A SEARCH FOR VERY HIGH ENERGY GAMMA RAYS FROM HERCULES X-1

by

PATRICK OLSON SLANE

A thesis submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

(Physics)

at the

UNIVERSITY OF WISCONSIN—MADISON

1988

A SEARCH FOR VERY HIGH ENERGY GAMMA RAYS FROM HERCULES X-1

Patrick Olson Slane

Under the supervision of Professor Ugo Camerini

The Haleakala Gamma Ray Telescope is used to detect Extensive Air Showers by utilizing the Atmospheric Čerenkov Technique. The resulting data is analyzed in an effort to detect the presence of γ-ray initiated showers above a threshold energy of approximately 200 GeV. Hercules X-1, a neutron star member of an accretion-powered X-ray binary system, was observed for 104 hours in May through July 1986. On May 13 a burst of activity of approximately 15 minutes duration was detected, with a period of $1^s.23593 \pm 0^s.00018$. Although this period differs significantly from the pulsar period expected from X-ray observations, it is consistent with measurements of the periodicity of VHE and UHE γ-rays by other observers operating during the summer of 1986. This surprising result indicates that the γ-ray production mechanism operates in a region well removed from the neutron star and may set constraints on models involving the interaction of the accretion disk with the magnetosphere associated with the neutron star.

Acknowledgements

The road which has led to the completion of this work has been populated by scores of friends and colleagues whose contributions have made my task enjoyable and intriguing. And while proper thanks can never be expressed in only a few lines, I am compelled to try.

It is only fair that I start by expressing my gratitude to the members of the Physics Department at U.W. Whitewater and especially to Ron Bergsten and Hugo Tscharnack. Without their efforts and continual support I may well have never studied physics beyond my freshman year. Similarly, thanks must go to Robert Greenler who took the time to show me that the beauty of physics is most readily found by remembering to put aside the books and look up from time to time.

I am very thankful for the contributions from all of the members of the Haleakala collaboration. The task of building the telescope, creating a workable laboratory (of sorts), and interfacing the complexities posed by a shoestring budget, an inexperienced group of astrophysicists, and pressure to produce new and exciting results has been difficult to say the least. That we have succeeded in some measure is a credit to all involved.

I would especially like to express my gratitude to my friend Andy Szentgyorgyi who has taught me a great deal of physics, introduced me to the finest beaches on Oahu, and without whom this experiment would have suffered greatly. My thanks also to Jim Matthews who, especially in my relative infancy, was instru-

mental in helping me understand the finer points of experimental physics.

Thanks must also go to Jack Fry for the many useful discussions we have had, and to Bob March whose efforts in bringing the results of my analysis on Her X-1 to press are especially appreciated. For countless hours of conversation during which, I believe, many important aspects of the experiment were scrutinized for the first time, I would like to thank Almus Kenter and John Finley. Thanks also to Mark Lomperski for the side by side work back in the early days and to Dick Loveless, Don Reeder, John Jennings and Gus Sinnis.

I would like to thank Dan Weeks and John Learned for their hospitality during my year in Hawaii. The fond memories of sailing, snorkeling, and TGIF will not soon fade.

To Jim Gaidos, whose efforts in keeping the telescope manned during my tenure as an operator were particularly appreciated, I am thankful for the many enjoyable hours of conversation, hard work, and Thai food. Give me a call the next time you are going to the Shakespeare pub. Thanks, too, to Pete Palfrey, Frank Loeffler, Glenn Sembroski, and Chuck Wilson for contributions varied but not few.

For reasons which I can only begin to enumerate, I am thankful to my friend and advisor Ugo Camerini. Although my purity remains intact (for better or worse), my outlook on physics, life, us, and them have been forever altered for the better as the result of lessons learned as his student.

For contributions of quite a different, but no less important nature, I am proud to be able to thank the members of my family as well as the Olson family,

for their encouragement and interest which have been such a source of motivation for me. For similar contributions, and for the ever important friendship which is needed for one to pursue a difficult task, I would like to thank Jim Solveson, Eric Switalski, Terry Gavin, Glen Goytowski, Brian Clark, and Tom and Jodi Goldsworthy.

Lastly, I would like to thank my wife Kathy for her years of patience, enthusiasm, companionship, and love. Were it not for the self-confidence she has helped instill in me, the goal of returning to the study of physics would have never been attempted. And without her continued support, especially during the trying times of the Qualifying and Preliminary exams, the goal would certainly not have been accomplished. With all my love, it is to her and our son Kevin that this work is dedicated.

This work was supported in part by the United States Department of Energy Contract DE-AC02-76ER00881.

Contents

Chapter 1

Introduction

The field of very high energy (VHE) γ-ray astronomy is relatively young. Its growth has been steady and, although the basic experimental technique used in the earliest such experiments is still in use today, numerous efforts to increase the efficiency of the technique have resulted in sophisticated telescope arrays which are still being improved. Unlike many other wavelength bands in astronomy, however, the VHE band has yet to find its "standard candle" source by which experiments can be calibrated and tuned. Instead, the field has developed through the slow accumulation of rather weak signals from a small collection of varied sources. The object of this thesis is to conduct a search for such signals from the X-ray binary system Her X-1/HZ Her. In order to accomplish such a task, it is necessary to understand the subtleties of the Atmospheric Čerenkov Technique (ACT) which is used to identify the presence of high energy photons (or particles) impinging on the atmosphere. Further, one must investigate the analysis techniques pertinent to the identification of signals in the presence of a very large cosmic ray background as well as the physics associated with the binary system in which the γ-rays are produced in order to plan an efficient strategy for the search for such signals. In this chapter we present an outline of

the thesis and summarize the content of the chapters to follow.

1.1 Sources of Energetic Radiation

As a general class, the compact objects such as neutron stars and black holes are likely candidates for production of extremely energetic radiation. Because of the very large gravitational fields associated with these objects, and often the presence of large rotating magnetic fields, the energetics are favorable for such production. In Chapter 2 we review the physics of the environment in the vicinity of rotating magnetic neutron stars. When such objects are located in binary systems in which the companion objects are normal stars, possibly which have evolved toward the red giant stage, we find that accretion of matter from the companion onto the neutron star can result in the release of huge amounts of energy. The accretion disks which form in such binary systems drive the energetics of the system and determine the state of the outer regions of the magnetosphere associated with the neutron star. We shall see that such an environment lends itself to acceleration of charged particles to very high energies. The production of γ-rays from such particles may occur in a variety of ways although the exact mechanism is still unknown. Studies of the periodicity of the γ-rays relative to that of the neutron star rotation could provide crucial information about the site of γ-ray production.

1.2 VHE γ-Ray Astronomy

The technique used to detect γ-rays of energies in the range 0.1-10 TeV involves collecting the Čerenkov photons which are produced by electrons and positrons in the particle cascade that the γ-rays initiate upon entering the atmosphere. In Chapter 3 we discuss this technique and, through Monte Carlo simulations, the characteristics which describe Extensive Air Showers (EAS). The Čerenkov radiation produced in such showers results in a photon density on the ground of about 135 m^{-2} TeV^{-1} (evaluated in the optical range between 300 nm and 475 nm and at an observation height of $\sim$ 3 km). This photon density decreases nearly exponentially with distance from the shower core and is reduced to e^{-1} of its maximum value at an impact parameter of about 35 m. The photons arrive in a "pancake" which is several nanoseconds in duration. Showers initiated by hadrons (primarily protons in the energy range under consideration), which provide the nearly overwhelming background to such an experiment, contain less light for a given primary energy because a reasonable fraction of the primary energy is used in the production of charged pions which do not contribute to the electron component of the shower.

Experimental techniques for rejecting showers which are hadronic in origin in an effort to enhance the γ-ray signal-to-noise ration have taken several forms. Because hadronic showers are characterized by pion production early in the development, the eventual pattern of photons on the ground tends to be less uniform than that from a γ-ray initiated shower. Further, because the charged pions eventually decay into muons, the hadronic showers tend to be muon rich com-

pared with their γ-ray counterparts. The muons, which are very penetrating, will reach observation level with sufficient energy to produce Čerenkov radiation. This radiation will form small rings around the muon axis which might be discernable if small enough detail is resolvable in the focal plane of the telescope used to collect the Čerenkov light.

Because the region of sky over which the Čerenkov radiation from a shower develops is rather large, correspondingly large apertures ($\sim 1°$) are necessary to optimize the collection of the light (for example, apertures of $0°.75$ are used on the Haleakala γ-ray telescope). However, the use of a large aperture increases the number of oblique proton-induced showers (*i.e.* those from directions different from that of the source under observation) which may be detected by the telescope. Because these showers are isotropically distributed, the number which become background events increases as the square of the aperture angle. Because these showers approach from directions different from that of the telescope axis, their images are different from those which come from directions near the axis. In addition, the light from such oblique showers will be spread over a somewhat larger time than that for showers arriving along the axis.

Various experimental groups have developed different techniques for attempting to discriminate γ-ray induced showers from proton-induced showers. The Whipple Observatory group, for example, uses a many-pixel array of photomultiplier tubes at the focal plane of a large segmented mirror in an effort to reject hadronic showers based on the details of the shower images. The Haleakala group, opting to collect all of the focal plane image on only several photomul-

tipliers, uses fast electronic timing to attempt to reject oblique showers on the basis of their longer development time (see Chapter 3). To date, none of the techniques has clearly established the ability to discriminate between γ-ray and proton initiated showers, largely because of the large fluctuations in the development of individual showers. To be fair, however, improvements and refinements currently underway may well prove successful in the future.

1.3 Analysis Techniques

Detecting γ-rays from the direction of a particular source in the presence of the large cosmic ray background mentioned in the previous section is a difficult task. In order to recognize an overall excess of events from the direction of the source at the low signal levels thus far identified with such sources requires a very precise understanding of the background behavior. This is a very difficult prospect except, possibly, during large outbursts of activity. In the case of a periodic signal, variations of basic Fourier techniques can be used to improve detectability. These techniques are reviewed in Chapter 4 where we investigate the Rayleigh test in particular and, through Monte Carlo simulations, illustrate the various pitfalls characteristic of this test when performed on data which are dominated by random background. The statistics associated with period searches are also investigated as are the barycentering techniques necessary to remove the effective Doppler shifts associated with motions of the Earth and of the γ-ray source.

1.4 VHE γ-Rays from Hercules X-1

Hercules X-1 is a highly magnetized neutron star with a rotational period of $\sim 1^s.24$. It is located in a binary system with the $\sim 2.3 M_\odot$ A-type star HZ Herculis. The characteristics of this binary system are outlined in Chapter 5 where we summarize the results of optical, X-ray, and γ-ray studies and review the basic model which best describes the details of these studies. VHE γ-rays from Her X-1 have been detected at various levels of significance by several groups and we review the details of these detections in order to guide our search techniques.

In Chapter 6 we explain in detail the results of a search for VHE γ-rays from Her X-1 in the data collected at the Haleakala Gamma Observatory between May and July of 1986. The identification of a burst of activity on 13 May, which is characterized by periodicity distinctly different from that of the rotating neutron star, is outlined and compared with results obtained by two other experimental groups operating that summer. The results of this study have been submitted for publication in the *Astrophysical Journal Letters*.

Chapter 2

Compact Objects: Sources of VHE γ-Rays

Twinkle twinkle little star.
How I wonder what you are...
Jane Taylor

Like a river to a stream
It's the famous final scene.
Bob Seger

VHE γ-rays are the signatures of extremely energetic processes. They generally owe their existence to the decay of neutral pions which are products of energetic collisions of protons with other matter or to radiation produced by ultrarelativistic electrons moving in strong magnetic fields. The conditions needed for such pion production and/or curvature radiation are found in the vicinity of compact astrophysical objects — in particular, neutron stars. Since such compact objects represent the endpoint of stellar evolution, it is instructive to consider the fundamental physics associated with such evolution. In addition, since many compact objects are members of binary systems, such a discussion will facilitate the understanding of the roles played by the companion stars in

such systems. We begin the discussion with a review of stellar structure and evolution followed by a discussion of the physics of neutron stars. Binary systems in which one member is a neutron star are then discussed with an emphasis on accretion of matter from the companion star. Finally, models depicting the production of VHE γ-rays in such systems are discussed.

2.1 Stellar Structure

Stars owe their existence to the mutual gravitational attraction between all of the individual matter constituents. The Virial theorem guarantees that, as gravitational potential energy V is converted into thermal energy K, we have

$$K = -\frac{V}{2} \tag{2.1}$$

where we may write

$$V = -\frac{\alpha G M^2}{R} \tag{2.2}$$

where R is the radius, M is the mass, and α is some proportionality constant. Assuming an ideal gas equation of state,

$$K = \frac{3}{2} N k T \tag{2.3}$$

where N is the number of atoms or molecules, T is the temperature, and k is Boltzmann's constant. The total energy of the star is

$$E = K + V = -\frac{\alpha G M^2}{2R}. \tag{2.4}$$

Thus, for constant mass, we see that as energy is lost (for example by radiation) the radius R decreases. However, from the expressions above, we find that

$$T = \frac{\alpha}{3Nk}\frac{GM^2}{R}. \tag{2.5}$$

Thus, as the radius decreases, the temperature increases. An increase in temperature, however, results in further radiation of energy which then causes R to decrease further. The star is thus destined to become a compact object. Only when internal pressures are large enough to halt the decrease of R will the configuration become stable. As we shall see, several such stability points are found as a star evolves, but most are merely temporary resting places along a path to final collapse.

The problem of determining the interior structure of a star requires relating the pressure, temperature, and density of the star with its mass and chemical composition. Since the time scale over which the evolutionary process occurs is generally long, it is reasonable to take static conditions as a first approximation. The basic structure equations are then the conservation of mass,

$$\frac{dM_r}{dr} = 4\pi r^2 \rho(r) \tag{2.6}$$

the condition for hydrostatic equilibrium (whereby internal pressure supports the weight of the exterior layers of the star),

$$\frac{dP}{dr} = -\frac{GM_r\rho(r)}{r^2} \tag{2.7}$$

and the condition for thermal equilibrium,

$$\frac{dL_r}{dr} = 4\pi r^2 \epsilon\rho. \tag{2.8}$$

Here, M_r is the mass interior to radius r, L_r is the net energy flux through a shell of radius r, $\rho(r)$ is the density, and ϵ is the power liberated by nuclear reactions in erg gm^{-1} s^{-1} (see sec. 2.1.3). To construct a model, then, these structure equations must be satisfied simultaneously with the equation of state $P(\rho, T)$, the transport equation $\nabla T = f(\rho, T)$, and the equations for energy production.

2.1.1 Equation of State

Temperatures in stellar interiors are high enough for the atoms to be almost completely ionized. Further, since the bulk of the matter is hydrogen, the overwhelming majority of the particles which constitute the star are free electrons and protons. The mean free path for these particles is small compared with the distance over which temperature changes occur so that we may assume local thermodynamic equilibrium. Thus, the pressure depends only on the density, temperature, and chemical composition. There are three basic contributions to the pressure: Gas pressure, radiation pressure, and electron degeneracy pressure.

- Radiation Pressure - By virtue of the momentum carried by photons, the pressure contributed is

$$P_{rad} = \frac{1}{3} a T^4 \tag{2.9}$$

  where $a = 7.55 \times 10^{-15}$erg cm^{-1} K^{-4}.

- Gas Pressure - Since interaction energies $\ll$ thermal energy we may assume ideal gas behavior. Thus,

$$P_{gas} = \nu k T \tag{2.10}$$

where ν is the number of particles per unit volume. Letting μ be the mean molecular weight, we have

$$P_{gas} = \frac{N_0 k}{\mu} \rho T \tag{2.11}$$

where N_0 is Avogadro's number. We can account for the chemical composition by defining

X = weight fraction of hydrogen

Y = weight fraction of helium

Z = weight fraction of heavier elements.

Then

$$\frac{1}{\mu} = 2X + \frac{3}{4}Y + \frac{1}{2}Z \tag{2.12}$$

where each term represents

$$\frac{\#\text{particles (nucleus + electrons)}}{\text{weight of nucleus (in AMU)}} \times \text{fractional abundance.}$$

Typical values found in Population I stars (*i.e.* stars, generally associated with spiral arms of galaxies, whose composition includes heavier elements synthesized in the evolution of earlier stars) are $X \simeq 0.6$, $Y \simeq 0.38$, $Z \simeq 0.02$.

- Degeneracy Pressure - The Pauli Exclusion Principle forbids the occupation of a given quantum state by more than one electron. As a result, if a collection of electrons are confined to a given volume, the available momentum states become fully populated; although there are an infinite

number of momentum states, the available energy dictates which of these are within the reach of the electrons. When all of the states up to some particular state are filled, the collection is said to be *degenerate*. Further reduction of the volume can be achieved only by supplying sufficient energy to populate higher momentum states; hence, compression is resisted by so-called *degeneracy pressure*. For complete degeneracy, the number of electrons as a function of momentum is

$$n_e(p) = \begin{cases} 0 & \text{if } p > p_0 \\ \frac{8\pi}{3h^3}p_0^3 & \text{if } p \leq p_0. \end{cases} \tag{2.13}$$

Normalizing to the total number of electrons,

$$n_e = \int_0^\infty n_e(p)dp = \frac{8\pi}{3h^3}p_0^3 \Rightarrow p_0 = \left(\frac{3h^3 n_e}{8\pi}\right)^{\frac{1}{3}}. \tag{2.14}$$

The associated pressure is then

$$P = \frac{1}{3}\int_0^\infty p v_p n(p)dp. \tag{2.15}$$

In the non-relativistic case, we have $v_p = p/m$ so that

$$P_{e,nr} = \frac{8\pi}{15mh^3}p_0^5 = \frac{8\pi}{15mh^3}\left(\frac{3h^3}{8\pi}n_e\right)^{\frac{5}{3}} \tag{2.16}$$

independent of temperature. The relativistic case and the case of partial degeneracy, although more complicated, can be approached in a similar manner. Degeneracy of protons and other nuclei do not occur at any appreciable level because of their larger masses. Thus, if the electrons are non-degenerate,

$$P_n + P_e = P_{gas}$$

where P_n is the pressure contribution from the nuclei and P_e is that from the electrons. In the degenerate case,

$$P_n = \frac{N_0 k}{\mu_n} \rho T$$

where μ_n is the weight (in AMU) of the nuclei only. The total pressure is thus

$$P = P_n + P_e + P_{rad}. \tag{2.17}$$

2.1.2 Energy Transport

The stellar interior is characterized by a temperature gradient ∇T. In its effort to achieve global thermodynamic equilibrium, the star may transport energy through radiative transfer, conduction, or convection. Neutrino emission, which also transports energy from the core of the star, is not characterized by interactive transport through the stellar matter; the process is not a function of temperature gradient. Conduction, whereby energy is transported directly through particle interactions, is much less efficient than radiative transfer (except in a degenerate electron gas) and, thus, does not represent a significant contribution.

- Radiative Transfer - Photons emitted thermally in hot regions and absorbed in cooler regions transport energy. The effectiveness of the process is a function of both the temperature gradient and the ability of the photons to travel from one region to another, *i.e.* the *opacity*. If we define κ to be the opacity (*i.e.* the mass absorption coefficient, in $cm^2\ gm^{-1}$) and F as the flux of radiation, then $\kappa \rho F$ is the energy flux absorbed. But from the

continuity equation,

$$\nabla \cdot \mathbf{S} = -\frac{\partial u}{\partial t} \tag{2.18}$$

where **S** is the Poynting vector and u is the energy density. By definition,

$$\frac{\partial u}{\partial t} = \kappa \rho F \tag{2.19}$$

and since the pressure is merely the momentum flux,

$$P = \frac{1}{c}\bar{S} \tag{2.20}$$

we have

$$\frac{dP}{dr} = \frac{dP}{dT}\frac{dT}{dr} = -\frac{\kappa\rho}{c}F. \tag{2.21}$$

But from Equation 2.9,

$$P = \frac{1}{3}aT^4 \Rightarrow \frac{dP}{dT} = \frac{4}{3}aT^3 \tag{2.22}$$

and since

$$F = \frac{L_r}{4\pi r^2} \tag{2.23}$$

we have the result

$$\left(\frac{dT}{dr}\right)_{rad} = -\frac{3}{4ac}\,\frac{\kappa\rho}{T^3}\,\frac{L_r}{4\pi r^2}. \tag{2.24}$$

The primary contributions to the opacity κ are photoionization, inverse-bremsstrahlung, and electron scattering. The relative importance of each mechanism is dependent upon both temperature and composition.

- Convective Transfer - Thermonuclear reactions in the central regions of a star are the source of energy that is ultimately radiated from the surface of the star. If the temperature gradient is sufficiently large, the rate at which

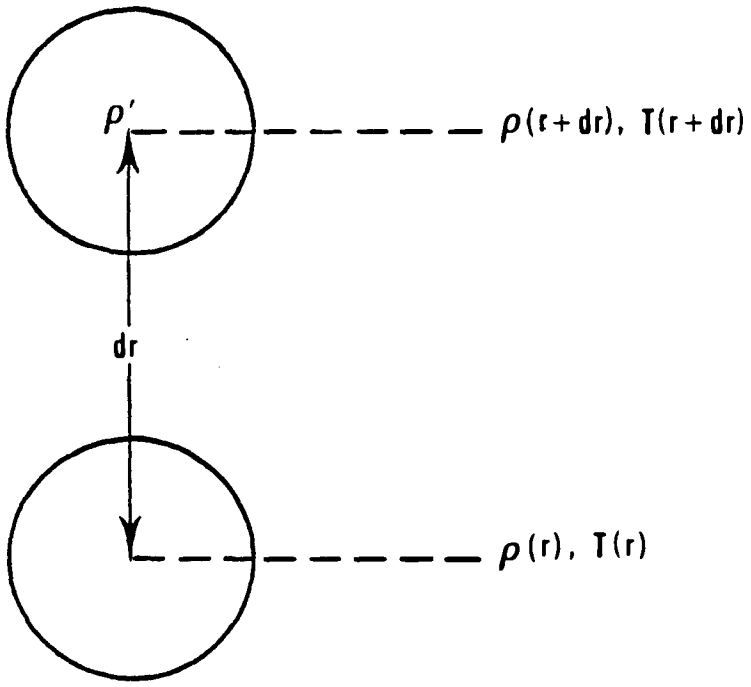


Figure 2.1: Convective instability occurs when the density of a fluid element which has been raised adiabatically by a distance *dr* is less than the density of the surrounding material.

radiation carries energy to the surface may be insufficient. In this case, the star becomes unstable against convection and energy is transported through bulk motion of stellar matter. The conditions for convective instability are illustrated in Figure 2.1. If a fluid element in the star is raised adiabatically from a radius r to $r + dr$, its density will change from $\rho(r)$ to

some value ρ'. Convective instability occurs if

$$\rho' < \rho(r + dr). \tag{2.25}$$

We may define the adiabatic coefficient Γ_1 by

$$\left(\frac{dP}{d\rho}\right)_{ad} = \Gamma_1 \frac{P}{\rho}. \tag{2.26}$$

Then

$$\rho' = \rho(r) + (d\rho)_{ad} = \rho(r) + \frac{1}{\Gamma_1}\frac{\rho(r)}{P}\left(\frac{dP}{dr}\right) dr \tag{2.27}$$

while

$$\rho(r + dr) = \rho(r) + \left(\frac{d\rho}{dr}\right) dr. \tag{2.28}$$

Hence, we can write the condition for convective instability as

$$\frac{1}{\Gamma_1}\frac{\rho}{P}\left(\frac{dP}{dr}\right) < \frac{d\rho}{dr}. \tag{2.29}$$

Often the inequality for convective instability is met only over limited portions of the star. For example, low mass stars are characterized by a convective surface and a radiative core. This is the result of the relatively low surface temperature which produces a very high opacity (partially due to neutral hydrogen) in the subsurface layers. High mass stars, on the other hand, often have convective cores and radiative surfaces. This opposite behavior is due to the lower opacity at the surface and the much larger energy production rate at the core. In both cases, convection occurs in regions where radiation is not capable of transporting energy at a sufficient rate. It should be noted that convection can play an important role

in the development of the star because of mixing of heavier elements with the hydrogen in the vicinity of the core. Detailed stellar models, then, are quite sensitive to convection effects (which may be further complicated by the effects of rotation).

2.1.3 Energy Generation

The mechanism by which stars produce their energy consists of thermonuclear fusion in the central regions. Such nuclear reactions convert light elements into heavier elements with the excess binding energy per nucleon released in the interactions powering the star. In order for nuclei with charges $Z_1 e$ and $Z_2 e$ to be brought close enough for fusion to occur, the potential barrier

$$V = \frac{kZ_1 Z_2 e^2}{r} = \frac{1.44 Z_1 Z_2}{r(\text{fm})} \text{MeV} \tag{2.30}$$

must be overcome. However, at temperatures typical of those found in stellar cores ($T \sim 10^7$K) the average energy of the nuclei is

$$E \sim kT \approx 1\text{keV}.$$

Thus only through tunneling effects can barrier penetration be accomplished.

We may define the energy production in terms of ϵ, the amount of energy released per gram per second in nuclear reactions. The the total energy generated per cm^3 per second is

$$\rho\epsilon = \sum_{reac} r_{12} Q_{12} \tag{2.31}$$

where r_{12} is the rate per cm^3 for interactions of nuclei of type 1 with those of type 2 and Q_{12} is the energy released in each such reaction; the sum is over all

possible reactions. If n_1 and n_2 represent the number densities of the nuclei, then the reaction rate may be written

$$r_{12} = \frac{n_1 n_2}{1+\delta_{12}} < \sigma v >_{12} \tag{2.32}$$

where δ_{12} is the Kronecker delta and

$$< \sigma v >_{12} = \int_0^\infty f(v) v \sigma_{12}(v)\, dv. \tag{2.33}$$

Here $f(v)$ is the velocity distribution of the nuclei (generally assumed to be Maxwellian) and σ_{12} is the reaction cross-section. We may write the cross-section as

$$\sigma = \frac{S(E)}{E} \exp\left(-\frac{2\pi Z_1 Z_2 e^2}{\hbar v}\right) \tag{2.34}$$

where $S(E)$ contains the additional energy dependence. Often $S(E) \approx S_0$ is a reasonable approximation. In this limit, the reaction rate may be approximated by

$$r_{12} \simeq (2.6 \times 10^{29}) \frac{X_1 X_2}{1+\delta_{12}} \frac{\rho^2 S_0 \tau^2 e^{-\tau}}{A_1 A_2 A_{12} Z_1 Z_2} \text{ cm}^{-3}\ \text{s}^{-1} \tag{2.35}$$

where

$$A_{12} = \frac{A_1 A_2}{A_1 + A_2}$$

is the reduced mass, X_1 and X_2 are the atomic abundances of nuclei 1 and 2, and

$$\tau = 42.48 \left(\frac{Z_1^2 Z_2^2 A_{12}}{T_6}\right)^{\frac{1}{3}}. \tag{2.36}$$

Here T_6 is the temperature expressed in units of 10^6K.

Nuclear energy production generally consists not of a single interaction but, rather, of an interaction cycle. For example, in the majority of normal stars, the

primary energy production mechanism is the fusion of hydrogen to helium. It is clear, however, that one cannot merely take four protons and join them to make helium; rather, a process such as the PPI cycle must occur:

$$\begin{aligned} H^1 + H^1 &\rightarrow H^2 + e^+ + \nu_e \\ H^2 + H^1 &\rightarrow He^3 + \gamma \\ He^3 + He^3 &\rightarrow He^4 + 2H^1. \end{aligned} \tag{2.37}$$

In such a cycle, the total reaction rate is dominated by the rate of the slowest process (the first process, in this case, which has an exceedingly small cross-section).

Depending on the temperature and the abundance of He^4, the reaction

$$He^3 + He^4 \rightarrow Be^7 + \gamma$$

may dominate the consumption of He^3. The PPII and PPIII cycles may then operate:

$$\begin{aligned} He^3 + He^4 &\rightarrow Be^7 + \gamma \\ Be^7 + e^- &\rightarrow Li^7 + \nu_e \\ Li^7 + H^1 &\rightarrow 2He^4 \end{aligned} \tag{2.38}$$

or

$$\begin{aligned} He^3 + He^4 &\rightarrow Be^7 + \gamma \\ Be^7 + H^1 &\rightarrow B^8 + \gamma \\ B^8 &\rightarrow Be^7 + e^+ + \nu_e \\ Be^8 &\rightarrow 2He^4. \end{aligned} \tag{2.39}$$

In practice, all three chains operate simultaneously. Figure 2.2 illustrates the temperature dependence of the relative contributions of each cycle; in the sun, reactions are ~40% PPI, 60% PPII, with a very small fraction of PPIII.

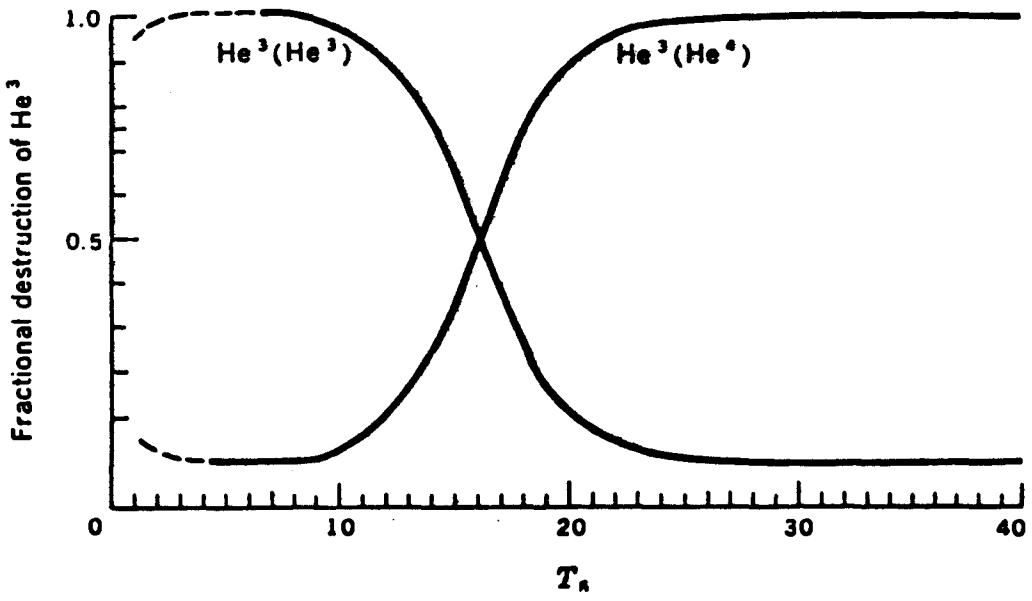


Figure 2.2: The fraction of the He^4 production due to PPI, PPII, and PPIII cycles.

If heavier elements are sufficiently abundant in the star, helium can be synthesized in reactions involving carbon and nitrogen nuclei with protons in the

so called CN cycle:

$$\begin{aligned} C^{12} + H^1 &\rightarrow N^{13} + \gamma \\ N^{13} &\rightarrow C^{13} + e^+ + \nu_e \\ C^{13} + H^1 &\rightarrow N^{14} + \gamma \\ N^{14} + H^1 &\rightarrow O^{15} + \gamma \\ O^{15} &\rightarrow N^{15} + e^+ + \nu_e \\ N^{15} + H^1 &\rightarrow C^{12} + He^4. \end{aligned} \tag{2.40}$$

Note that the C^{12} nucleus merely acts as a catalyst in the reaction. Bombardment of N^{15} nuclei with protons does not always produce C^{12}; the reaction

$$N^{15} + H^1 \rightarrow O^{16} + \gamma$$

occurs with a relative probability of $\sim 4 \times 10^{-4}$. This initiates the sub-cycle

$$\begin{aligned} N^{15} + H^1 &\rightarrow O^{16} + \gamma \\ O^{16} + H^1 &\rightarrow F^{17} + \gamma \\ F^{17} &\rightarrow O^{17} + e^+ + \nu_e \\ O^{17} + H^1 &\rightarrow N^{14} + He^4 \end{aligned} \tag{2.41}$$

The final production of N^{14} is then fed into the middle of the CN cycle. As a result, the entire process is generally referred to as the CNO bi-cycle. The reaction rates for the CNO cycle (which are extremely temperature sensitive) and its competition with the PP cycles are discussed thoroughly by Clayton (1968).

It is worth briefly noting that, because of the extremely large times required for photons to diffuse from the center of a star to its surface, the only direct test of the theoretical calculations regarding current energy production in the sun is the

measurement of the rate at which neutrinos (which escape virtually unimpeded) are produced. To date, the most extensive study in this area has been carried out by Davis *et al.* (1964) using a large tank of Cl^{37} located 1500 m underground in the Homestake mine in South Dakota. Unfortunately, the energy threshold for the experiment is about 0.81 MeV; the most plentiful neutrinos from the sun are produced in the first step of the PPI cycle and have a maximum energy of 0.42 MeV. The solar neutrinos for which the detector is most sensitive (those from the PPIII cycle) occur in only 0.02% of solar reactions. Nonetheless, after many years of operation the experiment has established a clear deficiency (by about a factor of 3) in the number of neutrinos expected. This puzzle is one of the most interesting and important issues in stellar astrophysics and the search for its solution has produced more than a few suggestions including deviations from steady-state behavior (Fowler 1972), non-Maxwellian distributions of the nuclei involved in the neutrino production process (Clayton 1974, Clayton *et al.* 1975a), and even the presence of a black hole at the center of the sun (Clayton *et al.* 1975b). Aside from deviations from the standard solar model, suggestions involving neutrino oscillations and neutrino decay have also been made. At present, however, there is no accepted explanation to the solar neutrino puzzle.

2.2 Stellar Evolution

For the purpose of developing the past history of compact objects, as well as to establish terminology relevant to further discussion concerning binary systems, let us review the standard model of stellar evolution. This subject is broad, of

course, and we shall merely summarize the more general features. More detailed discussion can be found in many standard texts (*e.g.* Clayton 1968, Cox and Guili 1969). Excellent review articles have also been written on most aspects of the problem (*e.g.* Iben 1974). For the sake of simplicity, we will divide the discussion into three stages; main sequence stars, post main sequence evolution, and binary evolution.

2.2.1 Main Sequence Stars

One will recall from introductory astronomy that a scatter-plot of the luminosity (or absolute magnitude) vs effective temperature (or spectral index) for stars, the so-called *Hertzsprung-Russell (H-R) diagram* results in several populated regions (Figure 2.3). A large percentage of the stars lie along the region referred to as the *main sequence.* Such stars are in a temporary steady state in which thermonuclear fusion of hydrogen to helium (as discussed in Section 2.1.3) provides sufficient pressure to halt gravitational contraction. The differences in luminosity and temperature of these stars are entirely due to differences in mass. Direct measurements of stellar masses are possible only for binary systems. Studies of such systems show that, for main sequence stars, a mass-luminosity relation exists;

$$\frac{L}{L_\odot} \simeq \left(\frac{M}{M_\odot}\right)^\eta \tag{2.42}$$

where η lies in the range of 3.5 to 4. The main sequence lifetime of a star is inversely related to its mass;

$$T_{MS} \approx 13 \times 10^9 \left(\frac{M}{M_\odot}\right)^{-2.5} \text{ yr.} \tag{2.43}$$

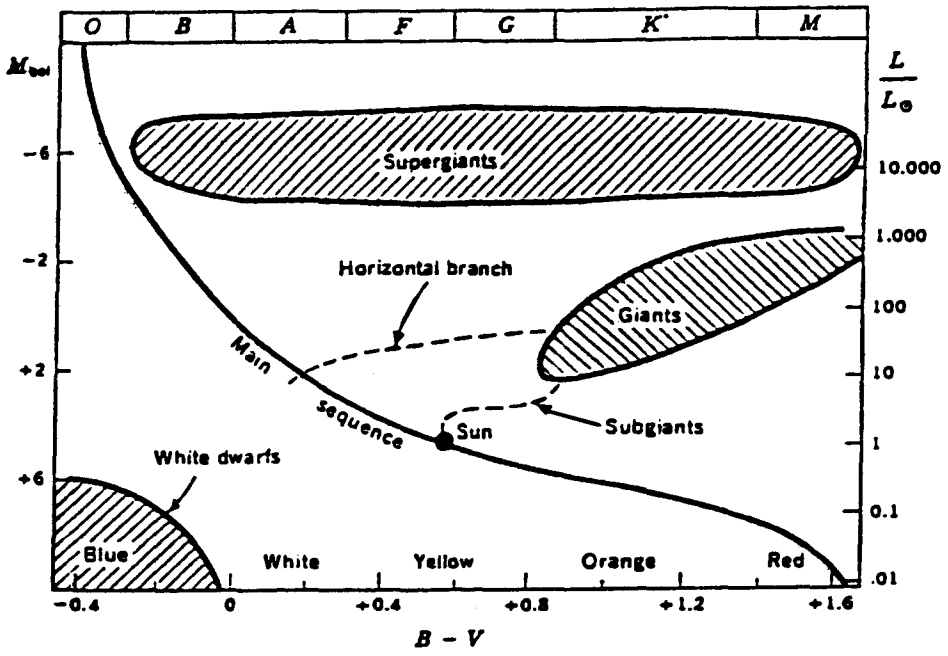


Figure 2.3: A schematic representation of the heavily populated ares in the H-R diagram. Here the absolute bolometric magnitude M_{bol} is plotted against color index $B - V$. A high percentage of stars lie near the main sequence. The next most populous groups are the white dwarfs and the giants.

The hot blue giants, which produce energy predominately through the CNO cycle, exhaust the supply of hydrogen at their cores in only a few million years while the cool red dwarfs, burning their fuel much more slowly (chiefly through the PPI cycle), survive for many billions of years. In either case, the supply of hydrogen in central regions eventually becomes depleted.

In that stars of different masses have quite different luminosities and lifetimes, it is not surprising that other general properties, such as mass loss, internal convection, or optical spectra, can also differ considerably. The spectral classes

of stars are, in fact, defined on the basis of these differences. Thus, spectra of O stars reveal lines of highly ionized atoms while those of the K and M stars display lines from neutral metals and even molecular bands of titanium oxide (which cannot exist at the higher temperatures of more massive main sequence stars).

The main sequence is generally divided into upper and lower regions where the upper main sequence is characterized by stars with a radiative surface and a convective core while the lower region contains stars with convective surfaces and radiative cores. Details regarding convection affect the future evolution of a star, in particular because of mixing in stellar interiors.

Mass loss during the main sequence phase occurs primarily in the form of a *stellar wind.* This steady, slow loss which is the result of heating of the tenuous outer layers of the star can play an important role in close binary systems (Section 2.4). Although the solar wind has been well measured, such winds in other stars are difficult to detect. Evidence suggests, however, that such mass loss can in some cases be considerably larger than that for the sun.

2.2.2 Post Main Sequence Evolution

Once hydrogen has been depleted in the central regions of a star there are no longer sufficient means of supporting the gravitational pressure. As a result, the core undergoes collapse. Energy is provided from the release of gravitational energy and from hydrogen burning in the shell surrounding the core. As the core collapses, its temperature, density, and pressure increase. The outer lay-

ers expand as a result of the energy released at the center. This expansion is accompanied by an increase in hydrogen shell burning so that, while the outer layers become cooler (and, thus, redder) the overall luminosity increases. This evolution from the main sequence follows a path to the upper right hand region of the H-R diagram; the star becomes a *red giant.* Its radius is enlarged by a factor of nearly 100.

For stars with sufficient mass ($> .25M_{\odot}$), the core temperature eventually becomes high enough to initiate the fusion of helium to carbon and oxygen. In stars with masses less than about $7M_{\odot}$, the helium core is degenerate when this process begins. Because degeneracy pressure depends only on density and not on temperature (Equation 2.11) the core cannot expand and cool in response to the temperature increase. Rather, because the nuclear reaction rates depend sensitively on temperature (for this process, $\epsilon \propto T^{40}$), the temperature rise results in a thermal runaway or *helium flash.* Only when temperatures become high enough to lift the degeneracy of the core can expansion and cooling occur. After this the star settles into a state whereby helium burning and hydrogen shell burning supply its energy.

The further evolution of a star follows much the same pattern as above. After helium is depleted, the core collapses until temperatures are high enough for carbon fusion accompanied by burning of shells of helium and hydrogen. The onset of such processes, which occur only in stars with sufficient mass for large enough core temperatures to be reached, may also be accompanied by carbon flashes. For stars with mass $\leq 8M_{\odot}$, the energy released in the carbon flash is

probably sufficient to explode the core and disrupt the entire star. For more massive stars, an oxygen burning cycle may be initiated after the completion of the carbon cycle. By this time, neutrino losses play a dominant role in the transport of energy from the core.

For less massive stars, core degeneracy will be sufficient to support the gravitational collapse before the onset of further fusion cycles. The remnants of such stars are white dwarfs. These remnants are stable up to the Chandrasekhar mass $M_c \approx 1.4 m_\odot$. For stars more massive than $\sim 12 M_\odot$, the core is pushed beyond this limit and further fusion cycles are ignited. However, the continual fusion of heavier and heavier elements cannot proceed past iron because beyond this point the reactions cease to produce energy. Hence, once an iron core is reached, the star cannot support itself against the influence of gravity and final collapse occurs. Depending on the final mass, sufficient pressure may be supplied after the core collapse has forced protons and electrons to form neutrons through inverse β-decay. Such collapse is rapid and violent resulting in a supernova explosion. The compact object left behind is a neutron star (see Section 2.3.1). In the case of extremely large mass, even the neutron star configuration does not supply sufficient pressure to balance gravity and the collapse continues to form a black hole.

2.2.3 Binary Evolution

Stars in binary systems can evolve along quite different paths than isolated stars [see review articles by Paczyński (1971) and Van den Heuvel (1978, 1987)]. This

is the result of mass transfer from one binary member to the other. In that the member which is initially the more massive of the two will evolve to the red giant stage more quickly, it is possible that, upon expansion of its outer layers, it will overfill its Roche lobe (see Section 2.4.1) and mass will flow onto the companion star. Depending upon orbital parameters and relative masses, this mass transfer can be sufficiently large to result in the companion star becoming the more massive member. Because post main sequence evolution is determined primarily by the core, not the stellar envelope, the evolved star will continue on its path toward final collapse, thus forming some type of compact object. Meanwhile, the companion star begins to pass much more rapidly through its main sequence stage as a result of its newly acquired mass. It is likely to possess a strong stellar wind which may be partially accreted onto the compact companion. As it evolves toward the red giant stage, large mass transfer onto the compact companion occurs. If the compact companion is a neutron star, such a configuration is likely to become an x-ray binary as discussed in Section 2.4. If it is a white dwarf, it is possible for the mass transfer to trigger a Type I supernova. In the end, assuming that no cataclysmic event disrupts the entire binary system, two compact objects will be remain.

2.3 Pulsars

The 1968 discovery of pulsars by Jocelyn Bell and Anthony Hewish was an unexpected result of an experiment designed to perform a large-scale survey of radio galaxies using the interplanetary scintillation technique (Hewish, 1975). In the

discovery paper (Hewish *et al.* 1968), several important model constraints were outlined. The very short duration of the pulses indicated that the source size could not exceed $\sim 4.8 \times 10^3$km. Absence of detectable parallax set a lower limit of 10^3 AU on the source distance while an upper limit of ~65pc was established on the basis of the measured dispersion of the radio signal (pulsar distances are still expressed in terms of the *dispersion measure*

$$DM \equiv \int_0^L n_e dl$$

where n_e is the electron number density and L is the pulsar distance). Models consisting of oscillating white dwarfs or neutron stars were also suggested. Indeed Cameron (1965) and Finzi (1965) had previously suggested that pulsation of neutron stars could play an important role in the development of supernova remnants through the generation of hydromagnetic shock fronts at the stellar surface. Thus it is not surprising that soon after the discovery of pulsars, models involving neutron stars were proposed. Pulsations, however, were ruled out partially because the observed periods seemed too long for neutron star pulsations (and too short for those of white dwarfs). Models involving orbiting white dwarfs, even in contact, could not produce periods short enough to match the data; the period for a satellite at the surface of an object with uniform density ρ can easily be calculated as

$$T = \left(\frac{3\pi}{\rho G}\right)^{\frac{1}{2}}. \tag{2.44}$$

Using a typical value of $\rho \sim 10^6$ gm cm^{-3} we see that periods no shorter than $\sim$ 10s can be accommodated by such a model. Orbiting neutron stars, on the other hand, would lose energy by gravitational radiation resulting in the rapid

decay of the orbit. Rotation of neutron stars was soon the model of choice and calculations by Gold (1968) indicated that pulse periods should show a secular decrease (which was soon confirmed with measurements made on the emission from the Crab pulsar). While other models still exist, it is the rotating neutron star model which has been most successful in explaining the observed features of pulsars. In order to understand the nature of the pulsars, a brief review of the models of neutron stars themselves is in order.

2.3.1 Neutron Star Structure

Shortly after Chadwick's discovery of the neutron, the suggestion was made by Baade & Zwicky that matter at greater-than-nuclear densities might be formed in the gravitational collapse of massive stars. Matter at such densities was predicted to consist primarily of neutrons. The first models of such neutron stars were calculated by Oppenheimer & Volkoff (1939). The matter was assumed to consist of a cold Fermi gas (note that, using Equation 3.9, the Fermi energy for heavy particles at nuclear densities is $E_0 \sim 80$MeV so that the Fermi gas is still "cold" at a temperature of $\sim 10^{10}$K) with degeneracy pressure balancing the gravitational forces of the star. A somewhat simple treatment led to the conclusion that such structures were unstable for masses greater than $\sim 75 M_\odot$. At the time it was believed that thermal radiation from such objects would be too weak to detect because of the small radii expected.

More recent models of neutron stars predict a range of possible masses from $\sim 0.1 - 2 M_\odot$. The typical radius of such objects (as mentioned in Section 2.2) is

about 10 km with densities reaching, and possibly surpassing, nuclear densities ($\rho_0 \sim 2.8 \times 10^{14}$ gm cm^{-3}). The resulting gravitational fields on such objects are nearly 10^{11} times that on earth. Since the matter of which the star is comprised is assumed to be highly conducting, the magnetic flux is conserved in the collapse of the progenitor star; the resulting fields are immense due to the compression of the flux. Observations indicate values as large as 10^{12}G.

Along with the suggestion that pulsar behavior could be related to radial oscillations in neutron stars came a strong interest in the equation of state for matter at extremely high densities. This research has continued so that, although many questions remain, we now have a fairly clear picture of the interior structure of neutron stars. The reader is referred to the review articles by Cameron (1970), Baym and Pethick (1975), and Canuto (1974, 1975) for a more thorough treatment. The matter inside neutron stars is relatively cold due to neutrino emission very shortly after birth. The cross-section of a typical neutron star is shown in Figure 2.4. At the surface, there is a shell that is likely to be composed of ^{56}Fe. Beneath this iron surface is a solid crust consisting of a lattice of bare nuclei surrounded by a degenerate electron gas. The lattice represents the configuration for which the energy due to the repulsion of the nuclei from one another (an effect which is generally screened by electrons) is a minimum. At greater depths, electron energies increase due to the increasing gravitational compression until they become relativistic. At this point, inverse β decay occurs thus combining electrons with protons (in the nuclei) to increase the neutron population. This process results in the formation of heavy, neutron-rich nuclei

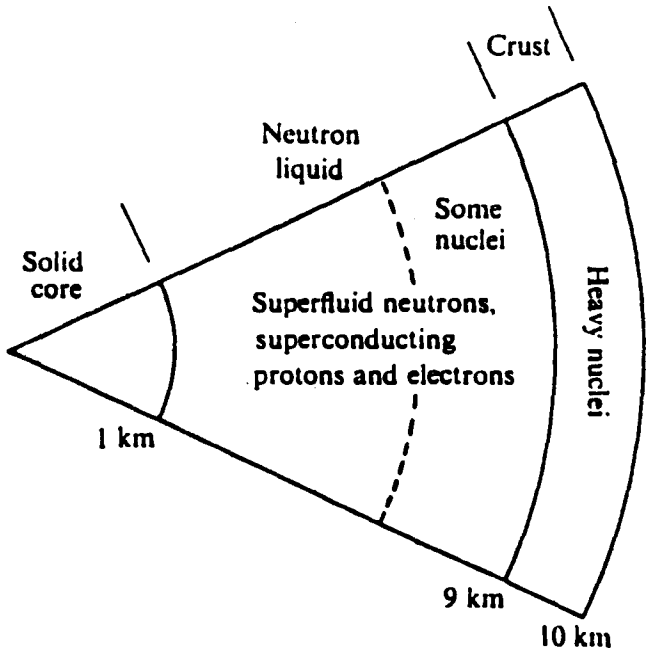


Figure 2.4: Cross-section of a typical neutron star.

including some species which are unstable in the laboratory (Baym *et al.*, 1971). At a density of $\rho \approx 4.3 \times 10^{11}$ gm cm^{-3}, the "neutron drip" point, the nuclei have become so neutron rich that further electron capture produces unbound neutrons. Thus the lattice becomes immersed in a sea of neutrons. As densities reach those of nuclear matter, the nuclei are nearly touching and the matter is composed of a liquid of neutrons, protons, electrons. When the Fermi level of the electrons exceeds the rest mass of a muon, it becomes energetically favorable to create negative muons with zero kinetic energy rather than pushing electron energies above the Fermi level. Thus, at sufficiently high densities, muons are added to the list of constituent particles. Similarly, as the Fermi level of the neutrons increases, Σ^- and then Λ^0 hyperons are created. At the great pressures

found in the deep interior, the final structure is still somewhat in question. It is possible that interactions between neutrons will form a solid neutron lattice although details concerning such interactions at very high energies, as well as those between hyperons, are still not well understood.

2.3.2 The Pulsar Magnetosphere

The physics involved with the exterior regions of a rotating neutron star are no less spectacular or puzzling than for the interior. Because of the large magnetic fields found in such objects, acting in concert with the rapid rotation, the opportunity for many types of energy loss exist. Since the neutron star cannot undergo further gravitational collapse, the energy released must be at the expense of rotational kinetic energy (except in the case of "glitches" whereby a resettling of the crust releases gravitational binding energy). As a result, energy loss is reflected in the secular change of the rotation period.

Suppose the neutron star possesses a magnetic dipole moment $m_\perp$ which rotates about an axis perpendicular to the dipole with an angular speed Ω. The power radiated is then

$$\dot{E} = \frac{\mu_0}{4\pi}\left(\frac{2m_\perp^2\Omega^4}{3c^3}\right). \tag{2.45}$$

Setting this equal to the loss of rotational kinetic energy

$$\dot{E} = \frac{d}{dt}\left(\frac{1}{2}I\Omega^2\right) \tag{2.46}$$

(where I is the moment of inertia of the neutron star) yields

$$m_\perp^2 = \frac{3c^3 I}{2\pi\mu_0} P\dot{P} \tag{2.47}$$

where we have used

$$P = \frac{2\pi}{\Omega}. \tag{2.48}$$

Using typical neutron star parameters ($m \approx 1.4 M_\odot$, $r \approx 10^4$m), and noting that

$$B = \frac{\mu_0}{4\pi}\frac{m_\perp}{R^3}, \tag{2.49}$$

we have

$$B = \left[\frac{3\mu_0 c^3}{80\pi^3}\frac{M}{R^4} P\dot{P}\right]^{\frac{1}{2}} \approx 3.4 \times 10^{19} (P\dot{P})^{\frac{1}{2}} \text{ Gauss.} \tag{2.50}$$

Measurements of P and $\dot{P}$ can thus provide information about the strength of the surface magnetic field. Of course, if other energy loss mechanisms are at work as well, such a calculation will overestimate the field strength. Further, since only the component of the magnetic moment which is normal to the rotation axis will radiate, the orientation plays a major role. Nonetheless, estimates based on such measurements suggest fields as large as $\sim 10^{12}$ G, as mentioned in the previous section.

In analyzing the electrodynamics of a rotating axisymmetric magnetic dipole, Goldreich and Julian (1969) noted that the presence of such large, rapidly rotating magnetic fields results in large electric fields as well. Because the stellar material has extremely high conductivity, the net force on any charges must vanish lest infinite currents be established. Thus,

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B}. \tag{2.51}$$

Using

$$\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{r} = r\Omega \sin\theta \hat{\phi}$$

at the surface, and

$$\mathbf{B} = \frac{\mu_0}{4\pi}\frac{m_\perp}{r^3}[2\hat{\mathbf{r}}\cos\theta + \hat{\theta}\sin\theta]$$

for the dipole magnetic field, we have

$$\mathbf{E} = \frac{\mu_0}{4\pi}\frac{m_\perp}{r^3} r\Omega[\hat{\mathbf{r}}\sin^2\theta - 2\hat{\theta}\sin\theta\cos\theta].$$

Thus, at the equator, the radial electric field is

$$E = \frac{\mu_0}{4\pi}\frac{m_\perp}{R^3} R\Omega \approx 6 \times 10^{13}\ \mathrm{V\ m^{-1}}$$

using typical values for a rotating neutron star. In comparison, the surface gravitational field is of the order

$$g = \frac{GM}{r^2} \approx 2 \times 10^{12}\ \mathrm{m\ s^{-2}}.$$

Thus, the ratio of electrostatic to gravitational force is about 10^{12} for electrons and about 10^9 for protons. As a result, particles will be removed from the surface of the star and a plasma-filled magnetosphere will be formed.

The charge density in the magnetosphere may be found using Maxwell's equations:

$$\frac{\rho}{\epsilon_0} = \nabla\cdot\mathbf{E} = -\nabla\cdot(\mathbf{v}\times\mathbf{B}). \tag{2.52}$$

Now

$$\nabla\cdot(\mathbf{v}\times\mathbf{B}) = \mathbf{B}\cdot(\nabla\times\mathbf{v}) - \mathbf{v}\cdot(\nabla\times\mathbf{B}) \tag{2.53}$$

where the second term vanishes in the absence of currents. But, since $\mathbf{v} = \boldsymbol{\Omega}\times\mathbf{r}$, we have

$$\nabla\times\mathbf{v} = -(\boldsymbol{\Omega}\cdot\nabla)\mathbf{r} + (\mathbf{r}\cdot\nabla)\boldsymbol{\Omega} - (\nabla\cdot\boldsymbol{\Omega})\mathbf{r} + (\nabla\cdot\mathbf{r})\boldsymbol{\Omega} = 2\boldsymbol{\Omega}. \tag{2.54}$$

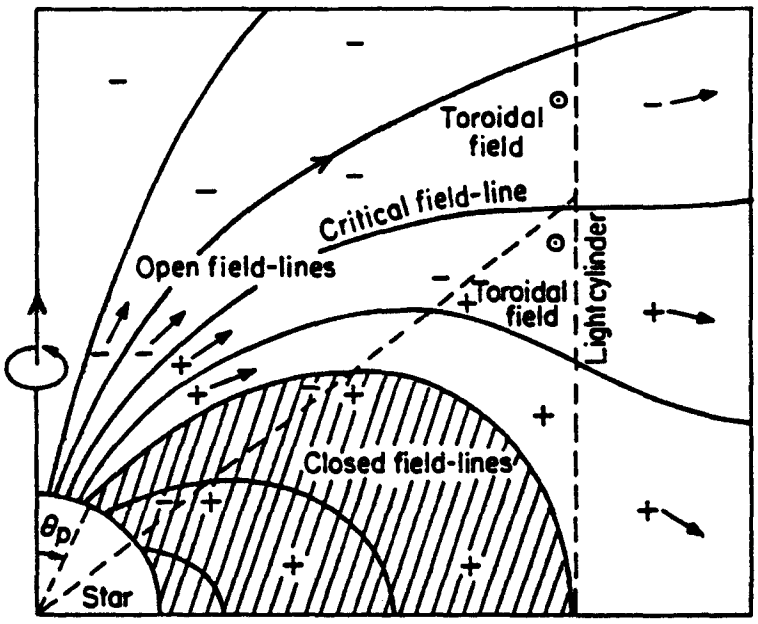


Figure 2.5: Schematic of Goldreich-Julian pulsar magnetosphere.

Thus,

$$\rho = (n_+ - n_-)e = -2\epsilon_0 \mathbf{B} \cdot \mathbf{\Omega} \tag{2.55}$$

or

$$n_- - n_+ = \frac{2\epsilon_0}{e} \mathbf{B} \cdot \mathbf{\Omega}. \tag{2.56}$$

This distribution of charges is locked to the magnetic field and, thus, corotates with the star out to a radius where the velocity becomes equal to c (the so-called "light cylinder"). The Goldreich-Julian model of the magnetosphere is shown in Figure 2.5. Two distinct regions are of particular interest, namely that in which the field lines close within the light cylinder and that which consists of open field lines. In the open-line region, charges can escape.

For a dipolar field, the field lines are described by

$$\frac{\sin^2\theta}{r} = \text{constant}. \tag{2.57}$$

Thus, for the last closed field line,

$$\frac{\sin^2\theta_p}{R} = \frac{\Omega \sin^2(\frac{\pi}{2})}{c} \tag{2.58}$$

where θ_p is the polar angle. Hence,

$$\sin\theta_p = \left(\frac{\Omega R}{c}\right)^{\frac{1}{2}}. \tag{2.59}$$

The locus of points for which $\mathbf{B}\cdot\mathbf{\Omega} = 0$ (shown as a dashed line in the figure) defines the separation of positive and negative charge density.

Solving Laplace's equation for the rotating dipole yields the electrostatic potential

$$\Phi = -\frac{B\Omega R^5}{6r^3}(3\cos^2\theta - 1). \tag{2.60}$$

Except for very near the light cylinder, we have $\mathbf{E}\cdot\mathbf{B} \approx 0$ along the field lines. Hence, they form equipotentials. The field line for which $\Phi = 0$ is defined by

$$\cos\theta_0 = \frac{1}{\sqrt{3}}. \tag{2.61}$$

For $\theta < \theta_0$, electrons can be accelerated toward the light cylinder while for $\theta_0 < \theta < \theta_p$ protons are pushed outward. The total potential difference across the polar cap is approximately

$$\triangle\Phi = \frac{B\Omega R^2}{2}\sin^2\theta_p = \frac{B\Omega^2 R^3}{2c}. \tag{2.62}$$

Using $B = 10^{12}$G and $R = 10^4$m, we may write this as

$$\triangle\Phi \approx 6\times 10^{12} P^{-2} \text{ volts} \tag{2.63}$$

where P is the rotation period in seconds. Hence, potential differences of 10^{16} volts are possible for rotation periods of $\sim 10^{-2}$s.

While the Goldreich-Julian model for pulsar magnetospheres provides a great deal of insight into the conditions surrounding rotating magnetic neutron stars, it is by no means a complete description. For example, the assumption that the rotation axis and the dipole axis are aligned, while a useful simplification, is not likely to be accurate. Further, as mentioned previously, the assumption that $\mathbf{E} \cdot \mathbf{B} = 0$ along the field lines is only approximate. In section 2.5 we will discuss results which suggest that vacuum gaps may form in certain regions of the magnetosphere. For a comprehensive review of pulsar magnetospheres, the reader is referred to Michel (1982).

2.4 X-Ray Binary Objects

A major breakthrough in the quest for detailed information about the properties of compact objects came in the form of the discovery of x-ray binaries. These objects consist of a compact star in gravitational association with a "normal" companion star. The systems span a vast range of orbital parameters, periods, and luminosities. The emission from those which contain neutron stars often exhibit pulsations believed to be associated with the rotation period of the stars. In such cases, orbital information can be ascertained just as in the case of spectroscopic binaries. Observations of the companion stars have provided estimates of the neutron star masses which are consistent with theoretical predictions. Studies of the temporal behavior of the pulse periods have placed constraints on

the accretion mechanisms believed to be acting to provide the x-ray luminosity and have provided information regarding the internal structure of neutron stars. Direct measurements of the surface magnetic fields of the neutron stars through x-ray cyclotron lines show that these stars retain considerable magnetic flux and provide information for models attempting to explain acceleration of high energy particles in the vicinity of these objects. In order to understand the energy production of these objects, we will review the important subject of accretion. First, however, a brief look at the properties of three distinct x-ray binary objects will illustrate the similarities and vast differences that have been observed from these objects.

2.4.1 Examples

While there are more than fifty x-ray binary sources known at present, only a handful have shown evidence, to date, of producing VHE γ-rays. Three of these sources — Cygnus X-3, 4U0115+63, and Her X-1 — display so many similarities, and also differences, that a brief discussion of their properties will illustrate the range of phenomena detected from such binary sources.

Cyg X-3 is one of the most luminous x-ray objects known. Its emissions are not confined to this band of the spectrum, however. Radio emission punctuated by outbursts in which the luminosity soars some three orders of magnitude above its quiescent level is observed. Infrared emission displays a 4.8 hour periodicity that is also seen is x-rays. This periodicity is presumed to be associated with the orbit of the system. Obscured by large quantities of galactic dust, it reveals

no optical information. In the γ-ray regime, emissions have been detected at energies ranging from several hundred TeV up to 10 PeV. Indeed the study of Cygnus X-3 is enough to fill volumes by itself. A claim of detection of 12.59 ms periodicity presumed to be associated with the rotation of a neutron star has yet to receive verification. If true, however, future observations may be able to establish much better system parameters. At present, even a rough estimate of the companion star mass does not exist.

The system designated 4U0115+63 consists of a neutron star orbiting its companion star with a period of 24.31 days. This is in stark contrast to the 4.8 hour period for Cyg X-3. The period of the "pulsar" in this system (which is not a pulsar at all, in the conventional sense, in that radio pulsations are not seen — this more the rule than the exception for such binary sources) is 3.61 s. Pulse timing of the x-ray emissions (which, unlike those from Cyg X-3 and Her X-1, occur sporadically) yields information suggesting a highly eccentric orbit ($e \approx 0.34$) whose projected semimajor axis is approximately 140 l-s. In addition, a secular decrease in the pulsar period is observed, $\dot{P} \approx -3 \times 10^{-11}$, corresponding to a spin-up that is expected to arise from the transfer of angular momentum from accreted matter. The companion star is this system, which has been optically identified, is believed to have a mass $> 5 M_{\odot}$.

The properties of Her X-1, the subject of this paper, are investigated in detail in Chapter 5. To summarize, this system has been studied at energies ranging from the infrared to the UHE γ-ray regions. The pulsar period of 1.24s is well established both in x-rays and γ-rays associated with the neutron star and in

optical pulses resulting from reprocessing of x-rays in the companion star HZ Her. Doppler delays in the pulse arrival times establish a circular orbit of radius 13.2 l-s whose period is 1.7 days. Combined with data obtained from HZ Her this suggests a companion star mass of $\sim 2.2 M_{\odot}$. An additional periodicity of $\sim$ 35 days is observed in x-rays. Correlations of this cycle with optical and VHE γ-ray emissions are also suggested. The origin of the 35 day periodicity is not known but it is believed to be associated with precession of an accretion disk, which is known to exist, or of the rotation axis of the neutron star.

We see from these examples that x-ray binary objects involve a broad range of physics. It is worth noting that claims (see *e.g.* Marshak *et al. 1985*) of detection of underground muons associated with these three objects, along with several others have be made. Whether or not such detections turn out to be correct, it is clear that these objects invite and require much further study.

2.4.2 Accretion

The physics of compact astrophysical objects is strongly determined by the gravitational environment in their immediate vicinity. While, by definition, the gravitational fields surrounding such objects are extremely large, the presence of matter which may be accelerated in the field is crucial for the conversion of gravitational energy into radiation. In particular, matter which undergoes accretion onto a compact object can produce detectable radiation of energies ranging from the infrared to the soft γ-ray regions.

Energy Considerations

We begin by considering spherical accretion onto an object of mass M and radius R. The energy-release available from accretion of matter of mass m is

$$\triangle E_{acc} = \frac{GMm}{R}. \tag{2.64}$$

For accretion onto a neutron star, for which $M \approx 1.4 M_{\odot}$ and $R \approx 10$ km, this results in a release of $\sim 10^{20}$ erg gm^{-1} or about 20% of the rest mass energy of the accreted matter. The corresponding luminosity due to such accretion is

$$L_{acc} = \frac{GM\dot{m}}{R} \tag{2.65}$$

where $\dot{m}$ is the accretion rate in gm s^{-1}. This results in a radiant energy flux

$$S = \frac{L_{acc}}{4\pi r^2}. \tag{2.66}$$

This flux will exert radiative forces on the particles in the accreting plasma. We may consider this plasma to consist primarily of protons and electrons. The force a given electron will feel as a result of the radiative flux is

$$F_{rad} = \sigma_T \frac{S}{c} \tag{2.67}$$

where

$$\sigma_T = \frac{8\pi}{3}\left(\frac{e^2}{m_e c^2}\right)^2 \simeq 6.7 \times 10^{-25} \text{ cm}^2 \tag{2.68}$$

is the Thomson cross-section. The radiative force on the protons is negligible by comparison because the cross-section is smaller by the factor $(m_e/m_p)^2$. Due to electrostatic forces, however, the protons are pulled outward along with the

electrons. The net force on such electron-proton pairs is thus

$$F \simeq \left(GMm_p - \frac{\sigma_T L}{4\pi c} \right) \frac{1}{r^2} \tag{2.69}$$

where m_p is the proton mass. This force vanishes at the so-called *Eddington luminosity*

$$L_E = \frac{4\pi c G M m_p}{\sigma_T} \approx 1.3 \times 10^{38} \left(\frac{M}{M_\odot} \right) \text{ erg s}^{-1}. \tag{2.70}$$

This luminosity, then, represents an upper limit beyond which radiation pressure overcomes the gravitational forces on the accreting matter. Such an upper limit imposes a constraint on the rate at which mass may be accreted. Setting $L_E \leq L_{acc}$ yields

$$\dot{m} \leq 9.7 \times 10^{11} R \text{ gm s}^{-1} = 1.5 \times 10^{-14} \text{ R } M_\odot \text{ yr}^{-1}. \tag{2.71}$$

For white dwarfs, this limits the accretion rate to $\sim 10^{21}$ gm s^{-1} while for neutron stars the rate must be less than $\sim 10^{18}$ gm s^{-1}. For Her X-1, the observed flux in ~ 5 keV x-rays is $\sim 1.8 \times 10^{-9}$ erg cm^{-2} s^{-1}. The source distance is $\sim$5 kpc yielding a luminosity of $\sim 5 \times 10^{36}$ erg s^{-1}, nearly the Eddington luminosity.

The radiation emitted as a result of the accretion process can be characterized by an effective temperature T_{rad}. For large optical thickness, the radiation is characterized by a black-body temperature

$$T_{bb} = \left(\frac{L_{acc}}{4\pi R^2 \sigma} \right)^{\frac{1}{4}} \tag{2.72}$$

while for small optical thickness, the thermal temperature

$$T_{th} = \frac{GMm_p}{3kR}. \tag{2.73}$$

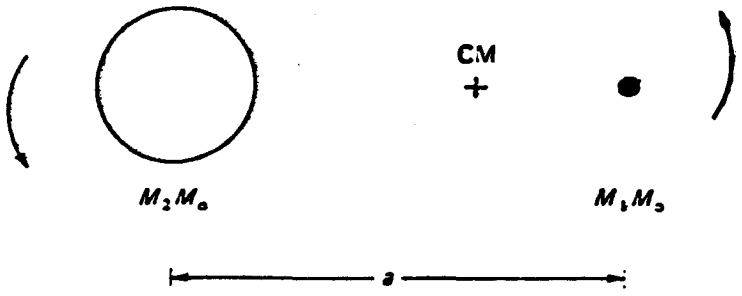


Figure 2.6: Schematic of binary star system.

is a more reasonable approximation. Thus, $T_{bb} \leq T_{rad} \leq T_{th}$ or, for accretion onto a neutron star, 1keV $\leq h\nu \leq$ 50MeV. It is reasonable, then, to expect the production of x-rays and low energy γ-rays. Lower energy photons resulting from secondary processes may also be produced.

Accretion in Binary Systems

Accretion is a particularly important subject when the characteristics of binary systems are considered. Consider a system consisting of stars of masses M_1 and M_2 separated by a distance a (Figure 2.4). The period of rotation is given by Kepler's third law:

$$4\pi^2 a^3 = G(M_1 + M_2)P^2. \tag{2.74}$$

The Roche potential for the system is given by

$$\Phi_R(\mathbf{r}) = -\frac{GM_1}{|\mathbf{r} - \mathbf{r}_1|} - \frac{GM_1}{|\mathbf{r} - \mathbf{r}_2|} - \frac{1}{2}(\omega \times \mathbf{r})^2 \tag{2.75}$$

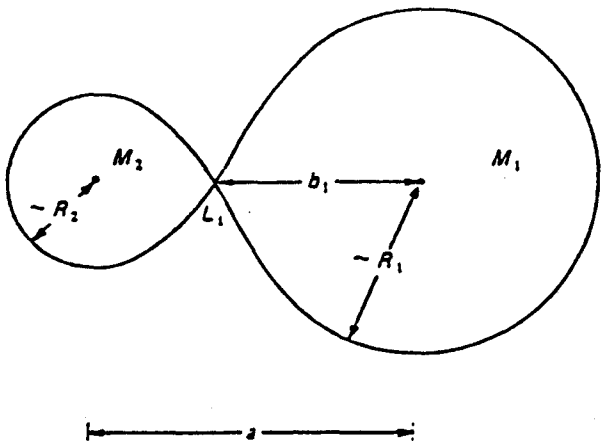


Figure 2.7: Roche lobe geometry.

where $\mathbf{r_1}$ and $\mathbf{r_2}$ are the position vectors of the two stars. The equipotential surface which contains the first Lagrangian point L_1 consists of two lobes as shown in Figure 2.5. The shape of the equipotentials is determined by the ratio

$$q = \frac{M_2}{M_1}. \tag{2.76}$$

As an approximation, each lobe can be represented by a sphere which contains the same volume as the lobe. The radii of such spheres can be approximated by the expressions (Paczyński 1971)

$$\frac{R_1}{a} = \begin{cases} 0.38 - 0.20 \log q, & 0.05 < q < 2 \\ 0.462(1+q)^{-\frac{1}{3}}, & 2 < q \end{cases} \tag{2.77}$$

and

$$\frac{R_2}{a} = \begin{cases} 0.38 + 0.20 \log q, & 0.5 < q < 20 \\ 0.462(1+q^{-1})^{-\frac{1}{3}}, & 0 < q < 0.5. \end{cases} \tag{2.78}$$

In addition, the distance from the primary star to L_1 is approximately

$$\frac{b_1}{a} \simeq 0.500 - 0.227 \log q. \tag{2.79}$$

As an example, the mass ratio for the Her X-1/HZ Her system is $q \simeq 1.68$ (Middleditch and Nelson 1976) so that $R_2 \approx 0.425a$ and $b_1 \approx 0.449a$. Using the orbital parameters of the system (see Chapter 5) we find $R_2 \approx 2.7 \times 10^6$ km or about four times that of the sun. As noted in Section 2.2, early B-type stars as well as evolved red giants can reach sizes considerably larger than this. Hence we see that it is quite possible for a star to fill its Roche lobe (as Middleditch and Nelson have shown is the case for HZ Her). Clearly, in such an instance, mass can easily be transferred from a star to its companion. However, because q changes as such mass transfer occurs, one must consider whether the result of the transfer will cause R_1 to decrease, in which case the process will continue, or to decrease, in which case the process will be quenched. The value $q = 1$ is the dividing line between the two scenarios; for $q > 1$, the process is self-perpetuating while for $q < 1$ the star must expand (as part of its evolutionary cycle, for example) in order for the transfer process to be maintained.

Even if the companion star does not fill its Roche lobe, accretion onto the compact star is possible via a stellar wind. Sufficiently massive upper main sequence stars may lose mass at a rate of $10^{-6} M_\odot$ yr^{-1} through such winds. Some fraction of this mass may be captured by the primary star. The captured gas will carry some angular momentum, but the accretion is likely to be more or less spherical; Roche lobe overflow, on the other hand, is likely to result in the formation of a disk of accreted material somewhat akin to the rings of matter which surround Saturn and other planets.

Accretion Disks

Let us consider the case of accretion through L_1 onto a compact star. The accretion appears, in a non-rotating reference frame at the compact star, to consist of streaming matter with velocities

$$\begin{aligned} v_\perp &\sim b_1\omega \\ v_\parallel &\sim c_s \end{aligned} \tag{2.80}$$

where ω is the angular velocity of rotation and c_s is the speed of sound in the accreting matter. For normal stellar surface temperatures ($< 10^5$K), $v_\parallel \sim$ 10 km s^{-1}.

Because circular orbits have the minimum energy for a given momentum, the accreted matter will tend to fall into circular Keplerian orbits whose radii R_c correspond to the specific angular momentum that the matter had in passing through L_1, *i.e.*

$$v_c = \left(\frac{GM_1}{R_c}\right) \tag{2.81}$$

subject to the constraint

$$R_c v_c = b_1 v_\perp = b_1^2 \omega. \tag{2.82}$$

Using Kepler's third law, this reduces to

$$R_c = \frac{b_1^4}{a^3}(1+q). \tag{2.83}$$

For Her X-1, then, $R_c \approx 0.24 b_1$; *i.e.* circularization occurs well inside the Roche lobe (although well outside of the light cylinder of the neutron star - see Section 2.3.2). This picture indicates a ring of matter. However, collisions will result in dissipation of energy (some of which is radiated). To meet this energy loss,

the matter will be pulled further into the gravitational well of the compact star. Such diminished orbits require loss of angular momentum. This is accomplished by the transfer of angular momentum to the outer regions of the ring. Thus, the inner portion of the ring is drawn in toward the compact star while the outer region is pushed out; the result is an *accretion disk.*

While a complete description of the structure of accretion disks is beyond the scope of this chapter, a summary of the main characteristics will be useful in further understanding the environment in which VHE γ-rays appear to be produced. The matter being accreted onto the compact star has virtually zero gravitational binding energy upon passing through the inner Lagrangian point. When circularization and viscosity effects have brought the matter near the surface of the star, it will have binding energy

$$E_{grav} = \frac{GM_1 m}{2r_1}. \tag{2.84}$$

Hence, the luminosity of the disk is

$$L_{disk} = \frac{GM_1 \dot{m}}{2r_1} = \frac{1}{2} L_{acc}. \tag{2.85}$$

The matter which comprises the accretion disk is characterized by highly subsonic radial velocities ($v_r \sim 3$ km s^{-1}) and supersonic tangential velocities ($v_\phi \sim 10^3$ km s^{-1}). The thickness of the disk is approximately given by

$$\frac{H}{R} \simeq \frac{c_s}{v_k(R)} \tag{2.86}$$

where $v_k(R)$ is the speed of Keplerian orbit at a distance R. The disk itself is thin in the z direction but the optical thickness is large, resulting in essentially blackbody behavior. The self-gravity of the disk is negligible in comparison with the

gravitational influence of the compact star. As a result, the disk is stable against breaking up into self-gravitating clumps of matter. The process of accretion onto the surface of the star is complicated by considerations of the magnetic field of the star. In the absence of a significant field, the matter falls into ever smaller Keplerian orbits as it approaches the star. However, the *corotation radius* of the system (*i.e.* the radius at which the period of the Keplerian orbit matches the rotation period of the compact star) is given by

$$r_{cor} = \left(\frac{GM_1}{4\pi^2}\right)^{\frac{1}{3}} P^{\frac{2}{3}} \tag{2.87}$$

which, for a neutron star with a period of $1s$, is about 10^6m. Inside of this radius, the matter will move faster. Hence, there must exist a region near the surface of the star where the angular velocity of the matter decreases in order to match that of the star (Figure 2.6). The matter in this region will form an optically thick boundary layer which will radiate as a black-body.

The weak field scenario described above may be a reasonable description of accretion onto a white dwarf (if the field is weak enough), but for a neutron star as the compact object the magnetic field contributions are quite likely to be much too large for this simple picture. We may define the Alfvén radius as that distance from the neutron star at which the magnetic pressure exceeds the gas/ram pressure of the matter in the accretion disk. The magnetic pressure is given by

$$P_{mag} = \frac{B^2}{2\mu_0} \tag{2.88}$$

while the disk pressure is given by

$$P_{disk} = \rho v^2. \tag{2.89}$$

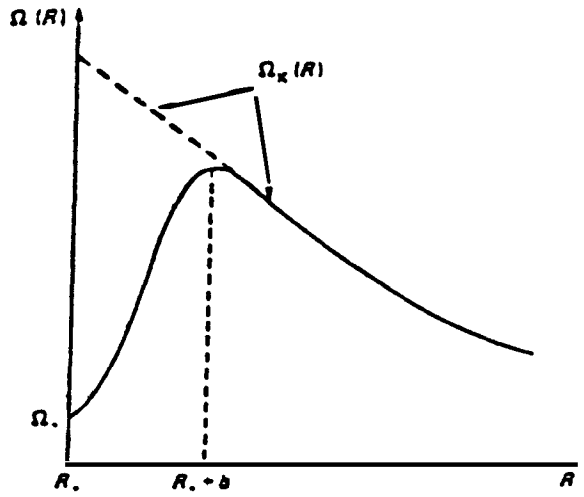


Figure 2.8: Angular velocity of accreting matter as a function of distance from the stellar surface. The region $R_1 < R < R_1 + b$ forms an optically thick boundary layer from which radiation is emitted.

We may approximate the disk pressure by considering a spherical mass of matter:

$$dm = \rho(4\pi r^2\, dr) \Rightarrow \dot{m} = 4\pi r^2 \rho v. \tag{2.90}$$

Solving for ρ and using

$$v = v_{ff} = \left(\frac{2GM}{r_m}\right)^{\frac{1}{2}} \tag{2.91}$$

and

$$B \simeq \frac{\mu}{r^3} \tag{2.92}$$

where μ is the magnetic moment of the neutron star, we obtain

$$r_m = \left(\frac{2\pi^2}{\mu_0^2}\right)^{\frac{1}{7}} M^{-\frac{1}{7}} \dot{m}^{-\frac{2}{7}} \mu^{\frac{4}{7}}. \tag{2.93}$$

Using Equation 2.46, then, we may write the Alfvén radius in terms of the luminosity:

$$r_m = \left(\frac{2\pi^2 G^2}{\mu_0^2}\right)^{\frac{1}{7}} M^{\frac{1}{7}} \mu^{\frac{4}{7}} L^{-\frac{2}{7}} R^{-\frac{2}{7}}. \tag{2.94}$$

For Her X-1, for which the luminosity in 2-6 keV x-rays is $\approx 5 \times 10^{36}$ erg s^{-1} and the magnetic field strength is $\approx 10^{12}$G, this corresponds to a radius of about 10^5m. More complicated calculations which treat the characteristics of the disk show that the effective radius is actually smaller than this by about a factor of two. Inside this radius, matter is forced to rotate with the field lines and, thus, with the same angular velocity as the rotating neutron star. Hence, as matter accretes toward r_m it is forced to reduce its angular velocity in order to fall into corotation. The result is a net transfer of angular momentum to the neutron star causing its period to decrease (assuming that the angular momentum of the accreting matter is in the direction of the neutron star rotation). On the other hand, viscous torques resulting from the shearing of the magnetic field lines in the vicinity of r_m will tend to oppose such a spin-up of the pulsar. Assuming that these torques are small, we can write the rate at which the angular momentum interior to r_m changes as

$$\frac{d(I\Omega)}{dt} = \dot{M} l(r_m) \tag{2.95}$$

where $l(r_m)$ is the specific angular momentum of accreting matter at r_m and Ω is the angular velocity of corotation. Using

$$\dot{I} = \frac{dI}{dM}\dot{M} \text{ and } \dot{\Omega} = -\frac{\dot{P}}{P}\Omega, \tag{2.96}$$

we find

$$\frac{\dot{P}}{P} = \frac{\dot{M}}{M}\left(\frac{M}{I}\frac{dI}{dM} - \frac{Ml(r_m)}{I\Omega}\right). \tag{2.97}$$

In general, the second term in this expression dominates. Noting that

$$l(r_m) = (GMr_M)^{\frac{1}{2}} \text{ and } L = \frac{GM\dot{M}}{R} \tag{2.98}$$

we can arrive at the relation

$$\dot{P} = -\left[\frac{\mu^2 R^6}{16G^6 m^3 I^7}\right]^{\frac{1}{7}}\left(PL^{\frac{3}{7}}\right)^2. \tag{2.99}$$

For compact stars of a given size and magnetic moment, then, the value of $\dot{P}$ is dependent only upon the period and luminosity.

As noted above, this expression neglects the effects of magnetic torques. Such torques can be important for systems with large values of the "fastness parameter"

$$\omega_s \equiv \frac{\Omega}{\Omega_K(r_m)} \tag{2.100}$$

where $\Omega_K(r_m)$ is the Keplerian velocity at r_m. For $\omega_s \simeq 1$, the magnetic torques oppose the spin-up due to the accreting material. As a result, during periods of low accretion rate (*i.e.* low luminosity) the star can even undergo spin *down*. Such variations in the behavior of $\dot{P}$ have been well documented in the case of Her X-1 (see Section 5.1.2).

2.5 VHE γ-Ray Production Mechanisms

Very high energy γ-rays are produced via interactions involving extremely energetic particles. The origins of such γ-rays, then, are related to the mechanisms by

which particles are accelerated to very high energies. We have seen, in Section 2.3.2, that electric fields in the vicinity of pulsars can be of sufficient magnitude to accelerate particles to energies $\sim 10^{16}$ eV (using the model of Goldreich and Julian, 1969). Gunn and Ostriker (1969) proposed a model in which the magnetic moment of the pulsar is assumed to be perpendicular to the axis of rotation. The resulting electric field associated with the dipole radiation can accelerate particles in much the same manner as that used in conventional man-made accelerators. The particles ride the wavefront and are accelerated in the direction of the propagation of the wave.

Ruderman and Sutherland (1975) extended the model of Goldreich and Julian by considering the case in which the magnetic moment of a spinning pulsar is antiparallel to the angular momentum vector. They conclude that ions in the polar regions will remain bound to the neutron star resulting in a net positive charge which will persist because of the inability of electrons to be drawn in through the light cylinder to discharge the region. The result is the formation of a charge-depleted polar gap region. The potential difference associated with the gap can be as large as $\sim 10^{12}$ volts. Such gaps are continually discharged by the formation of electron-positron pairs. These particles are accelerated in the gap and produce γ-rays through curvature radiation, synchrotron radiation, or inverse-Compton scattering on soft photons in the vicinity of the pulsar.

One problem with models in which γ-rays are produced near the pulsar is that absorption via pair-production in the strong magnetic field reduces the observable γ-ray flux. The "outer gap" model of Cheng *et al.* (1986) does

not suffer from such absorption problems in that acceleration occurs near the light cylinder where the magnetic field is greatly reduced. Such outer gaps are formed in the regions bounded by the "null surface" defined by $\mathbf{\Omega} \cdot \mathbf{B} = 0$ (see Equation 2.5.6), the last closed field line (refer to Figure 2.5), and the light cylinder. While charges within the closed magnetosphere cannot escape the region, nothing prevents charges from leaving the open magnetosphere through the light cylinder. Such charge flow results in a net charge in such regions which "pushes" the null surface toward the star. Again, as with polar gaps, a gap grows until discharge, and resulting acceleration, of electron-positron pairs occurs. The outer gap acceleration mechanism is most efficient for rapidly spinning pulsars such as the Crab and Vela pulsars but may fail for slowly rotating pulsars (such as Her X-1). Further, the presence of an accretion disk may modify the geometry of the field lines a such a way as to complicate the application of the model.

While we have seen that the accretion process can result in very efficient conversion of gravitational energy (Section 2.4.2), such a mechanism cannot produce TeV (or even GeV) particles or photons. However, a highly conducting accretion disk in the presence of a strong rotating magnetic field can result in large potential differences through a dynamo mechanism. Adapting a model proposed by Lovelace (1976) with reference to accretion onto massive black holes in active galactic nuclei, Chanmugam and Brecher (1985) propose a "unipolar inductor" mechanism which can result in sufficient potentials for acceleration to very high energies in accreting binary systems. They find that the potential across the disk

is given by

$$V_{12} = (GM)^{\frac{1}{2}} B_z(r_1) r_1^{\frac{1}{2}} \ln\left(\frac{r_2}{r_1}\right) \tag{2.101}$$

where r_1 and r_2 are the inner and outer radii of the accretion disk and B_z is the magnetic field at the inner edge of the disk. Using parameters associated with Her X-1, we may choose r_1 to be the Alfvén radius ($r_1 \sim 10^5$ m - see Equation 2.86) so that

$$B(r_1) = B\left(\frac{R}{r_1}\right)^3 \approx 10^5 \text{ T} \tag{2.102}$$

where B is the surface field and R the radius of the neutron star. Observations indicate that $r_2 \sim 2 \times 10^9$ m (see Chapter 5) so that $V_{12} \approx 4 \times 10^{18}$ volts.

In order to produce VHE γ-rays from such energetic particles, interaction with matter is required. It has been suggested that such "beam dump" interactions may occur in the limb of the companion star (Vestrand and Eichler, 1982) or in the rim of the accretion disk (Eichler and Vestrand, 1985). It must be noted, however, that the observations of VHE emission from Her X-1 do not appear to be correlated with orbital phase, and appear to occur when the disk is tilted toward the observer (thus making the disk rim a geometrically unlikely location for particle conversion).

While the models for acceleration in the vicinity of rotating neutron stars are many and varied, it seems clear that the basic energy requirements can easily be met. It is possible that, through observations of the sources of such radiation, model parameters will be refined and that a consistent picture may eventually emerge.

Chapter 3

The Atmospheric Čerenkov Technique

The way we look to a distant constellation
That's dying in the corner of the sky
These are the days of miracle and wonder...
Paul Simon

The detection of radiation in any band of the electromagnetic spectrum consists of measuring the secondary effects of photon interactions: radio photons produce time-varying potentials in antennas; optical photons liberate electrons from photocathodes of photomultiplier tubes or from silver- bromide atoms in photographic emulsions; and x-ray photons interact with target gasses within proportional counters to produce energetic electrons which go on to ionize other atoms in the detector. As photon energies increase, it becomes possible, in the presence of some medium, to pair-produce secondary particles. These particles then lose energy through a variety of mechanisms which will be discussed in this chapter. Detection of such high energy photons (γ- rays), then, consists of detecting the secondary particles. In order to determine the energy of the photons, it is necessary to investigate the entire energy-loss process of the secondaries.

For extraterrestrial γ-rays with energies above $\sim$ 1 GeV, a suitable interaction medium for such a process is the Earth's atmosphere. In this chapter, the technique by which the atmosphere is used as a Very High Energy (VHE) γ-ray detection medium is discussed.

3.1 Extensive Air Showers

If a γ-ray of energy $> 2m_e c^2$ enters the Earth's atmosphere, it is kinematically possible to form an electron-positron pair. For VHE γ-rays, such e^+e^- pairs will carry virtually all of the incident γ-ray kinetic energy. Various energy-loss mechanisms are then available to the electron:

- bremsstrahlung - radiation produced when a charged particle is accelerated (*e.g.* in the electric field of another charged particle). In the relativistic case, these losses are of the form

$$-\frac{dE}{dx} = \frac{E}{x_0} \tag{3.1}$$

  where x_0, the *radiation length*, is the path length through which the electron losses all but a fraction e^{-1} of its energy. In air, $x_0 \simeq 36.5$ gm cm^{-2}. On the average, a very high energy electron will deposit about half of its energy into one or two photons early in its travel through the atmosphere.

- ionization - process by which the electron transfers energy via the dislodging of an atomic electron. While bremsstrahlung losses dominate at higher energies, below a critical energy $E_c \sim 83$ MeV ionization losses begin to

dominate.

- multiple scattering - predominately Coulomb scattering of the electrons from the nuclei of surrounding matter. These losses are actually negligible at high energies. However, the multiple scattering results in deviations in the direction of the electron velocity. The angular distribution of the scattered electrons is governed by (Rossi)

$$\langle \theta^2 \rangle = \left(\frac{E_s}{\beta pc} \right)^2 n \tag{3.2}$$

where

$$E_s = m_e c^2 \sqrt{\frac{4\pi}{\alpha}} \approx 21.2 \text{ MeV},$$

θ is the scattering angle, and n is the number of radiation lengths traversed (here α is the fine structure constant).

At very high energies, then, we expect bremsstrahlung losses to dominate, thus producing additional γ-rays. As long as the γ-rays are above threshold, each will again pair-produce; the cross-sections for the two processes are approximately equal at very high energies.

3.1.1 Cascade Theory

Considering the above discussion, it is easy to give a qualitative description of the result of a high energy electron or γ-ray impinging on the atmosphere. A high energy photon (which, in the case of a primary electron, results from bremsstrahlung) pair-produces an electron and a positron each of which radiate high energy photons by bremsstrahlung. These newly created photons repeat

the pair-production process and the entire sequence unfolds to form a cascade of energetic particles called an Extensive Air Shower (EAS). This cascade (Figure 3.1) develops until the mean particle energy becomes too low for pair-production. After this, Compton scattering and ionization losses dissipate the shower.

Although a complete description of the shower development entails detailed calculations incorporating cross-sections for all relevant interaction processes, we can (after Longair 1981) give an illustrative discussion which demonstrates the salient features. In the ultra-relativistic case, the interaction length for pair-production is very nearly one radiation length. Thus, we can define a mean distance R for which the interaction probability for either pair-production (in the case of photons) or bremsstrahlung (in the case of electrons) is 1/2, *i.e.*

$$e^{-\frac{R}{x_0}} = \frac{1}{2}. \tag{3.3}$$

Further, assuming that the energy is roughly equally shared at each interaction, we see that a γ-ray of incident energy E_0 will, after travelling a distance R, produce an electron-positron pair each carrying energy $E_0/2$. Each of these will radiate a γ-ray of energy $E_0/4$ after travelling an additional distance R (Figure 3.1). Hence, at a distance nR, the total number of shower particles is 2^n and the average energy per particle is $E_0/2^n$; one-third of the particles are photons. The shower, as noted earlier, develops until the average energy per particle drops below the critical energy after which ionization losses dominate. From our approximate treatment we see that, at this point,

$$\frac{E_0}{2^n} = E_c \Rightarrow N = 2^n = \frac{E_0}{E_c} \tag{3.4}$$

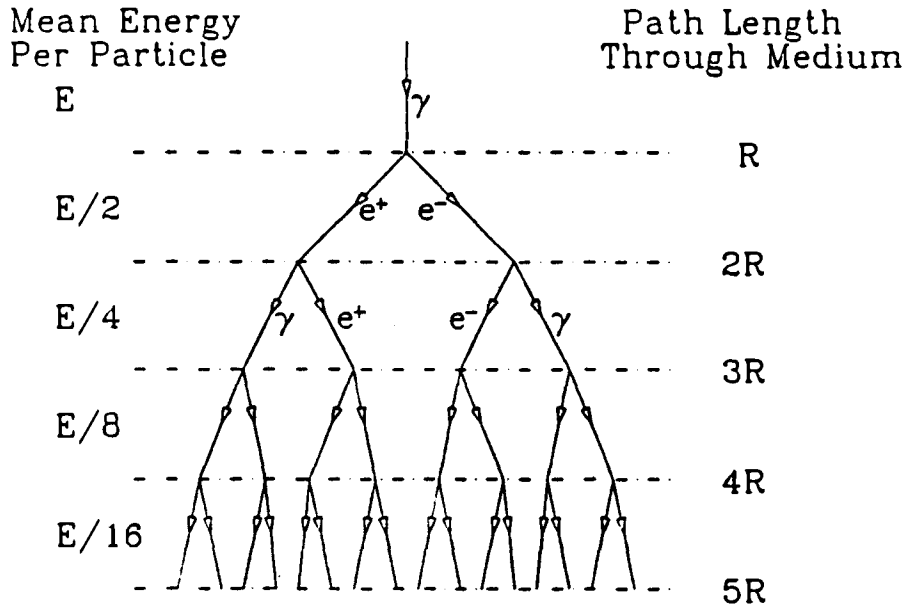


Figure 3.1: Simple geometrical development of EAS.

gives the total number of particles while

$$n = \frac{\ln \frac{E_0}{E_c}}{\ln 2} \tag{3.5}$$

yields the number of radiation lengths traversed to this point of "shower maximum." For example, for a 500GeV primary γ-ray, we would expect shower maximum to occur after $n \sim 12.5$ radiation lengths (corresponding to a depth of $\sim$ 23km for a first-interaction height of 30km) in which $n \sim 6 \times 10^3$ particles would be produced. In reality, the bremsstrahlung process has been grossly oversimplified by this model. Since the energy distribution of the photons radiated by the process is actually a continuum below the initial electron energy, considerable loss occurs in the form of low energy photons which are below the pair-production threshold (whereas we attributed all of the loss to a single pho-

ton). This results in fewer particles being produced than the above estimate suggests; a more reasonable approximation is that the number of electrons at shower maximum is given by

$$N_e \approx 0.5 E_{\mathrm{GeV}}.$$

While comparison of such calculations with those done by Monte Carlo simulations (see Section 3.3) shows our simple model to be very approximate indeed, the general features of EAS are illustrated; the growth of the shower is initially exponential and the yield of secondary particles at shower maximum is proportional to the energy E_0 of the primary.

Thus far, we have considered showers produced only from incident high energy electrons or γ-rays. EAS may also be initiated by other high energy particles. As our previous discussion would suggest, a shower initiated by a high energy electron will be indistinguishable from that initiated by a γ-ray. In either case, the shower is referred to as an electromagnetic cascade; such showers consist almost entirely of electrons and photons. Because electrons are light, they easily radiate energy and, as such, are not expected to contribute significantly to the flux of high energy particles that have travelled large distances through interstellar matter. Hadrons, on the other hand, are sufficiently massive to maintain considerable energy in their transport throughout the galaxy. When such particles are incident on the atmosphere, the first interaction generally produces a number of secondary pions, both charged (π^+, π^-), and neutral (π^0). The π^0's eventually decay into γ-rays which then go on to produce electromagnetic cascades. The

charged pions decay via

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \text{ or } \pi^- \rightarrow \mu^- + \bar{\nu}_\mu.$$

These resulting muons may provide a signature to differentiate between the otherwise similar electromagnetic and hadronic cascades. In showers induced by γ-rays, there are two types of processes involved in the generation of muons: photoproduction of pions (which then decay to produce muons) and direct creation of muon pairs. The cross-sections for both of these processes are extremely small and, thus, very few muons are expected from such showers (Stanev *et al.* 1985). There have been suggestions (Samorski & Stamm 1983) that electromagnetic cascades may, through some unexplained process in the shower development, contain considerable numbers of muons, but these claims have yet to be confirmed experimentally or explained theoretically. Because the number of pions produced in a hadronic shower will be roughly equally divided among the three varieties, the resulting electromagnetic component of the shower will contain approximately one-third the number of particles as a shower produced by a γ-ray of the same initial energy. On the other hand, because the hadronic particles are more penetrating than electrons, hadronic showers will contain more particles at large depths in the atmosphere than their electromagnetic counterparts. These statements are both quite energy dependent, however. For very high energies, heavy pionization (*e.g.* $\pi^+ N \rightarrow \pi^+\pi^-\pi^0 N$) results in the charged pions dumping much of their energy into neutral pions which then decay into γ-rays thus producing additional electromagnetic cascades. In the extreme case a large fraction of the charged pion energy may be reprocessed into the electromagnetic

component.

The angular spread of hadronic cascades is somewhat larger than for electromagnetic cascades due to the transverse momentum of the pions. For high energy collisions, the average transverse momentum is

$$\langle p_\perp \rangle \sim 0.4 \text{ GeV} \tag{3.6}$$

independent of energy. Roughly half of the initial proton energy is expended in making pions. The multiplicity at $E \sim 1$ TeV is ~ 10 with the production of kaons and baryon-antibaryon pairs less frequent by one and two orders-of-magnitude respectively (Perkins 1982). Thus,

$$\langle E \rangle_\pi \approx \frac{E_0}{2(10)} \approx p_{||}c \quad \Rightarrow \quad \theta = \frac{p_\perp}{p_{||}} \approx \frac{8 \text{ GeV}}{E_0}.$$

For a 1TeV proton shower, then, $\theta \sim 0°.5$.

3.1.2 Čerenkov Radiation

When a relativistic particle travels through a medium with a speed in excess of the local velocity of light, a form of electromagnetic shock-wave is produced in the medium resulting in the emission of so-called Čerenkov radiation. This process is analogous to the formation of pressure shocks during supersonic travel and results in the emission of visible light at an angle θ_c with respect to the particle's velocity; a cone of light is produced (Figure 3.2) with an apex half-angle given by

$$\theta_c = \cos^{-1} \frac{1}{\beta n} \tag{3.7}$$

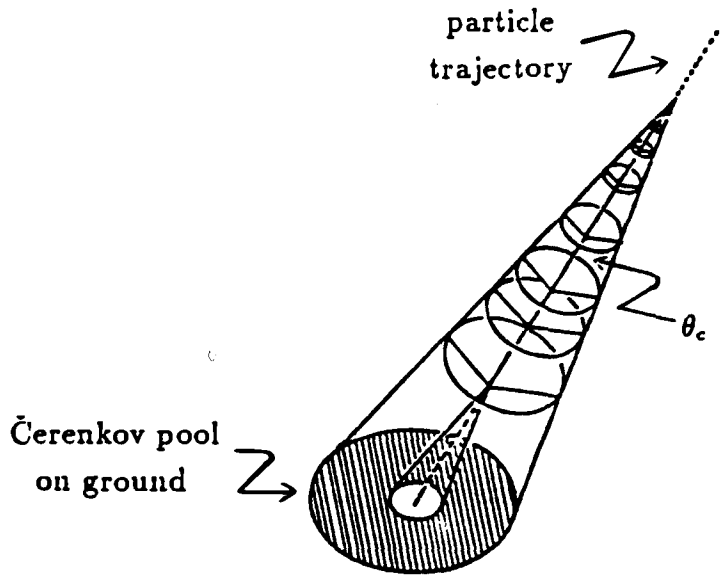


Figure 3.2: Geometry of Čerenkov radiation.

where n is the refractive index of the medium and $\beta = v/c$ is the particle velocity. For particles of very high energy, $\beta \approx 1$ so that

$$\theta_c \approx \cos^{-1}\frac{1}{n} \approx 1^\circ.3 \text{ in air at sea level.}$$

The rate of production of Čerenkov photons in the frequency range (ν_{min}, ν_{max}) is given by

$$N_c = \frac{1}{137c}\int_{\nu_{min}}^{\nu_{max}}(1 - \frac{1}{\beta^2 n^2})2\pi \; d\nu \text{ photons m}^{-1}. \tag{3.8}$$

Expressing this in terms of wavelength, and noting that

$$1 - \frac{1}{\beta^2 n^2} = \sin^2\theta_c$$

we have

$$\frac{dN_c}{d\lambda} = \frac{2\pi}{137}\frac{\sin^2\theta_c}{\lambda^2} \tag{3.9}$$

which, in the optical range (0.4nm< λ <0.7nm), integrates to

$$N_c \approx 5 \times 10^4 \sin^2 \theta_c \text{ photons m}^{-1}. \tag{3.10}$$

The preference for production at short wavelengths is fortuitous from the standpoint of photocathode efficiency in photomultiplier tubes and can also be used to reduce the relative contribution of background skylight (see Section 3.3.2).

There is a threshold velocity for the production of Čerenkov radiation given by

$$\cos \theta_c = \frac{1}{n\beta} = 1. \tag{3.11}$$

But since

$$\beta \equiv \sqrt{1 - \gamma^{-2}} \approx 1 - \frac{1}{2\gamma^2} \tag{3.12}$$

or

$$\gamma \geq \sqrt{\frac{n}{2(n-1))}} \tag{3.13}$$

we find a threshold energy:

$$E_{th} = \gamma m c^2 \geq m c^2 \sqrt{\frac{n}{2(n-1)}}. \tag{3.14}$$

For an atmospheric shower we may assume an exponential atmosphere for which

$$n = 1 + \eta(z) \tag{3.15}$$

where

$$\eta(z) = \eta_0 e^{-\frac{z}{z_0}}. \tag{3.16}$$

Here z is the atmospheric height and z_0 is the scale height. Using an isothermal approximation, $\eta_0 \simeq 2.7 \times 10^{-4}$ and $z_0 \simeq 8.2$ km. Thus, at sea level, $n \simeq 1.0003$

so that

$$E_{th} \simeq 41mc^2. \tag{3.17}$$

Corresponding threshold energies for Čerenkov production are 21 MeV for electrons, 4 GeV for muons, and 39 GeV for protons. Because these values scale exponentially with z it is clear that, even in hadronic showers, the bulk of the Čerenkov light is produced by the electrons.

Since the Čerenkov threshold for electrons is relatively low it is not surprising that, expecially in large cascades (initiated by very energetic particles), there are electrons with energies above threshold which penetrate deep into the atmosphere. Because these electrons travel at approximately speed c while the Čerenkov photons emitted at higher points in the atmosphere travel at speeds reduced by the index of refraction, there is a spread in time of arrival of photons produced at different altitudes. A simple calculation, neglecting additional delays due to path-length differences due primarily to multiple scattering, illustrates the effect. Suppose the first photons are produced at a height z_1. Since the photons travel at

$$v = \frac{dz}{dt} = \frac{c}{n} \tag{3.18}$$

we see that the time required to propagate to a height z_2 is

$$t = \frac{1}{c}\int_{z_2}^{z_1} n(z)dz \tag{3.19}$$

which, substituting Equation (3.11), yields

$$t = \frac{1}{c}\left[(z_1 - z_2) - \eta_0 z_0 \left(e^{-\frac{z_1}{z_0}} - e^{-\frac{z_2}{z_0}}\right)\right]. \tag{3.20}$$

Meanwhile, the electrons travel the same distance in a time

$$t_e = \frac{z_1 - z_2}{c}. \tag{3.21}$$

Thus, the difference in arrival times of photons produced at height z_2 and those produced at z_1 is

$$\Delta t = \frac{\eta_0 z_0}{c}\left(e^{-\frac{z_2}{z_0}} - e^{-\frac{z_1}{z_0}}\right). \tag{3.22}$$

As an example, suppose $z_1 = 11$ km and $z_2 = 5$ km (representing a reasonable spread about a typical height of 8km for shower maximum). The expression above then yields $\Delta t \approx 2.1$ ns. At observation level, then, the shower produces a "pancake" of Čerenkov radiation several nanoseconds in duration. In reality, the width is affected by multiple scattering and zenith angle considerations as well (see Section 3.1.3). This short duration is crucial in the detection of EAS by optical techniques in the presence of ambient starlight.

3.1.3 Monte Carlo Studies of EAS

In order to go beyond the simplified calculations describing the development of EAS outlined in the previous sections, it is necessary to consider the full details of the cascade. The generation of particles in such a process may be described by

$$\frac{dN_j}{dE_j dz} = -\frac{1}{\lambda_j}\frac{dN_j}{dE_j} + \sum_i \int_{E_j}^{\infty} \frac{1}{\lambda_i}\frac{dN_i}{dE_i}\frac{dN_{i\to j}}{dE_j} dE_i. \tag{3.23}$$

Here the first term accounts for the interaction or decay of particles of species j in the energy range $(E_j, E_j + dE)$ while the second term represents the production of particles of species j in the energy range $(E_j, E_j + dE)$ by particles of species

i from the energy range $(E_i, E_i + dE)$. The column density z is defined by

$$z = \int \rho \, dl \tag{3.24}$$

while the evolution length λ is a combination of the decay length λ_D and the interaction length λ_I:

$$\frac{1}{\lambda} = \frac{1}{\lambda_D} + \frac{1}{\lambda_I}. \tag{3.25}$$

The determination of the cascade parameters, then, requires the solution of the system of differential equations above. Generally this can be accomplished only by introducing approximations which simplify the solution (see, *e.g.*, Rossi and Greisen 1941, Rossi 1952, Greisen 1956, Misaki 1970).

While the above analytical solutions to the cascade equations concentrate on the particle component of the shower, our primary interest is in the Čerenkov radiation produced by these particles. The Čerenkov production can also be approximated analytically in a manner similar to that for the cascade development. An alternative to the analytical solution is to perform Monte Carlo simulations. In such simulations, all interactions are treated stochastically with formal integration over interaction probability densities replaced by discrete statistical "simulated interactions." Thus, a single calculation yields only one of many possible final results, but the distribution resulting from a large number of such calculations converges to that which would be obtained if the formal integration (if possible) was performed. Such a simulation program has been written by members of the collaboration (in collaboration with T. Stanev). The program is divided into two parts. First, given parameters describing the primary particle (particle type, energy, direction, etc.), each shower particle is followed until its

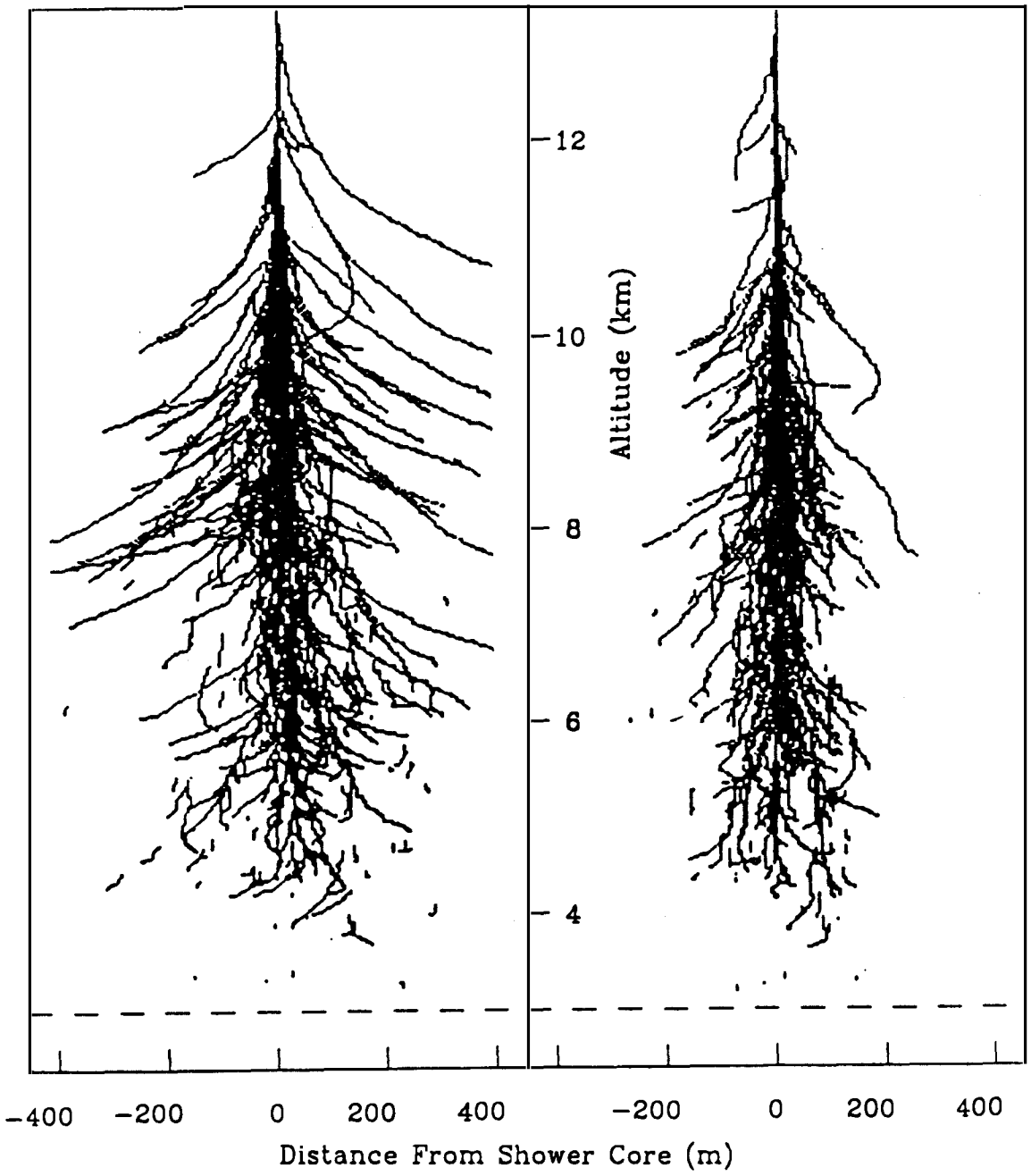


Figure 3.3: Spatial (a)East-West and (b)North-South development of EAS initiated by 100 GeV γ-ray.

energy reaches a particular cutoff (usually taken to be the threshold energy for production of Čerenkov radiation). The particle paths are divided into segments corresponding to no more than a particular number (generally 0.2) of radiation lengths. For each such segment, the direction cosines, Cartesian coordinates (relative to the projected position of the primary on the ground) and times (relative to the time the primary would have reached the ground without interacting) are recorded for the beginning and end of the path. The energy at the track midpoint is recorded as well.

The spatial development of a shower initiated by a 100 GeV γ-ray is illustrated in Figures 3.3(a) and (b). The effect of the magnetic field of the Earth is evident as a broadening of the shower in the East-West direction as compared with the North-South direction. Also apparent are the tracks of several low-energy electrons which have scattered through quite large angles. Because of these and other fluctuations present in the shower development, the general behavior is best represented by the average of many such showers. To illustrate the overall features, simulations of showers initiated by 500 GeV γ-rays have been performed. In the following, we present the averages of 50 individual showers. In Figure 3.4, the number of shower particles as a function of the number of radiation lengths traversed is plotted. The exponential increase in the growth of the shower is evident. Shower maximum occurs after about 8 radiation lengths corresponding to a typical altitude above sea level of 8 km. The number of particles (not including γ-rays) at shower maximum is $\sim$300. We note that the approximations reached in Section 3.1.1 are quite rough although the general be-

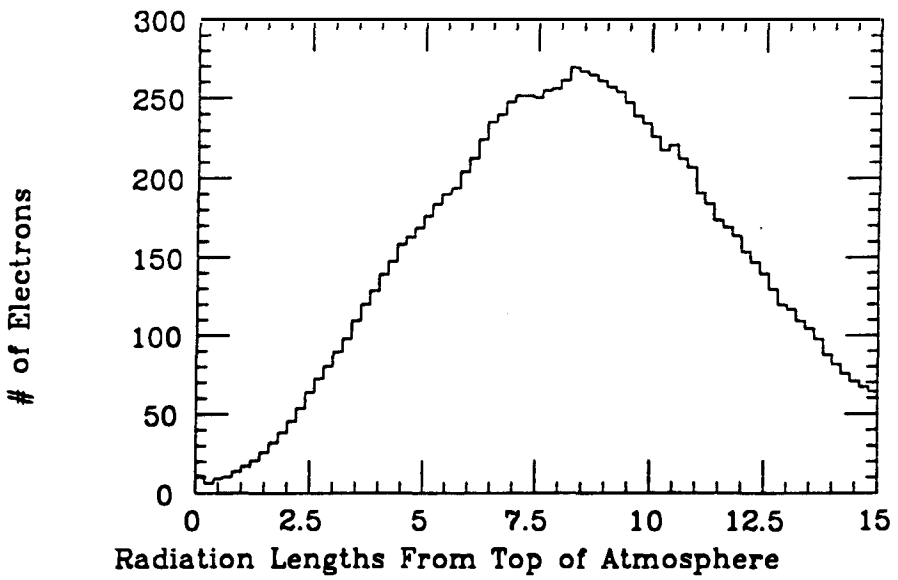


Figure 3.4: Number of electrons as a function of radiation lengths penetrated for EAS initiated by 500 GeV γ-ray.

havior of the shower development is similar to that described earlier. The energy profile of the showers is displayed in Figure 3.5. The trend toward lower mean energy with increasing shower depth is clear. Note, however, that a considerable spread in energies appears quite early in the shower development.

For comparison, Figures 3.6-3.7 contain similar plots to those above for showers initiated by 500 GeV protons. Note that fewer particles are produced in such showers, as expected. In Figure 3.7, a penetrating component to the shower is obvious. These particles, primarily muons, are an identifying feature of hadronic showers. They result in large localized fluctuations in the photon density. In addition, the transverse momentum of the pions produced early in the devel-

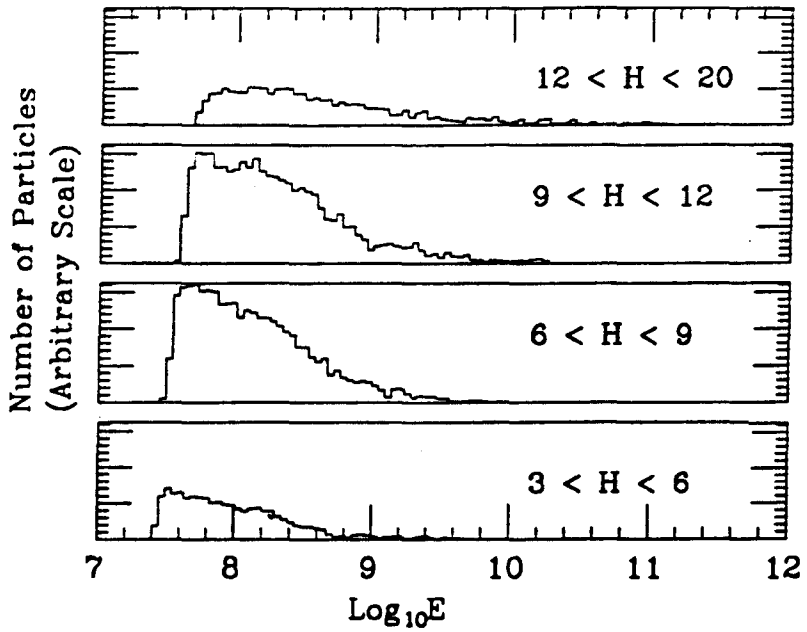


Figure 3.5: Particle energy as a function of penetration depth (500 GeV γ-ray primary).

opment of a hadronic cascade causes the resulting electromagnetic components to the shower to produce a somewhat uneven distribution of Čerenkov photons on the ground. Such a photon density profile may be useful in discriminating showers of hadronic origin from those initiated by γ-rays if imaging techniques or several independent telescopes are used. As shown if Figure 3.8, the ratio of electrons in γ-ray induced showers to those in proton showers is approximately 3 as suggested earlier.

The second part of the simulation program generates the Čerenkov photons from each of the particle path segments. The photons are propagated to the ground (with atmospheric extinction taken into account) and the final positions,

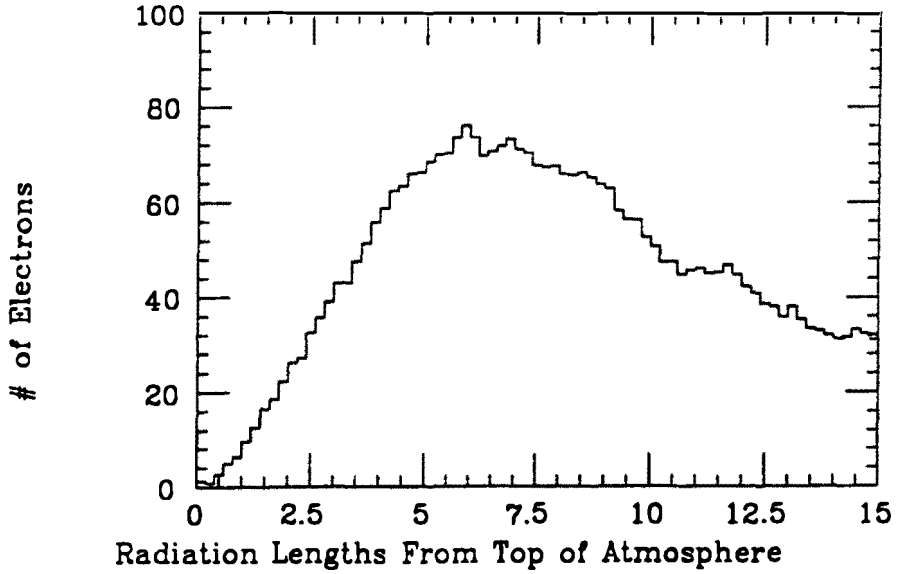


Figure 3.6: Number of electrons as a function of radiation lengths penetrated for EAS initiated by 500 GeV proton.

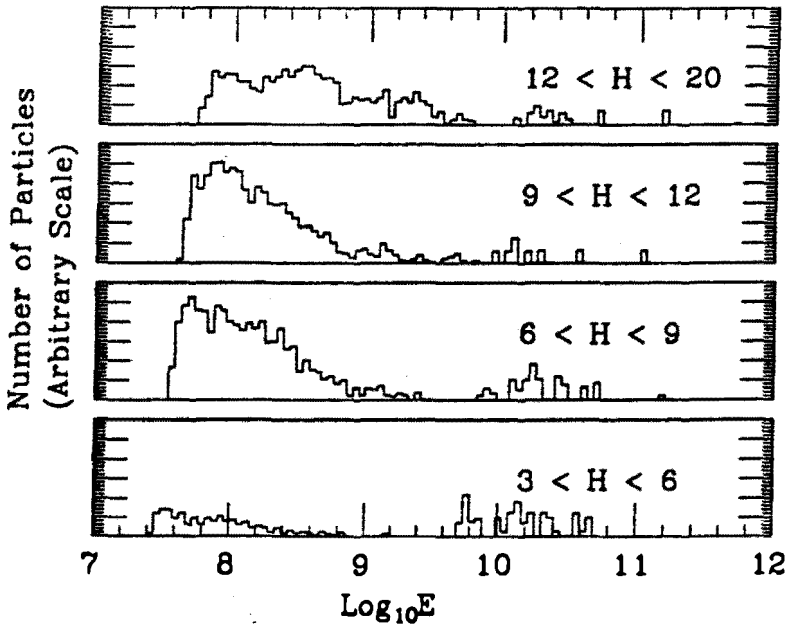


Figure 3.7: Particle energy as a function of penetration depth (500 GeV proton primary).

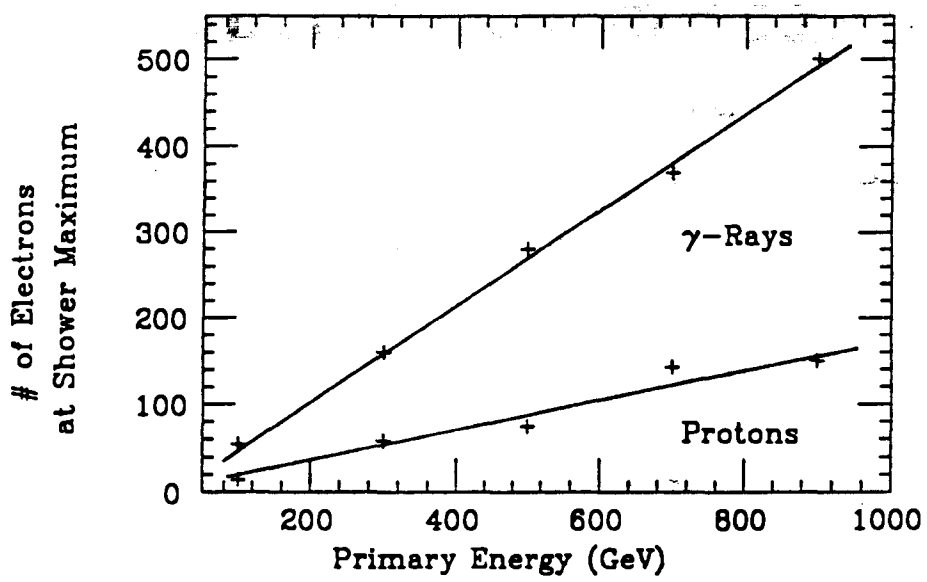


Figure 3.8: Ratio of number of electrons at shower maximum for γ-ray induced showers to that for proton induced showers.

directions, times, and wavelengths are recorded for each. Optionally, simulations of the telescope response characteristics can be invoked to study the distribution of photons actually detected. For purposes of direct comparison with the telescope of interest, only those photons which f ll in the bandpass 300 nm $\leq \lambda \leq$ 475 nm have been considered. Since the Čerenkov spectrum is well known, scaling to larger wavelength bands can be accomplished easily. Figure 3.9 is a plot of the photon density as a function of distance from the shower core for showers initiated by 500 GeV γ-rays. The shoulder at a radius of $\sim$ 80m is a geometric effect resulting from the size of the Čerenkov angle at shower maximum (where the largest amount of light is produced). While we shall see that the photon

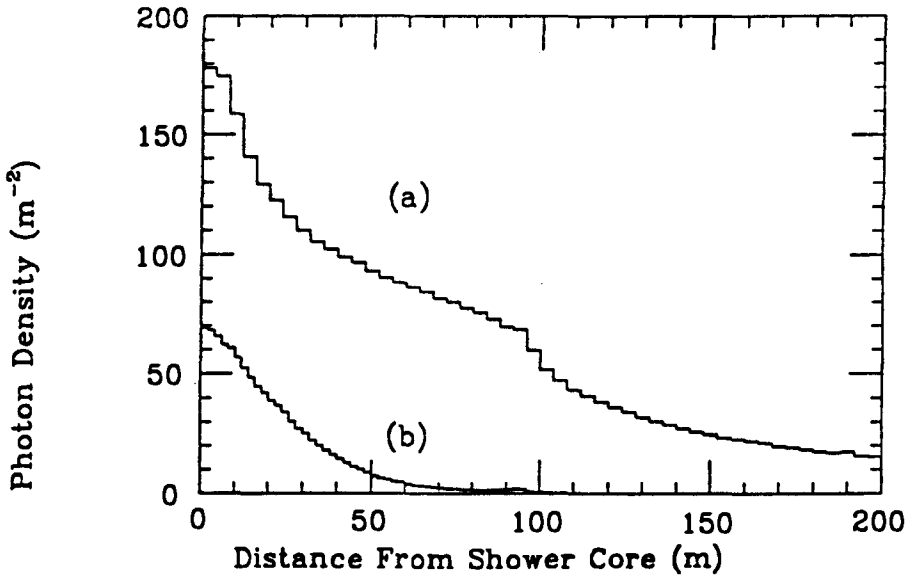


Figure 3.9: Čerenkov photon density as a function of distance from shower core: (a)all photons; (b)photons for which $2\theta < 0°.75$ (500 GeV γ-ray primary).

density is sufficiently high to trigger a telescope only within $\sim$ 50m of the shower core, light from the shower is spread over an appreciably larger distance. The photon density is decreased when an aperture constraint corresponding to the $0°.75$ full-width acceptance of the Haleakala telescope is imposed. The angular distribution of the photons, shown in Figure 3.10, has a width of $\sim 0°.75$ for $r < 50$m. This sets an upper limit to the useful telescope aperture size; using an angular acceptance larger than this will merely increase the amount of background skylight.

The distribution of arrival times for photons within both the angular range of acceptance and the radial range for sufficient density is shown in Figure 3.11.

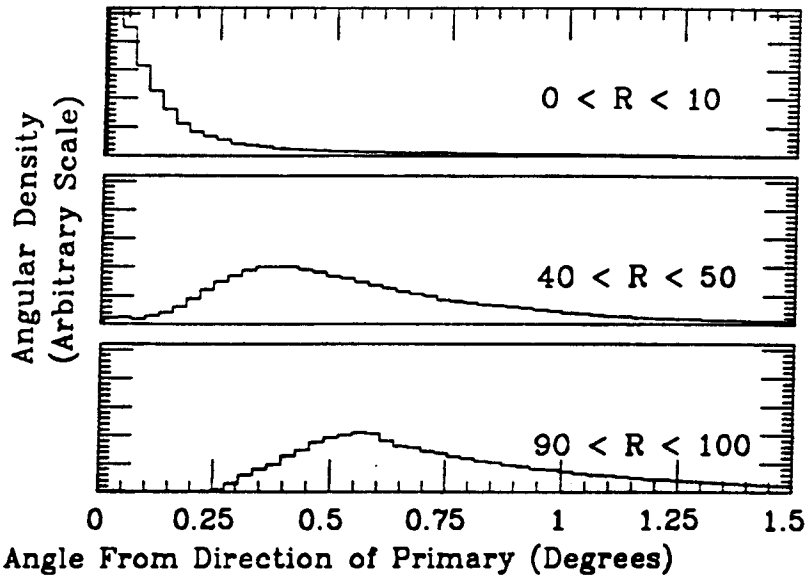


Figure 3.10: Angular distribution of Čerenkov photons (500 GeV γ-ray primary).

As suggested earlier, the duration is quite short ($\Delta t \leq 2.5$ns). A technique for detecting the light from these showers suggests itself; one can require the arrival of a large number of photons inside of a narrow time window (see Section 3.3.3). Figures 3.12-3.14 include corresponding plots for showers initiated by protons.

The above discussion permits us to quantify the shower characteristics sufficiently well to allow a reasonable calculation of the rate at which such showers will be detected given the characteristics of the telescope being used. The density of Čerenkov photons produced in a shower determine the ability of the telescope to trigger. Given a reasonable form for the density profile for showers, along with the rate at which showers are incident, a calculation of the expected trigger rate can be made. Let ρ be the photon density in photons m^{-2}. We may assume a

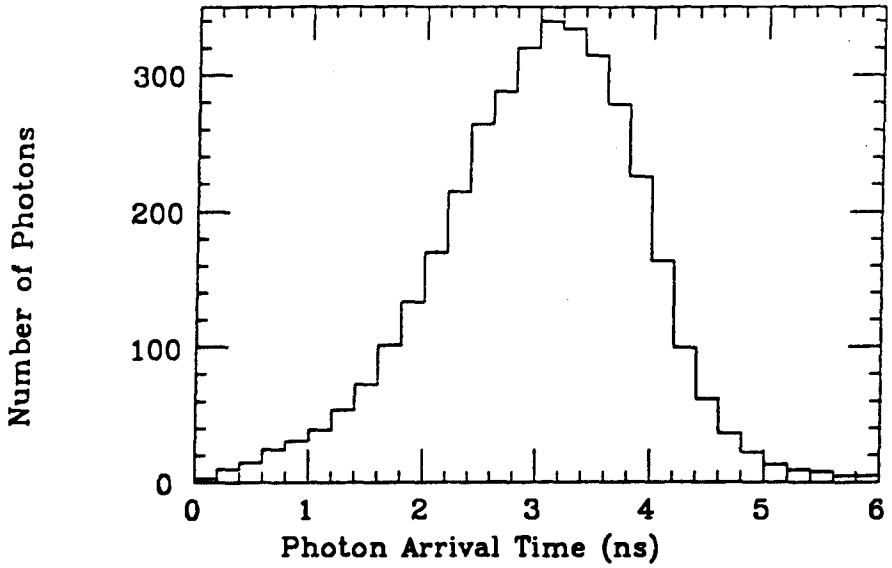


Figure 3.11: Distribution of arrival times for Čerenkov photons within Haleakala telescope angular acceptance and trigger radius (500 GeV γ-ray primary).

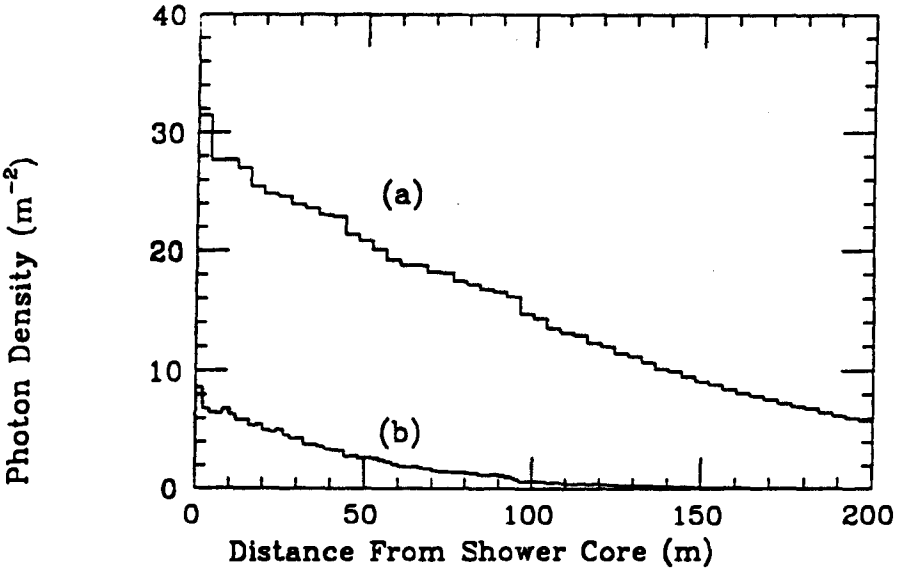


Figure 3.12: Čerenkov photon density as a function of distance from shower core: (a)all photons; (b)photons for which $2\theta < 0°.75$ (500 GeV proton primary).

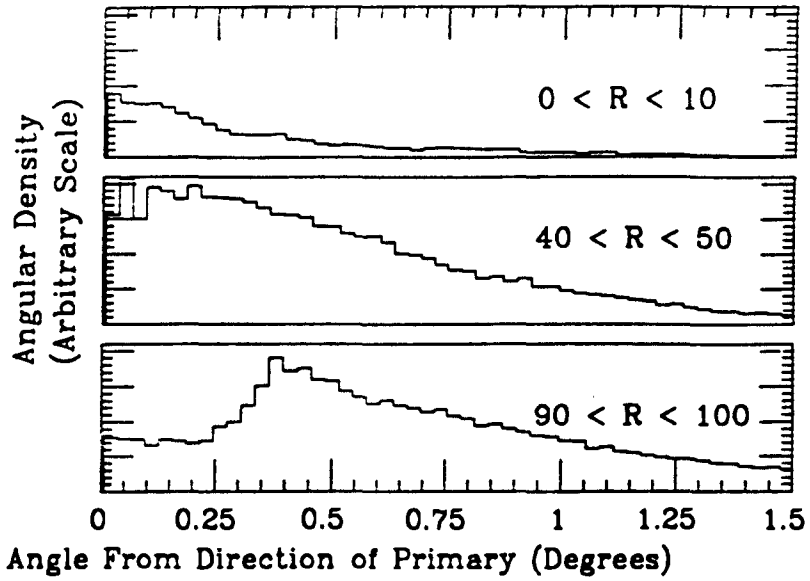


Figure 3.13: Angular distribution of Čerenkov photons (500 GeV proton primary).

density profile of the form

$$\rho = \alpha E^{\kappa} e^{-(\frac{r}{r_0})^{\epsilon}} \tag{3.26}$$

where r_0 is the characteristic radius at which the density is reduced by a factor e^{-1}. As shown in Figure 3.15, we can use $\kappa \approx 2$ and $\alpha \approx 24$ m^{-2} TeV^{-2} to a fair degree of accuracy. Similarly, Figure 3.16 indicates that $\epsilon \approx 1$ (the dotted curves represent the profiles given by equation 3.26 using $\epsilon = 1$). Thus,

$$\ln\left(\frac{\rho}{\alpha E^2}\right) = -\frac{r}{r_0} \Rightarrow r = r_0 \ln\left(\frac{\alpha E^2}{\rho}\right). \tag{3.27}$$

We can write the differential energy spectrum for primary protons as

$$J(E) = \beta E^{-\gamma} \text{ showers m}^{-2} \text{ sr}^{-1} \text{ s}^{-1}. \tag{3.28}$$

Then, the total shower rate for showers of energy E with photon density $\rho > \rho_0$

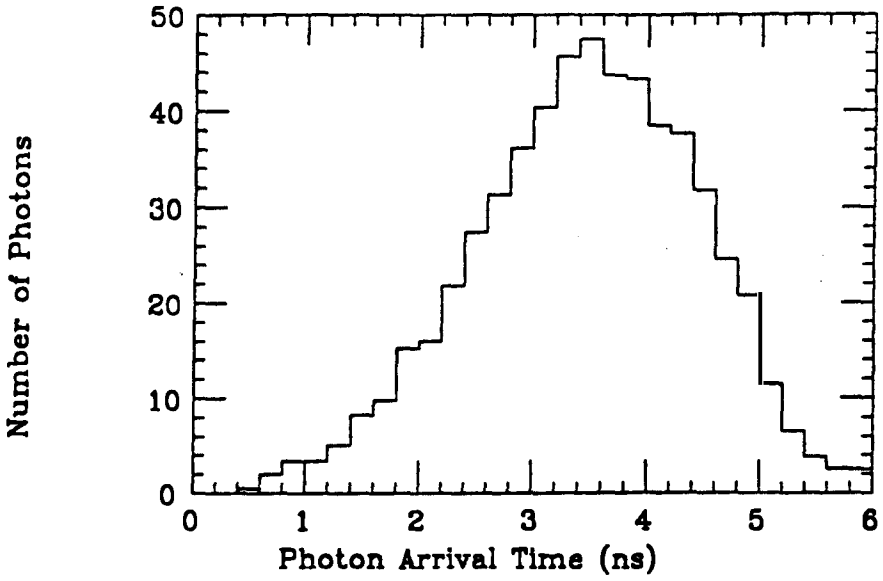


Figure 3.14: Distribution of arrival times for Čerenkov photons within Haleakala telescope angular acceptance and trigger radius (500 GeV proton primary).

is

$$R(>\rho_0, E) = J(E)A(>\rho_0)\Omega \tag{3.29}$$

where $A(>\rho_0)$ is the effective area for showers with $\rho > \rho_0$ and Ω is the solid angle accepted by the telescope aperture. Thus,

$$R(>\rho_0, E) = \beta E^{-\gamma}(\pi r^2)\Omega = \beta E^{-\gamma}\pi r_0^2 \left[\ln\left(\frac{\alpha E^2}{\rho_0}\right)\right]^2 \Omega \text{ showers s}^{-1}. \tag{3.30}$$

Hence, the total rate for showers with $\rho > \rho_0$ is

$$R(>\rho_0) = \int_{E_c}^{\infty} R(>\rho_0, E)dE \tag{3.31}$$

where

$$E_c = \left(\frac{\rho_0}{\alpha}\right)^{\frac{1}{2}} \tag{3.32}$$

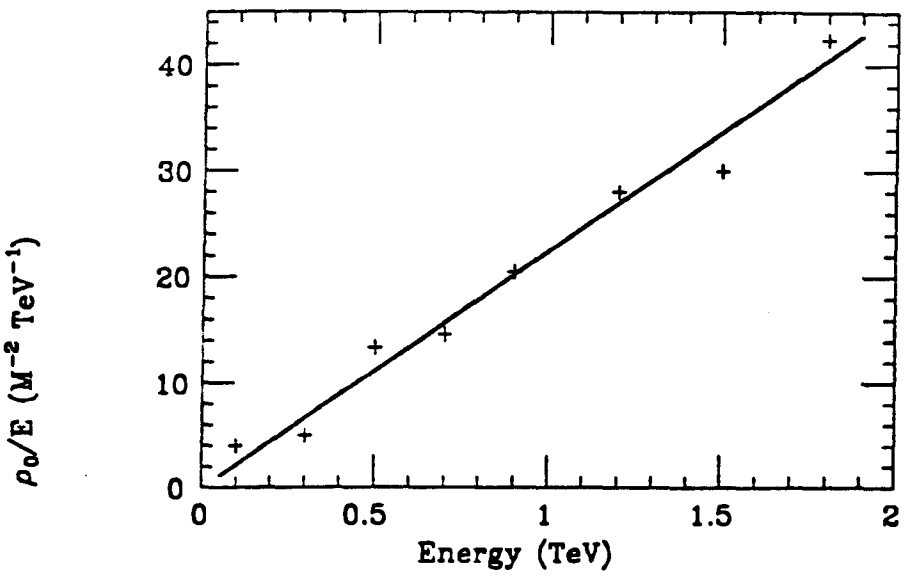


Figure 3.15: Ratio of photon density (at shower core) to energy as a function of proton primary energy (0°.75 aperture).

is the critical (minimum) energy for which density ρ_0 can be achieved. Therefore,

$$R(>\rho_0) = \beta\pi r_0^2\Omega \int_{E_c}^{\infty} E^{-\gamma}\left[\ln\left(\frac{\alpha E^2}{\rho_0}\right)\right]^2 dE. \tag{3.33}$$

Integrating,

$$R(>\rho_0) = \frac{8\beta\pi r_0^2\Omega}{(\gamma-1)^3}\left(\frac{\alpha}{\rho_0}\right)^{\frac{\gamma-1}{2}}. \tag{3.34}$$

We know the integral spectrum for primary cosmic rays (Lang 1980)

$$J(>E) = 0.16 E_{\rm TeV}^{-1.6} \text{ showers m}^{-2}\text{ s}^{-1}\text{ sr}^{-1}. \tag{3.35}$$

But

$$J(>E) = \int_E^{\infty} J(E')dE' = \frac{\beta E^{-\gamma+1}}{\gamma-1}. \tag{3.36}$$

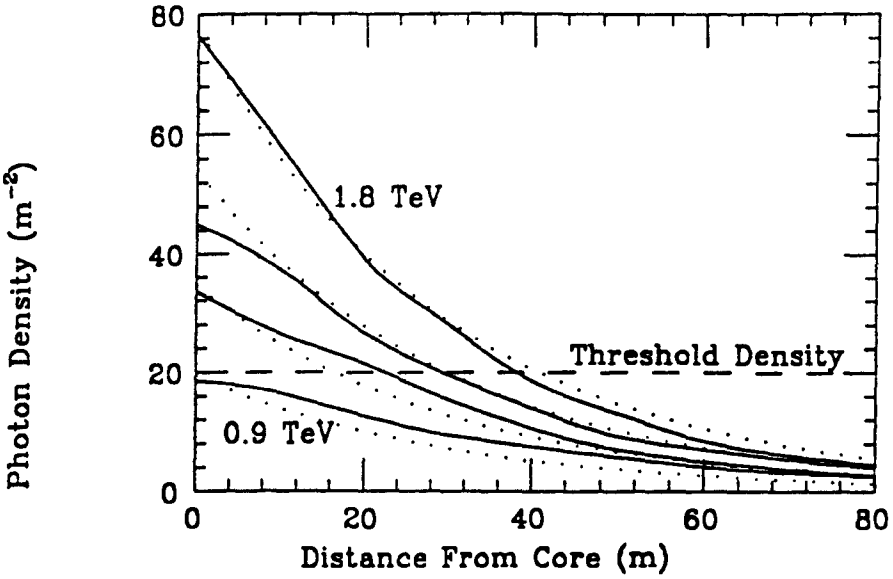


Figure 3.16: Photon densities for showers of various (proton) primary energies (in steps of 0.3 TeV). Effective trigger radii are defined by the critical trigger density line $\rho = \rho_c$. From this we can establish a trigger threshold of $\approx$ 900 GeV. The dotted curves represent the approximate density profile given by Equation 3.26 using parameters given in text.

Thus, we have $\gamma = 2.6$ and $\beta = .256$ m^{-2} sr^{-1} s^{-1} TeV^{-1}.

As described in Section 1.2 (see also Section 3.3.2) the aperture for the Haleakala γ-ray telescope has a full-angle of 0°.75 which corresponds to a solid angle $\Omega = 1.35 \times 10^{-4}$ sr and the critical density for triggering is $\rho_0 \approx$ 20 photons m^{-2}. From Figure 3.15, we saw that $\alpha \sim 24$ photons m^{-2} Tev^{-2} and from Figure 3.16 we see that $r_0 \sim 30$m. The trigger rate for the telescope is

then

$$R = \frac{8(.256)\pi(30)^2(1.35 \times 10^{-4})}{(1.6)^3}\left(\frac{24}{20}\right)^{0.8} \text{ s}^{-1} = 0.22 \text{ Hz}. \qquad (3.37)$$

It should be noted that this rate is somewhat of an underestimate in that the solid angle of acceptance is actually an energy dependent quantity. High energy showers can contain sufficient photon density at large angles to permit protons which are incident at angles larger than the aperature of the telescope to cause a trigger.

We can estimate the energy threshold of the telescope as well since

$$E_{th} \approx \left(\frac{\rho_0}{\alpha}\right)^{\frac{1}{2}} \approx 900 \text{ GeV}. \qquad (3.38)$$

This is consistent with the Monte Carlo results shown in Figure 3.16. Here the photon density is plotted for various values of primary energy and the effective radius for triggering is identified. The proton energy threshold is clearly about 900 GeV. Of course, the effective area A over which a shower can trigger the telescope is a function of energy. Thus it is useful to quote an effective energy defined as that for which the product $AE^{-\gamma}$ is a maximum. This will take into account the larger collection area available for higher energy showers along with the relative scarcity of such showers. Using equation 3.27 with the critical density $\rho_c = 20 \text{ m}^{-2}$, we have

$$AE^{-\gamma} = \pi r_0^2 \left[\ln\left(\frac{\alpha E^2}{\rho_c}\right)\right]^2 E^{-\gamma}.$$

Differentiating to determine the maximum, we find $E_{eff} \sim 2$ TeV (Figure 3.17). Because γ-ray induced showers of a given primary energy yield greater amounts of Čerenkov light that their proton induced counterparts, we expect the energy

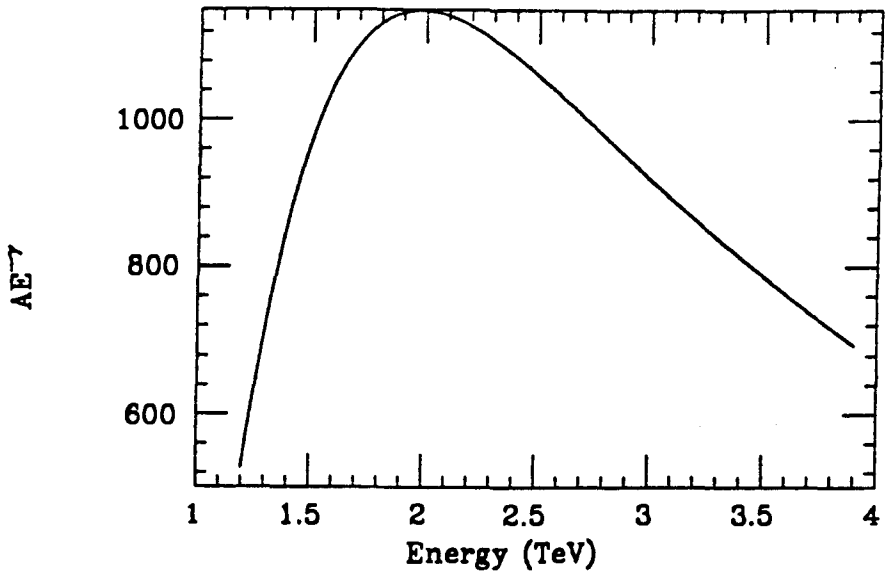


Figure 3.17: Effective energy at which telescope operates. The quantity $AE^{-\gamma}$ is plotted as a function of energy.

threshold for triggering on such showers to be lower. In Figure 3.18 we plot the photon density at the shower core as a function of energy while Figure 3.19 shows the density profiles for showers initiated by γ-rays of various energies. Clearly, the threshold is approximately 200 GeV. From these two figures we see that we may use the parameterization given by Equation 3.26 with $\kappa = 1$, $\alpha = 135$ m^{-2} TeV^{-1}, $\epsilon = 1$, and $r_0 = 30$ m. As with the proton induced showers, we may define an effective energy at which the telescope operates for γ-rays by considering the product $AE^{-\gamma}$. However, for primary γ-rays, the spectral index is unknown. Thus, in Figure 3.20 we plot the effective energy as a function of spectral index.

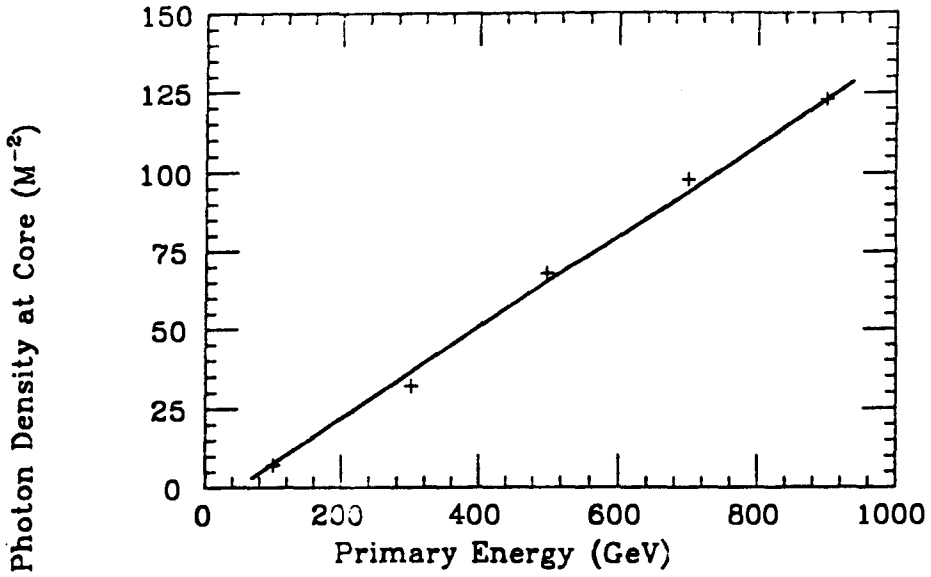


Figure 3.18: Photon density as a function of γ-ray primary energy (0°.75 aperture). Note that the density scales linearly with energy.

While it is clear that optical photons from EAS are in sufficient supply for the detection by a sensitive telescope, there still remains the task of identifying these photons in the presence of random night-sky background. We shall see that, although the Čerenkov photons from EAS make up only 10^{-4} of the night-sky brightness, we can use the distinct signature of the EAS photons to effectively discriminate against the background.

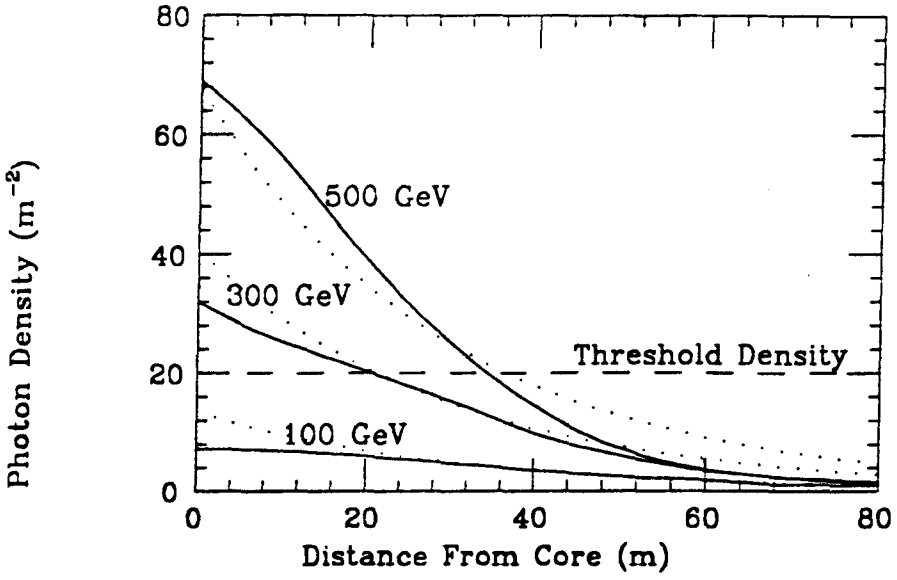


Figure 3.19: Photon densities for showers of various γ-ray primary energies. Effective trigger radii are defined by the critical trigger density line $\rho = \rho_c$. From this we can establish a trigger threshold of $\approx$ 200 GeV.

3.2 The Haleakala Gamma Ray Telescope

3.2.1 Introduction

The Haleakala gamma ray telescope is the result of a plan to introduce fast electronics to the standard Atmospheric Čerenkov Technique (ACT) in an effort to obtain energy thresholds nearly an order-of-magnitude lower, for a given area, than those attained by other facilities. Previous gamma ray telescopes have concentrated light from large mirrors onto relatively few phototubes. The resulting singles rates are quite high thus requiring the operation of the PMT's at low voltage to maintain reasonable anode currents. With the reduced voltage,

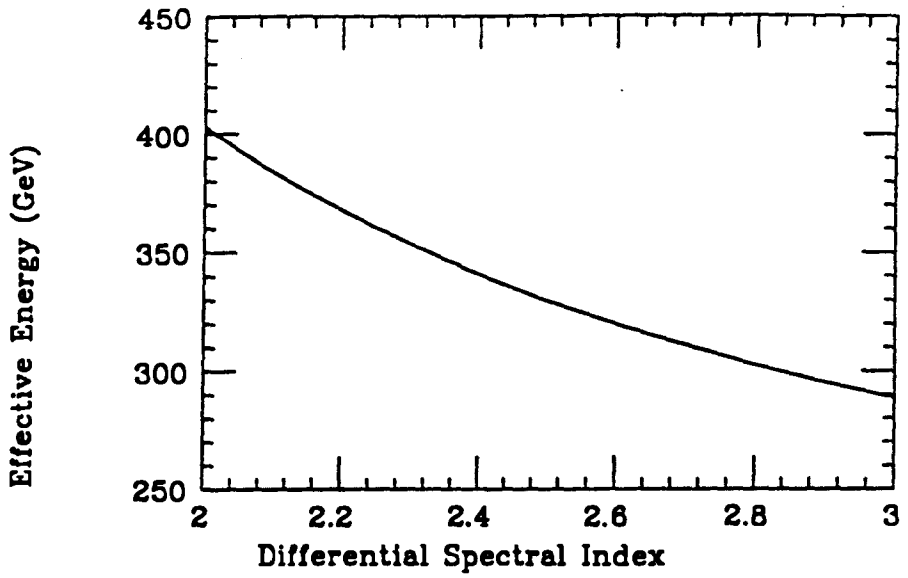


Figure 3.20: Effective energy for triggering on showers initiated by γ-rays as a function of spectral index.

many photoelectrons are required to provide a reasonable signal within a several nanosecond gate. By using fast PMT's which service a small mirror area, the Haleakala telescope operates at the single photoelectron level thus making more efficient use of the area. The result is that the time development of electromagnetic cascades can be studied on timescales of nanoseconds.

The telescope is located 10,000 feet above sea level on Mt. Haleakala, a dormant shield volcano situated on the island of Maui in Hawaii (east longitude $-156°15'30"$, latitude $20°42'36"$). The skies at the site are extremely clear with little sky shine from nearby cities and minimal cloud cover. The latitude allows observations of some southern hemisphere objects currently outside the viewing

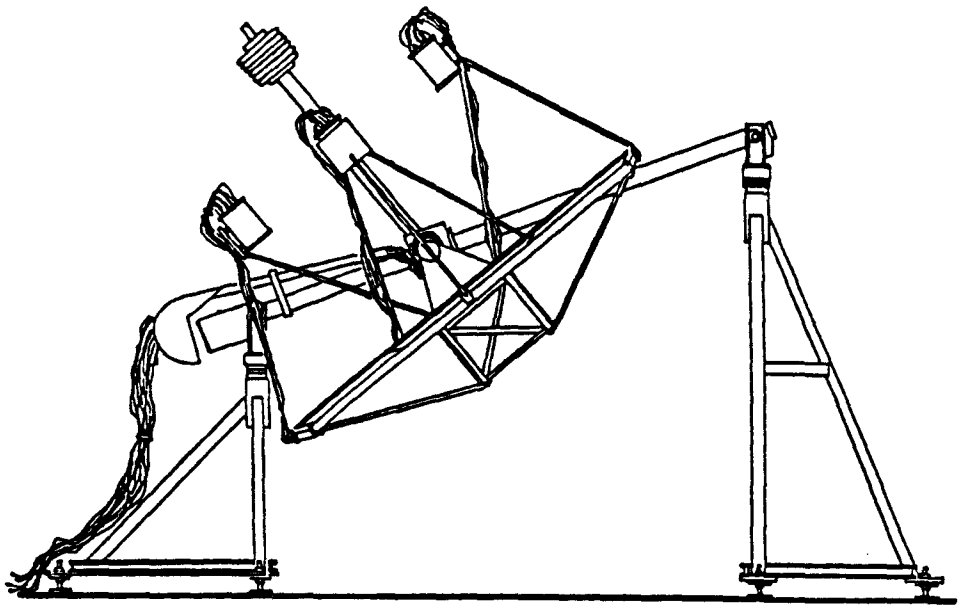


Figure 3.21: The Haleakala VHE γ-ray telescope (from Szentgyorgyi, 1986).

windows of many other experiments. A brief description of the apparatus and the data acquisition system follows. For a more complete discussion, the reader is referred to Szentgyorgyi (1986) and Resvanis *et al.* (1988).

3.2.2 Physical Description

The telescope consists of six 1.5m diameter, spherical mirrors situated on a digitally driven equatorial mount (Figure 3.21). The mirrors were fabricated by the University of Utah Cosmic Ray Group (for use in the "Fly's Eye" experiment) by slumping ordinary plate glass on a graphite mold. The optical quality of the mirrors (point-source spot size of $\sim 0^{\circ}.5$) is adequate for use in detecting air showers since the angular size of a shower is of approximately the same magnitude

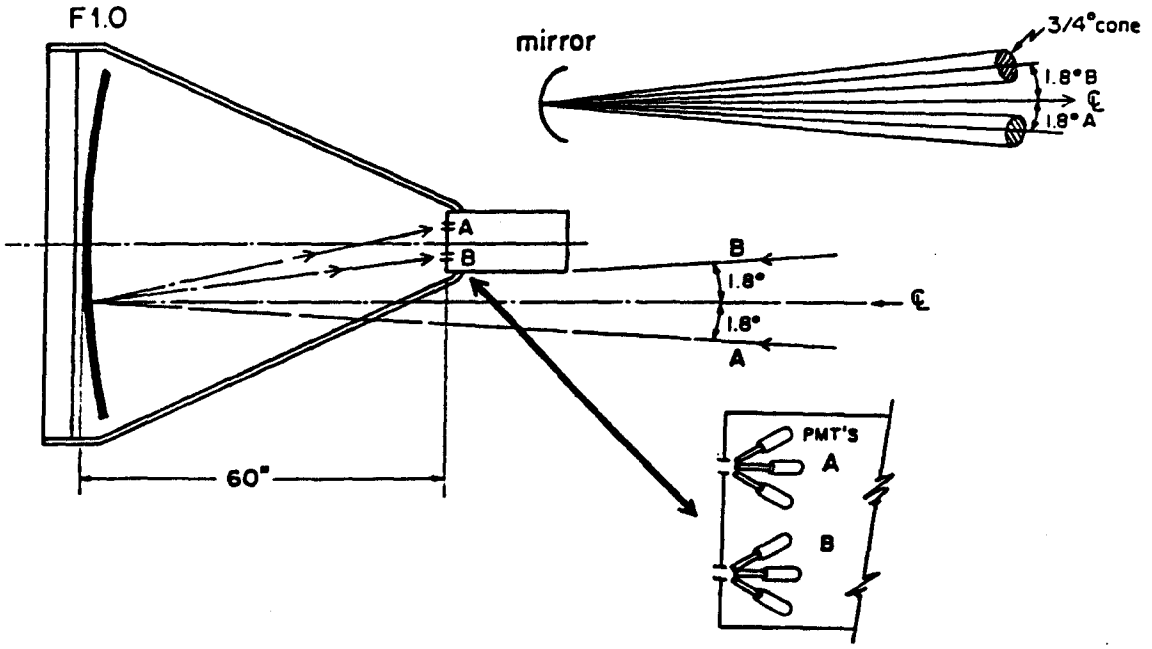


Figure 3.22: Schematic of mirror/PMT canister configuration illustrating the two separate apertures, each serviced by three PMT's.

as the mirror resolution. It should be noted, however, that future plans include the replacement of these mirrors with a large number of smaller, higher quality mirrors which are presently being fabricated by the collaboration.

Each mirror has, in its focal plane, an aluminum canister containing two 0°.75 apertures, separated by 3°.6 in declination (Figure 3.22). In practice, one aperture observes a source region while the other observes a "background" region. Often the role of on-source and off-source is switched periodically, by moving the declination axis, to eliminate systematic contributions to on-source/off-source differences. The mirrors each service six PMT's (three per aperture). Each PMT, then, views one-third of the mirror yielding a collection area of approxi-

mately 0.6 m^2. The focal plane scale of the apparatus is 2.62 cm deg^{-1}; apertures of 1.96 cm provide a 0°.75 acceptance cone. This size was chosen to strike a balance between sampling an ample portion of the Čerenkov pools and maintaining reasonable singles rates (produced by ambient skylight) in the PMT's. The 3°.6 separation of the apertures is large enough to eliminate cross-talk between the two channels while being small enough to minimize zenith-angle related differences between on-source and off-source.

The equatorial mount is driven by stepping motors through a gear drive. Each motor is under the control of a microprocessor which in turn receives its instructions from the main computer. During normal operation, the instructions provide for sidereal tracking through rotation of the right ascension axis as well as for periodic 3°.6 wobbles of the declination axis for on-source/off-source switching. Each axis is equipped with a shaft encoder whose angular resolution is 2.64 minutes of arc. The encoder values are written into each data record for both precautionary reasons and to facilitate identification of the wobble cycles. The mount is also equipped with a television camera which is used for rough siting and visual monitoring of the position of the telescope. The 0°.1 resolution of the camera is sufficient to guarantee proper pointing (through the monitoring of positions of guide stars) to well within the accuracy required by the ACT.

The photomultiplier tubes employed in the telescope are Hamamatsu model R1450. These 3/4-inch tubes have high quantum efficiency ($\sim$ 30% at 420 nm) and the 10-stage dynode system provides excellent time resolution ($\leq$ 1ns) with a gain of 10^6. Single photoelectrons produce anode signals of 2–3 mV amplitude

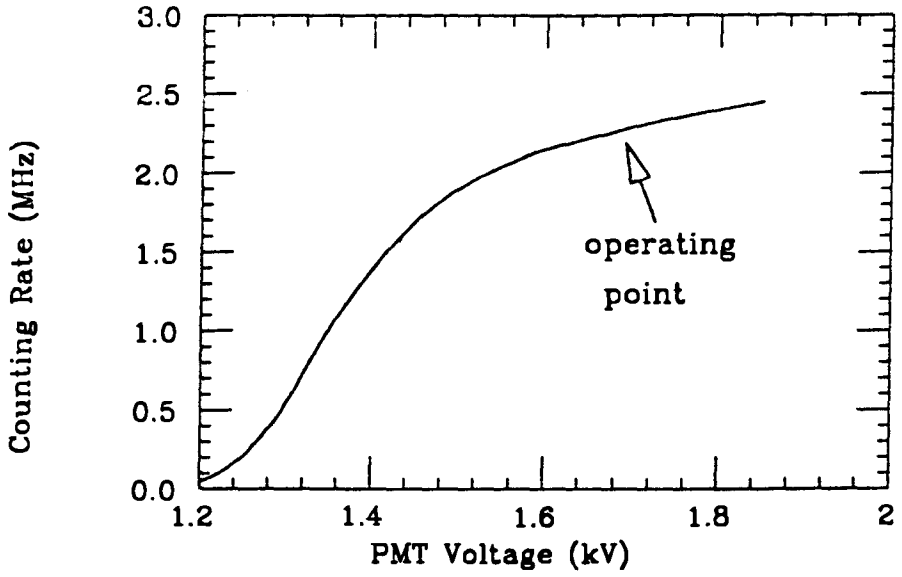


Figure 3.23: Typical "plateau" curve of PMT rate vs applied voltage.

with 3ns duration across a 50Ω load. Resolution is sufficient to identify the one- and two- photoelectron peaks in the PMT counting spectrum. The PMT's are operated at "plateau" voltages (Figure 3.23) to minimize gain fluctuations produced by any deviations in the operating voltages while providing maximum gain.

Optical filters (HOYA B-370), whose spectral characteristics are shown in Figure 3.24a, are used to confine the bandpass of the telescope to those wavelengths in which the Čerenkov radiation is most prominent. The wavelength spectra of both the Čerenkov radiation and the night sky background are illustrated in Figure 3.24b. From these plots, it is clear that the filters suppress the night sky background considerably. When the filter response is folded with

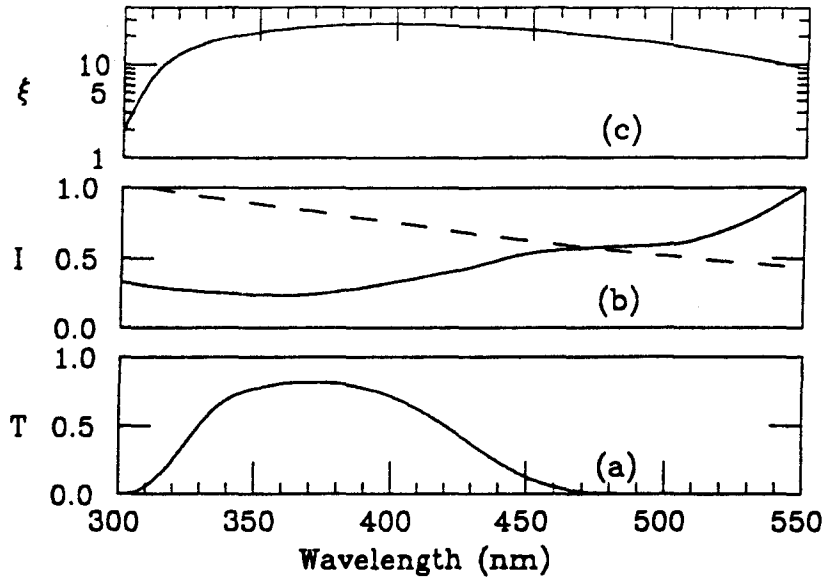


Figure 3.24: Spectral characteristics: a)Transmission coefficient T for HOYA B-370 filter; b)Čerenkov radiation (dashed curve) and night sky background intensity I (arbitrary scale); c)quantum efficiency ξ of Hamamatsu R-1450 PMT.

the quantum efficiency of the PMT's (whose spectral characteristics are shown in Figure 3.24c), the collection efficiency of the combination is $\sim$ 13% over the wavelength band 325nm $\leq \lambda \leq$ 425nm. The specular reflectivity of aluminum is $\sim$ 85%; a reasonable estimate for the mirrors on the telescope (which are directly exposed to the outside environment) is a reflectivity of $\sim$ 65%. Further, there are losses due to the incomplete overlap of the three PMT's in each aperture as well as due to the large angle of incidence of the photons which arrive from the outer regions of the $f/1$ mirrors. A reasonable overall estimate of the collection efficiency of the apparatus is $\sim$ 5%. Implementation of new mirrors in a configu-

ration requiring only one PMT per mirror could increase this efficiency by more than a factor of two.

Given the telescope parameters, we can determine the rate at which the PMT's will collect photons from the night sky background. At a dark observing site, the night sky brightness is approximately 22.5 magnitudes per square arcsecond, or -4.1 mag sr^{-1} (see, *e.g.* Zombeck, 1982). The site at Haleakala has very clear, dark skies so that this value should be a reasonable (or slightly optimistic) characterization of the background. We know that, if L represents the intensity in photons $m^{-2}\ s^{-1}\ \AA^{-1}$, then the associated magnitude is given by

$$m = -2.5 \log L + C \tag{3.39}$$

where C is a constant which is related to the bandwidth. By definition, $m = 0$ corresponds to an intensity of $(10^7\ m^{-2}\ s^{-1}\ \AA^{-1})\Delta\lambda$ centered at $\lambda = 560$ nm. Using the bandwidth of the telescope filters, $\Delta\lambda = 150$ nm, we find that $C = 25.5$. From Figure 3.24, however, we see that this bandwidth is centered at about 375 nm and that the night sky intensity at this wavelength is about a factor of 3 smaller than at 560 nm. Thus, the night sky intensity is approximately

$$L_{NS} = \left(\frac{1}{3}\right) 10^{\frac{C-m}{2.5}} = \left(\frac{1}{3}\right) 10^{\frac{25.5+4.1}{2.5}} = 2.3 \times 10^{11}\ m^{-2}\ s^{-1}\ sr^{-1}. \tag{3.40}$$

Given the efficiency of 5%, mirror area of 10.4 m^2, and solid angle of 1.3×10^{-4} sr, the expected rate in each of the 18 PMT's is about 1 Mhz. In practice, we find that the PMT's operate with rates between 1.5 and 2 Mhz, in good agreement with this approximation.

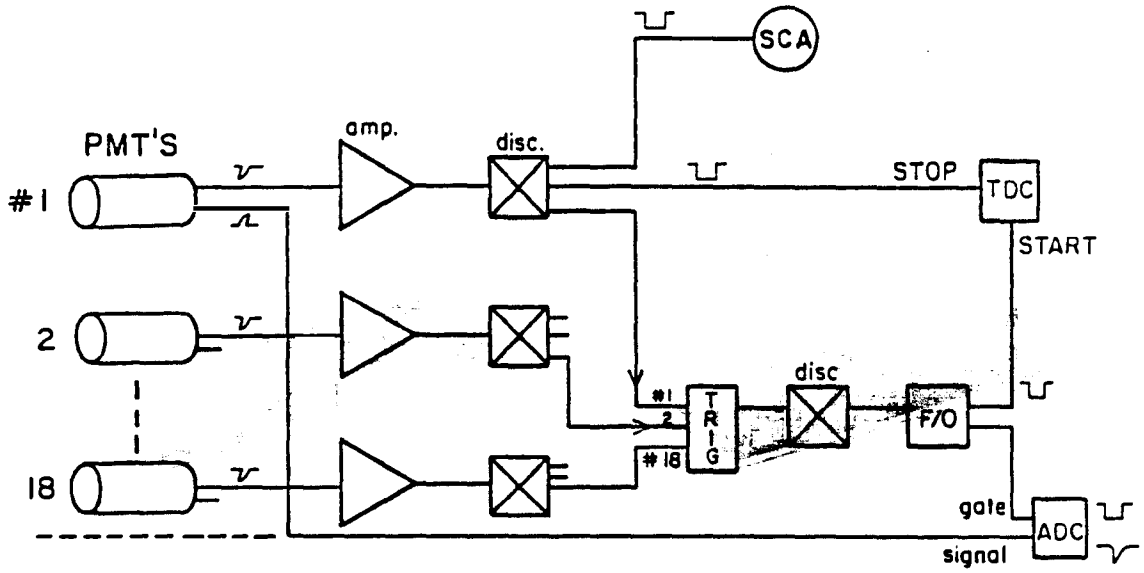


Figure 3.25: Schematic of data acquisition electronics.

3.2.3 Data Acquisition System

Trigger Processing

Each PMT in the telescope constitutes the start of a separate channel of electronics; the on-source and off-source groups of 18 channels are processed separately and identically. A schematic of the trigger electronics is shown in Figure 3.25. PMT pulses are directed from high speed amplifiers (bandwidth $\sim$ 1GHz, gain $\sim$ 40) to LRS 620BL discriminators where a threshold of 30 mV is used to pass only pulses corresponding to at least one photoelectron from the PMT cathode. The 18 discriminated pulses are then fed to a LRS 380 majority logic unit (MLU). The MLU is operated in a mode whereby the output has a pulse-height of $n \times 50$ mV where n is the number of overlapping input pulses. Because of the short du-

ration of the individual pulses, this overlap condition corresponds approximately to requiring the pulses to arrive within an effective gate of $\sim$ 7ns. A fan-out is then used to direct the MLU output to various discriminator channels whose thresholds are set to correspond to multiplicities of 4, 5, 6 etc. It should be noted that the MLU output is actually a response to an integration of the overlapping input pulses; as such, there are small ambiguities in the strict interpretation of the multiplicity that corresponds to a particular output. The operating point for the telescope trigger is chosen such that no more than $\sim$ 50% of the triggers taken correspond to events in which night-sky background (possibly in combination with light generated by a low energy shower) has satisfied the multiplicity requirement. This point can be estimated statistically as follows.

Let m represent the number of PMT's which can contribute to the trigger (in our case, $m = 18$) and let ζ be the counting rate of each PMT due to ambient light (which we will take to be the same for all tubes for simplicity). Since the ambient light consists of photons which arrive at random, the time intervals between hits are Poisson distributed. Hence, the probability that a given tube will fire in a given time window τ is given by $p = 1 - e^{-\zeta t} \approx \zeta t$ (for $\zeta t \ll 1$). Given that one tube is required to effectively start the coincidence gate, the probability of obtaining an additional $n - 1$ counts within the coincidence window is

$$P_{n-1,m-1} = \left(\begin{array}{c} m-1 \\ n-1 \end{array} \right) p^{n-1} \left(e^{-\zeta t} \right)^{m-n} \tag{3.41}$$

where $\left(\begin{array}{c} k \\ j \end{array} \right)$ is the binomial coefficient. Since the gates are produced at a rate

$r \approx m\zeta$, the random trigger rate for multiplicities greater than or equal to n is

$$R_{mn}(\zeta,\tau) = m\sum_{i=n}^{m}\left(\begin{array}{c} m-1 \\ i-1 \end{array}\right)\zeta^{i}\tau^{i-1}\left(e^{-\zeta\tau}\right)^{m-i}. \tag{3.42}$$

Now suppose that the rate for showers of a given photon density ρ is s. The probability z for a given PMT to fire is

$$z = 1 - e^{-a\rho} \tag{3.43}$$

where a is some constant. Further, assuming that s scales with density with the same spectral index as that for the rate as a function of energy (although this may be an unfair approximation - see Section 3.1.3), we have

$$s = K\rho^{-\gamma} \tag{3.44}$$

where $\gamma = 2.6$. Thus, integrating over all densities $\rho > \rho_{min}$, we find the rate for showers of multiplicity $\geq n$:

$$S_{mn} = K\int_{\rho_{min}}^{\infty}\rho^{-\gamma}\sum_{i=n}^{m}\left(\begin{array}{c} m \\ i \end{array}\right)(1-e^{-a\rho})^{i}(e^{-a\rho})^{m-i}d\rho. \tag{3.45}$$

The operating point of the telescope is chosen such that $R_{mn} \approx S_{mn}$. In this way, the bulk of coincidences due to ambient light pile-up are rejected while the majority due to actual showers are recorded. Figure 3.26 is a plot of the trigger rate as a function of coincidence multiplicity as determined by varying the threshold of the discriminator used to set the operating point (recall that the MLU output is $\approx$ 50mV $\times$ multiplicity). The solid lines are the rates expected from the expressions given above: the steeper line represents the rate for ambient

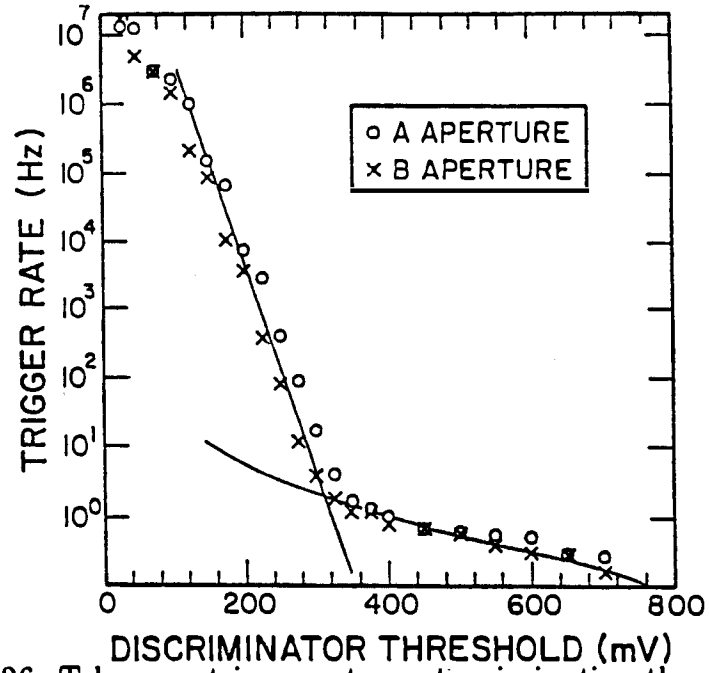


Figure 3.26: Telescope trigger rate vs discrimination threshold (multiplicity).

light triggers assuming $\zeta = 1.8$ Mhz, and $\tau = 8$ ns while the shallower line represents showers assuming a spectral index $\gamma = 2.6$.

In addition to being used for MLU triggering, the discriminated PMT pulses are used for time digitization of the individual pulse (with a TDC), for a separate timing latch (requiring an overlap with a hardware gate), and for recording the PMT singles rate (with a scaler). Once the MLU output passes its discriminator, a pulse is sent to an interrupt box which performs several functions. First, it sends an interrupt to the data-taking computer thus indicating both the existence and the origin (e.g. A-aperture, B-aperture, etc.) of the trigger; the computer then executes op-code based on the trigger type in order to read out the system clock and the data electronics. In addition, the interrupt box freezes

the electronics so that no additional triggers may be accepted until the current data has been read. Lastly, a register of environmental information (e.g. clock status, high voltage status) is passed through the interrupt box to the computer. The output of the discriminated MLU pulse is used to provide gates to the Latch registers and stop pulses to the TDC's. In addition, it provides gates to ADC's which integrate pulses taken from the last two dynodes of the PMT as a measure of the pulse height.

For a monitor of the system, environmental triggers are generated every second. In these triggers, the clock, encoders, and environmental registers are read as with any event. Scalers which are gated for 100 ms are used to monitor the PMT singles rates. Along with the PMT pulses, a 1Mhz signal is scaled to provide live-time measurements.

Low multiplicity events, as determined by discrimination of the MLU output at a low threshold, are also recorded to provide events with which to test off-line software algorithms designed to reject such triggers. These are usually taken at multiplicities $n \geq 5$; such events are pre-scaled by a large factor to maintain reasonable rates but are otherwise treated as normal triggers.

Data Logging

The data recorded by the acquisition electronics is logged onto tape through a DEC LSI11/73 computer under the control of RSX/MULTI. The CAMAC electronics are read through a Jorway CAMAC-11/73 interface via Direct Memory Access. The encoders and clock are read into memory through a direct DRV 11-

J parallel interface. The clock itself operates with a 1 Mhz cesium beam signal provided by the nearby Lunar Ranging Experiment; because this signal is periodically calibrated to UTC by Bendix Time Systems Inc., the clock is accurate to within several μs per year. The data is written onto a 9-track magnetic tape with a density of 1600 bpi.

On-Line Monitoring Facilities

To facilitate the operation of the system, various on-line monitoring procedures are employed. Through MULTI, watch blocks are used to warn the operator in the event that any PMT rate falls outside of a pre-set range or if the CAMAC branch driver notes that a problem has occurred in reading the electronics. On-line histograms are used to monitor the trigger rate and multiplicity distributions of the recorded triggers. Other histograms, such as TDC or ADC distributions can easily be generated by the operator and are frequently used to check the behavior of individual channels of electronics.

Individual PMT anode currents are monitored via scanning ADC's such that, if any exceeds a pre-selected threshold, the high-voltage to the PMT's is shut down, an alarm is sounded, and a bit is set for the data record as an indicator that the voltage is off. In this way the PMT's are spared from the excessive anode currents that can occur if, for example, nearby lights are accidently turned on or a bright star crosses the telescope aperture (*e.g.* during a wobble cycle).

The input to the clock is monitored so that a loss of the cesium signal automatically sounds an alarm and sets a flag in the data record. The alarm is

also set if the one-second marker that is provided with the 1 MHz signal is in disagreement with the clock. These precautions are primarily to watch for catastrophic events such as lightning glitches or power failures (which may occur while the operator is away). The μs timing can be checked by inspection of the times of environmental triggers; the time differences should be multiples of 1s down to the μs level.

3.2.4 Systematic Data Analysis

Inspection and Calibration

Prior to analyzing data for signatures of activity from the source under study, various calibrations and systematic checks are made to ensure that the data is of good quality. As a first step, individual data records are inspected. Occasionally we have found instances in which a particular CAMAC module has, for unknown reasons, not given a proper Q-response (*i.e.* "ready" response) when a read is attempted. Because CAMAC reads are performed with Q-scans (whereby modules are read sequentially until the specified number of words have been read) to reduce the time needed to read the data, any module failing to give a Q-response is merely skipped over. Unfortunately, data from the next available module is then stored at the position in the data record that would have normally contained information from the skipped module. This can be confusing but can be identified in several ways. First of all, the data record should not have the expected length if a module is skipped; such a signature is easy to monitor. In Figure 3.27 we plot the word lengths of individual data records for a particular

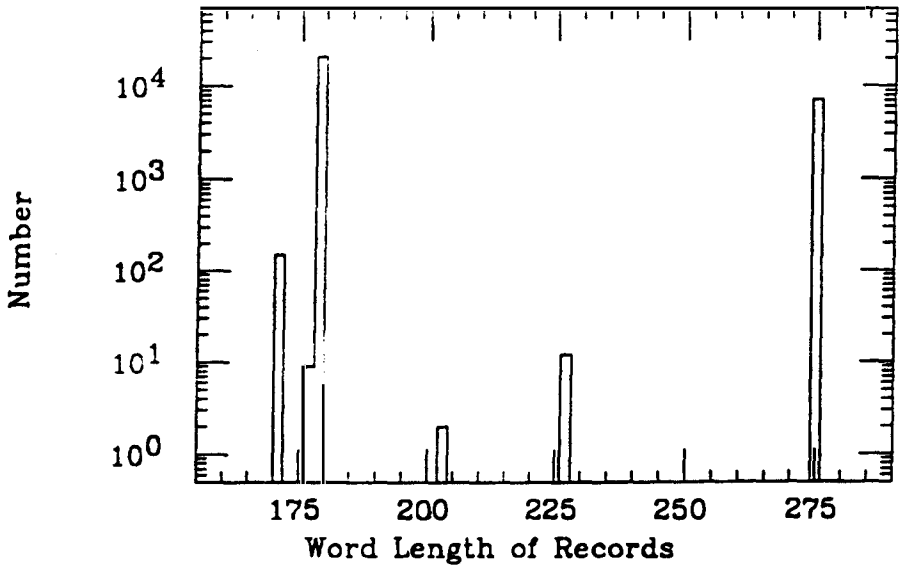


Figure 3.27: Word lengths of individual data records. Evidence (at the < 1% level) for module skipping is contained in the bins with incorrect word lengths (see text for explanation).

run. Normal events contain 179 words, pre-scaled events contain 177 words, and environmental records contain 275 words. Note that the bad events contain word counts that differ from the correct values by multiples of 8, 12, or 16 - corresponding to the 8 word TDC's, 12 word ADC's, and 16 word scalers. As an added precaution, pattern units have been hard-wired to establish a set configuration of 0's and 1's at several positions in the CAMAC crates so that if the pattern is not properly placed in the data record, the event can be skipped. Overall, these sorts of bad events constitute less than 1% of the data and pose no serious problem.

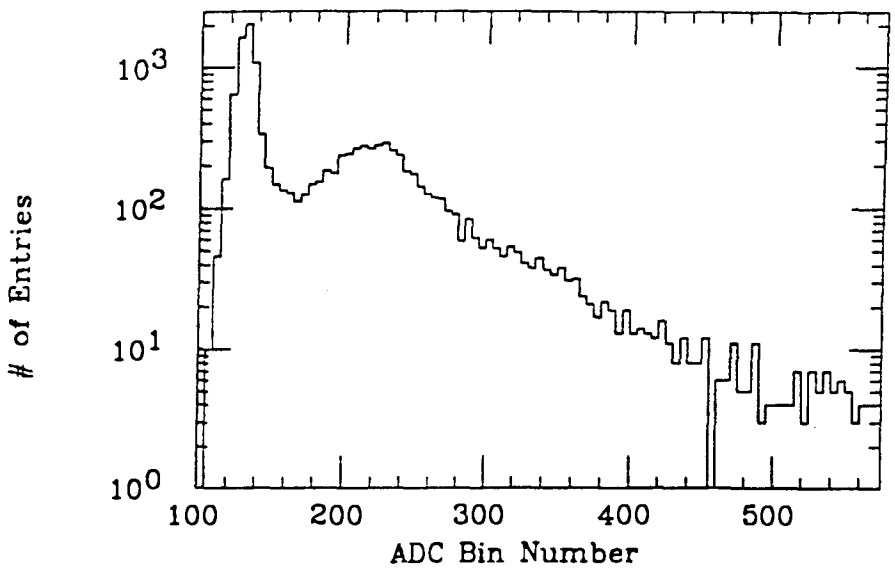
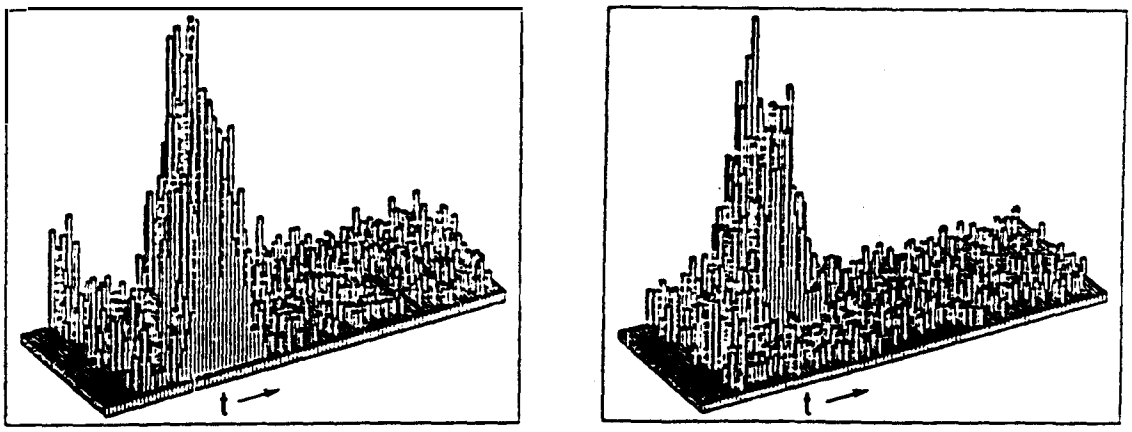


Figure 3.28: ADC pedestal and single-photoelectron peak for a typical channel.

For calibration purposes, the ADC pedestal and single photoelectron values are determined by taking averages over several runs in a given observation period. The pedestal values are determined by averaging the ADC values for PMT hits which occur in the off-trigger data (i.e. data from the aperture that did not produce the trigger) and which are accompanied by TDC values that indicate that no stop pulse occurred. This is performed on a tube-by-tube basis. Since any tubes that do record a hit in the off-trigger are most likely to have recorded only a single photon, the single photoelectron value may be determined by averaging the ADC values for latched off-trigger tubes. Figure 3.28 illustrates the positions of the pedestal and single-photoelectron values for a typical channel. These calculated values are quite stable from month to month unless PMT or amplifier

replacement has occurred.

Because TDC times are used to study the time-evolution of the hit pattern recorded by the telescope (in part to help identify triggers caused primarily by ambient light, as mentioned above) it is important that relative times as recorded by the TDC's are meaningful. While hardware timing has been carefully adjusted so that pulses which are physically in-time will arrive at the TDC's simultaneously, throughput differences and pulse-height slewing (large pulses, with their associated faster rise-times, will fire a discriminator faster than small pulses) can cause relative timing errors. These errors must be identified and removed in software. In Figures 3.29a) and b), the distributions of TDC times for individual channels are shown for triggers originating in the A and B apertures of the telescope, respectively. It is clear that, when a trigger occurs in a particular aperture, the TDC times for tubes corresponding to that aperture are grouped in a cluster whose duration is several nanoseconds while those for tubes in the other aperture have basically random distributions. Figure 3.30a shows TDC mean times for all PMT's for various multiplicities. Variations as large as 5 ns can be seen. Figure 3.30b shows the TDC mean times after slewing corrections based on the calibrated ADC's have been made. After correction, the relative timing is accurate to within about 1ns. One can note that the TDC means are larger for high multiplicity events than for low multiplicities. This is simply the result of the trigger forming more quickly for the high multiplicity events; since the delay lines carrying the TDC stop pulses are of fixed length, a faster start time yields a larger count in the TDC's before the stop pulses arrive.



(a) (b)

Figure 3.29: Distribution of TDC stop times for A aperture (a) and B aperture (b) events. Note that the times for tubes in the triggering aperture cluster together while those in the opposite aperture are randomly distributed.

Rates

Trigger rates as well as individual scaler rates are used for each run to identify effects related to clouds or nearby lights. Figure 3.31 shows the scaler rates for the various tubes as a function of time. Note that, although there are differences in the individual tube rates, the rates are quite stable. One can note a slow increase in the rates which is due to the first signs of dawn toward the end of a run. Also evident is a large drop very early in the run which was the result of a high-voltage trip in response to some momentary bright lights near the telescope. Figure 3.32a shows the trigger rate for events in the A aperture of the telescope.

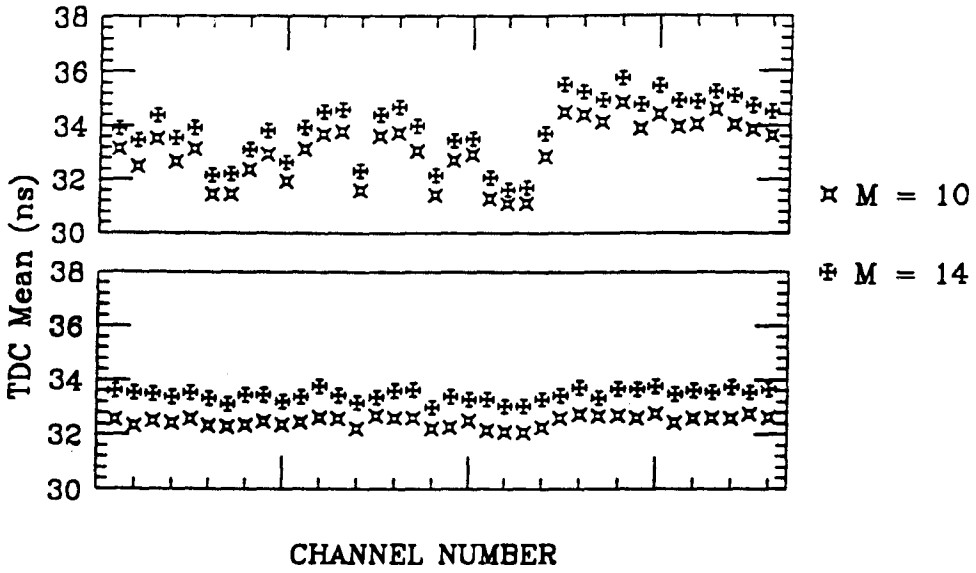


Figure 3.30: Means of TDC stop times during trigger for various multiplicities, before (a) and after (b) slewing corrections.

The rapid increase in the rate toward the end of the run is the result of increased ambient light caused by the rise of the sun.

Triggers which are due to cosmic ray showers can be separated from those caused by ambient light by requiring a tight coincidence of TDC times. Figure 3.33b shows the data from Figure 3.32a after a software cut requiring at least 9 of the 18 tubes to have fired within a 5ns window. Note that the sunrise effect is completely removed by such a cut. One should note that, because the TDC means are different for different multiplicities (as noted above) a sliding time window is used to look for tight clusters of TDC hits. The shower rate that results from the above multiplicity cut is quite steady from night to night. The

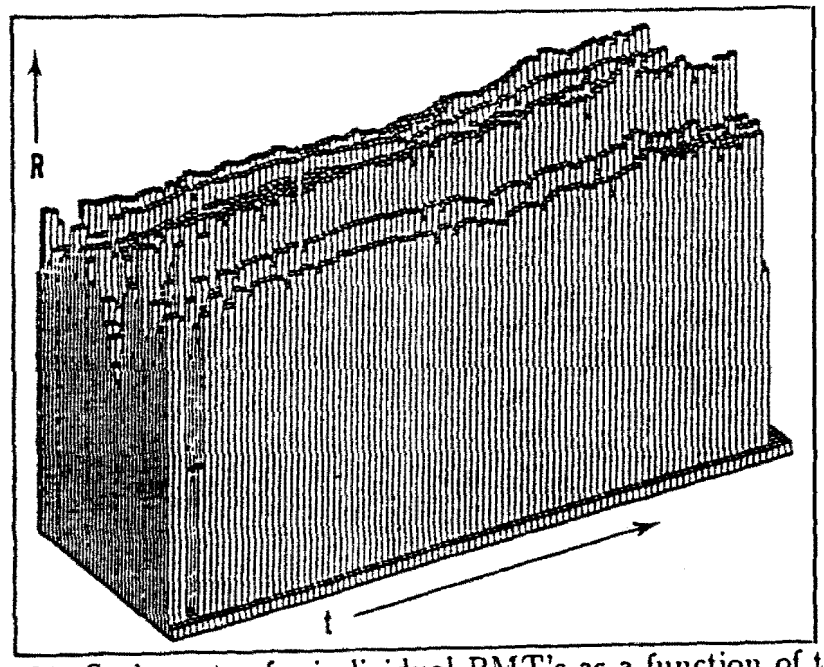


Figure 3.31: Scaler rates for individual PMT's as a function of time.

rate is a function of zenith angle as shown in Figure 3.33. In order to establish a baseline about which to identify rate excesses, the zenith angle dependence is fit to a three-term Legendre polynomial expansion as shown in the figure (along with the corresponding χ^2 for the fit). Using the fit, the expected rate for any given interval of data can be calculated (based on the zenith angle) and compared with the actual rate obtained. Inspection of Figure 3.33 indicates that, after performing the shower cut described, the behavior of the two telescopes (*i.e.* those corresponding to the two separate apertures) is very nearly identical. Given this cut as an operating point for the telescope, we can then define the critical photon density required from a shower. We have seen that the overall light collection efficiency of the apparatus is $\approx 5\%$. Because the average occupancy

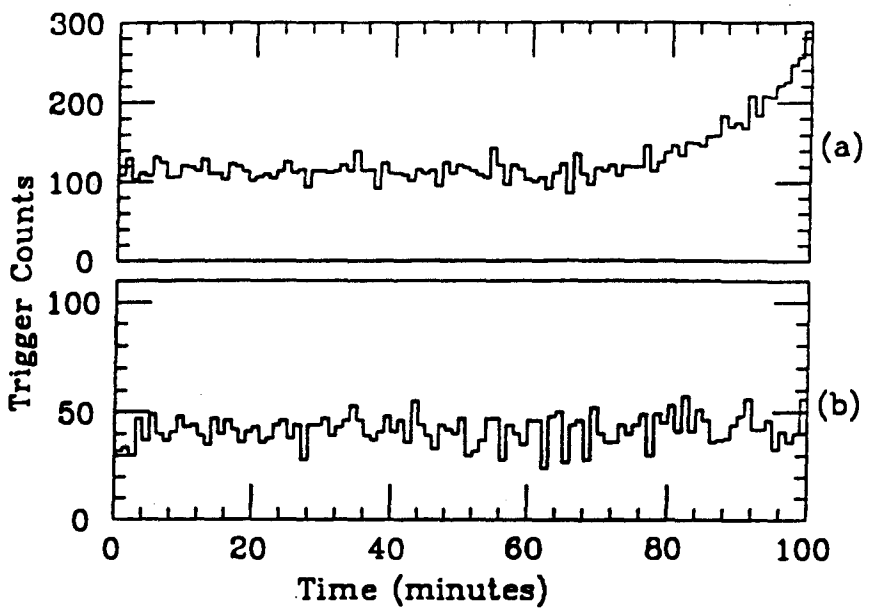


Figure 3.32: Telescope trigger rate: a)raw triggers; b)triggers surviving a cut requiring 9 or more TDC hits in a sliding 5ns window. The increasing rate in the raw triggers is due to a brightening of the sky at dawn; such contributions from random skylight are removed with the multiplicity cut.

of the PMT's is slightly greater than one for multiplicities larger than 9, we assume that we need at least 10 photoelectrons to yield a multiplicity sufficient to pass the shower cut. Given the 10.4 m^2 of telescope area, then, the minimum density is approximately 20 photons m^{-2} (which is the value which was used in the calculation of Section 3.1.3).

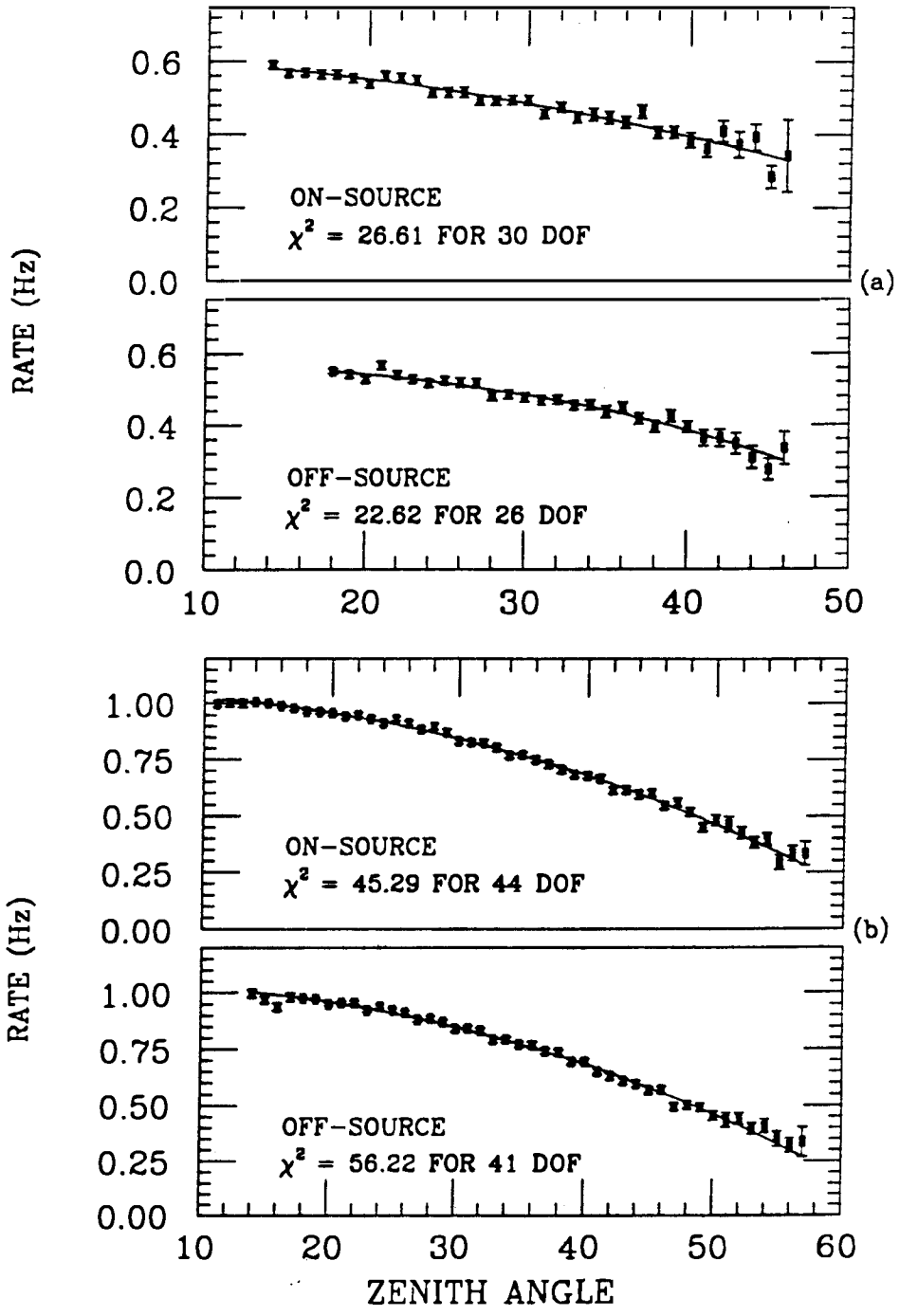


Figure 3.33: Zenith angle dependence of trigger rate for 0°.75 apertures (a) and 1° apertures (b). The χ^2 for the three term Legendre polynomial expansion used to fit the data is listed for each plot.

Chapter 4

Time Series Analysis for Periodic Behavior

What a tangled web we weave.
Go 'round in circumstance...
Someone show me how to tell
The dancer from the dance.
The Eagles

The search for periodicity in a set of arrival times has many angles of approach. This chapter summarizes several of these methods with particular attention paid to the statistical interpretation of the results of the tests as well as inherent difficulties associated with each. Before any attempt at detecting periodicity is made, however, it is crucial to remove any contributions to the measured arrival times that have been introduced by time-dependent relative motion between the source and observer. Our discussion, therefore, will begin with an investigation of the detailed procedure of reducing times to the respective barycenters of the source and observer systems.

4.1 Barycentering

The necessity of reducing times to the barycenter can be appreciated by considering the result of Doppler effects caused by the movements of source and observer. One may treat the variations in the period P of the source, caused by Doppler shifts, as an effective $\dot{P}_e$ (not to be confused with the period derivative related to the spin-up of the pulsar which, for the moment, is being ignored). Arrival times may be assigned a relative phase between 0 and 1 based on the period of the source:

$$\phi(t) - \phi(t_0) = \int_{t_0}^{t} \frac{dt'}{P(t')} \tag{4.1}$$

Treating $\dot{P}_e$ as constant over the time of observation, this becomes

$$\phi(t) - \phi(t_0) = \int_{t_0}^{t} \frac{dt'}{P_0 + \dot{P}_e t'} = \frac{1}{\dot{P}_e} \ln\left(P_0 + \dot{P}_e t\right)\Big|_{t_0}^{t} \tag{4.2}$$

We may arbitrarily set $t_0 = 0$ and $\phi(t_0) = 0$ to obtain

$$\phi(t) = \frac{1}{\dot{P}_e}\left[\ln(P_0 + \dot{P}_e t) - \ln P_0\right] = \frac{1}{\dot{P}_e} \ln\left(1 + \frac{\dot{P}_e t}{P_0}\right). \tag{4.3}$$

Now if $\dot{P}_e = 0$, the phase $\phi(t)$ is just

$$\phi(t) = \frac{t}{P_0}. \tag{4.4}$$

Thus, the difference in phase introduced by $\dot{P}_e$ is

$$\Delta\phi = \frac{t}{P_0} - \frac{1}{\dot{P}_0} \ln\left(1 + \frac{\dot{P}_e t}{P_0}\right) \approx \frac{1}{2}\dot{P}_e \frac{t^2}{P_0^2}. \tag{4.5}$$

As an example, consider data taken from Her X-1 ($P_0 \approx 1.24$s) at intervals separated by 48 hours; to maintain true phase to within $\Delta\phi \leq .1$, we must have

$$\dot{P}_e = \frac{.2P_0^2}{t^2} \leq 10^{-11}$$

The value of $\dot{P}_e$ introduced by the orbit of the earth may be approximated by

$$v = \frac{2\pi r}{T} \approx 3 \times 10^4 \text{ m s}^{-1}$$

$$\frac{\triangle P}{P} = \frac{v}{c} \Rightarrow \triangle P = \frac{v}{c} P \sim 10^{-4}\text{s}.$$

$\triangle P$ will vary sinusoidally with this amplitude:

$$\triangle P(t) \approx (10^{-4}\text{s}) \ \sin\theta$$

where θ is the angular position of the earth in its orbit. Thus,

$$\dot{P}_e = \frac{d(\triangle P)}{dt} = (10^{-4}\text{s}) \ \cos\theta \frac{d\theta}{dt}.$$

But

$$\frac{d\theta}{dt} = \frac{2\pi}{T} \Rightarrow \dot{P}_{emax} \approx 2 \times 10^{-11}.$$

Clearly, this is larger than our analysis can tolerate. While the data is generally broken into intervals much shorter than 48 hours, the barycentric corrections are necessary if phases are to be linked between different intervals. Similar calculations regarding the orbital motion of the source show that timing corrections to remove these effects are necessary as well.

4.1.1 Solar System Barycentering

Signal arrival times at the solar system barycenter are given by

$$t_b = t_s + \triangle t_r + \frac{1}{c}\mathbf{r} \cdot \mathbf{n} \tag{4.6}$$

where t_s is the observed arrival time, $\triangle t_r$ is a general relativistic correction to the clock time required to account for the variations in gravitational potential

around the Earth's orbit, r is the vector from the solar system barycenter to the observation site, and n is the unit vector from the barycenter in the direction of the source. The event times, as recorded by the observatory clock, are in coordinated universal time (UTC). As outlined in the *Astronomical Almanac*, to include the relativistic effects, that is, to convert to barycentric dynamical time (TDB) one needs to first convert to terrestrial dynamical time (TDT). This is accomplished by converting UTC to international atomic time (TAI) through the addition of a given number of leap seconds, depending upon the date, and then to TDT by adding exactly $32^s.184$ (an offset needed to provide continuity with a convention used prior to 1984);

$$TDT = UTC + 32^s.184 + \text{leap seconds.} \tag{4.7}$$

Then

$$TDB = TDT + 0^s.001658 \sin g + 0^s.000014 \sin 2g \tag{4.8}$$

where g is the mean anomaly of the Earth in its orbit,

$$g = 357^\circ.53 + 0^\circ.98560028(JD - 2451545.0) \tag{4.9}$$

and JD = Julian day (in TDT system) of the observation.

The values for the positions of the solar system barycenter and of the Earth have been calculated by a simultaneous numerical integration at the Jet Propulsion Laboratory in cooperation with the U.S. Naval Observatory. The correction term $\frac{1}{c}\mathbf{r} \cdot \mathbf{n}$, corresponding to propagation time to the barycenter, is determined using these values along with the coordinates of the source (precessed to epoch J2000.0, as per the IAU standard adopted in 1984, and consistent with the epoch

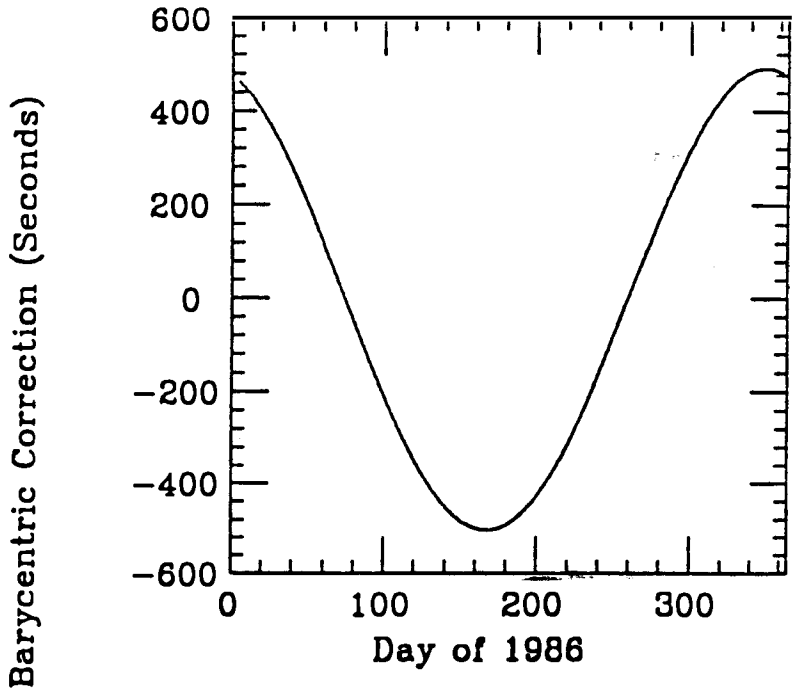


Figure 4.1: Barycentric correction as a function of time for plane waves arriving from the direction of Her X-1.

of the JPL positions). The correction has an annual sinusoidal variation with an amplitude of nearly 500s. Accurate positions for the source are thus imperative; an error as small as 0".1 can lead to errors as large as

$$t_2 - t_1 \sim 500\text{s}\ [\sin(\alpha + 0".1) - \sin\alpha] \simeq 500\text{s}\ \cos\alpha \sin(0".1)$$

so that

$$(t_2 - t_1)_{max} \sim 500\text{s}\ \sin(0".1) \approx 250\mu\text{s}.$$

A routine for determining the barycentric correction has been written by the collaboration; a plot of the correction as a function of time in 1986 is shown in

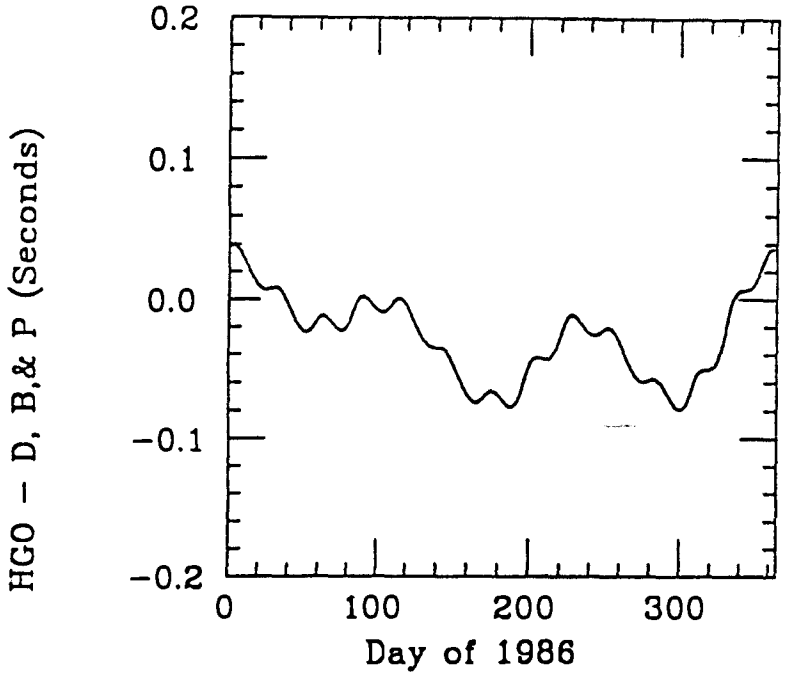


Figure 4.2: Comparison of barycenter routines. The difference between the correction as calculated by the Haleakala (HGO) routine and that of Deeter, Boynton, & Pravdo (DB&P).

Figure 4.1 for the direction of Her X-1 (RA $16^h\ 57^m\ 21^s$, Dec 35° 21′ 44″). An empirical formula for this correction has been introduced by Deeter, Boynton, and Pravdo (1981). As a check of the Haleakala routine, the difference between its calculated correction and that given by DBP is plotted in Figure 4.2. Lunar variations, as well as an annual and a ~120 day variation in the difference is obvious; these result from simplifications inherent in the DBP formula. As a further check, the routine was compared with an independent routine written at MIT (using the older MIT ephemeris); the differences are plotted in Figure

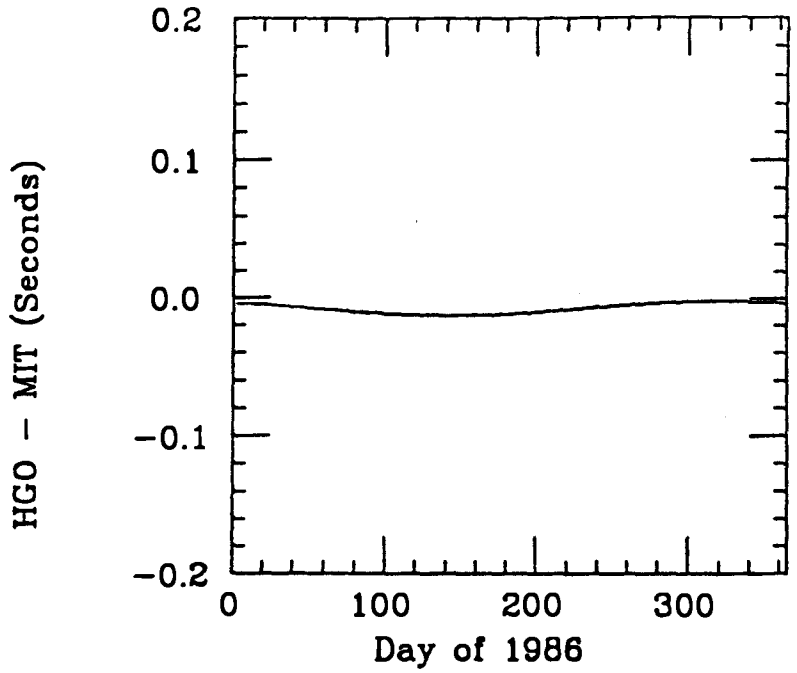


Figure 4.3: Comparison of barycenter routines. The difference between the correction as calculated by the Haleakala (HGO) routine and a routine using the MIT ephemeris.

4.3. For the purposes of this analysis, the small discrepancy between the two calculations is negligible. It should be noted, however, that the discrepancy is large enough to be worrisome for such calculations as the absolute phase of a millisecond pulsar.

4.1.2 Barycentering in the Her X-1/HZ Herculis System

In order to account for the changing position of the source in its orbit, an additional correction to event times must be included. For Her X-1, the problem is

simplified by the small eccentricity of its orbit ($e < 0.0003$). The general calculation for non-circular orbits is more complicated and is outlined in Appendix B. The orbital parameters for Her X-1 are listed in Table 5.1. Given an observation time t_{obs}, one can calculate the orbital phase of Her X-1;

$$\phi_{orb} = 2\pi(\frac{t_{obs} - t_0}{P}) \tag{4.10}$$

where t_0 is the epoch of zero phase (it is assumed that both t_{obs} and t_0 are solar-system barycentered times). The correction required to account for the time delay (or advance) relative to the barycenter of the binary system is then

$$\triangle t_{source} = a \sin i \cos(2\pi\phi_{orb}) \tag{4.11}$$

where $a \sin i$ is the radius of the Her X-1 orbit projected onto the plane of sight. Since $\phi_{orb} = 0$ is defined at the center of eclipse, $\triangle t_{source}$ must be subtracted from the observation time (*i.e.* at eclipse, the signal travel time is longer than the propagation time from the system barycenter).

It should be noted that the pulsar period P is not constant but, rather, shows the spin-up behavior characteristic of accreting binary systems (see Section 2.4.2). Thus, the pulsar phase calculations, even after barycentric corrections, must account for $\dot{P}$. This can be accomplished by using Equation 4.3 although problems may arise in the case of Her X-1 for which $\dot{P}$ can change significantly over relatively short periods of time (see Section 5.1) making phase-linking over intervals much longer than one month difficult.

4.2 Epoch Folding

In virtually any test for periodicity, time series data is folded modulo the period to be tested. This is conveniently done by assigning phases to the individual times. In the absence of any previous measurements of the source period, the choice of $\phi(t_0) = 0$ is arbitrary. In practice, it is convenient to choose t_0 as some fixed reference so that phases calculated for different data sets may be compared in a meaningful way. Of course, if an ephemeris for the source is already established, t_0 should be taken as the epoch of zero phase. Then we have

$$\phi(t_i) = \frac{t_i - t_0}{P} \tag{4.12}$$

assuming $\dot{P}$ is unknown. If $\dot{P}$ is known, the phase is determined by Equation 4.3. Once the times have all been converted to phases, a histogram of number vs phase bin may be plotted. Such a process is referred to as "epoch folding." In the absence of a signal, the resulting histogram should be uniformly populated within Poisson errors. Various statistical tests may be used to investigate the uniformity of the folded data. The χ^2 test, for example, can be used for a measure of the deviation of the folded data from that expected from a random distribution. Similarly, the histogram can be scanned for individual bins which deviate significantly from the mean. Both of these tests (which, one must note, are not mathematically independent) are attractive in that their statistical behaviors are well known and understood. They are not problem-free, however; as with any test which relies on binned data, the result can often be strongly dependent on the exact binning (*e.g.* the number of bins used). One must be careful to keep track of how many different arrangements of the data are used in

testing for a signal and weight the final statistics accordingly.

An advantage to the epoch folding approach is that one can test the folded distribution against non-uniform shapes as well. Thus, if the shape of the expected distribution is known (from a previous measurement, for example) one can perform least-squares analysis on the fit of the data to the expected distribution. Of course, proper phasing of the events relative to the epoch of the previous measurement then becomes crucial.

4.3 The Fast Fourier Transform

Any time domain function $f(t)$ can be decomposed into a combination of sinusoids of different frequencies ω. The Fourier transform $F(\omega)$ expresses this decomposition; $F(\omega)$ represents the amplitude of the sinusoid of frequency ω which is contained in $f(t)$. Mathematically,

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt. \tag{4.13}$$

In the case of discrete sampling of $f(t)$ with N data points in the interval $(0, T)$, we have the Discrete Fourier Transform (DFT)

$$F(\omega) = \int_{-\infty}^{\infty} a_j\delta(t - t_j)e^{-i\omega t}dt \tag{4.14}$$

where $a_j \equiv f(t_j)$. Hence,

$$F(\omega) = \sum_{j=0}^{N-1} a_j e^{-i\omega t_j}. \tag{4.15}$$

While the frequency spectrum can be calculated directly from the DFT, the process is computer intensive; if amplitudes are desired for N frequencies, the

computation time will be proportional to N^2. With large data sets the process can be quite time consuming. An algorithm for calculating the DFT which makes use of the cyclic nature of the $\exp(-i\omega t_j)$ factors (Cooley and Tukey, 1965) can reduce the computation by a factor of $\approx N/\log_2 N$. Since, for large values of N, this algorithm is much faster than the standard DFT, it is known as the Fast Fourier Transform (FFT).

The standard form of the FFT requires breaking the data set $f(t_j)$ into 2^n segments of equal length $\triangle t$ (although there exist modified algorithms which do not require that the number of segments be a simple power of 2). Such truncation and finite sampling in the time domain can cause spectral leakage in the the frequency domain [see, *e.g.* Brigham (1974)]. This can be understood in terms of the Fourier Convolution Theorem. The convolution of two functions $x(t)$ and $y(t)$ is defined by

$$h(t) = x(t) \circ y(t) \equiv \int_{-\infty}^{\infty} x(\tau)y(t-\tau)d\tau. \tag{4.16}$$

According to the convolution theorem, if $f(t) = f_1(t)f_2(t)$ then

$$F(\omega) = F_1(\omega) \circ F_2(\omega) \tag{4.17}$$

where F, F_1, and F_2 are the respective Fourier transforms of f, f_1, and f_2. This process is illustrated in Figure 4.4.

4.3.1 Aliasing

An immediate result of the convolution theorem occurs in the instance of a uniformly sampled waveform. Suppose the function $f(t)$ is sampled at regular

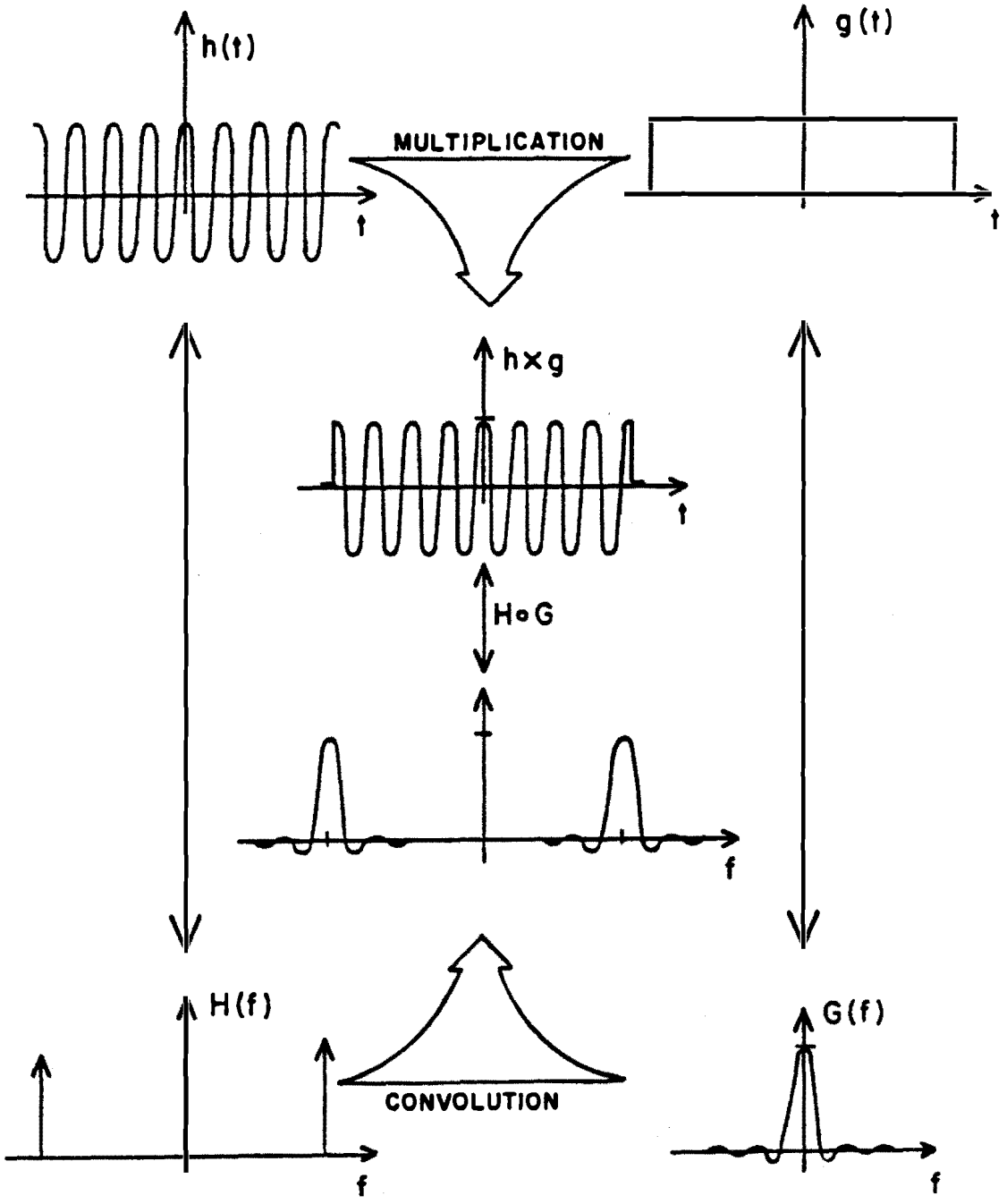


Figure 4.4: Graphical example of the frequency convolution theorem. Double-ended arrows connect Fourier transform pairs.

intervals of spacing Δt. The Fourier transform of such a sampling function in the time domain is a similar function in the frequency domain with discrete frequencies separated by

$$\Delta f = \frac{1}{\Delta t}.$$

When the transform of the unsampled $f(t)$ is convolved with this function in the frequency domain, the result is a superposition of $F(\omega)$ at each discrete frequency (Figure 4.5). Such distortion of the power spectrum is known as *aliasing* and is the result of choosing too large a sampling interval. Clearly, if Δt is chosen to be smaller, the resulting convolution will still yield power at frequencies outside of the true spectral range, but now the values in the true range are undistorted (Figure 4.6). The maximum spacing with which distortion by aliasing may be avoided is $\Delta t = 1/2f_c$ where f_c is the highest frequency component contained in $F(\omega)$. Of course, when the frequency spectrum of $f(t)$ is unknown, the choice of Δt is not so simple. In the FFT, the power spectrum is calculated only for frequencies up to the *Nyquist frequency*,

$$f_{Ny} = \frac{N}{2T} = \frac{1}{\Delta t} \tag{4.18}$$

where N is the total number of sample intervals and T is the duration of the full data set to be analyzed. Above the Nyquist frequency, aliasing occurs. As a result of this effective bandpass, Δt is generally chosen such that the Nyquist frequency is somewhat larger than the highest frequency of interest in the spectrum.

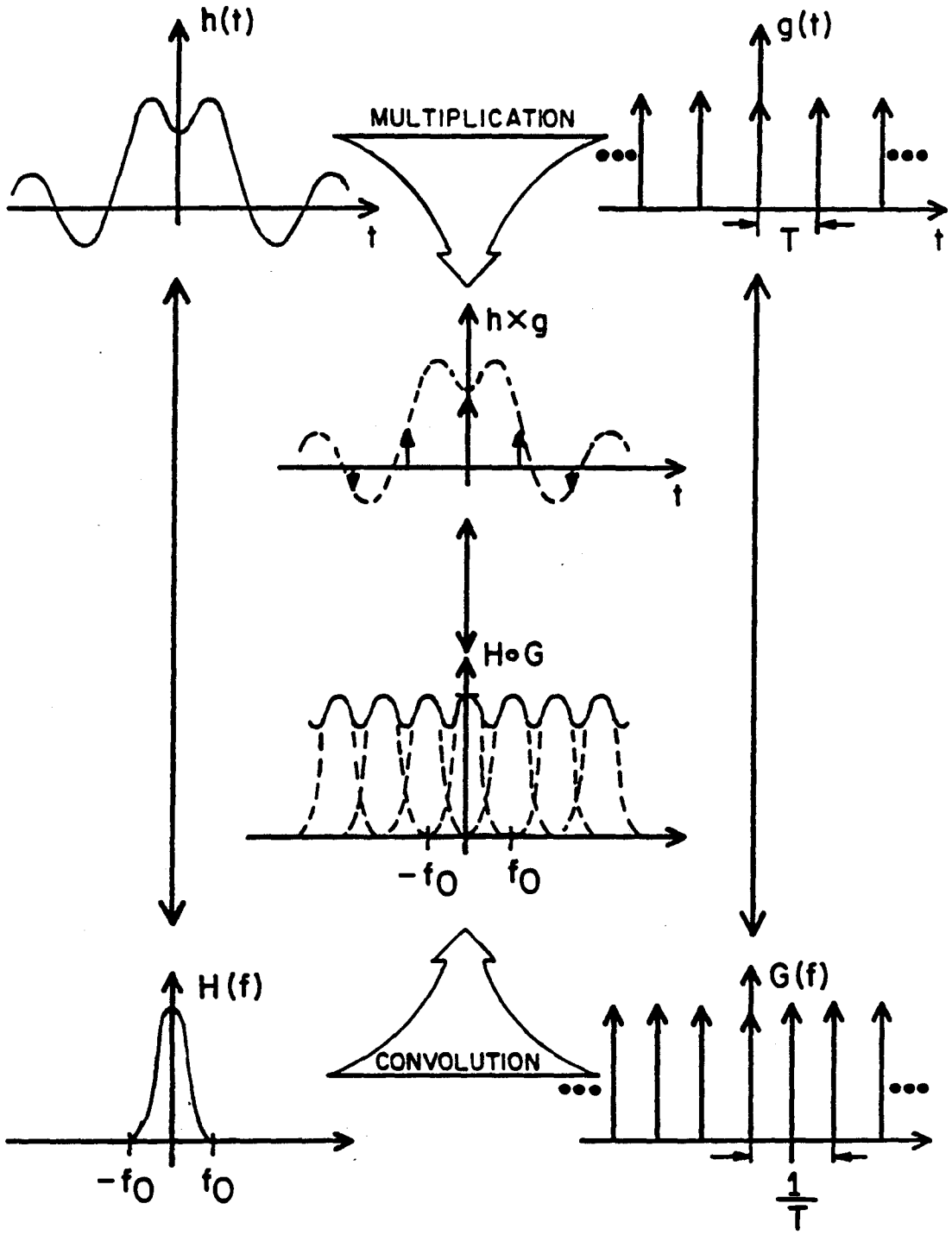


Figure 4.5: Graphical illustration of discrete, regular sampling of a function in the time domain resulting in *aliasing* in the frequency domain.

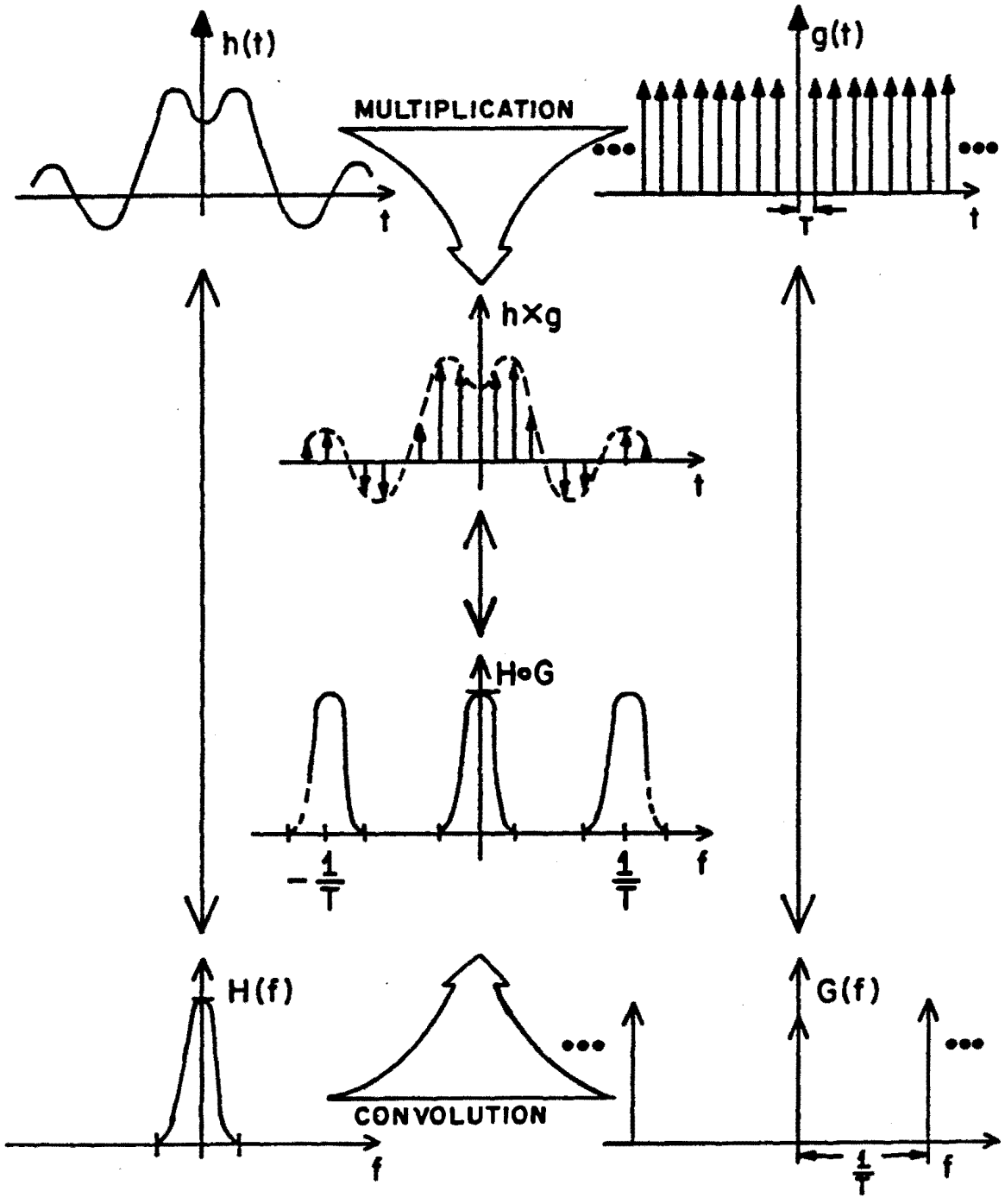


Figure 4.6: Aliased Fourier transform of a waveform sampled at a high enough rate to remove distortion in the frequency band of interest.

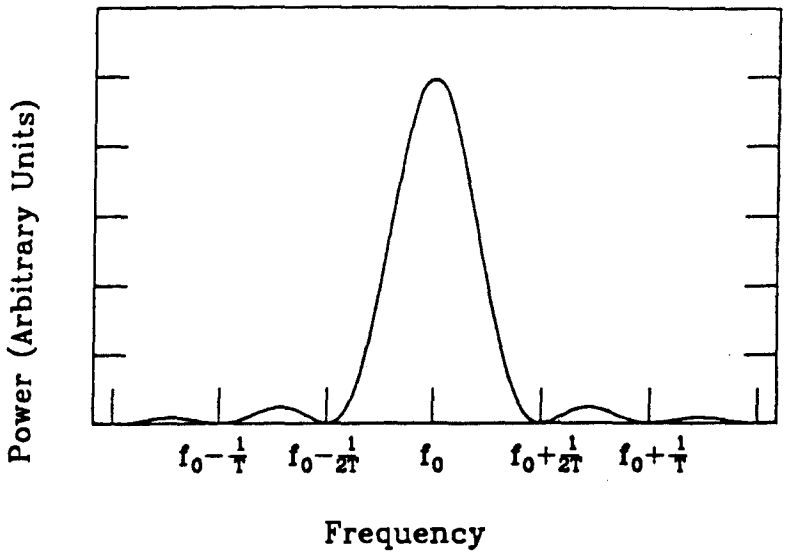


Figure 4.7: Illustration of *sidelobes* produced as a result of sampling for a finite duration in the time domain.

4.3.2 Sidelobes and Window Functions

Truncation of a waveform can also cause spectral leakage in the frequency domain. If the function $f(t)$ is truncated by a window of duration T, the resulting convolution in the frequency domain carries the transform of the window function. The result is the production of *sidelobes* at frequencies spaced at multiples of $1/2T$ about the true frequency (Figure 4.7). Because all real waveforms to be analyzed are of finite duration, the existence of sidelobes is unavoidable. One method for minimizing the power in the sidelobes consists of "tapering" the window function; a window that falls smoothly, rather that abruptly, to zero will contribute less power leakage to frequencies away from those actually contained

in the continuous waveform $f(t)$.

4.4 The Periodogram and Rayleigh Test

While the FFT yields the most efficient manner in which to calculate the DFT of a function $f(t)$, often one is not interested in the complete power spectrum but, rather, in the power contained at a given frequency (or in some limited frequency band). This is true for data obtained in γ-ray astronomy if the source has a known periodicity. While the information desired is certainly contained in the FFT, it is easier to go back to the DFT and calculate the power spectrum at just those frequencies of interest.

We note that the DFT as expressed in Equation 4.15 can be reduced to

$$F(\omega) = \sum_{j=0}^{N-1} a_j \cos \omega t_j + i \sum_{j=0}^{N-1} a_j \sin \omega t_j. \tag{4.19}$$

This represents a vector in the complex plane. We may associate the power associated with the frequency ω with the square of the length of this vector:

$$S(\omega) = \mid F(\omega) \mid^2 = \left(\sum_{j=0}^{N-1} a_j \cos \omega t_j \right)^2 + \left(\sum_{j=0}^{N-1} a_j \sin \omega t_j \right)^2 . \tag{4.20}$$

The classical periodogram is conventionally defined (Scargle, 1982) as

$$S(\omega) = \frac{1}{N} \mid F(\omega) \mid^2 \tag{4.21}$$

which, apart from normalization, is identical with the power as defined above.

The *Rayleigh Test* was first devised by Rayleigh in 1905 in order to describe the statistics of a random walk in two dimensions. It has found extensive use recently in the field of γ-ray astronomy (Gibson *et al.*, 1982). The test statistic

is derived by assigning a phase $\phi_j(\omega)$ to each time measurement (see Equation 4.12) and forming the sum

$$z(\omega) = \frac{1}{N}\left\{\left[\sum_{j=0}^{N-1}\cos\phi_j(\omega)\right]^2 + \left[\sum_{j=0}^{N-1}\sin\phi_j(\omega)\right]^2\right\}. \tag{4.22}$$

Since, for measurements in γ-ray astronomy, the coefficients a_j are all equal to one, we see that the Rayleigh statistic and the power defined by the periodogram are identical. If the data ϕ_j are random, then measurements of $S(\omega)$ yield a distribution which is proportional to a χ^2 distribution with two degrees of freedom (i.e. the sine and cosine sums act as independent variables). Thus, the probability of obtaining a power value larger than some fixed value z_0 is given by

$$W(z > z_0) = e^{-z_0}. \tag{4.23}$$

This distribution is valid only for large N, however; for N smaller than ~ 50, an approximate correction to the exponential should be used (Mardia 1972):

$$W(z > z_0) = e^{-z_0}\left[1 + \frac{2z_0 - z_0^2}{4N} - \frac{24z_0 - 132z_0^2 + 76z_0^3 - 9z_0^4}{288N^2}\right]. \tag{4.24}$$

4.4.1 Fourier Independent Frequencies

Because the periodogram is equivalent to the DFT, it is subject to the same spectral leakage problems associated with the DFT. An advantage of the periodogram is that the time series which constitutes the sampled waveform need not be binned (as with the FFT). Instead, the arrival times provide a very non-uniform sampling which reduces the effects of aliasing. The finite duration, T, of the data set, however, limits the frequency resolution. This is merely the result

of convolving the window function with the sampled waveform and is analogous to the optical resolution associated with single-slit diffraction. The result is that closely spaced frequencies are not statistically independent. The independent frequency spacing is given by

$$\triangle f = \frac{1}{T}. \tag{4.25}$$

Often one works in period space rather than in frequency space. Then

$$\triangle P = \frac{P^2}{T}. \tag{4.26}$$

If the time series is tested over a range of frequencies, one must account for the total number of independent frequencies tested in order to correctly assess the statistical results. It is generally useful to sample in frequency space at intervals somewhat smaller than the independent frequency spacing so as to avoid nulls in the spectrum created by the window function as well as to facilitate location of the peak of the spectrum; such a process is referred to as *oversampling*. It has been noted by de Jager (1987) that such oversampling results in the underestimation of the probability associated with the resultant value of the peak power (i.e. one calculates that the probability of obtaining a particular result is smaller than it really is). For an oversampling factor of ~ 5 the factor of underestimation, ψ, is between 2.5 and 3 (depending on the number of independent frequencies spanned).

The number of frequencies one needs to scan in searching for a signal from a periodic source depends on several factors. Clearly, if the period is not known, the range must be chosen based on physics assumptions about the likely behavior of the source. If the period is well known, however, one only needs to scan enough

frequencies to cover the uncertainty in the known period. For binary sources such as Her X-1, it is often desirable to scan a wide enough frequency band to account for any changes in the period that may result from orbital effects (which may be present even after barycentering if, for example, the site for production of the γ-ray signal differs from that of the x-ray signal).

We may now evaluate the statistical significance of obtaining a particular value z_0 from the Rayleigh test in a certain number of trials. Suppose all the data is divided into m_1 intervals of equal duration and that the Rayleigh test is used to span m_2 independent frequencies. Then, the total number of independent trials is $M = m_1 m_2$. Now we know that the probability of exceeding z_0 in a single trial is

$$p = e^{-z_0}. \tag{4.27}$$

Hence, the probability for obtaining no z larger than z_0 is $1 - p$. In M trials, this probability becomes $(1 - p)^M$ so that the probability of obtaining at least one value of $z > z_0$ in M trials is then

$$W(z > z_0 : M) = \psi \left[1 - (1 - p)^M\right] \approx \psi M p \tag{4.28}$$

where the approximation holds for $Mp \ll 1$ and the factor $\psi \approx 3$ is the underestimation factor mentioned earlier in this section. For M trials, the probability of obtaining n measurements $z > z_0$ is

$$W_n(z > z_0 : M) = \left(\begin{array}{c} M \\ n \end{array} \right) (\psi p)^n (1 - \psi p)^{M-n}. \tag{4.29}$$

We can immediately see the relation between the Rayleigh probability as defined above and the *level of confidence* ϵ. We consider two hypotheses, H_0

and H_1. Here H_0 is the hypothesis that no periodic signal is present in the data set (*i.e.* the *Null Hypothesis*) and H_1 is the alternate hypothesis. We may consider the set S of all possible results of a particular test and define a critical subset s in which H_0 is regarded as unlikely to be true. Then ϵ is defined as the probability of a measurement falling in s when H_0 is true. This is exactly what the Rayleigh probability represents. Thus, its identification with statistical significance is consistent. It should be noted, however, that it is possible to define the critical region mentioned above on the basis of more than one (independent) test statistic (see Section 4.4.3). If, for example, two independent statistics are to be combined, one must take care in properly interpreting the overall confidence level. Wallis (1942) has shown that if ϵ_1 and ϵ_2 represent the confidence levels associated with the two tests, then the overall confidence level is given by

$$\epsilon = \epsilon_1 \epsilon_2 \left[1 - \ln\left(\epsilon_1 \epsilon_2\right)\right]. \tag{4.30}$$

The derivation of this result is straightforward and is summarized in Appendix C.

4.4.2 Signal-to-Noise Characteristics

As suggested in Section 4.4, it is useful to consider the periodogram in terms of vectors. In principle, the phases used to calculate the periodogram can be separated into those due to the signal and those due to the random background. The periodogram may then be written as

$$S(\omega) = z(\omega) = \frac{1}{N}\left[R_s^2 + R_b^2\right] \tag{4.31}$$

where R_s represents the magnitude of the signal vector and R_b represents that of the background vector. In the case of pure random background, we know that the distribution of z values is

$$W(z) = e^{-z}. \tag{4.32}$$

Hence, the mean value for z is found from the first moment of the distribution:

$$\bar{z} = \frac{\int_0^\infty z e^{-z} dz}{\int_0^\infty e^{-z} dz} = 1 - \frac{z e^{-z} \,|_0^\infty}{\int_0^\infty e^{-z} dz}. \tag{4.33}$$

Thus,

$$\bar{z} = 1, \text{ or } \overline{R_b} = \sqrt{N_b}$$

where N_b is the number of random background events.

We can use this information to determine the average behavior of the Rayleigh test in the presence of a signal. Suppose the signal events are distributed in phase with probability density $H(\phi)$. The signal vector then has components

$$x_s = N_s \int_0^{2\pi} H(\phi) \cos\phi d\phi,$$

$$y_s = N_s \int_0^{2\pi} H(\phi) \sin\phi d\phi.$$

We may then write

$$R_s = \beta N s \tag{4.34}$$

where

$$\beta = \left\{ \left[\int_0^{2\pi} H(\phi) \cos\phi d\phi \right]^2 + \left[\int_0^{2\pi} H(\phi) \cos\phi d\phi \right]^2 \right\}^{\frac{1}{2}} \tag{4.35}$$

As an example, suppose the signal has the form

$$H(\phi) = \frac{1}{2\pi}[1 + \cos(\phi - \pi)].$$

Integrating, we find

$$R_s = \frac{N_s}{2}$$

so that $\beta = 1/2$. For a signal distributed as a δ-function, $\beta = 1$.

The angle θ between the signal and background vectors is arbitrary. The resultant value of z depends on this angle:

$$z = \frac{1}{N_s + N_b}\left[R_s^2 + 2R_sR_b\cos\theta + R_b^2\right] \tag{4.36}$$

so that, using the mean value for R_b, we have

$$z = \frac{1}{N_s + N_b}\left[R_s^2 + 2R_s\sqrt{N_b}\cos\theta + N_b\right]. \tag{4.37}$$

Let us now define the signal-to-background ratio

$$\alpha = \frac{N_s}{N_b}. \tag{4.38}$$

Then we may write

$$z = \frac{1}{1+\alpha}\left[\beta^2\kappa^2 + 2\beta\kappa\cos\theta + 1\right] \tag{4.39}$$

where

$$\kappa = \frac{N_s}{\sqrt{N_b}}. \tag{4.40}$$

We see, then, that for modest amounts of signal (so that α is reasonably small), z is primarily a function of κ and θ. Because θ is arbitrary, a given value of κ can yield a considerable range of values for z. Figure 4.8 shows the distribution of z with θ (neglecting the divisor $1 + \alpha$) for a signal distributed as

$$H(\phi) = \delta(\phi - \phi_0).$$

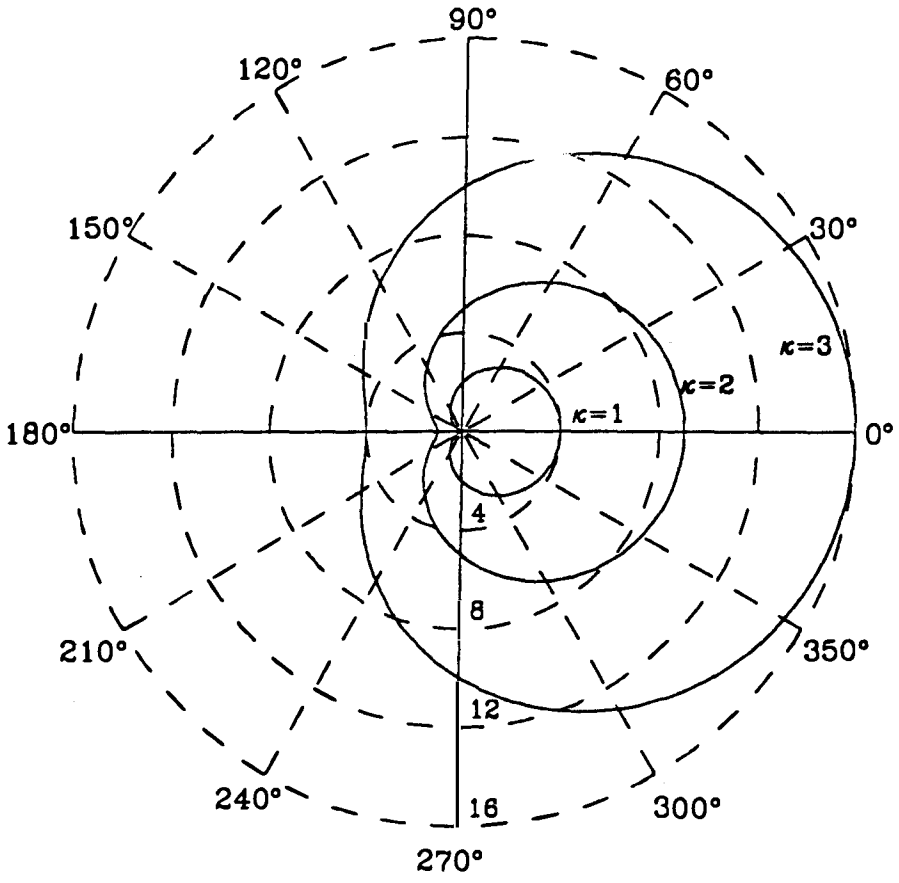


Figure 4.8: Distribution of Rayleigh statistic z as a function of the angle θ between the signal and noise vectors. (See text for a description of the approximations used to derive this result).

While this calculation neglects the fluctuations in R_b, we can make two qualitative statements: one need not expect a large value of κ to accompany a "detectable" value of z (*e.g.* for $\kappa = 2$ there are many values of θ for which $z > 6$ which has a corresponding probability of $\sim 2.5 \times 10^{-3}$); but a large value of κ does not guarantee a value of z which is large enough to claim detection. Monte Carlo simulations (see Section 4.4.3) show that these conclusions are indeed true.

Since the preceeding discussion indicates that z is primarily a function of κ, methods for maximizing κ must be devised. In the case of a steady signal, we can see that a search strategy should center on long integration times since

$$N_s = r_s t \text{ and } N_b = r_b t$$

where r_s and r_b are the respective signal and background rates. Thus,

$$\kappa = \frac{r_s t}{\sqrt{r_b t}} = \text{const} \times t^{\frac{1}{2}},$$

i.e. longer integration times will yield larger values of κ. Unfortunately, in VHE γ-ray astronomy signals often occur in bursts whose durations (and locations in time) are unknown. Clearly, if the burst time is short compared to the integration time, detectability is decreased.

4.4.3 Monte Carlo Simulations

In order to devise a coherent strategy for the search of periodic signals in a data set, it is useful to study the behavior of the chosen test statistics through simulations. To this end, Monte Carlo simulations of the expected time series have been performed with particular emphasis placed on the behavior of the

Rayleigh test in the presence of varying amounts of periodic signal. In these studies, a set of randomly distributed event times representing the background has been combined with "signal" times injected with a particular periodicity and distributed according to some chosen waveform. The signal distribution is that of a half sine-lobe which spans a chosen *duty fraction* δ of the period. The strength of the injected signal is characterized by the signal-to-background ratio α as defined in Equation 4.38.

Resolution

Although we know that the maximum resolution attainable using an FFT is given by the Fourier Independent Period spacing (Equation 4.26), it is not correct to assume that such resolution corresponds to the error in a period measurement using the Rayleigh test. Because of the ability to oversample the independent band, the resolution is actually increased. The price paid, however, is in the form of a statistical penalty associated with oversampling as mentioned earlier. The correct formulation of the error in a reconstructed period has been discussed by Middleditch (1976) and Gorham (1986) and warrants additional emphasis.

In order to form a standard 1σ error bar associated with a measurement of the period of a signal, one must determine the period range over which the probability drops from a value p' to a value $e^{-\frac{1}{2}}p'$. Thus, if we let p_0 represent the probability associated with z_0, the largest detected value of z, we need to find the period with Rayleigh power z' such that

$$p_0 = e^{-\frac{1}{2}}p' \tag{4.41}$$

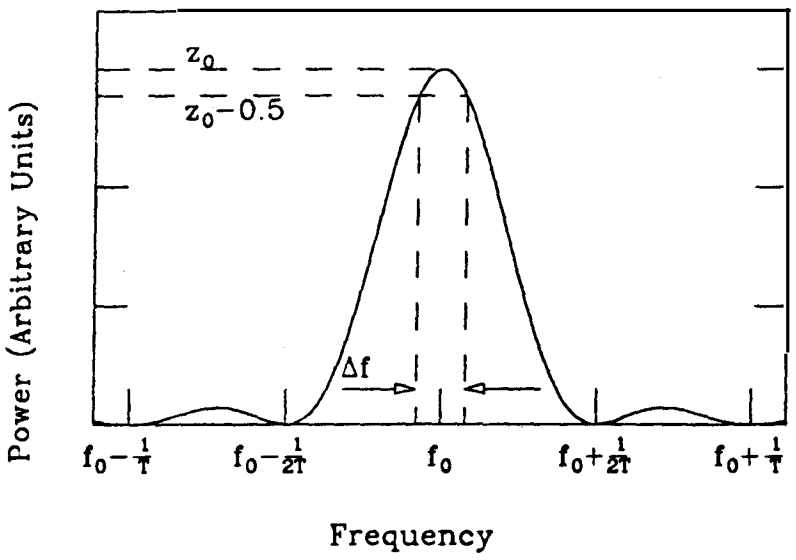


Figure 4.9: Schematic of power spectrum and associated error bars on reconstructed period.

or

$$e^{-z_0} = e^{-\frac{1}{2}}e^{-z'}. \tag{4.42}$$

Solving for z', we have

$$z' = z_0 - 0.5. \tag{4.43}$$

Thus, the range for the error bar on the period extends between values of P for which $z = z_0 - 0.5$ (Figure 4.9).

Such a formulation implies that the resolution of the Rayleigh test is dependent upon the value of z. This is intuitively obvious; the larger the signal strength, the better one should be able to resolve the period. In order to determine the range of the period resolution as defined above, it is necessary to

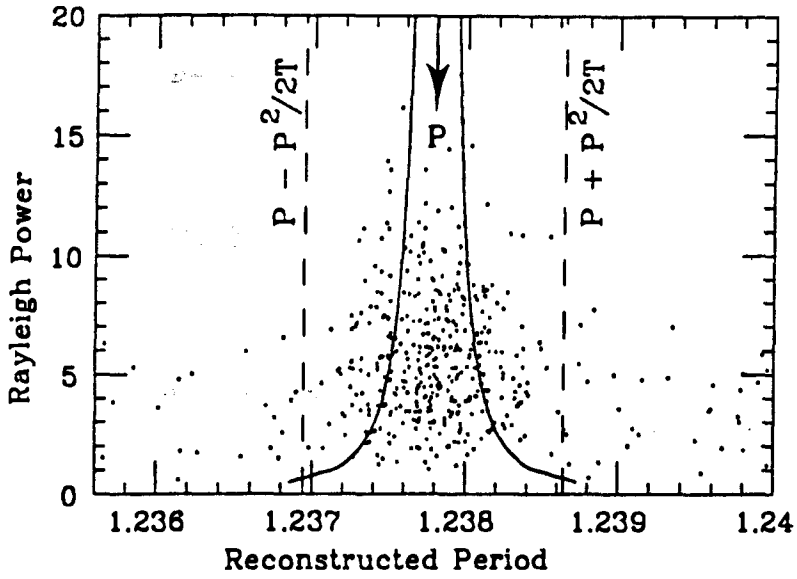


Figure 4.10: Resolution of Rayleigh test. Simulation of 500 data intervals with duration 900s and average background rate 0.7 Hz. The signal strength used was 10% distributed as a half sine-lobe with a duty fraction of 0.6. The injected period was chosen to coincide with that for Her X-1.

perform a graphical solution. Alternatively, one can assume a functional form for the power spectrum to obtain an analytical solution. Under the assumption that the power spectrum is described by a "sinc^{2n}" function, Middleditch (1976) has shown that the period resolution is given by

$$\Delta P = \frac{P^2}{2\pi T}\sqrt{\frac{6}{z}}. \tag{4.44}$$

In Figure 4.10 the results of a simulation of the resolving power of the Rayleigh test are plotted. Here signal with a period P corresponding to that of Her X-1 was added into a randomly distributed background. The duty cycle chosen for

the signal distribution was $\delta = 0.6$ (which is similar to that expected from Her X-1) and the signal strength was chosen to be 10%. The results of the simulation, however, are not significantly affected by the choice of these parameters. After injection of the signal, the Rayleigh test was performed over three Fourier independent periods and the value of the period which reconstructed with the largest power was plotted. From the figure it is clear that the distribution of reconstructed periods is significantly narrower than one independent period. The solid envelope shown is that corresponding the the formula for the resolution given above (Equation 4.44). It is clear that this envelope encompasses approximately two-thirds of the data points while the full independent period band encompasses a much larger fraction and, thus, corresponds to much more than a 1σ error bar.

While the preceeding discussion suggests that the period resolution for the Rayleigh test is always better than one Fourier independent period, this need not strictly be true. This is because the time window T defines the independent period spacing while the actual duration of the signal defines the the envelope of the power spectrum. Thus, if the signal occurs in a burst of duration $T' < T$, the envelope will be correspondingly larger and the resolution will be decreased (Figure 4.11).

Additional complications in resolving the period of a signal will result if periodicity occurs in several discrete bursts within the search window. If, for example two bursts of equal duration occur during a particular interval of data, the resulting power spectrum will correspond to the familiar double-slit interference

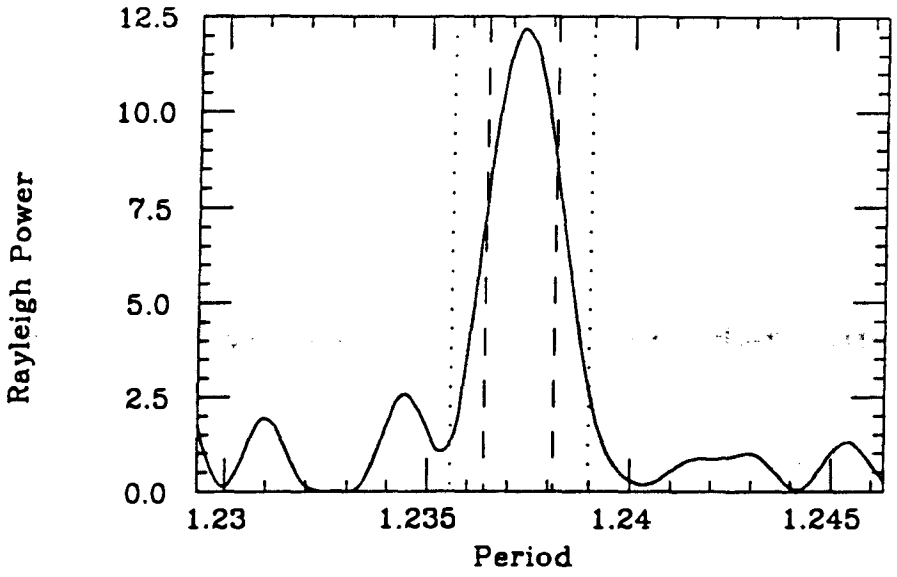


Figure 4.11: Power spectrum corresponding to a search window of 900s with the signal burst spanning only 450s. The dashed lines span one Fourier independent period assuming a duration of 900s while the dotted lines correspond to a duration of 450s. Note that the width of the peak, and thus the inherent resolution, is broader than expected from a 900s window.

pattern. Similarly, more complex arrangements of signal bursts result in correspondingly complex power spectra. In such cases, it may be helpful to study the temporal growth (and decay) of the Rayleigh statistic inside the search window. Such an exercise suffers from the normalization characteristics of the statistic whereby the Rayleigh vector is weighted inversely by the number of events. This normalization reduces the ability to identify bursts of signal unless the test is begun very near the beginning of the signal burst; accumulating a large num-

ber of random events preceeding a burst will diminish the resulting statistic in proportion to the number of such events.

Efficiency

Because of the nature of the Rayleigh test, it was suggested earlier that presence of a signal may go undetected because of cancellation from the background. Figure 4.12 illustrates this effect by plotting the value of z for fixed signal-to background ratio but varying values of the angle θ (compare with Figure 4.8). It is clear that fluctuations in the 'noise vector' can result in cancellation of the 'signal vector.' Using Equation 4.39, we can estimate the mean value of z for a given value of the signal fraction. For a signal distributed as a δ-function (*i.e.* very small duty fraction), we have β=1 so that averaging over all values of θ yields

$$\bar{z} = \frac{1 + \kappa^2}{1 + \alpha} \tag{4.45}$$

where α, β, and κ are as defined earlier (Equations 4.34, 4.38, and 4.40). Figure 4.13 illustrates this behavior along with the results from Monte Carlo simulations where we have used a background level corresponding to an average rate of 0.7 Hz tested over intervals of duration 900s. From the plot we see that if 500 such intervals are tested, the signal strength required to yield a $\bar{z}$ which corresponds to the 99% confidence level ($\bar{z} = 10.8$) requires $\kappa \approx 3.4$, or a signal-to-background ratio of about 14%.

The efficiency of a particular test for identifying a signal is dependent upon the type of signal being studied. Since the signal relevant to the analysis in

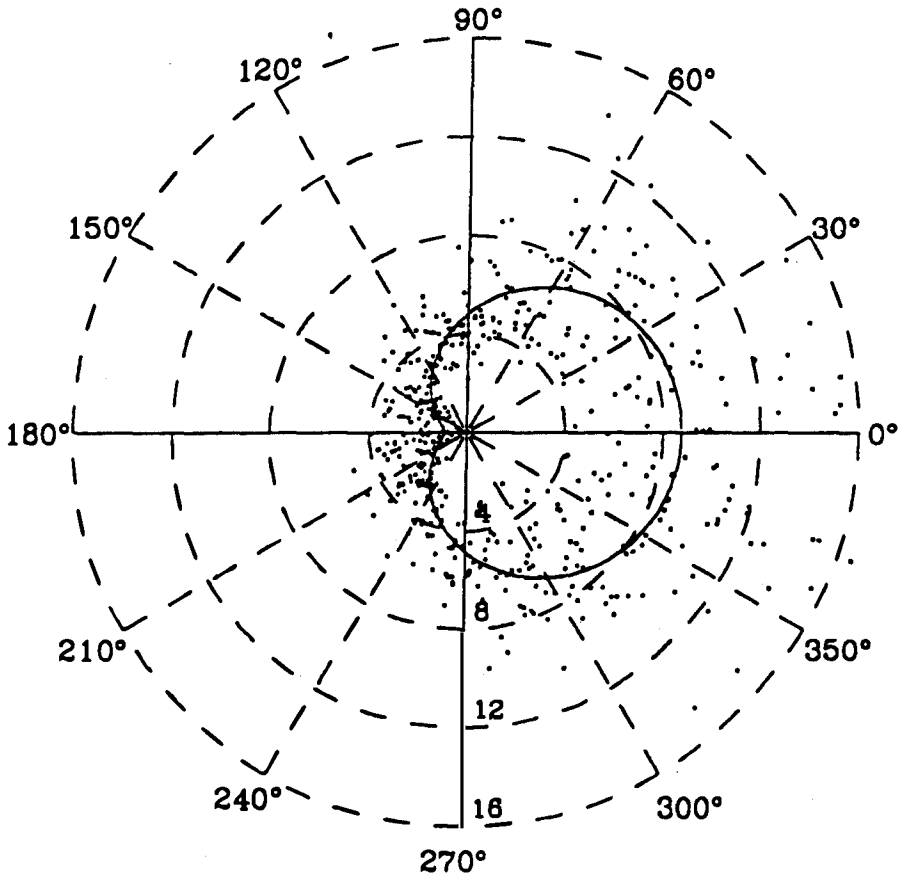


Figure 4.12: Scatter plot of z as a function of the angle θ between the background and signal "vectors" for fixed signal strength $\alpha = 0.1$ ($\kappa = 2.1$, $\beta = 1$).

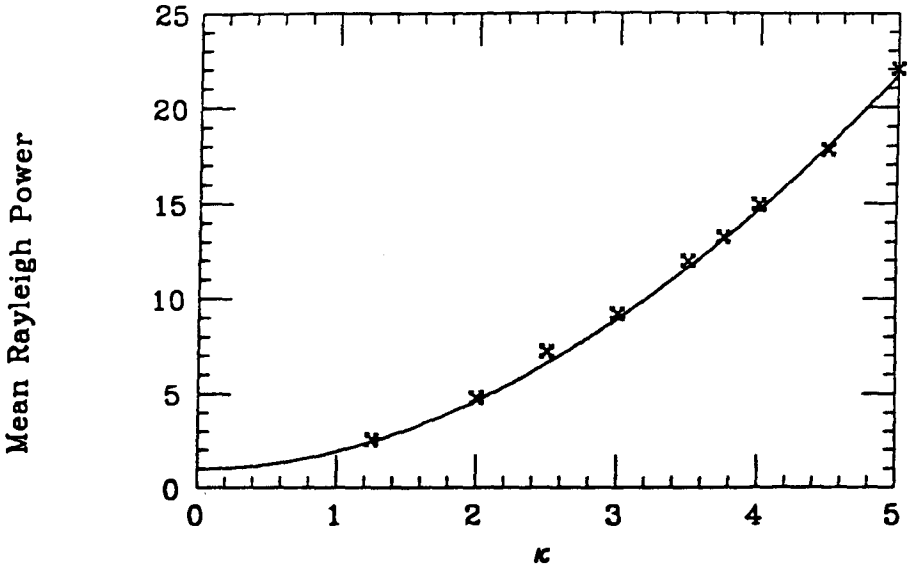


Figure 4.13: Average value of the Rayleigh power z as a function of signal strength using Equation 4.39. The data points are derived from Monte Carlo simulations using a light curve for which $\beta = 1$.

this study is likely to be sporadic, a reasonable definition of efficiency centers on the ability of a test to identify a single burst of signal in the presence of many intervals of pure background. Thus, in the case of the Rayleigh test, we must set a threshold in z beyond which we may conclude the presence of signal at the 99% confidence level. This threshold is dependent upon M, the total number of intervals searched since

$$P = 1 - (1 - e^{-z_0})^M. \tag{4.46}$$

Further, an additional factor must be included if the power spectrum is oversampled. For example, for P=0.01 when M=500, we have (assuming an oversampling

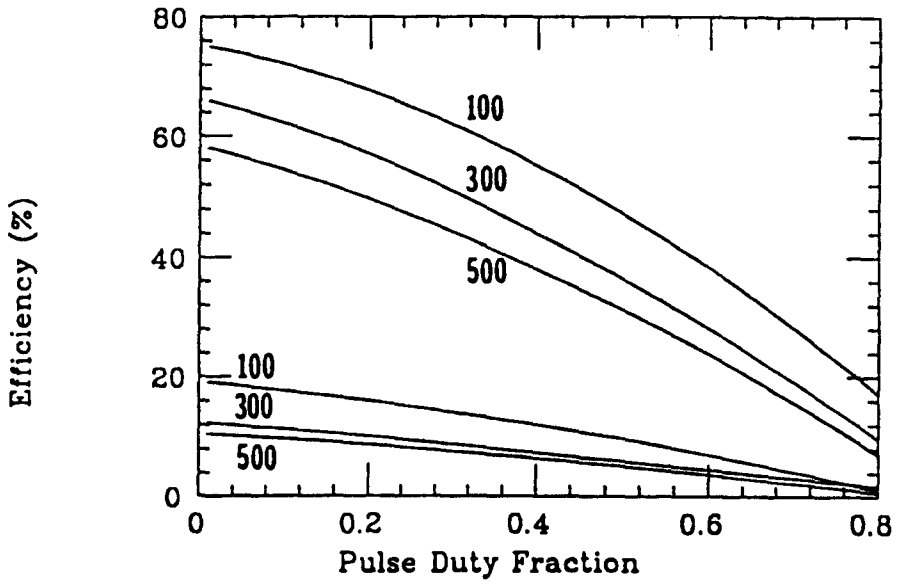


Figure 4.14: Efficiency of Rayleigh test as a function of signal pulse duty fraction. The upper family of curves correspond to a signal strength of 15% while the lower curves have 10% signal. The numbers along each curve indicate the number of such intervals searched (and, thus, the corresponding detection threshold z_0.

penalty $\psi = 3$) $z_0 = 11.9$. To determine the efficiency of the Rayleigh test, then, we may simulate data with a particular duty cycle and fraction of signal, perform the test on N such data strings, and find the percent which pass the detection threshold. The results of such simulations are shown in Figure 4.14 where we have plotted the efficiency as a function of pulse duty fraction for experiments consisting of 100, 300, and 500 test intervals. The simulation was performed for both 10% and 15% signal and, in each case, one Fourier independent period was sampled at 7 lattice points. The fraction of intervals that contained power

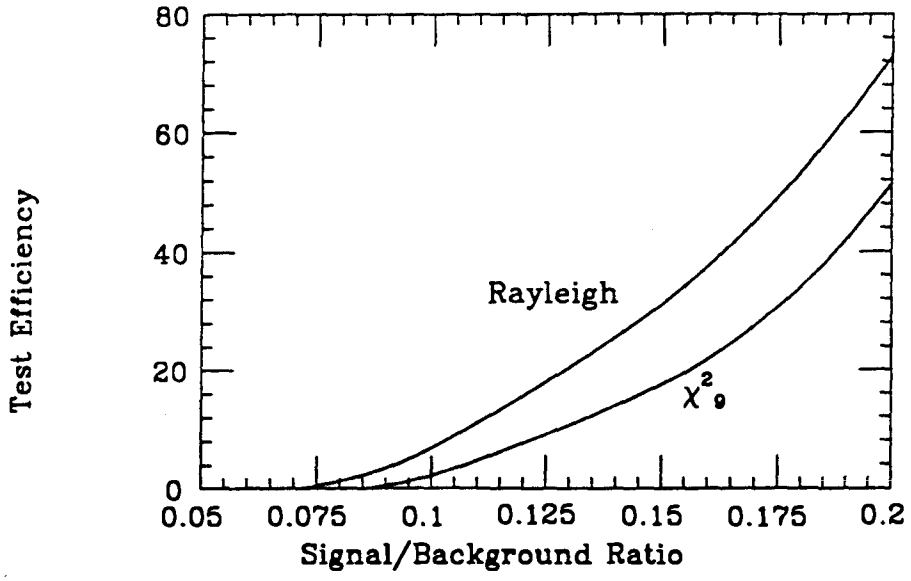


Figure 4.15: Comparison of Rayleigh test and χ^2_9 (for a 10 bin light curve) as a function of signal strength.

above z_0 then defines the efficiency. It is clear that the Rayleigh test favors narrow pulses (which agrees intuitively with the simple vector addition model) and that considerable signal is required before the likelihood of an isolated burst exceeding the threshold becomes large. For comparison purposes, the Rayleigh test and a χ^2 test on the folded data using 9 degrees of freedom were performed simultaneously. In this comparison, a duty fraction of 0.5 was used and 500 intervals were tested. In order to remove ambiguity due to oversampling in the two tests, only the period at which the signal was injected was tested. For such signal distributions, the Rayleigh test is more powerful than χ^2_9 for all values of signal strength (Figure 4.15).

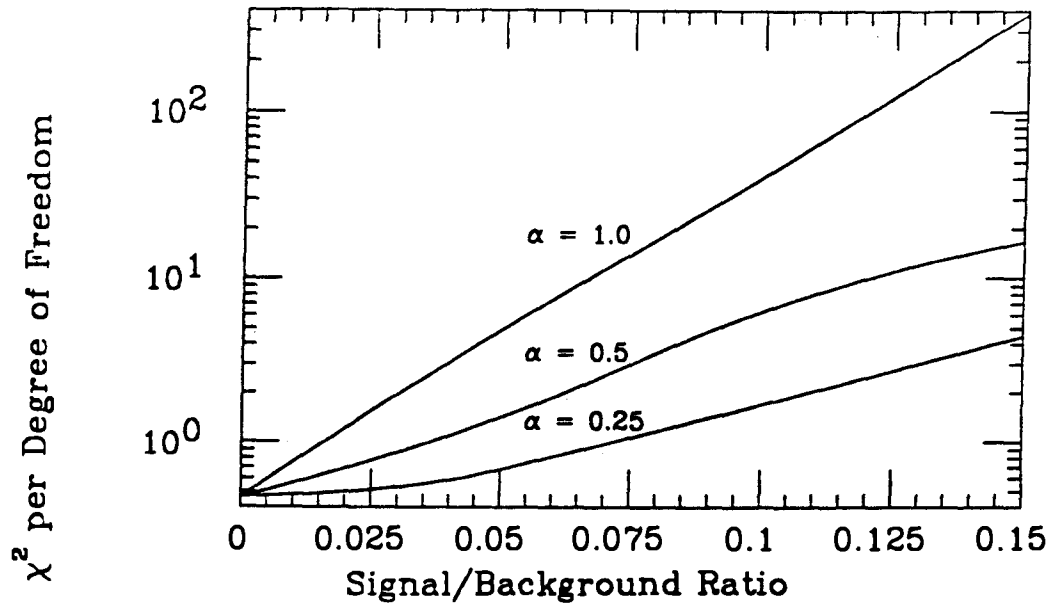


Figure 4.16: Results of χ^2 test of Rayleigh power distribution as a function of signal strength for 500 test intervals of 900s duration. The signal pulse duty fraction used was 0.5 and the average background rate in the intervals was 0.5 Hz.

Although the Rayleigh test has been represented as having rather low efficiency at low signal levels, if the signal is steady (rather than just occurring in a burst) the efficiency is quite high. In that the expected distribution for z in the absence of signal is known to be an exponential, one can test the distribution obtained from data against the random distribution. In Figure 4.16, the results of a χ^2 test of this distribution is plotted as a function of signal strength. Along the vertical axis, the value of χ^2 per degree of freedom is plotted. The number of degrees of freedom varies from about 20 at small signal strength to about 40 at larger values. The probability associated with these χ^2 values is vanishingly

small for all but the smallest amounts of signal. For comparison, the distribution was tested for the case in which the signal occurs only in a fraction α of the time window (although appearing in each window - a somewhat unrealistic model). It is obvious that continuous low-level pulsed signals can be detected with relative ease. Also apparent is the fact that choosing an interval size which is large compared with the signal duration significantly decreases the ability to detect the signal.

Correlation With Rate Deviations

When signal is introduced to a random background, a deviation in rate is expected. Of course, since the rate is subject to statistical variations, such deviations may not always be significant. As with the Rayleigh test, if a signal is continuously present (and the background is well known) the effect on the distribution of expected rate deviations will be obvious even at relatively small signal strength. In the case of a burst of signal, however, it is quite conceivable that the expected increase will meet with a statistical downward fluctuation in the background to reduce the overall significance of the signal. For this reason, it is useful to use both a periodicity test (such as the Rayleigh test) and a measure of rate excess to search for non-random behavior. As a simulation, intervals with fixed amount of signal (subject to fluctuations) were generated and tested for periodicity. At the same time, a rate excess based on the expectation from the known background was calculated. The distributions are plotted in Figure 4.17. In the case of no signal, it is clear that there is no correlation between

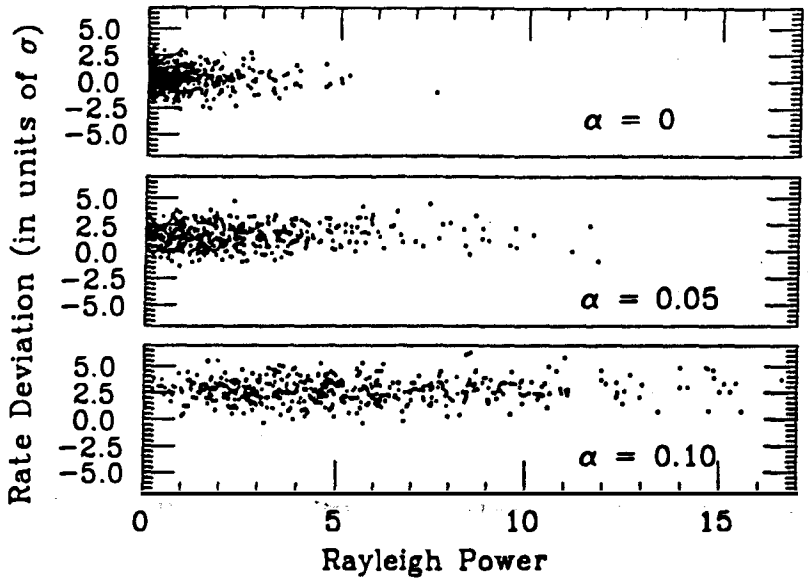


Figure 4.17: Scatter plots of rate deviation vs Rayleigh power for various values of signal strength. In this simulation, 500 intervals with duration 900s and average rate 0.5 Hz (prior to addition of signal) were used. The signal was distributed such that $\beta = 1$ (*i.e.* a very sharp light curve).

rate excess (here defined in standard deviations based on the mean background rate) and Rayleigh power. This is to be expected since the Rayleigh power is properly normalized to remove the contribution due to the number of events (for sufficiently large numbers). Hence, a statistical upward fluctuation in the rate of a set of random times should not show any non-random preference for periodicity. In that the time series to be analyzed is to be tested against the null hypothesis (*i.e.* the hypothesis that there is no signal present), we may define the probability associated with a given detection (z, ν), where ν is the number

of standard deviations corresponding to the rate deviation, as

$$W = \psi \left[1 - (1 - p_1)^M\right] p_2. \tag{4.47}$$

Here p_1 is the probability associated with the Rayleigh power, p_2 is that associated with the rate deviation (*i.e.* the probability against exceeding ν standard deviations), and M is the number of independent searches. The associated level of confidence, as explained in Section 4.4.1, is given by

$$\epsilon = W\,(1 - \ln W)\,. \tag{4.48}$$

From this it is clear that a signal may manifest itself as an excess with no apparent periodicity or as periodic with no significant rate excess. The power of both tests should be combined to identify deviations from random behavior.

Chapter 5

The Binary System Her X-1/HZ Herculis

Images of broken light
Which dance before me like a million eyes
They call me on and on
Across the Universe.
The Beatles

The binary system in which the periodic source Her X-1 resides has been studied extensively since its discovery in 1972. The picture that has unfolded presents views of both simple binary behavior and complex astrophysical processes. Pulse timing in electromagnetic bands extending from the infrared to UHE γ-rays has resulted in the identification of phenomena which has greatly improved the understanding of such binary systems. In addition, X-ray studies have provided a measure of the magnetic field strength in the vicinity of Her X-1 while optical studies have provided mass measurements for both stars as well as information pertaining to the accretion disk which powers the system. In this chapter a review of the studies of Her X-1 is presented. We begin with a discussion of the early studies in which the various periodicities of the system were identified. After discussing contributions made in the optical and infrared

studies of the companion star, we discuss the more recent VHE and UHE studies as well as models which have been proposed to explain the various features which have been identified in such studies.

5.1 X-Ray Studies of Her X-1

X-rays from Her X-1 were first discovered with the *Uhuru* satellite (Tananbaum *et al.*, 1972). Originally designated 2U1705+34, the source was immediately hypothesized to be part of an accretion-driven binary system. Periodicity of $1^d.7$, thought to correspond to orbital motion and associated eclipse of the X-ray source, was identified. Further, pulsations with a periodicity of $1^s.24$ were detected. Careful pulse timing of this periodicity revealed temporal variations in the period which were consistent with Doppler shifts due to an orbit with a $1^d.7$ period. By fitting these Doppler shifted values to an expression for the orbit, the path was found to be very nearly circular with a radius of about 13 light-seconds; the calculated eclipse was found to correspond in phase to the observed minimum in the $1^d.7$ periodicity. In addition, a long-term periodicity of approximately 35^d within which the X-ray source was observed for only about 9 days was noted. It was suggested that the source of this periodicity could arise from instabilities in the atmosphere of the companion star in the vicinity of the Roche lobe, thus resulting in decreased emission when the lobe ceases to be filled, or from precession of the pulsar beam. Although the nature of the long term cycle in this system is still in question, most of the original interpretations of the system parameters have turned out to be largely correct. The most recent

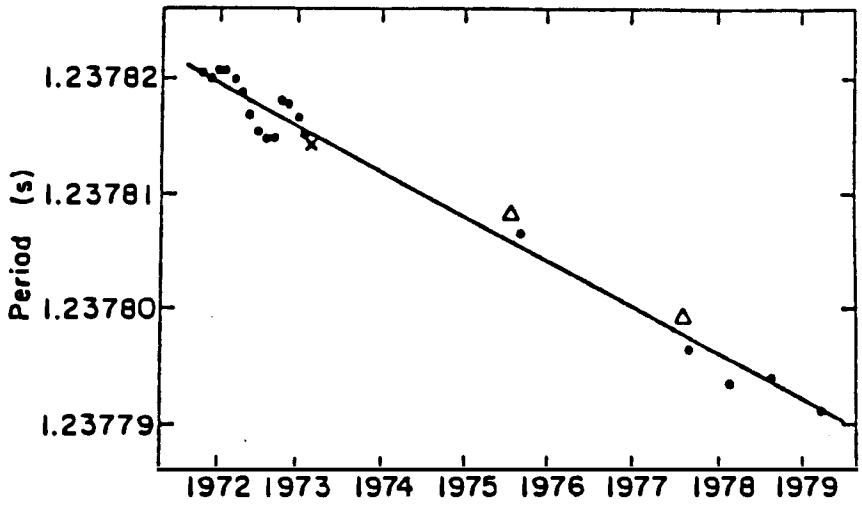


Figure 5.1: Secular changes of pulsar period of Her X-1. The overall decrease in the period is consistent with the suggestion that the system is powered by an accretion disk (which, thus, transfers angular momentum to the neutron star). Short term deviations from this overall trend are thought to result from torques applied by matter near the Alfvén radius.

ephemerides for the various cycles are listed in Table 5.1.

Further studies of *Uhuru* data (Giacconi *et al.*, 1973) show the pulsar period to be decreasing (Figure 5.1), consistent with the spin-up expected in accretion driven binary systems. The X-ray luminosity of $\sim 10^{37}$ erg s^{-1} (Forman *et al.* 1972) is consistent with such an accretion production mechanism. A connection between the orbital and long term periods have also been identified whereby the long term cycle (whose transition from off to on states was observed to be

Pulsar Period[a]	$P = 1^s.23779200 \pm 0^s.00000005$
Epoch of Period[a]	$T_0 = \mathrm{JD}2445778.56$
Period Derivative[a]	$\dot{P} = -2 \pm 1 \times 10^{-13}$
Orbital Period[b]	$P_{orb} = 1^d.700167788 \pm 0^d.000000011$
Projected Radius[b]	$\frac{a_x}{c} \sin i = 13^s.1831 \pm 0^s.0003$
Projected Velocity[b]	$v_x \sin i = 169.049 \pm 0.004$ km s^{-1}
Eccentricity[b]	$\epsilon < 0.0003$
Precession Period[c]	$P_{35d} = 34^d.928$
Epoch of Precession[a]	$T_{0,35d} = \mathrm{JD}2445788.0 \pm 0.5$
Her X $-$ 1 Mass[d]	$M_{\mathrm{Her\ X-1}} = 1.30 \pm 0.14 M_\odot$
HZ Her Mass[d]	$M_{\mathrm{HZ\ Her}} = 2.18 \pm 0.11 M_\odot$

(a) Ögelman *et al.* (1985)

(b) Deeter *et al.* (1981)

(c) Staubert *et al.* (1983)

(d) Middleditch and Nelson (1976)

Table 5.1: Her X-1 parameters.

very rapid) turns on at discrete values of the orbital phase, $\phi_{orb} \sim 0.2$ or 0.7. Further, dips in the X-ray intensity which are not associated with the eclipse of the source by the companion star have been observed. Such dips are thought to be produced by intervening gas. Variations of the intensity during such dips suggest that the column density associated with the intervening matter is not uniform. Even more interesting, however, is the fact that such intensity dips occur at progressively earlier orbital phases during the 35^d cycles. This is again indicative of some sort of precession whereby the intervening gas arrives along the line-of-sight somewhat earlier with each orbit. Brecher (1972) has suggested that precession of the neutron star could be responsible for the cycle while models for precession of the accretion disk have also been suggested (Roberts 1974, Katz 1973, Petterson 1975). Petterson (1977) has considered the "slaved-disk" model whereby the accretion disk precession results from precession of the axis of the companion star. He finds that, assuming that HZ Her corotates with the orbit and fills its Roche lobe, the precession of the axis results in a change in the angular momentum of the accreting matter thus producing a twisted accretion disk. Using this twisted disk along with a thin disk corona, a shell of matter at the Alfvén radius, and a time varying accretion rate, a model which is able to explain the various details of the cycle can be constructed. However, he concludes that such precession of HZ Her is hard to explain in that such effects should be removed in a time scale small compared to the orbit circularization time (note that the eccentricity of the orbit is extremely small - see Table 5.1). Crosa and Boynton (1986) have suggested that the absorption dips and discrete

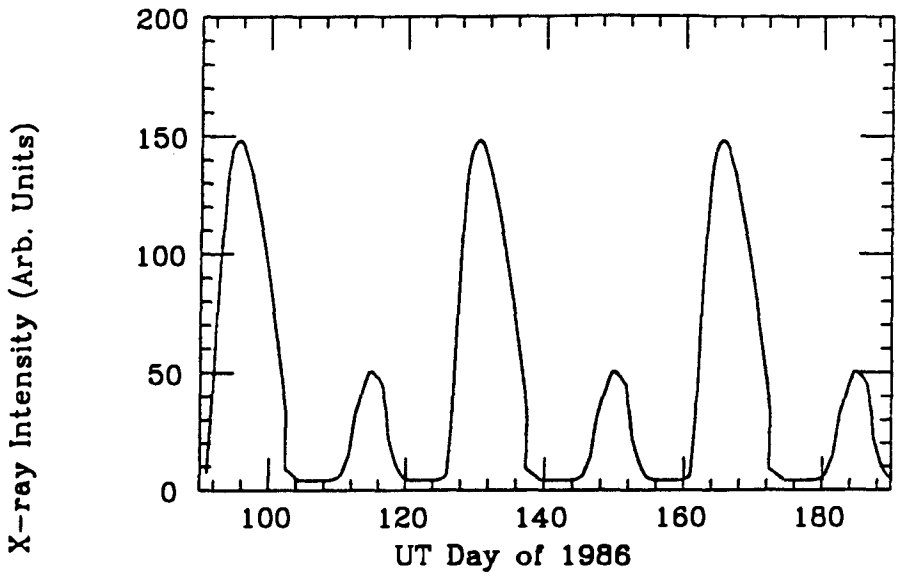


Figure 5.2: Variation of X-ray intensity due to 35^d cycle.

turn-on phases of the characteristic 35 day cycle can be explained in terms of periodic mass transfer associated with the passage of the X-ray shadow of the tilted accretion disk across the inner Lagrangian point of the system.

The identification of emission during the 35^d low state, displaying the characteristic $1^s.24$ period, eclipses, and absorption dips, (Fabian *et al.*, 1973; Cooke and Page, 1975; Jones and Forman, 1976) has extended the form of the intensity cycle to include both "High" and "Low" on-states separated by off-states. The form of the cycle is illustrated in Figure 5.2. The form of the bimodal intensity profile suggests that, in the model of a precessing disk, the X-ray source is partially obscured, possibly by a hot corona surrounding the disk, during the low-on state.

Using data from a balloon experiment, Trümper *et al.* (1978) have identified cyclotron lines in the X-ray spectrum which suggest a surface magnetic field of $\sim 4 \times 10^{12}$ Gauss. Such a measurement is consistent with values expected from models of stellar evolution to the neutron star state (see Chapter 2) and represents a most important parameter determination, especially in view of models invoked to explain the production of VHE and UHE γ-rays in such systems.

In 1983, an extended low state was discovered in Her X-1 (Parmar *et al.*, 1985) wherein EXOSAT observed the system over several 35^d cycles but measured no high increase in intensity above the that typical of the "off" portion of the cycle. This observation directly followed measurements made by the Japanese Hakucho and Tenma satellites which suggest that the secular spin-up of the pulsar had reversed, thus resulting in an increase of the period (Nagase *et al.* 1984). Such a spin-down could be related to a change in the accretion rate of the system although optical observations of the companion star made during the extended low state (Delgado *et al.* 1983) suggest that the X-ray source was still active. Rather, an increase in the size and/or density of the accretion disk may have been responsible for obscuring the X-ray source. By March, 1984 the extended low had ceased as EXOSAT once again observed transitions into the high intensity state. Evidence for neutron star precession based on a 180° phase shift of the main pulse feature in the EXOSAT data (Trümper *et al.* 1986) in passing from the high-on to the low-on state of the 35^d cycle has introduced further controversy in the interpretation of the nature of this long term variability.

The composite X-ray spectrum of Her X-1 (McCray *et al.* 1982) is shown

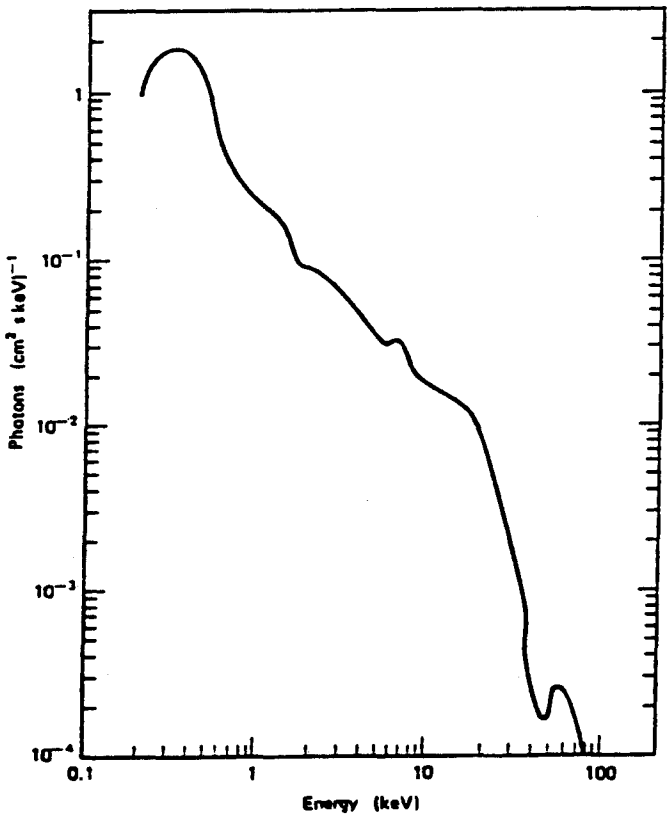


Figure 5.3: Composite photon spectra of Her X-1 from 0.1 to 100 keV.

in Figure 5.3. Below ~20 keV, the spectrum is relatively flat with a power-law index of $\beta \sim 1$. Above 20 keV the spectrum steepens to $\beta \sim 3.0$. Spectral features at ~6.4 keV (iron K-fluorescence line) and ~ 58 keV (corresponding to the cyclotron feature mentioned earlier) are clearly visible. The soft X-ray flux, from black body arguments, must originate from a distance $R > 10^8$ cm from the neutron star (Holt and McCray 1982). This distance very nearly coincides with the Alfvén radius (Section 2.4.2) where one expects a pile-up of accreting matter. A fraction of the hard X-ray flux is apparently intercepted by the matter and reradiated at lower energies.

Since the last EXOSAT data, there exists a gap in X-ray observations of Her X-1. The Japanese satellite *GINGA* (which has diverted much previously scheduled observing time in order to study SN1987A) and the X-ray facility "Kvant" (which is housed in the USSR's Mir space station) should provide new data in the near future, in particular with regard to the suggestion of the neutron star precession features.

5.2 Optical and Infrared Studies of Her X-1/HZ Her

Soon after the X-ray discovery of Her X-1, the ultraviolet-strong variable star HZ Herculis was suggested as a candidate for the companion star (Liller 1972). Identification of $1^d.7$ variations in the intensity of the star established its association with Her X-1 (Bahcall and Bahcall 1972). Optical pulsations with $1^s.24$ periodicity were soon identified (Davidsen *et al.* 1972). These pulsations were not detected in all of the observations, but appeared in both high and low X-ray intensity states. Indeed, the $1^d.7$ light curve shows no large dependence on the 35^d phase. This is consistent with the idea that X-rays from Her X-1 are absorbed, thermalized, and re-emitted in the atmosphere of HZ Her; one would expect the surface to receive some X-rays regardless of the 35^d phase.

The optical pulsations in HZ Her are found to occur primarily at discrete values of the binary phase. Middleditch and Nelson (1976) have identified three distinct regions of emission based on the frequency of these pulsations. By assuming that any matter which is reprocessing the X-rays is in corotation with the binary system, a measure of the (Doppler shifted) pulsation frequency and

its derivative yields a position of origin in the binary system. Such measurements suggest that pulsation occurs at binary phases near 0.25 and 0.75 from a source very near to the surface of the companion star while pulsations at phases near 0.85 originate from near the compact star. Models of the expected pulsations from HZ Her for various equipotential contours (*i.e.* various values of the fraction of the Roche lobe that HZ Her fills) indicate that the star must very nearly fill the Roche lobe. Thus, because the matter viewed at binary phase 0.5 is essentially "pointed" toward the observer, the reprocessed X-rays are not "reflected" in such a manner as to be observed. At the quadrature phases, however, the "reflections" off of the flatter regions of the distended stellar atmosphere are favorably directed toward the observer. By considering the expected Doppler shifts from these two regions for different values of the mass ratio of the two stars, the data is found to suggest a ratio of $\sim 1.5 - 2.0$. Given the duration of the X-ray eclipse and this mass ratio, a value for the inclination of the system can be inferred. This, along with the value of the mass function derived from X-ray data,

$$m \equiv \frac{(M_2 \sin i)^3}{(M_1 + M_2)^2} = \frac{4\pi^2}{G}\frac{(r_1 \sin i)^3}{P^2} \tag{5.1}$$

is sufficient to determine the individual masses of the components. The values obtained by Middleditch and Nelson are listed in Table 5.1.

Coordinated observations of optical and X-ray pulsations from Her X-1 (Joss *et al.* 1980) suggest that some of the optical pulsations may be emitted in-phase with the X-ray pulsations. Such optical photons are thought to originate in a region of enhanced density near the intersection of the accretion disk with an

accretion stream from HZ Her.

Simultaneous observations of optical and infrared pulsations from HZ Her (Middleditch *et al.* 1983) indicate that the spectrum is relatively flat in this region. These observations were made several orbits prior to a turn-on of the X-ray cycle. Later studies by Middleditch *et al.* (1985) successfully identified optical and infrared pulsations during the 1983 prolonged X-ray low state.

The optical data from this binary system is rich with information regarding the accretion disk and its effects relative to the 35 day cycle as well as the interplay between this cycle and the orbital period and its relation to details in the X-ray data. For detailed analysis the reader is referred to Boynton (1978), Crosa and Boynton (1980), and Middleditch (1983).

5.3 VHE and UHE γ-Rays from Her X-1

Observations of TeV γ-rays from Her X-1 have been made by several groups. The Durham group first reported such activity (Dowthwaite *et al.* 1984) when a three minute burst of activity pulsed at the $1^s.24$ period was observed. The time of the burst was consistent with a turn-on of the 35^d X-ray cycle. Because of the short duration of this burst, the resolution of the period measurement was quite poor. Several months later, pulsations from the source were detected by the "Fly's Eye" experiment (Baltrusaitis *et al.* 1985) at an energy > 500 TeV. The emission, however, apparently did not extend to lower energies in that simultaneous observations from the Durham group showed no such activity. In both 1984 and 1985, the Whipple Observatory group made several detections of

pulsations from Her X-1 (Gorham *et al.* 1986a, 1986b) including one episode that continued far into the eclipse of Her X-1 by HZ Her. Using the standard model of Eichler and Vestrand (1985), Gorham and Learned (1986) suggested that such behavior could result from beam steering in the magnetic field of HZ Her.

Upon investigating data which overlapped with the Whipple detections, the Durham group has reported a simultaneous detection of pulsations occurring on 4 April 1984 (Chadwick *et al.* 1986). On this date, the Whipple group reported a burst of activity in a 28 minute segment of data collected from the direction of Her X-1. The Durham group, analyzing an hour-long run containing this time interval, also observes periodic pulsations (at a somewhat less significant level). It appears, however, that the duration of the Durham signal is only $\sim$ 10 minutes and it is not clear from the literature whether this time interval is a direct overlap with the Whipple result.

The Haleakala group reported several episodes of activity displaying periodicity in 1985 (Resvanis *et al.* 1987). In this case, the light curves corresponding to the burst intervals were quite different from those generally seen from Her X-1. Rather than consisting of rather broad emission, these light curves showed activity confined to a narrow region of the pulsar phase. While consistent with the light curve from the Fly's Eye detection, this represents a considerable deviation from other TeV results. Additional episodes of activity in 1986 have been reported by several groups and will be discussed in Chapter 6.

When the orbital phase of the various TeV detections are plotted against the

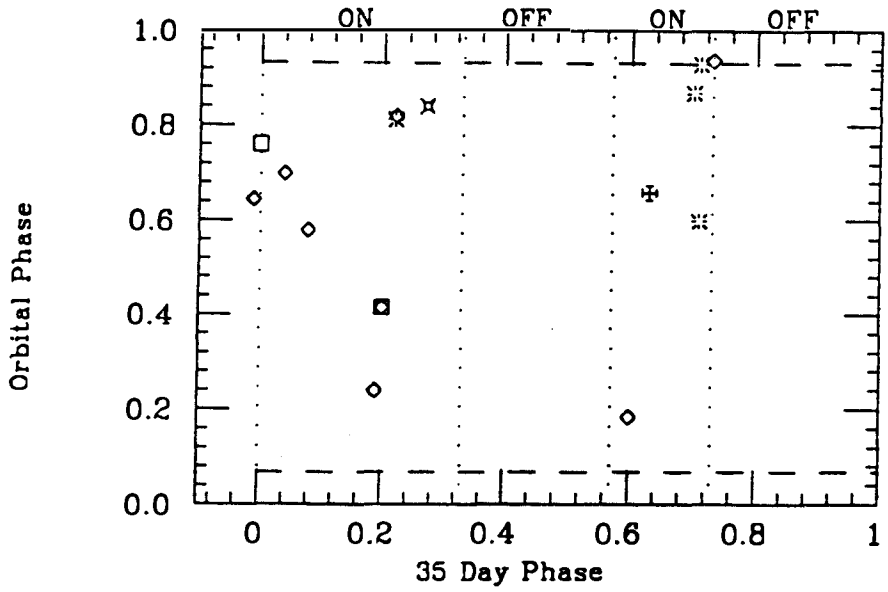


Figure 5.4: Scatter-plot of orbital phase vs 35^d phase for TeV observations of Her X-1. See Figure 5.5 for symbol key.

35^d phase (Figure 5.3) an obvious correlation is seen; the X-ray "off" portions of the phase are also void of VHE γ-ray detections. If the X-ray off cycles are the result of matter in the accretion disk blocking the view of the pulsar, one might expect enhanced γ-ray production because of the additional target material available for the conversion of charged particles. On the other hand, the particle acceleration regions are not well enough understood to claim that precession of the neutron star should result in a lack of particle beams directed toward the observer during such phases.

Possibly more interesting is the scatter of detected periods from TeV experiments. In Figure 5.4, these periods are plotted as a function of time. Note that

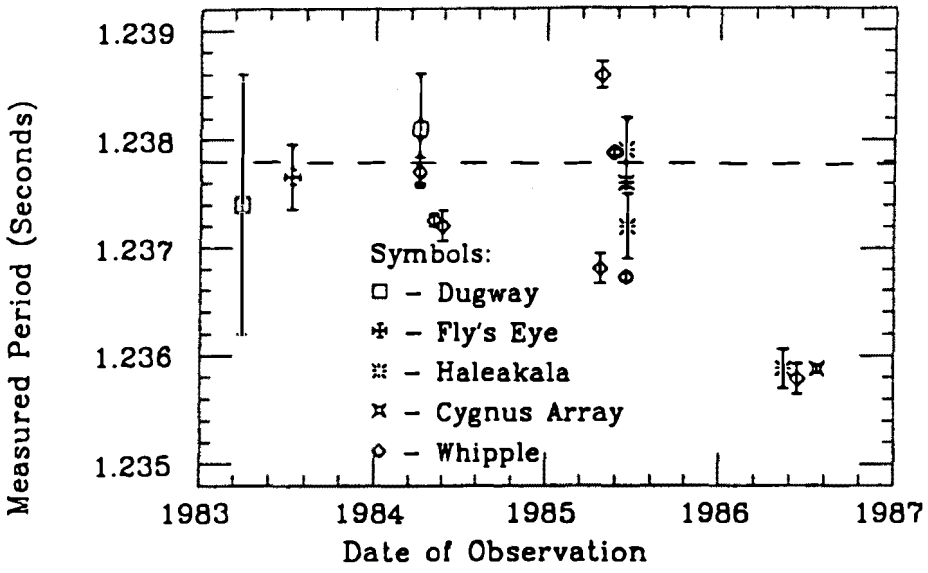


Figure 5.5: Measured period vs observation date for TeV γ-rays. Here the dashed line represents the expected X-ray period based on extrapolation from the most recent X-ray ephemeris.

while many of the observations produce periods which are consistent with the expected X-ray period, most are not. The 1986 observations create the illusion that the period is decreasing with time. This is an observational affect which results from the earlier practice of scanning a narrower range of periods in looking for signal (see Section 6.1). The deviations of periodicity from that of the pulsar implies that the γ-ray production occurs in regions removed from the neutron star. In models which use accelerated particles to produce the γ-rays, such deviations in the period may result from magnetic steering effects in combination with a moving beam dump target or from from a broad proton beam striking a

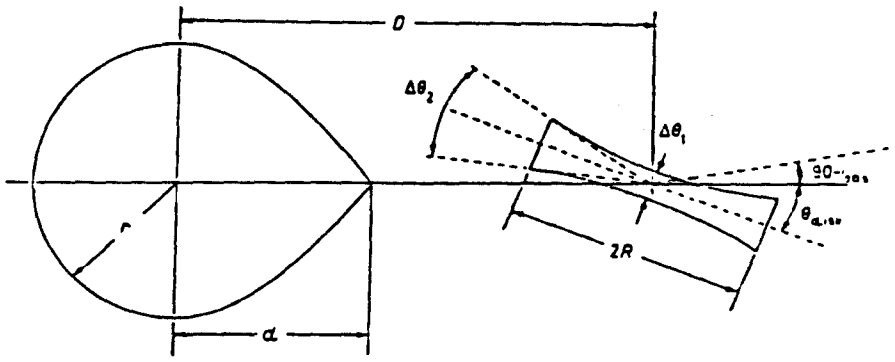


Figure 5.6: Schematic representation of the Her X-1/HA Her binary system.

target mass which has broken free from corotation.

5.4 Her X-1/HZ Her: A Synopsis

The many and varied studies of the Her X-1/HZ Her binary system have yet to provide a complete understanding of all the observed details. It is possible, however, to construct a reasonable picture which describes the approximate size, configuration, and energetics of the system. While such a picture will surely be deficient in some aspects and contain features which are questioned by some in the field, it will serve the purpose of illustrating the regions in which the characteristic radiation is produced.

The accretion-driven system is shown schematically in Figure 5.6. HZ Her-

culis, a $2.3M_{\odot}$ A-type star, fills its Roche lobe of nominal radius 2.7×10^6 km. Its binary companion Her X-1, located 6.4×10^6 km from its center, is a $1.4M_{\odot}$ neutron star with a surface magnetic field strength of $\sim 4 \times 10^{12}$ Gauss and a rotation period of 1.24s. The two stars orbit their relative center-of-mass, located 4×10^6 km from the neutron star, with a period of 1.7 days. Matter from HZ Her flows through the inner Lagrangian point, at a distance of 2.9×10^6 km from Her X-1, and, by virtue of its relative angular momentum, forms an accretion disk whose radius is $\sim 2 \times 10^6$ km - nearly that of the Roche lobe of the neutron star. The matter, which is accreted at a rate of $\sim 10^{-9} M_{\odot}$ yr^{-1}, gradually loses angular momentum thus falling into ever smaller Keplerian orbits. At a distance of $\sim 10^2 - 10^3$ km from Her X-1 (the Alfvén radius), the magnetic field of the neutron star is sufficient to force the matter into corotation. Angular momentum which is transferred to the neutron star upon accretion results in its overall spin-up. Occasionally instability regions develop in which matter which is orbiting with its Keplerian period couples to the magnetic field of the neutron star. The resulting torques may act to increase or decrease the rotation speed of the neutron star depending upon whether the instability develops inside of outside of the corotation radius. This behavior is evident in Figure 5.1.

Upon reaching the Alfvén radius, the accreting matter becomes controlled by the magnetic field lines whereby it is funneled toward the polar caps. The extensive heating which results is the source of X-rays. Because the polar caps rotate with the star, the X-rays appear pulsed to a distant observer. The accretion disk, which is tilted several degrees relative to the orbital plane, appears to

precess with a 35 day period. Because the orbital plane lies very nearly along our line of sight, the tilted disk regularly obscures the view of the X-rays. However, sufficient quantities reach the companion star to result in optical and infrared reprocessing which is modulated at the X-ray period.

Somewhere between the neutron star and the light cylinder (which is $\sim 6 \times 10^4$ km distant) electric fields sufficient to accelerate protons to energies in excess of 10^{15} eV are formed. These accelerated protons interact with surrounding matter, possibly in the accretion disk or in the outer regions of HZ Her, to form very high energy γ-rays. Such γ-rays have been observed only recently and the further study of such radiation is the subject of this paper.

Chapter 6

A Search for VHE γ-Rays from Her X-1

Well my feet they finally took root in the earth.
I got me a nice little place in the stars.
And I swear I found the key to the universe
In the engine of an old parked car.
Bruce Springsteen

The Haleakala Gamma Observatory logged 132 hours of observation time on Her X-1 between the months of May and July in 1986. After rejection of weather-affected data, 104 hours were selected for analysis. Here the details of this analysis are described and the results discussed. In particular, evidence of a burst of activity occurring on 13 May is presented. This burst is characterized by periodicity which is near, but distinctly different from, the pulsar period expected from extrapolation of the most recent X-ray ephemeris. This result is discussed in relation to similar results from other experimental groups in 1986 and suggestions as to the origin of such a result are discussed.

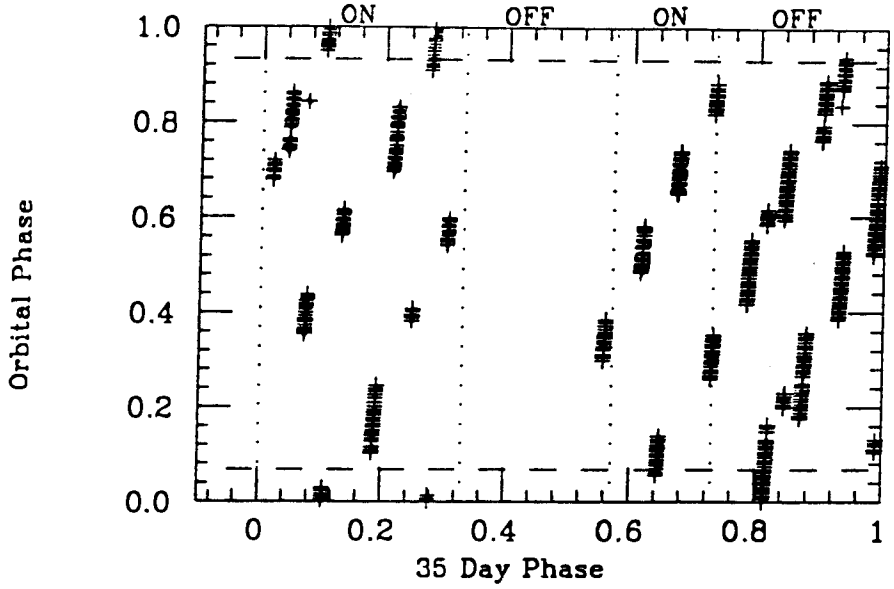


Figure 6.1: Orbital and 35^d phase distribution of HGO observation times in May-July 1986.

6.1 Data Preparation

In the summer of 1986, the Haleakala Gamma Observatory collected 132 hours of data from the direction of Her X-1. The data included observations covering a broad range of both orbital and 35^d phases (Figure 6.1). As always, an off-source region located 3°.6 from Her X-1 was monitored simultaneously and this off-source data was treated identically to the on-source data in performing a signal search. Because the durations of the data runs vary considerably, a decision was made to divide the data into 15 minute segments and to analyze each such interval for signs of γ-ray activity. Before compiling a set of such intervals, however, all data was inspected for overall quality. Any data which indicated the presence

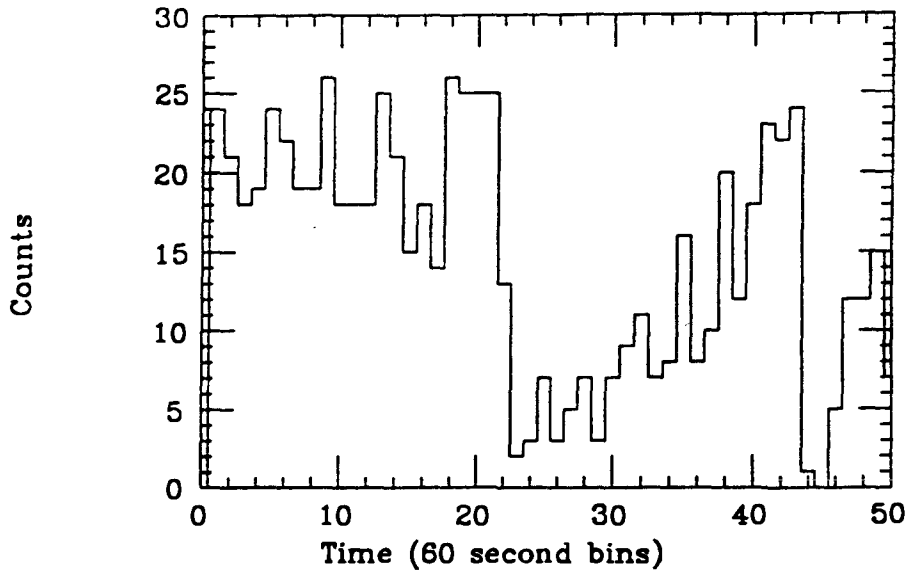


Figure 6.2: Example of data which has been affected by poor sky conditions. Here event counts are tallied in 60s bins.

of clouds (which are sometimes overlooked by observers) was rejected. Such identification was accomplished in a two step process. First a nominal shower cut (Section 3.2.4) was applied to all data. The resulting trigger rates were then inspected for "reasonable" behavior. Runs with rate fluctuations which were obviously related to sky conditions were identified and rejected (Figure 6.2). As a further check, the zenith angle dependence of the data was fit to a three-term Legendre polynomial expansion (as described in Section 3.2.4) and the deviation from the fit was recorded for each 15 minute interval. Runs in which a strong correlation between negative deviations in both the on-source and off-source apertures was apparent were rejected from the sample. This inspection process reduced the original 132 hours of data to a working set of 104 hours. Within this

set, the segments to be analyzed were chosen with particular attention paid to an effort to eliminate gaps in the data. Such gaps may occur, for example, when the high voltage to the PMT's shuts down in response to high base current in any particular channel. Although such gaps do not present any particular problem for periodicity analysis (except for the possible introduction of window effects), they are somewhat of a nuisance in terms of calculating deviations in rate and, as such, their elimination was deemed worthwhile. The end product of the data preparation was a set of 416 15-minute intervals.

As described in Chapter 3, the timing cut used to eliminate triggers which were caused by random light from the night sky background consisted of requiring a multiplicity of 9 or greater in a sliding 5ns time window. This requirement, of course, was made only after calibrating the data so that TDC times were corrected for slewing effects. The trigger times for each event passing the requirements of the shower cut were reduced to the solar system barycenter relative to the Her X-1/Hz Her barycenter as described in Chapter 4.

6.2 Period Search

The search for periodicity in the time series data selected to be analyzed was done with the Rayleigh test. Given the expected X-ray period, based on the ephemeris listed in Table 5.1, the Fourier-independent period spacing for the 15 minute data segments is $\triangle P = 1.702 \times 10^{-3}$s. Tests were made for 17 periods, centered on the expected X-ray period, and spanning 3 independent periods. This range corresponds to approximately ± 3.5 Doppler shifts in the binary system. While

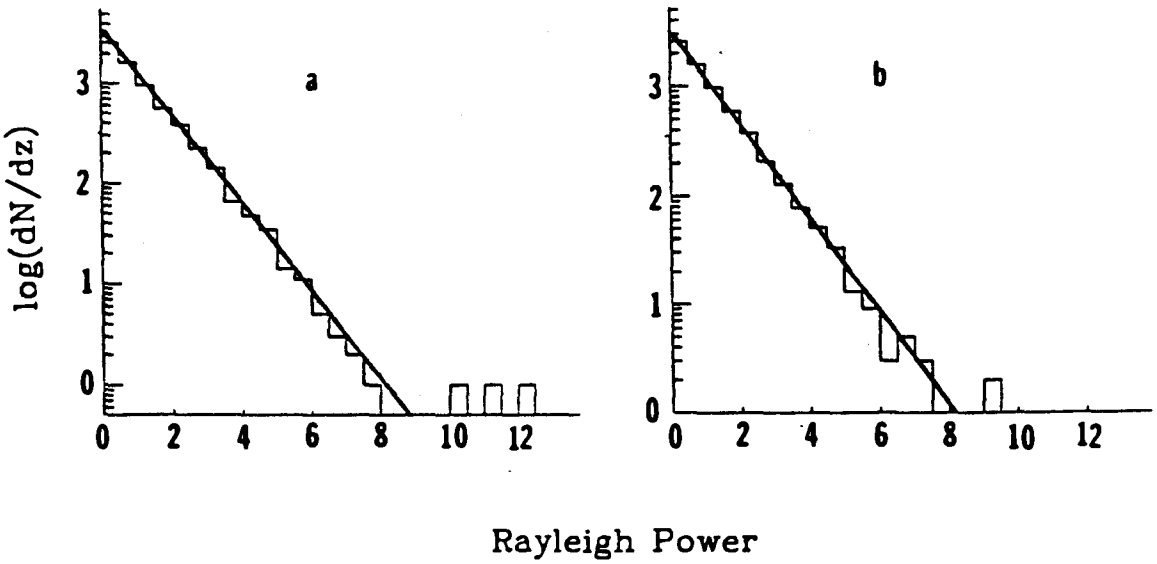


Figure 6.3: Distribution of Rayleigh powers obtained in periodicity search (a)on-source and (b)off-source. Note that the multiple entries of high power in the on-source data are the result of oversampling the burst interval.

this range is quite large, previous results indicate that VHE γ-ray emission may occur in a region different from that of X-ray emission (Gorham *et al.* 1986b). In that the production mechanism is not clearly understood, and since the presence of large magnetic fields can result in considerable steering of charged particle beams which may be responsible for the γ-ray production, such large shifts in the period are not unreasonable.

The distribution of Rayleigh powers for all intervals searched is shown in Figure 6.3. Note that such a distribution represents an oversampling of the independent periods. Thus, the three entries indicating high power in the on-

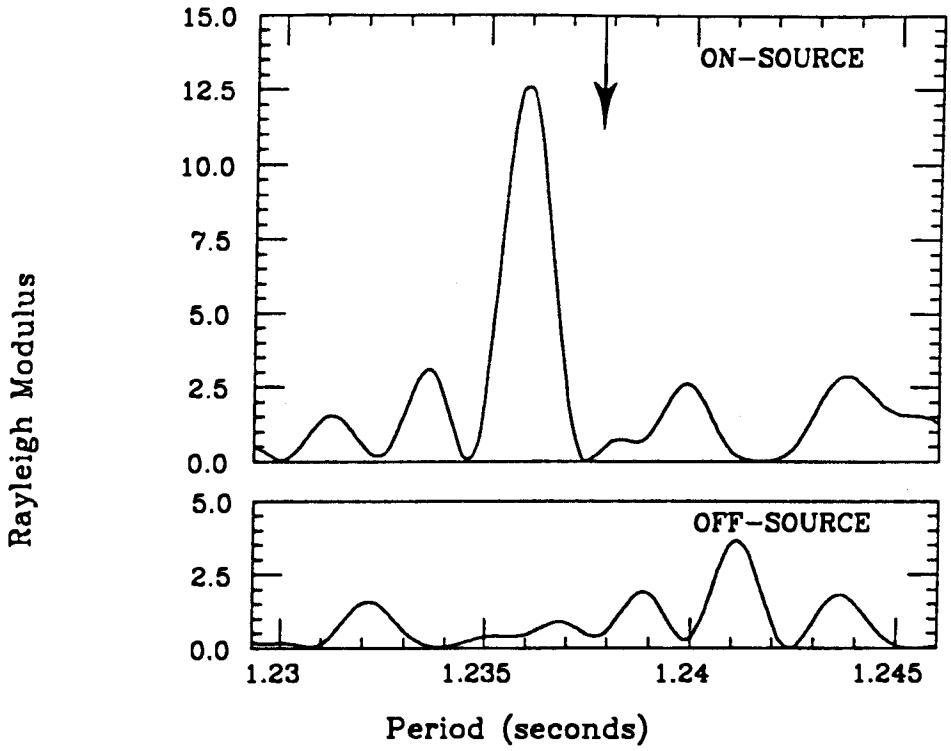


Figure 6.4: Power spectrum for burst interval at 13:35:40.98 UTC on 13 May 1986 (a)on-source and (b)off-source. The X-ray period is indicated by the arrow.

source data actually all correspond to the same time interval. Similarly, the two entries in the bin for which $z \approx 9$ in the off-source data correspond to a single oversampled peak. Comparison of the on-source and off-source distributions suggests no obvious low-level constant periodicity (*i.e.* the distributions are both fairly well described by the expected exponential fit). However, the on-source data clearly contains an interval of unexpectedly high power. This power corresponds to an interval starting at $13^h35^m40^s.98$ UTC on 13 May, at orbital phase $\phi_{orb} = 0.81$ and 35^d phase estimated at $\phi_{35d} = 0.22$. The power spectrum

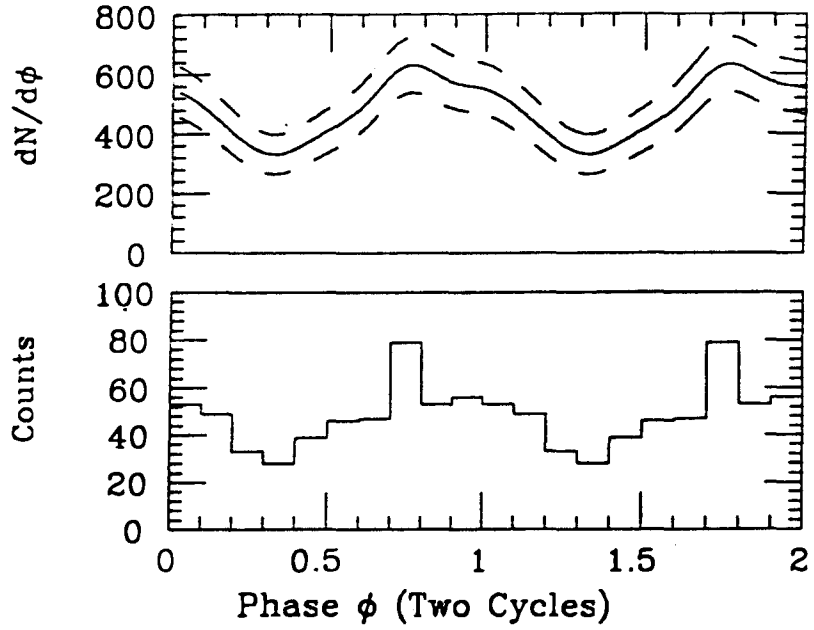


Figure 6.5: Light curve corresponding to VHE γ-ray burst on 13 May 1986. The upper curve was derived using the Kernel density estimator method of de Jager (1987).

for this interval is shown in Figure 6.4. Note that the spectrum is has been expanded to ten Fourier-independent periods for purposes of illustration. The detected period of $1^s.23593 \pm 0^s.00018$ is 0.15% lower than $1^s.237778$, the value expected by extrapolation from the most recent X-ray ephemeris. As we shall discuss in the next section, this surprising result is in excellent agreement with other TeV measurements made in 1986. The light curve for the burst interval is shown in Figure 6.5 along with the smoothed function as derived by the Kernel density estimator method of de Jager (1987). In an effort to determine whether the γ-ray activity extended beyond the 15 minute time window used for the

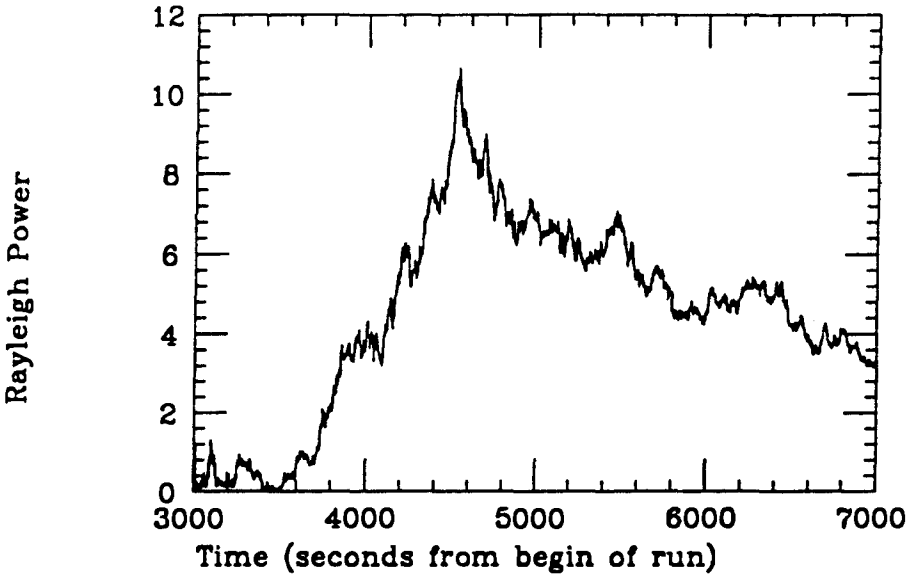


Figure 6.6: Rayleigh power as a function of time for data on 13 May 1986. The plot begins 3000 s into the data run. The search interval in which the burst was detected was begun at 3600 s (the beginning of the fourth 15 minute interval in the run).

interval, the Rayleigh power was calculated as a function of time. As explained in Chapter 4, the results of such a calculation suffer from the dependence upon starting time. Thus, events accumulated prior to the onset of the burst (which are included in the final weighting of the Rayleigh vector) tend to decrease the net power. In Figure 6.6, we plot the Rayleigh power as a function of time. Here we begin accumulating events 3000 seconds into the data run in an effort to investigate the time prior to the burst interval (which begins 3600 seconds into the data run). The onset of the burst appears to occur very near to the

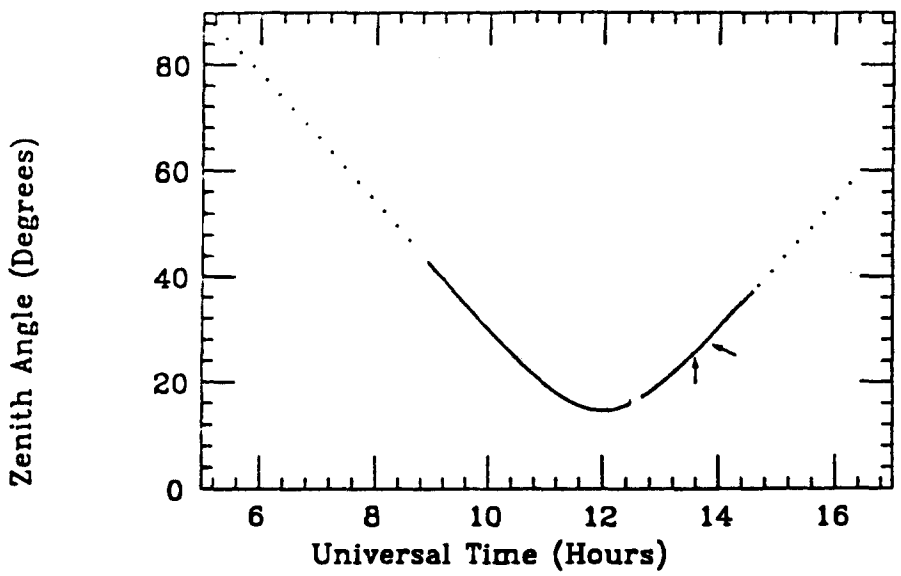


Figure 6.7: Observation window for Her X-1 on 13 May 1986. The solid curve represents the time interval over which data was taken while the interval in which the γ-ray burst occurs is indicated by arrows.

beginning of the search interval in question and the duration is very nearly 15 minutes. Note that the peak power is lower than the 12.79 expected because of the events accumulated in the times between 3000 s and 3600 s.

Figure 6.7 illustrates the observation window for Her X-1 on 13 May 1986 and indicates the time region in which the γ-ray burst occurred. The event rate for the interval containing the γ-ray burst is above the expected baseline by 6% or 1.4 standard deviations. Given the large Rayleigh power of the burst, this excess is lower than expected. As shown in Chapter 4, however, just as a real (pulsed) signal can manifest itself as a rate excess with very small Rayleigh power, one can

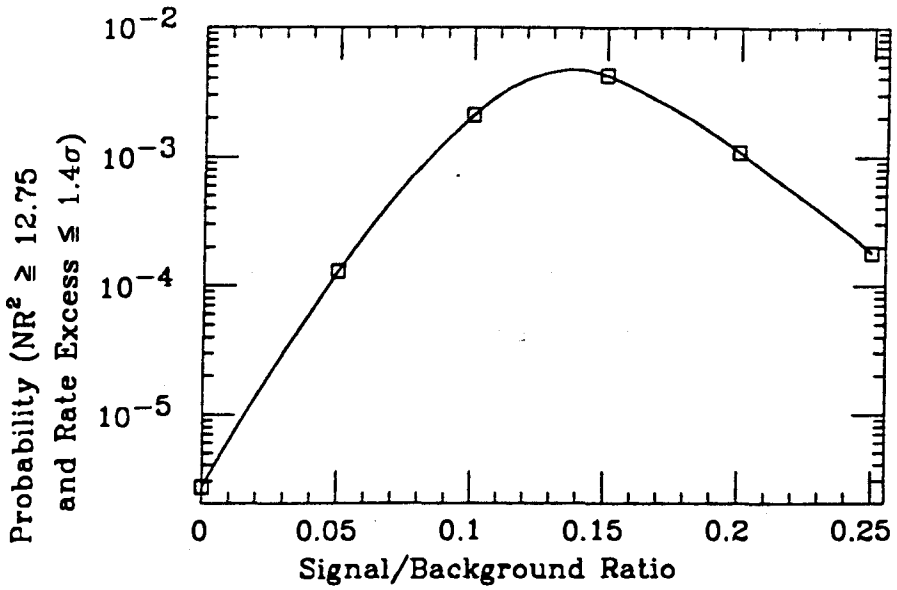


Figure 6.8: Probability of obtaining $z > 12.5$ *and* a rate deviation smaller than 1.4σ, as a function of signal-to-background ratio. Results derived from Monte Carlo simulations of intervals of 15 minute duration with a mean background rate of 0.7 Hz. The injected signal was a half-sine lobe of duty fraction 0.5.

also obtain large Rayleigh power with little overall rate excess. To be sure, such a combination is less likely than a combination of both large power and large rate excess. In Figure 6.8, we have plotted the probability, as determined by Monte Carlo (except for the case of zero signal, which was calculated directly), of obtaining a Rayleigh power larger than $z = 12.5$ *and* a rate deviation *less* than 1.4σ, as a function of signal-to-background ratio α. It is clear that such a combination is much more likely to arise from signal ($\alpha \approx 0.1$) than from pure random background. The γ-ray flux associated with the burst, based on the

number of excess counts accumulated, is approximately 10^{-9} cm^{-2} s^{-1} above an effective energy of about 400 GeV.

As discussed in Chapter 4, in the absence of signal, Rayleigh power and rate excess are uncorrelated. While this statement should be intuitive, it must be emphasized here in that calculations of the probability of our result having occurred from purely random background must make use of this fact. Thus, in testing against the null hypothesis, if we let p_1 represent the probability of obtaining a Rayleigh power larger than z_0, *i.e.*

$$p_1 = e^{-z_0}, \tag{6.1}$$

and p_2 represent that of obtaining a rate excess larger than n standard deviations, the total probability given N trials is

$$W = \left[1 - (1 - p_1)^N\right] p_2. \tag{6.2}$$

Here, we must consider the number of intervals searched, the number of independent periods scanned in each such interval, and the oversampling factor associated with such scanning. Given 416 intervals searched over 3 Fourier- independent periods, we have

$$N = 3 \times 3 \times 416 = 3744 \tag{6.3}$$

where the additional factor of three is associated with the oversampling. Using the values obtained from the interval of interest, we find $W = 8.4 \times 10^{-4}$. As explained in Chapter 4 and Appendix C, the confidence level associated with this combined probability is then

$$\epsilon = W(1 - \ln W) = 6.8 \times 10^{-3}. \tag{6.4}$$

By contrast, the lowest value of ϵ for any of the off-source intervals was 0.385.

To determine the degree, if any, to which the events contained in the burst interval differ from those in other intervals, the multiplicity distribution has been compared with the overall average with no statistical difference apparent. In addition, for each event the sum of the ADC's for participating PMT's was divided by the multiplicity of the event as an additional test for enhanced light in the events from the burst. Comparison of the distribution for the burst events with the average distribution again yields no indication of significant difference.

As a further check for indications of periodicity, the incoherent sums of the Rayleigh powers for each interval were formed as a function of test period. These sums were then treated as χ^2 variables with $2M$ degrees of freedom, where $M = 416$ is the number of intervals. The resulting distribution, plotted in Figure 6.9, does not provide any further evidence of periodicity. In a similar effort, the Rayleigh distributions shown in Figure 6.3 were tested against the expected exponential distributions using a χ^2 statistic. In neither case is the deviation from a random distribution indicated by this statistic.

6.3 Other Results from 1986

The observed periodicity in VHE γ-rays discussed in the previous section is complimented by the results of two other experimental groups which operated during the summer of 1986. On 11 June, at 7^h02^m UTC, the Whipple Observatory group observed a burst of activity from Her X-1 which persisted for at least 25 minutes (Lamb *et al.* 1988). The measured period of $1^s.2358$ is in excellent agreement

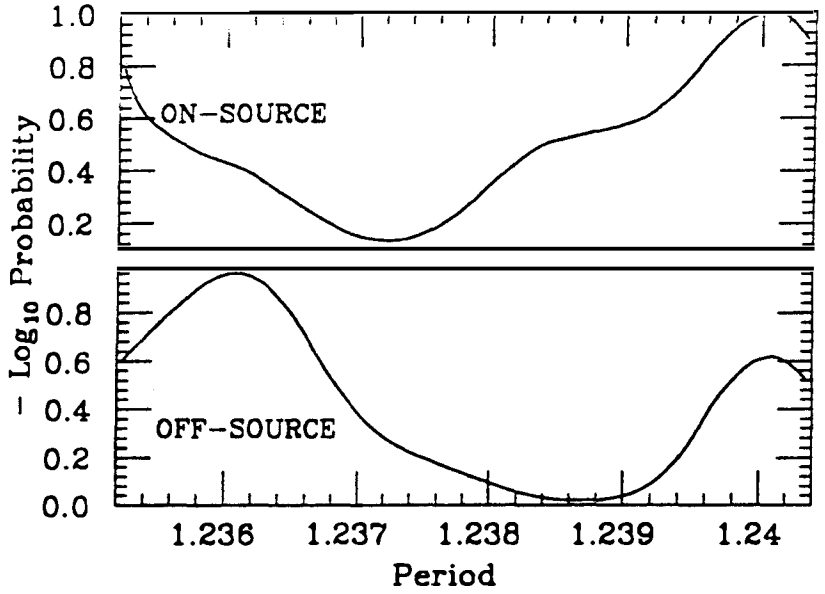


Figure 6.9: Results of incoherent sum of Rayleigh powers as a function of period. Here the Rayleigh powers for each search interval were added and the sum (for each test frequency) was treated as a χ^2 variable with $2M$ degrees of freedom, where M is the total number of intervals.

with the results obtained at Haleakala. The reported flux of 2×10^{-10} cm^{-2} s^{-1} above 1 TeV is compatible with that quoted in the previous section given the difference in threshold energies.

Operating at the much higher threshold of about 100 TeV, the Cygnus Array group (Dingus *et al.* 1987) observed a burst of activity from the direction of Her X-1 on 23 July 1986. The burst was characterized by an increase in rate which, when investigated for periodicity, was found to be pulsed with a period of $1^s.2358$, again in remarkable agreement with the Haleakala result. The three results are compared in Figure 6.10.

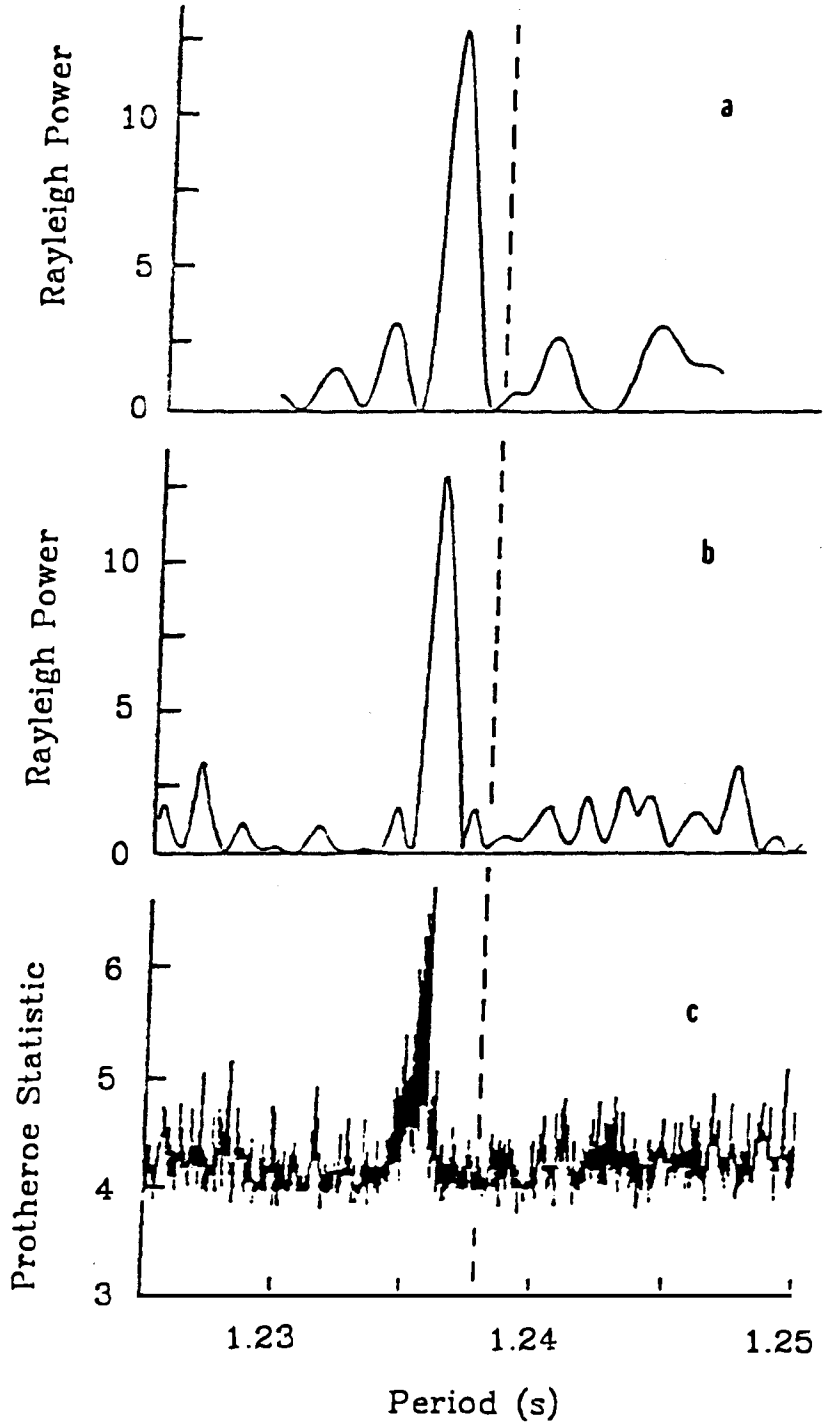


Figure 6.10: Power spectra for observations of Her X-1 in summer 1986: a)Haleakala Gamma Observatory, b)Whipple Observatory, c)Cygnus Array.

Optical observations of the Her X-1/Hz Her system during the summer of 1986 (Middleditch 1987) yield periodicity which is consistent with that expected from extrapolation of the most recent X-ray ephemeris (see Table 5.1). Thus, the deviations in the period measured in VHE and UHE γ-rays are not indicative of a change in the rotational period of Her X-1. Rather, the fact that the period differs from that expected indicates that the γ rays are not produced in the same region as the X-rays. This suggestion has been made previously (Gorham *et al.* 1986b) in reference to an observation of γ-ray activity which continued into the eclipse of Her X-1. Thus, these measurements provide important information regarding the physical conditions in the binary system.

6.4 Concluding Remarks

The production of pulsed γ-rays at a period differing from the rotational period of the neutron star can undoubtedly occur in a variety of ways. Details of the burst observed at Haleakala could set constraints on possible models. Here we consider a scenario which could pertain to the observed activity. We assume the presence of a charged particle beam which is produced in the vicinity of the neutron star and which rotates with the neutron star. The γ-ray production is assumed to result from the interaction of this particle beam with matter being accreted onto the neutron star.

As discussed in Section 2.4.2, the bulk of the material in an accretion disk orbits the central star with Keplerian velocities. In the region of the magnetopause, however, the magnetic field strength becomes sufficient to disturb this

pattern. As a result, matter may become locked onto magnetic field lines, thus being forced to corotate with the central star. While in a Keplerian orbit, however, the accreting matter may act as a moving target for a particle beam. In the process of becoming locked onto a field line, the matter may be channeled up out of the plane of the disk to a sufficient degree that the particle beam is obscured from the direction of the solar system (recall that the disk around Her X-1 is tilted) so that γ-rays may be directed toward the earth.

In Figure 6.9 we depict a broad particle (of width $\triangle\theta$) impinging on matter which is orbiting Her X-1 with a period P_K which is smaller than the rotational period of the neutron star P_{NS}. In such a scenario, γ-ray production occurs for as long as the target matter stays within the beam. This duration is merely determined by the angular width of the beam and the difference in periods of the orbiting matter and the rotating beam:

$$\tau = \frac{\triangle\theta}{2\pi}\left(\frac{1}{P_K} - \frac{1}{P_{NS}}\right)^{-1}. \tag{6.5}$$

Throughout this time, γ-rays will be beamed toward the earth only when the target matter is directly along the line-of-sight. Thus, the γ-ray period will be that of the orbiting matter, not that of the neutron star.

The period detected in the burst observed at Haleakala suggests a Keplerian orbit of radius $\approx 1.9 \times 10^6$ m which is very near the corotation radius. Let us consider the structure of the disk in this region. From equation 2.86, we know that the thickness of the disk at r_K is given by

$$\frac{H}{r_K} \simeq \frac{c_s}{v_K(r_K)}$$

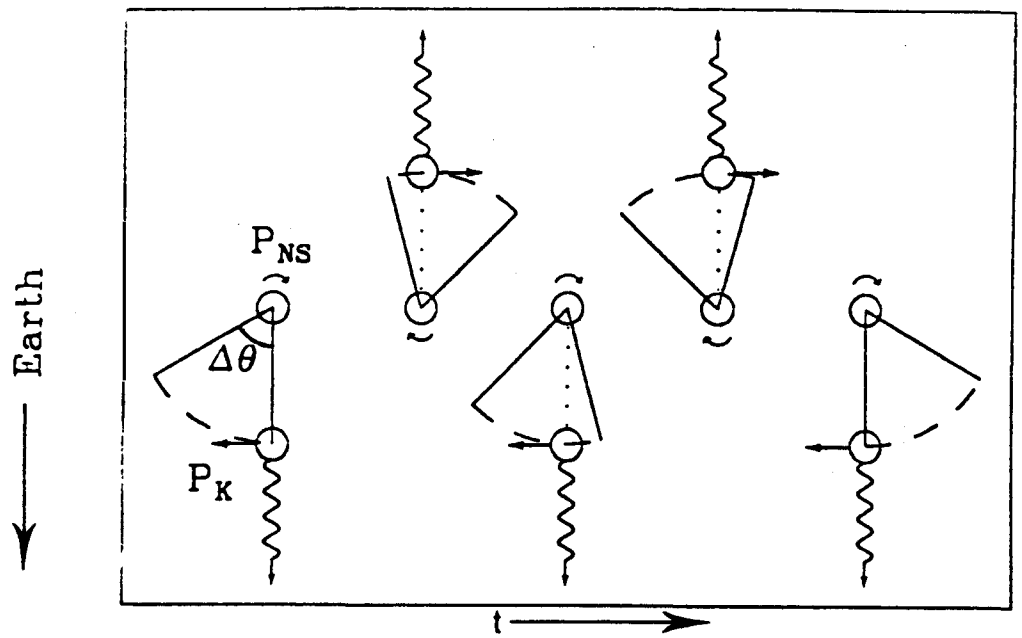


Figure 6.11: Schematic diagram depicting the interaction of a broad particle beam with target matter orbiting neutron star with a period smaller that the rotational period of the star (and particle beam). As time progresses, the target advances with respect to the beam.

where c_s is the speed of sound in the disk material and v_K is the Keplerian speed at r_K. For typical disk temperatures of $\sim 10^5$K, we have $c_s \sim 5 \times 10^4$ m s^{-1}. Since

$$v_K(r_K) = \frac{2\pi r_K}{P_K}$$

we have

$$H \approx \frac{cP_K}{2\pi} \sim 10^4 \text{ m}.$$

The density of the disk material in this vicinity (see, *e.g.* Frank *et al.* 1985) is of the order $\rho \sim 10^{-1}$ kg m^{-3}. In order for a region of target matter to act as

an efficient production source for γ-rays, the mean free path must be sufficient for production but small enough to guarantee that the γ-rays escape without considerable attenuation. Using, as an estimate, a path length of 1 gm cm^{-2} = 10 kg m^{-2}, we need $\sim 10^2$ m for the dimension of the target. This is considerably smaller than the disk thickness. It is not difficult to imagine that such tiny "droplets" of matter will be removed from the disk in regions of instability such as that of the magnetopause.

Given the existence of a target of radius $\sim 10^2$ m orbiting Her X-1 at a distance of $\sim 10^6$ m, with a period of $1^s.23593$, we can predict the approximate characteristics of the associated γ-ray signal. The angular size of the target is

$$\theta \sim \frac{10^2 \text{ m}}{10^6 \text{ m}} \approx 6 \times 10^{-3} \text{ degrees.}$$

As such, we would expect the duty fraction of the associated light curve to be quite small. This is in conflict with the relatively broad light curve observed. Given the duration of the observed activity, $\tau \sim 900$ s, the angular width of the proton beam must be at least

$$\triangle\theta = 2\pi\tau\left(\frac{1}{P_K} - \frac{1}{P_{NS}}\right) \approx 360^\circ.$$

If the beam is this broad, the burst duration is clearly determined by some other factor (possibly the dissipation of the target upon capture by the field lines).

We see, then, that although the above scenario can produce a γ-ray signal whose period is distinctly different from that of the neutron star, there are details which are difficult to reconcile with the data. Certainly features such as the period derivative which may be associated with the transition between Keplerian and corotation orbits and the steering of the particle beam by the field

of the neutron star can be expected to complicated matter considerably. Still, it is clear that details associated with the interaction of the inner regions of the accretion disk with the neutron star may play a major role in determining the characteristics of VHE γ-ray bursts in such systems. Further studies of such activity may provide direct information on the structure of such systems in addition to providing clues as to the nature of particle acceleration and the origin of the cosmic rays.

Appendix A

Calculation of Muon Trigger Rate

Because muons can penetrate deep into the atmosphere, they represent a potential source of triggers for the VHE γ-ray telescope. In this Appendix, the rate at which such triggers are expected to occur for the Haleakala telescope is calculated.

Let N_c be the number of Čerenkov photons produced per pathlength in the wavelength band to which the telescope is sensitive.

$$N_c = \frac{1}{137c}\int_{\nu_{min}}^{\nu_{max}} \sin^2\theta_c 2\pi d\nu \tag{A.1}$$

Integrating,

$$N_c = \frac{2\pi}{137}\sin^2\theta_c\left(\frac{1}{\lambda_{min}} - \frac{1}{\lambda_{max}}\right). \tag{A.2}$$

The effective bandpass of the telescope (see Figure 3.23) is 325 nm $\leq \lambda \leq$ 425 nm. Thus,

$$N_c = \frac{2\pi}{137}\left(\frac{1}{325\text{ nm}} - \frac{1}{425\text{ nm}}\right)\sin^2\theta_c = 3.3\times 10^4 \sin^2\theta_c \text{ photons m}^{-1}.$$

The refractive index at an altitude h is given by

$$n = 1 + (2.7\times 10^{-4})e^{-\frac{h}{h_0}} \tag{A.3}$$

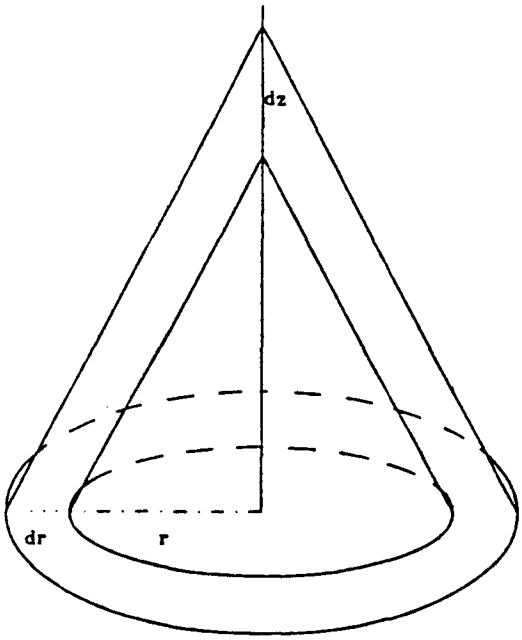


Figure A.1: Geometry of Čerenkov radiation.

where $h_0 \approx 8.2$ km. At Haleakala, $h = 3$ km so that $n = 1.00019$. The Čerenkov angle is given by

$$\theta_c = \cos^{-1}\left(\frac{1}{\beta n}\right) \approx 1^\circ.12 = 1.95 \times 10^{-2} \text{ rad.} \tag{A.4}$$

Thus, $N_c \sim 12.6$ photons m^{-1}.

From the geometry of the Čerenkov radiation (Figure A.1), light produced at a height z hits the ground at a radius r from the muon point of impact where

$$r = \theta_c z \Rightarrow dr = \theta_c dz. \tag{A.5}$$

The photon density ρ produced by the muon is then

$$\rho(r) = \frac{N_c\, dz}{2\pi r\, dr} = \frac{N_c}{2\pi r \theta_c} = \frac{1.03 \times 10^2}{r}\text{m}^{-1} \tag{A.6}$$

The telescope must process ~ 8 photoelectrons in order to trigger. The average quantum efficiency of the PMT's over the effective bandpass is $\sim$ 20% while the average transmission of the filters is $\sim$ 67%. The reflectivity of the mirrors is more difficult to judge but is probably $\sim$ 65% (it is about 85% when the aluminum is perfect). There are also losses due to reflections at the PMT surface because of the relatively large angle with which photons approach from the outer portions of the mirrors as well as overall losses due to incomplete coverage of the apertures by the PMT's. An estimate of the overall efficiency of the optics system is $\sim 5\%$ (or slightly lower). Thus, to get 8 photoelectrons, $\sim$ 160 photons on the $\sim$ 10.4 square meters of mirror area are needed. This yields a critical density $\rho = 15\mathrm{m}^{-2}$. The effective trigger radius is then

$$r_e = \frac{1.03 \times 10^2}{\rho_c} = 6.8 \text{ m}.$$

The rate for vertical muons with $E \geq 2\mathrm{GeV}$, at $h = 3.2$ km, has been measured (Shen & Chang 1979) as

$$I_\mu(\geq 2\mathrm{GeV}) = 49 \text{ m}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}. \tag{A.7}$$

However, the threshold for production of Čerenkov radiation is $\sim$5.4GeV for muons at h = 3km. The muon energy spectrum is illustrated in Figure A.2. Above 5 Gev, the integral spectral has the form E^{-2}. Thus, we have

$$I_\mu(\geq 5.4\mathrm{GeV}) \approx 6.7 \text{ m}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}. \tag{A.8}$$

The solid angle of acceptance for muons to trigger the telescope is just the

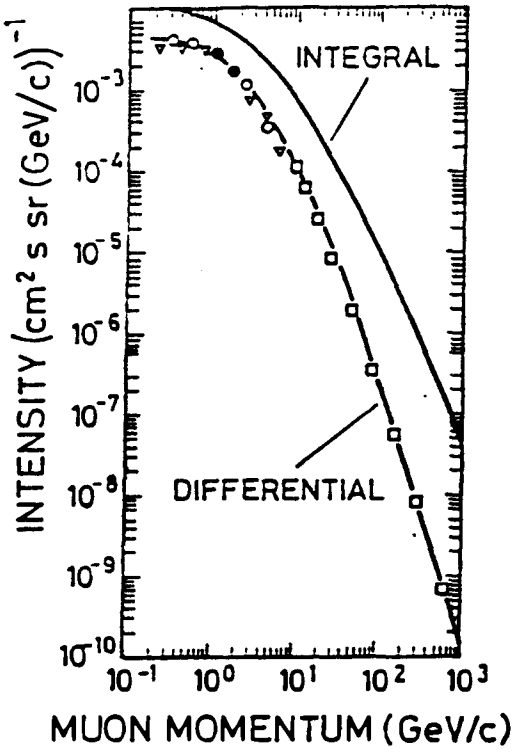


Figure A.2: Vertical differential and integral momentum spectra of muons at sea level.

solid angle of the telescope aperture:

$$\Omega = 5.4 \times 10^{-4} \text{ sr.}$$

The muon trigger rate is thus

$$R_\mu = I_\mu \pi r_e^2 \Omega = 0.1 \text{ Hz.} \tag{A.9}$$

One can see from the density calculation that a non-uniformity of photons on the mirrors exists (although, since we are operating in a very low-statistics regime whereby the bulk of our triggers have an average of 1 or 2 photoelectrons per mirror, such nonuniformity may be buried in the noise). The spread in

the arrival times of the Čerenkov light from a single muon is very small. This feature may be difficult to utilize in that large showers, with correspondingly bright leading edges, will also yield very narrow triggers. Although the trigger rate calculated above is likely to represent an overestimate, efforts to identify such muon triggers are clearly worthwhile and are in progress.

Appendix B

Arrival Time Corrections for Eccentric Orbits

In Chapter 4 we summarized the necessity for correcting arrival times to the barycenter of the binary system in which the signals originate. Such a procedure is particularly simple in the case of a circular orbit and has been outlined in Section 4.1.2. For binary systems whose eccentricity is not small (*e.g.* the system containing 4U 0115+63), the procedure is more complicated.

The orbital elements of a binary system are illustrated in Figure A.1 which represents a celestial sphere centered on the primary star A. The directions of east and north are indicated by the points E and N. The *position angle of the ascending node*, Ω, is defined by the arc NL while the orbital inclination i is defined by the spherical angle VLS. The *periastron* (point P) is the position of the pulsar when it is closest to the primary. The angle LAP' defines the *argument of periastron*, α. If the pulsar is, at some time t, located at point B, then the angle $P'AB'$ is called the *true anomaly*, β. The remaining orbital elements are the semimajor axis a, the orbital eccentricity e, the epoch of periastron τ, and the orbital period T. The seven parameters $(a, e, i, \alpha, \Omega, \tau, T)$, then, define the orbit.

Orbital parameters for X-ray binaries are determined by observing Doppler

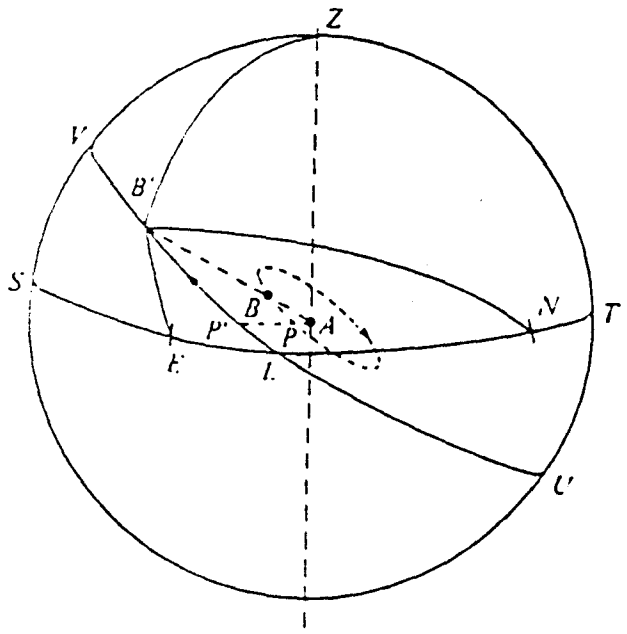


Figure B.1: Orbital elements of a binary star system.

shifts in the pulsar period P (just as those for spectroscopic binaries are determined by observing such shifts in a spectral line of one of the stars). With such observations, however, the full set of orbital elements cannot be determined. Instead, one can only obtain measurements of $(a \sin i, e, \alpha, \tau, T)$. These elements being determined, one can then reverse the process to determine the Doppler shifts expected at some other time.

When one views an elliptical orbit which is not confined to the observation plane, the apparent orbit is still an ellipse with the primary star located at a focus. Further, the center of the apparent orbit is also the center of the true orbit and, as such, allows one to determine the position of periastron (see Figure A.2). We may then, for a time t, define a mean anomaly

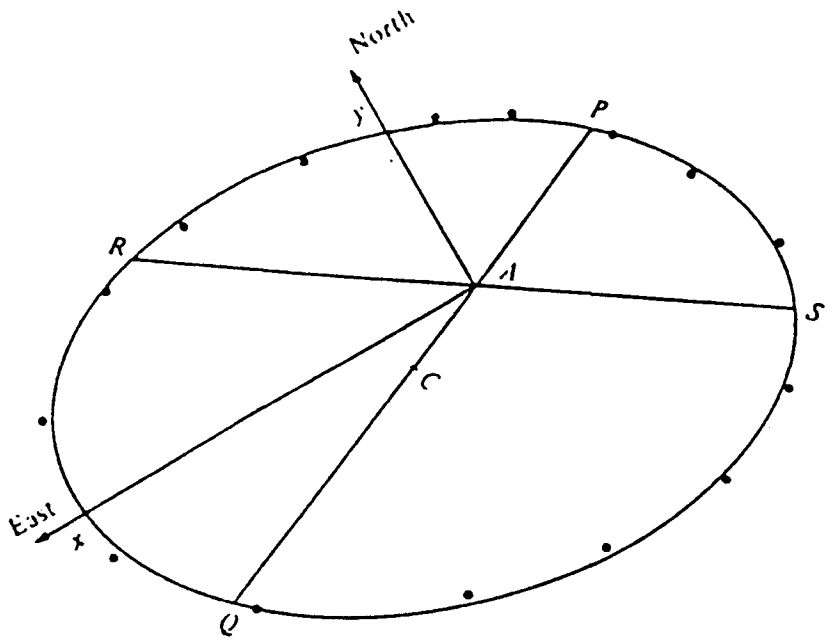


Figure B.2: Apparent orbit of an inclined binary system. Because the primary star is located at a focus and the center of the ellipse of the apparent orbit is also the center of the actual orbit, the location of the periastron point P (which lies on the extension of the line from the center through the primary star) can be determined.

$$M = \frac{2\pi}{T}(t - \tau). \tag{B.1}$$

The *eccentric anomaly* E is defined such that it is zero at periastron and increases by 2π in the course of one orbit. It is related to the mean anomaly by Kepler's equation

$$M = E - e \sin E. \tag{B.2}$$

Hence, given a time t, the mean anomaly may be determined and Kepler's equation can be solved numerically to yield the eccentric anomaly. The true anomaly

is then determined by

$$\tan\frac{\beta}{2} = \left(\frac{1+e}{1-e}\right)^{\frac{1}{2}} \tan\frac{E}{2}. \tag{B.3}$$

Knowing β, the correction required to account for the time delay (or advance) relative to the barycenter of the binary system is given by

$$\Delta t_{source} = \frac{a \sin i}{c} \frac{1-e^2}{1+e\cos\beta} \sin(\alpha+\beta). \tag{B.4}$$

The sign convention for the binary orbit is defined such that this correction is to be subtracted from the observation time.

Appendix C

Confidence Level for Combined Probabilities

In order to properly interpret the statistical significance of the result of a particular test, it is necessary to carefully assess the selection criterion used to characterize intervals of interest. In particular, when two or more independent measurements are used in combination to define an overall test statistic, the resulting confidence level is not the simple product of the probabilities associated with the individual tests. To understand this distinction between confidence level and probability, we must review the definition of confidence level and interpret this definition properly in the multi-variable case.

Let S represent the set of possible results α from the overall test chosen to characterize a set of data. Let s represent the subset of S which consists of test results that have been identified as being sufficient to reject a particular hypothesis H_0. Then the *confidence level* ϵ associated with the test is defined as the probability that α falls in s given that H_0 is true; *i.e.* it represents the probability of incorrectly rejecting the hypothesis H_0. As an example, in the Rayleigh test, the confidence level associated with $z = 10$ in a single test is 4.54×10^{-5}; if this value of z is used to reject the hypothesis that the data from which it is derived is purely random (the null hypothesis), one will incorrectly

reject the hypothesis about 4 or 5 times in 10^5 trials.

Suppose the region S is defined by results α_1 and α_2 of two independent tests. Let ϵ_1 and ϵ_2 represent the confidence levels associated with the two tests. Then the probability of simultaneously satisfying ϵ_1 and ϵ_2 is certainly the product $\epsilon_1\epsilon_2$. However, the overall confidence level ϵ is **not** given by this product. Rather, we must seek the probability of incorrectly rejecting H_0 based on the fact that the combined test statistic lies within the critical region defined by $\epsilon_1\epsilon_2$. However, there are many values of ϵ_1 and ϵ_2 which yield the same product. Hence, the probability of obtaining a result in the critical region defined by ϵ is actually greater than $\epsilon_1\epsilon_2$. Let

$$q = \epsilon_1\epsilon_2 \tag{C.1}$$

represent the value of the combined statistic which defines the critical region s. Then the confidence level associated with q is given by the fraction of the area of S represented by the critical region s. Wallis (1941) points out that the region s is defined by the hyperbola

$$\epsilon_2 = \frac{q}{\epsilon} \tag{C.2}$$

(see Figure C.1) so that the confidence level is merely given by

$$\epsilon = q + \int_q^1 \frac{q}{\epsilon_1} d\epsilon_1. \tag{C.3}$$

Thus,

$$\epsilon = q(1 - \ln q). \tag{C.4}$$

In general, when the critical region is defined by N independent probabilities

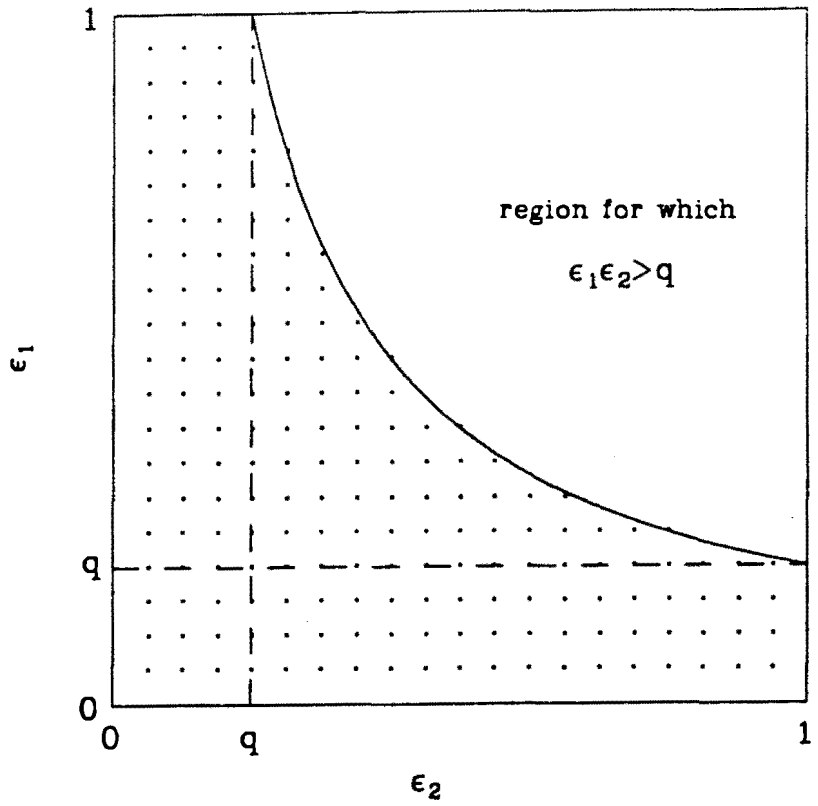


Figure C.1: Critical region defined by two independent statistics. The confidence levels ϵ_1 and ϵ_2 are used to define the region $\epsilon_1\epsilon_2 \leq q$. The overall confidence level is given by the area indicated.

whose product is q, the overall confidence level is given by

$$\epsilon = q \sum_{i=0}^{N-1} \frac{-\ln q}{k!}. \tag{C.5}$$

References

Astronomical Almanac for the Year 1986, Nautical Almanac Office, U.S. Gov. Printing Office (Washington: 1986).

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

Bahcall, J.M., and Bahcall, N.A. 1972, IAU Circ. No. 2427.

Baltrusaitis, R.M. it et al. 1985, **Apl J. Lett. 293**, L69.

Baym, G., and Pethick, C. 1975, **Ann. Rev. Nuc. Sci. 25**, 27.

Baym, G. *et al.* 1971, **Ap. J. 170**, 299.

Boynton, P.E. 1978, in *Physics and Astrophysics of Neutron Stars and Black Holes* (R. Giacconi and R. Ruffini eds.), North-Holland Publishing Co., Amsterdam.

Brecher, K. 1972, **Nature 239**, 325.

Brigham, E.O. 1974, *The Fast Fourier Transform*, Prentice-Hall Inc., New York, New York.

Cameron, A.G.W. 1965, **Nature 205**, 787.

Cameron, A.G.W. 1970, **Ann. Rev. Astron. Astrophys. 8**, 179.

Canuto, V. 1974, **Ann. Rev. Astron. Astrophys. 14**, 167.

Canuto, V. 1975, **Ann. Rev. Astron. Astrophys. 15**, 335.

Chadwick, P.M. *et al.* 1987, in *Very High Energy Gamma Ray Astronomy* (K.E. Turver ed.), D. Reidel Publ., Dordrecht, Holland.

Chanmugan, G., and Brecher, K. 1985, **Nature 313**, 767.

Cheng, K.S. *et al.* 1986, **Ap. J. 300**, 500.

Clayton, D.D. 1968. *Principles of Stellar Evolution and Nucleosynthesis*, McGraw-Hill, New York, New York.

Clayton, D.D. 1974, **Nature 249**, 131.

Clayton, D.D. *et al.* 1975a, **Ap. J. 199**, 494.

Clayton, D.D. *et al.* 1975b, **Ap. J. 201**, 489.

Cooley, J.W., and Tukey, J.W. 1971, **Math. Computation 19**, 617.

Cooke, B., and Page, C. 1975, **Nature 256**, 712.

Cox, J.P., and Guili, R.T. 1968, *Principles of Stellar Structure, I & II*, Gordon and Breach.

Crosa, L., and Boynton, P.W. 1980, **Ap. J. 235**, 999.

Davis, R. 1964, **Phys. Rev. Lett. 12**, 303.

de Jager, O.C. 1987, *The Analysis and Interpretation of VHE Gamma Ray Measurements* - Ph.D. Thesis (Potchefstroom University).

Delgado, A.J. *et al.* 1983, **Astr. Astrophys. 127**, L15.

Davidsen, A. *et al.* 1972. **Ap. J. Lett. 177**, L97.

Dreter, J.E. *et al.* 1981, **Ap. J. 247**, 1003.

Dowthwaite, J.C. *et al.* 1984, **Nature 309**, 691.

Eichler, D., and Vestrand, W.T. 1985, **Nature 318**, 345.

Fabian, A. *et al.* 1973, **Nature 244**, 212.

Finzi, A., 1965 **Phys. Rev. Lett. 15**, 599.

Forman, W. *et al.* 1972, **Ap. J. Lett. 177**, L103.

Fowler, W.A. 1972, **Nature 238**, 24.

Frank, J., King, A.R., and Raine, D.J. 1985, *Accretion Power in Astrophysics*, Cambridge University Press.

Giacconi, R. *et al.* 1973, **Ap. J. 184**, 227.

Gibson, A.I. *et al.* 1982, **Nature 296**, 833.

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

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

Gorham, P.W. 1986, *Ground-Based Detection of Gamma-Rays Above 250 GeV From Hercules X-1*, Ph.D. Thesis (University of Hawaii).

Gorham, P.W. *et al.* 1986a, **Ap. J. 309**, 114.

Gorham, P.W. *et al.* 1986b, **Ap. J. Lett. 308**, L11.

Gorham, P.W., and Learned, J.G. 1986, **Nature 323**, 422.

Greisen, K. 1956, in *Progress in Cosmic Ray Physics* Vol III (ed. J.G. Wilson), North Holland, Amsterdam.

Gunn, J.E., and Ostriker, J.P. 1969, **Phys. Rev. Lett. 22**, 728.

Hewish, A. 1975, **Rev. Mod. Phys. 47**, 567.

Hewish, A. *et al.* 1968, **Nature 217**, 709.

Holt, S.S., and McCray, R. 1982, **Ann. Rev. Astron. Astrophys. 20**, 323.

Iben, I. 1974, **Ann. Rev. Astron. Astrophys. 12**, 215.

Jones, C., and Forman, W. 1976, **Ap. J. Lett. 209**, L131.

Joss, P.C. *et al.* 1980, **Ap. J. 235**, 592.

Katz, J.I. 1973, **Nature 246**, 87.

Lamb, R.C. *et al.* 1988, **Ap. J. Lett.** (in press).

Lang, K.R. 1980, *Astrophysical Formulae*, Springer-Verlag, Berlin.

Liller, W. 1972, IAU Circ. No. 2415.

Longair, M.S. 1981, *High Energy Astrophysics*, Cambridge University Press, Cambridge.

Lovelace, R.V.E. 1976, **Nature 262**, 649.

Marshak, M. *et al.* 1985, **Phys. Rev. Lett. 54**, 2079.

McCray, R. *et al.* 1982, **Ap. J. 262**, 301.

Michel, F. 1982, **Rev. Mod. Phys. 54**, 1.

Middleditch, J. 1975, Ph.D. Thesis, LBL Report No. 3639.

Middleditch, J. 1983, **Ap. J. 275**, 278.

Middleditch, J. 1987, private communication.

Middleditch, J., and Nelson, J. 1976, **Ap. J. 208**, 567.

Middleditch, J. *et al.* 1983, **Ap. J. 274**, 313.

Middleditch, J. *et al.* 1985, **Ap. J. 292**, 313.

Misaki, A. 1970, in *Proc. 11th Int. Conf. on Cosmic Rays, Budapest 1969* Vol 3 (ed. A. Somogyi), Akademiai Kiado, Budapest.

Nagase, F. *et al.* 1984, **Publ. Astron. Soc. Japan 36**, 667.

[illegible] H. *et al.* 1985, **Space Sci. Rev. 40**, 347.

Oppenheimer, J.R., and Volkoff, G.M. 1939, **Phys. Rev. 55**, 374.

Paczynski, B. 1971, **Ann. Rev. Astron. Astrophys. 9**, 183.

Parmar, A.N. et al. 1985, **Nature 313**, 119.

Perkins, D.H. 1982, *Introduction to High Energy Physics*, Addison-Wesly Publishing Company, Reading, MA.

Petterson, J.A. 1975, **Ap. J. Lett. 201**, L61.

Petterson, J.A. 1977, **Ap. J. 218**, 783.

Rayleigh, Lord 1905, **Nature 72**, 318.

Resvanis, L.K. *et al.* 1987, in *Very High Energy Gamma Ray Astronomy* (K.E. Turver ed.), D. Reidel Publ., Dordrecht, Holland.

Resvanis, L.K. *et al.* 1988, **Nucl. Instr. and Meth.** (in press)

Roberts, W.J. 1974, **Ap. J. 187**, 575.

Rossi, B. 1952, *High Energy Particles*, Prentice-Hall Inc., New York, New York.

Rossi, B., and Greisen, K. 1941, **Rev. Mod. Phys. 13**, 240.

Ruderman, M.A., and Sutherland, P.G. 1975, **Ap. J. 196**, 51.

Samorski, M., and Stamm, W. 1983, **Ap. J. 268**, L17.

Scargle, J.D. 1982, **Ap. J. 263**, 835.

Shen, C., and Chiang, I. 1979, **PICRC 10**, 8.

Stanev, T. *et al.* 1985, **Phys. Rev. D 32**, 1244.

Staubert, R. *et al.* 1983, **Astr. Astrophys. 117**, 215.

Szentgyorgyi, A. 1986, *A Search for Very High Energy Gamma Ray Emission from Cygnus X-3*, Ph.D. Thesis, University of Wisconsin.

Tananbaum, H. *et al.* 1972, **Ap. J. Lett. 174**, L143.

Trümper, J. *et al.* 1978, **Ap. J. Lett. 219**, L105.

Trümper, J. *et al.* 1986, **Ap. J. Lett. 300**, L63.

Van den Heuvel, E.P.J. 1978, in *Physics and Astrophysics of Neutron Stars and Black Holes* (eds. R. Giacconi and R. Ruffini), North Holl. Publ. Co. Amsterdam, p.828-871.

Van den Heuvel, E.P.J. 1987, in *High Energy Phenomena Around Collapsed Stars*

(ed. F. Pacini), D. Reidel Publ. Co., Dordrecht, Holland, p.1-32.

Vestrand, W.T., and Eichler, D. 1982, **Ap. J. 261**, 251.

Wallis, W.A. 1942, **Econometrica 10**, 229.

Zombeck, M.V. 1982, *Handbook of Space Astronomy and Astrophysics*, Cambridge University Press, Cambridge.