside the range, $q_{out} < q < q_c$. Thus, the largest scales in the medium are $\sim r_{out}$, the "outer" scale, and the smallest scales are $\sim r_c$, the "cut-off" or "inner" scale.

The correlation function corresponding to (2.3.3) for $r_c \ll r \ll r_{out}$ and $2 < \beta < 6$ (e.g. GN) is

$$C(r) \approx \frac{Q_0}{2\pi k^2(\beta-2)} \left[\frac{2^{3-\beta}\Gamma[(6-\beta)/2]}{(\beta-4)\Gamma(\beta/2)} r^{\beta-2} + q_{out}^{2-\beta} \right] \qquad (\beta < 4)$$
$$\approx \frac{Q_0}{2\pi k^2(\beta-2)} \left[\frac{2^{3-\beta}\Gamma[(6-\beta)/2]}{(\beta-4)\Gamma(\beta/2)} r^{\beta-2} + q_{out}^{2-\beta} + \frac{q_{out}^{4-\beta} r^2}{2(\beta-4)} \right] \qquad (\beta > 4) \tag{2.3.4}$$

For $r \gg r_{out}$, $C(r)$ is negligibly small. On the other hand, for $r \leq r_c$, $C(r)$ can be expressed as a convergent series in r^2 with the term in $q_{out}^{2-\beta}$ as the leading term. We have omitted terms from (2.3.4) that vanish as $q_{out} \to 0$ and $q_c \to \infty$.

2.4 Correlation Function and Power Spectrum of the Flux

The normalized correlation function of the flux,

$$W(\vec{R}) \equiv \langle F(0)F(\vec{r})\rangle / \langle F\rangle^2 \tag{2.4.1}$$

contains information about the typical amplitude of flux variations and the time scales on which they occur. Its two-dimensional Fourier transform can be obtained from (2.2.1)-(2.2.4) [Rumsey (1975)]:

$$
\begin{aligned}
\tilde{W}(\vec{q}) &\equiv \int d^2\vec{R}\, e^{i\vec{q}\cdot\vec{R}} W(\vec{R}) \\
&= \int d^2\vec{r}\, e^{i\vec{q}\cdot\vec{r}} \exp\left\{ \frac{1}{2}\left[D(\vec{r}+r_F^2\vec{q}) + D(\vec{r}-r_F^2\vec{q}) - 2D(\vec{r}) - 2D(r_F^2\vec{q}) \right] \right\},
\end{aligned}
\tag{2.4.2}
$$

where

$$
D(\vec{r}) \equiv \left\langle [\phi(\vec{r}) - \phi(0)]^2 \right\rangle = 2[C(0) - C(\vec{r})]
\tag{2.4.3}
$$

is the structure function of ϕ. One can interpret $q^2\tilde{W}(q)$ as the mean-square amplitude of flux variations $\Delta F/F$ per octave on the length scale $R = 2\pi/q$ in the observer's plane. If v is the relative velocity of the earth and the pulsar transverse to the line of sight, then R corresponds to a time scale $t = R/v$.

The second and third terms in equations (2.3.4) express the large contribution of scales near r_{out} to the variance of the phase and, when $\beta > 4$, to the phase gradient as well. They *cancel* from (2.4.2) because only second and higher derivatives of ϕ are responsible for focussing and de-focussing the rays. However, these terms can contribute importantly to timing noise (Lovelace 1970; Armstrong 1984; Blandford, Narayan and Romani 1984). For questions concerning the flux alone, the precise value of r_{out} is unimportant provided that it is larger than all relevant length scales in the screen: in particular, r_{out} should be larger than the refractive scale r_{ref}, defined by

$$
r_{ref} = \theta_{rms} z \equiv 2\pi/q_{ref},
\tag{2.4.4}
$$

where θ_{rms} is the r.m.s. instantaneous size of the scatter-broadened image. We shall assume henceforth that $r_{out} \gg r_{ref}$.

For the purposes of evaluating (2.4.2), it is convenient to make the dependence of D on r_c simple and explicit. We assume

$$
\begin{aligned}
D(\vec{r}) &= -\frac{Q_0'}{2}\left[(r^2 + r_c^2)^{(\beta-2)/2} - r_c^{\beta-2} \right] && (\beta < 4) \\
D(\vec{r}) &= -\frac{Q_0'}{2}\left[(r^2 + r_c^2)^{(\beta-2)/2} - r_c^{\beta-2} + \frac{q_{out}^{4-\beta}\,\Gamma(\beta/2)}{2^{4-\beta}\,\Gamma((6-\beta)/2)} r^2 \right] && (\beta > 4)
\end{aligned}
\tag{2.4.5}
$$

where Q_0' is the coefficient of $r^{\beta-2}$ in (2.3.4), and has the sign of $(\beta - 4)$. This form of $D(\vec{r})$ has the right properties to describe the power spectrum (2.3.3), and corresponds to a particular choice for the functional form of the cut-off of $Q(q)$ outside the power-law range.

GN have discussed the case $r_c = 0$ under the assumption that $|A| \gg 1$, where

$$A \equiv Q_0' r_F^{\beta-2}. \qquad (2.4.6)$$

This is equivalent to assuming strong scintillation since $|A|$ is essentially the typical phase difference across the Fresnel length r_F. Their results for $q^2 \tilde{W}(q)$ for the cases $(A, \beta) = (-10^3, 11/3)$ and $(10^3, 4.3)$ are shown by the solid lines in Figures 2 and 3, where lengths have been normalized to r_F (*ie.*, $\eta = q r_F$). We see that when $r_c = 0$ and $|A| \gg 1$, almost all of the power in $\Delta F/F$ resides in two prominent peaks.

The high spatial frequency peak in Figures 2 and 3 reflects the power in *diffractive* scintillation. Rays from different parts of the scatter-broadened image interfere to form chaotic fringes on the observer's plane with characteristic linear separation $r_{dif} \sim 1/k\theta_{rms}$. Comparing with (2.4.4) and (2.2.2), we find

$$r_{dif} = r_F^2/r_{ref} \equiv 2\pi/q_{dif}. \qquad (2.4.7)$$

The total power in diffractive scintillation is always ~ 1 for $|A| \gg 1$.

The low spatial frequency peak in Figures 2 and 3 is due to *refractive* effects. Phase fluctuations with wavelength $\sim r_{ref}$ on the screen focus or defocus rays toward the observer, and give rise to flux variations on the observer's plane at the same scale. GN show that the power in refractive fluctuations decreases with $|A|$ if $\beta < 4$ but depends only on β if $4 < \beta < 6$. In general, refractive flux variations are more pronounced for $\beta > 4$ then for $\beta < 4$, as shown by Figs. 2 and 3.

2.5 Power in Flux Variations in the Presence of an Inner Scale

Using techniques similar to those described by GN, the spectrum of flux variations $\tilde{W}(q)$ can be calculated for the case $r_c \neq 0$. The details are given in Appendix A. The two dashed lines in Figs. 2 and 3 show $q^2 \tilde{W}(q)$ for a large inner scale, $r_c = 0.1 r_{ref}$, and for a smaller cut-off, $0.01 r_{ref}$. In Fig. 2, it was necessary to normalize Q_0 in proportion to $(r_c/r_{dif})^{4-\beta}$ in order to maintain the positions of the diffractive and refractive peaks. This is because for $\beta < 4$ the image size θ_{rms} is determined by the phase fluctuations on the scale r_c, so that $\theta_{rms} \sim k^{-1}[D(r_c)]^{1/2}/r_c \propto r_c^{(\beta-4)/2}$ from (2.4.5). An important consequence of the renormalization of Q_0 is that the refractive peak increases in height, an effect that has been noted by Coles *et al.* (1987). For $\beta > 4$, the phase

gradient is dominated by the largest scales, hence by r_{out}; but since the r.m.s. phase *curvature* decreases with scale ($\propto r^{(\beta-6)/2}$), the effect of the power at r_{out} is mainly to shift the image without distorting it, provided that $r_{out} \gg r_{ref}$. The scale that controls θ_{rms} is then the one whose focal length $\approx z$, which implies that its linear size must be $\approx z\theta_{rms} = r_{ref}$. (These arguments have been given by BN and are further developed in §3.) From this observation one can determine θ_{rms} self-consistently in terms of z and Q_0 (cf. BN), but for now it is sufficient to note that varying r_c does not affect θ_{rms} for $\beta > 4$, unless $r_c \geq r_{ref}$.

A dramatic effect of the finite inner scale r_c is that the "valley" in $q^2\tilde{W}(q)$ between the refractive and diffractive peaks is filled in with a region of constant power per logarithmic interval. According to the results of Appendix A, the filled region should extend over the range of q from

$$q_{int} \equiv r_F^{-2}|Q_0'|^{-1/2}r_c^{(4-\beta)/2} \equiv 2\pi/r_{int} \tag{2.5.1}$$

at the left to

$$q_{cdif} \equiv 2\pi r_c/r_F^2 \equiv 2\pi/r_{cdif} \tag{2.5.2}$$

at the right. The "intermediate" length scale r_{int} can be expressed in terms of the other scales as

$$r_{int} = r_{ref}, \quad \beta < 4 \tag{2.5.3}$$

$$r_{int} = r_{ref}\,(r_c/r_{ref})^{(\beta-4)/2}, \quad \beta > 4. \tag{2.5.4}$$

The length scale r_{cdif} is the diffractive scale ($1/k\theta$) associated with an image of projected size r_c on the thin screen. The height of the filled region in Figs. 2, 3 is determined by the height of the $r_c = 0$ spectrum at $q = q_{cdif}$. Hence the total power contributed by the filled region is approximately

$$\left[\frac{\Delta F}{F}\right]_c^2 = \int\limits_{q_{int}}^{q_{cdif}} qdq\tilde{W}(q) = q_{cdif}^2\tilde{W}(q_{cdif})\ln(q_{cdif}/q_{int}). \tag{2.5.5}$$

We can estimate the importance of the contribution (2.5.5) by noting that if $r_c = 0$, the total power in diffractive scintillations is always unity and that in the range $r_F^{-1} \leq q \leq q_{dif}$,

$$\tilde{W}(q) \equiv \begin{cases} \tilde{W}(q_{dif}) & 2 < \beta < 4 \\ \tilde{W}(q_{dif})(q_{dif}/q)^{\beta-4} & 4 < \beta < 6, \end{cases} \tag{2.5.6}$$

(cf. GN). Hence, since $r_{dif}r_{ref} = r_c r_{cdif} = r_F^2$,

$$\left[\frac{\Delta F}{F}\right]_c^2 \sim 2\,(r_c/r_{int})^2 \ln\left(r_c r_{int}/r_F^2\right) \tag{2.5.7}$$

Note that $[\Delta F/F]_c$ increases rapidly as $r_c \to r_{ref}$.

Our results seem paradoxical, particularly for $\beta > 4$, where we have no need to renormalize $|A|$: by *decreasing* the power in ϕ, we have *increased* the power in $\Delta F/F$. The explanation lies in the development of caustics, about which we have more to say in §3.4.

3. Physics of Caustics

In the previous section we found that the presence of an inner scale in the phase fluctuations introduces a new component in the flux spectrum. We associated this fluctuation power with caustics, which we study in greater detail here. In §3.1 we give a geometrical optics description of caustics, borrowing freely from the notions of catastrophe theory. We classify the kinds of caustics that are allowed and identify the ones that could be relevant for interstellar radio scintillation. We find that the pure geometrical optics approach leads to infinite flux for an observer located exactly on the caustic, whereas in reality this divergence would be controlled by the finite wavelength of the radiation. This leads us to investigate the effect of diffraction on caustics. §3.2 introduces certain useful physical concepts about the wavefront and the nature of images. In particular, we describe a way of estimating the flux due to an image. Using these ideas, we discuss in §3.3 the effect of wave optics on caustics, considering in detail the fold caustic. A higher order caustic, the cusp, is described in Appendix B. Finally, we show in §3.4 that caustics do indeed produce the new component in the flux variation spectrum.

3.1 Caustics and Catastrophe Theory – Geometrical Optics

As in section 2, we consider a point source at infinity and a thin phase-changing screen at a distance z from the observer. The wave amplitude $\psi(X, Y)$ received by an observer at $\vec{R} = (X, Y)$ is given by (2.2.3), (2.2.4). Note that the phase $\phi(x, y) \propto k^{-1}$ because of dispersion in the interstellar medium, and the second geometrical term is $\propto k$. The integral (2.2.3) is dominated by points of stationary phase on the screen that satisfy

$$
\begin{aligned}
\frac{\partial \zeta}{\partial x} &= \frac{\partial \phi}{\partial x} + \frac{(x - X)}{r_F^2} = 0, \\
\frac{\partial \zeta}{\partial y} &= \frac{\partial \phi}{\partial y} + \frac{(y - Y)}{r_F^2} = 0.
\end{aligned}
\tag{3.1.1}
$$

For the moment we will ignore the dispersion in ϕ and consider the geometrical optics limit $k \to \infty$ where we can talk of "rays". In this limit, we can say that the only rays that an observer at (X, Y) receives are those that emerge from points on the screen satisfying (3.1.1). Each such ray represents an image of the point source. In fact (3.1.1) provides a many to one mapping between each point (x, y) on the screen at which a ray from the source is incident, and the point (X, Y) on the observer plane at which the ray is received. Keeping this mapping in mind, we consider below the trajectories of a bundle of rays emanating from the point source.

Rays emanating from a point source can cross, so that an observer located at the crossing point may see two or more images of the source. As the observer moves, or as the conditions of observation change, two or more rays may coalesce. The point, line or surface where this occurs is a *caustic*. It is an envelope formed by the family of rays from the point source. In the geometrical optics approximation a point source observed from a caustic will subtend a finite angular size and therefore will be infinitely amplified. Although there are an infinite number of ways that rays can coalesce on a caustic, they can be classified into a few elementary forms or *catastrophes* whose important features and scaling laws are structurally stable to small perturbations. In general there is a hierarchy of catastrophes corresponding to the local phase being flat to successively higher order in a polynomial expansion of the excess phase about points satisfying (3.1.1). The classification of structurally stable catastrophes according to such a hierarchy was developed by Thom (1975), Arnol'd (1976) and others, and has found application in a number of fields (e.g. Poston and Stewart 1982). Berry and Upstill (1980) have summarized the application to wave propagation in random media. We present here a simplified discussion of the main features, using the concepts introduced in the previous sections.

One way to trace rays is to evaluate the optical distance or phase $\zeta(s_i) = \omega \int n dr / c$ where $n(\mathbf{r})$ is the refractive index along a virtual path parametrised by a set of *state variables* s_i (e.g. coordinates x, y on the scattering screen). Actual rays correspond to extrema of $\zeta(s_i)$ (eq. 3.1.1), according to Fermat's principle, and moving the observer represents changing *control variables* C_i (e.g. observer coordinates X, Y). Finding the ray(s) that pass through some point defines a gradient map from C_i to s_i and caustics are singularities of this map, where the Hessian $\mid \partial^2 \zeta / \partial s_i \partial s_j \mid$ vanishes. In the case of equation 2.2.4, at least one of the principal curvatures of the phase surface $\zeta(x, y; X, Y)$ passes through zero at a caustic.

If we confine attention to catastrophes with up to five rays there are seven elementary forms. These are conventionally described analytically by expanding the phase in a power series around the caustic using the minimum number of state variables (one or two) and control variables (up to four) needed to exhibit the catastrophe. In mathematical language, two catastrophes are said to be equivalent if one can be transformed into the other using a diffeomorphism of the control variables and, at each point in control space, a diffeomorphism of the state variables.

The classification of elementary catastrophes can be understood (though not properly derived) from basic considerations. In the simplest case (the fold) just two rays coalesce. If we regard the angle between the rays that represent the two images as being a function of the C_i , then just one control variable need be varied in the generic case to cause this angle to vanish. When n images coalesce, $n - 1$ angles must vanish and we must search for the catastrophe in a control space of dimensionality (codimension) $n - 1$. That is to say, we need $n - 1$ control variables to describe an elementary catastrophe involving n rays. *

Catastrophes that can be described using one state variable are called cuspoids and are known successively as the fold ($n = 2$), cusp ($n = 3$), and swallowtail ($n = 4$). The fold involves two images and so the phase must have two turning points, i.e. it must be cubic in the state variable (see Fig. 4). It is conventional to choose the coefficient of the linear term to be the single control variable, and as we are using the simplest analytical description of the catastrophe we remove the quadratic term by an elementary change of variable and drop the constant. Hence the phase can be expanded locally as a standard polynomial.

$$\bar{\zeta}_{fold} = s^3/3 + Cs \qquad (3.1.2)$$

for a suitably defined state variable s and control parameter C. Similarly, the standard form of the cusp is

$$\bar{\zeta}_{cusp} = s^4/4 + C_2 s^2/2 + C_1 s \qquad (3.1.3)$$

and so on.

---

* In fact, we don't necessarily need to vary $n - 1$ control variables to make n images coalesce. In the case of the swallowtail catastrophe, discussed in Fig. 5, four images can be made to vanish by varying only two control parameters. However, we do need to vary $n - 1 = 3$ independent parameters to find the full singularity.

One state variable is not always adequate as we demonstrate in Fig. 5. Elementary catastrophes requiring two state variables are known as umbilics; by sketching topologically distinct nestings of contours it is easy to verify that a minimum of 4 rays must be involved for such events. There are two distinct possibilities when $n = 4$, the hyperbolic and the elliptic umbilic (see Fig. 5). The character of the remaining two elementary five ray catastrophes is similarly derived.

Although this examination of state variable space has enabled us to classify the elementary catastrophes, we are more interested in the behaviour of the observed images as we vary the control parameters. Referring to equation (3.1.2) for the fold we see two images when $C < 0$, merging at $C = 0$ at the caustic and vanishing for $C > 0$. In the case of a cusp, three images can become one image as the two control parameters C_1, C_2 are varied. (Fig. 4) In this way we can subdivide the control space into distinct regions identified by the number of images of a point source observable from within them. As an observer crosses the boundary from one region to another, he will generally encounter a fold catastrophe and the number of images will change by ± 2. However occasionally he may encounter a higher order catastrophe, when the phase surface will be flatter to higher order. The images will appear to brighten and, in the case of the swallowtail and the hyperbolic umbilic, the number of images may change by ± 4.

It is in determining the types of caustic that will be encountered that the power of catastrophe theory becomes apparent. If, as is the case for a radio telescope passing through the wave field formed by interstellar irregularities, there is only one control variable (time or equivalently one observer coordinate, say X), a single control variable is all that can be explored. We therefore only expect to encounter folds (though we may pass close to a cusp). However, if we were able to explore two observer coordinates, e.g. by taking a photograph or, as in the present application, by varying the wavelength of observation, then we can also expect to encounter cusps. In order to encounter higher order catastrophes, we should have to explore three or more coordinates, *eg.* X, Y and z. These conclusions are independent of the complexity of the ray paths (i.e. the number of state variables necessary to specify them.)

Powerful as these results may be, they do not allow us to quantify the contribution of caustics to the flux variations. To do this, we must revert to a specific model of the propagation and the conditions of observation and we use the thin screen model of §2.1. This also allows us to illustrate the general principles we have just developed. We first consider the fold catastrophe.

Our procedure is to expand $\zeta(x, y; X, Y)$ in a Taylor series around the singularity in the mapping, translating and rotating the coordinate system for convenience and retaining only those terms that can dominate the total phase and its first two derivatives over the region of interest. We fix the origin of the (x, y) coordinates on the fold and orient the axes along the principal axes of the phase curvature tensor with the curvature vanishing along the x direction. The (X, Y) axes are oriented parallel to the (x, y) axes so that the caustic lies along the Y axis. The phase $\zeta(x, y; X, Y)$ then takes the form

$$\zeta(x, y; X, Y) = \frac{1}{3}ax^3 + \frac{1}{2}bx^2y + \frac{1}{2}cy^2 - \alpha x - \beta y, \qquad a > 0, \tag{3.1.4}$$

where $a = \phi_{xxx}/2$, $b = \phi_{xxy}$, $c = \phi_{yy} - k/z$. The state variables are x, y, and the control variables are $\alpha = kX/z$, $\beta = kY/z$. The term in x^2 vanishes with the curvature, the axis rotation has removed the xy term and higher order terms, xy^2, y^3 are demonstrably sub-dominant in the vicinity of the stationary points.

The rays are located at the stationary points of ζ, i.e.

$$
\begin{aligned}
x &\sim \pm \left(\frac{\alpha}{a}\right)^{1/2} + \frac{\alpha b^2}{4a^2c} - \frac{\beta b}{2ac} \\
y &\sim \frac{\beta}{c} - \frac{\alpha b}{2ac}
\end{aligned}
\tag{3.1.5}
$$

plus higher order terms. Two images can be seen when $\alpha > 0$. The Hessian is

$$
\begin{vmatrix}
\partial^2\zeta/\partial x^2 & \partial^2\zeta/\partial x \partial y \\
\partial^2\zeta/\partial x \partial y & \partial^2\zeta/\partial y^2
\end{vmatrix}
= (2ax + by)c - b^2x^2 \sim \pm 2ac(\alpha/a)^{1/2}
\tag{3.1.6}
$$

retaining only the leading terms. As the magnification is the reciprocal of the Hessian, it diverges near the origin on the critical line $y = -2ax/b$ which (since a and b are of the same order) has an arbitrary orientation with respect to the principal axes of the curvature tensor.

We wish to relate the specific form of the fold catastrophe produced by a single screen (eq. 3.1.4) to the standard polynomial (eq. 3.1.2). To do this, we must first augment the polynomial with the simplest terms describing the extra state variable and control variable C_2, i.e.

$$\bar{\zeta} = s^3/3 + C_1S + t^2/2 + C_2t. \tag{3.1.7}$$

The state and control variables are then related, to lowest order, through

$$s = a^{1/3}x$$

$$t = c^{1/2}y$$

$$C_1 = -a^{-1/3}\alpha$$

$$C_2 = \alpha b/(2ac^{1/2}) - c^{-1/2}\beta \tag{3.1.8}$$

In other words, the observer coordinate system must be sheared to bring the catastrophe into standard form. In Appendix B, we repeat this exercise for the cusp.

3.2 Wavefronts, Images, and Power Spectra

We begin our discussion of diffraction in caustics by introducing in this section certain key concepts related to the diffraction integral (2.2.3) and the power spectrum (2.3.3). We will make frequent use of the length scales r_F, r_{ref}, r_{dif}, r_c, r_{int} and r_{cdif} introduced in §2 and listed in the Table.

A crucial concept in what follows is that of the "image" of a point source as projected on the thin scattering screen. Consider first the case when the excess phase $\phi(x, y)$ vanishes. The diffraction integral (2.2.3) involves only the geometrical contribution to $\zeta(x, y, X, Y)$ which has a stationary point at $(x(X, Y), y(X, Y))$. The dominant contribution to $\psi(X, Y)$ arises from a coherent region of radius $\sim r_F$ around this point, and the contribution from the rest of the screen cancels out due to the oscillating phase. We then say that the image of the source is centered on (x, y) and has a coherent "size" $\sim r_F$. A more detailed discussion of the physics of the diffraction integral can be found in standard texts (e.g. Born and Wolf 1980, Hecht and Zajac 1974), which formalize the idea of the "coherent patch" by means of a geometrical construction in the complex plane using the Cornu spiral.

Consider now the case of non-vanishing $\phi(x, y)$. Provided a particular restriction that is discussed below is satisfied, it is still correct to say that the diffraction integral (2.2.3) is dominated by regions of stationary phase, i.e. regions in the vicinity of points (x, y) that satisfy equations (3.1.1). For the strong scintillation regime in which we are interested, there will in general be a large number of points on the screen that satisfy these conditions. We will therefore call the set of such points the

image of the source, and each individual point a *sub-image*. To calculate the flux contributed by a single sub-image located at (x_I, y_I), we make a local quadratic expansion of $\zeta(x, y)$. Let us assume that the x, y axes are aligned with the principal axes of the curvature tensor of ζ. We then have

$$\zeta(x, y) \sim \zeta(x_I, y_I) + \frac{1}{2}\frac{\partial^2 \zeta}{\partial x^2}(x - x_I)^2 + \frac{1}{2}\frac{\partial^2 \zeta}{\partial y^2}(y - y_I)^2. \tag{3.2.1}$$

As before, we can say that the dominant contribution to the flux arises from a coherent region around (x_I, y_I), but now the coherent patch is in general an ellipse, whose semi-axes ℓ_x, ℓ_y are given by the condition that ζ changes by ~ 1 radian, i.e.

$$\ell_x \sim \left(\frac{\partial^2 \zeta}{\partial x^2}\right)^{-1/2}, \qquad \ell_y \sim \left(\frac{\partial^2 \zeta}{\partial y^2}\right)^{-1/2}. \tag{3.2.2}$$

The flux F_s contributed by the sub-image is the square of the wave-amplitude, which leads to

$$F_s \sim \ell_x^2 \ell_y^2 / r_F^4. \tag{3.2.3}$$

This discussion in terms of individual sub-images is valid so long as the quadratic expansion (3.2.1) extends over an area much larger than the coherent central patch. In other words, the power series of ζ must be valid for changes in $\zeta \gg 1$. As we show below this condition holds in all of the situations considered in this paper because of the presence of the cut-off in the spectrum of phase fluctuations. In the case of an extended spectrum with no cut-off, more care is needed. One should identify a wave-vector q_{max} such that the rms phase fluctuations due to $q > q_{max}$ is $\lesssim 1$ radian (e.g Gapper and Hewish 1981, RNB), and the power from these large wave-vectors should be filtered out. The residual $\zeta(x, y)$ can then be approximately analyzed as above.

We now consider the effect of the power spectrum $Q(q)$ given in (2.3.3). For the particular power law form assumed, the mean square phase fluctuation varies with q as $\sim k^{-2}Q_0 q^{2-\beta}d(\ln q)$. Let us define ϕ_c to be the rms phase fluctuation due to scales near the cut-off scale r_c, i.e.

$$\phi_c \equiv k^{-1}Q_0^{1/2}r_c^{(\beta-2)/2}. \tag{3.2.4}$$

Hills and valleys in $\phi(x, y)$ tend to have linear sizes $\sim r_c$, and ϕ_c measures the typical phase difference between a neighbouring hill and valley.

The geometrical scattering angle at any point on the screen is given by $k^{-1}\overrightarrow{\nabla\phi}$, and so the rms scattering angle due to scales with wave-vector q is $\theta(q) \sim k^{-1}Q_0^{1/2}q^{(4-\beta)/2}$. We then see that the intermediate-scale r_{int} defined in (2.5.1) is related to the scattering angle at the cut-off scale by

$$r_{int} \sim z\theta(q_c) \sim r_F^2 \phi_c / r_c. \tag{3.2.5}$$

The total mean square scattering angle from all scales is

$$\langle \theta^2 \rangle \sim \frac{k^{-4}}{(2\pi)^2} \int \int d^2 q q^2 Q(q) = Q_0 \frac{k^{-4}}{2\pi} \int\limits_{q_{out}}^{q_c} d(\ln q) q^{4-\beta}. \qquad (3.2.6)$$

The behaviour of this integral depends critically on the sign of $4 - \beta$. When $\beta < 4$, $q^{4-\beta}$ is maximal at the upper limit q_c; hence

$$\theta_{rms} \sim \theta(q_c) \sim \left(Q_0/k^4 r_c^{4-\beta} \right)^{1/2}, \quad \beta < 4. \qquad (3.2.7)$$

This is the angular size of the image, and associated with it is the refractive scale (cf eq. 2.4.4)

$$r_{ref} = r_{int} \sim r_F^2 \phi_c/r_c, \quad \beta < 4. \qquad (3.2.8)$$

When $\beta > 4$, the refractive scale is no longer given by (3.2.8) since the scattering angle is dominated by small q, i.e. large length scales. In fact, the scattering angle diverges unless there is an "outer scale" to the spectrum. However, the divergence does not concern us since the largest scales only steer the mean position of the image without contributing to its broadening. The image broadening is in fact determined by the largest scale which can focus radiation at the observer, i.e. which satisfies the condition $r_F^2 \nabla^2 \phi \sim 1$. Since the contribution to $\partial^2 \phi / \partial \bar{r}^2$ from wavevectors $\sim q$ varies as $\left(Q_0 k^{-2} q^{6-\beta} \right)^{1/2}$, therefore we obtain

$$r_{ref} \sim \left(r_F^2 \phi_c \right)^{2/(6-\beta)} r_c^{-(\beta-2)/(6-\beta)}, \quad \beta > 4. \qquad (3.2.9)$$

Note that (3.2.8) and (3.2.9) agree for the marginal $\beta = 4$ case. Note also that (3.2.5) and (3.2.9) are consistent with (2.5.4), which gives the relation between r_{int} and r_{ref} for $\beta > 4$ (the two scales are equal for $\beta < 4$).

The theory of section 2 and Appendix A requires three conditions to be valid, which can be written as

$$r_{ref} \gg r_F, \qquad (3.2.10)$$

$$r_{ref} \gg r_c, \qquad (3.2.11)$$

$$r_c \gg r_F^2/r_{int}. \qquad (3.2.12)$$

The first condition ensures strong scattering, and the second ensures multi-path propagation, both of which are crucial for the occurrence of strong scintillation. The third condition can be rewritten with the help of (3.2.5) as

$$\phi_c \gg 1. \qquad (3.2.13)$$

But this is precisely the condition we required earlier for the concept of sub-image to make sense. We thus see that the filling up of the valley in $q^2 \tilde{W}(q)$ that we discussed in §2 occurs only when sub-images can be distinguished individually.

Let us now estimate the mean flux of sub-images. From (2.2.4), we have

$$\ell_x, \ell_y \sim \left(\frac{\partial^2 \zeta}{\partial \vec{r}^2}\right)^{-1/2} \sim \left[\left(\frac{\partial^2 \phi}{\partial \vec{r}^2}\right)_{rms} - \frac{1}{r_F^2}\right]^{-1/2}. \qquad (3.2.14)$$

Applying eq.(3.2.11) we see that

$$r_F^2 \left(\frac{\partial^2 \phi}{\partial \vec{r}^2}\right)_{rms} \sim \left(\frac{r_{ref}}{r_c}\right) \gg 1, \qquad \beta < 4,$$

$$\sim \left(\frac{r_{ref}}{r_c}\right)^{(6-\beta)/2} \gg 1, \quad \beta > 4. \qquad (3.2.15)$$

Thus $(\partial^2 \phi / \partial \vec{r}^2)_{rms} \gg 1/r_F^2$ in all regimes of interest, and so the contribution of the latter to ℓ_x, ℓ_y can be neglected. Equation (3.2.3) then gives the flux from each sub-image as

$$F_s \sim (r_c/r_{ref})^2, \quad \beta < 4, \qquad (3.2.16)$$

$$F_s \sim (r_c/r_{ref})^{(6-\beta)}, \quad \beta > 4. \qquad (3.2.17)$$

Since the mean total flux is conserved, this implies that the mean number of sub-images in the image is

$$N \sim (r_{ref}/r_c)^2, \quad \beta < 4, \qquad (3.2.18)$$

$$N \sim (r_{ref}/r_c)^{(6-\beta)}, \quad \beta > 4. \qquad (3.2.19)$$

Equation (3.2.18) has a transparent interpretation; it says that the image of size r_{ref} is uniformly covered with sub-images with mean separation $\sim r_c$. Equation (3.2.19) is more interesting. The wavefront still has hills and valleys of typical linear size r_c, but the number of sub-images is now $< (r_{ref}/r_c)^2$. The reason is that since the rms scattering angle due to perturbations on any given scale is larger than the broadening due to scattering by all smaller scales, the subimages are clustered at each scale. This results in a reduction of the number of images within r_{ref}^2 and gives the image a hierarchical fractal structure (e.g. GN).

For the discussion of caustics in §3.3 we need estimates of higher derivatives of ζ. The geometrical term does not contribute, and so

$$\left(\frac{\partial^3 \zeta}{\partial \vec{r}^3}\right)_{rms} \sim \frac{\phi_c}{r_c^3}, \qquad (3.2.20)$$

$$\left(\frac{\partial^4 \zeta}{\partial \vec{r}^4}\right)_{rms} \sim \frac{\phi_c}{r_c^4}. \tag{3.2.21}$$

It is assumed that the cut-off of the spectrum beyond q_c is at least as steep as q^{-10} for these estimates to be accurate; weaker cut-offs will give somewhat different coefficients.

3.3 Diffraction Effects in Caustics

As we discussed in §3.1, a caustic is obtained when one or both of the principal curvatures of $\zeta(x, y)$ vanishes at a sub-image. Equations (3.2.2), (3.2.3) show that the flux then diverges within the quadratic approximation. Indeed, in the limit of pure geometrical optics, the image becomes infinitely bright. However, for a finite wavelength, diffraction controls this divergence and there is a finite magnification of each subimage. To treat this case, we will need to generalize the discussion of §3.2 by introducing the effects of higher order terms in the polynomial expansion of ζ. The idea of the coherent patch is still correct and so the flux is determined by the coherent contribution from the region inside the $\Delta\zeta \sim 1$ contour around the sub-image. However, when the curvature in a particular direction vanishes, as it does at a caustic, the dimension of the coherent patch along that direction is no longer determined by the curvature but instead by the first non-vanishing term in the Taylor series.

Let us consider the fold caustic. From equation (3.1.4), the phase ζ takes the form

$$\zeta(x, y, X, Y) = \frac{1}{3}ax^3 + \frac{1}{2}bx^2y + \frac{1}{2}cy^2 - \frac{(Xx + Yy)}{r_F^2}, \quad |x|, |y| \ll r_c, \tag{3.3.1}$$

$$|a| \sim |b| \sim \phi_c/r_c^3, \quad |c| \sim \phi_c/r_c^2. \tag{3.3.2}$$

In the discussion that follows, we restrict our attention to the sub-images found in the small area, $|x|, |y| \lesssim r_c$. There are, of course, $\sim N$ other sub-images in the image, but they do not influence the behaviour of the caustic. The locations of sub-images in this "local" area for a given observer position (X, Y) are described by (3.1.5) which, to lowest order, gives

$$X = r_F^2\left(ax^2\right) \sim r_{int}\left(x/r_c\right)^2, \tag{3.3.3}$$

$$Y = r_F^2\left(cy + bx^2/2\right) \sim r_{int}\left((y/r_c) + (x/r_c)^2/2\right), \tag{3.3.4}$$

where we have used (3.3.2), (3.2.5). The caustic is defined by the vanishing of the determinant of the curvature tensor (3.1.6), which gives a locally straight line at arbitrary orientation on the

(x,y) screen and, to lowest order, the line $X = 0$ on the ground. This line in the observer plane is perpendicular to the x-axis, viz. the direction on the screen along which the curvature vanishes. Globally, however, the quadratic term in (3.1.6) indicates that the caustic is curved, with a radius of curvature $\sim r_{int}$. In fact, we can state on the basis of this result that each patch on the scattering screen of size $\sim r_c \times r_c$ produces a fold caustic *loop* on the observer plane of "radius" $\sim r_{int}$, as shown in Fig. 6. Note that, since $r_{int} \propto z$, the length of these caustic loops (and hence the frequency of fold crossings) will increase with z while the number of loops remains constant. This is a crucial difference between one and two-dimensional screens. In the former case the mean fold spacing is just $\sim r_c$, independent of z.

We now investigate the variation of received flux for observer positions in the vicinity of the caustic. The qualitative features are shown in Fig. 4. It is sufficient to consider the line $Y = 0$ on the observer plane, locally perpendicular to the caustic. The behaviour we determine here can be applied elsewhere on the caustic by simple translation. To the order at which we are working, $Y = 0$ implies $y = 0$ (eq. 3.3.4), and so (3.3.1) simplifies to the standard form (cf. eq. 3.1.7)

$$\zeta(x, y, X) = \frac{1}{3}ax^3 + \frac{1}{2}cy^2 - \frac{Xx}{r_F^2}, \qquad a > 0. \tag{3.3.5}$$

The maximum flux is clearly received at $X = 0$, and the corresponding image is centered on $x = y = 0$. At this point, the curvature of ζ parallel to x vanishes and so the size of the coherent patch along x is determined from the third derivative of ζ. The area of the coherent patch on the screen is thus $\sim (a^{-1/3}) \times (c^{-1/2}) \sim r_c^2/\phi_c^{5/6}$. The received flux is thus

$$F_{f,max} \sim \frac{r_c^4}{\phi_c^{5/3} r_F^4} \sim \phi_c^{1/3} \bar{F}_s \sim \phi_c^{1/3}/N. \tag{3.3.6}$$

We see that the maximum flux is greater than the mean flux of a sub-image by the factor $\phi_c^{1/3}$. For $X \neq 0$, equation (3.3.3) with $y = 0$ can be rewritten as

$$x = \pm \left(\frac{X}{ar_F^2}\right)^{1/2} \sim \pm \left(\frac{X}{r_{int}}\right)^{1/2} r_c, \quad y = 0. \tag{3.3.7}$$

Equation (3.3.7) reveals the characteristic feature of the fold caustic, viz. that there are two sub-images for observer positions on one side of the caustic (the side $X > 0$ when a is positive), and no sub-images on the other side ($X < 0$). Note also the square-root mapping between the image and observer positions.

For non-zero positive X, the extremum at $x = +(X/ar_F^2)^{1/2}$ is at a minimum of ζ with respect to variations in x, while the other extremum is at a maximum. The phase difference between the two extrema is

$$\Delta\zeta = \frac{4}{3}ax^3 \sim \phi_c \left(\frac{X}{r_{int}}\right)^{3/2} \tag{3.3.8}$$

When $\Delta\zeta < 1$, the two extrema cannot be distinguished as different sub-images, and we must think of the image as a single coherent patch with flux $\sim F_{f,max}$. For $\Delta\zeta > 1$, the two images separate and can produce interference fringes. Thus, the width X_f of the central fringe in the fold caustic is given by the condition $\Delta\zeta \sim 1$, which leads to

$$X_f \sim r_{int}/\phi_c^{2/3}. \tag{3.3.9}$$

Note that this is smaller than the size of the caustic loop, which is $\sim r_{int}$.

For $X > X_f$, we can compute the fluxes of the two images through (3.2.2), (3.2.3). We find

$$\ell_x \sim (ax)^{-1/2} \sim \phi_c^{-1/2}(X/r_{int})^{-1/4}r_c,$$
$$\ell_y \sim c \sim \phi_c^{-1/2}r_c, \tag{3.3.10}$$

and this leads to an observed flux variation of the form

$$F(X) \sim \left(\frac{X_f}{X}\right)^{1/2} F_{f,max}. \tag{3.3.11}$$

Thus the flux falls off as $X^{-1/2}$ away from the caustic, in agreement with the geometrical optics result. However, (3.3.11) only determines the *mean* flux. The two images will in fact interfere with each other, leading to flux oscillations with separation $S(X)$ varying as

$$S(X) \sim (d\Delta\zeta/dX)^{-1} \sim \left(\frac{X_f}{X}\right)^{1/2} X_f. \tag{3.3.12}$$

The fringe spacing decreases away from the caustic. For negative X, there is no extremum in the phase function; hence there are no images and the flux falls off rapidly on a scale $\sim X_f$.

The entire discussion above was for the case of the fold caustic. A similar analysis can be made for higher order caustics. Appendix B gives the results for the cusp caustic.

3.4 Spectrum of Flux Variations due to Caustics

We now make the connection between the results of section 2 and what we have discussed so far in section 3. In the asymptotic regime ($\phi_c \to \infty$) considered in section 2, the fold caustic will dominate over the cusp caustic, and so we limit ourselves to estimating the contribution of folds to the flux spectrum.

We have seen that folds generally form loops in the observer plane of size $\sim r_{int}$; hence the smallest q at which the fold can contribute is $\sim q_{int}$. The smallest length scales in the flux variation arise from the fine interference fringe structure in the fold. From eq. (3.3.14) we see that for $X \sim r_{int}$, this gives $q_{max} \sim r_{int}^{1/2}/X_f^{3/2} \sim \phi_c/r_{int} \sim q_{cdif}$. These are precisely the two limits between which the new component in the flux spectrum occurs (see equations 2.5.1, 2.5.2).

Let us now consider the shape of the spectrum. We substitute (3.3.8) into (2.2.3) and do the integral over y to obtain the wave-amplitude at $(X, 0)$

$$\psi(X, 0) = \frac{1}{\sqrt{cr_F^2}} \int dx \, \exp\left[\frac{i}{3} ax^3 - i\frac{k}{z} Xx\right].$$ (3.4.1)

The Fourier transform of the one-dimensional flux pattern along X is

$$\begin{aligned}
\tilde{I}(q) &\equiv \int_{-\infty}^{\infty} dX \psi^*(X, 0)\psi(X, 0)e^{iqX} \\
&= \frac{1}{\sqrt{c}} \int_{-\infty}^{\infty} dX \, \exp\left[\frac{i}{3} a \left(\frac{3z}{k} qX^2 + \frac{3z^2 q^2}{k^2} X + \frac{z^3 q^3}{k^3}\right)\right].
\end{aligned}$$ (3.4.2)

This gaussian integral can be done exactly and gives a power-law scaling in one dimension

$$|\tilde{I}(q)|^2 \propto q^{-1}.$$ (3.4.3)

In two dimensions, we get $\tilde{I}(\vec{q}) \propto q_x^{-1}\delta(q_y)$ if the fringes are parallel to the Y-direction, as in our calculations. Averaging over all orientations of the fold, we then get

$$|\tilde{I}(q)|^2 \propto q^{-2},$$ (3.4.4)

which is identical to the form found in section 2 and Appendix A for the extra contribution due to the cut-off.

It has been shown by Berry (1977) that for light rays reflecting off a randomly corrugated surface, the flux variations vary as

$$(\Delta F/F)^2 \sim \ln(k)$$ (3.4.5)

as the limit of geometrical optics is approached. Berry further shows that the logarithmic divergence is due specifically to fold caustics. To approach this limit in our problem, we cannot simply take $k \rightarrow \infty$, because in view of the cold-plasma dispersion relation, this would entirely suppress scintillation. Instead, we must hold θ_{rms} fixed as $k \rightarrow \infty$, which translates to fixing $k^{-2}C(r)$; in other words, we maintain a constant amplitude in the fluctuations of the optical path length. When the limit is taken in this way the intensity modulation due to the intermediate length scales, $(\Delta F/F)_c^2$ from (2.5.7), obeys the scaling (3.4.5).

This completes our proof that the extra flux variation power introduced on scales between r_{diff} and r_{ref} by a cut-off in the phase spectrum is due to the effect of caustics, primarily fold caustics.

4. Observation of Individual Caustics

In section 2 we discussed the contribution of the ensemble of caustics to the intensity fluctuation spectrum. In this section and the next, we apply the results of the previous sections to the interstellar medium. In particular, we estimate the required magnitude of r_c if individual caustics are to have observable effects. Unless otherwise stated, we assume $\beta < 4$.

We first estimate the consequences of imposing an inner scale r_c on the density perturbation spectrum for the flux variation power spectrum described in §2. We have a typical observing wavelength $\lambda_m \equiv (\lambda/1m) \sim 1$ and a distance to the effective scattering screen of $D_{kpc} \equiv (D/1\mathrm{kpc}) \sim 1$, giving a Fresnel scale $r_F \sim 2 \times 10^{11}(\lambda_m D_{kpc})^{\frac{1}{2}}$ cm. We can express the strength of the scattering Q_0 in terms of a conventional C_{-4} (Armstrong, et al. 1981, BN) where $Q_0 \propto C_{-4}$ and $C_{-4} \sim 1$ for nearby pulsars. If there were no inner scale, then from the measured scatting angles the refractive scale would be given by

$$
\begin{aligned}
r_{ref} &\sim 3 \times 10^{13}(C_{-4}D_{kpc}^{\beta-1}\lambda_m^{\beta})^{1/(\beta-2)}\mathrm{cm} \qquad \beta < 4 \\
r_{ref} &\sim 3 \times 10^{13}(C_{-4}D_{kpc}^{3}\lambda_m^{4})^{1/(6-\beta)}\mathrm{cm} \qquad \beta > 4
\end{aligned}
\tag{4.1}
$$

(cf. RNB). Imposing a cut-off scale at r_c implies that for a given scattering strength (C_{-4}) the refractive scale would be reduced (for $\beta < 4$). Thus, as discussed in §3, we must increase our estimate of C_{-4} by a factor $(r_c/r_{dif})^{4-\beta}$, $(\beta < 4)$, in order to keep fixed the measured scattering size. For $\beta > 4$, the large scales dominate, so r_{ref} is not affected. RNB also show that for $\beta = 11/3$ and no inner scale we have $(\Delta F/F)_{ref}^2 = 0.015C_{-4}^{-2/5}D_{kpc}^{-11/15}\lambda_m^{-17/15}$, while for $\beta > 4$, GN show

that $(\Delta F/F)^2_{ref} = \beta - 4$. With these numbers, for $D_{kpc} = \lambda_m = C_{-4} = 1, \beta = 11/3$, we have $(\Delta F/F)^2_{dif} \sim 1$ and $(\Delta F/F)^2_{ref} \sim 0.015$. As an example, if we now posit an inner scale $r_c = 7 \times 10^{12}$ (ie. adequate to give prominent individual folds, see below) then, from the discussion in §2.5, $(\Delta F/F)^2_{ref}$ is increased to ~ 0.26 and from (2.5.7) the caustics contribute $(\Delta F/F)^2_c \sim 0.93$. Adding the non-diffractive contributions in quadrature we get $\Delta F/F|_{rms,non-dif} \sim 1.1$, sufficient to explain even the most strongly varying pulsars.

A somewhat larger inner scale may be required to resolve the individual diffraction catastrophes. Nevertheless, the presence of large non-gaussian fluctuations in the slow modulation of pulsar flux suggests that large, detectable events may be occasionally present. To estimate this possibility, a reasonable criterion for detectability is that, at maximum, an individual catastrophe provide as much flux as from the time-averaged image as a whole. Suppose, more generally, we require a fold to have a peak flux α times the mean value; then from (3.3.6) and (3.2.5), in order for the merging images to be observable, we require that

$$r_c \gtrsim 7 \times 10^{12} \alpha^{3/7}(r_{int}/3 \times 10^{13}\text{cm})^{5/7}(r_F/2 \times 10^{11}\text{cm})^{2/7}\text{cm} \tag{4.2}$$

For a cusp, a similar criterion on the peak flux (B.8) gives

$$r_c \gtrsim 4 \times 10^{12} \alpha^{2/5}(r_{int}/3 \times 10^{13}\text{cm})^{3/5}(r_F/2 \times 10^{11}\text{cm})^{2/5}\text{cm} \tag{4.3}$$

From these expressions it is clear that the inner scale should be substantially larger than the Fresnel scale for individual catastrophes to be prominent. Equations (4.2), (4.3) are valid for all values of β. When $\beta < 4$, $r_{int} = r_{ref}$. However, when $\beta > 4$, r_{int}/r_{ref} is given by equation (2.5.4). Thus, the required value of r_c in (4.2) or (4.3) is slightly less in this case—but only very slightly, since observations indicate $|\beta - 4| \leq 0.3$ (GN).

Equations (4.2) and (4.3), together with the condition that some diffractive scintillation be present (viz. $r_c < r_{ref}$), imply that caustics can only cause a strong over-brightening (ie. $\alpha \gtrsim 3$) for a restricted range of cut-off scales and observing frequencies. As shown in figure 7, there is in fact a minimum wavelength below which (for a given α, distance and scattering strength) caustics will not be observed; since at this wavelength the cut-off scale will be roughly the same as the scattering disk, there is an associated minimum inner scale. For $\beta \sim 4$, these are

$$\begin{aligned} \lambda^* &\sim 0.03(C_{-4}D^2_{kpc})^{-1/3} \times \begin{cases} \alpha & \text{(fold)} \\ \alpha^{2/3} & \text{(cusp)} \end{cases} \quad \text{m} \\ r_c^* &\sim 3.5 \times 10^{10}(C_{-4}D_{kpc})^{-1/6} \times \begin{cases} \alpha^2 & \text{(fold)} \\ \alpha^{4/3} & \text{(cusp)} \end{cases} \quad \text{cm} \end{aligned} \tag{4.4}$$

For longer observing wavelengths and larger cut-off scales, there will be a region over which caustics of a given strength may be observed, shown in Figure 7 as the shaded regions between $\lambda_{min} = (r_c/r^*)^{1/2}\lambda^*$ and $\lambda_{max} = (r_c/r_c^*)^{7/11}\lambda^*$ (fold), $(r_c/r_c^*)^{5/7}\lambda^*$ (cusp).

Coles *et al.* have analysed the scintillation data for a number of pulsars and have estimated bounds on an inner scale, based on purely 'refractive' modulation indices and the prevalence of scintillation over a range of observation frequencies. They conclude, on quite conservative grounds, that the average r_c for the various lines of sight can be $\gtrsim 10^{11}$cm, although a very large scatter makes estimates difficult. This is substantially smaller than our fiducial value, but even for such r_c caustic behaviour may be significant for certain lines of sight and observing frequencies, though they will not, in general, dominate the fluctuation statistics. We therefore concentrate on cut-offs of order several times 10^{12}cm.

If the distance to the screen is greater than the focal length for perturbations on the scale r_c then each r_c^2 scale patch should produce, on the average, one caustic. This caustic will form a fold loop in the observer plane of length $2\pi r_{int}$ with, in general, of order one associated cusp. Thus an observer moving across this plane will encounter $\sim r_{int}/r_c^2$ folds/cm. If this motion is due to a typical pulsar velocity of $10^7 v_7$cm/s and we use the standard values for the ISM scales given above, then the mean time between fold crossings is

$$\sim 2\ (r_c/7\times 10^{12}\text{cm})^2 (r_{int}/3\times 10^{13}\text{cm})^{-1} v_7^{-1} \text{days} \tag{4.5}$$

Taking the size of a cusp to be that of the diffraction-broadened maximum (of width $r_{int}/\phi_c^{3/4}$ (B.9)), we can estimate the frequency of cusp encounters to be

$$\frac{r_{int}(r_F{}^2/(r_c r_{int})\ )^{3/4}}{r_c^2}\ v\ s^{-1} \sim \left(\frac{r_{int}r_F{}^6}{r_c{}^7}\right)\frac{v}{r_c}s^{-1} \tag{4.6}$$

where we have, on the average, one cusp per r_c^2. With our standard values, this gives ~ 4 years between cusp encounters. If cusps become fairly bright ($\alpha \gtrsim 3$), then it should be possible to detect several diffraction fringes in the X direction.

While it is the second and higher local derivatives of phase on the screen which determines the sub-image intensity, to detect a given area on the screen as a sub-image there is also the condition that the phase gradient have the appropriate value to steer the rays towards the observer. For $\beta < 4$, this typical bending is the mean scattering angle, θ_{rms}, meaning that the mean position shift of the

observer from the spot directly below the caustic is $\theta_{rms}D = r_{ref}$. (For $\beta > 4$, the bending could in principle be $\gg \theta_{rms}$ if the outer scale were sufficiently large, but dynamic scintillation spectra place important limits on r_{out} in this case [GN].) Since $\theta_{rms} \propto \lambda^2$, this shift is a rather strong function of the observation frequency. This has two consequences. First, in order to resolve a small scale structure, such as the peak of a fold $\Delta X \sim r_F^{4/3} r_{ref}^{1/3} r_c^{-2/3}$, one must observe over a narrow band to avoid image blurring from chromatic aberration. Since $\Delta X_{rms}/r_{ref} \sim 2\Delta\lambda/\lambda$ this condition is (for $\beta \sim 4$)

$$\frac{\Delta\lambda}{\lambda} < \frac{r_F^{4/3}}{r_{ref}^{2/3} r_c^{2/3}} \tag{4.7}$$

With the numbers above and $r_c = 8 \times 10^{12}$cm this is $\Delta\lambda/\lambda < 3 \times 10^{-3}$. Since a typical decorrelation bandwidth for diffractive scintillation is $\sim 50 C_{-4}^{-1} \lambda_m^{-4} D_{kpc}^{-2}$kHz (RNB), i.e. $\Delta\lambda/\lambda \sim 10^{-4}$, it should be possible to filter out the scintillations (and improve the signal-to-noise) without blurring the caustics. The second consequence of the frequency dependence is that one may use the motion of the caustics to scan the observer plane. To find caustics, one might wish to observe for $t_{ref} \sim r_{ref}/v \sim$ 30 days at $\sim$meter wavelengths, monitoring continuously at a single frequency. Alternatively, one could observe ten times at three-day intervals, each time scanning a range $\Delta\lambda/\lambda \sim .05$. With our standard numbers, equation (4.5) predicts ~ 10 folds should be found by either of these procedures.

This frequency dependence coupled with the falloff of the caustic's flux as a function of distance [(3.3.11) for the fold, (B.10) for the cusp] suggests that the best way to search for caustics may be via matched filters in λ, t space. For folds, for example, one would monitor for both image creation events $I \sim t^{-1/2}$ and image destruction events $I \sim (-t)^{-1/2}$. Light curves for the fold and the cusp are shown schematically in Fig. 8. The flux maxima should lie along slopes in the λ, t plane (like dynamic scintillation spectra) with $\Delta\lambda/\Delta t \sim \lambda v/r_{ref}$.

Existing time series of compact, refractively scintillating radio sources show examples of fluctuations that can be interpreted as caustics. The long term pulsar intensity records of Cole, Hesse and Page (1970) and Helfand, Fowler and Kuhlman (1977) show several fluctuations apparently much larger than the rms deviation which might be ascribable to caustic events. A particular example of compact extragalactic radio source fluctuations apparently due to strong refractive focussing can be seen in the striking events observed for three radio quasars by Fiedler et al. (1987). The quasar light curves of these events show two resolved, but sharply peaked maxima bracketing a period of

suppressed flux, and in one instance four strong intensity spikes at a second frequency. The observations suggest an interstellar propagation effect and, with typical galactic distances and velocities, a focusing perturbation of size $\sim 10^{13} - 10^{14}$cm. This focusing can be interpreted as caustic events caused by excess power at large scales (option (iii) in the Introduction) in the form of a single $\sim$Gaussian profile scatterer producing two sets of two fold catastrophes, viewed from a distance of several focal lengths (Romani, Blandford and Cordes 1987). Since the density perturbation δn_e inferred by Fiedler et $al.$ is rather large, the single high density lensing perturbation should be in a special high pressure region ($eg.$ an old supernova remnant) or be magnetically confined. The condition that a single cloud of size $\sim r_{ref}$ dominate the accumulated small scale scattering along the line of sight can be estimated from (4.1) to be

$$n_e \gtrsim 20 \, C_{-4}^{1/2} D_{kpc}^{1/2} \quad \text{cm}^{-3} \tag{4.8}$$

for $\beta \sim 4$ (Lower electron densities are adequate if the density perturbation takes the form of a filament or a sheet). These observations argue strongly for the presence of $10^{13} - 10^{14}$cm-scale, high density focussing plasma structures which can occasionally dominate the distributed scattering of the general ISM. For this component, the statistics will be non-gaussian and a power-law description will not be adequate.

5. Dynamic Scintillation Spectra

The patterns of flux maxima in the (ν, t) plane known as dynamic scintillation spectra have been recognized as important probes of scales larger than r_{dif} for some time ($e.g.$ Hewish 1980). The presence of systematic patterns and periodicities imposed on the mean drift expected for the effect of a random collection of larger scales was clearly described by Roberts and Ables (1982) and has been recently studied further by Cordes et $al.$ (1987). Pulsars are often found to exhibit frequency drifts with bands modulated on a scale several times smaller than the typical band spacing. Occasionally, two or even more periodicities are seen, forming sets of periodic flux modulations at different angles to the mean drift (crossed, or herringbone patterns) as portrayed in Figure 9. One interpretation is that these are the interference patterns formed when there are two (simple periodicities) or three (crossed patterns) bundles of rays interfering at the observer. When there is a single thin screen these

ray bundles must arise from different directions and so the image of the source must be subdivided into a small ($\sim 2-3$) number of bright subimages. In general the interference pattern on the ground is a measure of the separation of the sub-images on the sky, although Cordes *et al.* (1986) have noted that for an extended medium, the correspondence may not be direct. In fact, Cordes and Wolszczan (1986) find the number of coherent fringes to be quite high, typically indicating a subimage spacing 3–5 times the subimage size.

This situation would seem difficult to realize in an extended $\beta < 4$ spectrum since the rms scattering angle is dominated by perturbations on a scale r_{dif}; large scales of the order of the sub image separation would not have enough power to divert the subimages by an angle of order their own angular size. Conversely, an extended $\beta > 4$ power law would have difficulty producing such well separated spots because intermediate scales would have sufficient power to break up the subimages. Indeed, as described in GN, the fractal dimension of the image flux is $6 - \beta < 2$. Hence one expects a hierarchy of scales within the image, and the separation of patches on each scale should be less than or of order the patch size. In addition, such a spectrum must, of course, be truncated at long wavelengths to avoid arbitrarily large drift slopes and non-linear density perturbations.

To account for the form of the periodicities, it again seems necessary to postulate either (i) a cut-off at small scales or (ii) extra power at a large scale. While in some cases the latter picture seems most applicable, at high frequencies, with $r_F \sim r_c$, the inner scale effects can dominate.

If there is an inner scale $r_c \sim r_{ref}$ then there should be only a few sub-images per image spot. In this case the interference patterns could arise from the members of a single caustic; a merging pair (fold) for the case of single periodicities, or a merging triple (cusp) for herringbone patterns. Since r_c is a specific scale in the ISM this will only work for a limited range of frequencies for which the total image size is close to but slightly larger than this scale. Such a picture may be too restricted to account for the apparently common occurrence of such patterns.

A smaller inner scale is possible if each of the bright subimages contributing to the fringe pattern is itself a fold caustic. In this model, the *mean* number of subimages is $>> 1$, but a few are highly-magnified pairs. The typical separation between two pairs is $\sim r_{int}$, since the pairs are independent. Also, since the minimal separation between the components of a merging pair is $\sim r_{int}/\phi_c^{2/3} \sim r_{int}(r_F^2/r_c r_{int})^{2/3} \sim 3 \times 10^{-3} r_{int}$ for the standard numbers above, the number of coherent fringes is potentially very large. Let us consider the case of a herringbone pattern produced

by the interference of three folds. The fringes will be obvious and striking if the peak flux of the interference pattern is twice the mean flux from the rest of the image. Let the mean number of normal subimages be N, which is $\sim (r_{int}/r_c)^2$ according to equation (3.2.18), and let the flux of each of the three folds be F_c, with an average total flux from all subimages normalized to unity. Then, since the amplitudes add coherently in the fringes, we require $3^2 F_c > 2(N-3)/N$, i.e. $F_c \geq 0.2$. Now it follows from (3.3.11), (3.3.9), and (3.3.6) that for a large number of sub-images

$$F_c(X) \sim \left(\frac{r_{int}}{X}\right)^{1/2} \bar{F}_s, \tag{5.1}$$

where X is the distance to the fold line in the observer's plane, and $\bar{F}_s \sim 1/N$ is the mean subimage flux. If the interference pattern is to be frequently visible, X must be no larger than the average distance to the nearest fold line, r_{int}/N [cf. the derivation of equation (4.5)]. Hence $N \lesssim 25$. Putting $X = r_{int}/N$ in (3.3.7), we find that the projected separation between the components of the pair is $\sim r_c/\sqrt{N} \sim r_{int}/N$. So $r_c \sim r_{int}/5 \sim 6 \times 10^{12}$ cm for our standard numbers and the angular separation between bright folds should be of order 25 times the "size" of each fold.

In this picture one expects patterns only for observing frequencies and screen distances such that there are the appropriate number N of subimages within the scattering disk. Moreover, the mean flux of the pulsar should be brighter when drifting patterns are observed, typically 20% above normal for periodic patterns and 40% above normal for herringbone patterns. Finally, although patterns can be observed over typically an octave in observation frequency, the form of any individual pattern should be correlated only over a limited range of t and ν. In fact a position shift of $\sim r_{int}/N$ (the distance to the nearest fold line) should destroy the arrangement of the caustics; this corresponds to $t \sim 1 v_7^{-1}$ d or from the frequency dependence above $\Delta\lambda/\lambda \sim 10^{-2}$. Thus herringbone patterns should rearrange from day to day. We note that Hewish, Wolszczan, and Graham (1985) found that dynamic scintillation spectra not only rearranged on $\sim$day timescales, but that the coherence bandwidth of the periodicities was quite small ($\Delta\lambda/\lambda \lesssim 10^{-3}$ for the case of PSR 1642-03).

Cordes and Wolszczan (1986) have proposed that the subdivision of the image is effected by extra power superposed on a Kolmogorov spectrum at a large scale. The size r_{int} of each subimage is determined by the $\beta = 11/3$ scattering while the extra scale at $\sim 5r_{int}$ has sufficient power to steer images separated by this scale to a common focus. As in the multiple caustic picture, periodic drifts should only occur for a special octave of observing frequency for which there are a few well

separated images. However, here one expects the patterns to be much more persistent in both ν and t. The patterns should be similar over an octave in ν while the variation timescale should be several times the refractive timescale of order 30 days. Observationally, since drifting patterns can change quite rapidly ($\lesssim 1$ day) but then may persist for months, it is not clear whether such slow variation of the rapid caustic behaviour is indicated. The mean flux of the image in the extra scale picture also varies differently than in the multiple caustic scenario. Here all of the flux comes from a few subimages, so when an interference pattern is set up the increase over the mean flux level should be of order unity. In addition, there should be caustics associated with the large extra scale fluctuations, albeit degraded by the Kolmogorov fluctuations. In this scenario, when subimages approach one another the spacing of the periodic interference pattern increases and the subimages brighten as they approach the fuzzy, large scale caustic. In contrast, for the multiple caustic picture, merger events are independent and so fringe spacing and overall flux are not correlated in this fashion. Again, for a complete description of the observed dynamic spectra it may be necessary to invoke the existence of suitably-sized high-density clouds as deterministic scatterers.

6. Conclusions

In summary, if there is an inner scale in the density perturbation spectrum of the ISM comparable to the size of the scattering disk of pulsars and compact extragalactic radio sources, then caustics can have a dramatic effect on the flux modulations and individual folds and cusps may even be resolvable. Since the size of the scattering disk $r_{ref} \gtrsim \lambda^2$, this condition can only be satisfied over a narrow range of frequencies for an individual pulsar. However, despite these restrictive conditions, there are indications from the observations of non-gaussian spikes in the light curves and quasi-periodic patterns in the dynamic scintillation spectra that caustics are, in some cases, important. It is likely that a more complex model than a truncated power-law density fluctuation spectrum in a thin screen will be necessary to reconcile theory and observation. A possibly relevant modification would involve non-power-law density fluctuations; $ie.$ a highly non-gaussian distribution of density perturbations. In particular, if the scattering is relatively localized, a few or even a single large cloud with an enhanced density fluctuation can dominate the scattering. Such enhanced scattering from clumps of scale $\gtrsim r_{int}$ might be found in HII regions, supernova remnants, magnetically confined

plasma clouds, *etc.* Numerical simulation may provide the best means of studying such localised scatterers. Such simulations could probe the effects of relaxing the thin screen assumption and the statistics of higher-order catastrophes, as well. (It can be difficult to realize a *sharp* cut-off in the angular spectrum of density perturbations from a specific physical scale in an extended medium; caustics should be correspondingly weakened.) The computation of higher flux moments and comparison with the behaviour of pulsar flux records would also be useful in determining the importance of higher order caustic events.

From an observational standpoint, it will be important to determine if large intensity spikes and periodic dynamic scintillation spectra occur often over a restricted range of frequencies for each specific pulsar as suggested above, or if such phenomena are less common, but present over a wide range of observation frequency for most pulsars. The latter case would require a more complicated model of the spectrum of ISM density perturbations than a truncated power-law. Further, as described by Cordes *et al.* (1987), dynamic scintillation spectra are important tools for determining when pulsars are subject to multiple imaging. Sources displaying strong, varying periodicities are clearly good candidates for bright caustics. Monitoring of such sources over a range of time lags would allow determination of the variation timescale for the drifting bands and detection of the flux correlation with drift parameters described above. Data from Cordes and Wolszczan (1986) are useful as a first attempt to quantify these phenomena; such data show that in a number of cases dramatic periodicities are present, providing an important arena for testing the caustic scenario.

Acknowledgements

References

Armstrong, J.W. 1984, *Nature*, **307**, 527.

Armstrong, J.W., Cordes, J.M. and Rickett, B.J., 1981, *Nature*, **291**, 561.

Arnol'd, V.I. 1975, *Russ. Math. Survs.*, **30**, 1.

Berry, M.V. 1977, *J. Phys. A: Math. Gen.*, **10**, 2061.

Berry, M.V. and Upstill, C. 1980, *Progress in Optics*, vol 18, ed E. Wolf (Amsterdam: North-Holland), 257.

Blandford, R. and Narayan, R., 1985, *Mon. Not. R. astr. Soc.*, **213**, 591 (BN).

Blandford, R., Narayan, R. and Romani, R.W. 1984, *J. Astrophys. Astron.*, **5**, 369.

Born, M. and Wolf, E. 1980. *Principles of Optics*, (New York:Pergamon Press).

Cordes, J.M., Pidwerbetsky, A., and Lovelace, R.V.E. 1986, *Astrophys. J.*, in press.

Cordes, J.M., and Wolszczan, A. 1986, NAIC preprint 219.

Cole, T.W. , Hesse, H.K. and Page C.G., 1970, *Nature*, **225**, 712.

Coles, W.A., Frehlich, R.G., Rickett, B.J., and Codona, J.L. 1987, *Astrophys. J.*, submitted.

Fiedler, R.L., Dennison, B., Johnston, K.J. and Hewish, A. 1987, *Nature*, submitted.

Frehlich, R.G. 1986, preprint.

Gapper, G.R. and Hewish, A., 1981, *Mon. Not. R. ast. Soc.*, **197**, 209.

Goodman, J. and Narayan, R., 1985, *Mon. Not. R. astr. Soc.*, **214**, 519 (GN).

Hecht, E. and Zajac, A. 1974, *Optics*, (Reading, Mass.:Addison-Wesley).

Helfand, D.J, Fowler, L.A. and Kuhlman, J.V., 1977, *Astron. J.*, **82**, 701.

Hewish, A., 1980, *Mon. Not. R. astr. Soc.*, **192**, 799.

Hewish, A., Wolszczan, A. and Graham, D.A., 1985, *Mon. Not. R. astr. Soc.*, **213**, 167.

Lovelace, R.V.E., 1970, Ph.D. Thesis, Cornell University.

Poston, T. and Stewart, I.N., 1987,*Catastrophe Theory and its Applications* (London: Pitman).

Rickett, B.J., 1977, *Ann. Rev. Astron. Astrophys.*, **15**, 479.

Rickett, B.J., Coles, W.A. and Bourgois, G., 1984, *Astron. Astrophys.*, **134**, 390.

Roberts, J.A. and Ables, J.G., 1982, *Mon. Not. R. astr. Soc.*, **201**, 1119.

Romani, R.W., Narayan, R. and Blandford, R. 1986, *Mon. Not. R. astr. Soc.*, **220**, 19 (RNB).

Romani, R.W., Blandford, R. and Cordes, J. 1987, *Nature*, submitted.

Rumsey, V.H., 1975, *Radio Sci.*, **10**, 107.

Salpeter, E.E., 1967, *Astrophys. J.*, **147**, 433.

Shapirovskaya, N. Ya., 1978, *Soviet Astron.*, **22**, 544.

Shishov, V.I., 1974, *Sov. Astron.*, **17**, 598.

Sieber, W., 1982, *Astr. Astrophys.*, **113**, 311.

Tatarskii, V.I. and Zavorotnyi, V.U. 1980, *Progress in Optics* **XVIII**, 204.

Thöm, R., 1975, *Structural Stability and Morphogenesis* (Reading, Mass.: Benjamin).

Appendix A

We consider a phase structure function of the form (2.4.5), $D(r) \propto [|r^2 - r_c^2|^{\alpha/2} - r_c^{\alpha/2}]$, $\alpha = \beta - 2$. The limiting forms of this expression are

$$
\begin{aligned}
D(r) &\propto |r^2|^{\alpha/2} \qquad r \gg r_c \\
D(r) &\propto r_o^{\alpha-2} \qquad r \ll r_c
\end{aligned}
\tag{A.1}
$$

as appropriate for a structure function $\langle [\phi(r) - \phi(0)]^2 \rangle$ calculated from (2.3.2) and a fluctuation spectrum with a gaussian cutoff $Q_0 q^{-\beta} e^{-(qs_o)^2}$, $\beta < 4$. This choice of $D(r)$ yields

$$
\tilde{W}(\vec{q}) = \int d^2\vec{r} e^{i\vec{q}\cdot\vec{r}} \exp\Big\{ A r_F^\alpha \Big[2|(qr_F^2)^2 + r_c^2|^{\alpha/2} + 2|r^2 + r_c^2|^{\alpha/2}
$$

$$
-|(r + qr_F^2)^2 + r_c^2|^{\alpha/2} - |(r - qr_F^2)^2 + r_c^2|^{\alpha/2} - 2r_c^\alpha \Big] \Big\}
\tag{A.2}
$$

where A is (see GN)

$$
A = \frac{2^{2-\beta}}{\pi} \frac{\Gamma((6-\beta)/2)}{(\beta-4)(\beta-2)\Gamma(\beta/2)} \Big[Q_0 r_F^{\beta-2} \Big]
\tag{A.3}
$$

Note that, for $2 < \beta < 4, A < 0$ and that the quantity in brackets is essentially the amplitude of the phase fluctuations on the Fresnel length. As described in section 2, when $r_c > r_{dif}$, A must be replaced by $A(r_F/r_c)^{\beta-4}$ for $\beta < 4$ to maintain the strength of the scattering.

For strong scattering and fair separation of the various length scales, we have three conditions

$$
|A| \gg 1 \qquad \text{i.e. } \phi \gg 1
\tag{A.4}
$$

$$
|A|(r_c/r_F)^\alpha \gg 1 \qquad \text{i.e. } r_c \gg r_{dif}
\tag{A.5}
$$

$$
|A|(r_c/r_F)^{\alpha-4} \gg 1 \qquad \text{i.e. } r_c \ll r_{ref}
\tag{A.6}
$$

Under these assumptions, we wish to investigate the form of $\tilde{W}$ (A.2) in the regime $q_{int} \ll q \ll q_{cdif} = r_c/r_F^2$. The integral will be dominated by $|r| \ll r_c$ so we can expand the second exponential to lowest order in r, giving

$$
\tilde{W}(q) = \int d^2r e^{iq\cdot r} \exp[-\alpha(2-\alpha)|A|(r_c/r_F)^{\alpha-4}q^2 r^2]
\tag{A.7}
$$

The exponent in brackets will cutoff the integral when $r \gtrsim (|A|(r_c/r_F)^{\alpha-4})^{-1/2}q^{-1}$, which, by (A.6), gives $r \ll q^{-1}$. Accordingly the first exponential can be replaced by one and the gaussian integration carried out, yielding

$$
\tilde{W}(q) = \frac{\pi}{2^{1/2}\alpha(2-\alpha)|A|(r_c/r_F)^{\alpha-4}} q^{-2}
\tag{A.8}
$$

which gives the constant power at intermediate scales in Figures 2 and 3.

Similar analyses can be conducted for the other regimes, giving

$$r_F^{-2} \tilde{W}(q) =$$

$$
\begin{cases}
2^\alpha \pi \alpha (2-\alpha) \frac{\Gamma((\alpha+2)/2)}{\Gamma((4-\alpha)/2)} |A| (q r_F)^{2-\alpha} \exp[-(q/q_{ref})^2/2] & q < q_{int} \\[2mm]
\frac{\pi}{\alpha |A| (r_c/r_F)^{\alpha-2}} \exp[-(q/q_{dif})^2/2] & q \gg q_{cdif}
\end{cases}
\qquad (A.9)
$$

where $q_{dif} \equiv q_{ref}^{-1} r_F^{-2} \equiv (2\alpha |A| (r_c/r_F)^{\alpha-2})^{1/2}$. The scalings for the extended power law spectra (solid lines in Figures 2 and 3) have been given by GN.

Appendix B

In this appendix we show how to relate the specific expression describing the local phase variation from a thin screen observed at position (X, Y) in the vicinity of a cusp to the standard polynomial equation (3.1.3). Again, we must enlarge the standard form to admit two state variables by writing

$$\bar{\zeta} = \frac{1}{4} s^4 + \frac{1}{2} C_2 s^2 + C_1 s + \frac{1}{2} t^2. \qquad (B.1)$$

We follow the same procedure that we described in §3.3 for the fold. We allow ourselves to translate and rotate the (x, y) and (X, Y) coordinate systems and only retain the terms that can be dominant in the vicinity of the stationary points. We obtain

$$\zeta(x, y; \alpha, \beta) = \frac{1}{4} a x^4 + \frac{1}{2} b x^2 y + \frac{1}{2} c y^2 - \alpha x - \beta y ; \quad a > 0 \qquad (B.2)$$

where $a = \phi_{xxxx}/6$, $b = \phi_{xxy}$, $c = \phi_{yy}$. $\alpha = kX/z$, $\beta = kY/z$ are control variables. Again, this expansion is only valid for $|x|, |y| \ll r_c, r_{int}$. The stationary points of equation (B.2) are given by the real roots of the equation

$$\left(a - \frac{b^2}{2c} \right) x_i^3 + \frac{b\beta}{c} x_i - \alpha = 0 \qquad (B.3)$$

together with the associated y coordinate.

$$y_i = -\frac{(b x_i^2/2 - \beta)}{c}. \qquad (B.4)$$

There are three roots corresponding to three rays when

$$f(\alpha, \beta) = \left(\frac{\alpha}{2} \right)^2 \left(a - \frac{b^2}{2c} \right) + \left(\frac{b\beta}{3c} \right)^3 < 0, \qquad (B.5)$$

and one ray when the inequality is reversed. $f(\alpha, \beta) = 0$ is the equation of the caustic in the observer plane. The image of the caustic on the screen, i.e. the location of the coalescing images, is the parabola

$$H(x, y) = bcy + (3ac - b^2)x^2 = 0. \qquad (B.6)$$

$H(x, y)$, the Hessian, is the reciprocal of the sub-image amplification.

The procedure that we follow in relating the specific expression for the phase variation, equation (B.2), to the standard form, equation (B.1), is to assume that there exists the phase relation $\zeta = \bar{\zeta} + g(a, \beta)$ where $g(\alpha, \beta)$ is a function to be determined. The three functions $C_1(a, b, c; \alpha, \beta)$, $C_2(a, b, c, \alpha, \beta)$, $g(\alpha, \beta)$ can be determined by comparing the phases at the three stationary points (a one to one mapping is guaranteed) and solve for $s(x, y)$, $t(x, y)$. We obtain

$$C_1 = \alpha(a - b^2/2c)^{-1/4}$$

$$C_2 = \beta bc^{-1}(a - b^2/2c)^{-1/2}$$

$$g = -\beta^2/2c$$

$$s = -(a - b^2/2c)^{1/4}x$$

$$t = c^{1/2}(y - \beta/c + bx^2/2c). \qquad (B.7)$$

Equations (B.7) allow us to map a cusp created by a thin phase screen onto the standard polynomial equation. They therefore can be used to compute the diffraction pattern from an individual cusp, given the low order derivatives of the phase ϕ at the cusp.

The intensity behaviour and fringe scales can be estimated for a cusp, following section (3.3), although the complete description of the fringes given by the Pearcey function (*eg.* Berry and Upstill 1980) is rather complex. At the cusp, $x = y = 0$, we have from (B.2) the area of the first coherent zone is $\sim (a^{-1/4}) \times (c^{-1/2}) \sim r_c^2/\phi_c^{3/4}$ and so, in analogy to (3.3.6) the maximum flux is

$$F_{c,max} \sim \frac{r_c^4}{\phi_c^{3/2}r_F^4} \sim \phi_c^{1/2}\bar{F}_s \sim \phi_c^{1/2}/N. \qquad (B.8)$$

From (B.3), the sub-image separation for an observer at $Y = 0, X \neq 0$ is $\delta x \approx \pm(X/r_{int})^{1/3}r_c$. Thus the x-width of the central fringe of a cusp caustic, given by the distance over which the phase varies by order unity, is

$$X_c \sim r_{int}/\phi_c^{3/4}. \qquad (B.9)$$

From (3.2.2), (3.2.3) and (B.2) we can estimate the variation in the *mean* flux with distance X from the cusp point as

$$F_c(X) \sim \left(\frac{\partial^2 \zeta}{\partial x^2}\right)^{-1} \left(\frac{\partial^2 \zeta}{\partial y^2}\right)^{-1} r_F^{-4} \sim \left(\frac{\phi_c(X/r_{int})^{2/3}}{r_c{}^2}\right)^{-1} \left(\frac{\phi_c}{r_c{}^2}\right)^{-1} r_F^{-4} \sim \left(\frac{X_f}{X}\right)^{2/3} F_{c,max}.$$

$$(B.10)$$

An approximate scale for the decrease in fringe spacing within this envelope for observer shifts in the X direction, following (3.3.12) is

$$S_c(X) \sim (d\Delta\zeta/dX)^{-1} \sim \left(\frac{X_c}{X}\right)^{1/3} X_c.$$

$$(B.11)$$

There should be no coherent fringes in the Y direction beyond the central maximum of width $\sim r_F{}^2/l_y$ (3.2.2).

Table Caption

Table of important lengthscales for refractive scintillation from a density perturbation power spectrum with an inner scale. Scale are introduced and/or defined by the indicated equation in the text and are listed in order of size (for $r_F < r_c < r_{int}$).

Table of Length scales

Name	Symbol	Definition	Equation
diffractive	r_{dif}	$= r_F^2/r_{ref} \sim 1/k\theta_{rms}$	(2.4.7)
cut-off diffractive	r_{cdif}	$= 2\pi/q_{cdif} = r_F^2/r_c$	(2.5.2)
Fresnel	r_F	$= (z/k)^{1/2}$	(2.2.2)
inner or cut-off	r_c	$= 2\pi/q_c$	(2.3.3)
intermediate	r_{int}	$= \begin{cases} r_{ref}, & \beta < 4 \\ r_{ref}(r_c/r_{ref})^{(\beta-4)/2}, & \beta > 4 \end{cases}$	(2.5.3),(2.5.4)
refractive	r_{ref}	$= z\theta_{rms}$	(2.4.4)
outer	r_{out}	$= 2\pi/q_{out}$	(2.3.3)

Figure Captions

Fig. 1

Scattering geometry. The scattering medium is idealized as a thin screen at a distance z_{os}, with the source at distance z_{ps} beyond the screen. The transverse coordinates on the observer plane and scattering screen are denoted (X, Y) and (x, y), respectively, with the coordinate origin taken to be the screen intersection with the optical axis PSO.

Fig. 2

Flux variation power spectrum for $\beta = 11/3$ and $r_c = 0.0$ (solid line), $r_c = 0.01 r_{ref}$ (short dashes) and $r_c = 0.1 r_{ref}$ (long dashes). Spatial frequency is in units of the Fresnel scale $\eta = q r_F$. Note that, for $r_c \neq 0$, to maintain the separation of η_{ref} and η_{dif}, the strength of the scattering $\propto |A|$ has been renormalized and this has enhanced the refractive (low q) peak (see §2.5).

Fig. 3

Flux variation power spectrum for $\beta = 4.3$ (as in Figure 2). Note that, even for $r_c = 0$, the refractive fluctuations are markedly stronger than in the $\beta < 4$ case. Since for $\beta > 4$ large scales dominate the scattering, it has not been necessary to rescale $|A|$ for $r_c \neq 0$.

Fig. 4

a) Fold and cusp catastrophes exhibited in one dimension as a function of the state variable s. The vertical coordinate denotes excess phase ζ and sub-images are found at the extrema of ζ. The width of an extremum at a depth ~ 1 describes the sub-image patch size and, hence, intensity. As the control parameter X is varied approaching the singularity, the extrema (sub-images) broaden and merge. Crossing the singularity, the three cusp images merge, leaving a single extremum; there are no extrema on the dark side of a fold.

b) Loci of the fold and cusp caustic on the observer plane. For the cusp, the filled circle represents an observer near the cusp point, while the open circles represent observers approaching fold caustics from the bright side.

c) Loci of the fold and cusp caustics on the scattering screen. The open circles in b) form two images each, elongated and hence magnified. At the cusp point three images merge tangentially to the line of the fold asymptotes. For a source inside the cusp there will (globally) three images on the scattering screen; a source encountering the cusp from the outside will produce a single bright image.

Fig. 5

Catastrophes of codimension 3.

a) Swallowtail exhibited in one dimension. There are four turning points corresponding to four rays which can be made to merge simultaneously.

b) When there are two state variables (*ie.* coordinates x, y on the screen), we can plot contours of constant phase ζ to exhibit, in this example of a swallowtail, two maxima (H) and two saddles (S). Note that only one state variable (which can be chosen to advance along the dashed line) is necessary to exhibit the catastrophe.

c) Contours of constant phase $\zeta(x, y)$ near a hyperbolic umbilic catastrophe. Just as was the case with the swallowtail, there are a four stationary points (two saddles S, a maximum H and a minimum L). However this catastrophe is distinct from the swallowtail because two state variables must be used to describe it.

d) Contours of constant phase $\zeta(x, y)$ near an elliptic umbilic catastrophe. This nesting of contours cannot be continuously deformed into either the swallowtail or the hyperbolic umbilic and therefore represents a topologically distinct catastrophe.

Fig. 6

Caustic sheets. Beyond a focal length from the scattering screen rays can merge and caustics will form. The surface on which two rays coalesce describes the fold sheets, whose intersection with the observer plane forms fold loops. The lines on these sheets where three rays merge intersect the plane in cusp points.

Fig. 7

For a given over-brightening α, strong caustics can be found for only a restricted range of cut-off scale and observation wavelength. In particular, for any r_c, λ must be greater than λ_{min}, shown by the line $r_c < r_{int}$ (see §4). The maximum wavelengths are denoted by the lines "fold" and "cusp"; together these lines give the minimum wavelength λ^* and cut-off scale r_c^*, here for $\alpha = 1$. The top and right hand axes give numerical values for r_c in cm, λ_m in m, and $C_{-4} = D_{kpc} = \alpha = 1$.

Fig. 8

Light curve of a point source near fold and cusp catastrophes as a function of the single control parameter, t. The solid line depicts the geometrical optics variation of the mean flux: $|t|^{-1/2}$ for the fold, $|t|^{-1/3}$ for the cusp. The dotted lines depict the fringe behaviour for finite wavelength

observations; fringes are found on the bright side of the fold while the flux drops exponentially to zero after the two sub-images annihilate. For the cusp, fringes can only be seen by observers within the apex of the cusp. In the specific example shown, as the observer passes outside the caustic near the cusp point, two sub-images annihilate and the remaining sub-image fades smoothly, generating no fringes.

Fig. 9

Crossed patterns in dynamic scintillation spectra generated by interfering sub-images. Here we show schematically in the ν, t plane maxima of flux lying on striated bands caused by the interference of two bright sub-images, which drift in ν due to the varying refraction. Nearly orthogonal, finer scale fringes are superposed due to the regular interference pattern of a more distant third sub-image.

a)Phase   b)Observer Plane   c)Scattering Screen

Fold $\Delta\zeta\sim1$

Cusp $\Delta\zeta\sim1$

Scattering Screen

Observer Plane

z

r_{int}

x

y

X

Y

fold

cusp

t
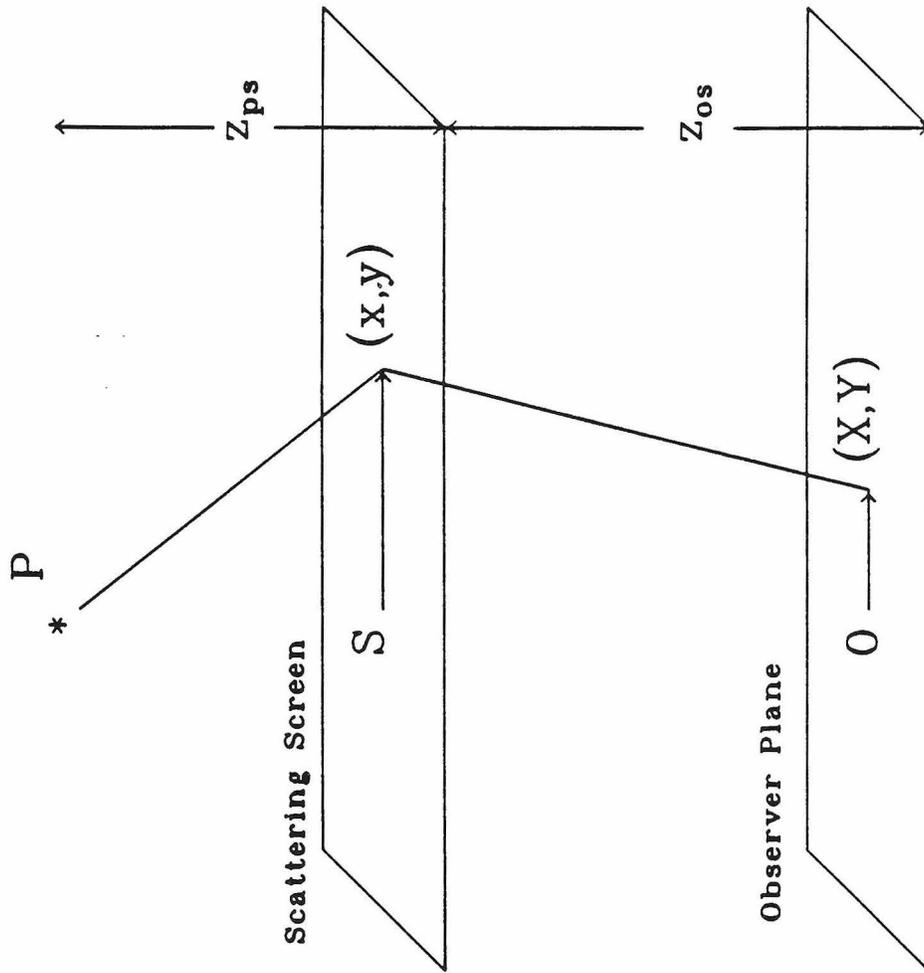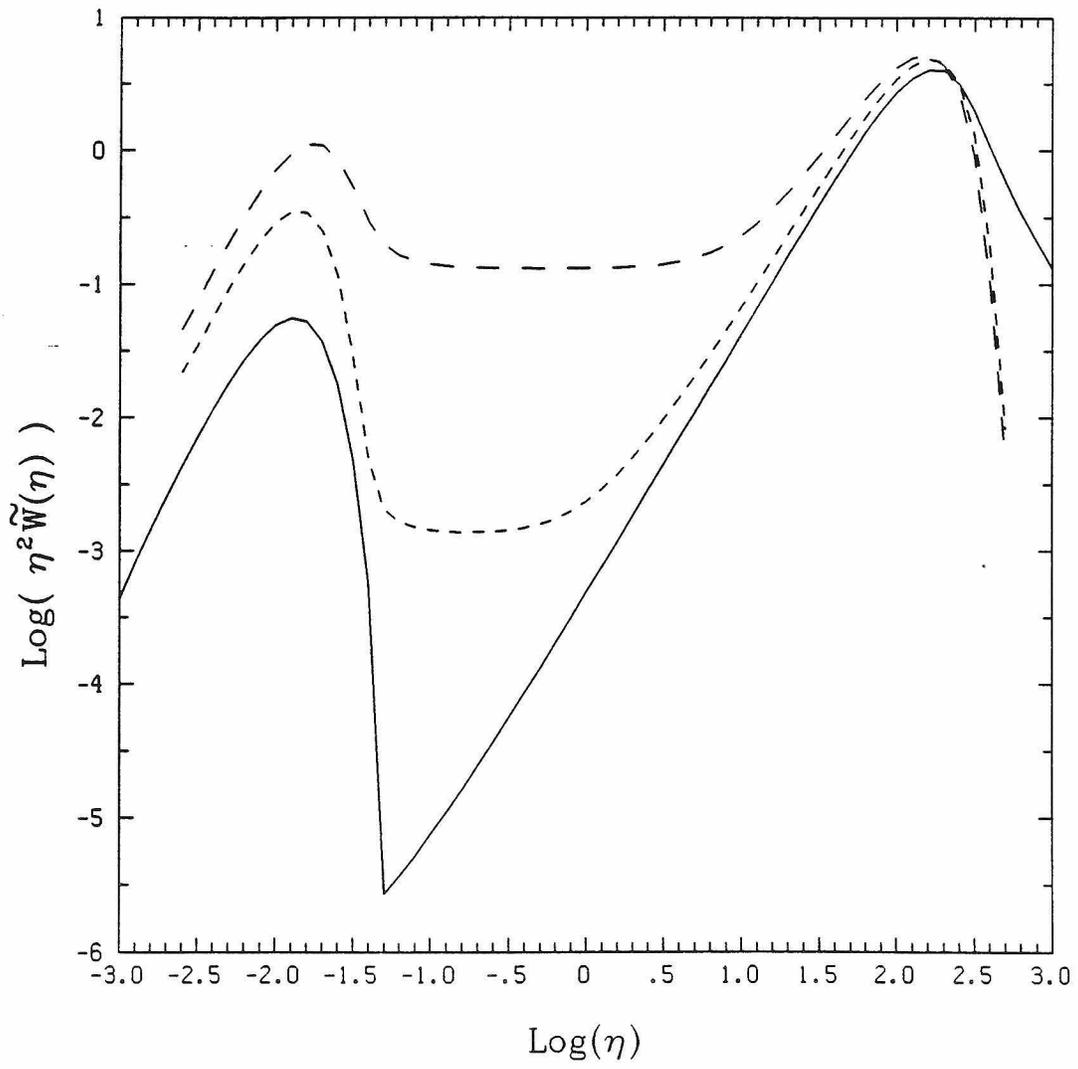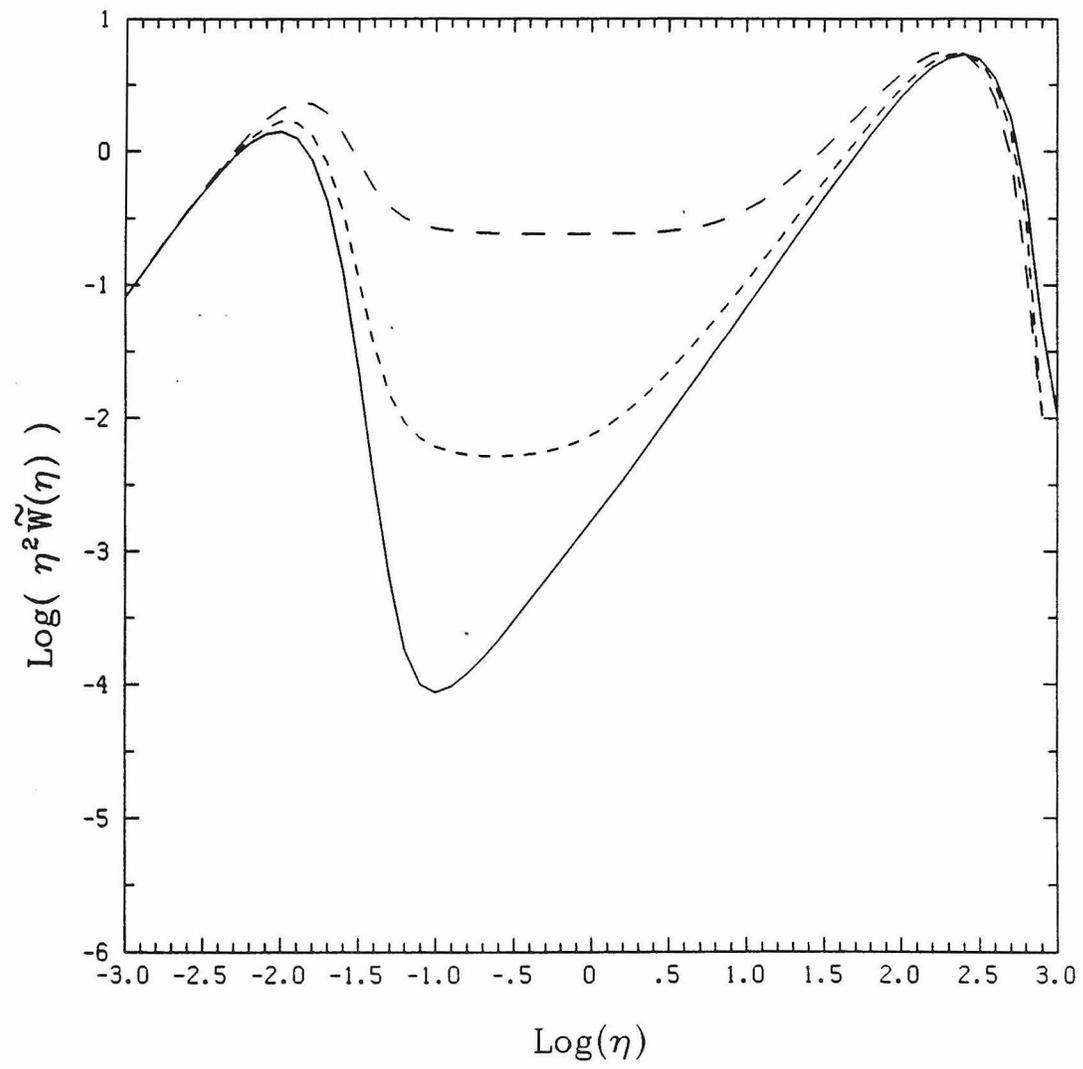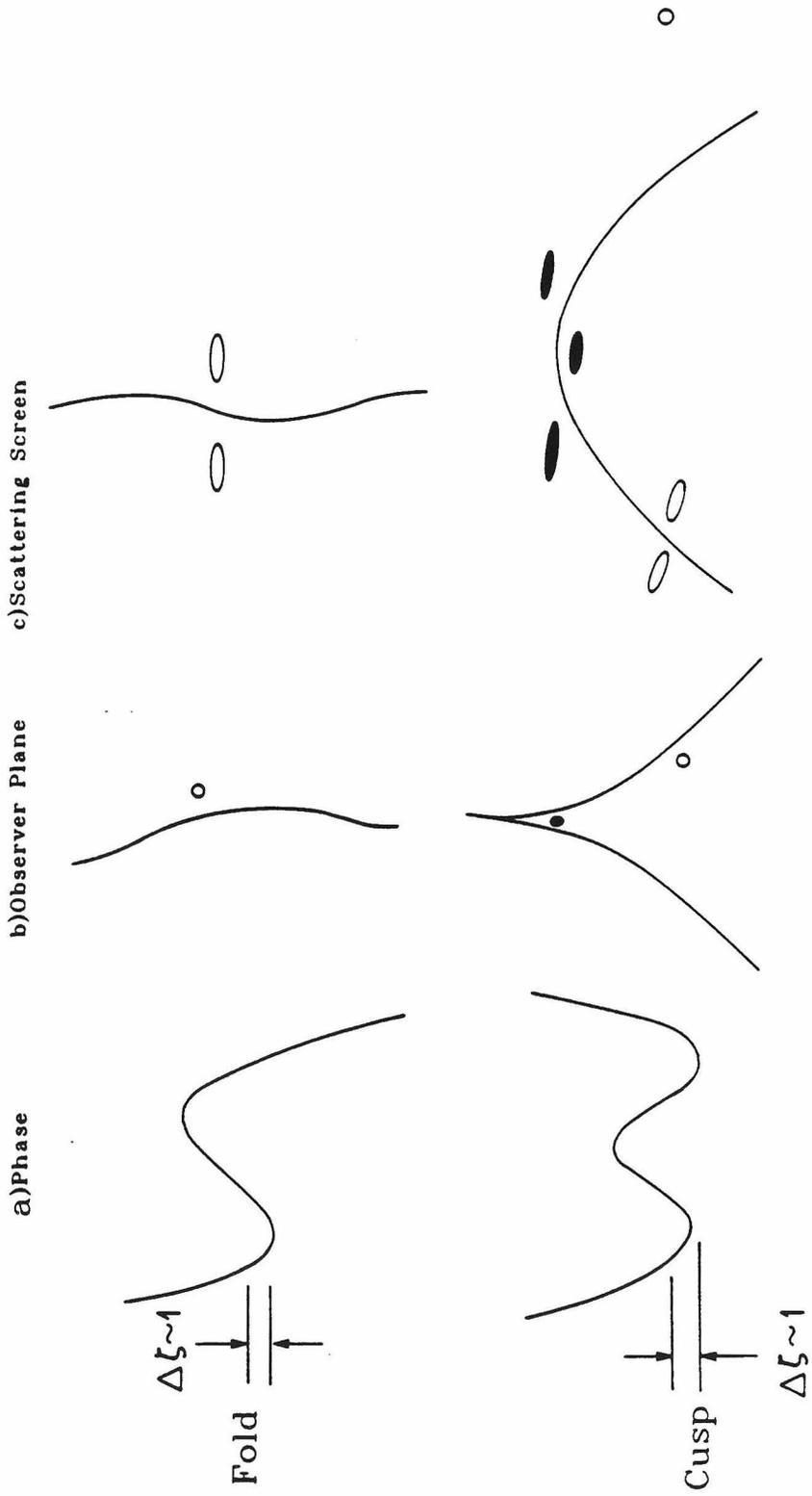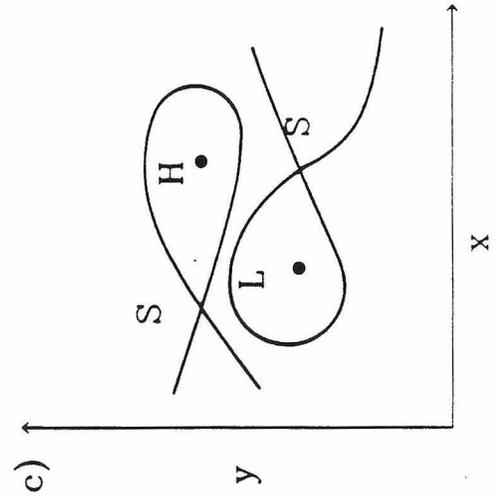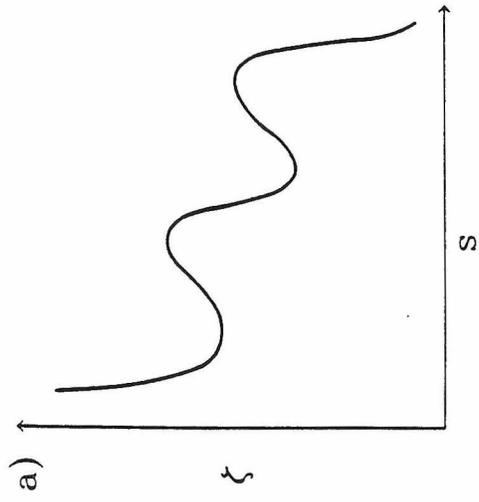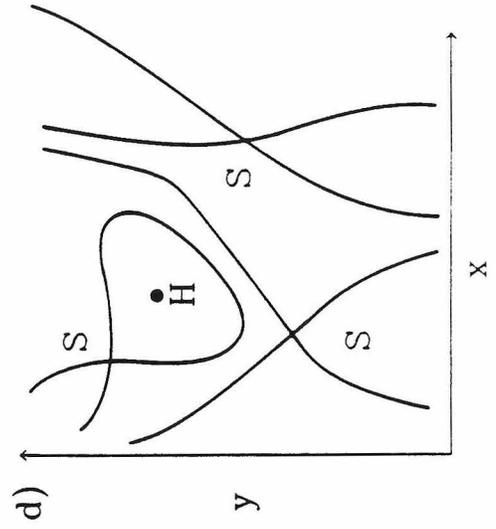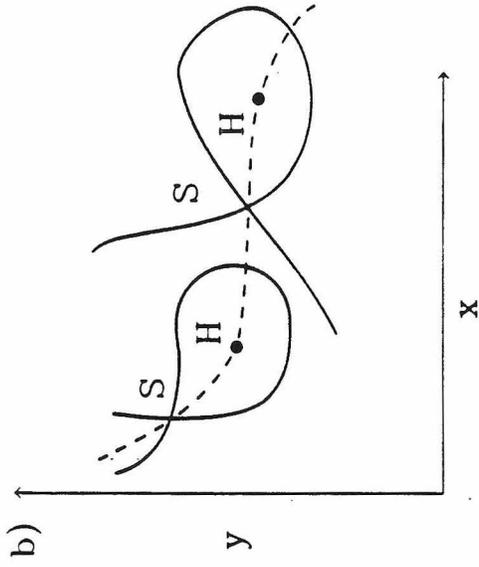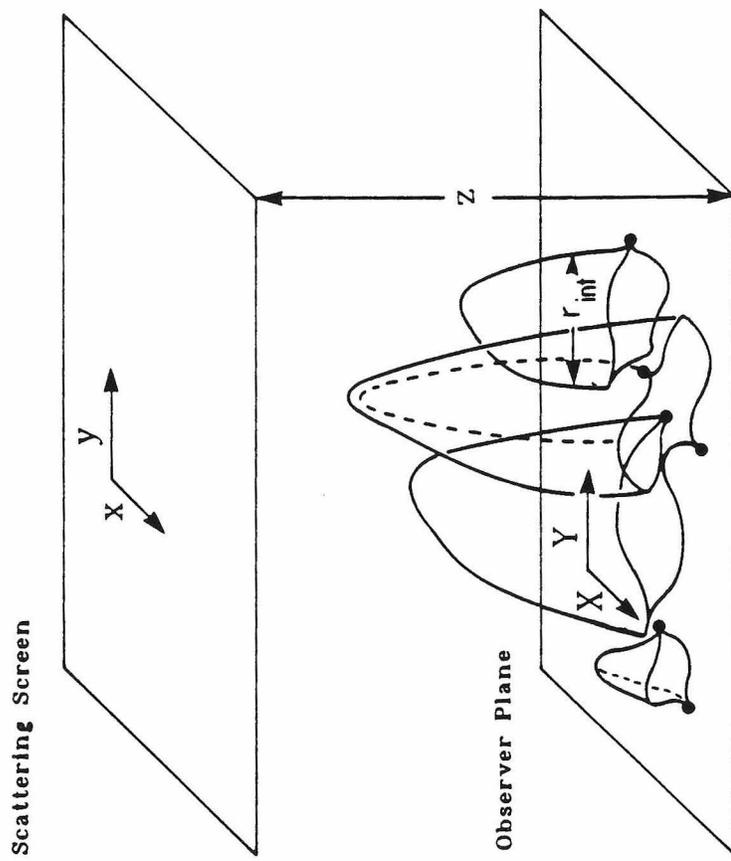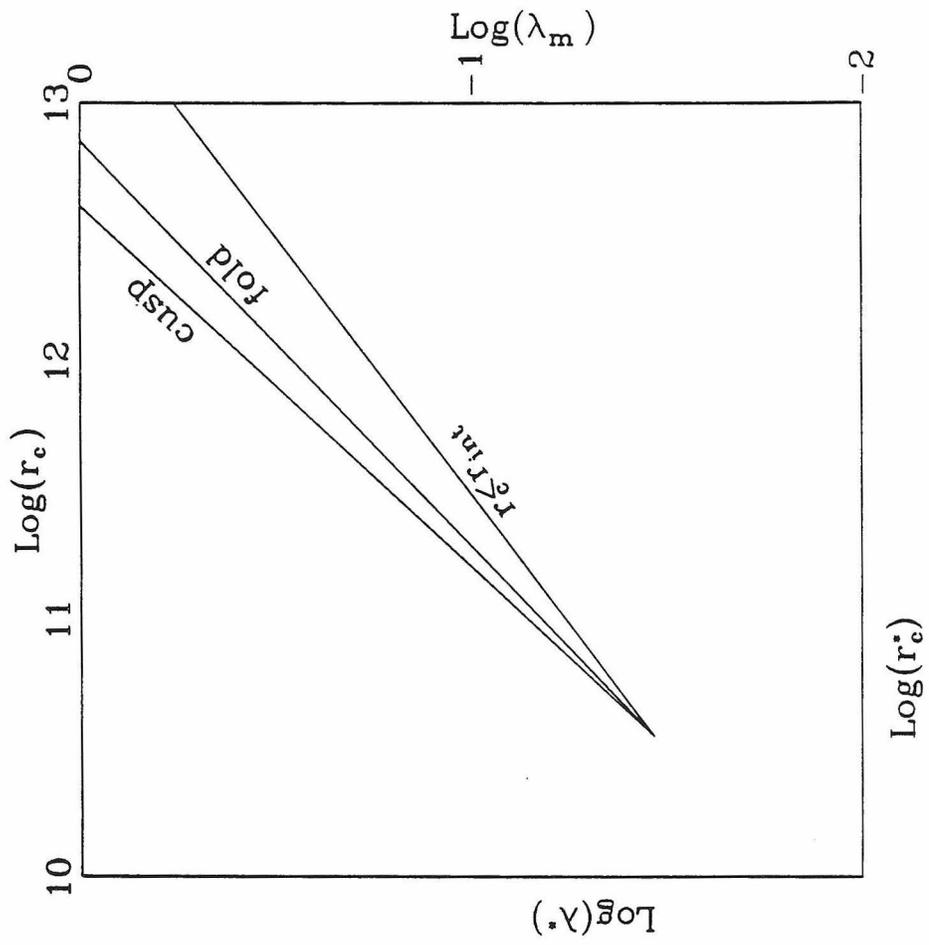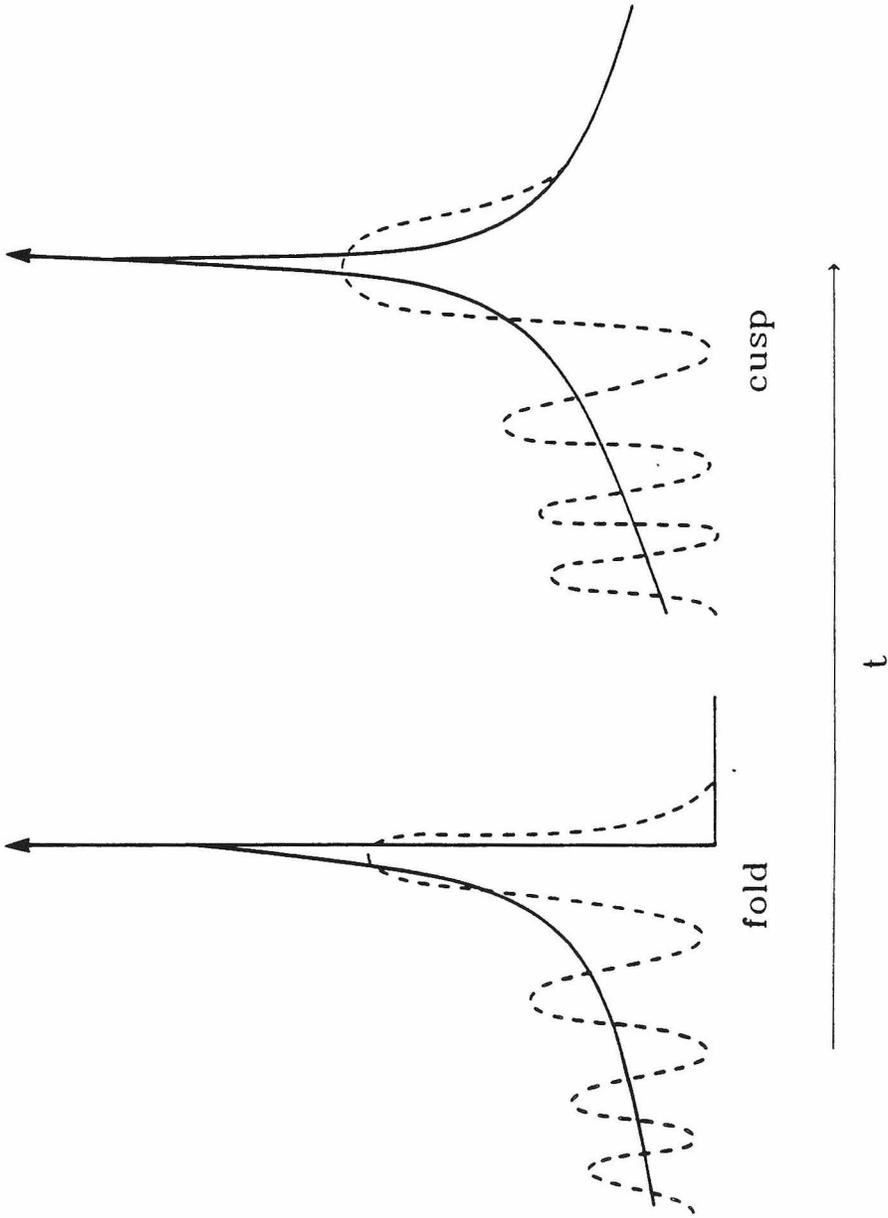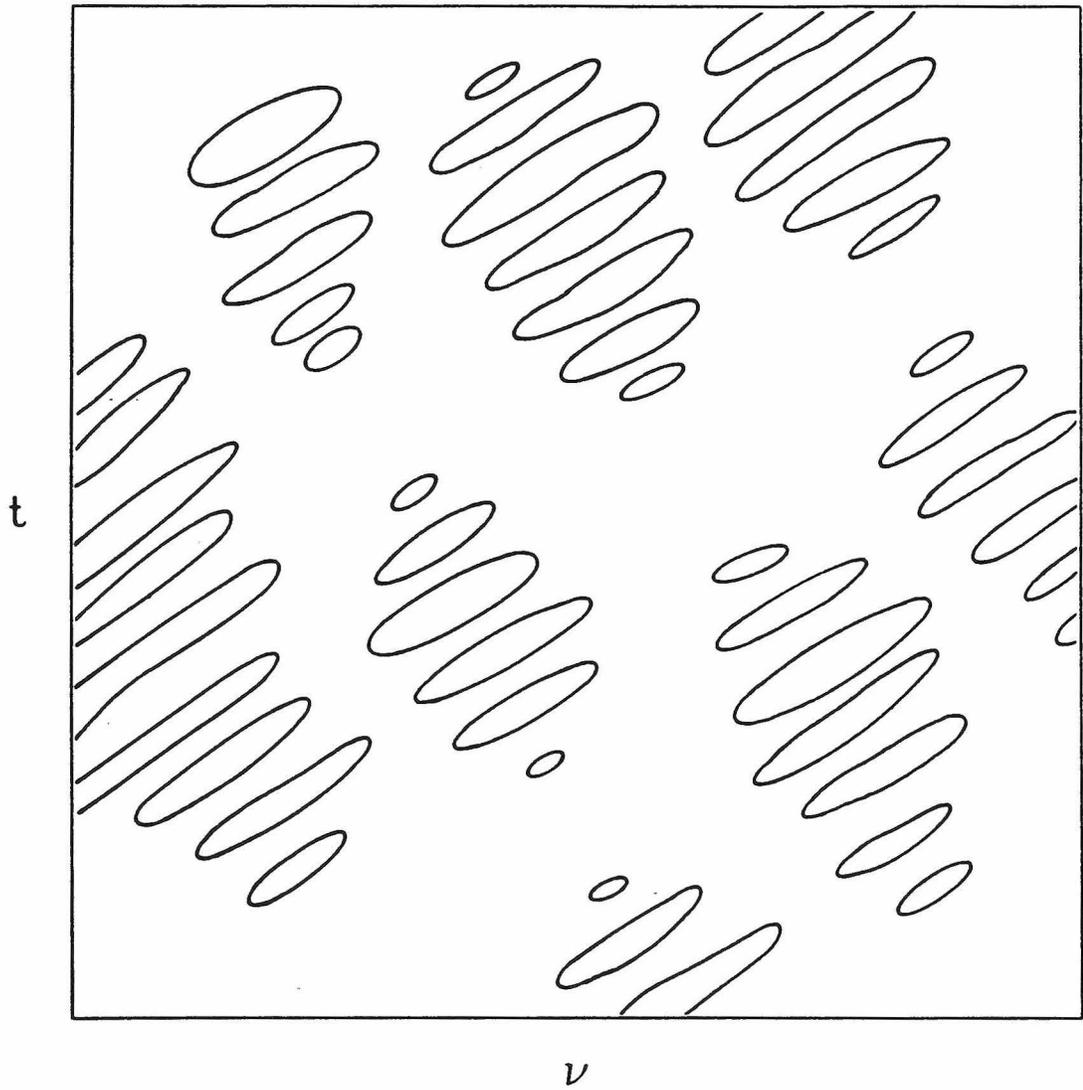
Chapter 8:

RADIO CAUSTICS FROM LOCALIZED ISM PLASMA STRUCTURES

Roger W. Romani, Roger D. Blandford,

Theoretical Astrophysics, California Institute of Technology

Pasadena, CA 91124

and *James M. Cordes*

Astronomy Department and NAIC, Cornell University

Ithaca, NY 14853

Summary. In a study of 36 extragalactic radio sources observed over 7 years with the Green Bank Interferometer at 2.7GHz and 8.1GHz, Fiedler, *et al.* (1987) have detected unusual variations in the light curves of several sources. The most dramatic event, found in the flux record of the quasar 0954+658, showed large modulations at both frequencies. At 8.1GHz, there are four spikes over a period $\sim$ 80 days in each of which the intensity increases roughly threefold. At 2.7 GHz, two broad maxima of width $\sim$30 days bracket the high frequency event, with a total duration of $\sim$120 days. At least two other sources show similar 2.7GHz variations. Fiedler, *et al.* argue cogently that these variations are unlikely to be intrinsic to the sources and that refractive scintillation in the interstellar medium is probably responsible. They postulate the existence of large scale $(10^{13}-10^{14}\text{cm})$ localized plasma density inhomogeneities in the interstellar medium (ISM) to explain these observations. Independent evidence for strong local fluctuations in the ISM electron density on these scales comes from the unusually large amplitude of low frequency quasar variation (Romani, Narayan and Blandford 1986), the presence of periodicities in pulsar dynamic scintillation spectra (Cordes and Wolszczan 1986) and non-Gaussian spikes in pulsar light curves (Goodman, *et al.* 1987). All three phenomena have been interpreted as multiple imaging and focussing by large-scale refractors. In this letter we present a more detailed interpretation of the event in 0954+658, propose possible sites for the refracting clouds and suggest some future observations.

The 0954+658 event can be understood in terms of geometrical optics as the signature of a strong focussing event or *caustic*. We picture the line of sight as passing through a localized plasma structure having a simple Gaussian electron density profile along the path of relative motion. Such a structure is a diverging radio-wave lens, causing ray crossing in the far field (analogous to the case of converging gravitational lenses postulated to explain multiple quasar imaging). The locus of points for which rays cross (in the two dimensions time, t and lens distance, D) forms four fold caustics which meet pair-wise in two cusp points at a distance D_c (Fig. 1). A point source will be infinitely magnified in the geometrical optics limit for an observer located on these caustics. As the line of sight crosses each fold line there is an image pair creation/merger event and the total received flux passes through a sharp maximum. We interpret the four spikes seen at 8.1GHz as fold crossings by an observer at a distance $D = \alpha D_c$, $\alpha \gtrsim 3$. The characteristic $\sim$10d rise times of these spikes are due to the convolution of the narrow caustic with a finite intrinsic source size ($\theta_{int} \sim$0.4 mas at 5GHz, Pearson and Readhead 1981). The refraction angle, and hence α, vary with frequency $\propto \nu^{-2}$ so that at 2.7GHz the caustics separate further. The intrinsic source size should, however, also increase. (Models in which the source size is dictated by synchrotron self absorption give $\theta_{int} \lesssim \nu^{-1}$.) Convolution with this size at 2.7GHz smooths over the fold pairs associated with each cusp and

creates two broad ($\tau \sim 30$ day) maxima, as observed. Numerical computations with similar lens density profiles show that the general form of the light curves is quite robust for a simple clump, and that perturbations from the exact Gaussian form can reproduce the observed asymmetry of the 8.1GHz peaks. For the two other observed events, we presume that α is too small and/or θ_{int} is too large to resolve the folds at 8.1GHz.

The plasma lenses should lie at typical distances of $D \sim 1$kpc and the combined motion of the earth and the ISM should give a relative velocity of $100v_7$km/s. If at an observation frequency ν, the event lasts $100\tau_2$ days and the spacing of the peaks implies $D = \alpha D_c$, then the inferred refraction angle is $\theta_{ref} \sim 6\alpha v_7\tau_2/D_{kpc}$mas, arising from a plasma structure of size $a \sim 10^{14}v_7\tau_2$cm. The associated electron column density variation is $\delta N_e \sim 2\pi\theta_{ref}a\nu^2/(r_ec^2)$. For the dramatic event in 0954+658 we infer $\alpha \gtrsim 3$ at 8.1GHz and the required column density is $\delta N_e \sim 10^{19}$cm^{-2}, while for a more typical event $\alpha \sim 1$ at 2.7GHz and $\delta N_e \sim 5 \times 10^{17}$cm^{-2}. For a spherical lens the electron density is $n \sim \delta N_e/a$; less extreme densities are required if the plasma structures are elongated along the line of sight. If the elongation factor is $\eta \sim 100\eta_2$, then for a typical event the required plasma density is $n \sim 50\eta_2^{-1}\alpha(v_7\tau_2)^2/D_{kpc}$cm^{-3} while for 0954+658, $n \sim 10^3\eta_2^{-1}\alpha(v_7\tau_2)^2/D_{kpc}$cm^{-3}. As described by Fiedler, *et al.* the duty cycle for these events in their data set is $f \sim 0.005$. If a caustic event is produced by a plasma filament with an aspect ratio η, aligned along the line of sight, then the cross section for such end-on events is η^{-3} that for filament intersections at average angles of incidence. By contrast, if the structures are locally planar, caustic events should occur η^{-2} times as often as face-on intersections. We consider the latter geometry as more probable, since it gives the greatest caustic event duty cycle for a given η and space density of the plasma lenses.

Requiring that such single high-density features occasionally dominate the distributed scattering along the line of sight contrasts markedly with the conventional description of scintillation. However, it is in accord with models of the ISM incorporating high-density, cold clouds embedded in a tenuous inter-cloud medium heated by supernovae (McKee and Ostriker 1977), and we regard ionization fronts and/or cooling instabilities in old supernova remnants as reasonable sites for the lensing plasma sheets (Hester, Parker and Dufour 1983; Hester 1987). This interpretation becomes more plausible if the remnant scale height is ~ 1kpc, somewhat larger than that usually inferred from observations (Caswell and Lerche 1979). Using the fiducial parameters for a typical plasma lens given above, the pressure in the refracting structures will be $\sim 7 \times 10^{-11}$dyne cm^{-2}, for T$\sim 10^4$K; such pressures can be found in remnants of radius $R \sim 35$ pc. As these remnants can have a covering

fraction $C \sim 0.2$ at moderate galactic latitude, the probability that we are looking through an edge-on sheet within an old remnant should be $\chi \sim f/c \sim 0.02$, so there are $\sim \chi\eta(R^2/\eta a^2) \sim 10^{11}$ sheet facets per remnant, containing $\lesssim 0.01$ of the remnant mass.

An average line of sight through an old remnant will pass through $\sim \chi\eta^2 \sim 100$ sheets face-on. If these are uniform slabs, the associated scattering angle at $\sim$meter wavelengths will be very small. More realistically, however, the sheets will have some transverse structure; for example, the density might have a corrugated modulation of width $\sim a$ (which could be organized by magnetic fields). With our fiducial values, the incidence of remnant-crossing lines of sight is then consistent with the oscillations seen roughly 10% of the time in pulsar dynamic scintillation spectra. These oscillations are interpreted as the beating between two or more subimages of the pulsar (Cordes and Wolszczan 1986; Roberts and Ables 1982; Hweish, Wolszczan and Graham 1985); the subimage splitting angles seen are a few m.a.s. at 0.4 GHz, $ie.$ too large to be caused by the general scattering medium. For an average remnant-crossing line of sight, however, the scattering of a fraction $\sim 1/\eta$ of the corrugated face-on sheets adds incoherently in a given transverse dimension, giving a mean scattering angle $\sim (\eta\chi)^{1/2}(\theta_{ref}/\eta)$. With observed angles for our typical caustic event and fiducial parameters, this gives refraction angles of 1-10mas at $\lesssim$GHz observing frequencies, appropriate to the periodic frequency drifts and strong pulsar focussing mentioned above. We thus have a unified picture of thin sheet-like lensing structures causing the dramatic caustic events at $\sim$cm wavelengths when seen edge-on, and the more common sub-GHz pulsar modulations when they intersect the line of sight at average inclinations.

As an alternative model for the production of the caustics, we can suppose that the lenses are confined magnetically, most plausibly in twisted magnetic ropes or filaments of diameter $\sim 10^{14}$cm. At this scale, a helical field of $\sim 5 \times 10^{-4}$G allows confinement of densities $\sim 10^4$cm^{-3} where interstellar UV ionization and heating can balance radiative cooling (Reynolds 1984). Here $\eta \sim 1$ and the required density of $\sim$pc-length flux tubes is ~ 1pc^{-2}. These ropes could be spun off by solar-type stars, reconnecting and combining as they cross one another. Filamentary structure, albeit on larger scales is observed in the galactic center (Yusef-Zadeh, Morris and Chance 1984).

These structures alone will not dominate the total energy, line radiation, pulsar dispersion measure and low latitude intense scattering of the ISM. However, they have far reaching implications for VLBI, variability of compact radio sources, and precision timing of pulsars, although the modest duty cycle means that these events will be infrequently seen for a given source. The most extraordinary caustic events (such as the edge-on event seen from 0954+658) have associated splitting angles that are resolvable with intercontinental VLBI. Near the caustic, variation of the sub-images with

time and observing frequency will be rapid. This variation of the bright, elongated sub-images can mimic a large superluminal motion in a 'core-jet' source, although such motion will be unrelated to the large scale source structure and can be distinguished by its strong frequency dependence. Sustained imaging (with moderate bandwidth at several frequencies) during a caustic event, feasible with the VLBA, should yield the detailed profile of column density through the refractor. More generally, further studies with VLBI and intensity monitoring will yield important information on the interstellar medium and this component of interstellar seeing, which can be an important limitation for QUASAT. Because single perturbation, small α events highlight the caustic variation, the most discerning observations will be at high (e.g. $\gtrsim$2 GHz) frequencies and moderate to high galactic latitudes.

A number of studies have addressed the influence of the distributed scattering component of the interstellar medium on pulsar arrival times (Armstrong 1984; Blandford, Narayan and Romani 1984; Cordes, Pidwerbetsky and Lovelace 1986). Arrival time perturbations induced by strong refraction events include those due to dispersion measure fluctuations, geometrical path length variations and position errors in the timing fit. For a typical caustic event the DM and path length variations will dominate the arrival time perturbation, giving $\delta t \sim 0.4\nu_{GHz}^2$ ms (or $\delta DM \sim 0.1$pc cm^{-3}); if sheet structures are responsible face-on events may be detectable in millisecond pulsar (*eg.* PSR1937+21) timing at $\delta t \sim 4\mu$s (Rawley 1986). Rotation measure changes may also be detectable in polarization monitoring programs (Aller, Aller and Hodge 1981; O'Dea, *et al.* 1983; Hamilton, *et al.* 1979). For standard parameters and a magnetic field $\sim 10^{-5}B_{-5}$G, position angle variations of $\sim 0.1B_{-5}$rad will be seen at 1 GHz in caustic events. If the field is important in organizing the structures, then even larger variations with characteristic signatures may be seen.

In conclusion, the observations of Fiedler, *et al.* along with low frequency variability of quasars and multiple imaging of pulsars point to *high-density localized* ISM electron density fluctuations in a dense ionized medium (DIM) on scales $\sim 10^{14}$cm which can *individually* dominate the refractive scattering due to the general ISM along a given line of sight. The proposed observations should provide extensive probes of this component of the ISM and clean tests of the scenarios outlined above.

Acknowledgements

References:

Aller, H.D., Aller, M. & Hodge, P.E. *Astron. J.* **86**, 325-339 (1981).

Armstrong, J.W. *Nature* **307**, 527-528 (1984).

Blandford, R.D., Narayan, R. & Romani, R. W. *J. Astrophysics and Astronomy* **5**, 369-388 (1984).

Caswell, J.L. & Lerche, I. *Mon. Not. R. ast. Soc.* **187**, 201-216 (1979).

Cordes, J.M., Pidwerbetsky, A. & Lovelace, R.V. *Astrophys. J.* **310**, 737-767 (1986).

Cordes, J.M. & Wolszczan, A. *Astrophys. J.* **307**, L27-L31 (1986).

Fiedler, R.L., Dennison, B. Johnston, K.J. & Hewish A. *Nature,* **326**, 675-678 (1987).

Goodman, J.J., Romani, R.W., Blandford, R.D. & Narayan, R. *Mon. Not. R. ast. Soc.*, submitted (1987).

Hamilton, P.A., McCulloch, P.M., Manchester, R.N., Ables, J.G. & Komesaroff, M.M. *Nature* **265**, 224-225 (1979).

Hester, J.J. *Astrophys. J.* **314**, 187-202 (1987).

Hester, J.J., Parker, R.A.R. & Dufour, R.J. *Astrophys. J.* **273**, 219-242 (1983).

Hewish, A., Wolszczan, A. & Graham, D. *Mon. Not. R. ast. Soc.* **213**, 167-179 (1985).

McKee, C.F. & Ostriker, J.P. *Astrophys. J.* **218**, 148-169 (1977).

O'Dea, C.P., Dent, W.A., Balonek, T.J. & Kapitzky, J.E. *Astron. J.* **88**, 1616-1625 (1983).

Pearson, T.J. & Readhead, C.S. *Astrophys. J.* **234**, 61-81 (1981).

Rawley, L. Princeton University Thesis (1986).

Reynolds, R.J. *Astrophys. J.* **282**, 191-196 (1984).

Roberts, J. & Ables, J. *Mon. Not. R. ast. Soc.* **201**, 1119-1138 (1982).

Romani, R.W., Narayan, R. & Blandford, R.D. *Mon. Not. R. ast. Soc.* **220**, 19-49 (1986).

Yusef-Zadeh, F., Morris, M. & Chance, D. *Nature* **310**, 557-561 (1984).

Figure Caption:

Caustic formation by a localized plasma structure. A Gaussian profile of excess electron density (b) causes ray crossing beyond $D_c \sim 1$ (c). The loci of the resultant cusp point and fold line caustics is shown schematically by the solid curves (c) for 8.1GHz. After convolution with the intrinsic source diameter (a) an observer line of sight cutting across the structure at at distance D will produce a light curve (solid line in d) similar to that of 0954+658. In regions (i) and (iii) there will be three images; in (ii) one de-amplified image is present. At 2.7GHz the refraction is stronger and the caustics (dotted lines in c) form closer to the plasma lens. The increase in θ_{int} (a) gives a light curve (d) with two broad maxima shown by the dotted line.

199