NASA Reference Publication 1386

September 1997

Gamma-Ray Astrophysics

New Insight Into the Universe

Second Edition

Carl E. Fichtel
Jacob I. Trombka
Goddard Space Flight Center
Greenbelt, Maryland

National Aeronautics and Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771
1996

This publication is available from the NASA Center for AeroSpace Information,
800 Elkridge Landing Road, Linthicum Heights, MD 21090-2934, (301) 621-0390.

PREFACE

During the 15 years that have passed since the first edition of this book was published, there has been a major increase in our knowledge of gamma-ray astronomy. Much of this advance arises from the extensive results that have been forthcoming from the Compton Gamma-Ray Observatory. There has been the discovery of a new class of gamma-ray objects, namely high-energy gamma-ray-emitting blazars, a special class of Active Galactic Nuclei, whose basic high-energy properties now seem to be understood. A much improved picture of our galaxy now exists in the frequency range of gamma rays. The question of whether cosmic rays are galactic or metagalactic now seems settled with certainty. Significant new information exists on the gamma-ray properties of neutron star pulsars, Seyfert galaxies, and gamma-ray bursts. Substantial new insight has been obtained on solar phenomena through gamma-ray observations. Hence, this seemed to be an appropriate time to write a new edition of this book to add the important scientific implications of these many new findings.

The special importance of gamma-ray astrophysics had long been recognized by many physicists and astronomers, and theorists had pursued many aspects of the subject well before the experimental results began to become available. The slower development of the experimental side was not because of a lack of incentive, but due to the substantial experimental difficulties that had to be overcome. Thus, as the gamma-ray results became available in much greater number and detail, it was possible to build upon the theoretical work that already existed and to make substantial progress in the study of many of the phenomena involved. Consequently, a much better understanding of many of the astrophysical phenomena mentioned here and others is now possible.

Our principal aims in writing this book are the same as they were for the first edition: to provide a text which describes the significance of gamma-ray astrophysics and to assemble in one place a treatment of gamma rays emitted from bodies in the solar

system, from objects in our galaxy, as well as from interactions between cosmic rays and the interstellar medium, and from beyond our galaxy. Thus, this book is intended for those in astrophysics who wish to have the opportunity to learn more about the evolving field of gamma-ray astronomy and its relationship to the high-energy, evolutionary processes occurring in the universe.

We have assumed that the readers will have had basic undergraduate courses in physics, including electromagnetic, atomic, and nuclear theory, and will be familiar with the rudiments of astronomy. We have attempted to write those portions where reference to the content of more advanced astronomy and physics courses is appropriate in such a way that the general flow of thought will not be interrupted in a significant way for readers unfamiliar with the details of these particular theories, yet those with a full understanding of the physics involved will see the validity and qualifications clearly.

After considering whether to discuss the gamma-ray production processes together near the beginning of the book or to describe each as it occurred in the description of individual astrophysical subjects, we chose the latter approach. This decision resulted in part from what appeared to be a more natural development and partially from the students', and our own, enthusiasm for the study of the astrophysical phenomena and their desire to learn about each subject in depth as it was presented to them.

The last three chapters of the book provide a general discussion of the experimental aspects of the field that seemed best treated together, separately from the astrophysical aspects of gamma-ray astronomy that are discussed in the first ten chapters. A list of symbols and a subject index are provided at the end for your reference.

The book would not have been possible without the cooperation of a very large number of individuals. We gratefully acknowledge the help of many scientists with whom we had discussions over the years, who provided us with material —some of which had not appeared in the literature at the time— and who were always ready to answer questions about their work.

CONTENTS

1 INTRODUCTION TO GAMMA-RAY ASTRONOMY 1

2 GAMMA-RAY OBSERVATIONS OF THE SOLAR SYSTEM

2.1 Introduction 11
2.2 Models of Solar System Evolution 11
2.3 Solar System Gamma-Ray Astronomical Observations 14

3 PLANETS, COMETS, AND ASTEROIDS

3.1 Introduction 18
3.2 Interaction Processes 19
3.3 Solar X-Ray Fluorescence 20
3.4 Charged-Particle X-Ray Fluorescence 27
3.5 Natural Radioactivity 27
3.6 Primary and Secondary Galactic Cosmic-Ray Interactions 30
3.7 Solar Proton Interactions 45
3.8 Experimental Results 46
3.9 Apollo X-Ray Spectrometer Results 48
3.10 The Apollo Gamma-Ray Spectrometer 53

4 SOLAR OBSERVATIONS

4.1 Introduction 63
4.2 Interaction Processes 65
4.3 Charged- and Neutral-Particle Interactions 65
4.4 Continuum Emission 75

4.5 Decay of π° Mesons 76
4.6 Positron Annihilation and Formation of Positronium 77
4.7 Experimental Results 79
4.8 Solar-Flare Observation in the Energy Region 0.1 to 10 MEV 79
4.9 High-Energy Continuum Radiation 84

5 DIFFUSE GALACTIC EMISSION, COSMIC RAYS, AND THE INTERSTELLAR MEDIUM

5.1 Introduction 90
5.2 Cosmic Rays, Interstellar Matter, Photons, and Magnetic Fields 91
5.3 Interaction Processes for Continuum Emission 98
5.4 Gamma-Ray Lines 103
5.5 Galactic Dynamic Balance and the Cosmic-Ray Distribution 105
5.6 Current High-Energy Gamma-Ray Results and Their Interpretation 107
5.7 Current Low-Energy Gamma-Ray Results and Their Interpretation 118
5.8 Summary 121

6 GALACTIC COMPACT OBJECTS

6.1 Introduction 124
6.2 Individual Neutron Stars and Pulsars 124
6.3 Gamma-Ray Pulsar Overview 146
6.4 Accretion-Driven Pulsars 150
6.5 Supernovae 153
6.6 Black Holes 156
6.7 Mini Black Holes 158
6.8 Other High-Energy Gamma-Ray Sources 160

7 NORMAL GALAXIES

7.1 Introduction 167
7.2 The Small Magellanic Cloud 168
7.3 The Large Magellanic Cloud 170

7.4 Supernovae in Other Galaxies 171
7.5 Some Other Relatively Close Galaxies 173

8 ACTIVE GALAXIES

8.1 Introduction 176
8.2 Seyfert and Radio Galaxies 177
8.3 Blazars 184
8.4 The Evolution of Gamma-Ray-Emitting Blazars 198
8.5 Two Types of Blazars 203
8.6 The Origin of the Gamma-Ray Emission From Blazars and the Requirements for the Parent Relativistic Particles 208
8.7 Particle Acceleration 211
8.8 Summary 214

9 DIFFUSE RADIATION

9.1 Introduction 219
9.2 Observations 220
9.3 Cosmology 224
9.4 The Origin of the Diffuse Radiation 230
9.5 Summary 235

10 GAMMA-RAY BURSTS

10.1 Introduction 238
10.2 Distribution of Sources in Direction and Intensity 238
10.3 Energy and Temporal Distributions in the Low-Energy Gamma-Ray Region 242
10.4 High-Energy Gamma-Ray Emulsion 244
10.5 Possible Sources and Mechanisms 246

11 GAMMA-RAY INTERACTION PROCESSES

11.1 Introduction 250
11.2 Photoelectric Effect and Fluorescence Radiation 253
11.3 Pair Production and Annihilation Radiation 254
11.4 Compton Scattering 262
11.5 Electron Leakage and Bremsstrahlung 265

12 DETECTORS FOR ENERGIES LESS THAN 25 MeV

12.1 Introduction 267
12.2 Properties of Detectors 268
12.3 Detector Response to Monoenergetic Radiation 286
12.4 Detector Efficiency 296
12.5 Radiation Damage, Annealing 299
12.6 Sources of Gamma-Ray Background in the Space Environment 301
12.7 Detector Flight Systems 322

13 DETECTORS FOR ENERGIES GREATER THAN 25 MeV

13.1 Introduction and History 343
13.2 The High-Energy Gamma-Ray Interaction 348
13.3 The 20 MeV to 50 GeV Region 351
13.4 The Greater Than 20 GeV Region 359
13.5 Angular Resolution 361
13.6 Energy Measurement and Resolution 364

SUBJECT INDEX 368

CHAPTER 1

GAMMA-RAY ASTRONOMY IN PERSPECTIVE

As we begin the study of gamma-ray astronomy and the astrophysical phenomena whose secrets these highest-energy photons best reveal, it is important to remember that, throughout most of history, our view of the heavens was restricted to the visible portion of the electromagnetic spectrum. Our concept of the Universe even now is strongly influenced by knowledge gained from this very important, but relatively quite small, wavelength band. Many major advances in our understanding of the Universe have come in this century as optical telescopes have grown larger. For example, in the 1920s it became clear that many galaxies beyond our own exist and that the Sun is not even at the center of our galaxy, but is, in fact, far from the Galactic center. These findings destroyed the prevalent belief that our solar system was the center of the Universe. Soon after the discovery of the existence of other galaxies, of which there are now known to be over a hundred billion, it was found that they are systematically receding from each other. This discovery lead to the present "big bang" theory of the Universe. It was further realized that our Sun is a rather ordinary star and that many galaxies release far more energy than our own. Still, our present knowledge of the Universe would be much more restricted if the visible wavelength range had remained the sole source of information. Many phenomena can be understood or even revealed only by the investigation of emission in other parts of the electromagnetic spectrum.

The first region of the spectrum beyond the optical to be explored was the radio range. Although radio emission from the Galaxy was detected in the 1930s, it was not until the 1950s that radio astronomy began to have a major impact. Studies of the 21 cm line, resulting from the hyperfine splitting of the ground state of atomic hydrogen, showed the full spiral structure of our galaxy. The 21-cm line emission, and the continuum radio emission generally

accepted to be synchrotron radiation, which indicates the presence and distribution of relativistic electrons, are just two examples of astrophysical information to come from a wavelength region outside the familiar, visible one.

As a result of the advent of the space age, which permitted instruments to be carried above the Earth's atmosphere, most of the remainder of the electromagnetic spectrum has now been examined, at least in a preliminary way. Some portions, such as the ultraviolet, have been extensively studied. Studies in the infrared have expanded our knowledge of molecules in interstellar space and the cosmic background, and extreme ultraviolet and x-ray observations have revealed much about the hotter regions.

Gamma-ray astronomy was the last major wavelength range to yield its wealth of information. This late development grew from a combination of factors, including the need to place gamma-ray telescopes above the Earth's atmosphere, the requirement to develop rather complex instruments, and the relatively low intensity of gamma-ray photons, particularly in relation to the charged-particle cosmic-ray intensity, rather than from a failure of scientists to realize its unique and significant potential contributions. Even though the photon intensity is low, the energy emitted in the gamma-ray range may be, and in several cases is, quite high, because each photon carries a very large amount of energy, and the gamma-ray frequency range is very broad.

There are several aspects of the Universe to which gamma-ray astronomy speaks more clearly than observations at other wavelengths, and, for some astrophysical studies, gamma rays carry unique information about the source. Gamma-ray astronomy permits the direct study of the largest transfers of energy occurring in astrophysical processes, including rapid expansion processes, explosions, high-energy particle acceleration, gravitational accretion onto superdense objects, the fundamental process of the building of the elements, and even particle-antiparticle annihilation. Synthesized nuclear material can reveal its presence through the emission of characteristic gamma-ray lines from certain excited nuclei. The nucleonic cosmic rays, representing one of the three approximately equal expansion pressures (the others resulting from the magnetic fields and the kinetic motion of matter) in the Galaxy, reveal themselves uniquely through the high-energy gamma rays emitted during the decay of π^0 mesons, which are formed in nuclear

interactions between cosmic rays and interstellar matter. High-energy cosmic-ray electrons reveal themselves through interactions with matter and photons. The region around a black hole is predicted to emit characteristic gamma rays. Matter-antimatter annihilation produces another specific type of gamma-ray spectrum.

Gamma-ray astronomical observations also find important applications in studies of the evolution of our solar system, the nature of high-energy processes in the Sun's atmosphere, and their relation to the basic problems of solar activity. A knowledge of the overall elemental composition of a given body can be related to the mechanisms of condensation and accumulation of materials from the primordial solar nebula. Analyses of characteristic x-rays and gamma rays from planetary surfaces can be used to infer global distribution of major and, in some cases, minor and trace elements.

Another attractive feature of gamma-ray astronomy is that the universe is largely transparent to gamma rays. They can reach the solar system from the Galactic center, distant parts of the universe, and dense regions near the centers of active-galaxies regions, which cannot be viewed in the optical or low-energy x-ray region. However, in contrast to optical photons, which penetrate easily through the Earth's atmosphere, only the total amount of matter, and not its form, is relevant for gamma-ray interactions. Thus, for example, the Earth's atmosphere has too much total matter for gamma rays to reach the Earth's surface. (It is primarily the dust of interstellar space which absorbs the optical photons in the Galactic disk.) Further, cosmic-ray interactions in the upper atmosphere produce gamma rays copiously, creating an undesirable background and adding to the need to place gamma-ray detectors above the Earth's atmosphere. It is also valuable to have the gamma-ray telescope in a region of space where the particle radiation is relatively low. Hence, a satellite orbit below the charged particles trapped in the geomagnetic field of the Earth is desired.

Once these difficulties are removed, the Universe is found to be marvelously clear when viewed in the gamma-ray frequency range. A more specific illustration of the penetrating power of this radiation is the following: a high-energy gamma ray passing through the diameter of the central plane of the Galactic disk has about a 1 percent chance of interacting, for a typical path. By contrast, an optical photon can only penetrate about one-tenth the distance from the Galactic center to the Earth in the central plane of the disk. A

high-energy gamma ray also can typically travel to the Earth from the outer part of the universe with less than a 1-percent chance of interacting; in this case, there is a red shift. This remarkable window begins at a few times 10^7 eV, below which it begins to close slowly as the energy decreases, so that, as the x-ray region is reached, the distant parts of the Galaxy are quite opaque. It extends to 10^{15} eV, at which point there begins a one-to-two-decade region in energy, wherein gamma-ray interactions with the blackbody radiation are important.

The scientific potential of gamma-ray astronomy and its direct relationship to the highest-energy processes occurring in astrophysics were recognized by theoreticians well before fruitful experiments became possible. As early as 1952, Hayakawa noted the effect of meson-producing nuclear interactions between cosmic rays and interstellar gas. In the same year Hutchinson (1952) discussed the production of bremsstrahlung radiation by cosmic rays. Even earlier, Feenberg and Primakoff (1948) examined the astrophysical significance of the Compton effect in regard to cosmic-ray electrons. Morrison (1958) wrote a very extensive article describing many of the gamma-ray production processes that are still thought to be important in astrophysics, and several papers discussed the importance of gamma-ray lines in studying rapid nucleosynthesis in supernovae. As the first results on Galactic gamma rays became available in the 1970s, there was a great expansion of the number of theoretical papers on the interrelationship between Galactic structure, cosmic-ray origin, cosmic-ray pressure in the Galaxy, and gamma rays.

As has been noted, the experimental development of gamma-ray astronomy has been slow. Several experimental attempts to detect gamma rays with relatively simple instruments flown on balloons in the years before 1960 were unsuccessful. The failure to obtain positive results led experimenters to develop the larger, more sophisticated detector systems of the 1960s and fly them on the larger, high-altitude balloons that were being developed during this period. The increase in size was dictated by the realization that the intensities of gamma rays were quite low, and the increased complexity was demanded both by the need to detect the rare gamma rays in a high background, primarily of charged cosmic-ray particles, and by the inherent nature of gamma-ray interactions. The first certain detection of high-energy, celestial gamma rays with a satellite experiment was made by Kraushaar et al. (1972). They

observed gamma rays with energies above 50 MeV from the Galactic disk, with a peak intensity toward the Galactic center, using an instrument on the third Orbiting Solar Observatory (OSO-3). However, the limited spectral and spatial resolution of this pioneering experiment left many unanswered questions. About the same time, a few confirmed positive results were obtained with high-altitude balloon experiments, including detection of the Crab pulsar (Browning et al., 1971) and the diffuse radiation in the central region of the Galaxy (Kniffen and Fichtel, 1970).

On November 15, 1972, a gamma-ray telescope with approximately twelve times the sensitivity of the third Orbiting Solar Observatory (OS0-3) telescope and an angular resolution of a few degrees, was launched on the second Small Astronomy Satellite (SAS-2). This instrument provided results that led to a much better picture of the gamma-ray sky, including fair detail on the Galactic plane (e.g., Fichtel et al., 1975; Bignami et al., 1979). Another gamma-ray instrument with approximately equal sensitivity and angular resolution, Cosmic-Ray Satellite (COS-B), was launched on August 8, 1975, and provided information which further expanded our knowledge (e.g., Mayer-Hasselwander et al., 1980). With the results from these satellites, gamma-ray astrophysics proceeded from the discovery phase to the exploratory phase. These data showed the rich character of the Galactic-plane diffuse emission with its potential for the study of the forces of change in the Galaxy, the study of the origin and expansion of the cosmic-ray gas, and the study of the Galactic structure. When examined in detail, the longitudinal and latitudinal distribution appear generally correlated with Galactic structural features, including spiral arm segments. With the observations of discrete sources, some of which were associated with pulsars, point-source, gamma-ray astronomy also began.

Before proceeding to the recent major advances of the Compton Gamma-Ray Observatory, consider briefly the early results on gamma-ray lines. Although nuclear gamma-ray lines whose origin lies outside our solar system have great astrophysical significance, their detection is especially difficult, because the predominant interaction process in this energy range is the Compton process and because there is a high level of locally produced gamma rays and neutrons. Solar nuclear gamma-ray lines were detected as early as 1972, however, by Chupp et al. (1973), from the large solar flares of August 4, and 7, 1972, with an instrument flown on OS0-7.

Also, the detection of a gamma-ray line from the Galactic center was reported (Johnson and Haymes, 1973; Leventhal, MacCallum, and Strang, 1978).

Successful measurements of x-ray, gamma-ray, and α-particle emissions from the Moon, Mars, and Venus have been carried out during both U.S. and Soviet spaceflight missions. The most extensive orbital and *in situ* measurements have been conducted at the Moon. Rangers 3, 4, and 5 in 1961 and 1962 carried the first gamma-ray spectrometer into space. Results were obtained on the spacecraft and cosmic background (Metzger et al., 1964). The first measurement of a gamma-ray spectrum from the Moon was obtained by the Soviet Luna orbiter in 1966, which recognized a compositional distinction between mare and terra provinces (Vinogradov et al., 1967). The first *in situ* measurement of composition was the alpha-backscatter experiment carried by Surveyor 5 in 1967 (Turkevich et al., 1969). During the Apollo 15 and Apollo 16 flights in 1971 and 1972 gamma-ray, x-ray, and α-particle spectometers were carried on the orbiting Command and Service Modules. Approximately 20 percent of the lunar surface was mapped for magnesium, aluminum, thorium, potassium, uranium, silicon, titanium, and oxygen (Metzger et al., 1973; Adler et al., 1972; and Bjorkholm et al., 1973). The two Soviet Lunakhod roving vehicles carried x-ray fluorescence experiments (Kocharov and Viktorev, 1974). A gamma-ray spectrometer aboard the Soviet Mars-5 orbiter obtained several hours of data, from an altitude of about 2300 km, indicating that the average potassium, uranium, and thorium content of the regions surveyed corresponds to that of terrestial rocks like oceanic basalts (Surkov et al., 1976). Measurements of gamma-ray emission from Mars were carried out during the Russian Phobes 2 mission (Surkov et al., 1993; Trombka et al., 1992). The data were collected with a gamma-ray spectrometer when the spacecraft was within 1500 km of the Martian surface. The two NASA Viking Landers carried an x-ray spectometer for the fluorescence analysis of selectable soil and rock samples (Clark et al., 1977). The finely divided samples of the two landing sites were remarkably similar, with high iron and sulfur content and little potassium. Three Soviet spacecrafts, Venera 8, 9, and 10 carried gamma-ray spectrometers to measure potassium, uranium, and thorium at the surface of Venus by detecting the decay of these naturally radioactive elements (Surkov, 1977). In addition,

Venera 10 had a second gamma-ray instrument on board to measure the bulk density of the surface materials.

Also, in the low-energy gamma-ray region of the spectrum, there has been the discovery of the low-energy gamma-ray bursts (Klebesadel, Strang, and Olsen, 1973). These bursts are short, typically observable for a fraction of a second to somewhat over 100 seconds, and have energies concentrated in the low-energy gamma-ray region. The origin and nature of these gamma-ray burst sources remain an unsolved puzzle, although considerably more is now known about them.

The knowledge of gamma-ray astronomy expanded dramatically in the years following the launch of the Compton Gamma-Ray Observatory in April, 1991. This satellite carried into space a set of four experiments each with an order of magnitude more sensitivity than space instruments in their range had had previously, The four experiments were called the Oscillating Scintillation Spectrometer (OSSE), the Compton Telescope (COMPTEL), the Energetic Gamma Ray Experiment Telescope (EGRET), and the Burst and Transient Source Experiment (BATSE). OSSE concentrated on the low-energy end of the spectrum, COMPTEL on the intermediate energy region, and EGRET on the high-energy range of the gamma-ray energy range. Together, they covered the energy region from below 0.1 MeV to about 3×10^4 MeV. BATSE was aimed at the detection of low-energy bursts and transients.

The results of these experiments have not only been very impressive, but have affected a very broad range of astrophysical subjects. They constitute a major share of the experimental information on gamma-ray astronomy in this book and have provided the basis for much of the more current theoretical thinking in the many areas that they effect. Some of the highlights include:

1. the finding of a new class of objects — high-energy, gamma-ray-emitting blazars, whose basic high-energy properties have now been measured,
2. a very clear separation of the gamma-ray properties of blazars and seyferts,
3. the increase in our knowledge of gamma-ray bursts,
4. the observation of an increased fraction of the pulsar electromagnetic radiation being emitted as gamma rays as the

age of the pulsar increases up to a million years, and the detailed knowledge of their spectra,

5. the determination with high certainty that cosmic rays are Galactic,
6. the detailed mapping of the Galactic diffuse radiation, including the aluminum line and the measurement of the pion bump in the high-energy gamma-ray spectrum,
7. the detection of gamma-ray lines from SN1987A and Cas A,
8. the absence of microsecond bursts and its implication for certain unification theories,
9. the existence of energetic particles near the sun for over 10 hours following a flare and the associated implication for the shock acceleration theory, and
10. the measurement of the spectrum of the diffuse, presumably extragalactic, gamma radiation with a flat spectrum in the high-energy region consistent with a blazar origin.

In the very high-energy TeV region of the gamma-ray spectrum, ground-based, Cerenkov light-reflector telescopes have evidence of unpulsed gamma-ray emission from the direction of two pulsars, B0531+21 (Helmken, Grindlay, and Weekes, 1975; and Gupta et al., 1978), and B1706-44 (Kifune et al., 1995), and two BL Lac objects, Mrk 421 (Punch et al., 1992) and Mrk 501 (Weekes et al., 1996).

In the following chapters of this book, the role that gamma-ray astronomy plays in astrophysics will be studied, starting with the closest objects, those of the solar system, expanding to the Galaxy and the stellar objects in it, and onward to other galaxies, cosmology, and the diffuse gamma radiation. The last chapters of the book will discuss the approaches to the detection of astrophysical gamma rays. Production and absorption processes associated with gamma rays will be developed individually as the need arises.

In general, the gamma-ray region will be defined as the energy realm above 0.1 MeV, although the division between x-rays and gamma rays is not sharp. There are some phenomena such as gamma rays from π^0 mesons produced in high-energy interactions, nuclear gamma-ray lines, and matter-antimatter annihilation gamma rays, which belong solely to the gamma-ray realm; however, other phenomena, such as bremsstrahlung from electrons interacting with matter, span the two regions. Results from other wavelength ranges

will be discussed as appropriate for the understanding of the astrophysical problems under consideration.

References

Adler, I., Trombka, J., Gerard, J., Lowman, P., Schmadebeck, R., Yin, L., Blodget, H., Eller, E., Lamothe, R., Gorenstein, P., and Bjorkholm, P., 1972, Science, 175, 436.

Bhat, P. N., Gupta, S. K., Ramanamurthy, P. V., Sreekantan, B. V., Tonwar, S. C., and Viswanath, P. R., 1980, A&A, 81, L3.

Bignami, G. F., Fichtel, C. E., Hartman, R. C., and Thompson, D. J., 1979, ApJ, 232, 649.

Bjorkhom, P., Golub, L., and Gorenstein, P., 1973, Proc. Fourth Lunar Planet. Sci. Conf., 3, 2803.

Browning, R., Ramsden, D., and Wright, P. J., 1971, Nature, 232, 99.

Chupp, E. L., Forrest, D. J., Higbie, P. R., Suri, A. N., Tsai, C., and Dunphy, P. P., 1973, Nature, 241, 333.

Clark, B. C., III, Baird, A. K., Rose, H. J., Jr, Toulmin, P., III, Christian, R. P., Kelliher, W. C., Castro, A.J., Rowe, C. D., Keil, K., and Huss, G. R., 1977, J. Geophys. Res., 82, 4577.

Colgate, S. A., and White, R. H., 1966, ApJ, 143, 626.

Feenberg, E., and Primakoff, H., 1948, Phys. Rev., 73, 449.

Fichtel, C. E., Hartman, R. C., Kniffen, D. A., Thompson, D. J ., Bignami, G. F., Ogelman, H. B., Ozel, M. E., and Turner, T., 1975, ApJ, 198, 163.

Grindlay, J. E., Helmken, H. F., Brown, H. R., Davis, J., and Allen, L. R., 1975, ApJ (Letters), 197, L9.

Gupta, S. K., Ramanamurthy, P. V., Sreekantan, B. V., and Tonwar, S. C., 1978, ApJ, 221, 268.

Hayakawa, S., 1952, Prog. Theor. Phys., 8, 571.

Helmken, H. F., Grindlay, J. E., and Weekes, T. C., 1975, Proc. of the 14th Int. Cosmic Ray Conf., 1, 123.

Hutchinson, G. W., 1952, Phil. Mag., 43, 847.

Johnson, W. N., III, and Haymes, R. C., 1973, ApJ, 182, L85.

Kifune, T., et al., 1995, ApJ, 438, L91.

Klebesadel, R. W., Strong, I. B., Olsen, R. A., 1973, ApJ 182, L85

Kniffen, D. A., and Fichtel, C. E., 1970, ApJ, 161, L157.

Kocharov, G., and Viktorev, S. V., 1974, Doklady Academy of Sciences, USSR, 214, 71.

Kraushaar, W. L., Clark, G. W., Garmire, G. P., Borken, R., Higbie, P., Leong, C., and Thorsos, T., 1972, ApJ, 177, 341.

Leventhal, M., MacCallum, C. J., and Strang, P. D., 1978, ApJ (Letters), 225, L11.

Mayer-Hasselwander, H. A., Bennett, K., Bignami, G. F., Buccheri, R., D'Amico, N., Hermsen, W., Kanbach, G., Lebran, F., Lichti, G. G., Masnou, J. L., Paul, J. A., Pinkau, K., Scarsi, L., Swanenberg, B. N., and

Wills, R. D., 1980, Annals of the New York Academy of Sciences, Proc. of the Ninth Texas Sym., 336, 211.

Metzger, A. E., Anderson, E. C., Van Dilla, M. A., Arnold, J. R., 1964, Nature, 204, 766.

Metzger, A. E., Trombka, J. I., Peterson, L. E., Reedy, R. C., and Arnold, J. R., 1973, Science, 179, 800.

Morrison, P., 1958, Nuovo Cimento, 7, 858.

Punch, M., et al., 1992, Nature, 358, 477.

Surkov, Y. A., 1977, Proc. Eighth Lunar Planet. Sci. Conf., 265.

Surkov, Y. A., Moskalyova, L. P., Kirnozov, F. F., Khyarkukova, V. P., Manvelyan, O. S., and Shchelov, O. P., 1976, Space Research XVI (Berlin:Akademie-Verschlag, COSPAR, 1976).

Surkov, Y. A., Moskaleva, L. P., Kharyukov, V. P., Manvelyan, O. S., and Golovin, A., 1993, in "Remote Geochemical Analysis: Elemental and Mineralogical Composition", ed. C. Pieters and P. Englert, Cambridge University Press, 413-425.

Trombka, J. I., et al., 1992, Proc, Lunar Planet, Science, 22, 23-29.

Turkevich, A. L., Franzgrote, E. J., and Patterson, J. H., 1969, Science, 165, 277.

Vinogradov, A. P., Surkov, Y. A., Chernov, G. M., Kirnozov, F. F., and Nazarkina, G. B., 1967, Cosmic Res., 5, 741.

Vladimirsky, B. M., Neshpor, Yu. I., Stepanian, A. A., and Fomin, N. P., 1975, Proc. of the 14th Int. Cosmic Ray Conf., 1, 118.

Weekes, T. C., et al., 1996, submitted to A&A Suppl.

CHAPTER 2

GAMMA-RAY OBSERVATIONS OF THE SOLAR SYSTEM

2.1 Introduction

The question of the origin of the solar system has attracted much attention. Such interest is evident throughout the history of mankind. Many models have evolved, but when the basic assumptions underlying such models, no matter how mathematically elegant, cannot be related to actual observed phenomena, there is no method to discriminate between truth or fiction. In the last three decades, however, observational data have been obtained from both space flight programs and meteorite studies that allow certain constraints to be imposed on the theoretical models for the origin and evolution of the solar system. Further, various theoretical approaches now can be evaluated in terms of their observational tests, and more rigorous models can be developed.

The present structure of the solar system should be understood as the product of a long series of complicated processes. A major aim of the solar system exploration program and the subsequent attempt to develop theoretical models is to reconstruct the early history of the solar system based on present-day observations. In order to comprehend the relationship between gamma-ray astronomical measurements and the development of theoretical models, one must understand, at least in broad outline, the general hypothesis relating to solar system evolution.

2.2 Models of Solar System Evolution

Regarding the evolution of the solar system, the approaches fall into two general categories: the dualistic view, or the planetesimal approach; and the monistic view, or the nebular hypothesis. The dualistic approach has led to the development of a number of

models. One model suggests that the Sun passed very near another star at some time and drew out material from it. These filaments of material ultimately condensed into separate planets. Another dualistic model hypothesizes that the Sun was originally a member of a binary system. The other star, having suffered a direct collision with a third star, broke up, and the planets were formed from the resulting fragments.

A number of objections can be raised with respect to both dualistic views just considered. In the first case, when considering materials being drawn from the Sun by the attraction of a passing star, there is a failure to account for the peculiar distribution of angular momentum in the solar system, which is mainly carried by the outer planets of far less mass than the Sun. Considering the binary system hypothesis, rather high temperatures would be expected, causing the material to stream out into space rather than to condense into planets. The dualistic view, in one form or another, seemed the most plausible explanation for the genesis of the solar system, for, until recently, it was believed that the Sun and planets could hardly represent the remains of independent condensing in a primeval nebula, their masses being too unequal to be suspect of any such explanations.

Recent evidence of the following has revived interest in the monistic view, the nebular hypothesis: the common age of solar and planetary evolution derived from the distribution of materials, and the comparison of that age to stellar evolution derived from astrophysical observation.

Using this hypothesis as a basis for our understanding, the following is a brief outline of a proposed model of the formation and evolution of the solar system. For more detailed expositions see, for example, Levy and Lunine, 1993; Cameron, 1988; Lissauer, 1993. A two-phase process is considered: the formation of a protoplanetary disk, followed by the formation and evolution of our planetary system.

The following outline describes a possible sequence in the formation of the protoplanetary disk from which our planetary system emerged. Observations of molecular cloud complexes have indicated the presence of large numbers of distinct regions with high-density molecular cores that seem to be on the verge of gravitational collapse. Thus, the existence of young stars

embedded deep within these molecular cloud cores may be evidence for the formation of stellar and planetary systems from a gravitational contraction of dense cloud cores. It is believed that these cloud cores are initially supported by magnetic fields, which may slowly lose this support through the motion of these field lines relative to the neutral bulk of the cloud. This loss of support could lead to the gradual contraction of the cloud cores and possibly to a rapid inward collapse during a period on the order of a million years. However, with respect to our solar system, there is evidence from the presence of extinct radioactivities in the early solar system that this collapse was accelerated by shock waves ejected from a near-by dying star. The dense cloud then begins to collapse under its own self-gravity. The cloud flattens and spins up, and its angular momentum is conserved. The lower angular-momentum matter moves toward the center of the cloud and forms the protosun. The protosun will then emit a high-velocity, bi-polar wind along its rotational axis. The higher angular-momentum gas and dust will fall into the disk to form the solar nebula. Once the infall ceases, the bulk of the nebular gas is accreted onto the protosun, thus clearing out the cloud envelope.

Planetary formation then begins as dust grains begin to grow in the solar nebula. The growth is initiated when interstellar dust grains of ~0.1 micron size are processed in the solar nebular accretion shock. These particles grow progressively larger and larger. Brownian motion or gas turbulence will produce collisions between dust grains. The gas grains will grow and gravity will cause them to migrate towards the nebula's mid-plane, and gas drag will cause them to spiral inward. Different size particles will move at different speeds, and chances for mutual collisions will be enhanced. The collisions must be rather gentle so that the objects will accumulate rather than fragment. Large objects are less affected by drag and will suffer little inward drift. Furthermore, these larger objects will be able to grow by accumulating the smaller bodies that are drifting toward them. Chemical, collisional, and thermal processes will modify the structure of the grains. Once the object sizes reach about 10 km, they are considered to be "planetesimal." The growth in this phase is mainly caused by accumulation forces which are mainly nongravitational. An era of so-called runaway growth is now believed to occur. There is now a period of rapid growth because the growth of these larger bodies will be enhanced by greater gravitational attraction. Also, their lower relative velocity will

enhance their ability to deflect smaller colliding bodies. This period of rapid growth is believed to occur over 10^5 years. When these planetesimals grow to about the size of the Moon or Mercury, they are considered "planetary embryos." These objects are then believed to be in nearly circular, coplanar orbits. Gas removal from the inner planets (terrestrial planets) occurs on a time scale of about 10^6 to 10^7 years. Mutual gravitational perturbations cause these initially circular orbits to become more and more eccentric. Asteroids are believed to be relics of similar processes in the present asteroidal region that failed to complete the runaway growth phase as a consequence of either gravitational or collisional removal of the other bodies in this region.

In the preceding, a possible scenario for the formation of the inner planets and asteroid belt was considered. A similar scenario will now be outlined for the formation of the outer planets and other solar system objects such as the comets. Again, the process begins as the dust grains in this region grow into about 10 km diameter "plantesimals." A runaway growth of these planetesimals takes place to form very large cores of the four outer planets on a time scale of 10^6 to 10^7 years for Jupiter and Saturn and 10^8 for Uranus and Neptune. There is then believed to have been gravitational capture of the nebula gas by Jupiter and Saturn on a 10^7 year time scale. Comets, as well as Pluto- and Triton-like objects may be relics of bodies in the outer solar system that failed to be incorporated into planets during the runaway phase. Most of these bodies may have been ejected into interstellar space by the outer planets. Residual comets may thus have been stored in the Oort cloud and in the Kuiper Belt.

Further detailed discussions of the various theories concerning solar system evolution can be found in such references as Adler and Trombka (1970), Sagan (1975), Cameron (1975), Alfven and Arrhenius (1976), Greeley and Carr (1976), and Pieters and Englert (1993).

2.3 Solar System Gamma-Ray Astronomical Observations

Considering the proposed models of the dynamic evolution of the solar system, how can gamma-ray astrophysical observations contribute to our understanding of solar system evolutions? Gamma-ray spectroscopic observations of planetary bodies will allow for the global determination of elemental surface

composition. The problem of the origin and evolution of the solar system can thus be considered in a geochemical context. The problems have been examined in three works, Adams et al. (1967, 1979) and Pieters and Englert (1993).

In order to determine the stages in the evolution of the solar system when chemical fractionation took place, some questions must be examined. Are the individual terrestrial planets and satellites chemically uniform or nonuniform? Did the final accretion result in the present array of planets and satellites or in an array that was subsequently altered? Was the cloud chemically homogeneous at the time of final accretion? What was the state of the Sun-cloud system when it first became an identifiable unit? Were there large-scale elemental and isotropic nonuniformities in the contracted nebula? Answers depend ultimately on chemical measurements as well as on mineralogy and petrology. Assumptions cannot be made a priori that a given planet is chemically homogeneous and that an analysis at a given site is representative of the whole planet. Thus, the remote orbital gamma-ray and x-ray remote sensing system will address these questions.

The only accessible part of a planet—its near-surface—may be totally nonrepresentative of the planet as a whole. It is necessary to have a general understanding of the planet, and this can be achieved effectively only by studying the processes that may have caused—and may yet be causing—redistribution of materials within the planet. Only if these processes are understood can valid general conclusions be drawn regarding the significance of specific measurements. The sampling or type of measurements made on a planet, therefore, should be guided by the need to understand geologic processes.

In the above discussion, surface properties are considered. Of equal importance in understanding solar system evolution is the relationship of the surface to the planets, internal heterogeneities and conditions. Chemical determinations of materials from the interiors obviously cannot be made directly; nevertheless, inferences are possible from measurements of physical properties. The value of any particular physical parameter is not important, since it depends on local conditions; its importance lies in the fact that it places limits on chemical composition and internal conditions, and provides data on internal processes. Stratigraphy is concerned directly with documentation of crystal heterogeneity.

Even though the surface rocks constitute a minute part of the total mass of the planet, it is the part on which a great many measurements are made. Thus, an understanding of surface materials is crucial, and how variegated the surface rocks are and how they are formed must be known. The mode of formation is particularly important, for different processes vary in the extent to which they cause chemical change. Consequently, to understand the broader implications of surface analyses, we should know the distribution and mode of formation of the rocks analyzed, as well as the degree to which they typify all other materials of the planet.

Of course, the central body of interest of the solar system is the Sun. There can be no theory for the origin and evolution of the solar system without an understanding of the nature of the Sun, its chronology, and present-day dynamics. Furthermore, the Sun is of great astrophysical interest, because it is the star that the astronomer can study in greatest detail. The study of solar transient phenomena yields significant information on solar dynamics. Therefore, gamma-ray astronomy, as applied to the observation of the Sun, has the specific objective of considering high-energy processes that take place in the outer region of the Sun's atmosphere and the relationship of these phenomena to the basic problems of solar activity. A measurement of the spectra of discrete gamma-ray line emission will reveal the detailed dynamics and time structure of solar flares and energetic particle accretion and release. Such measurements should also yield qualitative, and in some cases, quantitative, information on the composition of specific ambient or transient nuclides in the outer regions of the Sun's atmosphere.

In the next two chapters, remote gamma-ray geochemical sensing systems for planetary surface observations and solar gamma-ray observations during flares will be discussed in detail.

References

Adams, J. B., Conel, J. E., Dunne, J. A., Fanale, F., Holstrom, G. B., and Loomis, A. A., 1967, "Scientific Strategy, Objectives and Experiments for Mariner Missions from 1970 through 1973," Calif. Inst. Tech., Jet Prop. Lab. Doc. 900-94 (internal document).

Adler, I., and Trombka, J. I., 1970, Physics and Chemistry in Space, Vol. 3, ed. J. G. Roderer and J. Zahringer, (New York: Springer-Verlag, 1970).

Alfven H., and Arrhenius, G., 1976, Evolution of the Solar System, NASA SP-345 (Washington, D.C.: U.S. Government Printing Office, 1976).

Cameron A. G. W., 1975, Scientific American, 233, 33.

Cameron, A. G. , 1988, "Origin of the Solar System," Annual Reviews of Astronomy and Astrophysics , 26:441-472, Annual Reviews Inc., Palo Alto, Calif., .

Greeley, R., and Carr, M. H., (eds.), 1976, A Geological Basis for the Exploration of the Planets, NASA SP-417 (Springfield, VA.: National Technical Services, 1976).

Levy, E. H., and Lunnine J. I., (eds.), 1993, Protostars and Protoplanets III, University of Arizona Press, Tucson Ariz.

Lissauer, J. J., 1993, "Planet Formation," Annual Reviews of Astronomy and Astrophysics 31:129-174, Reviews Inc., Palo Alto, Calif.

Pieters, C. M., and Englert, P. A. J., (eds.), 1993, "Remote Geochemical Analysis: Elemental and Minerological Composition", (New York, Cambridge University Press, 1993).

Sagan, C., 1975, Scientific American, 233, 23.

CHAPTER 3

PLANETS, COMETS, AND ASTEROIDS

3.1 Introduction

An important part of the investigation of planetary evolution is the determination of the chemical composition of the surfaces of planets, comets, and asteroids. The abundance of certain elements with different condensation temperatures and with various types of geochemical behavior can provide important data relative to the origin and evolution of these bodies. The global planetary composition is characteristic of the nature of its formation during the accretion from the solar nebula. Furthermore, early condensation processes are indicated by certain key elements, e.g., uranium (a refractory element), iron (condensed as metallic iron-nickel), and magnesium (the first silicate formed). In the course of planetary evolution, the elemental distribution is greatly modified by such processes as core and crust formation during differentiation and later magma formation and emplacement. Evolutionary processes can be studied, for example, by determining the K/U ratio (remelting of the primordial condensates) and the Tl/U ratio (relative abundance of volatiles). A detailed discussion of the use of geochemical information in the study of the formation and evolution of planetary bodies can be found, for example, in Anders (1977) and Pieters and Englert (1993).

In this chapter, methods for remote sensing of elemental composition of surfaces will be considered. A number of processes that modified the elemental composition must be understood in order to interpret the compositional maps determined from the remote sensing observations. The surfaces of most planets usually involve a regolith consisting of most fragments and reworked material such as glasses. The extent to which regolith material has been transported away from its origin influences the geochemical distribution. Transport mechanisms possible on various planets

include volcanism, meteoroid impact, water and wind movement and particle transport caused by electrostatic charge processes. Remote geochemical information obtained by x-ray and gamma-ray spectroscopic measurements combined with photogeological maps can be used to infer information about both horizontal and vertical elemental composition relative to the problems just considered.

Both x-ray and gamma-ray remote sensing techniques will be considered. It is the combination of these techniques which is required for a more complete understanding of the remote sensing method. Combinations of x-ray and gamma-ray spectrometers or a gamma-ray spectrometer alone have been flown successfully on a number of planetary spaceflight missions. In U.S. spaceflight programs such remote sensing orbital spectrometers were included aboard Apollo 15 and Apollo 16, the Mars Observer, the Near Earth Asteroid Rendezvous (NEAR), and the Lunar Prospector (Adler et al., 1972; Metzger et al., 1973; Boynton et al., 1993; Trombka et al., 1996; Feldman, 1996), and in the Soviet space program on a number of Luna missions, Mars 4 and Mars 5, the Phobos 2 mission, and the Mars '96 mission (Vinogradov et al., 1966; Surkov et al., 1976, 1993; Surkov 1996; and Mitrofonav, 1996). A number of lander missions have included gamma-ray spectrometers, e.g., the Russian Venera missions Penetrator (Surkov et al., 1976).

3.2 Interaction Processes

Measurements of discrete line x-ray and gamma-ray emission from condensed bodies in space can be used to obtain both qualitative and quantitative elemental composition information. Discrete line emission in this energy domain (~0.9 keV to ~10 MeV) can be attributed to a number of processes, such as natural radioactivity, solar x-ray fluorescence, and cosmic-ray primary and secondary induced activity. Figure 3.1 summarizes the major processes for x-ray and gamma-ray emissions from the Moon. The elements that are listed produce significant x-ray and gamma-ray fluxes because of their abundance, interaction cross sections, and magnitude of exciting flux. Also indicated in the figure are the depths at which the interactions occur and the depth from which the discrete line x-ray and gamma-ray photons can escape through the planetary surface and be detected from orbit. Both alpha-particle emission and neutron albedo can be used to determine elemental composition. Each of the major emission processes are considered in the next section.

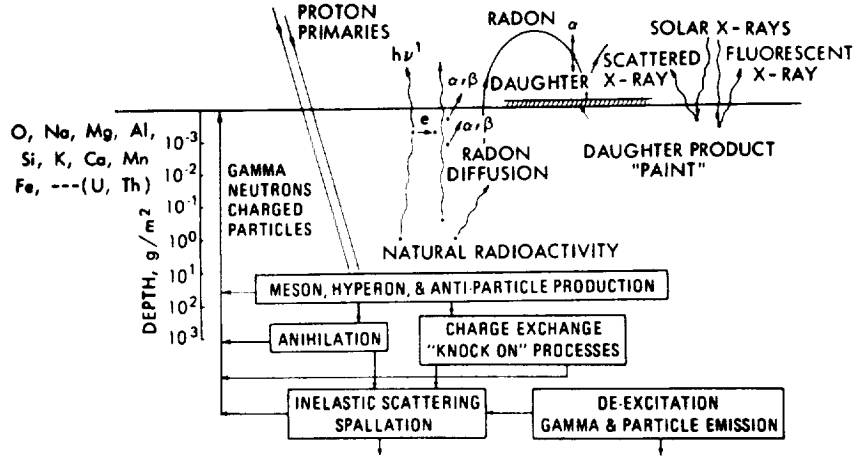

Figure 3.1. Radiation evironment at the lunar surface.

3.3 Solar X-Ray Fluorescence

The major radiation source for remote orbital x-ray elemental analysis is that produced by characteristic x-rays following the interaction of solar x-rays with the surface of the given body being studied. Figure 3.2a shows a so-called "quiescent" solar x-ray emission spectrum characteristic of that observed during Apollo 15. The spectrum shown was produced by a semiempirical method. Data obtained from the solar observatory, "Solrad," were combined with a theoretical calculation based on a model developed by Tucker and Koren (1971). The calculation was predicated on combining a coronal temperature of ~1.5 x 10^8 K with a hot-spot temperature of about ~3 x 10^6 K in proportions determined from Solrad measurements during Apollo 15.

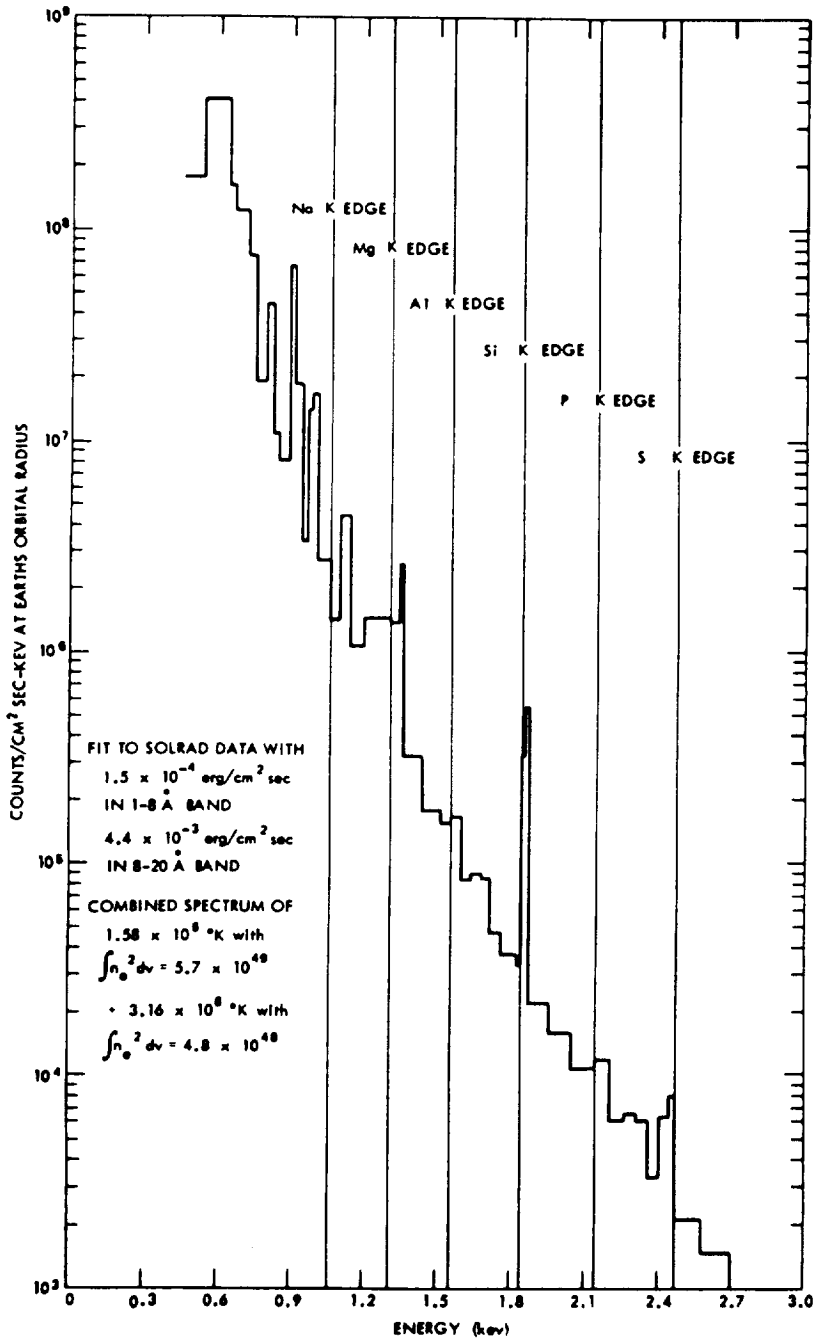

Figure 3.2a. Solar x-ray spectrum at the Earth's orbital radius (IAU).

The expected fluorescence emission can be calculated using the following equation (Jenkins and DeVries, 1967):

$$I_j = \left[w_j g_j \left(\frac{r_j - 1}{r_j} \right) \frac{d\Omega C_j}{4\pi} \right] \times \left\{ \int_{\lambda_{min}}^{\lambda_{max}} \frac{J(\lambda)\mu_i(\lambda)d\lambda}{\sum_{1}^{n} C_i \left[\mu_i(\lambda) + \mu_i(\lambda_j) \left(\frac{\cos\alpha}{\cos\beta} \right) \right]} \right\}, \quad (3\text{-}1)$$

where I_j is the expected fluorescent flux for element j in the irradiated surface in photons per second per fraction solid angle,

$J(\lambda)$ is incident for exciting differential energy flux as a function of wavelength;
$\mu_i(\lambda)$ is the mass absorption coefficient as a function of wavelength for element i;
C_i is the percent of atomic concentration for element i;
ω_j is the fluorescence yield for element j;
r_j is the absorption jump for element j;
g_j is the probability of a particular electron transfer (e.g., K-alpha) in a particular shell (e.g., K-series);
$d\Omega/4\pi$ is the differential fraction of a solid for emission (assumed isotropic in these calculations);
λ_{edge} is the wavelength at the absorption edge;
λ_{min} is the shortest wavelength or highest energy for the incident or exciting flux;
$\mu_i(\lambda_j)$ is the mass absorption coefficient for element *i* for the excited x-ray fluorescence of wavelength λ_j;
α is the complement of the angle of the incident flux;
β is the complement of the so-called take-off angle (angle of emission) of the secondary flux; and
n is the number of elements in the sample.

This equation holds for the case of a thick target, infinite in terms of the absorption of the wavelengths of interest, for example, remote x-ray sensing.

A major source of background in remote x-ray analysis can be attributed to the x-rays emitted from an irradiated surface because of the coherent scattering of the incident flux S (θ,λ). This component can be calculated from the following relationship (Hubbell, 1969):

$$S(\theta,\lambda) = 0.239[\cos^2\theta + 1]J(\lambda)\frac{\sum_{1}^{n}\left[\frac{C_i}{W_i} f_i^2\right] d\Omega}{\sum_{1}^{n}\left[\mu_i(\lambda)\left(1+\frac{\cos\alpha}{\cos\beta}\right)\right]}, \qquad (3\text{-}2)$$

where $S(\theta,\lambda)$ is the differential angular and wavelength scattered flux,

- θ is the backscatter angle (only x-rays scattered out of the surface will be detected); the constant 0.0239 is the product of N(Avogadros number) and $r^2/2$ where r is the electron radius;
- W_i is the atomic weight of element i; and
- f_i is probability for coherent scattering.

The calculated lunar fluxes assuming the solar spectrum shown in Figure 3.1 and using Equations (3-1) and (3-2) were consistent with those observed in orbit during the Apollo 15 and Apollo 16 missions (Adler and Trombka, 1977).

The x-ray emission is strongly dependent on the elemental composition and on the distance from the Sun. It should also be noted that the higher Z elements are not shown. The expected emission flux from these elements would be too low to be detectable from orbit. Table 3.1 shows the calculated elemental x-ray flux (counts/sec) and elemental ratio for a 25 cm^2 proportional counter at 35 km from the center of mass of an asteroid of the size of 433 Eros. Four different meteorite composition types are assumed and three different solar conditions are assumed (Figure 3.2b). The asteroid is assumed to be 1.4 AU from the Sun.

Table 3.1.

The calculated elemental x-ray flux (counts/sec) and elemental ratio for a 25 cm^2 proportional counter at 35 km from the center of mass of 433 Eros.

Ordinary Chondrite Type H

		Flux (c/s) Solar Condition				Elemental Ratio Solar Condition	
Element	% Comp	B1	C1	M1	Elements	B1	C1
Mg	14.2	2.0	9.1	47.0	Mg/Fe	16.7	3
Al	1.0	0.13	0.83	4.6	Mg/Al	15	11
Si	17.0	2,3	18.0	106.0	Fe/Si	1005	0.17
Fe	28.0	0.012	3.0	33.0			
				Stony Iron Palisites			
Mg	11.9	1.4	6.1	30.0	Mg/Fe	82	2
Al	0.2	0.12	0.14	0.7	Mg/Al	70	43
Si	8.0	0.94	7.4	44.0	Fe/Si	0.018	0.57
Fe	55.0	0.019	4.2	48.0			

Table 3.1 (continued).

Achondrite Eucrite

Element	% Comp	Flux (c/s) Solar Condition B1	C1	M1	Elements	Elemental Ratio Solar Condition B1	C1
Mg	15.0	0.7	3.2	16.0	Mg/Fe	104	2
Al	23.0	1.2	6.7	37.0	Mg/Al	0.6	0.5
Si	6.6	3.2	26.0	154.0	Fe/Si	0.002	0.07
Fe	4.3	0.007	1.7	19.0			
			Carbonaceous Chondrite Type 1				
Mg	9.6	1.5	6.7	34.0	Mg/Fe	150	3
Al	0.9	0.14	0.8	4.5	Mg/Al	11	8
Si	10.0	1.6	13.0	76.0	Fe/Si	0.006	0.19
Fe	18.0	0.010	2.5	29.0			

*Note: Four meteorite compositions and three solar flare comditions (Figure 3.2b) are used in the calculations.

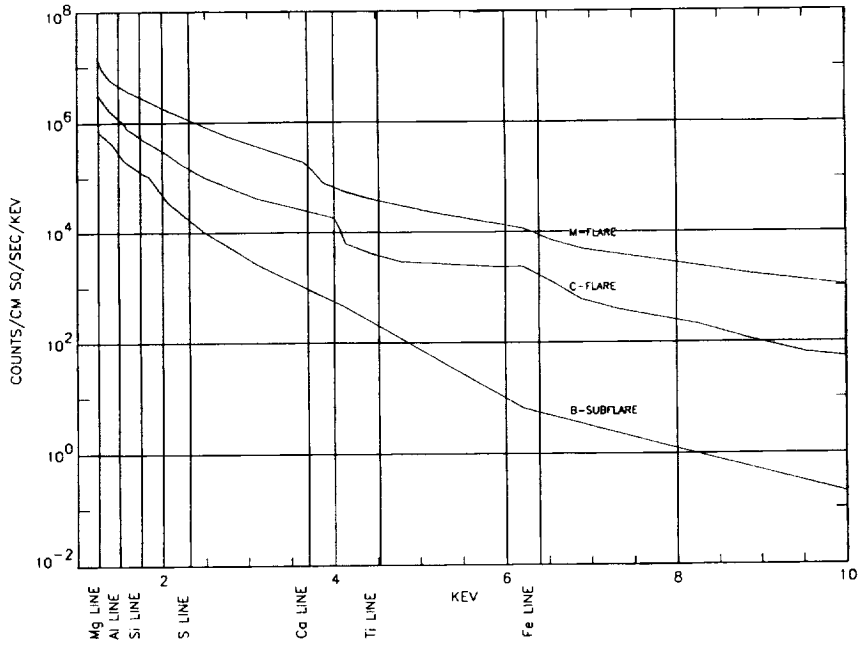

Figure 3.2b. Expected solar flux conditions during solar maximum at 1.4 AU. Three solar flux types are shown: B-subflare; C-flare; and M-flare.

An increase in solar activity and possibly in temperatures tends to harden the solar spectrum. This hardening enhances the intensities of the emissions of the heavier elements relative to the lighter elements. The increase in intensity also increases the absolute magnitude of the fluorescent x-ray flux. Whenever xray orbital measurements are used to obtain elemental composition, simultaneous observations must be made of the incident solar x-ray spectrum and the planetary x-ray emission spectrum in order to infer both qualitative and quantitative information.

Characteristic x-ray emission can be produced by charged particle interactions with condensed matter. These interactions may be of interest for possible remote chemical-composition sensing of the Galilean satellites of Jupiter. X-ray fluorescence from such key elements as magnesium, aluminum, silicon, and iron may be excited by the interactions of charged particles (e.g., the Jovian radiation belt) with the satellite surfaces. Of particular interest will be electron

and proton interactions. The processes of interest will be briefly considered here. The preponderance of useful characteristic x-rays are generated by the inner shell ionization caused by electron impact. The inner shell ionization cross sections from protons can become comparable to electron cross section if the protons have energies about 1000 times that of electrons. For example, both 5 keV electrons and 4 MeV protons have a cross section of ~2 x 10^{-20} cm^2 in the K-shell ionization of an aluminum target (Brandt, 1972; Toburan, 1972).

Finally, the mean tree path of these soft x-rays is very short with respect to the density of materials found on planetary surfaces. Thus, the elemental composition obtained by utilizing the x-ray fluorescent method is characteristic of the composition for only about 100 μ in depth. Furthermore, any significant atmosphere around a planetary body will absorb the solar x-ray flux, thus eliminating any chance of surface fluorescence. Therefore, the technique can only be used for atmosphereless bodies such as our Moon, comets, asteroids, and Mercury.

3.4 Charged-Particle X-Ray Fluorescence

Progressing to the outer planets, it is found that the solar x-ray flux has decreased to such an extent that the emission produced by solar x-ray excitation is negligible. For the outer planets, and specifically for the Galilean satellites of Jupiter, charged-particle x-ray fluorescence, as discussed previously, may produce sufficient fluxes of characteristic x-rays so as to allow for remote analysis.

3.5 Natural Radioactivity

One of the major sources of gamma-ray emission from planetary surfaces is attributable to the decay of the so-called natural radioactive elements. These are the naturally occurring primordial radioactive elements ^{40}K, ^{138}La, C ^{176}Lu, and the uranium and thorium decay. The gamma-ray energies, half-lives, and yields (i.e., gammas per disintegration of the parent radionuclide) for these nuclides are shown in Table 3.2 (Reedy, 1978).

Not all the energies listed in Table 3.2 can be observed at orbital altitudes. The major lines are used to infer the distribution of the nuclear species enumerated. For this case, it can be assumed that

the gamma rays are emitted at a steady and predictable rate. A so-called secular equilibrium is assumed among the parent and daughter products, an equilibrium that can be disturbed by radon emanation which can change the ratio of parent-to-daughter gamma-ray emission. This disequilibrium can be used to measure the chronology and time extent of such emission processes. In order to obtain information on disequilibrium, gamma-ray lines from both the parent and daughter nuclides must be detected.

Table 3.2.

Natural Radioactive Elements of Interest in Remote Gamma-Ray Sensing of Planetary Bodies

Element	Nuclide	Energy	Yield	Half-Life
K	^{40}K	1,4608	0.1048	1.250×10^{9}y
La	^{138}La	1,4359	0.671	1.12×10^{11}y
	^{138}La	0.7887	0.329	1.12×10^{11}y
Lu	^{176}Lu	0.3069	0.94	3.79×10^{10}y
	^{176}Lu	0.2018	0.85	3.79×10^{10}y
Th	^{208}Ti	2.6146	0.360	3.0527m
	^{228}Ac	1.6304	0.019	6.13h
	^{212}Bi	1.6205	0.016	60.60m
	^{228}Ac	1.5879	0.037	6.13h
	^{228}Ac	1.4958	0.010	6.13h
	^{228}Ac	1.4592	0.010	6.13h
	^{228}Ac	0.9689	0.175	6.13h
	^{228}Ac	0.9646	0.054	6.13h
	^{228}Ac	0.9111	0.290	6.13h
	^{208}Ti	0.8605	0.045	3.0527m
	^{228}Ac	0.8402	0.010	6.13h
	^{228}Ac	0.8356	0.018	6.13h
	^{228}Ac	0.7948	0.048	6.13h
	^{212}Bi	0.7854	0.010	60.60m
	^{228}Ac	0.7721	0.016	6.13h

Table 3.2 (continued).

Natural Radioactive Elements of Interest in Remote Gamma-Ray Sensing of Planetary Bodies

Element	Nuclide	Energy	Yield	Half-Life
	^{228}Ac	0.7552	0.011	6.13h
	^{212}Bi	0.7271	0.070	60.60m
	^{208}Ti	0.5831	0.307	3.0527m
	^{228}Ac	0.5623	0.010	6.13h
	^{208}Ti	0.5107	0.083	3.0527m
	^{228}Ac	0.4630	0.046	6.13h
	^{228}Ac	0.4094	0.022	6.13h
	^{228}Ac	0.3384	0.120	6.13h
	^{228}Ac	0.3280	0.034	6.13h
	^{212}Pb	0.3000	0.031	10.643h
	^{208}Ti	0.2774	0.024	3.0527m
	^{228}Ac	0.2703	0.038	6.13h
	^{224}Ra	0.2410	0.038	3.665d
	^{228}Ac	0.2094	0.045	6.13h
	^{214}Bi	2.4477	0.016	19.7m
	^{214}Bi	2.2041	0.050	19.7m
	^{214}Bi	2.1185	0.012	19.7m
	^{214}Bi	1.8474	0.021	19.7m
	^{214}Bi	1.7645	0.159	19.7m
	^{214}Bi	1.7296	0.031	19.7m
	^{214}Bi	1.6613	0.012	19.7m
	^{214}Bi	1.5092	0.022	19.7m
	^{214}Bi	1.4080	0.025	19.7m
	^{214}Bi	1.4015	0.014	19.7m
	^{214}Bi	1.3777	0.040	19.7m
	^{214}Bi	1.2810	0.015	19.7m
	^{214}Bi	1.2381	0.059	19.7m
U	^{214}Bi	1.1552	0.017	19.7m
	^{214}Bi	1.1203	0.150	19.7m
	^{214}Bi	0.9341	0.032	19.7m

Table 3.2 (continued).

Natural Radioactive Elements of Interest in Remote Gamma-Ray Sensing of Planetary Bodies

Element	Nuclide	Energy	Yield	Half-Life
	^{214}Bi	0.8062	0.012	19.7m
	^{214}Pb	0.7859	0.011	26.8m
	^{214}Bi	0.7684	0.049	19.7m
	^{214}Bi	0.6655	0.016	19.7m
	^{214}Bi	0.6093	0.461	19.7m
	^{214}Pb	0.3519	0.371	26.8m
	^{214}Pb	0.2952	0.192	26.8m
	^{214}Pb	0.2419	0.075	26.8m
	^{226}Ra	0.1860	0.055	1599y
	^{235}U	0.1857	0.54	7.038×10^{8}y

Note: y = years, d = days, h = hours, m = minutes, s = seconds.

3.6 Primary and Secondary Galactic Cosmic-Ray Interactions

The bombardment of a body of condensed matter by the primary Galactic cosmic rays produces many interactions and numerous secondary particles. The spectra of primary and secondary cosmic rays are discussed in detail by Reedy and Arnold (1972). The approximately 90 percent proton component of the cosmic-ray flux is of most interest in this section. Numerous secondary particles are produced by the primary cosmic-ray interactions. Most secondary charged particles have low enough energies that they are stopped before they can induce nuclear reactions. However, this is not true for the secondary neutrons produced by ionization energy losses. Thus, these neutrons become the major flux of particles capable of exciting gamma-ray emission.

The majority of secondary neutrons produced in this manner have energies from ~0.5 MeV to ~20 MeV. They can be further slowed down in the planetary surface to produce significant fluxes of

thermal neutrons. Calculations of the secondary neutron flux produced by the incident primary cosmic-ray flux and moderated by the planetary material have been carried out by a number of investigators (Lingenfelter et al., 1961, 1972; Korblum and Fireman, 1974; Lapides et al., 1980; Spergel et al., 1980). The neutron production mechanisms used in the calculation involve such processes as evaporation and knock-on. Detailed descriptions of these mechanisms can be found in Lingenfelter et al. (1961, 1972), and Korblum and Fireman (1974).

Neutrons procuced in solar-system bodies, as described above, can also escape from the body producing a neutron albedo. These neutrons can be observed at orbital altitudes and also can be used to study the composition of the planet's surface (Lingenfelter et al., 1961; Drake et al., 1988; Haines and Metzger, 1984). Escaping neutrons below a certain energy are gravitationally bound to a planet, and some low-energy neutrons can return to the planet's surface if they do not beta decay while in space. This effect is negligible for bodies of the size of the Moon or smaller, but can be important for such bodies as Mars. The effect of gravity on both the equilibrium surface neutron flux and the leakage neutron flux has been calculated for Mars (Feldman et al., 1989). Gamma emissions from Mars and other solar system bodies can be found in Boynton et al. (1993) and Trombka et al. (1996). In the following, we consider only the lunar case.

Results for thermal neutron distributions are shown in Figures 3.3 and 3.4 (Lapides et al., 1980; Lapides, 1981). In Figure 3.3, the spatial distribution for thermal neutrons is shown for the Apollo 11 landing-site-type material and a Carbonaceous Chondrite material (CCN). The curve-marked source indicates the spatial dependence of a given energy group. The magnitude of this curve has been chosen only to be able to compare the spatial dependence of the Apollo 11 and CCN curves. The source term reflects the cosmic-ray-produced secondary neutrons. The depth is measured in terms of gm cm^{-2} in order to generalize the solution. As can be seen, the spatial dependence of the thermal flux near the surface of a planetary body will be quite different from a source term because of leakage from the surface, but at depths greater than about 150 gm cm^{-2} the source, Apollo 11, and the CCN spectrum seem to have the same spatial dependencies. The flux intensity is presented in terms of flux per neutron produced. For this case the magnitude of the

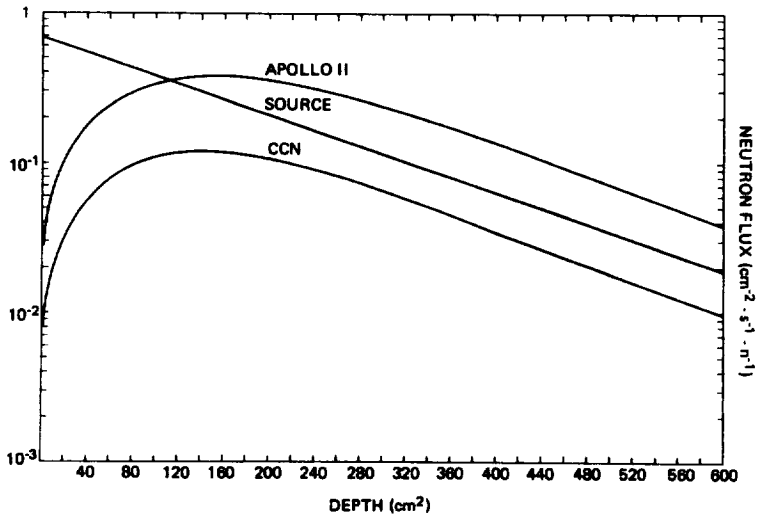

Figure 3.3. Spatial dependence of the thermal neutron flux in a lunar-type material (Apollo 11) and a carbonaceous chondrite material (CNN). The source neutrons are produced by cosmic-ray interactions in the surface materials. The spatial dependence of the source neutrons is also indicated. The shape is shown only as a comparison for the Apollo 11 and CCN (Labides, 1981). The magnitude is arbitrary and depends on the particular energy group.

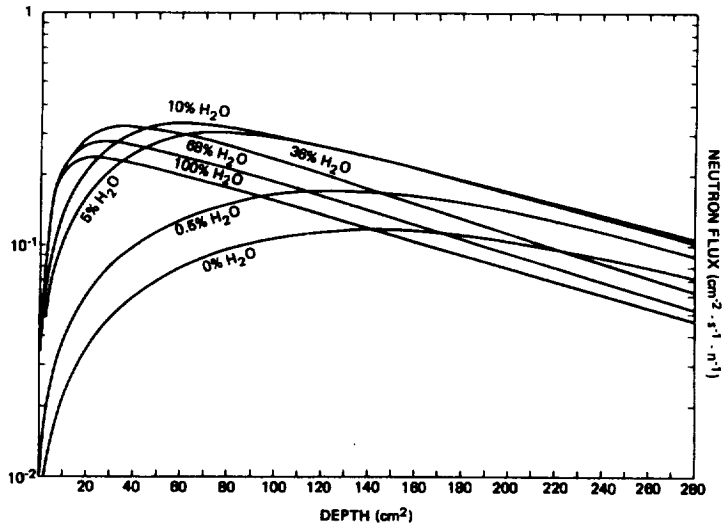

Figure 3.4. Calculated thermal flux in a planetary surface as a function of depth in the surface. The effect of the addition of hydrogen (in the form of water) is shown for a number of different hydrogen concentrations.

thermal flux in the Apollo l l-type material is larger than in the CCN. The macroscopic neutron absorption cross section is larger in the CCN material, thus leading to the depression in thermal neutron flux.

Figure 3.4 shows the spatial dependence of the thermal flux as a function of hydrogeneous content. The thermal flux reaches a maximum closer to the surface of a planet for those cases in which the hydrogen content increases. This is caused by the increase in thermalization due to the presence of hydrogen. Hydrogen is also an excellent absorber of thermal neutron. Thus, although the thermalization increases, there is a point where the absolute magnitude of the flux decreases with increasing hydrogen, as can be seen in Figure 3.4. The spatial shape and the flux magnitude of the thermal neutrons affect the shape and magnitude of the gamma-ray emission from the surface. A knowledge of the exciting flux will be required to obtain geochemical information from a determination of the gamma-ray emission. Methods for performing such analyses can be found, for example, in Reedy and Arnold (1972), Reedy et al. (1973), Lapides et al. (1980), and Lapides (1981). Experimental neutron flux determinations have been obtained during the Apollo 17 mission by Woolum et al. (1975). These results are consistent with those shown in Figure 3.3.

Neutrons with energies greater than or equal to 1 MeV can produce gamma-ray exciting levels of nuclei through inelastic scattering reactions of the (n, xγ) type, where x is usually a neutron but can be another particle. For the case where x is a neutron, the reaction can be described as follows:

$$ {}^{A}_{Z}Y + {}^{1}_{0}n \rightarrow {}^{A}_{Z}{}^{*}Y + {}^{1}_{0}n', \tag{3-3a} $$

then

$$ {}^{A}_{Z}{}^{*}Y \rightarrow {}^{A}_{Z}Y + \gamma - \text{ray}, \tag{3-3b} $$

where

Y is a given nuclide;
Z is the atomic number;
A is the atomic weight;

${}_0^1n$ is a neutron;

Y is the excited state of the nuclide; and

${}_0^1n'$ stands for the neutron with energy reduced from the incident energy.

The emission in this case is prompt. The major gamma-ray lines expected from materials of interest in planetary exploration are listed in Table 3.3. The yields are strongly dependent on the spectral distributions of the exciting neutrons. That is, the gamma-ray energy produced depends on the energy levels that can be excited. Assume a neutron spectral distribution corresponds to the lunar case; the yields are then the fraction of gamma rays with given energy produced per de-excitation of the excited level. Another source of gamma-ray emission is also included in Table 3.3: the primary Galactic cosmic-ray interactions with the surface materials. These interactions are due to the very-high-energy protons (above ~100 MeV). This type of interaction leads to the production of radionuclides by spallation. The latter process is of much greater importance when considering the solar gamma rays and activation of gamma-ray detectors during spaceflight.

Secondary neutrons, with energies below the first excited level of the nuclei, in a condensed medium irradiated by the incident cosmic rays can be elastically scattered by various nuclei, and can either escape the media or be captured by other nuclei.

Another important process leading to gamma-ray emission can be described as follows:

$$ {}_Z^AY + {}_0^1n \rightarrow {}_Z^{A+1}Y^* \rightarrow {}_Z^{A+1}Y + \gamma\text{-ray}, \qquad (3\text{-}4a) $$

then, for example, if ${}_ZY^{A+1}$ is an unstable nucleus,

$$ {}_Z^{A+1}Y \rightarrow {}_{Z+1}^{A+1}Y^* + {}_{-1}^{0}e, \qquad (3\text{-}4b) $$

Table 3.3.

Major Discrete Line Gamma-Ray Emission Caused by Neutron Inelastic Scatter and Primary Galactic Cosmic-Ray Interactions With a Lunar-Type Surface

Element	Source	Energy	Yield	Half-Life (where applicable)
C	$^{12}C(n,n\gamma)$	4.4389	1.00	
N	$^{14}N(n,n\gamma)$	5.1049	0.80	
	$^{14}N(n,\alpha\gamma)$	4.4441	1.00	
	$^{14}N(n,n\gamma)$	2.3127	1.00	
	$^{14}N(n,\alpha\gamma)$	2.1245	1.00	
	$^{14}N(n,n\gamma)$	1.6348	1.00	
O	$^{16}O(n,n\gamma)$	8.8691	0.072	
	$^{16}O(n,n\gamma)$	7.1170	1.00	
	$^{16}O(n,n\gamma)$	6.9172	1.00	
	$^{16}O(n,n\gamma)$	6.1294	1.00	
	^{16}N	6.1294	0.69	7.13s
	$^{16}O(n,n\gamma)$	4.949	0.40	
	$^{16}O(n,n\alpha\gamma)$	4.4389	1.00	
	$^{16}O(n,n\gamma)$	4.161	0.44	
	$^{16}O(n,\alpha\gamma)$	3.854	0.691	
	$^{16}O(n,n\gamma)$	3.833	1.00	
	$^{16}O(n,\alpha\gamma)$	3.6842	1.00	
	$^{16}O(n,\alpha\gamma)$	3.088	1.00	
	$^{16}O(n,n\gamma)$	2.7408	0.76	
	$^{16}O(n,n\gamma)$	1.753	0.126	
F	$^{19}F(n,n\gamma)$	1.3569	.93	
	$^{19}F(n,n\gamma)$	1.2358	.97	
	$^{19}F(n,n\gamma)$	0.1971	1.00	

Table 3.3 (continued).

Element	Source	Energy	Yield	Half-Life (where applicable)
Na	$^{23}Na(n,n\gamma)$	2.6396	1.00	
	$^{23}Na(n,n\gamma)$	1.6364	1.00	
	^{20}F	1.6337	.91	10.996s
	^{22}Na	1.2746	1.00	2.602y
	$^{23}Na(n,n\gamma)$	0.4399	1.00	
Mg	$^{24}Mg(nn\gamma)$	4.238	0.79	
	$^{24}Mg(nn\gamma)$	3.8671	0.983	
	^{24}Na	2.7539	0.9995	15.030h
	$^{24}Mg(n,n\gamma)$	2.7539	1.00	
	$^{26}Mg(n,n\gamma)$	1.8087	1.00	
	$^{25}Mg(n,n\gamma)$	1.6117	1.00	
	$^{24}Mg(n,n\gamma)$	1.3686	1.00	
	^{24}Na	1.3686	1.00	15.03h
	^{22}Na	1.2746	0.9994	2.602y
	$^{26}Mg(n,n\gamma)$	1.1297	0.90	
Al	$^{27}Al(n,n\gamma)$	4.580	0.75	
	$^{27}Al(n,n\gamma)$	4.409	0.59	
	$^{27}Al(n,n\gamma)$	3.9556	0.84	
	$^{27}Al(n,n\gamma)$	3.2103	0.86	
	$^{27}Al(n,n\gamma)$	3.004	0.89	
	$^{27}Al(n,n\gamma)$	2.981	.97	
	^{24}Na	2.7539	0.9995	15.030h
	$^{27}Al(n,n\gamma)$	2.734	0.22	
	$^{27}Al(n,n\gamma)$	2.2997	0.77	

Table 3.3 (continued).

Element	Source	Energy	Yield	Half-Life (where applicable)
	$^{27}Al(n,n\gamma)$	2.2104	1.00	
	^{26}Al	1.8087	0.997	$7.16x10^5y$
	$^{27}Al(n,n\gamma)$	1.8087	1.00	
	$^{27}Al(n,n\gamma)$	1.7195	0.76	
	^{24}Na	1.3686	1.00	15.030h
	^{22}Na	1.2746	0.9994	2.602y
	$^{27}Al(n,n\gamma)$	1.0144	0.97	
	$^{27}Al(n,n\gamma)$	0.8438	1.00	
Al	^{27}Mg	0.8438	0.71	9.462m
Si	$^{28}Si(n,n\gamma)$	7.4162	0.94	
	$^{28}Si(n,n\gamma)$	6.8777	0.64	
	$^{28}Si(n,n\gamma)$	5.6012	0.62	
	$^{28}Si(n,n\gamma)$	5.1094	1.00	
	$^{28}Si(n,n\gamma)$	5.0992	0.33	
	$^{28}Si(n,n\gamma)$	4.4972	0.92	
	$^{28}Si(n,n\gamma)$	3.2000	1.00	
	$^{28}Si(n,n\gamma)$	2.8387	1.00	
	^{24}Na	2.7539	0.9995	15.030h
	$^{30}Si(n,n\gamma)$	2.2354	1.00	
	^{26}Al	1.8087	0.997	$7.16x10^5y$
Si	$^{28}Si(n,n\gamma)$	1.7788	1.00	
	^{28}Al	1.7788	1.00	2.2405m
	^{24}Na	1.3686	1.00	15.03h
	^{22}Na	1.2746	0.9994	2.602y

Table 3.3 (continued).

Element	Source	Energy	Yield	Half-Life (where applicable)
	$^{29}Si(n,n\gamma)$	1.2733	1.00	
P	$^{31}P(n,n\gamma)$	2.2337	1.00	
	$^{31}P(n,n\gamma)$	1.2661	1.00	
S	$^{32}S(n,n\gamma)$	2.2301	1.00	
Cl	$^{35}Cl(n,n\gamma)$	1.7632	1.00	
	$^{35}Cl(n,n\gamma)$	1.2194	1.00	
Ar	$^{40}Ar(n,n\gamma)$	1.4608	1.00	
K	$^{39}K(n,n\gamma)$	2.8137	1.00	
	$^{39}K(n,n\gamma)$	2.5225	1.00	
Ca	$^{40}Ca(n,n\gamma)$	5.2486	0.81	
	$^{40}Ca(n,n\gamma)$	3.9044	1.00	
	$^{40}Ca(n,n\gamma)$	3.7366	1.00	
	$^{40}Ca(n,\alpha\gamma)$	1.6112	1.00	
	$^{40}Ca(n,\rho\gamma)$	1.1589	0.81	
	$^{40}Ca(n,n\gamma)$	1.1569	1.00	
	$^{40}Ca(n,\rho\gamma)$	0.8916	1.00	
	$^{40}Ca(n,\rho\gamma)$	0.7705	1.00	
Ti	$^{48}Ti(n,n\gamma)$	1.3112	1.00	
	$^{48}Ti(n,n\gamma)$	0.9834	1.00	
	$^{46}Ti(n,n\gamma)$	0.8892	1.00	
	^{46}Sc	0.8892	1.00	83.80d
Cr	$^{52}Cr(n,n\gamma)$	1.5308	1.00	
	$^{52}Cr(n,n\gamma)$	1.4342	1.00	

Table 3.3 (continued).

Element	Source	Energy	Yield	Half-Life (where applicable)
	$^{52}Cr(n,n\gamma)$	1.3338	.98	
	$^{55}Mn(n,n\gamma)$	1.5289	1.00	
	$^{55}Mn(n,n\gamma)$	1.1660	1.00	
	$^{55}Mn(n,n\gamma)$	0.8585	1.00	
	^{54}Mn	0.8348	1.00	312.20d
	$^{55}Mn(n,n\gamma)$	0.1260	1.00	
Fe	$^{56}Fe(n,n\gamma)$	3.6019	0.69	
	$^{56}Fe(n,n\gamma)$	2.601	1.00	
	$^{56}Fe(n,n\gamma)$	2.5231	0.85	
	$^{56}Fe(n,n\gamma)$	2.1129	1.00	
	$^{56}Fe(n,n\gamma)$	1.8109	.98	
	$^{54}Fe(n,n\gamma)$	1.4077	1.00	
	$^{56}Fe(n,n\gamma)$	1.3164	1.00	
	$^{56}Fe(n,n\gamma)$	1.2383	1.00	
	$^{56}Fe(n,2n\gamma)$	1.0380	.99	
	$^{56}Fe(n,2n\gamma)$	0.9312	1.00	
	$^{56}Fe(n,n\gamma)$	0.8467	1.00	
	^{54}Mn	0.8348	1.00	312.20d
Ni	$^{58}Ni(n,n\gamma)$	1.4544	1.00	
	$^{60}Ni(n,n\gamma)$	1.3325	1.00	
	^{58}Co	0.8106	1.00	70.78d
Sr	$^{88}Sr(n,n\gamma)$	1.8360	1.00	
	^{88}Y	1.8360	0.994	106.61d
	$^{89}Y(n,n\gamma)$	1.7445	1.00	

Table 3.3 (continued).

Element	Source	Energy	Yield	Half-Life (where applicable)
	$^{89}Y(n,n\gamma)$	1.5074	1.00	
Zr	$^{90}Zr(n,n\gamma)$	2.1865	1.00	
	$^{92}Zr(n,n\gamma)$	0.9345	1.00	
	$^{94}Zr(n,n\gamma)$	0.9182	1.00	
	^{89}Zr	0.9092	1.00	78.43h
Ba	$^{138}Ba(n,n\gamma)$	1.4359	1.00	

Note: y = years, d = days, h = hours, m = minutes, s = seconds.

and possibly

$$^{A+1}_{Z+1}Y^{*} \rightarrow {}^{A+1}_{Z+1}Y + \gamma - \text{ray}. \qquad (3\text{-}4c)$$

The gamma-ray emission accompanying the process in Equation (3-4a) is prompt and will be referred to as a prompt capture process. The gamma ray emitted by the decay of the radioactive nuclide shown in Equations (3-4b) and (3-4c) follows the decay of the electron ($^{0}_{-1}e$). The energy involved in the prompt capture is comparable to the binding energy of the nucleus and thus, in most cases, is higher than the energy of the gamma rays with the given energy produced by inelastic interaction or by the decay of radionuclides. Table 3.4 is a compilation of prompt capture lines and activation lines. The yield is the fraction of gamma rays with the given energy produced per excitation of the excited level.

For the activation products in Table 3.4, the yield is the fraction of gamma rays emitted per decay of the radionuclide. The half-lives of the radionuclides are also indicated. It must be remembered that the nature of the primary cosmic-ray flux, and thus the secondary particle flux can be strongly influenced by the presence of a magnetic field and atmosphere around a planetary body. Therefore, calculation of induced gamma-ray emission will be different for various planetary bodies in our solar system.

Table 3.4.

Major Discrete Line Emission caused by Prompt Capture and Neutron activation in Lunar-Type Material

Element	Source	Energy	Yield	Half-Life (where applicable)
H	$^{1}H(n,\gamma)$	2.2233	1.00	
N	$^{14}N(n,\gamma)$	10.8295	0.135	
	$^{14}N(n,\gamma)$	6.3225	0.184	
	$^{14}N(n,\gamma)$	5.5334	0.185	
	$^{14}N(n,\gamma)$	5.298	0.217	
	$^{14}N(n,\gamma)$	5.2693	0.311	
Na	$^{23}Na(n,\gamma)$	6.395	0.20	
	^{24}Na	2.7539	0.9995	15.030h
	^{24}Na	1.3686	1.00	15.030h
	$^{23}Na(n,\gamma)$	0.4723	1.00	
Mg	$^{24}Mg(n,\gamma)$	3.918	0.48	
	$^{24}Mg(n,\gamma)$	2.8285	0.36	
Al	$^{27}Al(n,\gamma)$	7.724	0.30	
	$^{27}Al(n,\gamma)$	7.694	0.045	
	$^{27}Al(n,\gamma)$	4.735	0.057	
	$^{27}Al(n,\gamma)$	4.2600	0.06	
	$^{27}Al(n,\gamma)$	4.1341	0.065	
	$^{27}Al(n,\gamma)$	3.0345	0.08	
	$^{27}Al(n,\gamma)$	2.960	0.09	
	^{28}Al	1.7788	1.00	2.2405m
Si	$^{28}Si(n,\gamma)$	7.200	0.08	
	$^{28}Si(n,\gamma)$	6.381	0.13	
	$^{28}Si(n,\gamma)$	4.934	0.61	
	$^{28}Si(n,\gamma)$	3.5395	0.66	
Si	$^{28}Si(n,\gamma)$	2.0931	0.20	

Table 3.4 (continued).

Element	Source	Energy (MeV)	Yield	Half-Life (where applicable)
	$^{28}Si(n,\gamma)$	1.2733	0.19	
S	$^{32}S(n,\gamma)$	5.424	0.55	
S	$^{32}S(n,\gamma)$	2.379	0.40	
Cl	$^{35}Cl(n,\gamma)$	7.791	0.09	
	$^{35}Cl(n,\gamma)$	7.415	0.11	
	$^{35}Cl(n,\gamma)$	6.621	0.09	
	$^{35}Cl(n,\gamma)$	6.111	0.21	
	$^{35}Cl(n,\gamma)$	1.951	0.20	
K	$^{39}K(n,\gamma)$	7.770	0.07	
	$^{39}K(n,\gamma)$	5.7525	0.07	
	$^{39}K(n,\gamma)$	5.697	0.07	
	$^{39}K(n,\gamma)$	5.381	0.09	
	$^{39}K(n,\gamma)$	0.7705	0.58	
Ca	$^{40}Ca(n,\gamma)$	6.4201	0.40	
	$^{40}Ca(n,\gamma)$	5.9005	0.07	
	$^{40}Ca(n,\gamma)$	4.419	0.18	
	$^{40}Ca(n,\gamma)$	2.0015	0.18	
	$^{40}Ca(n,\gamma)$	1.9427	0.80	
Ti	$^{48}Ti(n,\gamma)$	6.7615	0.40	
	$^{48}Ti(n,\gamma)$	6.557	0.04	
	$^{48}Ti(n,\gamma)$	6.419	0.28	
	$^{48}Ti(n,\gamma)$	4.967	0.03	
	$^{48}Ti(n,\gamma)$	4.882	0.05	
	$^{48}Ti(n,\gamma)$	1.7620	0.045	
Ti	$^{48}Ti(n,\gamma)$	1.5853	0.09	
	$^{48}Ti(n,\gamma)$	1.4983	0.04	

Table 3.4 (continued).

Element	Source	Energy (MeV)	Yield	Half-Life (where applicable)
	$^{48}Ti(n,\gamma)$	1.3815	0.82	
	$^{48}Ti(n,\gamma)$	0.3419	0.38	
Cr	$^{53}Cr(n,\gamma)$	9.719	0.10	
	$^{53}Cr(n,\gamma)$	8.8841	0.24	
	$^{52}Cr(n,\gamma)$	7.939	0.11	
Mn	$^{55}Mn(n,\gamma)$	7.244	0.10	
	$^{55}Mn(n,\gamma)$	7.058	0.09	
	^{56}Mn	1.8109	0.29	2.5785h
	^{56}Mn	0.8467	0.988	2.5785h
	$^{54}Fe(n,\gamma)$	9.299	0.034	
	$^{56}Fe(n,\gamma)$	7.6457	0.22	
	$^{56}Fe(n,\gamma)$	7.6313	0.24	
	$^{56}Fe(n,\gamma)$	7.279	0.05	
	$^{56}Fe(n,\gamma)$	6.019	0.08	
	$^{56}Fe(n,\gamma)$	5.921	0.08	
	$^{56}Fe(n,\gamma)$	4.810	0.018	
	$^{56}Fe(n,\gamma)$	4.2185	0.04	
	$^{56}Fe(n,\gamma)$	1.725	0.09	
	$^{56}Fe(n,\gamma)$	1.6126	0.07	
	$^{56}Fe(n,\gamma)$	0.6921	0.08	
Ni	$^{58}Ni(n,\gamma)$	8.999	0.37	
	$^{58}Ni(n,\gamma)$	8.534	0.18	
Y	$^{87}Y(n,\gamma)$	1.8360	0.45	
Sr	$^{89}Sr(n,\gamma)$	6.080	0.41	
Nd	$^{143}Nd(n,\gamma)$	0.697	0.65	
Sm	$^{149}Sm(n,\gamma)$	0.4395	0.41	
	$^{149}Sm(n,\gamma)$	0.3340	0.71	

Table 3.4 (continued).

Element	Source	Energy (MeV)	Yield	Half-Life (where applicable)
Eu	^{156g}Eu	1.409	0.15	15.11d
	^{156g}Eu	1.113	0.09	15.11d
Gd	^{157}Gd(n,γ)	6.747	0.02	
	Gd(n,γ)	1.187	0.11	
	^{157}Gd(n,γ)	0.945	0.07	
	^{157}Gd(n,γ)	0.182	0.22	

Note: y = years, d = days, h = hours, m = minutes, s = seconds.

Table 3.5 is a tabulation of spallation products of interest in planetary exploration.

Table 3.5.

Spallation Products of Interest in Planetary Exploration

Element	Source	Energy (MeV)
Mg	^{20}Ne*	1.6337
Al	^{23}Na*	1.6364
	^{20}Ne*	1.6337
	^{24}Mg*	1.3686
Si	^{27}Al*	2.2104
	^{20}Ne*	1.6337
	^{24}Mg*	1.3686
	^{27}Al*	1.0144
Ca	^{39}K*	2.8137
	^{39}Ca*	2.793
	^{32}S*	2.2301
	^{36}Ar*	1.9704
Fe	^{52}Cr*	1.4342

3.7 Solar Proton Interactions

Both charged and neutral particles are released from the Sun during strong solar flares. About 90 percent of these particles are protons. For the purpose of this discussion, only proton interactions will be considered. The spectral shape and intensity of the solar proton flux arriving at a planetary surface will depend on the distance from the Sun, on the presence of a magnetic field, and on an atmosphere about the planetary body. The solar proton flux is much lower in average energy than the cosmic-ray flux. Thus, the induced gamma-ray emission can be determined from an analysis of the primary-particle interaction, since the secondary particles produced will not be significant.

The solar protons will travel only a few centimeters in planetary materials. They will slow down by such interaction as inelastic scatter. Capture can also occur. Prompt and delayed gamma-ray emission can be produced. In terms of remote gamma-ray analysis, only the delayed gamma-ray emission produced by the decay of a radioactive nuclide will be detectable, since the prompt emission will be coincidental with the arrival of the solar proton flux. During large solar events, protons will tend to saturate the gamma-ray detectors. Observation can therefore be made only after the proton flux has stopped.

The magnitude and spectral distribution of gamma-ray emission caused by the solar proton flux activation of a planetary surface will depend on not only the most recent event leading to proton irradiation, but also on the previous history of all other such events. From the half-lives of the activated species one can infer how far back in time the particular event produces nuclides of importance in remote gamma-ray sensing. Some of the more important activated species observed on the Moon during Apollo 15 and 16 are shown in Table 3.6. The target nucleus is indicated, as well as the activated radionuclides. Details for calculation of the solar proton flux and the radionuclides produced in the Moon can be found in Reedy and Arnold (1972), Reedy et al. (1973), and Reedy (1978).

Table 3.6.

Solar Proton-Induced Gamma-Ray Emission Observed During Apollo 15 and Apollo 16

Element	Nuclide	Energy (MeV)	Half-Life
Mg	^{22}Na	1.2745	2.602y
Al	^{26}Al	1.8087	7.16×10^5y
	^{22}Na	1.2745	2.602y
Si	^{26}Al	1.8087	7.16×10^5y
	^{22}Na	1.2745	2.602y
Fe	^{56}Co	1.2383	78.76d
	^{56}Co	0.8467	78.76d
	^{54}Mn	0.8348	312.20d

Note: y = years, d = days.

3.8 Experimental Results

One of the early objectives of the space program has been understanding the origin, development, and present-day dynamics of the solar system. Remote orbital and in situ observations of x-rays, gamma rays, and particle emission from the surfaces of planetary bodies has provided new information relative to these objectives. In Chapter 2 a number of questions regarding the understanding of the evolution of the solar system were presented. It was indicated that the determination of a global distribution of the major, minor, and trace elements in such bodies as the planets, comets, and asteroids can contribute significantly to the development of evolutionary models of the solar system. For example, a knowledge of the overall elemental composition of a given body can be related to the mechanism of condensation and accumulation from the primordial solar nebula. Subsequent to accretion, the bulk distribution of elements has usually been affected by evolutionary processes. Thus, inferences can be made concerning early processes from a study of the present distributions.

During the manned and unmanned missions to the Moon and planets, x-ray and gamma-ray remote sensing systems have been

flown and successful measurements achieved. The combination of remote sensing, in situ analysis, and returned-sample analysis has allowed us to obtain extensive global geochemical data on the Moon. In addition, the combination of orbital measurements and surface sampling has enabled us to verify the remote analysis measurement system.

A brief review of the missions in which successful measurements of x-ray, gamma-ray, and α-particle emission were accomplished will now be presented. Some of the more significant results relating to the development of models of solar evolution will also be indicated. Because it will be impossible to detail the results from all such flights, results from the remote sensing x-ray and gamma-ray experiments aboard Apollo 15 and Apollo 16 will be considered in greater detail. It is hoped that in this way the power of these remote sensing techniques can be demonstrated and understood.

Rangers 3, 4, and 5 (1961-1962) carried the first gamma-ray spectrometers in space. Although the spacecrafts did not come close enough to the Moon to measure emissions from the surface, results were obtained on the spacecrafts and cosmic backgrounds. The first indication of an isotropic diffuse cosmic gamma-ray background was found as a result of an analysis of the Ranger gamma-ray data (Metzger et al., 1964). In 1966, the Soviet lunar orbiter mission, Luna, carried a gamma-ray spectrometer, and the first gamma-ray spectrum near the Moon was obtained. Detailed spectral information was not obtained because of the poor energy resolution of the detector and the background problems, but compositional distinctions between mare and terra provinces were recognized (Vinogradov et al., 1967). The first successful in situ measurement on the Moon was the α-backscatter experiment carried by Surveyor 5 in 1967. The surface composition at the landing site was found to be basaltic in nature. Measurements using this experiment on a number of Surveyor missions consistently indicated that the mare areas at the landing site were also basaltic (Turkevich, 1969). During the Apollo 15 and 16 flights in 1971 and 1972, gamma-ray, x-ray, and charged-particle spectrometers were carried on the orbiting Command and Service Modules. Approximately 20 percent of the lunar surface was mapped for magnesium, aluminum, thorium, potassium, uranium, silicon, titanium, and oxygen (Adler et al., 1974; Metzger et al., 1973; and Bjorkholm et al., 1973). The results obtained for the gamma-ray and x-ray spectrometer

experiments will be considered in detail in the following sections. Two Soviet Lunakhod roving vehicles carried x-ray fluorescence experiments. These detection systems were used to screen materials for selection for return flights to Earth (Kocharov and Viktorev, 1974).

Remote nuclear techniques have been applied to Mars and Venus, though less extensively than to the Moon. A gamma-ray spectrometer aboard the Soviet Mars 5 orbiter obtained several hours of data from an altitude of about 9300 km, indicating that the average potassium, uranium, and thorium content of the regions surveyed corresponds with that of terrestrial rocks like oceanic basalts (Surkov et al., 1976). The two Viking Landers carried an x-ray spectrometer for the fluorescence analysis of selected soil and rock samples (Clark et al., 1977). The finely divided samples at the two landing sites were remarkably similar, with high iron and sulfur content, and little potassium.

Three Soviet spacecraft, Venera 8, 9, and 10, have carried a gamma-ray spectrometer to measure ^{40}K, uranium and thorium at the surface of Venus by detecting the decay of these naturally radioactive elements (Surkov, 1977). In addition, Venera 10 had a second gamma-ray instrument onboard to measure the bulk density of the surface materials. High abundances were found for potassium, uranium, and thorium at the Venera 8 site, the potassium value being comparable to granitic rocks on Earth, while the uranium and thorium values were similar to those of alkali basalts. The Th/U ratio agrees with that found for major rock types on Earth, on the Moon, and in chondritic meteorites. The subsequent landings of Venera 9 and 10 measured sites with much lower gamma-ray activity, with values for potassium and uranium similar to those of terrestrial oceanic basalt rocks.

3.9 Apollo X-Ray Spectrometer Results

The Apollo x-ray spectrometer consisted of three large-area proportional detectors, a set of large-area filters for energy discrimination among the characteristic x-rays of aluminum, silicon, and magnesium, collimators, and data-handling electronics for obtaining 8-channel, pulse-height spectra. These three detectors were pointed at the lunar surface while taking measurements. A fourth proportional counter that looked in the direction opposite to

the three large-area proportional counters was used as a solar monitor. The characteristic x-ray emission flux variation reflects the changes in nature of the different surface elemental composition overflow. A number of other factors can also produce variation in the observed x-ray emission flux from orbit. The variation in intensity caused by changes in the incident solar flux was monitored by measuring the solar differential energy spectra incident on the lunar surface. Correction methods for solar-flux variation have been developed (e.g., Clark, 1979). Other sources of variation can be attributed to such factors as matrix effects, solar illumination angle, and surface roughness. Those effects are about the same for the magnesium, aluminum, and silicon x-ray emission processes. These other factors can be removed substantially by the use of the intensity ratios, Al/Si and Mg/Si (Adler et al., 1972), which provide information on the magnesium and aluminum variation, because the silicon variation is relatively small over the lunar surface covered by Apollo 15 and 16.

When considering the results obtained, three other factors must be kept in mind. First, the depth of surface materials sampled corresponds to tens of micrometers. Second, the collimators used on the detector system allowed observations to be made with spatial resolution of about 10 km on the lunar surface when the Apollo orbiter was at about 100 km from this surface; most of the measurements were obtained at about this altitude. Finally, observations can be obtained on the sunlit portions of the Moon, since only x-rays from the Sun produce the fluorescence.

Figures 3.5 and 3.6 show Al/Si and Mg/Si ratio profiles along a northern track observed during the Apollo 15 mission (Adler et al., 1972). The ratios for various analyzed materials are shown along the right-hand axis for reference. Each point on the graph represents 16 seconds of data accumulation. A number of observations can be drawn from a study of these figures.

- The Al/Si ratios are high in the highlands and are considerably lower in the mare areas. The extreme variation is about a factor of two. The Mg/Si concentration ratios generally show the opposite relationship.

- There is a general tendency for Al/Si ratio values to increase from the western mare areas to the eastern limb highlands.

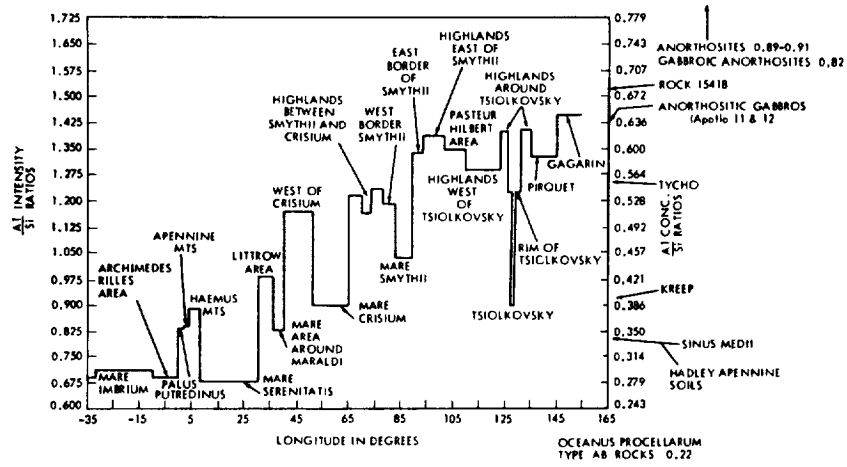

Figure 3.5. Al/Si ratio as a function of longitude for an Apollo 15 ground track. The values for some reference materials are indicated on the right-hand axis.

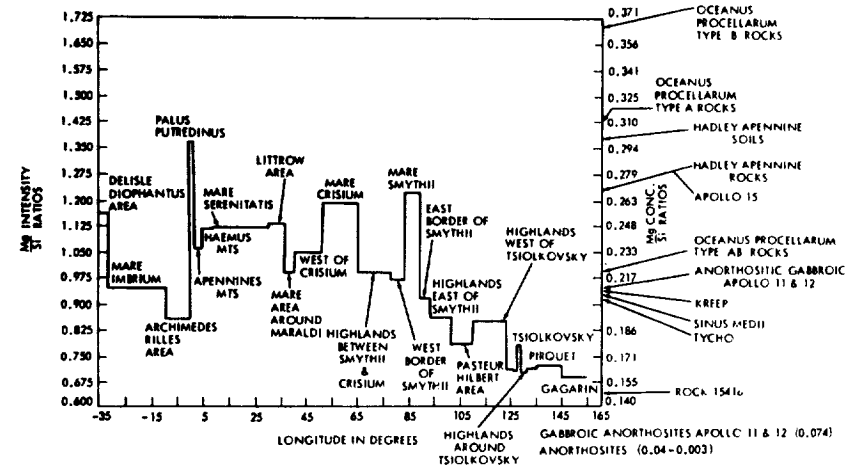

Figure 3.6. The Mg/Si ratio as a function of longitude for an Apollo 15 ground track. The values for some reference material are indicated on the right-hand axis.

• There are distinct chemical contrasts between such features as the small mare basins and the highland rims (note, for example, the crater Tsiolkovsky in Figure 3.5). The rim areas are intermediate between the mare areas and the surrounding highlands.

• There was agreement between the Al_2O_3 content of the returned Apollo samples and that inferred from the x-ray measurements. This demonstrated that the orbital measurements are a reliable guide to at least this aspect of the lunar surface chemistry. Furthermore, a remarkable factor can be inferred regarding the vertical profile — that the dust surface is characteristic of the composition below.

The figure represents merely a few of the results obtained, but with these results the following observations can be made: This experiment clearly demonstrated that the Moon's crust is chemically differentiated. The global features can be used to define the extent of differentiation and to clearly distinguish the mare basalt (low aluminum) from the highland anorthosites (high aluminum material). On a medium scale, it is seen that the terra and mare are increasingly aluminous from west to east, suggesting an inhomogeneous lunar crust.

Figure 3.7 shows a color coded map of all the x-ray data obtained during Apollo 15 and 16 missions. The Al/Si ratio is plotted (Andre et al., 1977). If one carefully studies this map, the impact-ray structure caused by meteor impact can be seen. This information can be used to obtain data on the vertical distribution of elemental composition, since these craters can disclose layers deep below the surface.

Finally, on a smaller scale (~20 km), details of cratering on the Moon can be investigated. Figure 3.8 shows impact cratering effects near the Mare Crisium area. Here the Mg/Al ratio is plotted. Because of the intermittent inverse correlation between these elements, the change in elemental composition is emphasized. Such craters as Picard and others, and a significant chemical contrast relative to the surroundings (Andre et al., 1978) are shown. This contrast is a surface expression of the excavation of the subsurface material of different compositions. Higher magnesium is shown as an increase in red, while higher aluminum is shown as an increase in blue. The dark red area indicates a deep penetration into the surface. The surrounding material is blue, characteristic of highland material, as compared with the basaltic material characterized by the red.

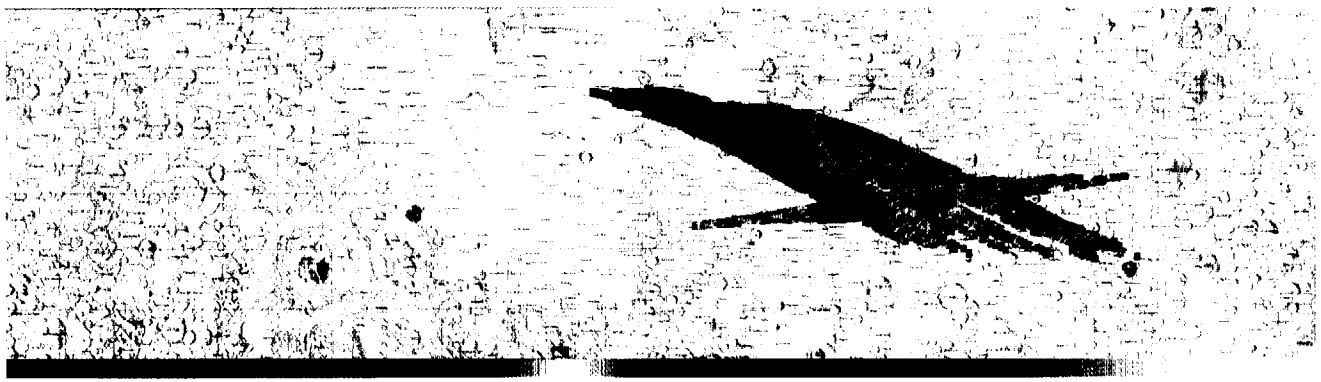

Figure 3.7. Color-coded Al/Si map of all of the x-ray fluorescence data obtained during Apollo 15 and Apollo 16 (Andre et al., 1977, and Clark, 1977).

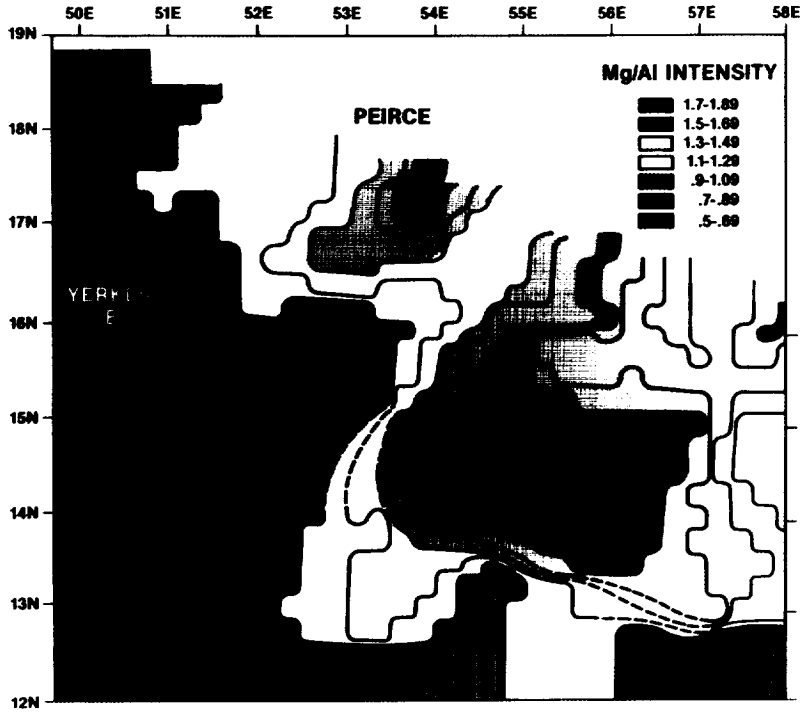

Figure 3.8. Color-coded Mg/Al ratio map in the Mare Crisium area showing lunar surface cratering on a ~20 km resolution scale (Andre et al., 1978).

3.10 The Apollo Gamma-Ray Spectrometer

The Apollo gamma-ray detector consisted of a 7 x 7 cm right cylindrical NaI(Tl) crystal detector surrounded by a plastic anticoincidence mantle for suppressing events caused by charged particles. The electronic processor consisted of a 512-channel pulseheight analyzer, including an amplifier, but no memory. The information was transmitted channel by channel, either in real time or stored on magnetic tape for subsequent transmission. Spectra were obtained by sorting the received pulses for various time periods. Approximately 30 minutes of data accumulation are required to obtain a statistically significant spectrum. The detector was deployed on a boom some 8.3 m from the spacecraft in order to minimize the background gamma radiation, produced on the

spacecraft by cosmic-ray interactions. The instrument design, the data analysis, and the background problems will be discussed in Chapters 12 and 13. Details on detector design can be found in papers by Arnold et al. (1972) and Harrington et al. (1974).

Figure 3.9 is an energy-loss or pulse-height spectrum of the discrete line emission from the Moon. By the energy loss or pulse height we mean the discrete line spectrum as observed by the scintillation detector. There have been no corrections for detector response and efficiency as a function of energy. The conversion to energy from pulse height is 19.5 keV/channel. The energy region up to approximately channel 140 is dominated by the gamma rays produced by the natural radioactive components ^{40}K, thorium, and uranium. Thus, the integrated counts up to this energy will reflect the variation in these radioactive elements. Figure 3.10 is a color-coded map indicating the variation in intensity of natural radioactive emission of the Moon for those areas observed during Apollo 15 and Apollo 16 (Metzer, 1993, and Clark, 1997).

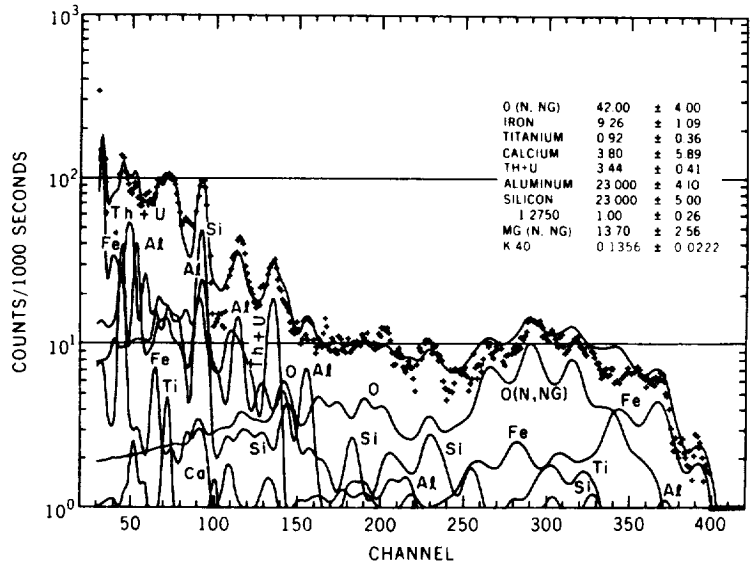

Figure 3.9. Discrete line spectrum measured by the Apollo gamma-ray spectrometer during Apollo 15. The background has been subtracted. The characteristic gamma-ray signature for each of the elements contributing a significant photon flux is also shown. The energy to pulse-height conversion is approximately 19.5 keV per channel.

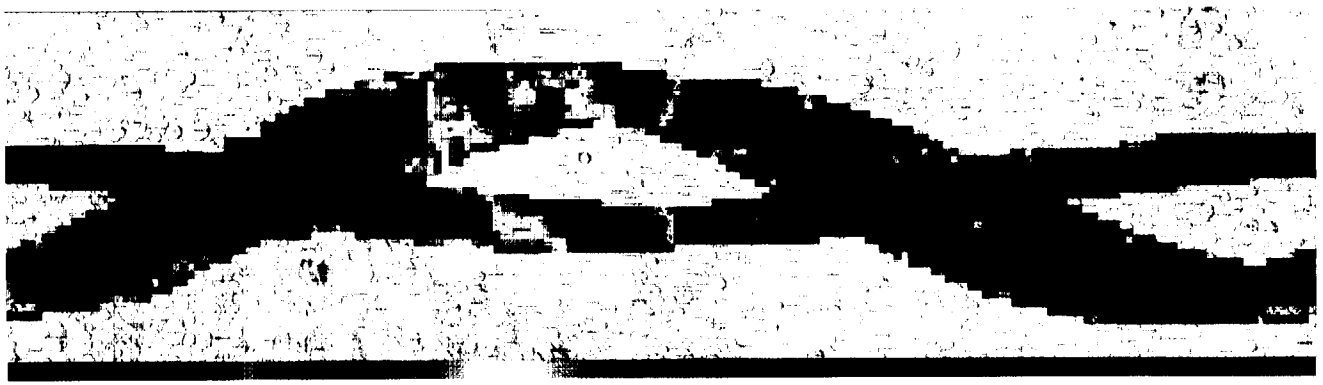

Figure 3.10. Distribution of naturally radioactive element potassium over the region overflown by the Apollo 15 and Apollo 16 Command and Service Modules. The count rate which is proportional to the concentration of the radioactive elements is color coded: blue for low count rate to red for the highest count rate.

Low radioactivity is indicated in blue and increases in intensity to red. Because of the higher count rate, the best spatial resolution maps can be obtained for these components. Similar maps have been obtained for iron and lithium. More complete elemental compositions have been obtained over much larger areas of the Moon. Table 3.7 shows a number of more complete elemental analyses over a few areas of the Moon. As can be seen in Figure 3.10, the area measured by the gamma-ray spectrometer system is much larger than that covered by the x-ray system. Remember that the gamma-ray emission is caused by natural radioactive and cosmic-ray-induced activation and thus is independent of the position of the Sun relative to the Moon.

Table 3.7.

Elemental Composition of Selected Lunar Sites as Determined from the Apollo 15 Remote Sensing Gamma-Ray Experiment

Lunar Feature	Coordinate Boundary Region	Element Concentration in				
		Fe(%)	Mg(%)	Ti(%)	Th(ppm)	K(ppm)
Van de Graff	168°W - 168°E	7.7	3.8	0.1	2.3	1600
Highland East	88°E - 60°E	6.5	4.5	1.3	1.0	940
Mare Feccunditatis	60°E - 42°E	11.3	7.0	2.2	1.2	1400
Mare Imbrium	15°W - 39°W	13.6	6.2	1.4	5.8	1700
Aristarchus Region	54°W - 81°W	9.6	4.9	2.2	6.9	2500
Highland West	81°W - 68°W	5.7	3.5	1.5	0.4	950
Average over all of Apollo 15 Orbit		8.7	4.8	1.5	2.2	1230

Note: ppm = parts per million.
Source: Trombka et al. (1977).

A number of characteristics of the elemental distributions shown in Figure 3.10 and Table 3.7 should be pointed out. Some of the factors are considered relative to the analysis of returned lunar samples. Details of the analysis carried out on the lunar samples can be found in the yearly Proceedings of the Lunar Science Conferences from 1968 to the present. The articles are too numerous to list here. The observations cannot be extended to the whole Moon; however, they represent the areas overflown during the Apollo 15 and Apollo 16 missions only.

The regions within and bounding the western maria (Oceanus Procellarum) show higher levels of radioactivity than any others elsewhere on the lunar surface. There is a striking contrast between this region and the rest of the Moon, particularly the eastern maria. Further, there is a detailed structure in the distribution of radioactivity within the high-radioactivity regions. The highest concentration, shown in red in Figure 3.10, was in the Aristarchus region in the high ground west of the Apollo 15 landing site and south of Archimedes, and in the area south of the Fra Mauro crater. The Fra Mauro site overflown is about 7^{o} south of the Apollo 14 landing site. The soil from the landing sites showed levels comparable to the orbital data.

The eastern maria showed evidence of local enhancement, although the radioactivity was lower than that in the western maria. The highland regions showed relatively low radioactivity except on the borders of the western maria, where lateral mixing may have occurred. The eastern far side highlands (180^{o} to 90^{o} east) were measurably more radioactive than the western highlands (90^{o} to 180^{o} west). A small maximum in activity was found near Van de Graff, where a major magnetic anomaly exists.

Laser altimeter profiles were obtained during the Apollo mission (Sjorgren and Wollenhaupt, 1973). Comparison between the Al/Si ratio data shown in Figure 3.5, the distribution of radioactivity (Figure 3.10), and the laser altimeter data (Trombka et al., 1973) reveals the following observations: the Al/Si ratio profile directly correlates with the laser altimeter data. With the exception of the Imbrium Procellarium area, marked inverse correlation can be seen between the natural radioactivity and the elevation. While this inverse correlation is also broadly true of the Imbrium Procellarium region, it does not hold up with regard to the detailed structure in the

variation in radioactivity. On the far side, this inverse correlation extends to an observation of greater east-west asymmetry (around 180° for the Apollo 16 trajectory) than that for Apollo 15: an asymmetry which exists for both elevation and natural radioactivity. This correlation and the inverse correlation appear to reflect the nature and extent of major lunar differentiation processes. If the Moon is in an isostatic equilibrium, then the more extensive the early anorthositic differentiation (characterized by lower densities and lower concentrations of the naturally radioactive nuclides, compared to the mare), the higher the aluminum concentration and the lower the radioactivity.

The major depression, which occurs in the vicinity of the crater Van de Graff, exhibits the sharpest contrast in elevation to the adjacent highlands, a difference on the order of about 8 km. The same area is also the site of the only major far-side enhancement in natural radioactivity. A strong magnetic feature is also observed in this area by magnetometers aboard the Apollo subsatellite (Coleman et al., 1973). Thus, over this portion of the lunar surface, as scanned up to date, are the largest surface remnant magnetic fields, one of the deepest depressions, and the only significant far-side enhancement in radioactivity. All such features have been observed within about 150 km of each other. This Van de Graff area is quite notable and, to date, quite a singular exception to the general conditions prevailing on the lunar far side. When understood, it is likely to contribute significantly to our understanding of lunar evolution. From the chemical composition shown in Table 3.7, the Van de Graff shows a difference from any thus far observed on the Moon. The other mare and highland-like compositions have been found abundantly in returned lunar samples. For the Van de Graff composition, the major elements are highland-like, though the iron concentration is a little high. The concentration of potassium and thorium are very similar to the eastern mare. One returned lunar sample, a "granite-like rock" (12013) contained what might be similar to the Van de Graff composition, even though this region is on the far side of the Moon. The origin of this rock may be from this region.

Early interpretations of the gamma-ray results were presented in terms of three components: mare basalts, KREEP (mare-type materials high in potassium (K), rare earths (RE) and phosphorus (P)), and a low radioactivity highland component (Metzger et al., 1973). The Van de Graff component does not fit this model, but

otherwise the components shown in Table 3.7 do follow this pattern. The so-called KREEP materials may have their origin in the region of the hot spots shown in Figure 3.10. Most returned sample compositions can be inferred from a mixture of these basic components.

The large highland regions are not entirely uniform. The iron shows a notable east-west asymmetry. The iron concentration is higher in the eastern regions. The values of titanium found on the far-side highlands are also high and are possibly another example of inhomogeneity in broad highland areas. There are also titanium differences in the maria and KREEP regions.

A most interesting result can be found in the ratio of K/Th. Potassium is a volatile material, and thorium is refractory. Thus, this ratio can be used to measure the volatile-to-refractory material variation. In both missions, this ratio varied around the Moon and was found to be consistently lower than that found on Earth. This reflects a global depletion of volatiles for the Moon as compared to the Earth.

In the last two sections, a review of the results of the Apollo remote sensing gamma-ray and x-ray results has been presented. Further analyses of the results from the Apollo missions and the Russian missions can be found in Evans et al. (1993); Yin et al. (1993); Feldman et al. (1993); Metzger (1993); Surkov et al. (1993). The only successful x-ray and gamma-ray remote mapping of a planetary body up to this time has been on the Apollo 15 and 16 missions. There have been a number of attempts in the past few years, e.g., the U.S. Mars Observer, the Russian Phobos, and the Mars '96 mission, but most have ended in failure. The Near Earth Asteroid Rendezvous (NEAR) X-Ray/Gamma-Ray Spectrometer system was launched in February 1996 and will rendezvous with the Asteroid 433 Eros in 1999 and spend one year in orbit around this body. A gamma-ray and neutron spectrometer system will be launched aboard the U.S. Lunar Prospector spacecraft in late 1997 and will spend one year in orbit around the Moon. Finally, remote sensing gamma-ray and neutron spectrometer systems will be included aboard the U.S. Mars Surveyor `01 mission with plans for launch in late 2001. With these three missions, it is hoped that global elemental composition maps will be obtained for the Moon, 433 Eros, and Mars.

References

Adler, I., et al., 1972, Preliminary Science Report, NASA SP-289, p. 436.

Adler, I., and Trombka, J. I., 1977, Phys. and Chem. of the Earth, 10, 17.

Adler, I., Podwysocki, M., Andre, C., Trombka, J. I., Eller, E., Schmadebeck, R., Lin, L., 1974, "The Role of Horizontal Transport as Evaluated from the Apollo 15 and 16 Orbital Experiments," Proc. Lunar Sci. Conf. 5th, p 975-979.

Anders, E., 1977, Phil. Trans. Roy., Soc. London, A285, p. 93.

Andre, C. G., Bielefeld, M. J., Eliason, E., Soderblom L. A., Adler, I., and Philpotts, J. A., 1977, Science, 197, 986.

Andre, G. W., Wolfe, R. W., and Adler, I., 1978, in Mare Crisium: A View from Luna 24, ed. R. Merrill (New York: Pergamon Press, 1978), p. 1.

Arnold, J. R., Peterson, L. E., Metzger, A. E., and Trombka, J. I., 1972, Preliminary Science Report, NASA SP-989.

Bjorkholm, P., Golub, L., and Gorenstein, P., 1973, Proc. Fourth Lunar Planet, Sci. Conf., 3, 2803.

Boynton, W. V., Evans, L. G., Reedy, R. C., and Trombka. J. I., 1993, Determination of Planetary Composition by In-situ and Remote Gamma-Ray Spectrometry," in Remote Geochemical Analysis: Elemental and Mineralogical Composition, ed. C. Pieters and P. Englert, Cambridge University Press, p. 395-411.

Brandt, W., 1972, Proc. Int. Conf. on Inner Shell Ionization Phenomena, USAEC Conf. 720404, 2, 948.

Clark, B. C., Baired, A. K., Rose, H. J., Jr., Toulmin, P., Christian, R. P., Kelliher, W. C., Castro, A. J., Rowe, C. D., Keil, K., and Huss, G. R., 1977, J. Geophys. Res., 82, 4577.

Clark, P. E., 1979, "Corrections, Correlation and Theoretical Considerations of Lunar X-Ray Fluorescence Intensity Ratios," Ph. D. Diss., University of Maryland.

Coleman, P. J., Jr., Lichtenstein, B. R., Russell, C. T., Schubert, G., and Sharp, L. R., 1973, Apollo 16 Preliminary Science Report, NASA SP-315, p. 23-1.

Evans, L. G., Reedy, R. C., and Trombka, J. I., 1993, "Introduction to Planetary Remote Sensing", in Remote Geochemical Analysis: Elemental and Mineralogical Composition", ed. C. Pieters and P. Englert, Cambridge University Press, p. 167-196.

Feldman, W. C., Boynton, W. V., and Drake, D. M., 1993, in Remote Geochemical Analysis: Elemental and Mineralogical Composition , ed. C. Pieters and P. Englert, Cambridge University Press, p. 213-243.

Feldman, W. C., Binder, A. B., Hubbard, G. S., McMurray, R. E., Jr., Miller, M. C., and Prettyman, T. H., 1996, "The Lunar Prospector Gamma-Ray Spectrometer," in: Lunar and Planetary Science XXVII (Lunar and Planetary, Institute, Houston, TX) p. 355-356.

Harrington, T. M., Marshall, J. H., Arnold, J. R., Peterson, L. E., Trombka, J. I., and Metzger, A. E., 1974, Nucl. Instr. and Meth, 118, 401.

Hubbell, J. H., 1969, Photon Cross-Sections, Attenuation Coefficients, and Energy Absorption Coefficients from 10 keV to 100 GeV, NSDRS-NBS 29 (Washington, D.C.: National Bureau of Standards,1969).
Jenkins, R., and DeVries, J., 1967, Practical X-Ray Spectrometry (New York: Springer-Verlag, Inc., 1967).
Kocharov, G., Ye, and Viktorev, S. V., 1974, Doklady Academy of Sciences, USSR, 214, 71.
Kornblum, J. J., and Fireman, E. L., 1974, Preprint No. 51 (Washington, D.C.: Smithsonian Astrophysical Observatory, Center for Astrophysics, 1974).
Lapides, J. R., 1981, Ph. D. Dissertation, University of Maryland.
Lapides, J. R., Spergel, M. S., Lazareth, O. W., Levy, P. W., Reedy, R. C., and Trombka, J. I., 1980, Abstracts Lunar and Planet. Sci. Conf: XI (Houston, Texas: NASA, L. B. Johnson Space Center, 1980), p. 605.
Lingenfelter, R. E., Canfield, E. H., and Hess, W. N.,1961, J. Geophys. Res., 66, 2665.
Lingenfelter, R. E., Canfield, E. H., and Hampel, V. E., 1972, Earth and Planetary Sci. Ltr., 16, 355.
Metzger, A. E., Anderson, E. C., Van Dilla, M. A., and Arnold, J. R., 1964, Nature, 204, 766.
Metzger, A. E., Trombka, J. I., Peterson, L. E., Reedy, R. C., and Arnold, J. R., 1973, Science, 179, 800.
Metzger, A. E., 1993, in Remote Geochemical Analysis: Elemental and Mineralogical Composition , ed. C. Pieters and P. Englert, Cambridge University Press, p. 341-363.
Mitrofanov, I. G., et al., 1996, "High Precision Gamma-Ray Spectrometer PGS for Russian Interplanetary Mission to Mars, Adv. Space Res. 17 (#12), 51-59.
Pieters, C. and Englert, P., 1993, in Remote Geochemical Analysis: Elemental and Mineralogical Composition , ed. C. Pieters and P. Englert, Cambridge University Press.
Reedy, R. C., 1978, Proc. Ninth Lunar Planet. Sci. Conf., 2961.
Reedy, R. C., and Arnold, J. R., 1972, J. Geophys. Res., 77, 537.
Reedy, R. C., Arnold, J. R., and Trombka, J. I., 1973, J. Geophys. Res., 78, 5847.
Sjorgren, W. L., and Wollenhaupt, W. R., 1973, Science, 179, 275.
Spergel, M. S., Lazareth, O. W., Slatest, L. A., and Levy, P. W., 1980, Abstracts Eleventh Lunar Planet, Sci. Conf. (Houston, Texas: NASA, L. B. Johnson Space Center, 1980), p. 1069.
Surkov, Y. A., 1977, Proc. Eighth Lunar Planet. Sci. Conf, 3, 2665.
Surkov, Y. A., Moskalyova, L. P., Kirnozov, F. F., Khyarkukova V. P., Manvelian, O. S., and Shchelov, O. P., 1976, Space Research XVI (Berlin: COSPAR Akademie-Verlag, 1976).
Surkov, Y. A., et al., 1993, in Remote Geochemical Analysis: Elemental and Mineralogical Composition , ed. C. Pieters and P. Englert, Cambridge University Press.
Surkov, Y. A., 1996, private communication.
Touren, L. H., 1972, Phys. Rev., A5., 2482.

Trombka, J. I., Arnold, J. R., Reedy, R. C., Peterson, L. E., and Metzger, A. E., 1973, Proc. Fourth Lunar Planet. Sci. Conf., Geochim Cosmachim Acta, 3, 2847.

Trombka, J. I., Arnold, J. R., Adler, I., Metzger, A. E., and Reedy, R. C., 1977, The Soviet American Conference on Cosmochemistry of the Moon and Planets, ed. J. H. Pomeroy, and N. J. Hubbard, NASA SP-370, p. 153.

Trombka, J. I., et al., 1996, submitted to JGR.

Tucker, W. H ., and Koren, M. , 1971, ApJ., 168, 283.

Turkevich, A. L., Franzgrote, E. J., and Patterson, J. H., 1969, Science, 165, 277.

Vinogradov, A. P., Surkov, Y. A., Chernov, G. M., Kirnozov, F. F., and Nazarkina, G. B., 1966, Geochemistry No. 8 (Moscow: Vernadsky Institute of Geochemistry and Analytical Chemistry, 1966), p. 891.

Vinogradov, A. P., Surkov, Y. A., Chernov, G. M., Kirnozov, F. F., and Nazarkina, G. B., 1967, Cosmic Res. USSR, 5, 741.

Woolum, D. S., Burnett, D. S., Furst, M. and Weiss, J. R., 1975, The Moon, 12, 231.

Yin, L. I., Trombka, J. I., Bielefeld, M., 1993, "X-Ray Remote Sensing Techniques for Geochemical Analysis of Planetary Surfaces", in Remote Geochemical Analysis: Elemental and Mineralogical Composition, ed. C. Pieters and P. Englert, Cambridge Univeristy Press, 199-212.

CHAPTER 4

SOLAR OBSERVATIONS

4.1 Introduction

Gamma-ray astronomy as applied to observations of the Sun has the specific objective to investigate high-energy processes that take place in the Sun's atmosphere, and the relationship of these phenomena to the basic problems of solar activity. Gamma-ray emission exhibits characteristics of the conditions in regions where accelerated high-energy particles interact. A number of gamma-ray production mechanisms, both particle field and particle matter, are of interest in such astrophysical considerations. The types of particle-field interactions which are of specific interest are the Compton effect, magnetobremsstrahlung, and photomeson production yielding $\pi^o \rightarrow 2\gamma$. Examples of particle-matter interactions of interest are as follows: bremsstrahlung; production by electron-ion interaction; π^o meson production by proton-proton interaction or by proton-antiproton annihilation; nuclear deexcitation; fission; electron-positron annihilation; and neutral or charged-particle radiative capture or inelastic scatter. In the case of the Sun, the original gamma rays at the source can have their energy modified by Compton scattering and/or Doppler broadening. A calculated composite solar-flare spectrum (Dennis, 1997) is shown in Figure 4.1. The energy domain considered extends from 1 keV to 100 MeV for a flare temperature corresponding to 2×10^7 K. The contribution of the various major interaction processes are indicated. The region of interest in this chapter is the spectrum above 300 keV. Background problems significantly limit the gamma-ray flux that can be observed at energies greater than 100 keV; in fact, with presently available gamma-ray-detection systems solar gamma-ray emission can only be detected during periods of solar flares. The nature of the gamma-ray background problems are considered in later chapters of this book.

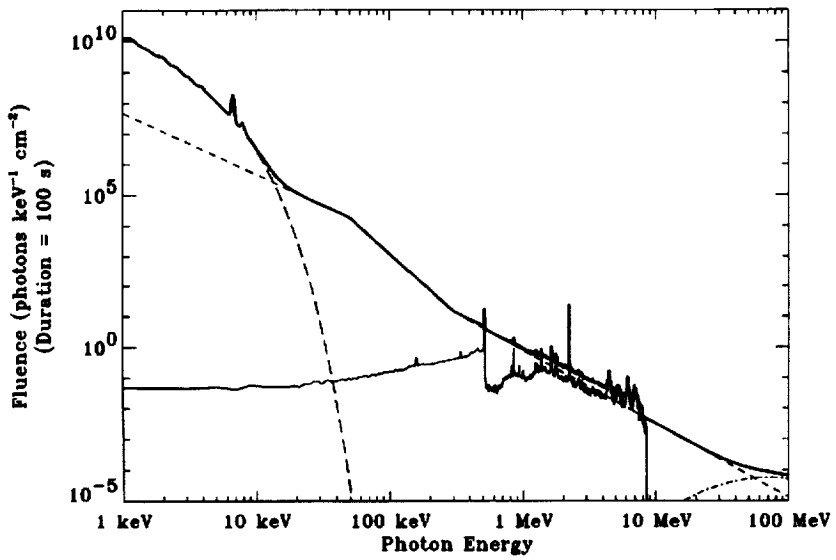

Figure 4.1. Composite solar-flare spectrum from 1 keV to 10 MeV (Dennis, 1997).

Solar flares allow us to study astrophysical energy release and particle-acceleration processes in great detail. Measurements of a combination of solar-flare charged-particle, gamma-ray and neutron spectra and fluxes yield important information concerning these processes. Details of both solar observations and theoretical models can be found in Ramaty et al. (1996). The results can be briefly summarized as follows: (1) charged-particle and gamma-ray observations have shown that stochastic acceleration due to gyroresonant interactions with the plasma waves constitute an important mechanism in solar flares; (2) coronal-mass-ejection (CME) driven shocks accelerate the bulk of the particles observed in space and such events are often associated with gradual flares; (3) accelerated MeV/nuclear ions may contain a large fraction of the energy released in flares, similar, in fact, to that contained in accelerated electrons; (4) GeV ions are either trapped at the Sun or accelerated for hours after some very large flares; (5) nuclear gamma-ray lines observations can be used to infer elemental abundances in the solar atmosphere; (6) observations of behind-the-limb solar flares have shown that gamma rays can be produced at high altitudes or at large distances from the original flare site.

4.2 Interaction Processes

Various interaction processes occur during solar flares. High-energy particle interaction, producing solar gamma-ray emission, continuum emission, π^0 mesons and, possibly, positronium, will be of major interest in terms of gamma-ray observations. These processes are described in Sections 4.3 – 4.6 and in later chapters of this book.

4.3 Charged- and Neutral-Particle Interactions

Solar gamma-ray emission can be stimulated by nuclear reactions between particles energized in solar flares and the abundant constituents of the ambient medium. The energetic particles include both primary charged particles and secondary particles which may be charged or neutral. Of particular interest for discrete gamma-ray line production are those interactions involving protons, α particles, carbon, oxygen nuclei, ions of all abundant elements in the solar atmosphere, and neutrons. Capture, inelastic scattering, and spallation reaction are most important. Neutron prompt capture (n,γ) and inelastic scatter $(n,n'\gamma)$ were considered in Chapter 3 for the case of remote elemental analysis of planetary surfaces.

Most of the interactions involve accelerated protons hitting ions at the ambient temperature of the interaction region, but significant numbers of heavier ions of all elements in the solar corona are also accelerated and they also hit the thermal ions. In the latter case, significant Doppler shifts can be expected, as the ion in the excited state is moving at high speed with respect to the observer. Such Doppler shifts can yield important information about the direction of the accelerated ions.

There are a number of prompt, discrete gamma-ray lines produced during solar flares that should be observable by detector systems. The lines are primarily attributable to proton and α particle interactions. The inelastic interactions of interest are

$$ {}^{A}_{Z}Y + {}^{1}_{1}p \rightarrow {}^{A}_{Z}Y^{*} + {}^{1}_{1}p \qquad (4\text{-}1) $$

and

$$ {}^{A}_{Z}Y^{*} \rightarrow {}^{A}_{Z}Y + \gamma\text{-ray}, \tag{4-1a} $$

that is the excited nucleus decays to a ground state (g.s.).

The nucleus at excited level (G) can also decay to a second excited state (K) emitting a gamma-ray of energy (G-K). That is,

$$ {}^{A}_{Z}Y^{*(G)} \rightarrow {}^{A}_{Z}Y^{*(K)} + \gamma\text{-ray (G - K)}, \tag{4-1b} $$

where Z is the atomic number, A is the atomic weight, Y is the nucleon, and p is a proton. The excited state can decay to a lower excited state and then further decay by gamma-ray emission to a ground state. Multiple decay may be required before the final ground state is reached. The number of levels excited depends on the interaction cross section and energy of the particle.
Competing modes of decay are also possible. The proton interaction indicated in Equation (4-1) will be written as (p, p'). In a similar way, inelastic scattering α-particle interactions will be noted as (α, α'). Other interaction processes such as (α, n), (α, p) and (p, n) will produce prompt gamma-ray emission. For example, the (α, n) reaction can be written as follows:

$$ {}^{A}_{Z}Y + {}^{4}_{2}\alpha \rightarrow {}^{A+3}_{Z+2}X + {}^{1}_{0}n. \tag{4-2} $$

In the interactions considered above the excited levels in nuclei were populated by direct excitation reactions. Another important reaction method is spallation. In this reaction protons or α-particles break up heavier nuclei into lighter fragments that emerge from the reaction in excited states.

Examples of such interaction are (p, 2p), (p, 2pn), (p, pn), (p, 3pn), and (p, pα). The reaction (p, 3pn) can be written, for example,

$$ {}^{A}_{Z}Y + {}^{1}_{1}p \rightarrow {}^{A-3}_{Z-2}X + 3\,{}^{1}_{1}p + {}^{1}_{0}n \qquad (4\text{-}3) $$

Rather complex spallation reactions will be denoted by (p,—). Table 4.1 lists some of the more prominent lines produced by such prompt gamma-ray interaction (Ramaty, Kozlovsky and Lingenfelter, 1975). The relative intensity of the gamma-ray lines for various solar-flare conditions has been calculated and the results can be found in Ramaty, Kozlovsky and Lingenfelter, (1975).

It will be noted that only lines of gamma rays with energies less than 8 MeV have been indicated in Table 4.1. There are two principal methods by which gamma rays can be produced with energies of 8 MeV or greater. First, in neutron capture (n, γ), gamma-ray energies corresponding to the binding energy per nucleon are obtained. Some of these interactions are shown in Table 4.2. Because of the low concentration of the heavier elements in the solar atmosphere in regions where some solar flares are produced, these types of interactions should not produce significant fluxes of gamma rays with energies above 8 MeV regions. Second, high-energy states can also be excited by the interaction of particles with sufficient energy to excite these levels.

Most of the higher-energy states are particle unstable and lead to little or no gamma-ray emission. Such interactions have been studied (Crannell, Crannell, and Ramaty, 1979). These higher-energy lines have not been deteced as yet in any flares, but may be of interest in future missions when detectors with sufficient sensitivity to this high energy may be available. The strength of gamma-ray line emission is proportional to the product of three factors: interaction cross section, branching ratio, and nuclear abundance. Gamma-ray-emission lines of possible interest with energies greater than 8 MeV are indicated in Table 4.2. Excitation due to both capture and high-energy interactions are included.

Table 4-1
Prompt γ-Ray Lines of Interest in Solar Flares

Energy MeV	Interaction Process	Decay Mode	Half (where applicable)
0.431	$^{4}He\ (\alpha, n)^{7}Be^{*}$	$^{7}Be^{*0.431} \rightarrow {}^{7}Be + \gamma\text{-Ray}\ (0.431\ MeV)$ to g.s.	
0.478	$^{4}He\ (\alpha, p)^{7}Li^{*}$	$^{7}Li^{*0.478} \rightarrow {}^{7}Li + \gamma\text{-Ray}\ (0.478)$ to g.s.	
	$^{4}He^{7}\ (\alpha, n)^{7}Be^{*}$	$^{7}Be \xrightarrow{12\%} ECT\ {}^{7}Li^{*0.478}$ $Li^{*0.478} \rightarrow {}^{7}Li + \gamma\text{-Ray}\ (0.478)$ to g.s.	53.3d. (^{7}Be)
0.72	$^{12}C\ (p, 2pn)^{10}B^{*}$	$^{10}B^{*0.72} \rightarrow {}^{10}B + \gamma\text{-Ray}\ (0.72)$ to g.s.	
	$^{16}O\ (p, -)^{10}B^{*}$	$^{10}B^{*0.72} \rightarrow {}^{10}B + \gamma\text{-Ray}\ (0.72)$ to g.s.	
0.845	$^{56}Fe\ (p, p')^{56}Fe^{*}$	$^{56}Fe^{*0.845} \rightarrow {}^{56}Fe + \gamma\text{-Ray}\ (0.845)$ to g.s.	
1.24	$^{56}Fe\ (p, p')^{56}Fe^{*}$	$^{56}Fe^{*2.08} \rightarrow {}^{56}Fe^{*0.845} + \gamma\text{-Ray}\ (1.24)$	
1.387	$^{24}Mg\ (p, p')^{24}Mg^{*1.37}$	$^{24}Mg^{*1.37} \rightarrow {}^{24}Mg + \gamma\text{-Ray}\ (1.37)$ to g.s.	
1.63	$^{20}Ne\ (p, p')^{20}Ne^{*}$	$^{20}Ne^{*1.63} \rightarrow {}^{20}Ne + \gamma\text{-Ray}\ (1.63)$ to g.s.	
	$^{14}N\ (p, p')^{14}N^{*}$ $^{16}O\ (p, 2pn)^{14}N^{*}$	$^{14}N^{*3.94} \rightarrow {}^{14}N^{*2.31} + \gamma\text{-Ray}\ (1.63)$	
	$^{24}Mg\ (p, 2p)^{23}Na^{*}$	$^{23}Na^{*2.07} \rightarrow {}^{23}Na^{*0.44} + \gamma\text{-Ray}\ (1.63)$	
1.78	$^{28}Si\ (p, p')^{28}Si^{*}$	$^{28}Si^{*1.78} \rightarrow {}^{28}Si + \gamma\text{-Ray}\ (1.78)$ to g.s.	
1.99	$^{12}C\ (n, pn)^{11}C^{*}$	$^{11}C^{*1.99} \rightarrow {}^{11}C + \gamma\text{-Ray}\ (1.99)$ to g.s.	
2.31	$^{14}N\ (p, p')^{14}N^{*}$	$^{14}N^{*3.94} \rightarrow {}^{14}N^{*2.31} + \gamma\text{-Ray}\ (1.63) \rightarrow {}^{14}N + \gamma\text{-Ray}\ (2.31)$ to g.s.	
	$^{14}N\ (p, n)^{14}O$	$^{14}O \rightarrow {}^{0}_{+1}e + {}^{14}N^{*2.31}, N^{*2.31} \rightarrow {}^{14}N + \gamma\text{-Ray}\ (2.31)$ to g.s.	70.59s (^{14}O)

Source: Ramaty et al. (1975).

Table 4-1 (continued)
Prompt γ-Ray Lines of Interest in Solar Flares

Energy MeV	Interaction Process	Decay Mode	Half (where applicable)
	$^{16}O\ (p, 2pn)^{14}N^*$	$^{14}N^{*2.31} \rightarrow {}^{14}N + \gamma$-Ray (2.31) to g.s.	
2.75	$^{16}O\ (p, p')^{16}O^*$	$^{16}O^{*8.88} \rightarrow O^{*6.14} + \gamma$-Ray (2.75)	
~3.62	$^{16}O\ (p, 3pn)^{13}C^*$	$^{13}C^{*3.68} \rightarrow {}^{13}C + \gamma$-Ray (3.84) to g.s.	
	$^{12}C\ (p, -)^{6}Li^*$	$^{6}Li^{*3.56} \rightarrow {}^{6}Li + \gamma$-Ray (3.56) to g.s.	
3.84	$^{16}O(p, 3pn)^{13}C^*$	$^{13}C^{*3.84} \rightarrow {}^{13}C + \gamma$-Ray (3.84) to g.s.	
4.43	$^{12}C\ (p, p')^{12}C^*$		
	$^{12}C\ (\alpha, \alpha')^{12}C^*$	$^{12}C^{*4.43} \rightarrow {}^{12}C + \gamma$-Ray (4.43) to g.s.	
	$^{16}O\ (p, p\alpha)^{12}C^*$		
~5.3	$^{16}O\ (p, pn)^{15}O^*$	$^{15}O^{*5.26} \rightarrow {}^{15}O + \gamma$-Ray (5.26) to g.s.	
	$^{16}O\ (p, 2p)^{15}N^*$	$^{15}N^{*5.28} \rightarrow {}^{15}N + \gamma$-Ray (5.28) to g.s.	
	$^{16}O\ (p, 2p)^{15}N^{*5.31}$	$^{15}N^{*5.31} \rightarrow {}^{15}N + \gamma$-Ray (5.31) to g.s.	
6.14	160 (p, p,)160*	$^{16}O^{*6.14} \rightarrow {}^{16}O + \gamma$-Ray (6.14) to g.s.	
	160 (q, q,)160*		
6.33	$^{16}O\ (p, 2p)^{15}N^{*6.33}$	$^{15}N^{*6.33} \rightarrow {}^{15}N + \gamma$-Ray to g.s.	

Table 4-1 (continued)
Prompt γ-Ray Lines of Interest in Solar Flares

Energy MeV	Interaction Process	Decay Mode	Half (where applicable)
~6.7	$^{12}C\ (p, 2p)^{11}B^*$	$^{11}B^{*6.76} \rightarrow {}^{11}B + \gamma$-Ray (6.76) to g.s.	
	$^{12}C\ (p, 2p)^{11}B^*$	$^{11}B^{*6.81} \rightarrow {}^{11}B + \gamma$-Ray (6.81) to g.s.	
	$^{12}C\ (p, pn)^{11}C^*$	$^{11}C^{*6.50} \rightarrow {}^{11}C \rightarrow \gamma$-Ray (6.50)	
7.12	$^{16}O\ (p, p')^{16}O^*$	$^{16}O^{*7.12} \rightarrow {}^{16}O + \gamma$-Ray (7.12) to g.s.	

Note: g.s. = ground state, e.s. = excited state.

Table 4-2

Greater than 8 MeV γ-Ray Transitions of Interest in the Study of Solar Flares

Excited Nuclide	Gamma Ray Energy in MeV	Interaction Process	Transition
^{11}B	8.56	^{12}C $(p, 2p)$	g.s.
^{12}C	10.67	^{12}C (p, p') and ^{16}O (p, p')	e.s.
	12.71		g.s.
	15.11		g.s.
^{15}N	8.31	^{16}O $(p, 2p)$	g.s.
	9.05		
	9.15		
	9.76		
	9.93		
	10.07		
^{20}Ne	8.32	^{20}Ne (p, p')	e.s.
^{27}Al	8.05	^{28}Si $(p, 2p)$	g.s.
^{28}Si	8.33	^{28}Si (p, p')	g.s.
	8.43		e.s.
	8.49		g.s.
	8.53		e.s.
	8.76		e.s.
	8.89		e.s.
	8.90		g.s.
	9.17		e.s.
	9.48		g.s.
	9.50		g.s.
	10.31		g.s.
	10.59		g.s.
	10.72		g.s.
	10.90		g.s.
	11.45		g.s.
^{32}S	8.13	^{32}S (p, p')	g.s.

Source: Crannell et al. (1979).

Table 4-2 (continued)
Greater than 8 MeV γ-Ray Transitions of Interest in the Study of Solar Flares

Excited Nuclide	Gamma Ray Energy in MeV	Interaction Process	Transition
	8.59		e.s
	10.63		e.s.
	11.23		g.s.
^{23}Na	8.36	^{24}Mg (p, $2p$)	g.s.
	8.65		g.s.
	8.67		g.s.
^{24}Mg	8.44	^{24}Mg (p, p')	g.s.
	8.69		e.s.
	9.00		g.s.
	9.15		g.s.
	9.36		e.s.
	9.83		e.s.
	9.84		e.s.
	9.97		g.s.
	10.73		g.s.
	8.29		g.s.

Note: g.s. = ground state, e.s. = excited state.

In the discussion above, primary interactions by protons and α-particles were considered. We shall now discuss secondary production of neutrons and interactions with nuclear species in the solar flare environment. Neutron production by accelerated charged particles has been studied in detail (Lingenfelter et al., 1965; and Lingenfelter and Ramaty, 1967). The most important neutron-producing reactions and their threshold energies are listed in Table 4.3 (Lingenfelter and Ramaty, 1967). Only the number of neutrons is indicated in some cases: the total interactions are not shown. The neutrons are born at high energies. Calculations of the distribution of neutrons released in the chromosphere or corona have yielded the following results (Lingenfelter and Ramaty, 1967). If the neutrons are released above the photosphere, any initially upward moving

neutrons escape from the Sun. Of course, unlike charged ions and electrons, neutrons are not constrained by the coronal magnetic field and hence they can escape much more easily. Some of the downward moving neutrons can also escape after being backscattered elastically by ambient protons, but most of these neutrons either are captured or decay at the Sun. Because the probability for elastic scattering is larger than the capture probability, the majority of neutrons are thermalized before they are captured. The bulk of the thermal neutrons at the Sun are captured either in ^{1}H

$$ {}^{1}_{0}n + {}^{1}_{1}H \rightarrow {}^{2}_{1}H + \gamma\text{-ray}, \tag{4-4} $$

or on ^{3}He

$$ {}^{1}_{0}n + {}^{3}_{21}H_{e} \rightarrow {}^{3}_{1}H + {}^{1}_{1}p. \tag{4-5} $$

The gamma-ray energy produced during the interaction indicated by Equation (4-4) is 2.223 MeV. The reaction indicated in Equation (4-5) is radiationless. Thus, the magnitude of gamma-ray emission at 2.223 MeV depends on the ratio $^{3}He/H$ in the interacting medium.

As we have indicated above, the positrons and neutrons produced during a solar flare are thermalized by the ambient medium before they are captured or annihilated. Therefore, any information concerning directional anisotropies in the primary particles would be lost before the gamma rays at 0.511 MeV and 2.223 MeV are produced. On the other hand, the ^{12}C and ^{12}O nuclei, which emit the gamma-ray lines at 4.438 MeV and 6.126 MeV, are excited by protons with energies of several tens of MeV, which also impart kinetic energy to these nuclei. Because the gamma rays are subsequently emitted in a short time interval, compared with the slowing down time of the nuclei, any directional anisotropy in the primary particles would cause a Doppler shift of the energies of the lines. It has been shown in Ramaty and Crannell (1976) and Kozlovsky and Ramaty (1977) that observable Doppler shifts result from particle anisotropies in solar flares. These lines can also be thermally broadened, but the Doppler shift broadening, caused by particle anisotropies, would be expected to be an order of magnitude higher than that expected for the highest known temperature in the solar atmosphere.

Table 4-3
Neutron Production Modes During Solar Flares

Reaction	Threshold (MeV/nucleon)
1. $p + {}^{1}H \rightarrow n + p + {}^{+}$	292.30
2. $p + {}^{4}He \rightarrow {}^{3}He + p + n + (\pi)$	25.70
$\rightarrow {}^{2}H + 2p + n + (\pi)$	32.60
$\rightarrow 3p + 2n + (\pi)$	35.40
3. $p + {}^{13}C \rightarrow n + \ldots$	19.60
$p + {}^{13}C \rightarrow n + \ldots$	3.20
4. $p + {}^{14}N \rightarrow n + \ldots$	6.30
5. $p + {}^{16}O \rightarrow n + \ldots$	16.60
$p + {}^{18}O \rightarrow n + \ldots$	2.50
6. $p + {}^{20}Ne \rightarrow n + \ldots$	15.90
7. $p + {}^{56}Fe \rightarrow n + \ldots$	5.50
8. $\alpha + {}^{4}He \rightarrow {}^{7}Be + n$	9.50
9. $\alpha + {}^{12}C \rightarrow n + \ldots$	2.80
$\alpha + {}^{13}C \rightarrow n + \ldots$	-
10. $\alpha + {}^{14}N \rightarrow n + \ldots$	1.50
11. $\alpha + {}^{16}O \rightarrow n + \ldots$	3.80
$\alpha + {}^{18}O \rightarrow n + \ldots$	0.21
12. $\alpha + {}^{20}Ne \rightarrow n + \ldots$	2.16
$\alpha + {}^{22}Ne \rightarrow n + \ldots$	0.15
13. $\alpha + {}^{56}Fe \rightarrow n + \ldots$	1.37
14. $\alpha + {}^{25}Mg \rightarrow n + \ldots$	-
$\alpha + {}^{26}Mg \rightarrow n + \ldots$	-
15. $\alpha + {}^{29}Si \rightarrow n + \ldots$	0.43

Note: (-) is exoergic

Source: Ramaty et al. (1975).

Detailed calculations of these effects caused by anisotropic acceleration of particles in solar flares can be found in Murphy, Kozlovsky and Ramary (1988). Line broadening and shifts are considered for the case of accelerated carbon and heavier nuclei interacting with the ambient H and He. The 0.429 MeV line from ^{17}Be, the 0.478 MeV from ^{17}Li, and the 4.438 MeV from ^{12}C are significant because their profiles are sensitive to the angular distribution. The following results were obtained. The 0.429 MeV and the 0.4178 MeV lines profiles for isotropic, perfect fan beam, and downward-beam distributions were found to be quite different. The mere presence of two distinct lines would immediately rule out isotropic distributions. When viewed from the limb, the downward-beam distribution produces line profiles very different from other distributions. It was also found that line profiles resulting from isotropic and perfect fan-beam distributions are different from each other, but the differences are small and may be hard to discriminate. High-energy resolution detector systems will be required to discern the differences. It is believed that the line splitting of the 4.438 MeV from ^{12}C is sufficient to distinguish an isotropic distribution from either a perfect fan beam observed at the center of the disk or a downward-directed beam viewed at the solar limb.

4.4 Continuum Emission

X-ray and gamma-ray continuum emission can be produced by three major interaction methods or processes: bremsstrahlung, Compton scattering, and synchrotron radiation. In terms of gamma-ray emission during solar flares, the latter two mechanisms are not important. However, they are important for other gamma-ray astrophysical processes and will be considered in detail in other chapters of this book.

Bremsstrahlung processes will also be considered in detail in other chapters of this book. In the case of a solar flare in the nonrelativistic domain, most of the gamma-ray emission is produced by the interaction of accelerated electrons with ambient ions. The inverse process, in which accelerated protons interact with ambient electrons, can produce bremsstrahlung, and the interaction is the electron-proton process, when the accelerated protons have the same velocity as the accelerated electrons. Bremsstrahlung can also be produced in electron-electron interactions. This process is negligible in the nonrelativistic domain, but becomes comparable to the electron-proton bremsstrahlung in the relativistic domain. Detailed

calculations of bremsstrahlung produced during solar flares have been carried out by many authors (e.g., Ramaty, Kozlovsky, and Suri, 1977; and Ramaty et al. (1993).

The solar gamma-ray continuum spectrum in the energy domain from 0.3 to 1 MeV can be most likely attributed to electron bremsstrahlung produced by electrons with energies greater than 1 MeV. See Figure 4.1. Both thin and thick target models have been used to describe the interaction processes. If one considers the thin target model, then electrons escape from the gamma-ray-emitting region in a time that is short when compared to their collisional lifetime. In this case, the spectra of the injected particles and of the escaping particles are the same. In the case of the thick target, the gamma-ray emission is produced by electrons which lose all of their energy through collisions in the interaction regions. Thus, the shape of the gamma-ray continuum is determined by the balance between the injection of newly accelerated electrons and the loss of electrons through collisions. Because of these differences, thin and thick target processes will produce different types of gamma-ray-emission spectra (for example, see Ramaty et al., 1993; Dröge, 1996, and Cliver, 1996).

4.5 Decay of π° Mesons

The π° mesons are produced by such interactions of high-energy particles as indicated in Equations 4.6 and 4-7:

$$p + p \rightarrow n\pi^{\circ}, \tag{4-6}$$

and

$$p + \alpha \rightarrow n\pi^\circ + ...,\tag{4-7}$$

where n is the multiplicity which may become significant during solar flares. The decay of the π° meson leads to gamma-ray emission, that is

$$\pi^\circ \rightarrow \gamma + \gamma. \tag{4-8}$$

The gamma-ray spectrum produced will be a broad continuum with a peak around 100 MeV. These gamma rays should dominate the very-high-energy portion of the spectrum. No discrete lines are

visible; thus, even detectors with poor energy resolution should be capable of detecting these emissions, though rather intense flares will be required for such observations. Details of the expected emission have been given in Ramaty, Kozlovsky, and Lingenfelter (1975) for the case of solar flares. Details of π° production will be considered later in this book.

4.6 Positron Annihilation and Formation of Positronium

The production of positrons during solar flares is believed to be caused by the decay of π^{+} mesons and that of radioactive nuclei, produced in nuclear reactions of flare-accelerated particles with constituents of the solar atmosphere (Lingenfelter and Ramaty, 1967; and Ramaty, Kozlovsky, and Lingenfelter, 1975). Table 4.4 lists the principal positron emitters, their half-lives, and maximum positron energies (Ramaty, Kozlovsky, and Lingenfelter, 1975).

As can be seen from Table 4.4, nuclear species produced during solar flares decay to produce positrons ranging in energy from several hundred keV to several hundred MeV.

Relativistic positrons can annihilate in flight (~10 percent, e.g., Wang and Ramaty, 1975), escape from the Sun, or decelerate to thermal energies through interactions with ambient matter and magnetic fields. Those relativistic positrons which annihilate in flight produce 0.511 MeV line emission. Because of Doppler broadening, these annihilation lines are not observed at 0.511 MeV, but appear as a gamma-ray continuum.

Those positrons that do not escape from the Sun or annihilate in flight can lose energy by collision and ionization and thus be thermalized. The thermal positrons can then either annihilate freely or form positronium. Two 0.511 MeV gamma rays are produced per free annihilation. Because the annihilation occurs at thermal energies, there is very little Doppler broadening, and these emissions can be observed as discrete gamma-ray lines.

Table 4.4

Principal Positron Emitters

Nuclear Species	Half-Life	Maximum Reactive Positron Energy (MeV)	Production Mode	Threshold (MeV)
μ	1.5×10^{-6}s	53	$p+^{1}H \rightarrow \pi^{+}$...	292.3
			$p+^{4}He \rightarrow \pi^{+}$...	
			$\pi^{+} \rightarrow \mu^{+}+\upsilon$	
^{10}C	19s	1.9	$\rho+^{16}O \rightarrow ^{10}C^{+}$..	41.4
			$\rho+^{14}N \rightarrow ^{10}C^{+}$..	17.1
			$\rho+^{12}C \rightarrow ^{10}C^{+}$..	34.4
^{11}C	20.5m	0.92	$\rho+^{16}O \rightarrow ^{11}C^{+}$..	27.5
			$\rho+^{14}N \rightarrow ^{11}C^{+}$..	3.1
			$\rho+^{12}C \rightarrow ^{11}C^{+}$..	17.9
^{12}N	0.011s	16.4	$\rho+^{12}C \rightarrow ^{11}N+n$	19.6
	10m	1.19	$\rho+^{16}O \rightarrow ^{13}N^{+}$..	5.5
			$\rho+^{14}N \rightarrow ^{13}N^{+}$..	9.0
		1.8 (99.4%)	$\rho+^{12}O \rightarrow ^{14}O^{+}$..	30.7
			$\rho+^{14}N \rightarrow ^{14}O+n$	6.3
			$\rho+^{16}O \rightarrow ^{15}O+$..	14.3
			$\alpha+^{12}C \rightarrow ^{15}O+n$	2.8
			$\alpha+^{16}O \rightarrow ^{19}Ne+n$	3.75

Note: m= minutes, s= seconds
Source: Ramaty, Kozlovsky, and Lingenfelter (1975)

The broadening of the line at 0.511 MeV, and the ratio of the line intensity produced at 0.511 MeV to the continuum, can be used to measure critical parameters relative to solar-flare dynamics. The instruments flown up to this time do not have the energy resolution required to observe the broadening. Large-area high-resolution Ge

detectors have the required energy resolution needed and could be used on future space flights to observe this line broadening.

4.7 Experimental Results

Gamma-ray emission studies have been carried out on spacecraft that have been in Earth orbit. During quiet Sun periods, the gamma-ray flux is too small to be observed by the detector systems flown in Earth orbit thus far. Gamma-ray flux observations have been reported during solar flares and for ≤ 10 hours after the few biggest flares. The gamma-ray observation in 1972 by the seventh Orbiting Solar Observatory (OS0-7) was the most definitive measurement made up to that time (Chupp, Forrest, and Suri, 1975). A number of possible solar gamma-ray events were reported earlier, but no detailed spectral and temporal information was obtained (see, for example, Kondo and Nagase, 1969; Hirasima, Okudaira, and Yamagami, 1969; and Koga, Simnett, and White, 1974). A gamma-ray spectrometer aboard the Solar Maximum Mission (SMM), launched on February 14, 1980, monitored the Sun and obtained detailed observations of solar gamma-ray emission. The SMM gamma-ray detector continued its operation through 1989 (Rieger, 1989; and Chupp, 1990). Important observations were carried out during both the declining portion of solar cycle 21 (1980-1984) and the rising portion of cycle 22 (1988-1989). During cycle 21, gamma-ray observations were also carried out using spectrometers aboard HINOTORI (Yoshimori, 1990). Finally, a most active period of solar gamma-ray observations occurred during the peak of solar cycle 22. Gamma-ray spectral measurements were made during solar flares with the following instruments: the OSSE (Murphy et al., 1993a), COMPTEL (Ryan et al., 1993), and EGRET (Schneid, 1994; and Ramaty et al., 1993) instruments on CGRO; the Phebus instrument on the GRANAT spacecraft (Barat et al., 1994); and the gamma-ray spectrometer on the YOHKOH spacecraft (Yoshimori et al., 1994).

4.8 Solar-Flare Observation in the Energy Region 0.1 to 10 MEV

In order to interpret the gamma-ray spectral observations obtained during the cycle-21 and -22 solar flares, it is important to first consider the theoretical calculation of expected gamma-ray solar emission. The spectrum shown in Figure 4.2 was calculated,

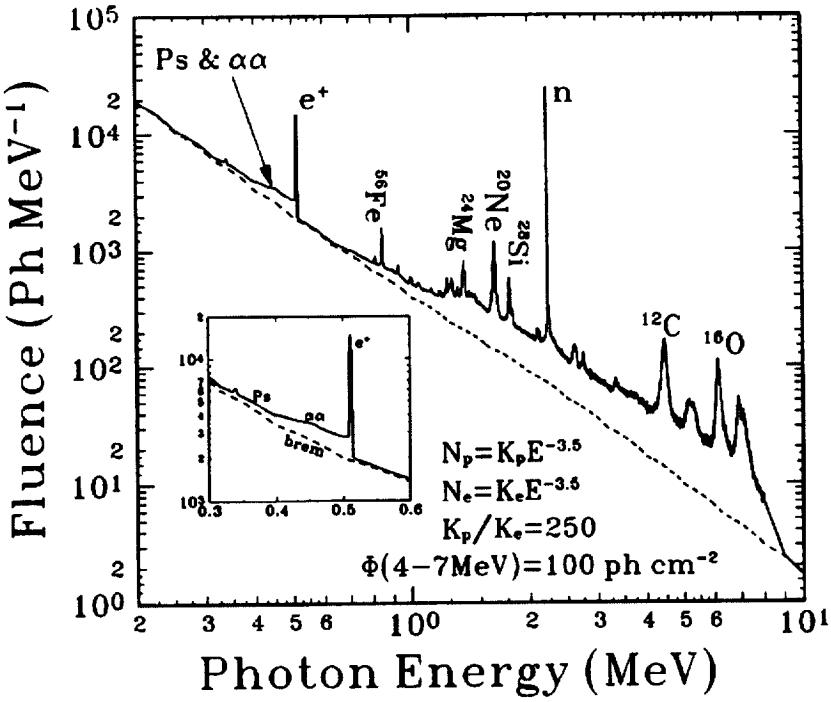

Figure 4.2. Theoretical calculation of a solar flare gamma-ray spectrum (Ramaty and Lingenfelter, 1995).

assuming gamma-ray emission is produced through the interaction of flare-accelerated ions and electrons with power-law spectra with an index of −3.5.

Furthermore, the spectrum was normalized to the flux observed in the 27 April, 1981, solar flare in the energy domain 4–7 MeV (Ramaty and Lingenfelter, 1994, and 1995).

The various mechanisms for gamma-ray emission were described in Sections 4.3 – 4.6, and in later chapters of this book. The following mechanisms and processes are considered here: electron-ion bremsstrahlung continuum up to ~1 MeV, nuclear line emission up to ~8 MeV, electron-ion bremsstrahlung above 8 MeV, and π^0 decay centered at 70 MeV. The strongest line at 2.223 MeV is caused by neutron capture on hydrogen; and the inelastic scatter

lines are the 6.129 MeV from ^{16}O, the 4.438 MeV from ^{12}C, 1.779 MeV from ^{28}Si, 1.634 MeV from ^{20}Ne, 1,369 MeV from ^{24}Na, and .847 MeV from ^{56}Fe. The dashed line represents the brehmsstrahlung contribution and the excess above the brehmsstrahlung can be attributed to broad lines and a number of unresolved lines. The line at 0.511 MeV is due to positron annihilation and the continuum under the line can be attributed to positronium annihilation. The assumption is made that there is an isotropic distribution of interacting particles and thus the ^{7}Li and ^{7}Be at 0.478 MeV and 0.429 MeV blend into a single feature (Koslovsky and Ramaty, 1974). It has also been shown that under certain conditions of anisotropic particle distributions, two individual lines will be seen (Koslovsky and Ramaty, 1977).

With this model in mind, the experimental results are now considered. It is believed that the neutron capture line 2.223 MeV on ^{1}H interaction takes place at a greater depth in the photosphere than the prompt inelastic scattering interactions. Thus the ratio of the fluence of the 2.223 MeV line to the prompt emission of the other deexcitation lines will depend on the position of the flare on the solar disk (Wang and Ramaty, 1974) and Hua and Lingenfelter 1987). The ratio becomes very small for flares near or at the limb, and can vanish for flares behind the limb. Gamma-ray observations of a flare on 1 June, 1991, occurring 10° behind the limb, showed the deexcitation lines while the 2.223 MeV line was absent (Barat et al., 1994). An exception to this prediction was found when the 2.223 MeV line was observed for a solar flare on 29 September, 1989, occurring $10^{\circ}+5^{\circ}$ behind the limb (Vestrand and Forrest, 1993). This very large flare produced a number of emissions, including protons up to 25 GeV (Swinson and Shea, 1990). The observation of the 2.223 MeV line implies that this line must have been produced by charged particles interacting on the visible hemisphere of the Sun. Cliver, Kahler, and Vestrand (1993) have suggested that these particles were accelerated by a coronal shock over a large volume, thereby producing an extended gamma-ray-emitting region visible from the Earth, even if the optical flare occurred behind the limb.

The use of solar gamma-ray observations to study the nature of solar-particle impulsive events is now considered. Interplanetary measurements of the electron to proton ratio (e/p) have been used to

distinguish two classes of solar particle events (Cane, McGuire and von Rosenvige, 1986; and Reames, 1990). There are impulsive events of relatively short duration and associated with soft x-ray emission that have large e/p ratios, and there are longer-lasting, gradual events, for which the x-ray emission lasts longer and the e/p ratio is smaller. Ramaty et al. (1993) used the ratio of the fluence of the bremsstrahlung continuum to that in the lines to find that the e/p ratio was comparable or higher than the values measured in interplanetary space. This held true for both the shorter impulsive and the gradual events. From these observations it has been suggested that the solar trapped particles which produce gamma rays, and the particles produced by impulsive events and observed in interplanetary space are accelerated by the same mechanisms. The mechanisms are most probably stochastic acceleration due to gyroresonant interactions with the plasma turbulence. Miller and Steinacker (1992) have shown that relativistic electron acceleration by such turbulence, particularly whistler waves, can be quite efficient. The gradual events are thought to be produced by accelerated particles from cooler, coronal gas, most likely a shock which does not accelerate the electrons efficiently.

The solar-flare gamma-ray spectrum obtained using the SMM scintillation spectrometer (NaI crystal detector) is now considered and shown in Figure 4.3 (Murphy et al., 1991). The following lines can be seen: the 6.126 MeV line from ^{16}O, the 4.438 MeV from ^{12}C, the 1.634 line from ^{20}Ne. Furthermore, the annihilation line possibly on top of the positronium continuum with the α-α lines may produce a broad feature above the strong brehmsstrahlung continuum. The position of the lines for ^{24}Mg and ^{54}Fe are shown, but are both weak and not resolved.

Some information concerning the composition of the ambient gas can be inferred from these data (e.g., Share and Murphy, 1995; Ramaty et al., 1995; and Ramaty et al., 1996). It has been shown that the Mg, Si, and Fe abundances relative to C and O are consistent with coronal abundances, but enhanced in comparison with the photospheric abundances. It has been suggested that the enhanced Mg, Si and Fe abundances could be attributed to a charge-dependent transport process from the photosphere to the chromosphere and corona, which favors collisionally ionized, low

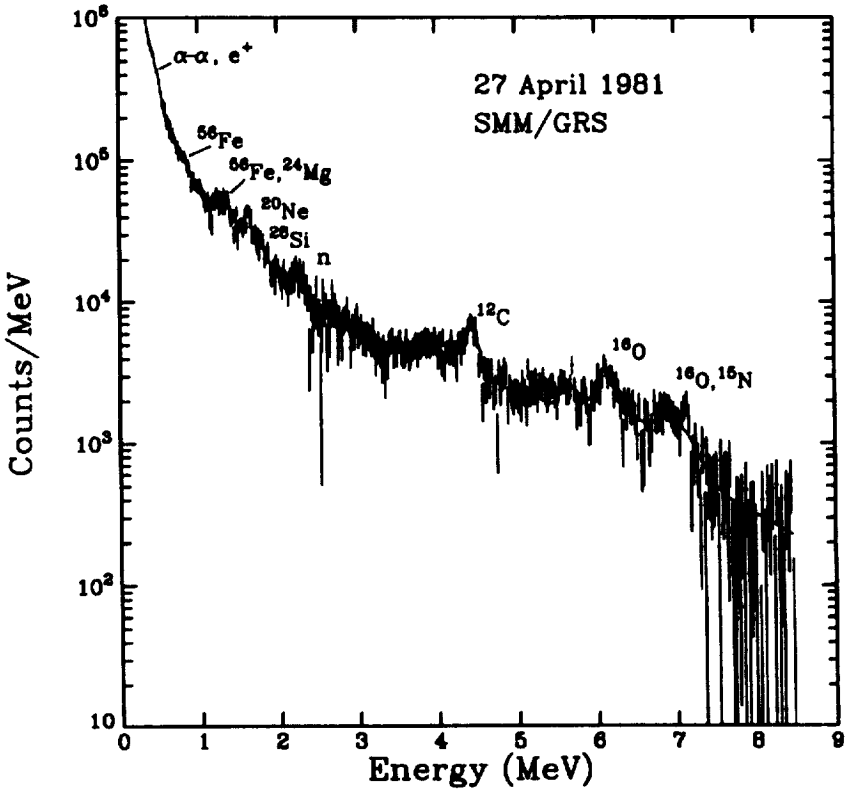

Figure 4.3. Observed solar-flare gamma-ray pulse height spectrum, measured using the SMM scintillation spectrometer. A theoretical fit to the data is also shown (Murphy et al., 1991).

first ionization potential elements, such as Mg, Si, and Fe in the photosphere (Meyer, 1989). A rather unexpected result from these solar gamma-ray observations is that the Ne abundance is higher than expected (Murphy et al., 1991). The Ne/O ratio derived from these observations is a factor of two higher than the coronal ratio. The coronal ratio seems to be consistent with the local Galactic Ne abundance (Meyer, 1989). The reason for this discrepancy is not understood, but two different interpretations have been suggested. First, the Ne abundance is representative of the photosphere, whose abundance is not well determined. A second interpretation suggests that the Ne enhancement could be caused by photoionization by soft x-rays (Shemi, 1991). Finally, observations of the time-dependent

flux of the 2.223 MeV line (Prince et al., 1983) were used to derive a photospheric ^{3}He/H ratio of ~2 x 10^{-5} (Hua and Lingenfelter, 1987). This derivation depended on a suggestion by Wang and Ramaty (1974) that the ^{3}He in the photoshere could capture as much as one half of the neutrons, and thus the flux-time profile of the 2.223 MeV line, resulting from the capture of the other half of 1H, would strongly depend on the photospheric abundance of ^{3}He. The abundance derived in this manner is consistent with the expected primordial nucleosynthesis ^{3}He abundance.

A most interesting set of gamma-ray observations on the solar flare of December 16, 1988, were carried out by the Gamma-Ray Spectrometer (GRS) on the SMM spacecraft. This was one of the most intense gamma-ray events ever recorded. There were several well-separated bursts during the period of the solar flare. Analyses were carried out using the temporal variation of the fluences in different energy bands. Using the analyses, the spectrum of the primary particles were inferred. These particle spectra changed from burst to burst, which indicated different particles acceleration mechanisms from burst to burst. A detailed analysis can be found in Rieger (1996). A further discussion will be found in the next section.

4.9 High-Energy Continuum Radiation

The solar activity observed on the Compton Gamma-Ray Observatory (CGRO) is of particular interest relative to the measurement of the high-energy (greater than 8 MeV) gamma-ray continuum, which permits the study of the relativistic solar particles.

Although the Sun's activity was low for most of the period during which the CGRO has been in space (it was launched in April, 1991), thus far (the end of 1996) five of the biggest flares ever recorded in a short time occurred in June, 1991. Some of the results of these observations have already been discussed, particularly the gamma-ray lines. As indicated in the previous section, continuum gamma-ray emission was observed for several hours following several intense solar flares. Furthermore, for several of the smaller flares, although the statistics were poorer, the subsequent continuum emission was found by Schneid et al. (1996) and Bertsch et al. (1996) to be very similar to that seen earlier by Forrest (1983),

Vestrand (1988), Murphy et al. (1990, 1993b), and Cliver, Crosby, and Dennis (1994). In particular, the data were in agreement with a contribution from lines and a continuum that was a power law in energy, consistent with bremsstrahlung, with a spectral index in the range of 1.8 to 2.5.

Of special interest are the observations by the EGRET telescope on the Compton Gamma-Ray Observatory of the high-energy gamma-ray emission observed on June 11, 1996, following the large X-class solar flare. Several important results were obtained. First, high-energy gamma-rays were seen with energies up to 2 GeV, indicating the presence of relativistic particles to even higher energies. Second, the gamma rays were observed for over ten hours following the unpulsed phase of the flare, indicating that either acceleration or storage of relativistic particles lasted for this period of time. Third, the spectrum in the 30 MeV to 2 GeV was not a simple power law in energy, but consistent with a combination of a bremsstrahlung power law and a nucleon-nucleon interaction spectrum. (The nucleon-nucleon interaction spectrum is often called a π^0 spectrum because of the dominance in the spectrum of the decay of that meson. See Section 5.3 for a further discussion of the nucleon-nucleon interaction gamma-ray spectrum.) This latter result showed clearly that both relativistic electrons and nucleons were present. Fourth, the combination of the bremsstrahlung and nucleon-nucleon interaction spectrum lasted throughout the ten hour period, as shown in Figure 4.4.

This last result is of special significance. The rate of loss of energy by electrons in the relevant energy range through bremsstrahlung is far greater than that of the energetic nucleons. If the relativistic particles were simply accelerated and trapped very near the Sun, the electrons would lose their energy relatively quickly and the bremsstrahlung portion of the gamma-ray spectrum might be expected to disappear more quickly than observed. As Figure 4.4 shows, this is not the case. Not only does the bremsstrahlung portion not disappear during the first hour, but the two components remain relatively equal over the ten hour period. These results imply that the ratio of the parent populations of electrons and nucleons has remained the same during the whole period.

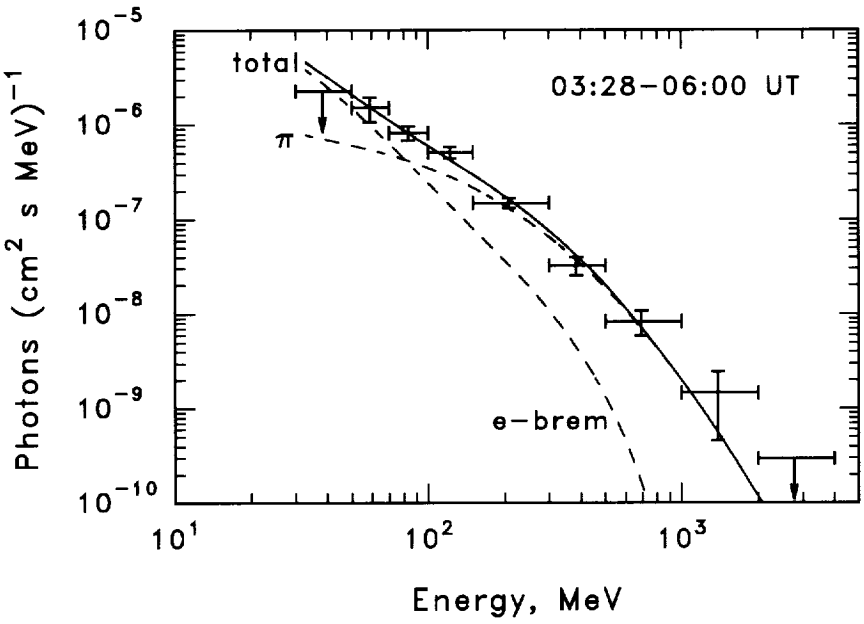

Figure 4.4. Photon differential energy spectrum for the 1991 June 11 flare event. Data were accumulated from 03:28 UT until 06:00 UT. The dotted curves are for solar proton and electron power-law spectra of indices −3.0, scaled so that their total, shown as a solid curve, is the best fit to the data (Bertsch et al., 1996).

To understand this conclusion and its significance, consider the question of the origin of the energetic solar particles. When high-energy particles were first observed in the vicinity of the Earth following a major solar flare, it was naturally assumed that they were somehow accelerated in association with the flare itself. However, flares, even the very large ones, typically only last for a fraction of an hour to at most a few hours; yet, solar particles are observed in space for days. Particles could be seen following flares from anywhere on the visible disk of the Sun or even behind it, even though the particles had to be substantially constrained by the field lines. The concept of trapping and diffusion across field lines was developed, but there were difficulties in detail. The alternate idea for part or most of the solar particle acceleration is that it occurs in association with a shock wave, driven by a coronal-mass ejection. In that case, particles are accelerated on a broad range of field lines as the shock propagates away from the Sun at velocities up to 2000

km/sec. While the importance of shock acceleration was recognized relatively early by Wild, Smerd, and Weiss (1963), it took time for further support of the concept to show its importance. The observations of the many coronal-mass ejections and the detailed properties of the energetic solar particles all suggested that there was a prolonged acceleration by the shock wave driven by the coronal-mass ejection. For detailed review of this subject, see, for example, Reames (1993), Gosling (1993), and Reames, Barbier, and Ng (1996).

Within the framework of the discussion of the last paragraph, it appears that the EGRET result for the June 11, 1996, solar event adds additional support to the theory that a shock wave driven by coronal-mass ejection is the primary energy source for the energetic solar particles. As noted earlier, if the flare was responsible for the acceleration and the particles were trapped near the flare, the bremsstrahlung portion of the high-energy gamma-ray spectrum would disappear quickly at the end of the flare, and it does not. If the acceleration is occurring continuously at the shock, then the electron and nucleon components would keep their relative abundances, at least approximately, and the observed result would be expected. Hence, the high-energy gamma-ray observation of the June 11, 1991 solar event adds to the evidence that the particle acceleration occurs in the shock driven by the coronal-mass ejection. However, this alternative is also not without difficulties in terms of the exact manner in which the nucleons and electrons would achieve their observed spectra; so the question must probably remain open for a while longer.

References

Barat, C., et al., 1994, ApJ, 425, L109.

Bertsch, D. L., et al., 1996, ApJ, submitted.

Cane, H. V., Mcguire, R. E., and Von Rosenvinge, T. T., 1986, ApJ, 310, 448.

Chupp, E. L., Forrest, D. J., and Suri, A. N., 1975, in Solar Gamma-, X-, and EUV Radiation, ed. S. R. Kane, IAU Symp. 68 (Dordrecht, Holland: D. Reidel Publishing Co., 1975), 341.

Chupp, E. L., 1990, Physica Scripta, T18, 15.

Cliver, E. W., 1996, High-Energy Solar Physics, ed. R. Ramaty, N. Mandzhavidze, and X. M. Hua, AIP Conf Proc., 374, 45-60.

Cliver, E. W., Crosby, N. B., and Dennis, B. R., 1994, High-Energy Solar Phenomena - A New Era of Spacecraft Measurements, ed. J. M. Ryan and W.T. Vestrand, AIP Conf. Proc. 294, 177.
Cliver, E. W., Kahler, S. W., and Vestrand,W., T., 1993, 23rd Intern. Cosmic Ray Conf. Papers, 3, 91.
Crannell, C. J., Joyce, G., Ramaty, R., and Werntz, C., 1976, ApJ, 210, 582.
Crannell, C. J., Crannell, H., and Ramaty, R., 1979, ApJ, 229, 762.
Dennis, B. R., 1996, private communication.
Dröge, W., 1996, High-Energy Solar Physics, ed. R. Ramaty, N. Mandzhavidze, and X. M. Hua, AIP Conf Proc., 374, 78-85.
Forrest, D. J., 1983, Positron-Electron pairs in Astrophysics, ed. M. L. Burns, A. K. Harding, and R. Ramaty, AIP, NY, 3.
Gosling, J. T., 1993, J. Geophy. Res., 98, 18,949.
Hirasima, Y., Okudaira, K., and Yamagami, T., 1969, Acta Phyz. Hungaria, Suppl. 2, 39, 683.
Hua, X. M., and Lingenfelter, R. E., 1987, ApJ, 319, 555.
Koga, R., Simnett, G. M., and White, R. S., 1974, in Proc. of the 9th ESLAB Symp., ed. B. G. Taylor, ESRO SP-106, 31.
Kondo, I., and Nagase, F., 1969, in Solar Flares and Space Research, ed. C. Delager and Z. Svestka, (Amsterdam: North Holland Publishing Co., 1969), p. 114.
Kozlovsky, B., and Ramaty, R., 1974, ApJ, 191, L43.
Kozlovsky, B., and Ramaty, R., 1977, ApJ, 19, L19.
Lingenfelter, R. E., Flamm, E. J., Canfield, E. H., and Kellman, S., 1965, J. Geophys. Res., 70, 4087.
Lingenfelter, R. E., and Ramaty, R., 1967, High-Energy Nuclear Reactions in Astrophysics, ed. B. S. P. Shen, W. A. Benjamin, (New York: 1967), p. 99.
Meyer, J-P., 1989, Cosmic Abundances of Matter, ed. C. J. Waddington, New York: AIP, 245.
Miller, J. A., and Steinacker, J., 1992, ApJ, 399, 284.
Murphy, R. J., Share, G. H., Letaw, J. R., and Forrest, D. J., 1990, ApJ, 358, 298.
Murphy, R. J., Ramaty, R., Kozlovsky, B., and Reames, D. V., 1991, ApJ, 371, 793.
Murphy, R. J., et al., 1993a, 23rd Internat. Cosmic Ray Conf. Papers, 3, 99.
Murphy, R. J., et al., 1993b, Compton Gamma-Ray Observatory Symposium, ed. M Friedlander, N. Gehrels, and D. J. Macomb, AIP Conf. Proc. 280, 162.
Prince, T. A., et al., 1983, 18th Internat. Cosmic Ray Conf. Papers, 4, 79.
Ramaty, R., Kozlovsky, B., and Lingenfelter, R. E., 1975, Space Sci. Rev., 18, 341.
Ramaty, R., and Crannell, C. J., 1976, ApJ, 203, 766.
Ramaty, R., Kozlovsky, B., and Suri, A. N., 1977, ApJ, 214, 617.
Ramaty, R., and Lingenfelter, R. E., 1994, High Energy Astrophysics, ed. J. M. Matthews, Singapore: World Scientific, 32.

Ramaty, R., and Lingenfelter, R. E., 1995. The Analysis of Emission Lines, ed. R. Williams and M. Livio, Cambridge University Press, 183-187.
Ramaty, R., Mandzhavidze, N., Kozlovsky, B., and Skibo, J., 1993, Adv. Spac. Res. 13.275.
Ramaty, R., Mandzhavidze, N., Kozlovsky, B., and Murphy, R. J., 1995, ApJ, 455, L193.
Ramaty, R., Mandzhavidze, N., Kozlovsky, B., 1996, High-Energy Solar Physics, ed. R. Ramaty, N. Mandzhavidze, and X. M. Hua, AIP Conf Proc. 374, AIP Press, Woodbury, N. Y..
Reames, D. V., 1990, ApJS, 73, 235.
Reames, D. V., 1993, Adv. Space Res., 13, No 9, 331.
Reames, D. V., Barbier, L. M., and Ng, C. K., 1996, ApJ, 466, 473.
Reeves, H., 1994, Modern Physics, 66, 193.
Rieger, E., 1989, Solar Phys. 121, 323.
Rieger, E., 1996, High-Energy Solar Physics, ed. R. Ramaty, N. Mandzhavidze, and X. M. Hua, AIP Conf Proc. , 374, 194-199.
Riegler, G. R., et al., 1981, ApJ, 248, L13.
Ryan, J. M., et al., 1981, ApJ, 244, L175.
Schneid, E. J., 1994, Paper presented at the 1994 AAS Winter Meeting, Crystal City, Va.
Schneid, E. J., et al., 1996, ApJ, to be published.
Share, G. H., and Murphy, R. J., 1995, ApJ, 452, 933.
Shemi, A., 1991, MNRAS, 251, 221.
Swinson, D. B., and Shea, M. A., 1990, Geophys. Lett, 17, 1073.
Vestrand, W. T., 1988, Solar Physics, 100, 465.
Vestrand, W. T., and Forrest, D. J., 1993, ApJ, 409, L69.
Wang, H. T., and Ramaty, R., 1974, Solar Phys.,36, 129.
Wang, H. T., and Ramaty, R., 1975, ApJ, 202, 532.
Wild, J. P., Smerd, S. F., and Weiss, A. A., 1963, ARA&A, 1, 291.
Yoshimori, M., 1990, ApJS, 73, 227.
Yoshimori, M., et al., 1994, ApJS, 90, 639.

CHAPTER 5

DIFFUSE GALACTIC EMISSION, COSMIC RAYS, AND THE INTERSTELLAR MEDIUM

5.1 Introduction

In picturing our galaxy, one is often to thinking of the stars that can be seen with the eye or in the many excellent images obtained with optical telescopes. In addition, there are many other important components besides the visible stars that play a major role in the Galaxy. These components include the interstellar gas, the magnetic fields, photons of a wide range of frequencies, and the cosmic rays. The latter, which are misnamed from a time in the past when their nature was unknown, are in fact charged particles, some of quite high energy. There is a tug of war occurring in the Galaxy between the cosmic rays, the gas particles, and the Galactic magnetic fields, all of which are trying to expand the Galaxy, and the gravitational attraction of the matter, which is trying to pull the whole together.

This latter statement has implicit in it the belief that the cosmic rays are Galactic and not from beyond our galaxy. In Chapter 7 it will be seen that, although this assumption had seemed reasonable for some time, recent high-energy gamma-ray measurements of the Magellanic Clouds have shown it to be correct with a high degree of certainty. Thus, the description of the cosmic rays must be consistent with their Galactic nature and the Galactic forces that come into play. Although on a smaller scale there are believed to be local perturbations caused by supernovae explosions, on a broad Galactic scale, the balance just described must be taken into account.

It is worth remembering at this point that, although we live in our galaxy and some measurements can be made very well, in some respects we know less about it than other galaxies. This situation results from our Sun being located very near to the plane of the Galactic disk and relatively far from the Galactic center (~10 kpc).

The interstellar matter prevents us from seeing very far into the Galaxy in either the optical band or the x-ray band. Also, it is not possible to step back and take a picture of it. There is no beautiful picture of our galaxy, such as the one of the Great Galaxy in Andromeda. For our galaxy, we see what is appropriately called the Milky Way. Thus, the overall structure has been deduced to a large degree from line-of-sight measurements in the plane, mostly from the radio band, and the deconvolution of this information is not without problems.

High-energy gamma rays suffer only negligible absorption for a typical path in the plane of the Galaxy. For some time, high-energy gamma rays have been recognized to be produced principally by the interaction of cosmic rays with the interstellar matter in our galaxy. By combining the gamma-ray information with radio data, a picture of the Galaxy not previously available is forthcoming. The primary goal of this chapter is to discuss what can be learned about our galaxy from gamma rays. To do so, it is desirable to begin in the next section by describing the nature of the principal components of our galaxy that are involved in the production of high-energy gamma rays. As is often the case in astrophysics, it is necessary to use information obtained from other wavelengths, particularly the radio band, as well as measurements of the cosmic-ray particles. This section is followed by a description of the relevant interaction processes to see how the gamma rays that are of interest are produced and to have the quantitative basis for calculations. The next section provides a discussion of Galactic dynamic balance, wherein the expansive pressures in the Galaxy are balanced by gravitational attraction. The subsequent sections describe the gamma-ray observations in our galaxy and their interpretation in terms of our current understanding of the processes involved.

5.2 Cosmic Rays, Interstellar Matter, Photons, and Magnetic Fields

Galactic gamma rays are primarily produced by the cosmic rays interacting with the interstellar matter and, to a lesser extent, interstellar photons. Hence, it is these three constituents of our galaxy whose properties will be summarized here.

Cosmic rays are bare nuclei and electrons, some of very great energy. The discovery of cosmic rays is generally attributed to Hess (1911, 1912), who showed with instrumentation carried on a

balloon that ionization began to increase with altitude once one was far enough away from the Earth's surface. The exact nature of the cosmic rays, however, remained largely unknown for some time, until very high-altitude balloons began to reveal their properties after the Second World War. Now, many details of their properties are known, but it is fair to say that even today, more than 80 years later, the origin of the bulk of the radiation, or to be more specific, the manner and location of the primary acceleration of the particles, is primarily a matter of speculation. The energy density of the cosmic rays in the Galaxy is about equal to that of starlight, the kinetic motion of interstellar matter, and the magnetic field. Hence, they are not a trivial phenomenon. (At a more personal level, cosmic rays bombard the top of the atmosphere, and some of their secondary products proceed all of the way down to sea level. There are a few secondary cosmic-ray particles, mostly muons, passing through each of you every second.) The nuclei are the dominant component of the cosmic rays and have an average energy of several GeV, although nuclei with energies of more than 10^{20} eV have been detected. Protons constitute the majority of the nuclear component, with helium nuclei making up about 10 percent of the total by number, and heavy nuclei approximately 1 percent. Electrons of both positive and negative charges exist, and their intensity is about 1 percent of that of the protons at the same energy.

Figures 5.1 and 5.2 show the composition and energy spectra of some of the major components of the cosmic-ray nuclei. Their composition is seen to be similar to that of the solar system, except for the large excess of Li, Be, B, and the nuclei just below the iron group. These excesses are generally believed to be secondaries produced in the interaction of the cosmic rays with the instellar matter. Although cosmic-ray electrons have an energy density much smaller than that of the protons, they are of importance for producing high-energy gamma rays, as will be discussed later.

In the figure showing the energy spectra, it is seen that the spectra of the cosmic-ray components are very similar within the range shown. At higher energies there is no information on the composition of individual nuclei, but from ground-level, secondary air-shower measurements, it is known that the spectrum continues with the same slope to about 10^{15}eV, at which point the spectra begin to steepen and the composition begins to change to one with a greater

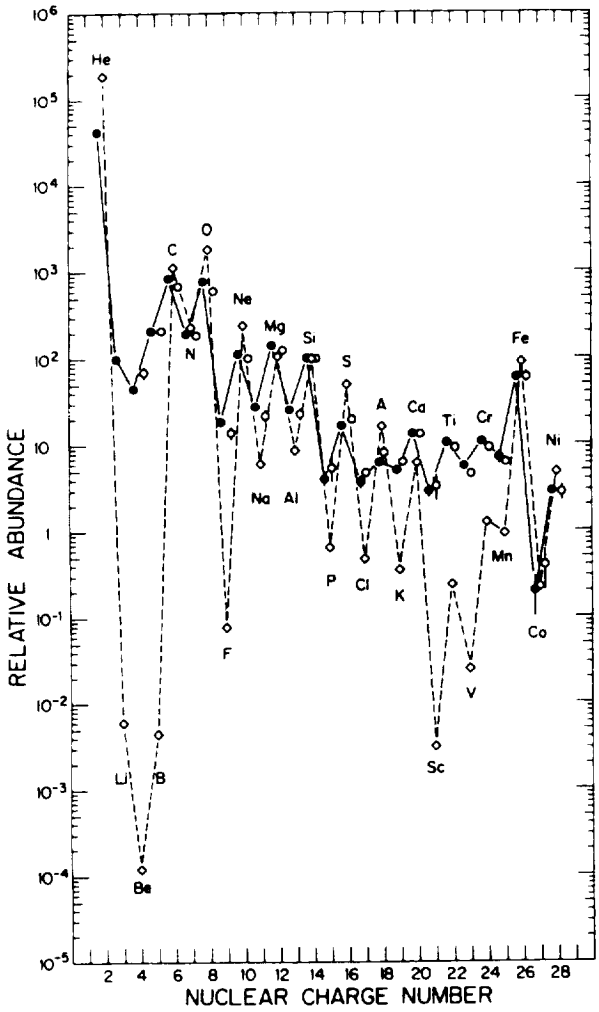

Figure 5.1. The cosmic-ray nuclear abundances (He-Ni) measured at 1 AU compared to the solar system abundances, all relative to Silicon (Si ≡ 100). The open circles represent a compilation of high-energy (1000 – 2000 MeV/n) measurements; the closed circles are solar system abundances. The figure is from Simpson (1983) and is reproduced with permission from "Annual Review of Nuclear and Particle Science", Vol. 33, c 1983 by Annual Review, Inc.

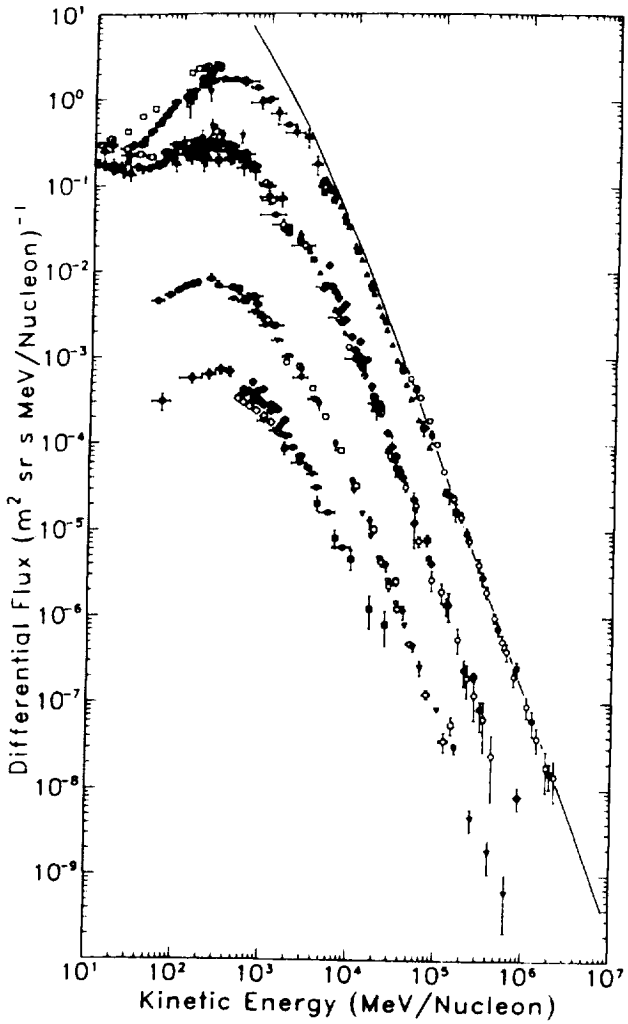

Figure 5.2. The energy spectra of cosmic-ray nuclei measured on many satellite and balloon-borne experiments near the time of solar minimum modulation. Differential energy spectra are for the nuclei of (from top) hydrogen, helium, carbon, and iron. The dashed curve shows the proton spectrum extrapolated to interstellar space eliminating the effects of solar modulation. The figure is from Simpson (1983) and is reproduced with permission from "Annual Review of Nuclear and Particle Science", Vol. 33, c 1983 by Annual Reviews Inc.

abundance of heavier nuclei. These characteristics are generally attributed to the fact that rigidities higher than that corresponding to about 10^{15} eV for a proton cannot be held in the Galaxy, and, for the same rigidity, heavier nuclei have a higher total energy. See, e.g., Fichtel and Linsley (1986). (It is also possible that sources in our galaxy cannot accelerate particles above this point.) At about 10^{18} eV, the spectrum begins to become less steep, suggesting the possibility of an extragalactic component.

An analysis of the secondary cosmic rays mentioned above suggests that the cosmic rays have typically passed through about 10 to 15 g cm^{-2}, although there is a distribution in path lengths, and that the higher-energy ($\sim 10^{11}$ eV) cosmic rays have passed through somewhat less material on the average than the lower-energy ($\sim 10^{9}$ eV) cosmic rays. Some secondaries, such as Be^{7}, are unstable and act as clocks giving an estimate of the age of the cosmic rays, now thought to be (1 to 2) x 10^{7} years. Combining the lifetime of the nuclei, the velocity, and the average amount of matter traversed shows that the average density of matter seen by the nuclear cosmic rays is (0.2 to 0.3) cm^{-3}. Hence, on the basis of the discussion in the last paragraph, they do not spend most of their time in the thin matter disk where the density is about 1 cm^{-3}, but, rather, they have a broader distribution relative to the Galactic plane.

The nonthermal continuum radio emission, which is generally attributed to the synchrotron radiation from cosmic-ray electrons interacting with the Galactic magnetic fields (e.g., Ginzburg and Syrovatskii 1964, 1965), provides information about the distribution of high-energy cosmic-ray electrons. Baldwin (1967, 1976) estimates the equivalent disk thickness for synchrotron emission to be about 750 pc., and some analyses have suggested that it is even larger. Significant nonthermal emission is even seen as high as 2 kpc above the plane. If it is assumed that both the electron density and the magnetic-field density have the same distribution on the average and that it is Gaussian, then, since the synchrotron radiation is proportional to the product of the two, the scale height of each individually is $\sqrt{2}$ x 0.75, or just over 1 kpc. It seems reasonable to assume that all cosmic-ray electrons, protons, and the magnetic fields, at least in our local region of the Galaxy, and probably elsewhere, have a scale height of this order. Most of the knowledge just described has come from direct measurements of the cosmic rays and, therefore, reflects the cosmic-ray properties in

the local vicinity of our galaxy. It is generally presumed that the energy spectrum and composition of the cosmic rays are similar elsewhere. The existing radio data on synchrotron radiation suggests that the electron spectrum is at least similar. The gamma-ray spectrum to be discussed later in this chapter supports the concept of uniform nucleon and electron spectra throughout the Galaxy. The question of the distribution of the cosmic rays in the Galactic plane will be addressed later in Section 5.5.

The interstellar matter that is relevant here is the diffuse interstellar matter with which the cosmic rays interact. It is in three forms: atomic, molecular, and ionized. All are known to be confined to a narrow disk of the order of 0.1 kpc or less. The radius of the disk does not have a sharp edge, but is of the order of 20 kpc with the Sun being about 8.5 kpc from the center and approximately in the middle of the disk.

Atomic hydrogen reveals its presence through the emission of the 21-cm line, which is produced by the hyperfine transition of this neutral atom. Observations of this line are usually expressed in terms of antenna or brightness temperature as a function of frequency Doppler shifted from the line's natural frequency. These measurements must then be converted to a density distribution using a rotational function for the Galaxy and including corrections for partial saturation, where appropriate. Although the translation of the observations into a Galactic spatial distribution is difficult, on a broad scale, the density profile is reasonably well known, even though details of arm structure are not always agreed upon by all workers in the field. A general spiral pattern does appear to emerge. The distribution of continuum radiation, gamma-radiation, HII regions, supernovae remnants, pulsars, and infrared emission is all consistent with the existence of spiral structure in the Galaxy.

The density distribution of molecular hydrogen cannot be measured directly, but must be inferred from other measurements. At present, the best approach appears to be through the observations of the 2.6 mm spectral line of ^{12}CO, which shows the distribution of cold interstellar material. The fact that ^{12}CO and ^{13}CO lines have the same general shape has been used to justify the belief that the ^{12}CO lines are associated with optically thin regions and that a Galactic distribution of CO may be derived accurately. The molecular hydrogen distribution was then originally estimated by relating the CO densities to those of H_2 on the basis of solar abundances. More

recently, a more accurate determination of this ratio has been possible, in part, through the use of the high-energy gamma-ray measurements. There are some problems in detail, especially in the general direction of the Galactic center, but a reasonable understanding seems to exist. However, the molecular hydrogen density distribution is still less certain than that of the atomic hydrogen because of the nature of its determination. The average Galactic radial distribution of molecular and atomic hydrogen deduced by Gordon and Burton (1976) shows that the molecular hydrogen to atomic hydrogen ratio is much larger in the inner Galaxy than it is in the outer Galaxy. This is consistent with the concept that the density and clustering of matter are greater in the inner Galaxy.

The CO observations indicate that the great majority of the molecular hydrogen is in clouds. There is also some evidence that atomic hydrogen tends to cluster on the peripheries of these clouds and in small clouds. The work of Solomon and Sanders (1980) has, in fact, suggested that the interstellar medium is dominated by massive cloud complexes; originally, it had not been clear whether molecular clouds were associated with spiral structure because of the difficulty of the measurements. On the basis of a high-sample survey and observations in both the first and second quadrants of the Galactic plane, Cohen et al. (1980) have shown the existence of the molecular counterparts of the five classical 21-cm spiral arms segments in these quadrants, namely, the Perseus arm, the Local arm, the Sagittarius arm, the Scutum arm, and the 4 kpc arm.

Departures from strict circular symmetry (e.g., gas streaming motions) introduce systematic errors in the Galactic matter distribution deduced from radio observations, and absolute mass determinations from integrated spectral line profiles also may be affected by optical depth and temperature effects. For a detailed discussion of these topics, see Mihalas and Binney (1981), Kulkarni and Heiles (1988), and Burton (1988). For a further discussion of the atomic and molecular hydrogen in the Galaxy, particularly as it relates to high-energy gamma-ray astronomy, see Bertsch et al. (1993). Helium and heavier nuclei also must be considered because they do contribute significant additions to the diffuse gamma-ray intensity. They represent about 10 percent and 1 percent of the hydrogen content, respectively, but their cross sections are larger. Since little is known about their composition throughout the Galaxy,

the local composition values are used and are included as multiplying factors in the production functions.

The contribution of other constituents of the interstellar medium is small. Ionized matter, grains, and dust are negligible from the standpoint of high-energy gamma-ray production, except for the possibility of gamma-ray lines from interstellar grains.

The relevant photons are the stellar ones of all wavelengths, but principally visible and near visible, and the blackbody radiation. They are important for the discussion here because cosmic-ray electrons interact with them and produce gamma rays. It is interesting to note that, once again, the energy density of these photons is about 1 eV in the plane.

Because of the limited information available, it is necessary to model the spatial distribution and spectrum of each of these photon populations. The Far-Infrared (FIR) photon radial distribution is taken from the cold dust emission curve of Cox, Krügel, and Mezger (1986); the decrease with height above the plane of the Galaxy is approximated by an exponential distribution characterized by a scale height of ~0.1 kpc. The distribution is then normalized so that the total integrated FIR luminosity of the Galaxy is 1.5×10^{10} L_o (Cox, Krügel and Mezger, 1986). The interstellar radiation fields in the near infrared, optical, and ultraviolet adopted here are taken from the work of Chi and Wolfendale (1991), who used the stellar model of Mathis, Mezger, and Panagla (1983) with four stellar components emitting in the range of $8 < \lambda < 1000$ μm. The distribution corresponding to a solar radius of 8.5 kpc (Kerr and Lynden-Bell, 1986) is adopted for use here.

The magnetic fields are important, not primarily because they lead to gamma rays (The gamma-ray production from cosmic-ray interactions with them is negligible.), but because they control the motions of the cosmic rays and are one of the expansive pressures in the Galaxy.

5.3 Interaction Processes for Continuum Emission

Within the Galaxy, cosmic rays may interact with matter, photons, and magnetic fields, in each case producing gamma rays. The cosmic-ray nucleon interactions give rise to gamma rays, primarily

through the decay of π^o mesons. Cosmic-ray electrons produce gamma rays through bremsstrahlung, but with a markedly different energy spectral shape. In addition, the positrons left over as nuclear debris slow down and annihilate with electrons to form a pair of 0.511 MeV gamma rays. Cosmic-ray electrons also interact with the interstellar starlight, optical and infrared, and the blackbody radiation through the Compton process. Finally, cosmic-ray electrons can interact with magnetic fields giving rise to synchrotron and curvature radiation, but these processes are very much less important than the others just mentioned for the Galactic diffuse gamma radiation and will not be discussed here.

High-energy physics has provided a wealth of information on nucleus-nucleus interactions, enough to permit a good calculation of the expected gamma-ray emission, although the calculations are long and tedious. The primary contribution comes from the proton-proton interaction with additions for the proton-helium, helium-protron, helium-helium, etc. interactions. The cosmic-ray composition is well known, and the interstellar medium composition, although less well known, is believed to be quite adequately known for this purpose. A very brief summary of the interaction process now follows.

Of the many mesons produced in the interactions of the cosmic rays with interstellar matter, the most commonly produced are π mesons, which can be either charged or neutral. The π^o mesons decay in about 10^{-16}s into two gamma rays, each with about 68 MeV in the rest frame. Many of the other mesons and hyperons also decay into π mesons, e.g., $K^{\pm} \rightarrow \pi^{\pm} + \pi^o$, $K_1^o \rightarrow 2\pi$, $K_2^o \rightarrow 3\pi^o$, $K_2^o \rightarrow \pi^+ + \pi^- + \pi^o$, and $\Lambda \rightarrow n + \pi^o$. The detailed calculations leading to the predicted intensity and energy spectrum of the gamma rays (based on the average numbers of mesons formed in an interaction, their angular distribution, and the resulting energy spectrum) are quite lengthy. This complexity is the result of the need to include the many different products, to take into account the different cosmic-ray species and interstellar nuclei in the correct proportions, to follow their decay, to integrate over all angles, and then to integrate over the cosmic-ray energy spectrum. These calculations, however, have been performed (See, e.g., Caravello and Gould, 1971; Stecker, 1971, 1979; Morris, 1984; Dermer, 1986). The resulting energy spectrum is shown in Figure 5.3. The

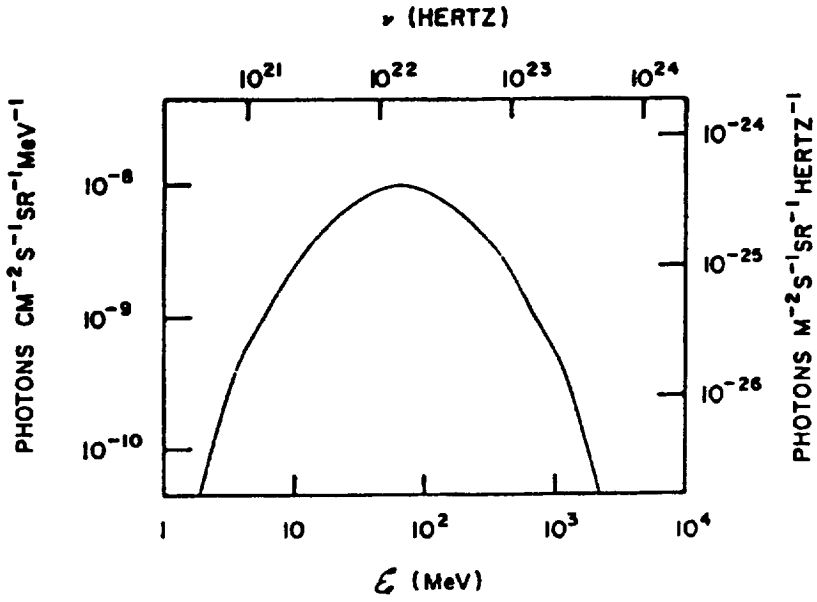

Figure 5.3. Typical differential spectrum expected from the decay of neutral pions producded in cosmic ray interactions with interstellar gas (Fichtel and Kniffen, 1974, adopted from Stecker, 1971). The curve is symmetric about E* since the abscissa is chosen to be log E.

total cosmic-ray nucleon, matter interaction source function is estimated to be 1.2 x 10^{-25} photons (E > 100 MeV) s^{-1} (hydrogen nucleon)$^{-1}$ and 3.0 x 10^{-26} photons (35 < E < 100) s^{-1} (hydrogen nucleon)$^{-1}$, assuming a helium-to-hydrogen ratio of 0.1 and a heavy-nuclei-to-hydrogen ratio of 0.01.

The energy spectrum of the gamma rays resulting from π^o decay is quite different from that resulting from most other astrophysical processes, such as bremsstrahlung, synchrotron radiation, and the inverse Compton effect.The origin of the unique spectrum is the decay of the π^o into two massless particles of equal energy in the rest frame. The probability distribution of the gamma rays in the laboratory system can be shown (e.g., Fichtel and Kniffen, 1974) to be a constant from $E_\gamma^*[(1-\beta)/(1+\beta)]^{1/2}$ to $E_\gamma^*[(1+\beta)/(1-\beta)]^{1/2}$ and zero outside this range, where E_γ^* is the energy of the π^o decay

gamma ray in the rest frame and β is the particle velocity divided by the speed of light. Since this range always includes E_γ^* and the integral is performed over a distribution of β's, the distribution must have a peak at E_γ^* and be symmetrical when plotted against ln E_γ, as seen in Figure 5.3.

As cosmic-ray electrons pass through the interstellar medium, they interact with the matter and produce photons over a wide range of energies. In the case of the energetic electrons, the resulting high-energy photon spectrum reflects that of the electrons, at least qualitatively, in the sense that it is a monotonically decreasing function with increasing energy. The largest gamma-ray intensity from electrons probably comes from bremsstrahlung. The calculation of the radiation from this mechanism in the region below 10^2 MeV is rather uncertain, because the interstellar cosmic-ray electron spectrum is not well known at low energies, where the electron spectrum observed near the Earth has undergone strong solar modulation, which suppresses the lower energies, particularly in the inner part of the solar system. The degree of the suppression is a function of the solar cycle. This problem did not arise in the case of cosmic-ray nucleon, matter interactions, because there the higher-energy cosmic rays predominate in the production of gamma rays, and the spectrum and general composition of these particles are well known and not affected markedly by solar modulation.

The cosmic-ray electron, matter gamma-ray production can be calculated using the bremsstrahlung cross-section formulas of Koch and Motz (1959). The calculations, in general, are very complex. However, to a good approximation, if the parent electron differential intensity has the form

$$j_e \equiv \frac{dJ_e}{dE_e} = KE_e^{-a}, \tag{5-1}$$

then the gamma-ray source function is given by

$$q(E_\gamma) = \Sigma_s \frac{\rho_s}{X_o}\left(\frac{4\pi}{E_\gamma}\right)\int_{E_\gamma}^{\infty} j_e dE_e , \tag{5-2}$$

where ρ_s is the mass density of the particular nuclear species, and X_o is the radiation length. The sum is over all target species. Assuming the interstellar composition to be 10 percent helium and 1 percent heavy nuclei, the result is

$$q(E_\gamma) \cong 4.7 \times 10^{-25}\, n_H \left(\frac{K E_\gamma^{-a}}{(a-1)}\right) \frac{\text{photons}}{\text{cm}^3\ \text{s MeV}} \quad , \qquad (5\text{-}3)$$

where n_H is the number of protons per unit volume in both atomic and molecular form. This last equation includes the effect of the nuclei heavier than the proton, as long as they are in the assumed proportions. Notice that the spectral index of the gamma-ray differential source function is the same as that of the differential electron spectrum. Although this source function is smaller than the cosmic-ray nucleon one for energies greater than 100 MeV, it dominates at lower energies and, in fact, becomes increasingly important the lower the gamma-ray energy.

Even for simple models, the cosmic-ray electron spectrum is not expected to have exactly the same shape everywhere in the Galaxy, because more secondary electrons are produced in regions of greater matter density. Specific gamma-ray spectral shapes will be discussed later in this chapter, following a discussion of the Galactic matter distribution and the current concepts of cosmic-ray confinement and propagation. For a further discussion on these source functions and those in the next subsection, see Bertsch et al. (1993) and Hunter et al. (1996).

Cosmic-ray electrons also interact with starlight photons, for which both the optical and infrared ranges are important, and with the 2.7° microwave blackbody radiation to produce Compton gamma rays. The source function of these interactions is much smaller in the Galactic plane in the vicinity of the solar system. The total contribution to the Galactic gamma radiation, however, is significant, although smaller than the others, because the cosmic-ray and stellar-photon scale heights above the Galactic plane are much greater than those of the matter, and, of course, the blackbody photon density is uniform. Hence, the integral intensity along a line of sight is closer to that of the bremsstrahlung than the source functions would imply.

The calculations associated with the production of Compton gamma rays are quite complex; however, they have been performed in some detail for the cases of astrophysical interest by Ginzburg and Syrovatskii (1964). A power law in energy for the electron spectrum of the form given by Equation (5-1) will produce a power-law gamma-ray source function of the form

$$q_{phi}\ (E_\gamma, r) = \frac{8\pi}{3}\,\sigma_T\,(\frac{4}{3}<E_{ph}>)^{(a-3)/2}\ (mc^2)^{(1-a)}\ K(r)E_\gamma^{-(a+1)/2}\ , \quad (5\text{-}4)$$

where E_{ph} is the photon energy before the electron interaction, and σ_T is the Thompson cross section. Whereas, in the case of bremsstrahlung, the astrophysical gamma rays come predominantly from electrons of energies similar to those of the gamma rays, or only a factor of several higher in energy, for the Compton radiation the average energy of the parent electron is given approximately by the equation

$$<E_\gamma> \cong \frac{4}{3}\,(\frac{E_e}{mc^2})^2\,<E_{ph}>. \quad (5\text{-}5)$$

Therefore, if the interaction is with starlight, where the typical photon energy is a few eV, in order for the gamma ray to be about a hundred MeV, E_e must be several GeV. For the 3° blackbody radiation, the typical photon energy is about 8×10^{-4} eV; therefore, the parent electrons must be in the range of 2×10^5 MeV to produce 100 MeV gamma rays. In the energy range above 1 GeV, there is no serious uncertainty in the electron spectrum because of a lack of knowledge of the solar modulation. However, in the few GeV range, the electron energy spectrum changes shape, so the calculation must be performed carefully. At energies as high as a few times 10^5 MeV, the electron energy spectrum is less well known, and an increased uncertainty in the calculated gamma-ray intensity is introduced.

5.4 Gamma-Ray Lines

In addition to the gamma-ray continuum existing in our galaxy, it is possible for distinct gamma-ray lines to be emitted in the intergalactic medium. These lines may be created in several of the following ways: (1) the interaction of low-energy cosmic rays with both interstellar gas and dust grains, leading to excited nuclei that

subsequently emit gamma rays through electromagnetic deexcitation of an excited state (often called "prompt emission"), (2) the annihilation of positrons that have been slowed down with interstellar electrons, and (3) the decay of excited nuclei that, in turn, emit gamma rays (called 'delayed emission").

For case (1), the strength of the line radiation contains information about the composition and nuclear excitation conditions in the radiating gas. At least some of the dominant lines formed in this way should be the result of deexcitation of the lower levels of the most abundant nuclei. In general, the main nuclear lines from direct excitation lie in the energy range 0.1 MeV < E < 7 MeV, where the lower range is affected by the contributions of higher excited states and less abundant nuclei. The lines from interstellar grains should be separable from those from the gas by their much narrower line widths, allowing the two constituents to be studied separately. The expected source strengths are low (Ramaty and Lingenfelter, 1979), but at least two of these lines now appear to have been detected.

Regarding case (2), positrons are produced in sizable numbers through the interaction of the cosmic radiation with interstellar matter, both through the decay of π^+ nuclei formed in high-energy cosmic-ray interactions and as the result of low-energy cosmic-ray interactions. There has been an observation of the 0.511 MeV line from the Galactic plane in the general direction of the Galactic center at an intensity level of approximately 1.0×10^{-3} photons cm^{-2} s^{-1} (Leventhal, MacCallum, and Strong, 1978); more recent information will be described in the next section. This flux is almost two orders of magnitude greater than would be expected from cosmic-ray interactions and may be from a single source at or near the Galactic center. Ramaty and Lingenfelter (1979) have proposed that positrons in large numbers may be coming either from black holes or from the decay of active nuclei produced in nucleosynthesis during supernovae explosions.

The subject of supernovae brings up case (3), the last of the three diffuse sources of gamma-ray lines to be discussed. If the half-life of the unstable parent nuclei formed in a supernova, nova, flare or elsewhere is very long, compared to the typical time interval between production (nucleosynthesis) events in the volume of space observed, the abundance tends to build up to a steady-state value. Thus, in principle, at least, gamma-ray-line astronomy provides a means for the study of cosmic explosions and nucleosynthesis

events through their residue in interstellar space, as well as through direct observation of the explosion itself. The most intense (for interstellar space as opposed to the supernova remnant itself) is the 1.809 MeV line from the decay of ^{26}AL (Ramaty and Lingenfelter, 1979). In view of the long lifetime of the isotope (mean life = 1.1 x 10^6 years), the emission is expected to be steady and distributed well beyond the supernova remnant. The current experimental status will be described in Section 5.7.

5.5 Galactic Dynamic Balance and the Cosmic-Ray Distribution

Several fundamental theoretical considerations place constraints on the cosmic-ray distribution. It is now generally accepted that the cosmic rays and magnetic fields are primarily Galactic and not universal. The recent important support for this belief from high-energy gamma-ray observations of the Magellanic Clouds will be described in Chapter 7. If these fields and cosmic rays are Galactic, they can only be constrained to the Galactic disk by the gravitational attraction of the matter (Biermann and Davis, 1960; Parker, 1966, 1969, and 1977). The local energy density of the cosmic rays (~1 eV/cm^3) is about the same as the estimated energy density of the magnetic field and that of the kinetic motion of matter. Together, the total expansive pressure of these three effects is estimated to be approximately equal to the maximum that the gravitational attraction can hold in equilibrium. Assuming the solar system is not at an unusual position in the Galaxy, these features suggest that the cosmic-ray density throughout the Galaxy may be as large as could be contained under near-equilibrium conditions.

Further theoretical support is given to this concept by the calculated slow diffusion rate of cosmic rays in the magnetic fields of the Galaxy and the small cosmic-ray anisotropy. These considerations lead to the hypothesis that the energy density of the cosmic rays is larger where the matter density is larger on a coarse scale such as that of the Galactic arms. On a smaller scale, the pressures of the cosmic-ray gas and magnetic fields cause the cosmic-ray gas and field system to expand between the large clouds through which the magnetic fields thread. As noted earlier, the radio continuum measurements and cosmic-ray results support the picture of the cosmic rays having a large-scale height relative to the matter and spending only a relatively small amount of time passing through the dense cloud region. This large-scale height is important in

determining the Compton contribution to the diffuse emission, as will be seen later. Also, on a small scale, it is believed that supernovae cause local perturbations and density fluctuations. The picture that seems to emerge then on a broad scale is that cosmic rays, at least below 10^{16} or 10^{17} eV/nucleon, are bound to the lines of force, and the field lines are closed (or else the cosmic rays would escape too quickly). The cosmic rays are constrained and not free to escape individually. Thus, their escape is the result of cosmic-ray group pressure inflating the magnetic field lines and pushing outward from the Galaxy and of diffusion.

The fact that the cosmic-ray density is approximately as large as can be contained suggests a plentiful source. What is this source? At present, it is generally believed that cosmic rays come from supernovae and flare stars, but may receive subsequent acceleration. If supernovae were to supply the total cosmic-ray energy, each supernova would have to produce between 10^{49} and 10^{50} ergs of cosmic rays, assuming a supernova rate of 1 every 30 years. This seems not to be a prohibitive amount of energy for a supernova to give to energetic particles. Even if supernovae are not the primary initial source, they may supply an important part of the total energy of the cosmic rays by shock-wave acceleration of existing cosmic rays (Blandford and Ostriker, 1980).

Within the context of the previous discussion, an unanswered question is the scale of the coupling between cosmic rays and matter. Synchrotron measurements suggest a scale height for the magnetic fields and cosmic rays perpendicular to the plane of about 1 kpc. Calculations of the diffusion rate of cosmic rays in the Galaxy also suggest that the scale may be of this order, or somewhat larger. Galactic arm widths and large clouds may also suggest a coupling scale in the range of a half to a few kpc in the plane. Disturbances, such as supernovae, suggest the scale is probably not smaller than about a half kpc. A purely theoretical approach to obtain an answer to the question of the scale size seems not to be possible, because too much is unknown. As will be seen, in the next section, the high-energy gamma-ray results now provide experimental information to obtain at least an approximate measure of the scale of the coupling. It is, as will be seen, consistent with the previous discussion.

5.6 Current High-Energy Gamma-Ray Results and Their Interpretation

The most intense celestial high-energy gamma radiation observed is that from the Galactic plane. This feature was observed first by the pioneering counter telescope flown on OSO-3 (Kraushaar et al., 1972), and the major features of this Galactic radiation were defined by measurements made with the SAS-2 satellite launched in 1972 (Fichtel et al., 1975; Hartman et al., 1979), and the COS-B satellite launched in 1975 (Bennett et al., 1977; and Mayer-Hasselwander et al., 1980). The results from these satellites suggested that the Galactic plane gamma-radiation is strongly correlated with Galactic structural features, especially when the known strong gamma-ray sources are subtracted from the total radiation. These important results now have been followed by the observations of the Energetic Gamma Ray Experiment Telescope (EGRET) flown on the Compton Gamma-Ray Observatory (CGRO). Because EGRET is far more sensitive than the gamma-ray telescopes on the earlier missions and has improved angular resolution, the EGRET results are now the ones to be considered and compared to theoretical expectations.

The gamma-ray sky appears quite different from the optical sky, even when the detector resolution is considered, as seen in Figure 5.4. The Galactic diffuse radiation dominates strongly, as seen even more clearly in Figure 5.5. To the best of our knowledge, the majority of the Galactic plane radiation is diffuse, with there being in addition three strong point sources and several weak ones, which will be discussed in the next chapter. The radiation is particularly intense and more narrow in a quarter of the Galactic latitude range centered about the Galactic center. It is least intense in the general direction of the anticenter, where it is also broader in Galactic latitude. Not surprisingly, in view of the discussion that has been presented earlier in this chapter, the gamma-ray sky is more similar to the radio sky, and particularly the 21 cm and CO intensity distribution, which also show a high intensity along the Galactic plane. The question that now must be addressed is, how well does the high-energy gamma radiation that is observed match that which is expected on the basis of the detailed considerations that have been described thus far in this chapter. The answer is: reasonably well, as will be seen as the two are compared. In order to make this comparison, it is necessary to proceed from the source functions discussed in Section 5.3 to the estimated number of

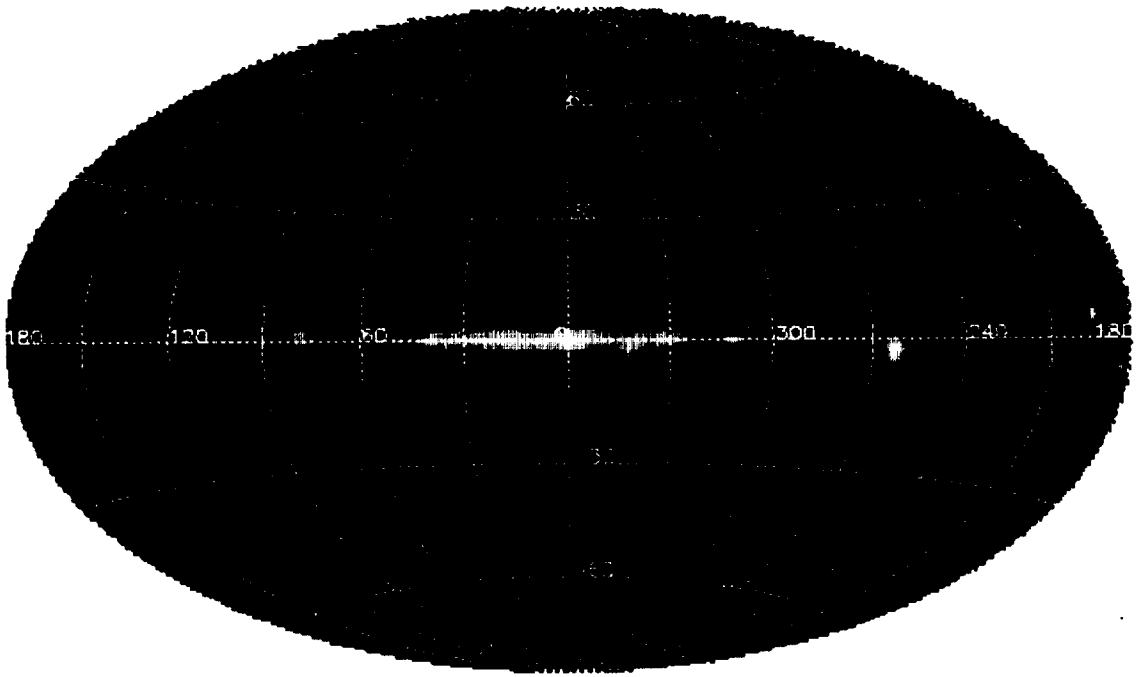

Figure 5.4. All sky intensity distribution for high-energy gamma rays (E>100MeV) determined by EGRET. (Courtesy of Carl Fichtel and the EGRET Instrument Science Team).

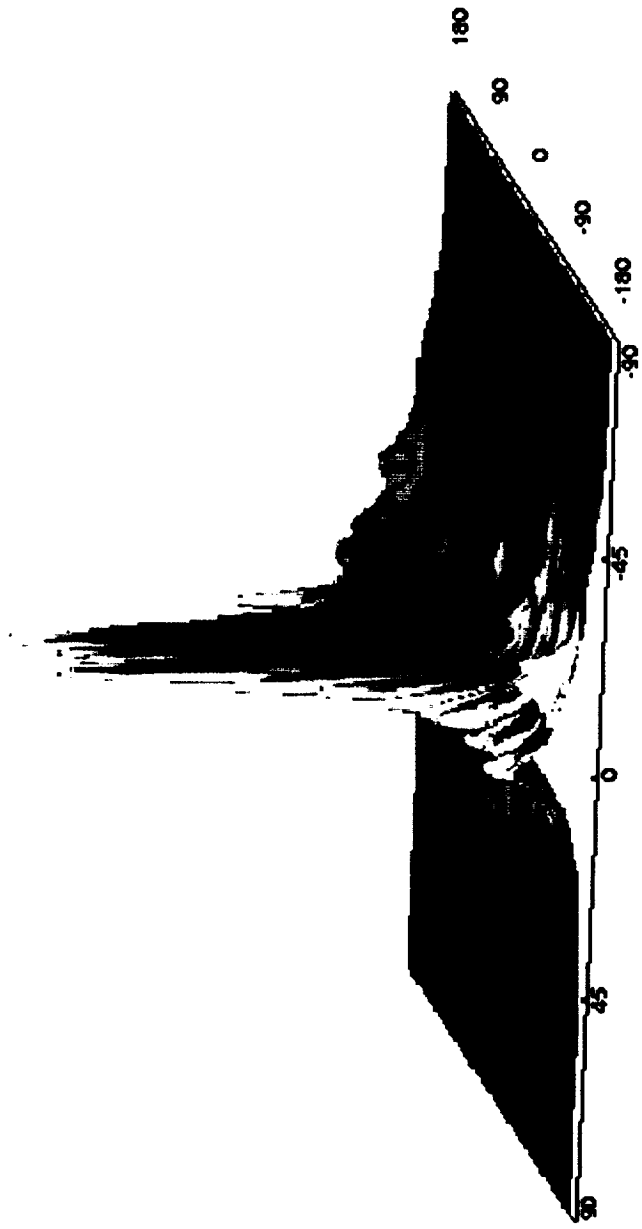

Figure 5.5. Three-dimensional portrayal of the high-energy gamma-ray intensity (E > 100 MeV) measured by EGRET. (Courtesy of Carl Fichtel and the EGRET Science Instrument Team).

gamma rays expected to be seen in a given direction within a small solid angle element.

The number of gamma rays per unit area and time within dE falling on a surface a distance r from a volume element dV with a source strength q is

$$\frac{1}{4\pi^2} q(E_\gamma,r)\, dV dE. \tag{5-6}$$

Hence, the differential intensity is

$$j = \frac{dJ}{dE} = \frac{1}{4\pi}\int_0^\infty q(E_\gamma, r)\, dr, \tag{5-7}$$

where j is the number of photons per unit time, area, solid angle, and energy.

Considering first the interactions between cosmic rays and matter,

$$\begin{aligned} j_s(E_\gamma,l,b) = \frac{1}{4\pi}\int dr\Big\{&\Big[q_{\gamma_n}(E_\gamma, r=0) g_n(r,l,b)\Big] * \\ &\Big[n_{HI}(r,l,b) + n_{HII}(r,l,b) + n_{H_2}(r,l,b)\Big] \\ &+\Big[q_{\gamma_e}(E\gamma, r=0) g_e(r,l,b)\Big] \\ &\Big[n_{HI}(r,l,b) + n_{HII}(r,l,b) + nH_2(r,l,b)\Big]\Big\}, \end{aligned} \tag{5-8}$$

where g_i is the ratio or the density of the cosmic-ray component of type i to the local value at a distance r from the Sun in the direction (l, b) to that at r = 0, q_{γ_n} represents the number of rays produced per second per volume element per energy interval in interactions of nucleonic cosmic rays (with the intensity and spectral distribution in the solar vicinity) with the interstellar gas, and q_{γ_e} and is a similar function for cosmic-ray electrons. Implicit in this approach is the assumption that the spectral shape of each component is unchanged throughout the Galaxy. This assumption is reasonable, as long as the source spectral shape in the solar vicinity is typical of that throughout the Galaxy, and energy losses remain within certain limits. The latter condition is true for the energy region and intensity levels relevant to the consideration here.

Equation (5-8) shows that the cosmic-ray primary electron component is proportional to the cosmic-ray nucleon component, as it is expected to be if the secondary electron component is small compared to the primary electron component, as the positron data suggest. The gamma-ray intensity is approximately proportional to the integral over the product of the ratios of the cosmic-ray intensity and matter density to their local values.

Similar expressions apply to the Compton radiation, except that u_{phi} refers to the photon density, and the q_{phi} refers to the Compton process, with i denoting the wavelength band. Including this process and making the assumption implied by the discussion of the last paragraph that $g_n = g_e = g_{cr}$ gives

$$j_T(E_\gamma,l,b) = \frac{1}{4\pi} \left\{ \int_0^\infty dr [q_{\gamma_n}(E_\gamma, r=0) + q_{\gamma_e}(E_\gamma,r=0)] g_{CR}(r,l,b) \right. *$$

$$[n_{HI}(r,l,b) + n_{HII}(r,l,b) + n_{H_2}(r,l,b)] +$$

$$\left. \sum_i \int_0^\infty dr\, q_{phi}(E_\gamma,r=0) g_{cr} u_{phi}(r,l,b) dr \right\}. \qquad (5\text{-}9)$$

There are three unknown parameters that must be determined by making a quantitative comparison between the predicted intensity and that which is observed. One is the level of the extragalactic radiation, presumably reasonably uniform. It is taken as being an unknown function of energy and a constant in direction. The second is the normalization of the molecular hydrogen density based on the CO observations discussed in Section 5.2. The third is the scale of coupling between the cosmic rays and matter. Fortunately, the current EGRET data are more than adequate to allow a determination of these parameters and still permit a good comparison between theory and experiment.

These detailed experimental results from EGRET just mentioned are in agreement with the pioneering results mentioned earlier, but provide much greater detail, to which the predictions of any model may be compared. Hence, the discussion here will be limited to these more recent results. There are a number of aspects of the

results that are of interest. These include the energy spectrum, which may be compared to the calculated one to determine if it is in agreement with that predicted by the source functions. Then there is the study of the radiation as a function of Galactic latitude and longitude, which shows whether the predicted distribution in the Galaxy is consistent with the observations. Finally, the distribution is, in fact, in good agreement with expectations, and estimates of the three parameters discussed earlier may be made.

Turning first to the energy spectrum, it has been found through an analysis of the EGRET results (Hunter et al., 1996) that the spectrum did not change with position in the Galaxy within the relatively small uncertainties. This result, of course, is consistent with the electron-to-proton ratio being essentially the same throughout the Galaxy. Therefore, it is also consistent with the theoretical considerations presented here.

With regard to the spectrum itself, the best spectrum is obtained in the general direction of the Galactic center, because the intensity is larger and more photons are collected. This spectrum is shown in Figure 5.6. Notice that one does see the bump in the spectrum caused by the neutral pion decays resulting from the nucleon-nucleon interactions. Generally, there is excellent agreement between the predicted spectrum and the observed one in both magnitude and shape. The exception is the deviation in the GeV region. This deviation increases from about 1 to 5 GeV and then remains fairly constant. The cause of this is not known. One possibility is that, for some reason, the cosmic-ray spectrum seen at the Earth is different at higher energies than that in the rest of the Galaxy. This explanation seems unlikely. Another possible explanation is that the theoretical calculations of the gamma-ray production are inadequate because of the approximations made in the case of the nucleon-nucleon interactions. This alternative is currently being investigated. Other possibilities are considered in the paper of Hunter et al. (1996), but they are believed to be less likely.

Consider next the longitude and latitude distributions. The distribution in latitude of the gamma-ray diffuse intensity along the Galactic plane is shown in Figure 5.7, and the distribution as a function of latitude is shown in Figure 5.8 for several different longitudes. Notice the obviously excellent agreement between the

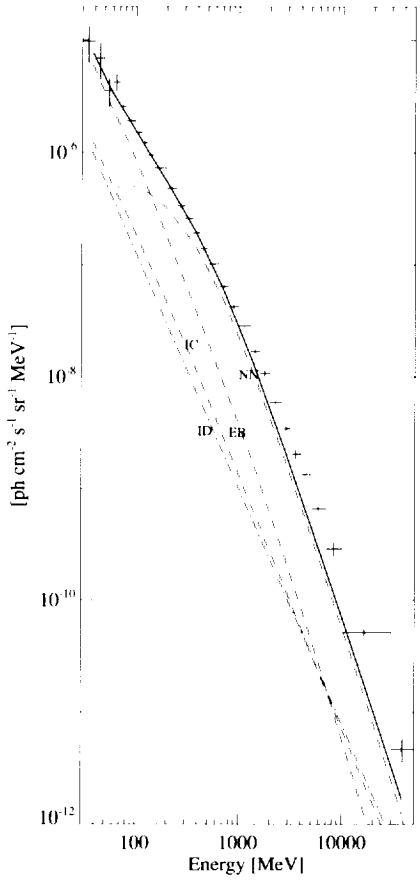

Figure 5.6. The average, diffuse gamma-ray spectrum of the Galactic center region, 300° < l < 60°, -10° < b < 10°, measured by the EGRET telescope. The contribution of point sources detected with more than 5σ significance has been removed. The curves show the expected contributions from the nucleon-nucleon interactions, bremsstrahlung, the inverse Compton interactions, and the sum of the three. (The figure is from Hunter et al., 1996).

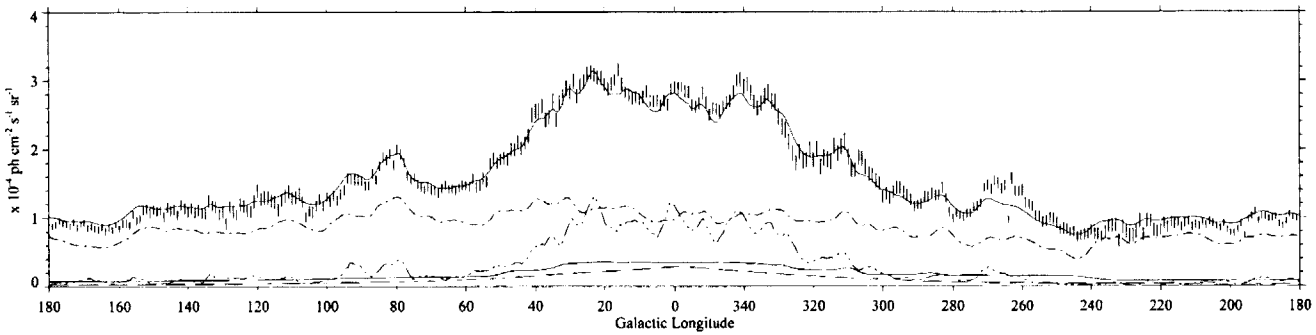

Figure 5.7. The diffuse gamma-ray intensity (E > 100 MeV) measured by EGRET as a function of Galactic longitude for the interval -10° < b < 10° compared to the expected distribution as explained in this chapter. The lower lines represent cosmic rays interacting with various components: dash-dot for atomic nuclei, dash-triple dot for molecular hydrogen, solid line for ionized hydrogen, and dashed line for inverse Compton radiation. The figure is from Hunter et al. (1996).

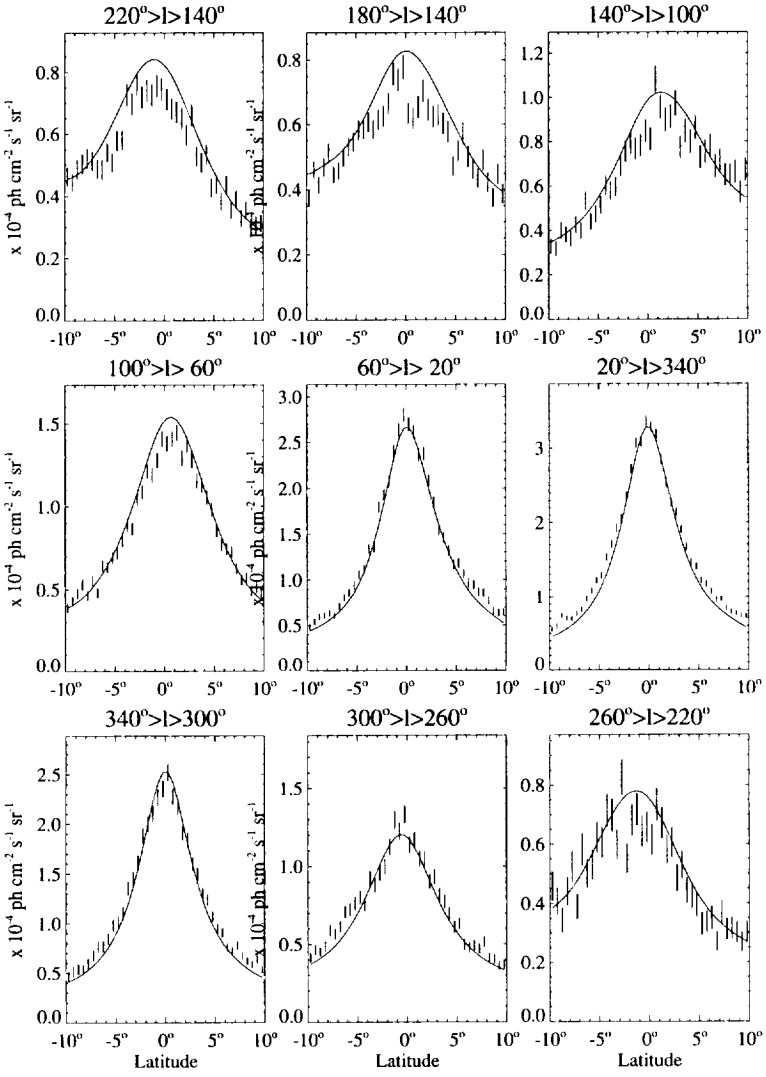

Figure 5.8. The diffuse gamma-ray intensity (E > 100 MeV), measured by EGRET as a function of Galactic latitude for several different longitudes. The figure is developed from the work of Hunter et al. (1996).

distributions and the predicted ones. Many additional comparisons are shown in the paper of Hunter et al. (1996) with similar fine agreement between the experimental results and the theoretical estimates. A major point-source contribution seems excluded by these results and is probably not expected based on the observed point sources as discussed in the next chapter of this book. The conclusion that may be deduced now is that the concept of dynamic balance and the corresponding cosmic-ray distribution leads to a prediction for the high-energy gamma-ray distribution in our galaxy that agrees well with the experimental results.

For most latitude cuts, there is a smooth, approximately symmetrically peaked distribution about the center, which is taller and relatively narrow in the center region and much lower and wider in the anticenter region. This results from the anticenter direction being dominated by the relatively local matter distribution, and the center direction being dominated by rather distant features. Since the width of the matter distribution is similar, the latter naturally has a much narrower distribution. The distributions are in good agreement with the prediction, both in shape and in intensity.

Notice in Figure 5.9 the plateau in the region ($-20^\circ < l < -10^\circ$), which is due to a large molecular cloud. The fact that this cloud and other clouds that have been seen are well explained in intensity by the model indicates that the cosmic-ray density is neither enhanced nor diminished on the scale by the cloud itself. This result suggests that the cosmic rays travel through these clouds relatively easily.

Finally, there is a matter of the three parameters that were determined in the process of the analysis of the diffuse gamma-ray distribution. The first is the value of X, the conversion factor to proceed from the CO data to the estimate of the molecular hydrogen normalization discussed in Section 5.2. It was found to be $(1.56 \pm 0.05) \times 10^{20}$ H-atoms cm^{-2} $(Kkms^{-1})^{-1}$, an improved estimate of this quantity, but consistent with earlier values, which had larger uncertainties. The value of the cosmic-ray, matter coupling scale was found to be (1.75 ± 0.20) kpc, consistent with the discussion in Section 5.4. The last constant to be determined is the value of the extragalactic diffuse radiation; it is relatively small and will be discussed in detail in Chapter 10.

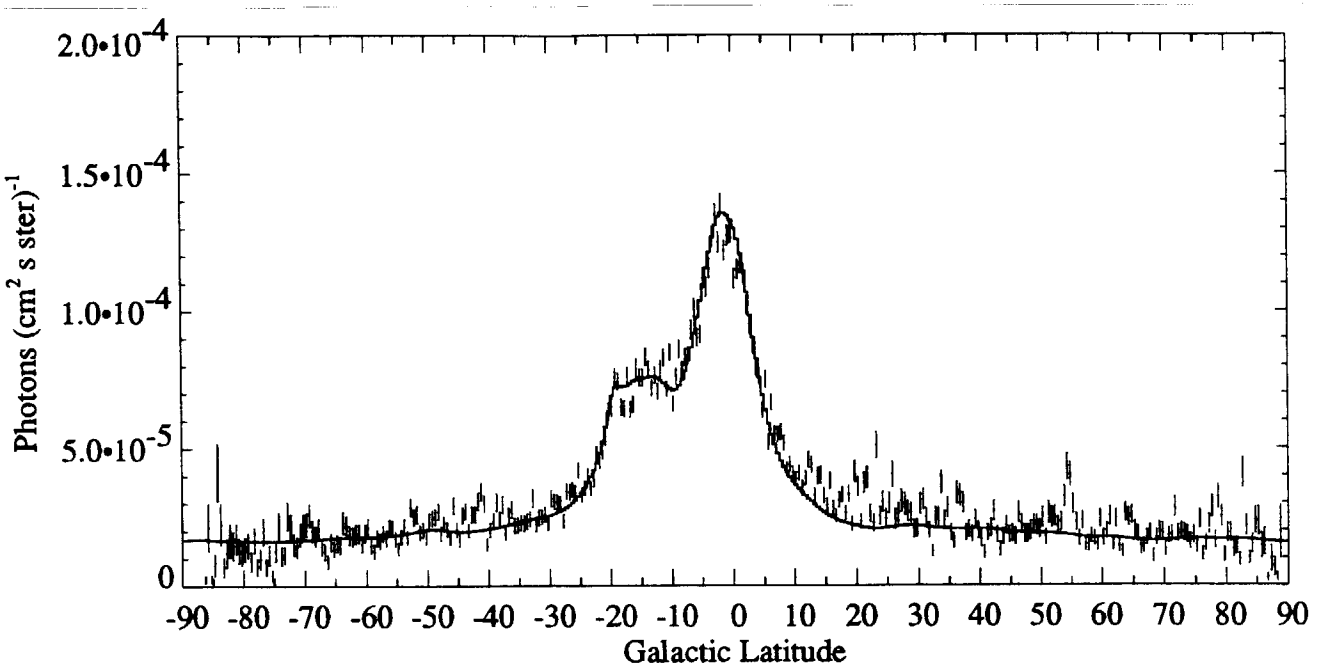

Figure 5.9.The diffuse gamma-ray emission (E > 100 MeV) measured by EGRET as a function of Galactic latitude for a 20° longitude interval centered at b = -210°. The figure is from the work of Hunter et al. (1996). Note the molecular cloud between b = -10° and b = -20°.

5.7 Current Low-Energy Gamma-Ray Results and Their Interpretation

A general diffuse Galactic continuum emission has now been observed by the COMPTEL instrument on the Compton Gamma-Ray Observatory (Strong et al., 1994). This is a more difficult energy range in which to measure the gamma radiation than the high-energy gamma-ray range discussed in the last subsection because of the basic nature of the gamma-ray conversion process described in Chapter 11. Nonetheless, the diffuse emission appears visible in the whole energy range covered by the telescope, although a significant contribution from point sources is possible. The energy spectrum is a smooth continuation of that observed at higher energies. The spectral shape is consistent with the radiation being from bremsstrahlung by cosmic-ray electrons. The intensity is somewhat above that predicted. This difference may be caused by either the electron spectrum at low energies being slightly more intense than thought or by unresolved point sources. The former is certainly possible, because the electron spectrum at low energies is not well known because of the strong solar modulation in that energy band and the difficulty of interpreting radio data precisely enough.

In addition to the continuum radiation in the low-energy gamma-ray region, there are also gamma-ray lines, of which two have been seen as apparent diffuse Galactic radiation. The 1.809 MeV line emission from ^{26}Al was first observed from the central region of the Galaxy by Mahoney et al., (1984) with a gamma-ray telescope on HEAO-3. The ^{26}Al line can be produced in a variety of sites, as noted well before it was observed by Ramaty and Lingenfelter (1979). Because of its long lifetime of 1.04×10^6 years, ^{26}Al will diffuse some from its source; it will also provide information during the last million years or more. The line has now been mapped along the Galactic plane by COMPTEL (Diehl et al., 1994, 1995). As shown in Figure 5.10, the radiation is not smooth like the continuum emission, but shows substantial evidence for localization. According to the authors, apart from a possible detection of the nearby Vela supernova remnant, essentially all of the other localized emission regions that can be seen in the figure must be caused by an ensemble of sources. At the moment, it seems not to be clear whether this radiation is primarily caused by Wolf-Rayet stars, supernova explosions, massive stars, or other possible sources.

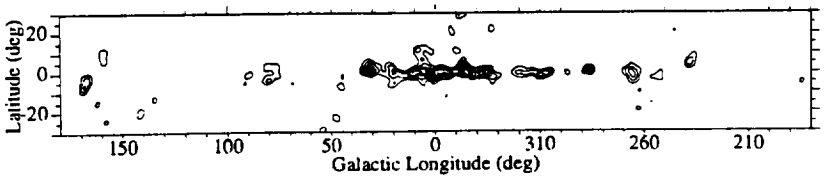

Figure 5.10. COMPTEL map of the Galactic ^{26}Al 1.809 MeV emission line, showing evidence for localized emission regions along the Milky Way. This figure is reproduced from Adv. Sp. Res. (1995), 15, 5, 123, with kind permission from Elsevier Science Ltd., The Boulevard, Langford Lane, Kidlington OXS 1GB, UK, and Diehl et al.

The amount of radiation seems to be fairly reasonable for either of the first two possibilities.

Another line that has been seen coming from the Galactic plane is the 0.511 MeV electron, positron annihilation line. This line has been reported in the general region of the Galactic center by a number of observers beginning with Leventhal, McCallum, and Strong (1978). Prior to the launch of the Compton Observatory, there was considerable difficulty in interpreting the data, because large differences in the opening angles of the instruments used to detect the radiation made it unclear as to whether variations in intensity were caused by time variations in the source, the size of the opening angle, or possibly other instrumental effects. For a summary of these early results, see, for example, Share et al. (1988), Gehrels et al. (1991), and Ramaty and Lingenfelter (1994a).

The OSSE instrument on the Compton Observatory has a relatively small solid angle compared to many of the earlier instruments, 3.8° x 11.4°. Extensive measurements were made of the Galactic center region with OSSE, wherein the radiation was mapped and a search was made for time variations. Purcell et al. (1993 and 1994) note that the 0.511 MeV line is described reasonably well by two components. The more intense component, with an intensity of 1.5 x 10^{-3} photons cm^{-2} s^{-1} in the inner region, is a spheroid centered on the Galactic center and with a scale size of about 1.2 kpc. The other contribution is a narrow disk component in both directions, but whose intensity at the Galactic center is only about 20% that of the spheroidal component. Alternatively, Tueller et al. (1995) have suggested that the source 1E 1740.7-2942 may account

for most of the central component. There is no evidence for variability in the narrow 0.511 MeV line in the OSSE data. While this result does not rule out long-term time variability, it does lead in the direction of making the time variability explanation of the early results less likely than the alternatives. Purcell et al. (1993, 1994) also show that the positronium fraction is 0.96 ± 0.04, in general agreement with the earlier, generally less precise, results.

The origin of the diffuse component of the 0.511 MeV line along the Galactic plane is general attributed to the decay products of various Galactic processes of nucleosynthesis, particularly supernovae. The most likely parents in the decay process are ^{56}Co, ^{44}Sc, and ^{26}Al. There seems to be reasonable quantitative agreement with theoretical estimates of the intensity. See, e.g., Ramaty and Lingenfelter (1994b), and also Ramaty (1996) and Timmes et al. (1996). The origin of the 0.511 Mev line in the bulge in the Galactic center is not certain. Several possibilities have been suggested, including a supernova, several stellar-mass-size black holes, and a massive black hole in the center of the Galaxy. The present data are not adequate to distinguish between these possibilities.

When low-energy cosmic-ray protons and helium nuclei interact with interstellar ^{12}C and ^{16}O, relatively narrow deexcitation lines at 4.44 and 6.13 MeV should be produced. Similarly, interactions of low-energy cosmic ray ^{12}C and ^{16}O nuclei should interact with interstellar hydrogen nuclei to produce broad lines at these same energies. Bloemen et al. (1994) reported a positive detection of these two lines from the direction of the Orion molecular cloud complex. The estimated line intensities were $(5\pm1) \times 10^{-5}$ gamma rays cm^{-2} s^{-1} for the 4.44 MeV line and $(3\pm1) \times 10^{-5}$ gamma rays cm^{-2} s^{-1} for the 6.13 MeV line. These flux numbers are within the rather broad range of theoretical predictions. On the other hand, Harris, Share, and Messina (1995) found only upper limits for these lines from radiation from the Galactic center and anticenter using the data from the Solar Maximum Mission. For the Galactic central radian, they obtained upper limits of about 9×10^{-5} gamma rays cm^{-2} rad^{-1} s^{-1} for narrow lines and 3.0×10^{-4} gamma rays cm^{-2} rad^{-1} s^{-1} for broad lines. Depending on the theoretical model used to interpret the data, there may or may not be a conflict in these results.

5.8 Summary

The high-energy gamma-ray sky is dominated by diffuse radiation from the Galactic plane. The distribution has some similarity to the 21-cm radio sky, but is even more enhanced in the central quadrant of the Galactic plane. The majority of the high-energy gamma radiation is explained well by the interaction of the Galactic cosmic rays with the interstellar matter and photons. The radiation is specifically consistent with the assumption that the cosmic rays are Galactic and vary with position in the Galaxy. The concept that the cosmic rays are Galactic is supported by a number of observations, but particularly by the recent measurement of the high-energy gamma-ray emission from the Small Magellanic Cloud, which shows that the bulk of the cosmic rays are Galactic, and hence play an important role in Galactic dynamic balance. The cosmic-ray density seems to be at the local value in nearby Galactic molecular clouds —that is, there appears to be nothing special about these clouds. The low-energy gamma-ray continuum radiation also is likely to be of Galactic cosmic-ray interaction origin. The 1.806 MeV line Galactic radiation, however, is likely to be from Wolf-Rayet stars or supernova explosions, and the 0.511 MeV line may be from novae, supernovae, pulsars, black holes, or some combination.

References

Baldwin, J. E., et al., 1967, Radio Astronomy and the Galactic System, IAU Symp., 31, 337.

Baldwin, J. E., 1976, The Structure and Content of the Galaxy and Galactic Gamma Rays, NASA CP-002 (Washington, DC), p. 206.

Bennett, K., Bignami, G. F., Buccheri, R., Hermsen, W., Kanbach, G., Lebrun, F., Mayer-Hasselwander, H. A., Paul, J. A., Piccinotti, G., Scarsi, L., Soroka, F., Swanenburg, B. N., and Wills, R. D., 1977, in Recent Advances in Gamma Ray Astronomy, ESA SP-124, p. 83.

Bertsch, D. L., Dame, T. M., Fichtel, C. E., Hunter, S. D., Sreekumar, P., Stacy, J. G., and Thaddeus, P., 1993, ApJ, 416, 587.

Biermann, L., and Davis, L., Jr., 1960, Zs, F., Ap., 51. 19.

Blandford, R. D., and Ostriker, J. P., 1980, ApJ, 237 , 793.

Bloemen et al., 1994, A&A, 281, L5.

Burton, W. B., 1988, Galactic and Extragalactic Radio Astronomy, 2d ed., ed. G. L. Verschuur and K. I. Kellermann (New York; Springer), 295.

Caravello, G., and Gould, R. J., 1971, Nuovo Cimento, 2B 77.

Chi, X., and Wolfendale, A. W., 1991, J. Phys. G., 17, 987.

Cohen, R. S., Cong, H., Dame, T. M., and Thaddeus, P., 1980 ApJ, 239, L53.

Cox, P., Krügel, E., and Mezger, P. G., 1986, A&A, 155, 380.
Dermer, C. D., 1986, A&A, 157, 223.
Diehl, R., et al., 1994, Proc. Second Compton Symposium, ed. C. E. Fichtel et al., AIP Conf. Proc., 304, 147.
Diehl. R., et al., 1995, Adv. Sp. Res., 15, 5, 123.
Fichtel, C. E., and Kniffen, D. A., 1974, in High Energy Particles and Quanta in Astrophysics, ed. F. B. McDonald and C. E. Fichtel, p. 428.
Fichtel, C. E., Hartman, R. C., Kniffen, D. A., Thompson, D. J., Bignami, G. F., Ögelman, H., Özel, M. E., and Tümer, T., 1975, ApJ, 198, 163.
Fichtel, C. E., Arnett, D., Grindlay, J., and Trombka, J., 1980, "Gamma Ray Astrophysics," in Astrophysics from Spacelab, ed. P. L. Bernacca and R. Ruffin (Dordrecht: D. Reidel Publ. Co., 1980), p. 145.
Fichtel, C. E., and Linsley, J., 1986, ApJ, 300, 474.
Gehrels, N., Barthelmy, S. D., Teegarden, B. J., Tueller, J., Leventhal, M., and McCallum, C. J., 1991, ApJ, 375, 613.
Ginzburg, V. L. and Syrovatskii, S. I., 1964, The Origin of Cosmic Rays (Oxford: Pergamon Press, 1964).
Ginzburg, V. L., and Syrovatskii, S. L., 1965, Ann. Rev. Astron., 3, 297.
Gordon, M. A., and Burton, W. B., 1976, ApJ, 208, 346.
Harris, M. J., Share, G. H., and Messina, D. C., 1995, ApJ, 448, 157.
Hartman, R. C., Kniffen, D. A., Thompson, D. J., and Özel, M. E., 1979, ApJ, 230, 597.
Hess, V. F., 1911, Physik Z., Vol. 12, 998.
Hess, V. F., 1912, Physik Z., Vol. 13, 1084.
Hunter, S. D., et al., 1996, ApJ, submitted.
Kerr, F. J., and Lynden-Bell, D., 1986, MNRAS, 221, 1023.
Kniffen, D. A., and Fichtel, C. E., 1974, High Energy Particles and Quanta in Astrophysics, ed. F. B. McDonald and C. E. Fichtel, MIT Press, Chapter 9, 428-470.
Kniffen, D. A. and Fichtel, C. E., 1981, ApJ, 250, 389.
Kniffen, D. A., Hartman, R. C., Thompson, D. J., and Fichtel, C. E., 1973, ApJ, 86, L105.
Koch, H. W. and Motz, J. W., 1959, Rev. Mod. Phys., 31, 920.
Kraushaar, W. L., Clark, G. W., Garmire, G. P., Borken, R., Higbie, P., Leong, C., Thorsos, T., 1972, ApJ, 177, 341.
Kulkarni, S. R. and Heiles, C., 1988, Galactic and Extragalatic Radio Astronomy, 2d ed., ed. G. L. Verschuur and K. I. Kellermann (New York; Springer), 95.
Leventhal, M., McCallum, C. J., Strong, P. D., 1978, ApJ (Letters), 225, L11.
Mahoney, W. A., et al., 1984, ApJ, 286, 578.
Mathis, J. S., Mezger, P. G., and Panagla, N., 1983, A&A, 128, 212.
Mayer-Hasselwander, H. A., Bennett, K., Bignami, G. F., Buccheri, R. D'Amico, N., Hermsen, W., Kanbach, G., Lebrun, F., Lichti, G. G., Masnou, J. L., Paul, J. A., Pinkau, K., Scarsi, L., Swanenburg, B. N., Wills, R. D., 1980, Annals of the New York Academy of Sciences, Proc. of the 9th Texas Symp., ed. J. Ehlers, J. J. Perry, and M. Walker, 336, 211.

Mihalas, D., and Binney, J., 1981, Galactic Astronomy, Structure and Kinematics, 2d ed. (San Francisco; Freeman).
Morris, D. J., 1984, Geophys. Res., 89, A12, 10685.
Parker, E. N., 1966, ApJ, 145, 811.
Parker, E. N., 1969, Space Sci. Rev., 9, 651.
Parker, E. N., 1977, "Cosmic Ray Propagation and Containment," in The Structure and Content of the Galaxy and Galactic Gamma Rays, NASA CP-002, (U.S. GPO, Wash, DC), p. 283.
Purcell, W. R., et al., 1993, ApJ, 413, L85.
Purcell, W. R., et al., 1994, AIP Conference Proceedings, 280, 70.
Ramaty, R., 1996, A&A (Third Compton Symposium), to be published.
Ramaty, R., and Lingenfelter, R. E., 1979, Nature, 278, 127.
Ramaty, R., and Lingenfelter, R. E., 1994a, in "High Energy Astrophysics, Models and Observations from MeV to EeV," ed. J. M. Matthews, World Scientific Publishing Co., p. 32.
Ramaty, R., and Lingenfelter, R. E., 1994b, in the "Analysis of Emission Lines", ed. R. E. Williams and M. Livio, Proceedings of the Space Telescope Science Institute Symposium, Baltimore, May 16-18, 1994, Cambridge University Press, 180.
Ramaty, R., 1996, A&A (Third Compton Symposium), to be published.
Share, G. H., Kinzer, R. L., Kurfess, J. D., Messina, D. C., Purcell, W. R., Chupp, E. L., Forrest, D. J., and Reppin, C., 1988, ApJ, 326, 717.
Simpson, J. A., 1983, Elemental and Isotopic Composition of the Galactic Cosmic Rays, Ann. Rev. Nucl. Part. Sci., 33, 323.
Solomon, P. M., and Sanders, D. B., 1980, "Giant Molecular Clouds as the Dominant Component of Interstellar Matter in the Galaxy," Molecular Clouds in the Galaxy, ed. P. M. Solomon and M. Edmunds (Oxford; Pergamon Press, 1980), p. 41.
Stecker, F. W., 1971, Cosmic Gamma Rays (Baltimore, MD: Mono Book Corp., 1971).
Stecker, F. W., 1979, ApJ, 228, 919.
Strong, A. W., et al., 1994, A&A, 292, 82.
Tmmes, F. X., Woosley, S. E., Hartman, D. H., and Hoffman, R. D., 1996, ApJ, 464, 332.
Tueller, J. et al., 1995, A&A Sup. (Third Compton Symposium), to be published.

CHAPTER 6

GALACTIC COMPACT OBJECTS

6.1 Introduction

Since the launch of the Compton Gamma-Ray Observatory, gamma-ray astronomy has emerged as a truly observational science, not only of the diffuse radiation, but also of compact objects, particularly in the area of extragalactic objects with the detection of a large number of active galaxies, as will be seen in Chapter 8. Although the growth in the number of observed compact objects which are clearly Galactic has not been as great, both the positive detections and the constraints on those not seen have provided new insight into our understanding of the most energetic processes occurring in association with Galactic astrophysical objects.

6.2 Individual Neutron Stars and Pulsars

The discovery of radio pulsars by Hewish et al. (1968) initiated an extensive search for these objects. These investigations have led to the detection of approximately 600 radio pulsars whose periods range from about a millisecond to several seconds. However, in spite of extensive searches, only 3 have been seen as pulsed radiation in the optical region (one of these after its gamma-ray identification). It was somewhat of a surprise that even a modest number (7 for certain, as of this time) of these pulsars would be observable in gamma rays. For the known pulsars, there is much more energy being released in the form of gamma rays, more than 10^{33} to 10^{35} ergs s^{-1} than in the radio region where typically 10^{31} ergs s^{-1} is being emitted for those pulsars also seen in the gamma-ray region.

Almost immediately after their discovery, pulsars were proposed to be associated with neutron stars (Gold, 1968), and this relationship

is now generally accepted. The large release of energy, the very fast period, and the remarkably small variation of the period seemed to dictate that the pulsed radiation must be from a massive object of small size. The very short length of the individual pulses indicates that the size of the emitting region is associated with something substantially smaller than normal stellar dimensions, or even white dwarfs. On the other hand, the periods, in general, are constant to one part in 10^8 or greater, indicating a massive object rather than a plasma phenomenon. If the period of the pulse is associated with a rotating body, then the object must be a neutron star, because the surface cannot move faster than the speed of light. Further, the period is probably too short to be associated with an oscillating phenomenon.

Long before the observation of pulsars, theoreticians had predicted the existence of neutron stars. Baade and Zwicky (1934) proposed the concept of a neutron star. Pacini (1968) showed that a rapidly rotating neutron star with a strong magnetic field would produce electric fields, which could provide the power for a supernova remnant like the Crab Nebula. It was, and still is, believed that neutron stars are formed by the violent collapse of the inner core of a star in a supernova explosion, wherein the outer part of the star is ejected. Neutron stars are believed to have masses somewhat larger than one $M_\odot$, radii of about 10 km, and a mean density of 10^{14} gm cm^{-3}. The observed association of some of the fastest pulsars with known supernovae remnants provided further evidence to support the identification of these objects as neutron stars.

It is now generally agreed that the ultimate source of the radiated energy is the rotational energy of the neutron star. As energy is radiated from the neutron star, the pulsar must slow down, and, indeed, the radio pulsars generally seem to have periods increasing with time. For PSR B0531+21, the change in rotational energy based on the pulsar period change is about -2×10^{38} ergs s^{-1}. In general, slower pulsars have smaller rates of change. Another generally accepted idea is that, as the neutron star was formed, its surface magnetic field strength became about 10^{12} gauss on the basis of calculations associated with the collapse of a body of 1 $M_\odot$ having an initial surface field similar to that of an ordinary star. Hence, the picture that has evolved for a pulsar is a rapidly rotating neutron star with a high surface magnetic field, configured such that there is beaming of the radiation.

With the strong magnetic field expected at a neutron star, coupled with the rapid rotation, large electric fields can be produced which may, in principle, accelerate charged particles to high energies (10^{12} eV or more). In this environment, electrons, in particular, can interact to produce gamma rays in several ways. One of these is synchrotron radiation. Another source of gamma rays is the inverse Compton radiation. Finally, there is curvature radiation; electrons moving along a curved magnetic field line can be treated with classical electrodynamics (Jackson, 1975) using the field line curvature as the radius for the synchrotron formula. There is still a great variety of opinions regarding the details of the pulsar model, including the specific manner and location of the relativistic particle acceleration, and no consensus seems yet to have evolved. It is very likely that all of these production mechanisms play a role at some level.

The value of $P/(2\dot{P})$ is generally accepted as being a moderately good estimate of pulsar age (e.g., Ostriker and Gunn, 1969; Ruderman and Sutherland, 1975), where P is the pulsar period. For the Crab pulsar, this apparent age is reasonably consistent with the known age of the pulsar. Consistency between apparent and true age may not hold for pulsars with $P/(2\dot{P})$ greater than 10^6 years (Lyne et al., 1975); also, some sudden changes in pulsar period have been observed. The radio luminosity of pulsars older than 10^6 years appears to decrease relatively rapidly, and various theoretical reasons for this turnoff have been suggested (Gunn and Ostriker, 1970; Ruderman and Sutherland, 1975; Lyne et al., 1975).

Assuming that the observed pulsar radiation comes from the rotational energy, and that the term containing the change in the moment of inertia is negligible compared to the term containing the period change, the rotational energy loss is given by the expression

$$\frac{dE_R}{dt} \cong I\,\Omega\dot{\Omega} = \left(\frac{4\pi^2 I}{p^3}\right)\dot{p}, \qquad (6\text{-}1)$$

where I is the pulsar moment of inertia, and Ω is the angular frequency of the pulsar. Using the measured values of p and $\dot{p}$ and a value for I of 10^{45} g cm^2 (Taylor and Manchester, 1975), an estimate of the values of (dE_R/dt) can be made. The values of the

rotational energy loss thus calculated are upper limits, assuming that all the energy that is lost comes from the rotational energy.

No consensus yet exists with regard to the exact model for the pulsed radiation, although ideas are plentiful. Two models appear to have emerged as the more promising at the moment. One is the polar cap model (See, e.g., Daugherty and Harding, 1994, and references therein). In this model the acceleration occurs near the neutron star surface at the magnetic poles—a picture that seems to follow from studies of radio emission properties. The outer gap model (e.g., Cheng, Ho, and Ruderman, 1986a, b; Cheng and Ding, 1994) assumes that the primaries are accelerated in the outer magnetosphere, in vacuum gaps that develop along null charge surfaces. In both models, electromagnetic cascades develop, creating many electron-positron pairs which radiate most of the observed gamma rays. In the former model, magnetic pair production cutoffs around a few to several GeV follow naturally from electron-photon cascades in the strong fields. In the latter, photon-photon absorption provides a cutoff in a similar energy range. Discussion of these two theoretical models will continue after a discussion of the experimental results.

Seven pulsars have now been seen in gamma rays, and three others are possiblities. The seven certain ones, plus B0656+14 because of its special interest, are given in Table 6.1, along with the Compton Observatory instruments that have seen them. Some of their observed and deduced parameters are given in Tables 6.2 and 6.3, which are modified versions of those in Ramanamurthy et al. (1995b). References to the papers giving the gamma-ray results are given in the following subsections on the individual pulsars. The pulsars that have been seen in gamma rays are generally the ones that are the most likely candidates. The four pulsars with a greater luminosity, that have not been seen, either are further away (B0540-69 and B1823-13) or have longer periods (B2334+61 and B0656+14). Although seven have been seen in high-energy gamma rays ($E > 100$ MeV), only four have been seen in low-energy gamma rays, namely the Crab (B0531+21), Vela (B0833-45), Geminga, and B1509-58. The last one is the only one observed in low-energy gamma rays, but not seen in high-energy gamma rays; it is also further away than any pulsar seen in high-energy gamma rays. The seven pulsars observed in gamma rays will now be discussed individually, followed by a summary of some of the general properties and implications.

Table 6.1.

Spin Down Pulsars Seen by the Compton Observatory

Pulsar	Period	Compton Observatory Instrument detecting the pulsars
CRAB	33 ms	EGRET, BATSE, COMPTEL, OSSE
GEMINGA	237 ms	EGRET, COMPTEL
VELA	89 ms	EGRET, BATSE COMPTEL, OSSE
B0656+14	385 ms	EGRET
B1055-52	197 ms	EGRET
B1509-58	151 ms	BATSE, COMPTEL, OSSE
B1706-44	102 ms	EGRET
B1951+32	40 ms	EGRET

PSR B0833-45 (The Vela Pulsar)

This pulsar, often called the "Vela" pulsar, was identified in the SAS-2 data (Thompson et al., 1975, 1977) not only as the brightest pulsar observed at the Earth in gamma rays, but also as the brightest discrete source in the high-energy gamma-ray sky. The two most striking features are the two gamma-ray pulses, as opposed to one in the radio region, and the fact that neither gamma-ray pulse is in phase with the radio pulse. These features were confirmed by the data obtained later from the COS-B satellite gamma-ray telescope (Buccheri et al., 1978), and the radiation has now been studied in much greater detail by EGRET (Kanbach et al., 1994). If this result were not enough to complicate attempts to find a satisfactory theoretical model following the detection of PSR B0833–45 in gamma rays, two peaks were found in the optical region by Wallace et al. (1977); neither peak was in phase with either the gamma-ray or the radio peaks. Subsequently, the periodic emission was seen

Table 6.2.

Observed Parameters of the Gamma-Ray Pulsars

Pulsar	l	b	P (s)	P (10^{-15} s/s	γ*	F_γ (erg cm^{-2} s^{-1}) (E_γ>100 MeV)
Crab	184.6°	-5.8	0.033	421	2.15	1.0×10^{-9}
Geminga	195.1°	4.3	0.237	11.0	1.50	3.7×10^{-9}
Vela	263.6°	-2.8	0.089	125	1.70	7.1×10^{-9}
B1055-52	286.0°	6.6	0.197	5.83	1.18	4.2×10^{-10}
B1509-58	320.3°	-1.2	0.150	1540	**	**
B1706-44	343.1°	-2.7	0.102	93	1.72	8.3×10^{-10}
B1951+32	68.8°	2.8	0.040	5.85	1.74	2.4×10^{-10}

* γ is the exponent of the gamma-ray differential energy spectrum.

** B1509-58 was not seen by EGRET at E > 100 MeV.

Table 6.3.

Inferred Parameters of the Gamma-Ray Pulsars

Pulsar	Dist. (kpc)	Age (yr)	B(G)	E(erg s^{-1})	L_γ(ergs^{-1}) $E_\gamma > 100$MeV	efficiency*,η $E_\gamma > 100$MeV
Crab	2.0	1.3×10^{3}	3.8×10^{12}	4.5×10^{38}	3.9×10^{34}	0.00009
Geminga	0.25	3.4×10^{5}	1.6×10^{12}	3.3×10^{34}	2.2×10^{33}	0.068
Vela	0.50	1.1×10^{4}	3.4×10^{12}	7.0×10^{36}	1.7×10^{34}	0.0024
B1055-52	1.5	5.3×10^{5}	1.1×10^{12}	3.0×10^{34}	9.3×10^{33}	0.31
B1509-58	4.4	1.5×10^{3}	1.5×10^{12}	1.8×10^{37}	**	**
B1706-44	1.8	1.7×10^{4}	3.1×10^{12}	3.4×10^{36}	2.6×10^{34}	0.0077
B1951+32	2.5	1.1×10^{5}	4.9×10^{11}	3.7×10^{36}	1.4×10^{34}	0.004

* η is proportional to the beaming factor assumed.

** B1509-58 was not seen by EGRET at E > 100MeV.

weakly in the x-ray region by Ögelman, Finley, and Zimmerman (1993), and the pattern of the radiation as a function of phase is different again from any of the others. The Vela pulsar also has been seen by the COMPTEL instrument on the Compton Observatory in the 1 to 30 MeV region (Hermsen et al., 1993) and in the 0.06 to 9 MeV region by the OSSE instrument, also on the Compton Observatory (Strickman et al., 1993).

The distribution of high-energy gamma-ray photons as a function of phase obtained by EGRET is shown in Figure 6.1. There are two clear, sharp peaks, and, in between in phase from about 0.15 to 0.50, there is a level of radiation that is well above the local Galactic diffuse background. From about 0.60 to 0.05 in phase, there is no measurable emission above the level of the Galactic diffuse radiation. There has been no detectable difference in this phase distribution from one EGRET observation to the next (Kanbach et al., 1994).

The gamma-ray energy spectrum obtained with the instruments OSSE, COMPTEL, and EGRET on the Compton Observatory is shown in Figure 6.2. From approximately 4 MeV to 4 GeV, 3 decades in energy, the spectrum seems to be well fit by a power law in energy with a slope of -1.7. Below 4 MeV, there is a clear departure from this spectral shape. There is also a departure above 4 GeV in the 4 to 10 GeV region, the highest-energy interval for which EGRET had a detectable flux. The Vela pulsar is sufficiently intense, and EGRET has sufficient sensitivity to allow the data to be divided into several intervals depending on phase. These intervals are shown in Figure 6.1 and the spectra are shown in Figures 6.3a and 6.3b as a function of phase. One sees at once that the spectra are different. The clean power-law spectrum seen in Figure 6.2 is actually the sum of several spectra of different slopes from the different phase intervals. Notice that the two peaks (P1 and P2) have relatively flat spectra, as do the adjacent sections, but the intervening area (I1 and I2) have relatively soft spectra.

Models of the emission from pulsars in general, and Vela in particular, have been presented by many authors. The two principal models are those noted earlier; namely, the polar cap (e.g., Harding and Daugherty, 1993) and the outer gap (e.g., Ho, 1993). The general spectra of both models are in qualitative agreement with the observations and can probably be made to agree in detail with

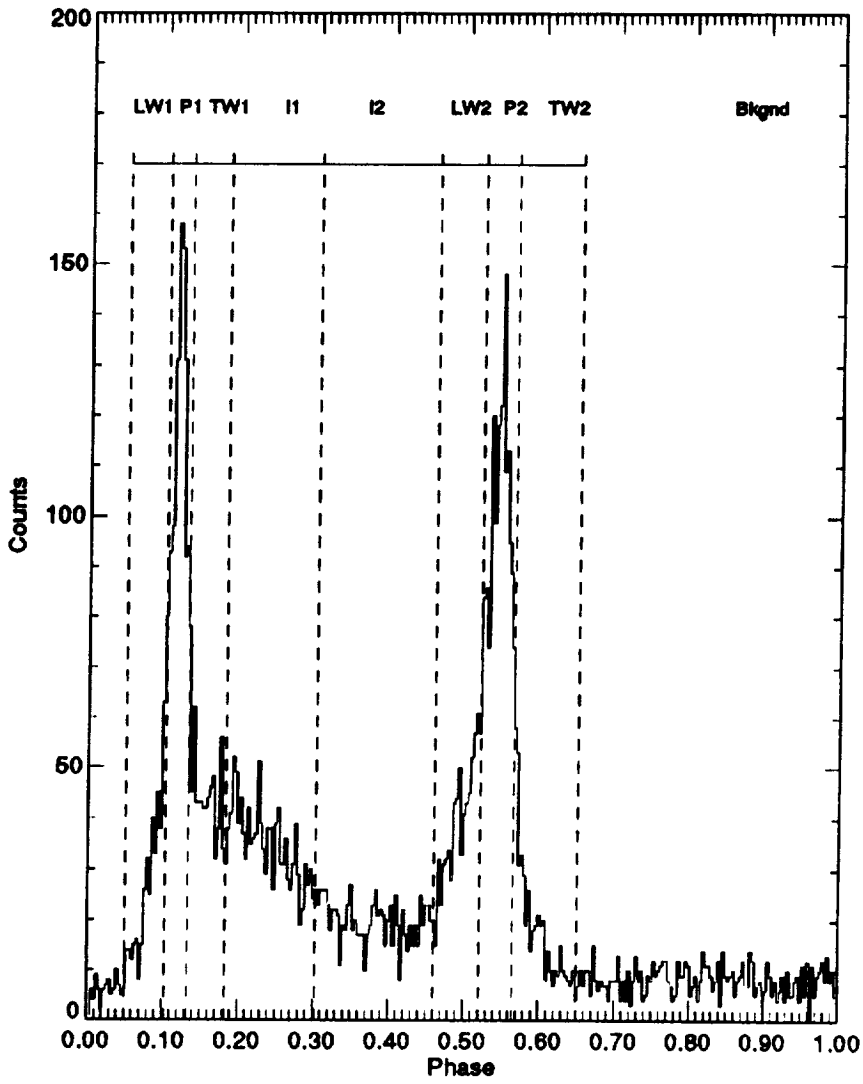

Figure 6.1. Light curve above 100 MeV from the sum of all available Vela EGRET observations through 1993. The figure is reproduced with the permission of Kanbach et al. (1994).

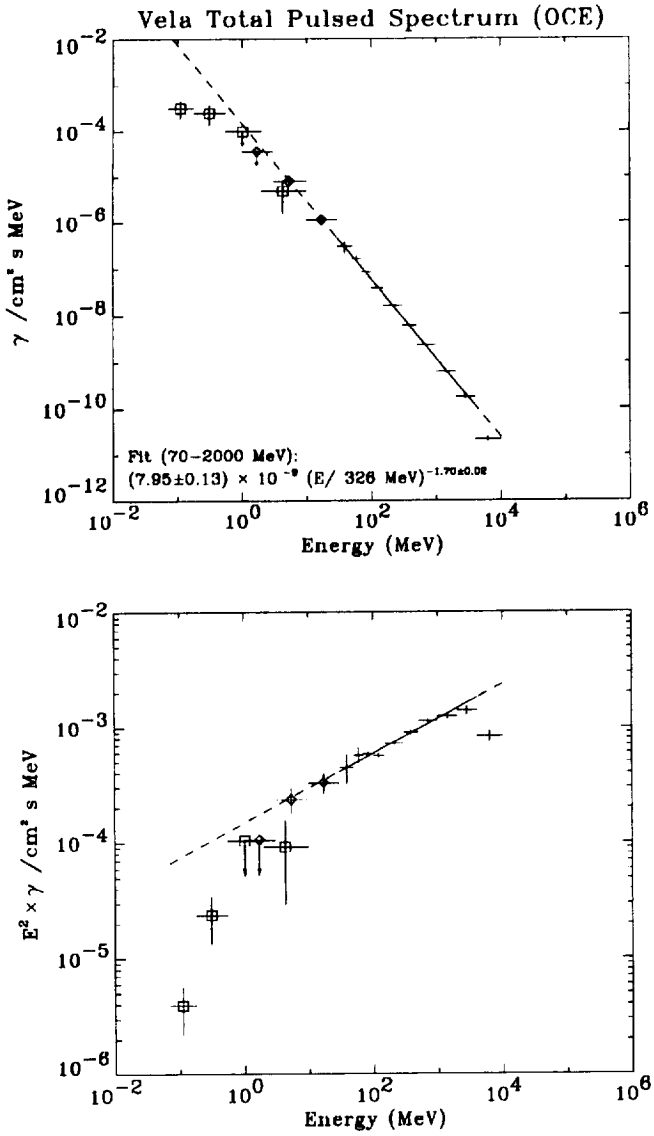

Figure 6.2. Spectrum of the pulsed emission from PSR B0833-45 as observed by EGRET (no symbols: Kanbach et al., 1994); COMPTEL (diamonds: Hermsen et al., 1993); and OSSE (squares: Strickman et al., 1993). The figure is reproduced with the permission of Kanbach et al. (1994).

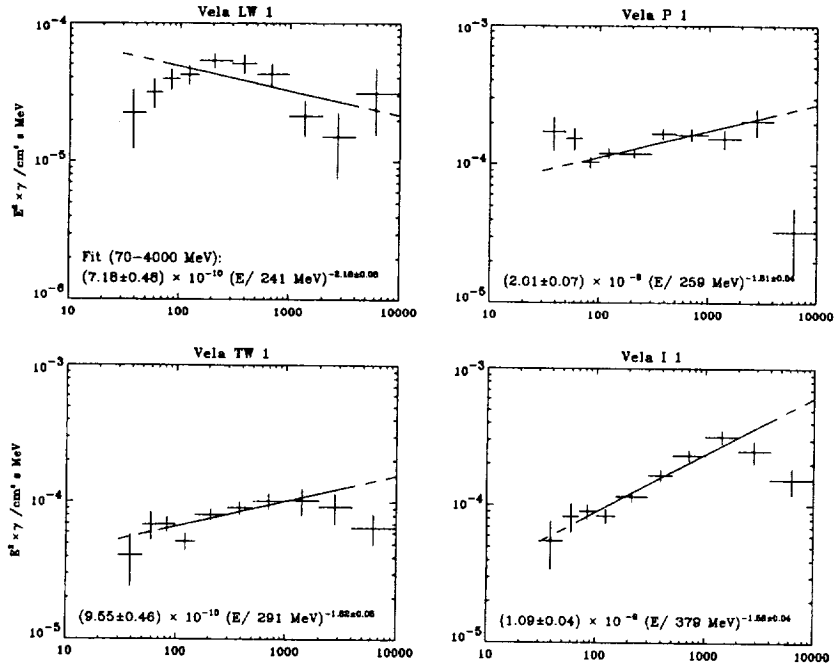

Figure 6.3a. Time-averaged spectra of four of the eight phase components defined in the pulsed emission of PSR B0833-45 measured with EGRET. The figure is reproduced with the permission of Kanbach et al. (1994).

appropriate adjustment of the parameters. Several detailed features found by Harding and Daugherty (1993) can be discerned in the results of the phase-resolved spectroscopy, particularly the emission peaks composed of rather hard spectra cores with softer spectra in the wings as a cascade model would naturally predict. The polar cap geometry also predicts (because of magnetic-field line bending) that the leading and trailing wings of the emission peaks should have slightly different spectra, with the trailing component extending to higher energies than the leading component. This is the trend for the observed spectra.

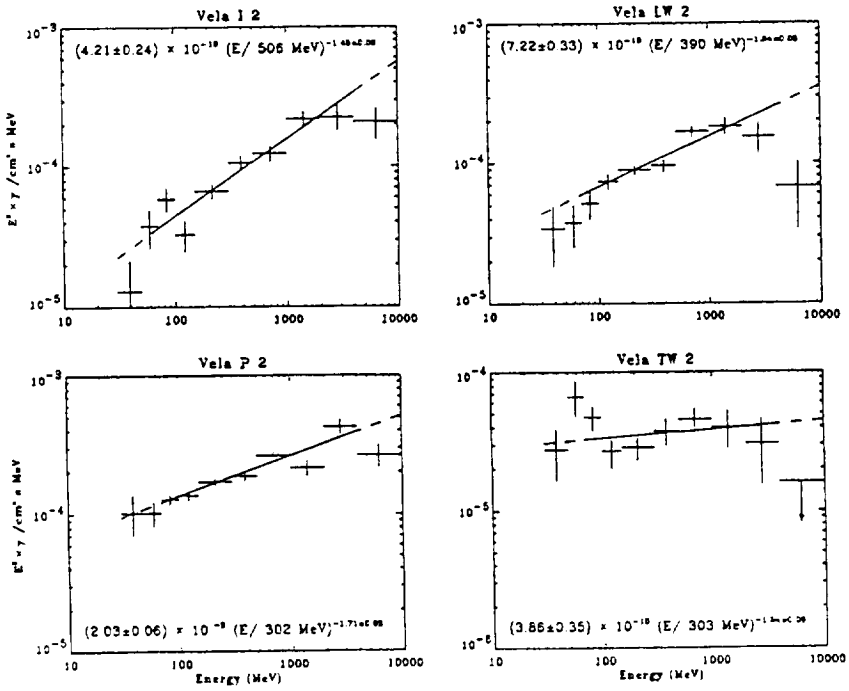

Figure 6.3b. Time-averaged spectra of four of the eight phase components defined in the pulsed emission of PSR B0833-45 measured with EGRET. The figure is reproduced with the permission of Kanbach et al. (1994).

PSR B0531+21 (The Crab Pulsar)

This pulsar was the first gamma-ray pulsar reported (Browning et al., 1971), even though it is not the strongest. The SAS-2 and COS-B satellite gamma-ray telescopes provided details of the high-energy gamma radiation (Kniffen et al., 1974; Clear et al., 1987). Recently, with the instruments on the Compton Gamma-Ray Observatory, gamma-ray measurements exist for over 5 decades in energy.

In one respect, the observational picture for the pulsar PSR B0531+21 in the Crab nebula is simpler to describe. It is seen with the double pulsed structure in the radio, optical, x-ray, and gamma-ray regions. Further, the pulses in each wavelength are in phase as shown in Figure 6.4. The Crab Pulsar is unique in this respect.

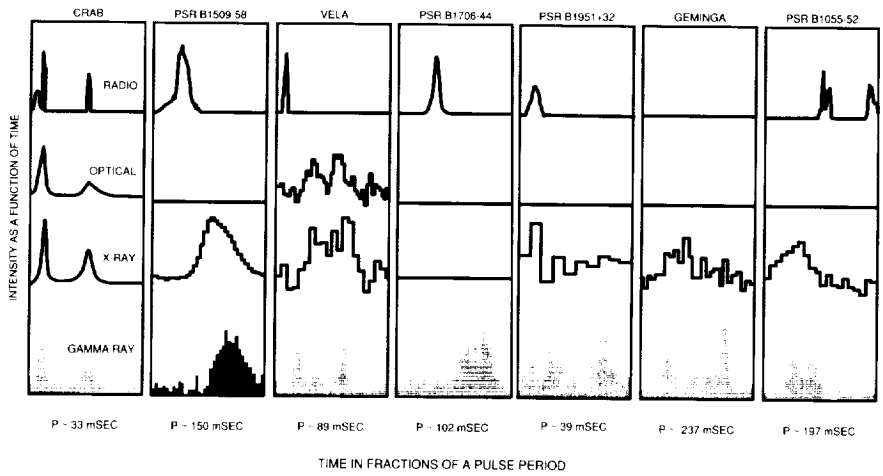

Figure 6.4. The seven gamma-ray pulsars detected with high certainty. For each pulsar, the brightness variation during one rotation of the neutron star is shown for four different wavelength ranges: radio, optical, x-rays, and gamma rays. References: Crab: Manchester, 1971 (radio); Groth, 1975 (optical); Rappaport et al., 1971 (x-ray); Nolan et al., 1993 (gamma ray). B1509-58: Ulmer et al., 1993a (radio and gamma); Kawai et al., 1991 (x-ray). Vela: Kanbach et al., 1994 (radio and gamma ray); Wallace et al., 1977 (optical); Ögelman, 1994 (x-ray). PSR B1706-44: Thompson et al., 1992. PSR B1951+32: Kulkarni et al., 1988 (radio); Safi-Harb et al., 1995 (x-ray); Ramanamurthy et al., 1995 (gamma ray). Geminga: Bertsch et al., 1992 (gamma ray); Halpern and Ruderman, 1993 (x-ray). PSR B1055-52: Fierro et al., 1993 (radio and gamma ray); Ögelman and Finley, 1993 (x-ray). Summary compiled by Thompson (1996a).

For every other pulsar detected in gamma rays, the pulses at different wavelengths are not in phase. The common in-phase double peak feature suggests the same mechanism for the pulsed radiation at all wavelengths for PSR B0531+21, which is perhaps to be expected, since it is by far the youngest known gamma-ray pulsar, approximately 10^3 years old. The older Vela pulsar, PSR B0833−45, apparently has a dominant high-energy component. In fact, whereas the pulsed luminosity ratio, L(PSR B0531+21)/L(PSR B0833−45), is about 5 above 100 MeV, it is almost 10^4 in the optical range. The pulsed radiation is seen with GRO instruments in gamma rays in the low (Ulmer et al. 1994), medium (Much et al., 1995a), and high (Mayer-Hasselwander et al.,

1994) energy range. The combined energy spectrum from OSSE, COMPTEL, and EGRET is shown in Figure 6.5 (Ulmer et al., 1995). As one can see, the spectrum is well fit by a smooth curve with a slowly changing slope, although there may be a dip near 10 MeV, which could be explained in the context of the outer gap model (Much et al., 1995b; Ho, 1993). The first pulse energy spectrum (Ulmer et al., 1993a) particularly shows a greater tendency towards the outer gap module.

The Crab nebula also has strong unpulsed, or constant emission. It is the only pulsar observed in high-energy gamma rays that has measurable unpulsed emission (Nolan et al., 1993; Kanbach et al., 1994; Mayer-Hasselwander et al., 1994). Ramanamurthy et al. (1995a) have studied the Crab emission when observed by EGRET in the 30 MeV to 60 GeV range during the period from April 1991 to early 1994. The pulsed and unpulsed emissions from the Crab behave differently with respect to time variability. The pulsed component does not show any variability in any of the energy bands studied, 30–100, 100–300, and $\geq$ 300 MeV. The unpulsed component, on the other hand, appears to be variable, with the respective reduced χ^2 values being 2.38, 2.19, and 1.39. The overall probability for this to happen under the null hypothesis is as low as 2.6×10^{-3}.

Consider next the energy spectra of these two components, shown in Figure 6.6, as determined from the EGRET data. One can see that the unpulsed/total ratio decreases with energy up to about 1000 MeV and then increases. The steep increase of the unpulsed component at lower energies ($E \leq 100$ MeV is interpreted by De Jager et al. (1996) in terms of the dominance by synchrotron emission by $e^{\pm}$ in the Crab nebula. Below 20 MeV, the unpulsed radiation spectrum flattens (Much et al., 1995a) and the unpulsed-to-total ratio drops to 0.2 or less. The trend seen at the highest EGRET energies is confirmed by the Whipple group at TeV energies where the unpulsed/total ratio is ~1 (see Vacanti et al., 1991), i.e., there is no measurable pulsed component. The higher-energy component of the constant emission is generally interpreted as inverse Compton radiation. The combined data from EGRET and Whipple and other observations are shown in Figure 6.7, along with the theoretical curves of De Jager et al. (1996).

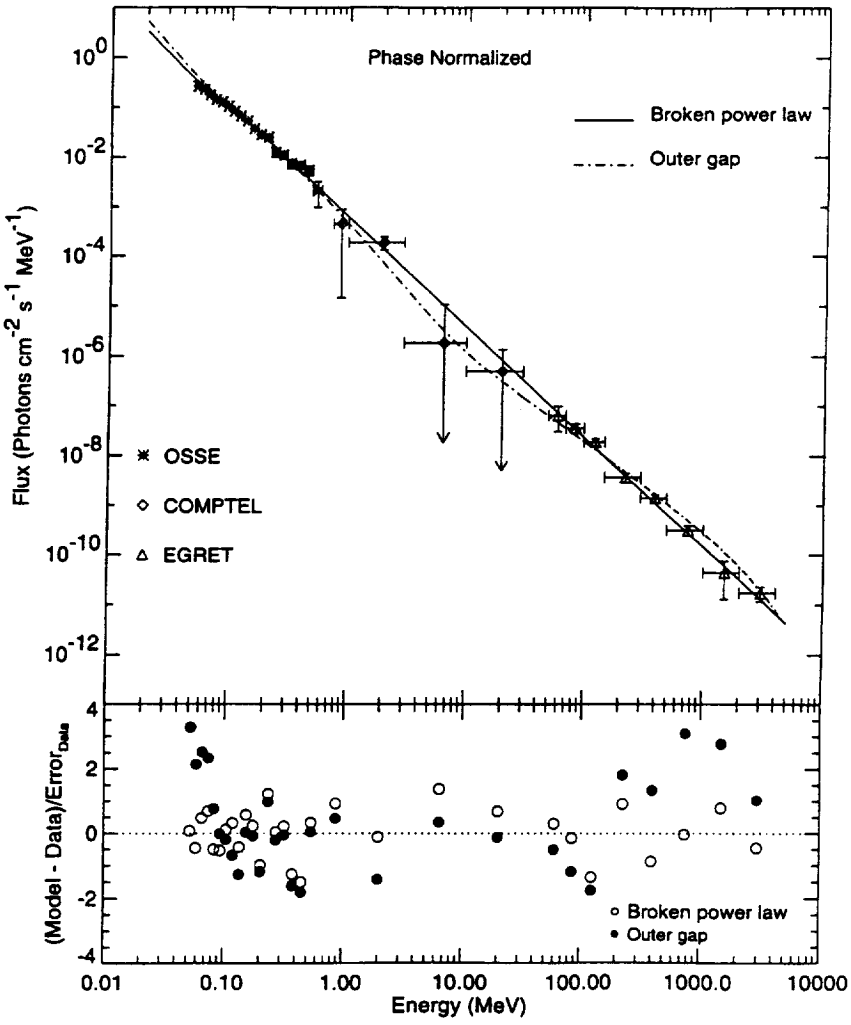

Figure 6.5. The phase-averaged, total Crab-pulsed spectrum from the data of OSSE, COMPTEL, and EGRET. The figure is reproduced with the permission of Ulmer et al. (1995); the solid line is the best-fit, broken power-law spectrum , and the dashed line is the best fit from the outer gap model. The lower figure shows the difference between the best fit and the data in units of the 1σ uncertainties in the data.

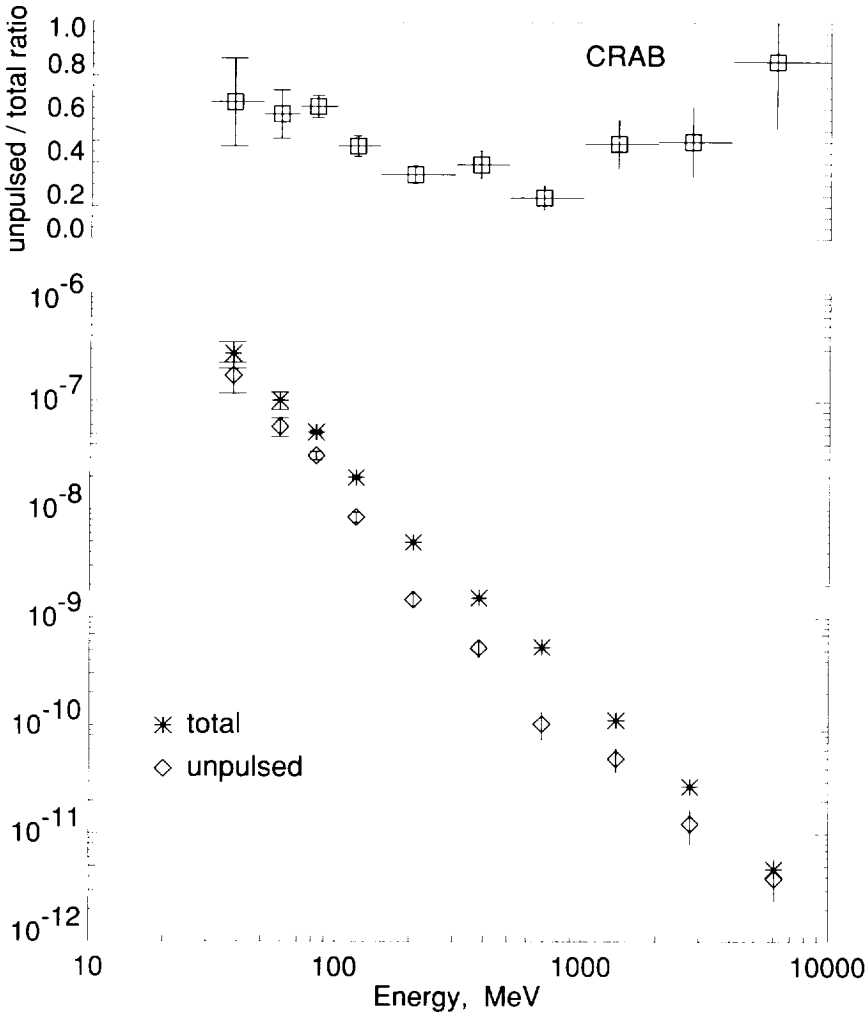

Figure 6.6. Differential unpulsed (diamonds) and total fluxes (asterisks) from the Crab determined from EGRET data, in units of quanta cm^{-2} s^{-1} MeV^{-1}, Their ratios are shown as a function of energy in the upper figure. This figure is reproduced with the permission of Ramanamurthy et al. (1995a).

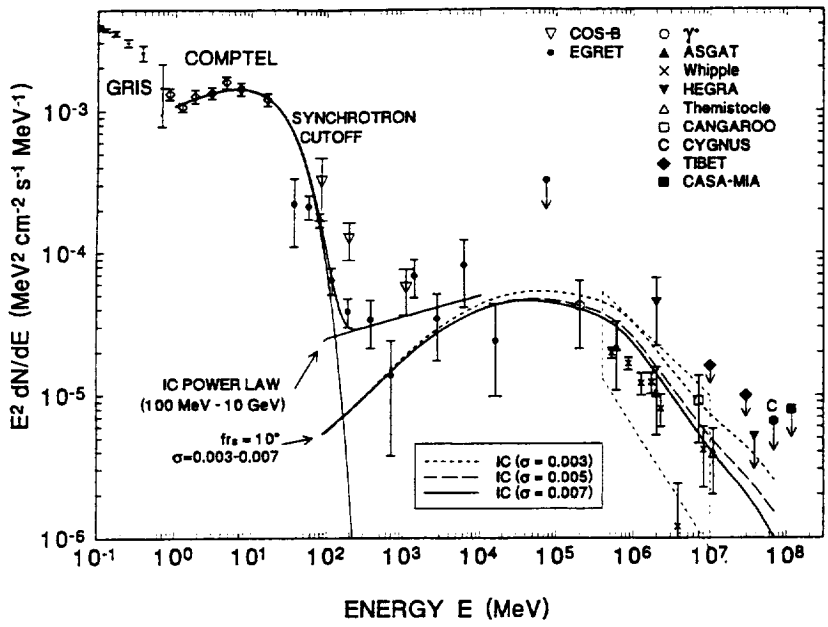

Figure 6.7. The Crab nebular unpulsed spectrum in terms of E^2 times the differential photon spectrum dN/dE for the energy range 100 keV to 300 TeV. The GRIS spectrum (100 to ~700 keV) is from Bartlett (1994) and the COMPTEL data, from Much et al. (1995a). The three COS-B points are from Clear et al. (1987) (50 - 200 MeV) and Mattox, De Jager and Harding (1992) (150-300 MeV; 300 MeV - 5 GeV). The EGRET results (De Jager et al., 1996) for the 30 MeV to 100 GeV range are shown as solid circles. The best-fit synchrotron cutoff and IC spectra of their work are indicated as "SYNCHROTRON CUTOFF" (1 to ~200 MeV) and "IC POWER LAW" (100 MeV to 10 GeV). The TeV results for the 0.2 TeV to 100 TeV range are summarized with references in De Jager et al. (1996). The dashed trapezium between 0.4 and 10 TeV (Weekes et al, 1994) indicates the systematic error in the Whipple spectrum (The Whipple data are indicated as crosses). The three curved lines between 100 MeV and 100 TeV (corresponding to three different σ-parameters as indicated in the Legend) represent the inverse Compton spectra corresponding to $fr_S = 10"$. The figure is reproduced with the permission of De Jager et al. (1996).

Regarding gamma-ray lines, although there were some early reports of detections, the OSSE group (Ulmer et al., 1994) found no evidence for lines and set limits well below earlier tentative reports.

Geminga

The history of this source is interesting in itself, reading a bit like a mystery novel wherein the proof of its nature came approximately 20 years after its discovery. The gamma-ray source Geminga (2CG195+04) was first seen by the SAS-2 satellite (Kniffen et al., 1975; Fichtel et al., 1975; Thompson et al., 1977), in observations made in 1972 and 1973. Then, as now, it was one of the three strongest high-energy gamma-ray sources in the sky, the other two being the Crab and Vela pulsars, PSR B0531+21 and PSR B0833+45. It was seen at roughly the same intensity both times that SAS-2 viewed the region of the sky where it was located. Subsequently, the COS-B satellite observed the source five times between 1975 and 1982 (Bennett et al., 1977; Masnou et al., 1981; Hermsen et al., 1990), again with essentially the same intensity each time.

After the original SAS-2 observations, Thompson et al. (1977) argued on the basis of the strength, its location, the absence of observations at other wavelengths, and other factors, that it was highly probable that the source was Galactic, and possible that it was an undiscovered pulsar. For over two decades, attempts to find a radio counterpart failed, even after the COS-B results provided a better location (Spoelstra and Hermsen, 1984). Other theories were put forward over the years, but the mystery remained.

The detection by the Einstein Observatory (Bignami, Caraveo, and Lamb, 1983) of the x-ray source 1E 0630+178 with unusual properties in the Geminga COS-B error box provided the first real hope of a counterpart at another wavelength. Bignami et al. (1987) then reported the detection of an optical object inside the Einstein error circle. Halpern and Tyler (1988) confirmed that this was a very blue object. They argued that this result and other factors support the original proposal of the SAS-2 group, that the object is a neutron star and further that the large gamma-ray to x-ray flux ratio requires that the gamma rays be produced far from the surface of the neutron star, as is required in most pulsar models.

The discovery of a 237.0974 ms period in 1E 0630+178 (Halpern and Holt, 1992) provided virtually certain evidence that the source is a neutron star. It also yielded the information needed to search for periodicity in the gamma-ray data; otherwise, the search for a period of this magnitude in high-energy gamma-ray data is difficult or impossible, because the paucity of photons leads to a substantial probability of finding chance apparent structure in the phase diagram. On learning of the periodicity and upper limit to the period derivative, Bertsch et al. (1992) were quickly able to find that the Geminga radiation was pulsed at the period measured in the x-ray region. The pulse profile from this paper, which first showed the periodic high-energy gamma-ray emission is shown in Figure 6.8. The pulsar was found to be remarkably stable. Bertsch et al. (1992) determined $\dot{p}$ to be $(11.4 \pm 1.7) \times 10^{-5}$ s s^{-1}. With their determination of p and $\dot{p}$, it was possible to see very quickly that the Geminga radiation was pulsed in the COS-B and SAS-2 data (Mattox et al., 1992; Bignami and Caraveo, 1992) taken from 20 to 11 years earlier because of the very stable nature of this pulsar. Another unique aspect of this pulsar is that the phase separation between the two pulses is almost exactly one-half, being 0.50 ± 0.02, with the uncertainty being largely associated with how one defines the peaks. The intensity and shape of the peaks, although not exactly identical, are very similar. As of this writing, a pulsed radio emission from the source has not been found, even though it is one of the three strongest gamma-ray pulsars.

Using the EGRET data and all of the earlier data, Mayer-Hasselwander et al. (1994) determined the following parameters for the Geminga pulsar:

$$f = 4.2176749979 \text{ Hz}, \tag{6-2}$$

$$\dot{f} = -1.9526(7) \times 10^{-13} \text{ Hz s}^{-1}, \tag{6-3}$$

$$\ddot{f} = < 2 \times 10^{-23} \text{ Hz s}^{-1} \text{ s}^{-1}, \tag{6-4}$$

where the number in parenthesis is approximately the 1σ uncertainty on the last digit. Mayer-Hasselwander et al. (1994) further conclude that all the radiation from the Geminga source is found to be pulsed.

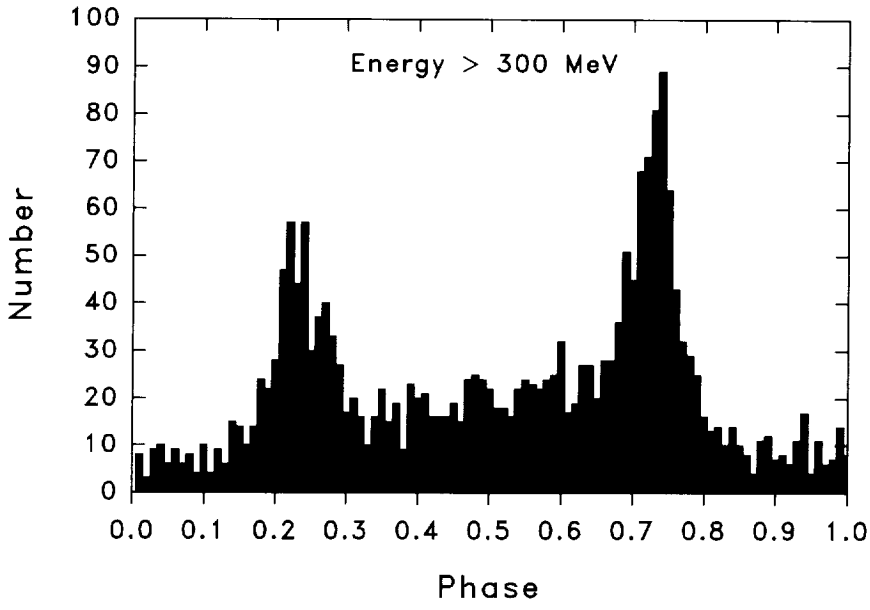

Figure 6.8. Light curve of gamma rays (E > 300 MeV) for the combined observation periods 22 April-7 May, 16-30 May, and 8-15 June 1991 for Geminga. The epoch period and period derivative values quoted in the text were used (Bertsch et al., 1992).

The total spectrum, with a very flat power-law index of -1.50 ± 0.08, does not exhibit a turnoff at low energies down to 30 MeV. At high energies, a pronounced roll off is observed above 2 GeV. The spectra for phase subintervals of the light curve vary in hardness between power-law exponents -1.22 ± 0.05 and -1.67 ± 0.23. Within the statistical accuracy available, the spectra for the phase subintervals show the same rollover at high energies as seen in the total spectrum. It has just been learned, as this book is going to press, that Geminga has been seen as a radio pulsar (Malofeev and Malov, 1997) at a frequency of 102.5 MHz. Its luminosity is quite weak.

PSR B1706-44

This pulsar has a period of 102 ms, a single, broad, high-energy gamma-ray peak, which might be two narrow peaks close together

(Thompson et al., 1992; Thompson et al., 1996b). The single radio peak comes earlier in phase than the gamma-ray emission. There is no evidence of an unpulsed component in the medium or high-energy gamma-ray region, but a constant component has been detected in the TeV range with the ground-based Cerenkov imaging telescope in Woomera, Australia (Kifune et al., 1995). All of the high-energy gamma-ray emission information for this source comes from the EGRET data.

The energy spectrum is unique among gamma-ray emitting pulsars in that there is a clear change in spectral slope of the phase-averaged spectrum at about one GeV, as shown in Figure 6.9. Such a spectral break of the phase-averaged spectrum was not predicted by any theory prior to its observation. One also notes from this figure that the observed spectrum extends to 30 GeV.

Between 100 MeV and 10 GeV, the observed energy flux is $(9.4 \pm 1.2) \times 10^{-10}$ erg cm^{-2} s^{-1} (Thompson et al., 1996b). For a distance of 2.4 kpc and a beaming factor of 0.1, this represents 6×10^{34} erg s^{-1} or more than 1 percent of the spin-down energy loss of 3.4×10^{36} erg s^{-1}, comparable to that seen in the Vela pulsar for similar assumptions.

PSR B1055-52

EGRET has also detected pulsed high-energy gamma radiation from this pulsar (Fierro et al., 1993), which has a period of approximately 197 ms. The pulsed radiation from PSR B1055–52 has a very hard photon spectral index and a high efficiency for converting its rotational energy into gamma rays. No unpulsed emission was observed. The pulsed spectrum is well represented between 100 MeV and 4 GeV by a single power law of the form

$$\frac{dN}{dE_\gamma} = (1.91 \pm 0.25) \times 10^{-10} \left(\frac{E_\gamma}{908 \text{ MeV}} \right)^{-1.18 \pm 0.16}$$
$$\text{photons cm}^{-2}\text{ s}^{-1}\text{ MeV}^{-1}. \qquad (6\text{-}5)$$

Above 4 GeV, there is a sharp drop in the spectrum. The time-averaged spectrum has a normalization factor of 0.35. This spectrum is harder than the pulsed spectrum of any of the four other high-energy gamma-ray pulsars seen by EGRET and discussed thus

far. The time-averaged pulsed photon flux between 100 MeV and 4 GeV is determined to be (2.4 ± 0.4) x 10^{-7} photons cm^{-2} s^{-1}.

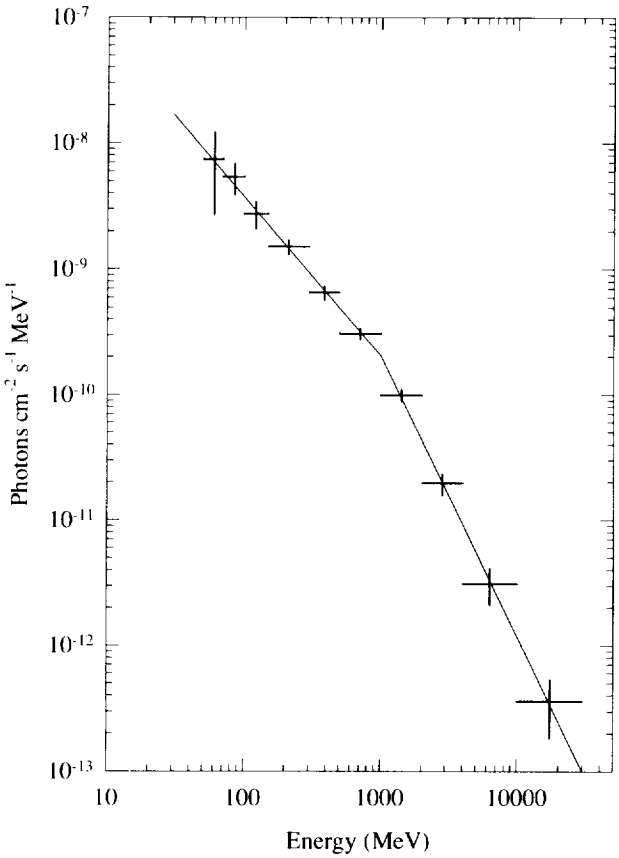

Figure 6.9. Phase-averaged energy spectrum of PSR B1706-44. The fit to a broken power law is described in the text. The uncertainties shown are statistical only. The figure is from Thompson et al. (1996b).

PSR B1951+32

This pulsar, with a period of 39.5 ms, is the most distant seen in high-energy gamma rays. (See Table 6.3.) The energy flux above 100 MeV is 2.4 x 10^{-10} ergs cm^{-2} s^{-1}. Two pulses are observed in the high-energy gamma-ray region, neither in phase with the radio peak. The pulsed radiation is consistent with a power law spectrum in energy with a photon spectral index of -1.74 ± 0.11.

PSR B0656+14

PSR B0656+14 has the lowest flux of any pulsar that has been detected in the high-energy gamma-ray energy range (Ramanamurthy et al., 1996), and its detection is still not considered absolutely certain. Special analysis techniques had to be used in order to see the source in the data (Hartmann, 1995). It is of special interest, however, because, as will be discussed later in this chapter, the intensity level measured is surprisingly low for one type of model and very high for the other; hence, it is a pulsar of unusual theoretical interest.

PSR B1509-58

This pulsar is unique among those seen in gamma rays. All the others were seen first in high-energy gamma rays as opposed to low-energy gamma rays, and only two of those seen in high-energy gamma rays are seen in the low-energy gamma-ray region. On the other hand, PSR B1509-58 was seen in very low-energy gamma rays by BATSE (Wilson et al., 1993), OSSE (Ulmer et al., 1993b; Matz et al., 1994), and COMPTEL, and has not been observed in high-energy gamma rays (Fichtel et al., 1994b). This pulsar has a double-humped structure, as seen in Figure 6.4. An examination of the observations and upper limits indicates that there must be a sharp downward break in the spectrum between about 8 and 30 MeV.

6.3 Gamma-Ray Pulsar Overview

At first consideration, the gamma-ray pulsars seem to differ considerably from each other. Except for the Crab, which has much the same appearance across the electromagnetic spectrum, the pulsars show a wide variety of pulse shapes and relative pulse arrival times. As an example, consider the Vela pulsar, which has a

single radio pulse, a double optical pulse, a double x-ray pulse, and a double gamma-ray pulse—no two of these seven pulses are exactly in phase. Three of the gamma-ray pulsars have a double-peak structure, but two do not. One of these pulsars is not seen in the radio, three are optically quiet, and one has no known x-ray pulsation. The magnetosphere of a pulsar can have a complex geometry; the different types of radiation may originate in different sites and with different directional patterns. No obvious similarities serve to link the light curves of these particular pulsars together. For a further comparison, see Fichtel and Thompson (1994).

However, there are some common features. All the gamma-ray pulsars radiate far more energy in the form of gamma rays than in radio waves. Because gamma rays are produced by direct interactions of the high-energy particles around the neutron stars, it is perhaps not surprising that they are important in the energy budget of the neutron stars. There are also other common features, which will now be considered, after giving two definitions in the next paragraph.

The observed energy flux F(in ergs/cm^2 s) of gamma radiation can be converted to a luminosity L at the source by assuming a solid angle Ω into which the pulsar beams the gamma radiation:

$$L = \Omega\, F d^2, \tag{6-6}$$

where d is the distance to the pulsar, which can be estimated from the dispersion of the radio signal as it propagates through the interstellar medium. The efficiency for the conversion of the rotational energy of the spinning neutron star into gamma radiation is the ratio of the gamma-ray luminosity to the total rate of energy loss of the pulsar E, as determined by Equation (6-l).

The high-energy gamma-ray pulsars that have been observed thus far show a trend in which the efficiency of gamma-ray production increases with apparent age. This effect can be explained either as an increasing current flow out of the neutron star with increasing age in the polar model gap (Harding, 1981), or as an increasingly large particle-acceleration region in the magnetosphere around the neutron star (Ruderman and Cheng, 1988).

In the polar gap model, there is predicted to be a linear increase with apparent age of the gamma-ray production efficiency multiplied by $\dot{p}^{-1.3}$. There is good agreement with the experimental observations as shown in Figure 6.10. Nel et al. (1996) have also looked at all the high-energy gamma-ray upper limits for other pulsars and find them to be near or above the line in Figure 6.10 for ages below 10^6 years. Beyond approximately this age, the production efficiency is expected to fall.

A possible corrallary of this concept is that the higher-energy pulsed radiation in the gamma-ray range, at least until the expected cut-off, is even more efficently produced than the lower-energy gamma radiation. If this should be true, the spectrum would be expected to be harder (the spectral slope less negative) in the high-energy

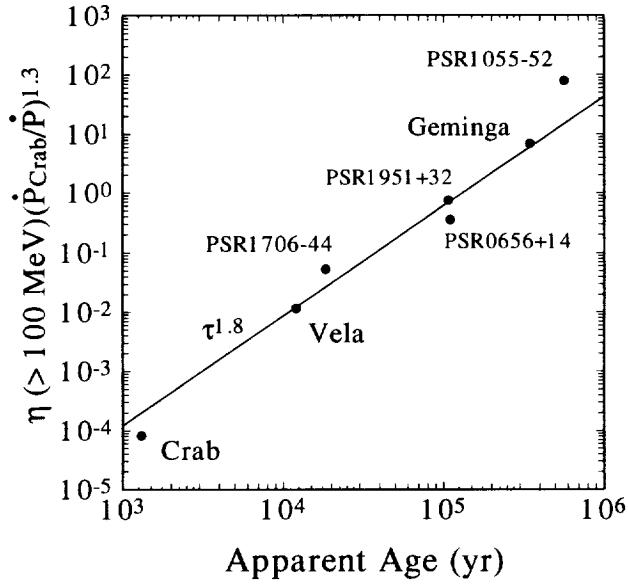

Figure 6.10. The efficiency of gamma-ray production multiplied by $\dot{p}^{-1.3}$ as a function of apparent age for pulsars seen in high-energy gamma rays (Harding, 1996). See text for discussion.

gamma-ray range as a function of apparent age, and this prediction indeed appears to be the case as illustrated in Figure 6.11.

In the polar gap model when the normalized predicted luminosities are compared to the observed ones, there is, with one exception, good agreement and there is no upper limit for high-energy pulsed gamma-ray emission below the expected level. See Figure 6.12. The one exception is B0656+14, whose observed intensity is well below that which is expected. Of the pulsars with observed high-energy gamma-ray emission, it is the one with the highest deduced magnetic field. It may be that the higher magnetic field is causing this effect, especially since the only pulsar seen at low gamma-ray energies and not seen by EGRET, namely B1509-58, has an even higher calculated magnetic field.

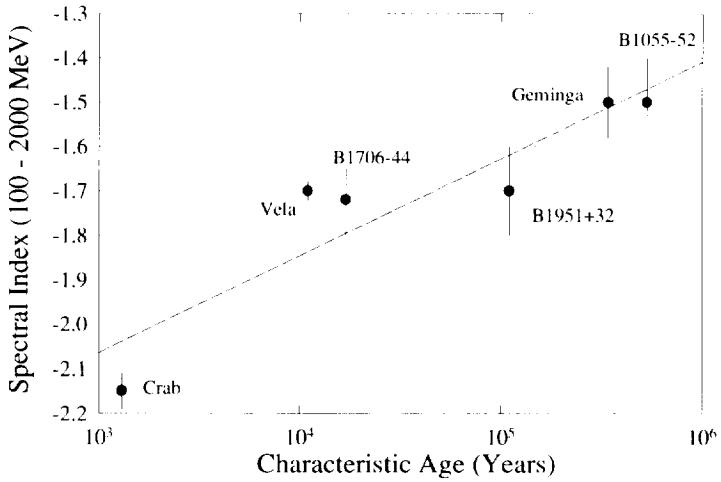

Figure 6.11. High-energy gamma-ray energy spectral index as a function of characteristic age (Thompson, 1996c).

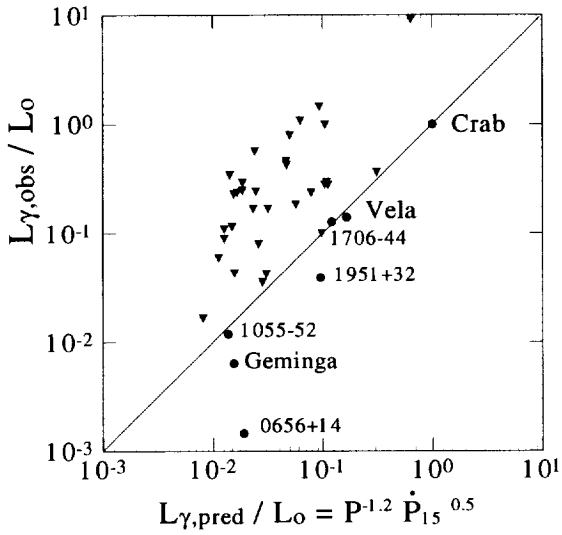

Figure 6.12. The gamma-ray luminosity as a function of the observed gamma-ray luminosity in the high-energy range(E > 100 MeV), both normalized to L_O. The figure is from Harding (1996), using results in the EGRET papers referenced here.

In spite of the successes of the polar cap model that have just been described, there are some other features which present difficulties for this model, specifically the phase differences between the ratio and gamma-ray peaks and the crab optical spectrum. The outer gap model, on the other hand, is able to explain the radio, gamma-ray phase differences on the basis of geometrical considerations and also is able to explain the Crab optical spectrum. However, the outer gap model has had less success in predicting the luminosities and spectra of pulsars other than the Crab. For review of these two models, including comparisons to the observations, see, e.g., Nel and De Jager (1995).

6.4 Accretion-Driven Pulsars

In addition to existing by itself, a neutron star may be a part of a binary system. Its companion may be one of several types of stars. The high-energy emission that has been observed is in the hard

x-ray or low-energy gamma-ray region. They are, in fact, often called "x-ray pulsars", and are more commonly thought of as a subject in x-ray astronomy. There are over 30 known accreting pulsar systems, and, thus far, 17 accretion-driven pulsars have been seen in low-energy gamma rays by OSSE or BATSE on the Compton Gamma-Ray Observatory. None have been observed in high-energy gamma rays by EGRET. Table 6.4 lists those that have been observed and gives the nature of the companion star.

The stellar companion of the high-magnetic-field neutron star transfers mass to the neutron star via Roche-lobe overflow or a stellar wind. The gravitational energy released by mass accretion yields thermal and nonthermal radiation. The observed radiation in the hard x-ray to low-energy gamma-ray range, therefore, provides

Table 6.4.

Accretion-driven Pulsars Seen by the Compton Gamma-Ray Observatory

Type of companion	Object
Low-mass companion	HER X-1
	GX 1+4
	4U 1626-67
High-mass companion	CEN X-3
High-mass/wind fed	OAO 1657-415
	VELA X-1
	GX 301-2
BE binaries	4U 0115+63
	A1118-616
	GRS O834-43
	A0535+26
	GRO J1008-57
	4U 1145-619
	GRO J1948+32
	2S 1417-624
	GRO J1750-27
	GRO J2014+34

information on the interaction of the relativistic particle wind with the gas of the companion star. The emission that is observed is periodic in the period of the binary system. An example is the emission from GRO J2014+34, observed by BATSE (Chakrabarty et al., 1995 and 1996), shown in Figure 6.13; this particular source has a 198 s period. The accreting pulsars provide the potential of studying the torques during magnetic accretion. The observed systems vary considerably one from the other, providing a considerable challenge for theorists. Reviews on the subject of accretion-driven pulsars include those of Nagase (1980) and Prince et al. (1994).

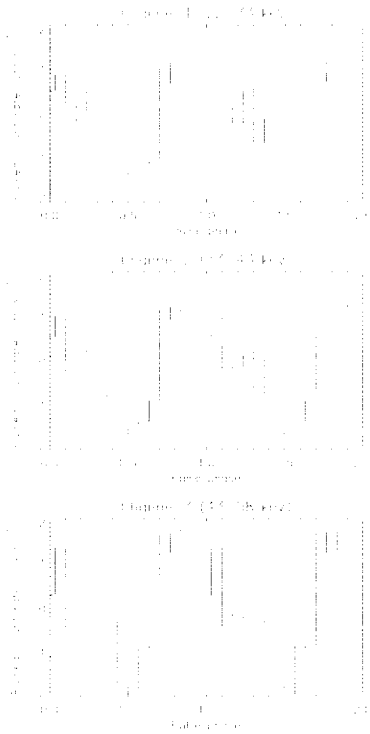

Figure 6.13. Pulse profiles as a function of energy for GRO J1948+32, during the interval MJD 49462-49468. Two pulses are shown for each energy channel. The energy edges for CONT channels 1-3 of BATSE detector 0 are indicated. These pulse profiles are overresolved by a factor of 1.7. Note that the pulse shapes are uncorrected for the rapid change in detector response as a function of energy in the 20-100 keV range. The figure is reproduced with the permission of Chakrabarty et al. (1996).

6.5 Supernovae

Supernovae, the most spectacular of stellar events in the optical region, have received much deserved attention by optical astronomy and were among the first objects considered by gamma-ray theorists. They are thought to be the source of nucleosynthesis of the heavy elements, and many feel they are the origin of most of the energy in cosmic rays. If nucleosynthesis has occurred during the explosive phase, specific gamma-ray lines with known energies and lifetimes should be emitted by unstable, heavy nuclei. Some of these have now been seen, as will be described here. The cosmic rays would reveal themselves through the emission of gamma rays produced in the interaction of cosmic rays with the surrounding material.

Exactly what occurs just before and during a supernova explosion is still uncertain in many respects. It is generally agreed that the temperature inside massive stars reaches the point at which nucleosynthesis occurs, and, after a number of steps, an "Fe" core is formed. There is then a sudden, implosive collapse of the core.

It is now thought to be more likely that the shock wave and ejection of matter result from an elastic bounce rather than from the deposition of neutrino energy in the outer layers. Most astrophysicists in this field, however, would agree that a completely satisfactory theoretical model of the process does not yet exist. For smaller stellar masses, the core remaining behind after the collapse of the rotating, magnetic stellar interior is the neutron star, now generally accepted as giving rise to the pulsar radiation discussed earlier in this chapter.

As the shock moves outward in the star, rapid nucleosynthesis occurs in the intermediate layers as a result of the shock heating. Following the early work of Fowler and Hoyle (1964), the abundances to be expected in the thermonuclear burning were calculated in substantial depth (e.g., Bodansky et al., 1968). During the explosive phase, there is the creation of many nuclear species which are unstable to beta decay, followed by a nuclear gamma-ray emission (Clayton et al., 1969). This situation provides an opportunity for gamma-ray astronomy to study the products of nucleosynthesis in supernovae at the level required to supply the observed quantities of some of the heavier elements.

Hydrodynamic calculations of these shocks have been performed, and the nucleosynthesis predicted in the matter to be ejected has been determined. The iron, cobalt, and nickel group is the most likely to be detected because of the relatively large number of atoms predicted. As is well known, there has not been a supernova detected in our galaxy for a long time, so there is no expectation for seeing the intense lines from unstable nuclei that decay in times of days to a few years. These lines have been seen in supernovae in other galaxies, and they will be discussed in the next chapter.

Within our galaxy, Cassiopeia A (Cas A) appears to be the best candidate for the emission of lines at a level that could be detected. It is estimated to be 3 kpc away, and it is about 340 years old. The explosion date has been estimated to be about AD 1658 ± 3 by van der Bergh and Kamper (1983). The 1.156 MeV gamma-ray line from the decay of ^{44}Ti with a lifetime of approximately 80 years is the most likely one to be seen, and, indeed, it has been observed by COMPTEL (Iyudin et al., 1994). The intensity observed by COMPTEL is within the expected range. It is a particularly interesting line, because it provides the possibility of detecting obscured supernova remnants of explosions that have occurred in the last few hundred years.

Formed in both hydrostatic and explosive carbon burning (Arnett, 1969), ^{26}Al, which decays to ^{26}Mg, is one of the best prospects for gamma-ray line emission (Arnett, 1977a; Ramaty and Lingenfelter, 1977). An excess of ^{26}Mg in the Ca-Al rich inclusion in the carbonaccous meteorite Allende has been interpreted as strong evidence for ^{26}Al in primitive solar system matter (Lee et al., 1977). In young supernova remnants, the large velocities smear the line; however, in older remnants better signal-to-noise ratios are more favorable. Diehl et al. (1996) have reported the detection of the 1.809 MeV line from the Vela region. The authors indicate that the location of the emission suggests an association with the Vela supernova remnant. The measured flux is consistent with theoretical estimates if the distance to the supernova remnant is below 500 pc, as noted in their paper.

As described in the last chapter, the ^{26}Al line has been seen in the diffuse galactic emission, presumably because of the gradual spreading of this material. Lines have been seen in association with supernovae in other galaxies, and they will be discussed in the next chapter.

In addition to the gamma-ray lines, it has long been thought possible that a gamma-ray continuum with the appropriate energy spectrum might be produced by the interaction of cosmic rays, created in association with the supernova, with the surrounding material, or by the further acceleration of particles in the supernovae shocks. The origin of the cosmic rays is a problem which has attracted attention for decades. Since it is generally believed that the cosmic rays are Galactic, and since supernovae appear to be the most energetic events occurring in the Galaxy, interest developed in supernovae as the primary sources of cosmic rays. An early concept was that the cosmic rays were actually the outermost thin layer of a supernova which was accelerated to cosmic-ray energies (Colgate and Johnson, 1960); more recently, however, hydrodynamic calculations make this possibility seem unlikely. There is some thought now that cosmic rays may receive some of their energy in the initial explosion or possibly in the shock and turbulence following the actual supernova explosion, and the rest, subsequently, in shock waves in interstellar space. These shock waves themselves may be from older supernovae. It has been noted in the last section that a continuum radiation has been observed for the Crab, but for no other pulsar. The low-energy gamma-ray Crab radiation is generally interpreted, however, as synchrotron radiation of relativistic electrons, whereas the high-energy radiation is thought to be inverse Compton radiation.

Except in a few cases, gamma rays would not be expected to be seen from most supernovae remnants with present experiments. Approximately a few times 10^{49} ergs per supernova in the form of relativistic particles are required if they are to be the source of cosmic rays or sights of acceleration providing much of the cosmic-ray energy. Both Esposito et al. (1994, 1995) and Sturner and Dermer (1995) have performed analyses of the EGRET data to see if there is a statistically significant result, if one compares the unidentified EGRET high-energy gamma-ray sources (Fichtel et al., 1994b) with known supernova remnants. Both obtain a positive but not absolutely compelling result. The results are consistent with a few close supernovae remnants having fluxes of the order of $(0.5-1.2)\times10^{-6}$ gamma rays (E > 100MeV) $cm^{-2}\,s^{-1}$, a flux range in agreement with the energy input required if they are the major source of the cosmic-ray energy. It also may be that the high-energy gamma rays are due to a pulsar undetected in the radio range, like Geminga, since the angular resolution could not resolve the

difference between the two explanations. Above 10^{10} eV, Allen et al. (1995) have searched for emission from supernova remnants (SNRs) and found none.

6.6 Black Holes

The intriguing theoretical prediction of the possible existence of black holes has excited astronomers and laymen alike. Masses ranging from just over one solar mass (the upper limit for a white dwarf is 1.4 $M_{\odot}$) to over a billion solar masses could be created in the Universe as it exists today, and mini black holes might have been created during the big bang. To pursue the discussion of the last section, the following is one way in which a black hole may form: If the outward shock following the collapse of a stellar core is not adequate to create an explosion blowing off the outer shells, the core will continue to accrete matter from the surrounding envelope, ultimately evolving into a black hole. It is worth noting that the formation of a black hole in this manner does not lead to a spectacular visible explosion, and, hence, does not leave an outstanding signature to be read by astronomers. Such objects would continue to collect matter that comes too close. However, the very massive black holes that are speculated upon are associated with the centers of galaxies.

Theorists originally believed that once a black hole was formed, it remained, and there was no chance of observing it. Now there are thought to be several ways the presence of a black hole might be revealed, although in some cases the observation of the radiation would not uniquely establish its source as a black hole. Much of the predicted radiation falls in the range of gamma rays. Thus, black holes are a very appropriate subject for discussion here, and several possible mechanisms for the emission of high-energy electromagnetic radiation from black holes will be described as follows.

Initially, the most interesting evidence suggesting the possible existence of a black hole came from optical and x-ray observations. A black hole in a binary system might pull gas from its companion in sufficient quantity so that the x-rays are emitted at a detectable level, while the gas is heated, so falling into the black hole. To be certain that the emission is from a binary with a black hole rather than a neutron star or a white dwarf, the unseen companion must be proven to have a mass clearly in excess of three solar masses.

Three black hole candidates, Cygnus X-1, the Galactic Center, and 1E 1740.7 – 2442 (1E 1740) have been reported to have gamma-ray emission. Gamma rays would be expected from black holes; Maraschi and Treves (1977) have noted that if the accretion flow onto the black hole is turbulent and dissipation maintains approximate equipartition among the different forms of energy, electrons can be accelerated by the induced electric fields. The resulting synchrotron energy spectrum is relatively flat to about 20 MeV, above which it falls steeply. There also would be a Compton contribution. Under the right conditions, observable gamma-ray fluxes would be generated. Collins (1979) has pointed out that matter falling onto a rotating black hole will be heated sufficiently, so that proton-proton collisions will produce mesons, including neutral pions which decay into two gamma rays. For massive ($> 10^3$ M) black holes, such as might exist in the Galactic center, the resulting gamma-ray luminosity might be detectable. Emission from black holes through the Penrose process is another mechanism, but it is most appropriately discusssed in relation to very large black holes, which may exist at the centers of active galaxies (see Chapter 8).

Beginning with 1E 1740, three transient events have been reported by the imaging gamma-ray telescope SIGMA on board the GRANAT spacecraft (Sunyaev et al., 1991). These transients from 1E 1740 were characterized as broad-line excesses in the 0.3 – 0.6 MeV range. They have been interpreted as arising from positron annihilation near a Galactic black hole. However, in at least one case, OSSE did not confirm the transient, and the upper limit is in conflict with the observation.

Transient emission of gamma rays in the MeV energy range has been reported from Cygnus X-1 (Ling et al., 1987) and the Galactic Center region (Riegler et al., 1985). These observations, however, are in contrast to the nondetection of such events in the 9 years of SMM data. The SMM upper limit for broadband transients at MeV energies, on time scales of 12 days or longer, was $(3\sigma) \leq 5 \times 10^{-3}$ photons cm^{-2} s^{-1} (Harris et al., 1993). Although the SMM data are not concurrent with SIGMA observations, the limits are not consistent with the frequency implied by all three SIGMA detections of gaussian-like spectral features from 1E 1740 (Harris et al., 1994).

Thus, for the moment, although interesting results have been reported, the question of gamma-ray emission from Galactic black holes larger than 1 $M_\odot$ appears best left as an open issue.

6.7 Mini Black Holes

For a very long time there has been an interest in making general relativity compatible with quantum mechanics. An attempt to achieve this marriage was made by Hawking (1974) and Page and Hawking (1976). A very specific prediction arises from this work that can be tested or at least constrained by high-energy gamma-ray observations. In the early Universe, small black holes might have been created if the Universe were chaotic or had a soft equation of state (Hawking, 1971; Carr and Hawking, 1974; Carr, 1975). They slowly evaporate and the temperature increases until the black hole begins radiating particles of higher-rest mass. Using the statistical bootstrap approach of Hagedorn (1968), the number of species of particles increases exponentially with energy; therefore, the black hole will lose its energy very quickly when it reaches the Hagedorn limiting temperature of 160 MeV. The heavy hadrons emitted by the black hole would decay rapidly and about 10 to 30 percent of their energy would appear as a burst of gamma rays with energies primarily between 10^2 and 10^3 MeV. The total energy in the form of gamma rays would be about 10^{34} ergs, and the time of emission would be 10^{-7} s, according to Page and Hawking (1976).

It should be noted that the exponential increase of the number of particles with energy is not predicted by the "Standard Model," which is based on a finite number of quarks, leptons, and gauge bosons. If the number of states is not infinite, then the GeV emission from the evaporation of a primary black hole will occur much more slowly. (See, e.g., MacGibbon and Webber, 1990, or Halzen, Zas, MacGibbon and Weekes, 1991, and references therein). In this alternative, although a burst of the order of seconds is expected for a threshold energy, E(t), of 10 TeV, the length of the burst in this model is proportional to $E(t)^{-3}$, so a group of gamma rays in the 10^2 to 10^4 MeV range would be extended so long in time that it would not be a burst. Cline and Hong (1993) have suggested that these bursts might be much shorter, a small fraction of a second. Cline and Hong (1992) have proposed still a third alternative leading to millisecond gamma-ray bursts.

Attempts have been made to see bursts of two of these types using ground-based telescopes which can record secondaries produced in the atmosphere. For the former 10^{-7} s 10^{2} MeV type burst, Porter and Weekes (1978) used a ground-based system, in their case a large Cerenkov telescope, to attempt to detect the low-energy photons resulting from an air shower that would be produced by such a high-energy gamma-ray burst. They set an upper limit of 4×10^{-2} decays pc^{-3} yr^{-1}. Alexandreas et al. (1993), using an ultra high-energy extensive air shower array, have set an upper limit for the proposed TeV bursts of a fraction of a second to a few seconds of 8.5×10^{5} pc^{-3} yr^{-1}, although this limit is theory dependent. Connaughton et al. (1993) have also set an upper limit for this type of burst.

It is possible to search for high-energy gamma-ray microsecond bursts by the direct observation of the high-energy gamma rays with EGRET. Because there is a time interval of approximately 6×10^{-7} seconds from the passage of the initial gamma ray in a group to the triggering of the spark chamber system, several gamma rays could be seen in an individual burst. Multiple gamma-ray events generated in the accelerator target were seen during the calibration of EGRET on the ground. Hence, unless the event is vetoed by the anticoincidence system, all gamma rays converting in the active volume of the telescope will be recorded, and one can look for events in which there is more than one gamma ray from a given direction in one event. The approach to processing the gamma-ray events is described in Thompson et al. (1993).

There was then no evidence for any multiple gamma-ray microsecond burst. Fichtel et al. (1994a) showed that the probability of missing a multiple gamma-ray event was small. That paper also describes the method for calculating the upper limit for the black hole emission density for a black hole with total energy E. In a later paper, Fichtel et al. (1996) showed that, with the analysis of the majority of the EGRET data complete, there is still no evidence for the existence of a microsecond burst of high-energy gamma rays. The energy spectrum of the gamma rays used in that work is that given by Page and Hawking (1976), which is peaked at 250 MeV. For $E = 10^{34}$ ergs, the energy suggested by Page and Hawking, the upper limit is 3×10^{-2} pc^{-3} yr^{-1}. Since approximately 85 percent of the effective collecting life of EGRET has now been expended, this limit will not change significantly,

unless, of course, a short multiple gamma-ray event is observed. Even then it might be caused by something else. The relevant distance range is from about 20 to 100 pc. Below 20 pc, the instrument would be overwhelmed and not record the event; beyond 100 pc, the probability of detection becomes quite small. For E = 10^{35} ergs, the upper limit is 1 x 10^{-3} pc^{-3} yr^{-1}, and the distance range is somewhat larger.

There remains no evidence within the work reported here for the existence of a microsecond burst of high-energy gamma rays, and severe upper limits have been set. As noted in the last paragraph, the upper limit of primordial black holes decay, set by an instrument that could detect them directly, is now 3 x 10^{-2} pc^{-3} yr^{-1}. Also, for 10^{34} ergs, the black holes that could have been detected would have been predominantly within 100 pc and, hence, within the local region of our galaxy. Page and Hawking noted that "one might expect that the [PBH] would be concentrated in the gravitational potential wells of galaxies." Hence, in view of the distant range, this limit is particularly relevant in terms of this model.

6.8 Other High-Energy Gamma-Ray Sources

In addition to the known gamma-ray pulsars, there are several as yet unidentified sources seen in the gamma-ray sky. Particularly in the Galactic plane, many of these may be localized excesses rather than stellar objects or point sources. The second COS-B catalog (Swanenburg et al., 1981) listed 22 such sources. Later analysis showed that about half of these could be explained as cosmic-ray interactions with gas clouds (Mayer-Hasselwander and Simpson, 1990), and one of those remaining was subsequently identified by EGRET as the B1706-44 pulsar discussed earlier in this chapter. At least six sources, however, have been confirmed by EGRET and remain unidentified, along with many new sources, providing a total of about seventy unidentified sources (Fichtel et al., 1994b; Thompson et al., 1995).

As seen in Figure 6.14, the distribution of the EGRET unidentified sources is not uniform on the sky. Approxiately 40% of them lie within 5° of the Galactic plane, even though this region represents only about 9% of the sky. This result is even stronger than these

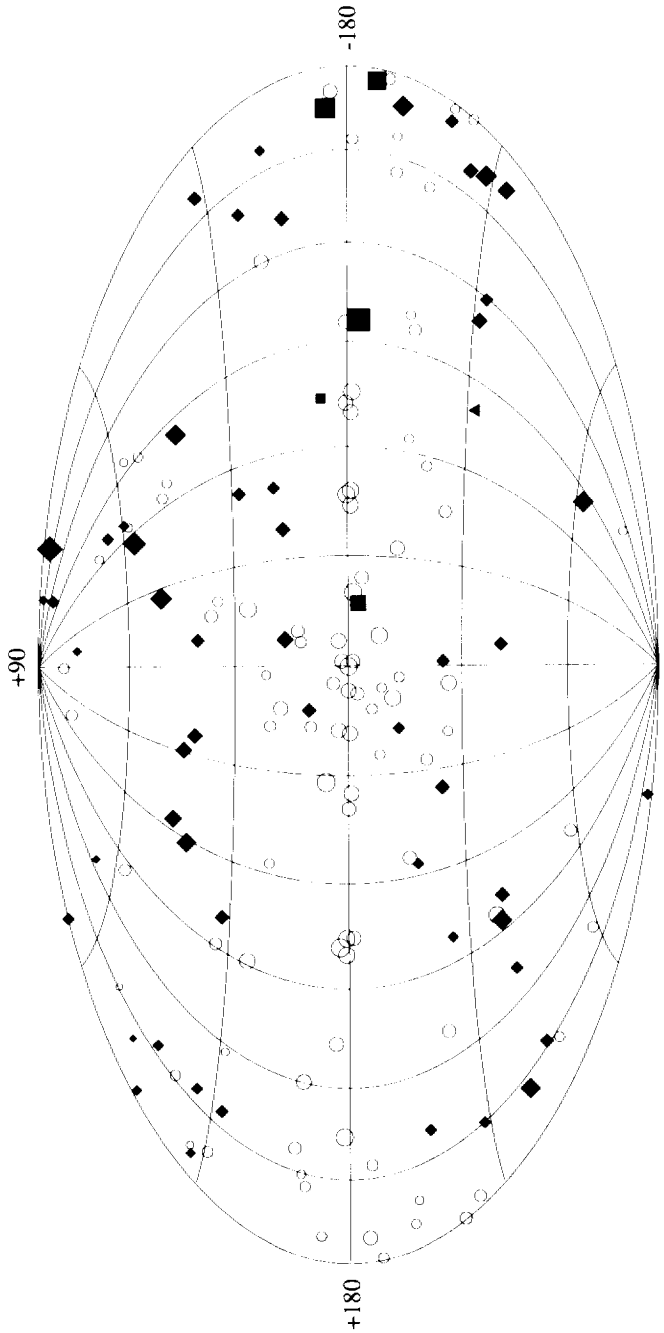

Figure 6.14. Distribution of the high-energy gamma-ray sources seen by EGRET; diamond nuclei; circles: unidentified EGRET sources; squares: pulsars; and triangle: LMC. The figure EGRET Catalog and its supplement (Thompson et al., 1995, 1996).

numbers might indicate, because the high level of the Galactic diffuse radiation makes it more difficult to detect point sources in this region of the sky than elsewhere. It is true that the exposure is greater in a broad band of about 25° on each side of the plane, but the number of sources does not increase linearly with exposure, and the number of sources per unit solid angle is a factor of more than three smaller for the two regions on each side of the plane that are between 5° and 15° from the plane. Approximate estimates, based on that spatial distribution only, suggest that almost half of all of the unidentified sources are Galactic.

What is the nature of these unidentified sources? It seems reasonable to assume that the larger half that are uniformly distributed over the whole sky are likely to be blazars with weaker radio emission. There are many suggested specific associations, but the greatly increased number of weaker radio sources relative to the strongest ones and the poorer source locations of the weaker gamma-ray sources cause one to be in the realm of source confusion, which does not permit highly probable identifications with FSRQs or radio-emitting BL Lacs. Hence, exact identifications are not possible, but there is every reason to think that the distribution of gamma-ray-emitting blazars extends down to less intense sources.

The unidentified Galactic sources are probably a mixture of types. Some may be unidentified pulsars, particularly radio quiet ones like Geminga. Some beyond those already discussed earlier in the chapter might be associated with supernova remnants. Some will probably be excesses caused by cosmic rays interacting with localized matter concentrations not presently adequately included in the existing models of the interstellar medium.

References

Alexandreas, D. E., 1993, Phys. Rev. Letter, 71, 2524.

Allen, G. E. , et al., 1995, ApJ, 448, L25.

Arnett, W. D., 1969, ApJ, 157,1369.

Arnett, W. D., 1977a, Annals of the New York Academy of Sciences, Proc. of the 8th Texas Symposium, ed. M. D. Papagiannis, 302, 90.

Baade, W., and Zwicky, F., 1934, Proc. Nat. Acad. Sci. USA, 20, 254 .

Bartlett, L. M., 1994, personal communication.

Bennett, K., et al., 1977, A&A, 56, 469-471.

Bertsch, D. L., et al., 1992, Nature, 357, 306.

Bignami, G. F., Caraveo, P. A., and Lamb, R. C., 1983, ApJ, 272, L9-L13.
Bignami, G. F., Caraveo, P. A., Paul, J. A., Salotti, L., and Vigroux, L., 1987, ApJ, 319, 358-361.
Bignami, G. F., and Caraveo, P. A., 1992, Nature, 357, 287.
Bodansky, D., Clayton, D. D., and Fowler, W. A., 1968, ApJS, 16, 299.
Brazier, K. T. S., et al., 1993, in preparation.
Browning, R., Ramsden, D., and Wright, P. J., 1971, Nature, 232, 99.
Buccheri, R., Caraveo, P., D'Amico, N., Hermsen, J. L., Wills, R. D., Manchester, R. N. and Newton, M., 1978, A&A, 69, 141.
Carr, B. J. and Hawking, S. W., 1974, Mon. Not. R. Astron. Soc., 168, 243.
Carr, B. J., 1975, ApJ, 201, 1.
Chakrabarty, D., et al., 1995, ApJ, 446, 826.
Chakrabarty, D., et al., 1996, private communication.
Cheng, K. S., Ho, C., and Ruderman, M., 1986a, ApJ, 300, 500.
Cheng, K. S., Ho, C., and Ruderman, M., 1986b, ApJ, 300, 522.
Cheng, K. S., and Ding, W. K. Y., 1994, ApJ, 431, 724.
Churazov, E., Gilfanov, M., Sunyaev, R., et al., 1993, ApJ, 407, 752.
Clayton, D. D., Colgate, S. A., Fishman, G. J., 1969, ApJ, 155.
Clear, J., Bennett, K., Buccheri, R., Greniar, I. A., Hermsen, W., Mayer-Hasselwander, H. A., and Sacco, B., 1987, A&A, 174, 85.
Cline, D. G., and Hong, W., 1992, ApJ, 401, L57.
Cline, D. G., and Hong, W., 1993, Gamma Ray Bursts, Second Workshop, AIP Conference Proceedings 307, ed. G. J. Fishman, J. J. Brainerd, K. Hurley, 470.
Colgate, S. A., and Johnson, M. H., 1960, Phys. Rev. Letters, 5, 235.
Collins, M. S., 1979, A&A, 74, 108.
Connaughton, V., et al., 1993, Gamma Ray Bursts, Second Workshop, AIP Conference Proceedings 307, ed. G. J. Fishman, J. J. Brainerd, K. Hurley, 470.
Cordier, B., Paul, J., Ballet, J., et al. 1993, A&A, 275, L1.
Daugherty, J. K., and Harding, A. K., 1994, ApJ, 429, 325.
De Jager, O. C. et al., 1996, ApJ, 457, 253.
Diehl, R., et al., 1996, A&AS, [Third Compton Symposium],to be published.
Esposito, J. A., Hunter, S. D., Kanbach, G., Sreekumar, P., 1994, Bulletin AAS, 26, 1518, 120, 04.
Esposito, J. A., Hunter, S. D., Kanbach, G., and Sreekumar, P., 1995, submitted to ApJ.
Fichtel, C. E., et al., 1975, ApJ, 198, 163-182.
Fichtel, C. E. et al., 1994a, ApJ, 434, 557.
Fichtel, C. E., et al., 1994b, ApJS, 94, 551.
Fichtel, C. E., and Thompson, D. J., 1994, High Energy Astrophysics Models and Observations from MeV to EeV, ed. James Matthews, World Scientific, 1.
Fichtel, C. E. et al., 1996, Gamma-Ray Bursts, 3rd Huntsville Symposium, Huntsville, AL, October 1995, AIP Conference Proceedings, 384, 368.
Fierro, J. M., et al., 1993, ApJ Letters, 413, L27.
Fowler, W. A. , and Hoyle, F., 1964, ApJS, 9, 201.

Gold, T., 1968, Nature, 731.
Groth, E. J., 1975. ApJS, 293, 431.
Grindlay, J. E., Helmken, H. F., and Weekes, T. C., 1976, ApJ, 209, 592.
Gunn, J. E., and Ostriker, J. P., 1970, ApJ, 160, 979.
Hagedorn, R., 1968, Nuovo Cim., 269, 1027.
Halpern, J. P., and Tyler, D., 1988, ApJ, 330, 201-217.
Halpern, J. P., and Holt, S. S., 1992, Nature, 357, 222-224.
Halpern, J. P., and Ruderman, M., 1993, ApJ, 415, 286.
Halzen, F., Zas, E., MacGibbon, J. H., and Weekes, T. C., 1991, Nature 353, 807.
Harding, A. K., 1981, ApJ, 245, 267.
Harding, A. K., and Daugherty, J., 1993, in Isolated Pulsars, ed. K. A. van Riper, R. Epstein, C. Ho, (Cambridge University Press), 279.
Harding, A. K., 1996, private communication.
Harris, M. J., Share, G. H., Leising, M. D., and Grove, J. E. 1993, ApJ, 416, 601.
Harris, M. J., Share, G. H., and Leising, M. D. 1994, ApJ, 433, 87.
Hartmann, D. H., 1995, A&A Rev., 6, 225.
Hawking, S. W., 1971, Mon. Not. R. Astron. Soc., 152, 75.
Hawking, S. W., 1974, Nature, 248, 30.
Helmken, H. F., Grindlay, J. E., and Weekes, T. C., 1975, Proc. of the 14th Int. Cosmic Ray Conf., 1, 123.
Hermsen, W., 1990, EGRET Science Symposium, NASA/GSFC 161-177.
Hermsen, W., et al., 1993, Compton Symposium St. Louis, ed. M. Friedlander, N. Gehrels, and D. J. Macomb, AIP Conf. Proc., 280, 204.
Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F., and Collins, R. A., 1968, Nature, 217, 709.
Ho, C., 1993, in Isolated Pulsars, ed. K. A. van Riper, R. Epstein, C. Ho, (Cambridge University Press), 271.
Iyudin, A., et al., 1994, A&A, 284, L1.
Jackson, D., 1975, Classical Electrodynamics 2nd Ed., New York, Wiley.
Kanbach, G., 1989, in Proc. Gamma Ray Observatory Science Workshop, ed. W. N. Johnson (NASA Goddard Space Flight Center), 2-1.
Kanbach, G., et al., 1994, A&A, 289, 855.
Kanbach, G., et al., 1996, A&A, [Third Compton Symposium], to be published.
Kifune, N., et al., 1991, ApJ, 383, L65.
Kifune, N., et al., 1995, ApJ, 438, L91.
Kniffen, D. A., Hartman, R. C., Thompson, D. J., Bignami, G. F., Fichtel, C. E., Tümer, T., and Ögelman, H., 1974, Nature, 251, 397.
Kniffen, D. A., et al., 1975, Proc. 14th ICRC, Conf. Vol. 1, 100-105, Munich, Germany.
Kulkarni, S. R., et al., 1988, Nature, 331, 50.
Lee, T., Papanastassiou, D. A., and Wasserburg, G. J., 1977, ApJ, 211, L107.
Ling, J. C., Mahoney, W. A., Wheaton, W. A., and Jacobson, A. S. 1987, ApJ, 321, L117.

Lyne, A. G., Ritchings, R. T., and Smith, F. G., 1975, Mon. Not. Roy. Ast. Soc., 171, 579.
MacGibbon, J. H., and Webber, B. R., 1990, Phys. Rev., D41, 3052.
Malofeev, V. M., and Malov, O. I., 1997, "Detection of Geminga as a Radio Pulsar," preprint.
Manchester, R. N., 1971, ApJ, 163, L61.
Maraschi, L., and Treves, A., 1977, ApJ, 218, L113.
Masnou, J. L., et al., 1981, Proc. 17th ICRC, Vol. 1, 177-180.
Mattox, J. R., et al., 1992, ApJ, 401, L23.
Mattox, J. R., De Jager, O. C., and Harding, A. K., 1992, in Isolated Pulsars, ed. K. A. van Riper, R. Epstein, and C. Ho, (Cambridge University Press: Cambridge), 298.
Matz, S. M., et al., 1994, ApJ, 434 288.
Mayer-Hasselwander, H. A., and Simpson, G. A., 1990, The Energetic Gamma-Ray Experiment Telescope (EGRET) Science Symposium, (eds. C. E. Fichtel, S. Hunter, P. Sreekumar, and F. Stecker), NASA Conf. Pub. 3071, p. 153.
Mayer-Hasselwander, H. A., et al., 1994, ApJ, 421, 276.
Much, R. P., et al., 1995a, COSPAR Conference, Hamburg, 1994.
Much, R. P., et al., 1995b, Adv. in Space Res., 15(5), 81.
Mukherjee, R., et al., 1995, ApJ, 441, L61.
Nagase, F., 1980, PASJ, 41, 1.
Nel, H. I., and De Jager, O. C., 1995, Astrophysics and Space Science, 230, 299.
Nel, H. I., et al., 1996, ApJ, 465, 898.
Nolan, P. L., et al., 1993, ApJ, 409, 697.
Oda, M., 1977, Space Sci. Rev., 20, 757.
Ögelman, H. B., and Finley, J. P., 1993, ApJ, 413, L31.
Ögelman, H. B., Finley, J. P., and Zimmerman, H. U., 1993, Nature, 361, 136.
Ögelman, H. B., 1994, in Lives of Neutron Stars, NATO ASI, Kemer, Turkey, in press.
Ostriker, J. P., and Gunn, J. E., 1969, ApJ, 157, 1395.
Özel, M. E., and Mayer-Hasselwander, H. A., 1984, in Data Analysis in Astronomy, ed. V. DiGesu et al. (New York: Plenum), 81.
Özel, M. E., and Thompson, D. J., 1996, ApJ, 463, 105.
Pacini, F., 1968, Nature, 216, 567.
Page, D. N. and Hawking, S. W., 1976, ApJ, 206, 1.
Porter, N. A., and Weekes, T. C., 1978, Mon. Not. R. Astron. Soc., 183, 205.
Prince, T. S., Bilksten, L., Chakrabarty, C., Wilson, R. B., and Finger, M. H., 1994, American Institute of Physics Conference Proceedings, 308, 235.
Ramanamurthy, P. V., et al., 1995a, ApJ, 450, 791.
Ramanamurthy, P. V., et al., 1995b, A&A, [Third Compton Symposium], to be published.
Ramanamurthy, P. V., et al., 1995c, ApJ, 447, L109.
Ramanamurthy, P. V., Fichtel, C. E., Kniffen, D. A., Sreekumar, P., and Thompson, D. J., 1996, ApJ, 458, 755.
Ramaty, R., and Lingenfelter, R. I., 1977, ApJ, 213, L5.

Rappaport, S., Bradt, H., and Mayer, W., 1971, Nature Phys. Sci, 229, 40.
Riegler, G. R., Ling, J. C., Mahoney, W. A., Wheaton, W. A., and Jacobson, A. S., 1985, ApJ, 294, L13.
Ruderman, M. A., and Sutherland, P. G., 1975, ApJ, 196, 51.
Ruderman, M. A., and Cheng, K. S., 1988, ApJ, 335, 306.
Safi-Harb, S., Ögelman, H. B., and Finley, J. P., 1995, ApJ, 439, 722.
Samimi, J., Share, G. H., Wood, K., Yentis, D., Meekins, J., Evans, W. D., Shulman, S., Byram, E. T., Chupp, T. A., and Friedman, H., 1979, Nature, 279, 434.
Spoelstra, T. A. Th., and Hermsen, W., 1984, A&A, 135, 135-140.
Strickman, M. S., et al., 1993, Compton Symposium St. Louis, ed. M. Friedlander, N. Gehrels, and D. J. Macomb, AIP Conf. Proc. 280, 209.
Sturner, S. J., and Dermer, C. D., 1995, A&A, 293, L17.
Sunyaev, R., et al. 1991, ApJ, 383, L49.
Swanenburg, B. N., et al., 1981, ApJ, 242, L69.
Taylor, J. H., and Manchester, R. N., 1975, AJ, 80, 794.
Thompson, D. J., Fichtel, C. E., Kniffen, D. A., and Ögelman, H. B., 1975, ApJ, 200, L79.
Thompson, D. J., et al., 1977, ApJ, 213, 252-262.
Thompson, D. J., et al., 1992, Nature, 359, 615.
Thompson, D. J. et al., 1993, ApJS, 86, 629.
Thompson, D. J., et al., 1995, ApJS, 101, 259.
Thompson, D. J., et al., 1996a, submitted to the ApJS.
Thompson, D. J., et al., 1996b, ApJ, 465, 385.
Thompson, D. J., 1996c, private communication.
Thorne, K. S., 1974, Scientific American, 234, 32.
Toor, A., 1977, ApJ, 215, L57.
Ulmer, M., 1993a, IAU Colloq. No. 142, Particle Acceleration Phenomena in Astrophysical Plasmas.
Ulmer, M. P., et al., 1993b, ApJ, 413, 738.
Ulmer, M. P., et al. 1994, ApJ, 432, 228.
Ulmer, M. P., et al, 1995, ApJ, 448, 356.
Vacanti, V, et al., 1991, ApJ, 377, 467.
van der Bergh, S., and Kamper, K. W., 1983, ApJ, 268, 129.
Wallace, P. T., et al., 1977, Nature, 266, 692.
Weekes, T. C., et al., 1994, in Proc. 2nd Compton Symposium, ed. C. E. Fichtel, N. Gehrels, and J. P. Norris, (New York: AIP Conf. Proc. 304), 270.
Wilson, R. B., et al., 1992, IAU Circular 5429.
Wilson, R. B., et al., 1993, in Isolated Pulsars, ed. K. A. van Ripen, R. Epstein, and C. Ho (Cambridge University Press), 257.

CHAPTER 7

NORMAL GALAXIES

7.1 Introduction

In the context of the discussion here, normal galaxies compose a very broad category, which includes most galaxies and, in particular, elliptical galaxies, normal spirals, and all their variations. It specifically excludes those galaxies that are of an exceptional nature in luminosity or variability, such as Seyfert galaxies, quasars, BL Lacertae objects, and intense radio galaxies.

As discussed in Chapter 5, our galaxy seems to be a typical, spiral galaxy, belonging to a small local group of about 20 galaxies, most of which have masses smaller than our own. The distribution of galaxies in the universe on a fine scale is irregular. Galaxies generally exist in small groups, which, in turn, are part of large clusters, the biggest of which may contain a thousand members or more. There is now even some evidence to suggest that the clusters of galaxies are not uniformly spaced, but tend to be grouped into superclusters of galaxies.

The two closest galaxies to our own are the Small and Large Magellanic Clouds (the SMC and the LMC). These will be examined here first, followed by a discussion of some of the others. As galaxies go, the Magellanic Clouds are not particularly impressive, but they are important for astrophysical studies in a number of ways, resulting in large part from the fact that they are close and, therefore, can be studied more fully. With regard to gamma rays in particlular, the LMC has been seen in high-energy gamma rays, and even gamma-ray lines from a supernova have been observed there. The SMC has also been the source of a very important result; the upper limit of the flux of high-energy gamma rays from it provided the convincing proof that cosmic rays are

Galactic in nature and not metagalactic. A few other galaxies close to our own will also be discussed in this chapter.

7.2 The Small Magellanic Cloud

The Small Magellanic Cloud is a relatively small galaxy compared to the average; however, it appears even smaller than it is, because it is relatively long and small in cross section like a cigar, and it is oriented such that we see it from the viewpoint of this small cross section. It also appears to be somewhat fragmented or disrupted. The high-energy gamma-ray observations from this object are of particular interest, because they are able to reveal the cosmic-ray content in the same way that the cosmic rays were studied in our own galaxy, as described in Chapter 5. Hence, they represent a very direct estimate of the cosmic rays that are present. Knowing the matter and approximate photon content, a direct calculation can be made of the expected gamma-ray emission, based on the cosmic-ray content.

A long-standing, critical question in the study of the dynamic balance of our own galaxy is whether the bulk of the cosmic rays are Galactic or metagalactic in origin. Although there was evidence indicating that the cosmic rays are primarily Galactic in origin, there were also counter-arguments for an extra-galactic origin, based on contributions from radio galaxies, quasi-stellar objects, and galactic winds. As early as the 1970s, Ginzburg (1972) and Ginzburg and Ptuskin (1976) had noted that a definitive test of whether the bulk of the cosmic rays is Galactic or universal is to compare the level of high-energy gamma-ray emission from these galaxies. Radio synchrotron observations provide information on the cosmic-ray electron component, although the major energy component is the nucleonic one. Sreekumar and Fichtel (1991) showed that, based on the synchrotron data, arguments regarding the expected magnetic-field strength, and the assumption of the same cosmic-ray electron-to-nucleon ratio as in our local solar neighborhood, the Small Magellanic Cloud (SMC) would be expected to have a cosmic-ray density well below that in our galaxy if the cosmic rays were galactic. This conclusion is consistent with the concept that the SMC is in a state of irreversible disintegration, in agreement with the independent experimental findings by Mathewson, Ford and Visvanathan (1986, 1988) and the tidal interaction model of Murai and Fujimoto (1980).

The expected gamma-ray emission from the Small Magellanic Cloud may be estimated for the cases of a disrupted state just mentioned, a quasi-stable equilibrium situation, and a metagalactic cosmic-ray density case, and these calculations have been performed by Sreekumar and Fichtel (1991). The results for the flux above 100 MeV are 2.4 x 10^{-7} photons cm^{-2} s^{-1} for the metagalactic cosmic-ray case, 1.2 x 10^{-7} photons cm^{-2} s^{-1} for the quasi-stable equilibrium case, and (2 to 3) x 10^{-8} cm^{-2} s^{-1} for the disrupted state described in the last paragraph. For the latter estimate, the range of values is associated with the question of what part of the matter is assumed to be relevant for the disintegrating case. Except for that issue, the uncertainty in all of these numbers is estimated to be about 20%.

The measurements made with the EGRET have been able to provide only an upper limit of 0.5 x 10^{-7} cm^{-2} s^{-1} (Sreekumar et al., 1993). This result shows that the high-energy gamma-ray emission from the SMC is neither consistent with a metagalactic cosmic-ray assumption nor a quasi-stable equilibrium situation in the SMC. First, this is the most conclusive evidence that cosmic rays are Galactic and not metagalactic, and is the strongest basis for proceeding with the assumption that the cosmic rays in our own galaxy are Galactic, as was done in Chapter 5. Given that the bulk of the cosmic rays in any galaxy then are presumably a product of sources in that galaxy, there cannot be a quasi-equilibrium state in the SMC as there is in our own galaxy. Hence, the situation described earlier in this section of a disrupted state must be the one that exists in the SMC. Until there can be a much more sensitive exposure to the SMC, it will not be possible to determine just how disrupted it is from a cosmic-ray point of view, but the estimated intensity provided by Sreekumar and Fichtel (1991) based on the radio data and other factors is clearly consistent with the observed limit.

Hence, the high-energy gamma-ray study of the SMC provides two pieces of information. It supports the evidence noted earlier that the SMC is in a disrupted state, presumably as the result of an encounter with the LMC in the past, and it provides very convincing evidence that the bulk of the cosmic rays are Galactic and not metagalactic.

7.3 The Large Magellanic Cloud

From the standpoint of gamma-ray astronomy, there are two aspects of the LMC that are of particular interest. One is the high-energy gamma-ray measurements and the implications for the cosmic-ray content, and the other is the detection of supernova 1987A at low energies, which will be discussed in the next section.

Since the results and discussion of the last section have shown with a high probability that the cosmic rays are Galactic, this assumption will be taken as a starting point, and it will be seen that there is nothing to indicate a disagreement with this concept. From the material presented in Chapter 5, it is known that the principal information that is needed to calculate the expected high-energy cosmic-ray density is the matter density distribution in the LMC. Considerable information exists on the matter density in the LMC, or at least the column density, which is all that will be required here. The atomic hydrogen density may, of course, be determined directly from the 21 cm data, and Mathewson and Ford (1984) have obtained extensive data on this line for the LMC. For the molecular density, a full survey of the central 6^o x 6^o region has been obtained by Cohen et al. (1988). There is some uncertainty in the factor to convert the CO density to that of molecular hydrogen, but the molecular hydrogen density is relatively small and concentrated in a relatively small region. Hence, it does not make a major contribution.

With this information it is possible to calculate the expected cosmic-ray distribution in the LMC under the assumption that quasi-stable equilibrium exists, following the approach described in Chapter 5, and Fichtel et al. (1991) have done this. The result is dependent on the assumed coupling scale size, but, if it is assumed that the value of the coupling constant is in the range of one to three kiloparsecs, the result for the expected gamma-ray emission is not sensitive to the choice of the coupling constant.

Alternately, the cosmic-ray density distribution may be calculated from the radio synchrotron information that exists for the LMC in the same manner as for the SMC, and Fichtel et al. (1991) have done this. From this, the gamma-ray emission distribution and the total emission may be calculated. The two approaches are found to be in good agreement. The estimated total intensity based on the dynamic balance concept within an intensity level contour of

1.0 x 10^{-5} photons (E > 100 MeV) cm^{-2} s^{-1} sr^{-1} is (2.3 ± 0.4) x 10^{-7} photons (E > 100 MeV) cm^{-2} s^{-1}, with the error primarily caused by the uncertainty in the estimated molecular hydrogen density. After the EGRET point-spread function for the angular accuracy of the instrument is taken into account, the appropriate intensity to be compared with the experimental determination is (2.0±0.4) x 10^{-7} in the same units, in good agreement with the measured value of (1.9±1.4) x 10^{-7} (Sreekumar et al., 1992). It should be remembered that there is also a small, but unknown, point-source contribution; however, the two numbers would still be considered in agreement even with that consideration, although there is a chance that the cosmic-ray density might be slightly below its equilibrium value. Certainly it is not consistent with the LMC being disrupted at the level of the SMC, and other evidence would support this conclusion.

Thus, it appears from the high-energy gamma-ray data that the cosmic-ray content of the LMC is consistent with what one would expect for the LMC being in quasi-stable equilibrium.

7.4 Supernovae in Other Galaxies

It is well known that a supernova has not been seen for a long time in our galaxy, as discussed in the last chapter. Fortunately, there have been supernovae in other galaxies since low-energy gamma-ray experiments have been in space, and low-energy gamma rays have been seen from some of these, the closer ones, as would be expected.

The rapid sequence of events leading to a supernova explosion creates a large number of nuclei of various types. Many of these nuclei are unstable and decay, emitting low-energy gamma-ray lines. The sites of the nucleosynthesis are initially deep enough in the interior that the gamma rays will not reach the surface and be seen. Hence, only after the expansion of the supernova has reached the point that the overlying material is at least partially transparent to the low-energy gamma rays, will they be visible in the vicinity of the Earth. Thus, only those gamma rays coming from unstable isotopes that are plentiful enough and have lives that are long enough will be seen. There are some of these nuclear species formed in supernovae, and gamma-ray lines from ^{56}Co and other freshly synthesized radio nuclei formed in supernovae were predicted by

Clayton, Colgate, and Fishman (1969). The most intense lines are those from $^{56}Co \rightarrow ^{56}Fe$, which are at energies of 0.8468 MeV, 1.2383 MeV, 1.5985 MeV, 1.7714 MeV, and 1.0378 MeV (Huo et al., 1987). This decay is in the chain $^{56}Ni \rightarrow ^{56}Co \rightarrow ^{56}Fe$. Other decays that are less intense, but still quite significant, are $^{57}Co \rightarrow ^{57}Fe$ and $^{44}Ti \rightarrow ^{44}Sc \rightarrow ^{44}Ca$.

As the supernova expands, these gamma-ray lines are Doppler-broadened by the velocity distribution in the expanding material. The gamma-ray lines reflect this velocity distribution, modified by the opacity along the line of sight. Therefore, observations of these gamma-ray lines with high-resolution instruments provides a direct means of testing not only the models of explosive synthesis, but also the dynamics of the supernova ejecta (Ramaty and Lingenfelter, 1995).

The Type II supernova 1987A in the LMC provided an opportunity to study a supernova with gamma-ray line instruments. Although the supernova is in another galaxy, its distance from Earth is about 50 kpc, still relatively close. The ^{56}Co lines at .847 and 1.24 MeV, indeed, have been seen by a large number of balloon-borne instruments and studied in detail with the SMM gamma-ray telescope (e.g., Matz et al., 1988; Tueller et al., 1990; Leising and Share, 1990, and references therein). The gamma-ray observations showed that the ^{56}Co is extensively mixed in the ejecta. Specifically, calculations had indicated that the lines would not be visible for about 600 days after the explosion (Chan and Lingenfelter, 1987) if the ^{56}Co were confined to the innermost layers, but, in fact, it was seen within 200 days (Matz et al., 1988). This result led to modifications of the supernova models to explain the effects of mixing on the explosion and ejecta dynamics (e.g., Nomoto et al., 1988; Bussard, Burrows, and The, 1989; Pinto and Forestini, 1991). Four years after the explosion, the 0.122 MeV line from $^{57}Co \rightarrow ^{57}Fe$, with a 1.1-year lifetime, was detected by OSSE on the Compton Gamma-Ray Observatory (Kurfess et al., 1992).

A more recent supernova, 1991T, has occurred in the Virgo cluster, which is, of course, more distant than the LMC. COMPTEL, on the Compton Gamma-Ray Observatory, has reported a detection of both the ^{56}Co 0.847 MeV and 1.24 MeV lines (Leising et al., 1995).

However, The et al.(1994), using data from OSSE, have obtained only an upper limit, which is in marginal disagreement with the COMPTEL result.

In spite of the current situation regarding 1991T, there is now solid evidence of recent nucleosynthesis in the observation of supernovae explosions shortly after their occurrence both in our galaxy, as reported in the last chapter, and for supernova 1987A, as described here. In addition, there is the observation of the ^{26}Al line, with the long lifetime, spread over the Galactic plane, which is also consistent with the prediction. Together, these findings indicate that the nucleosynthesis theory of supernovae is on reasonably firm ground.

7.5 Some Other Relatively Close Galaxies

Since it is assumed that the production processes involved are the same in other normal galaxies, the energy spectra are expected to be similar to that in our own. Estimating the gamma-ray emission in another galaxy similar to our own may be accomplished in a straightforward manner, using the method that has been described here, if the basic information on the galaxy is known. Estimating the luminosity of another galaxy that is quite different from our own would require considerable caution. There are four nearby galaxies besides the Magellanic Clouds that are worth considering here.

Of these four, M31 is the closest and is the simplest one from the standpoint of making an estimate of the high-energy gamma radiation. Özel and Fichtel (1988) have performed a calculation of the type described here and determined that the expected flux would be about 2×10^{-8} photons cm^{-2} s^{-1} for gamma rays above 100 MeV. This is a factor of a few below the upper limit obtained by EGRET (Sreekumar et al., 1994). Although the result is only an upper limit and consistent with what was expected, it does suggest that there is neither an unexpectedly high cosmic-ray intensity, nor a surprising gamma-ray source.

M87 is a giant elliptical galaxy, about thirty times farther away than M31, but its radio luminosity is approximately one-thousand times greater than M31. EGRET has also obtained an upper limit for this galaxy, which is 4×10^{-8} photons ($E > 100$ MeV) cm^{-2} s^{-1}. This number is consistent with the estimated flux calculated by

Sreekumar et al. (1994) of 1.5 x 10^{-8} photons (E > 100 MeV) cm^{-2} s^{-1}. Hence, again there is no indication of a cosmic-ray density that is extraordinary.

NGC 253 and M82 are the two nearest starburst galaxies at distances of 2.5 and 3.3 Mpc respectively. They are typical starburst galaxies, characterized by an unusually high star-formation rate. They are bright infrared objects and exhibit extended x-ray and radio emission. Bartel et al. (1987) estimated the supernova rates lie in the range of (0.1 – 0.3) per year, about ten to thirty times larger than the estimated rate in the Milky Way. This very high supernova rate might suggest the possibility of higher cosmic-ray density, possibly even higher than could be held in a stable configuration. Assuming a cosmic-ray density ten to a hundred times that in the local region of our galaxy leads to an estimated flux of (0.2 to 2.0) x 10^{-8} photons (E > 100 MeV), below the upper limit of EGRET. Estimates for M82 also have been made, and they fall below the upper limit obtained by EGRET (Sreekumar et al.,1994; Paglione et al., 1996).

Thus, although EGRET has set upper limits that are an order of magnitude or more below previous results, there are no positive detections of these or other normal galaxies, except for the LMC. Further, there are no surprises, since the upper limits are consistent with expectations for normal assumptions. The results do suggest that there are no extraordinary cosmic-ray densities in these objects.

References

Bartel, N., Ratner, M. I., Shapiro, I. I., Rogers, A. E. E., and Preston, R. A., 1987, ApJ, 323, 505.

Bussard, R. W., Burrows, A., and The, L. S., 1989, ApJ, 341, 401.

Chan, K. W., and Lingenfelter, R. E., 1987, ApJ, 318, L51.

Clayton, D. D., Colgate, S. A., and Fishman, G. J., 1969, ApJ, 155, 75.

Cohen, R. S., Dame, T. M., Garay, G., Montani, J., Rubio, M., and Thaddeus, P., 1988, ApJ, 331, L95.

Fichtel, C. E., Özel, M., Stone, R., and Sreekumar, P., 1991, ApJ, 374, 134.

Ginzburg, V. L., 1972, Nature, 239, 8.

Ginzburg, V. L., and Ptuskin, V. S., 1976, Rev. Mod. Phys. 48, 161.

Huo, J. D., Hu, D. L., Zhou, C. M., Han, X. L., Hu, B. H., and Wu, Y. D., 1987, Nuclear Data Sheets, 51, 1.

Iyudin, A. F. et al., 1994, A&A, 284, L1.

Kurfess, J. D. et al., 1992, ApJ, 399, L137.

Leising, M. D., and Share, G. H., 1990, ApJ, 357, 638.
Leising, M. D., et al., 1995, 450, 805.
Mathewson, D. J., and Ford, V. L., 1984, Structure and Evolution of the Magellanic Clouds, ed. S. van den Berg, and K. S. de Boer, 125.
Mathewson, D. S., Ford, V. L., and Visvanathan, N., 1986, ApJ, 301, 664.
Mathewson, D. S., Ford, V. L., and Visvanathan, N., 1988, ApJ, 333, 617.
Matz et al., 1988, Nature, 331, 416.
Murai, T., Fujimoto, M., 1980, Publ. Astronomical Society of Japan, 32, 581.
Nomoto, K. et al., 1988, in Physics of Neutron Stars and Black Holes, ed. Y. Tanaka, (Tokyo: Universal Academy Press), 441.
Özel, M., and Fichtel, C. E., 1988, ApJ, 335, 135.
Paglione, T. A. P., Marscher, A. P., Jackson, J. M., and Bertsch, D. L., 1996, ApJ, 460, 295.
Pinto, G., and Forestini, M., 1991, in Gamma Ray Line Astrophysics, ed. Durouchoux and Prantzos, (New York: AIP), 183.
Ramaty, R., and Lingenfelter, R. E., 1995, The Analysis of Emission Lines, ed. R. E. Williams, and M. Livid, Cambridge: Cambridge Univ. Press.
Sreekumar, P., and Fichtel, C. E., 1991, Astron. Astrophys., 251, 447.
Sreekumar, P. et al., 1992, ApJ, 400, L67.
Sreekumar, P. et al., 1993, Phys. Rev. Letters, 70, 127.
Sreekumar, P. et al., 1994, ApJ, 426, 103.
The, L. S., Bridgman, W. T., and Clayton, D. D., 1994, ApJS, 93, 531.
Tueller, J., Barthelmy, S., Gehrels, N., Teegarden, B. J., Leventhal, M., and MacCallum, C. J., 1990, ApJ, 351, L41.

CHAPTER 8

ACTIVE GALAXIES

8.1 Introduction

Active galaxies are much less common than ordinary galaxies, but have a far greater luminosity. Active Galactic Nuclei (AGN) are defined by their having highly luminous and variable emission from their cores, which remain unresolved even on the milliarsecond level. They include Seyfert galaxies, radio galaxies, and BL Lacertae objects (BL Lacs). The bolometric luminosities, assuming istropic emission, are in the range from from 10^{42} tp 10^{49} ergs s^{-1}. The extraordinary energy source is now generally believed to be accretion onto a supermassive black hole with a mass in the range of 10^6 to 10^{10} $M_\odot$.

The contribution of gamma-ray astronomy to the study of AGN has increased remarkably in recent years. In 1980, when the writing of the first edition of this book was finished, there were three known gamma-ray emitting AGN. This number grew to just under twenty by the launch of the Compton Gamma-Ray Observatory in April, 1991; however, only one had been seen in the high-energy gamma-ray energy region. At the end of 1995, about 25 had been seen in the low-energy gamma-ray region, and over 50 have been seen in the gamma-ray energy range above 100 MeV.

When the gamma-ray emission from the AGN is examined, it becomes apparent that those observed fall into two groups. There are the Seyfert galaxies (some of which are also classified as radio galaxies), which are only seen in the low-energy gamma-ray region. Generally, the spectrum of these begins to fall rapidly above about 0.1 MeV. Only one is seen above 1 MeV, and none are seen above 10 MeV. In addition, the low-energy gamma-ray AGN are

relatively close on a cosmological scale, having z values of about 0.06 or less.

The other group consists of the high-energy gamma-ray emitters observed by EGRET. Here the high-energy gamma-ray emission dominates, and only about a half a dozen of the more than 50 that have been detected are even seen at energies below 30 MeV. The objects emitting high-energy gamma rays appear to belong to the class of objects called blazars, including BL Lac objects and flat-spectrum, radio-loud quasars (FSRQs). Among their other properties, blazars are believed to have jets directed toward Earth. For the observed gamma-ray-emitting blazars, the z distribution is similar to the radio-loud, flat-spectrum quasars, with over half having z values greater than 1, and a few having z values greater than 2. The luminosity, spectral features, z distribution, and association with different types of AGN suggest different types of emission processes. The association of the gamma radiation from blazars with a jet origin is a natural hypothesis, which, in fact, now has observational support, as will be discussed. Because of this natural separation into two groups, the discussion of the gamma-ray-emitting AGN will now be divided into two sections, which follow.

8.2 Seyfert and Radio Galaxies

Seyfert galaxies have very bright nuclei with strong emission lines in their spectra. Their energy release typically far exceeds that of a normal galaxy. Through a traditional optical telescope, a Seyfert galaxy appears as a bright star surrounded by a faint envelope. In addition to their high luminosity, the broad emission lines are also an indication of an exceptionally high level of activity. Seyfert galaxies have been found to range from ordinary to barred spirals, but there seem to be few ellipticals. One of the most striking aspects of Seyfert galaxies is the variability in their luminosity. The relatively small radiating volume implied by this variability, combined with the high luminosity, places severe constraints on theoretical models for the radiation and is the basis for the belief that there is a massive black hole at the center.

Radio galaxies comprise a rather broad class of galaxies that are strong radio emitters without exceptional optical properties. The majority of the identified radio galaxies are extended and show little evidence of self-absorportion or radio emission. The source of the

radio emission is often thought to be synchrotron radiation. Centaurus-A (NGC 5128), the closest and most intense radio galaxy observed, has been detected in all frequency bands from the radio through the medium-energy gamma-ray range.

Prior to the launch of the Compton Gamma-Ray Observatory, just over a dozen Seyfert and radio galaxies had been seen in the hard x-ray to low-energy gamma-ray region, specifically 0.020 to 0.165 MeV, by HEAO-1, Exosat, and Ginga. The OSSE instrument on the Compton Gamma-Ray Observatory has added several more in the 50-150 keV region. Only one Seyfert or radio galaxy, Centaurus-A, has been seen above 1 MeV, and none were seen above 30 MeV by EGRET. A list of these objects is given in Table 8.1.

NGC 4151, a Seyfert 1 galaxy, is the nearest Seyfert and the brightest radio-quiet AGN in the low-energy gamma-ray region. NGC 4945 is the brightest Seyfert 2 galaxy (at 0.1 MeV). An early result of the OSSE observations (Maisack et al., 1992a) showed that NGC 4151 had a thermal-like spectrum, contrary to some early suggestions that its spectrum was harder. Johnson et al. (1994) deduced an average spectrum for Seyferts by summing spectra of weak Seyferts observed by OSSE over a period of more than 2 years. The resulting spectrum, which is also well fit by a thermal spectrum, is shown in Figure 8.1 and compared to the NGC 4151 spectrum. It is seen that the two spectra are very similar in shape. Zdiarski et al. (1995) have shown that Seyfert 1, Seyfert 2, and radio-loud Seyfert 1 galaxies may all be fitted by absorbed power-law spectra with exponential cutoffs combined with a reflection and Fe line contribution. The contributions of the different components, however, appear to be different for the 3 types of Seyferts. A good example of this type of spectral analysis exists for IC 4329A (Madejski et al., 1995) and is shown in Figure 8.2. The Seyfert galaxies show strong neutral absorption, and there is an indication that they are intrinsically harder. The associated model is Comptonization in a relativistic, optically-thin, thermal corona above the surface of an accretion disk. The differences in spectral shape are associated with the spectral index and the amount of reflection.

These results taken together are in agreement with the currently popular concept that AGN are powered by accretion onto a massive

Table 8.1.

Seyfert Galaxies Detected Above 50 keV

Source	Other name	Type	z	20-165 keV	50-150 keV	1-30 keV	>100 MeV
0316+413	NGC 1275	Sy2	0.0183	Y	Y		
0415+379	3C 111	Sy1	0.049		Y		
0430+052	3C 120	Sy1	0.0334	Y	Y		
0551+467	MCG +8-11-11	Sy1	0.0205	Y	Y		
0609+710	MRK 3	Sy2			Y		
0945-307	MCG -5-23-16	Sy2	0.0082	Y	Y		
1020+201	NGC 3227				Y		
1136-374		Sy1	0.0092	Y			
1208+396	NGC 4151	Sy1	0.0033	Y	Y		
1223+129	NGC 4388	Sy2	0.0087		Y		
1232-396	NGC 4507	Sy2			Y		
1302-492	NGC 4945	Sy2			Y		
1322-428	Cen A	Sy2	0.0008	Y	Y	Y	
1333-340	MCG -6-30-15	Sy1	0.0078		Y		
1346-300	IC 4329A	Sy1	0.0157		Y		
1351+695		Sy1	0.031	Y			
1410-029	NGC 5506	Sy2			Y		
1415+253	NGC 5548	Sy1	0.017	Y	Y		
1845+797	3C 390.3	Sy1	0.0569	Y	Y		

Table 8.1 (continued).

Source	Other name	Type	z	20-165 keV	50-150 keV	1-30 keV	>100 MeV
1939-104	NGC 6814	Sy1	0.0053	Y	Y		
2041-109	MRK 509	Sy1	0.0355	Y	Y		
2159-321	NGC 7172	Sy1			Y		
2300+086	NGC 7469	Sy1			Y		
2302-090	MCG -2-58-22	Sy1	0.048		Y		

Sy= Seyfert

Note 1: With some exceptions, the columns refer to detections by instruments as follows:
20-165 keV: HEAO-1, Exosat, and Ginga; 50-150 keV: OSSE and Sigma; 1-30 MeV: COMPTEL; > 100 MeV: EGRET. Most of those detected in the 50-150 keV range have been seen in hard x-rays by HEAO-1, Exosat, or Ginga. See Dermer and Gehrels (1995).

Note 2: A "Y" indicates a detection of significance greater than 3.5σ.

Note 3: Upper limits for the >100 MeV column are typical (0.3 to 1.5) $\times 10^{-7}$ ph cm^{-2} s^{-1}. (Fichtel et al., 1994; Lin et al., 1993; von Montigny et al., 1995a; Thompson et al., 1995; Thompson et al, 1996).

Note 4: The above table with significant modifications is from Dermer and Gehrels (1995); references to the very large number of papers on the individual objects are given there. The principal modifications are the inclusion of the more recent low-energy gamma-ray data from OSSE (Kurfess, 1996).

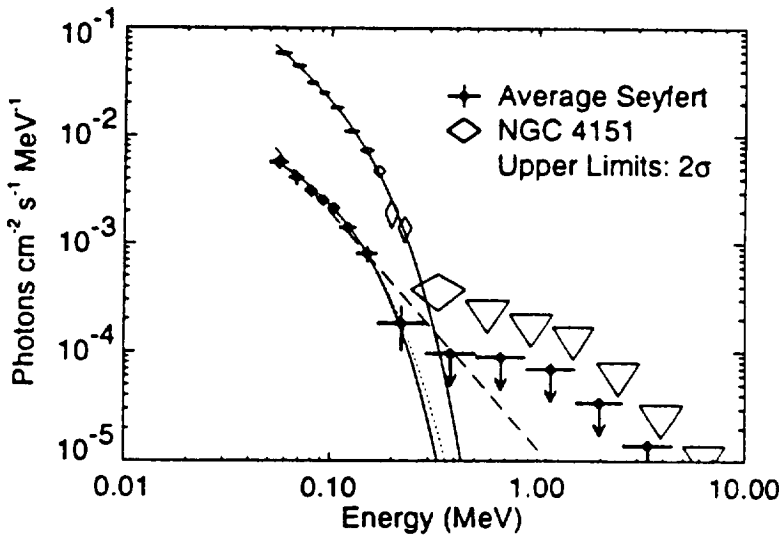

Figure 8.1. Average spectrum of weak Seyfert galaxies observed by OSSE. The spectrum of NGC 4151 is shown for comparison. The figure is from Kurfess (1995); the data on the average Seyfert is from Johnson et al. (1993). The figure is reproduced with their permission.

black hole, and many of the observed characteristics are determined by the observer's line of sight, relative to the thick torus which is postulated to envelop the central object. In this model, the distinction between Type 1 and Type 2 Seyferts is determined by whether or not the line of sight is obstructed by the torus, with the Type 2 being obstructed. The line of sight is not aligned, or nearly aligned with the jet in either case, if a jet exists at all in these cases. The sharp decrease in the spectrum near 0.1 MeV also severely limits the contribution of the Seyfert galaxies to the diffuse radiation, to be discussed in Chapter 9.

Attention will now be turned to Centaurus-A (1322-42; NGC 5182). It is appropriate that it should come last in this section, because it is a somewhat peculiar source that may belong in between the category of this section and the next. It is the brightest radio source of all, and was one of the first radio sources identified to be extragalactic.

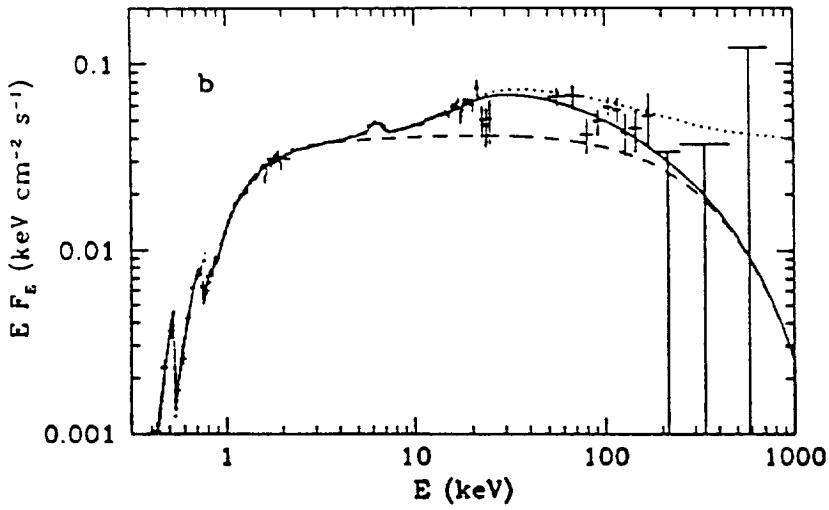

Figure 8.2. The broad-band spectrum (in EF_E) of IC 4329A. The crosses give the data; the solid curves represent the best-fit power law with an exponential cutoff, reflection, Fe K line, and absorption; the dashed curves represent the power-law component with the cutoff, but without the contribution from reflection. It is apparent that the reflection contribution is important in the OSSE energy range. For comparison, the dotted curves show the best-fit spectrum for the model without the cutoff. The upper limits are 2σ. ROSAT (< 2 keV), OSSE (> 50 keV), the nonsimultaneous Ginga data (2-25 keV) are shown. The model without the cutoff is ruled out as the spectral index is fixed by the Ginga data. The figure is reproduced with the permission of Madejski et al. (1995), and references to the many data points are given there.

It is the closest radio galaxy with a z of 0.0018. The radio structure is very complex, showing two outer lobes which extend over several degrees, and an inner core component plus two inner lobes with bent structure. The optical appearance resembles an elliptical galaxy. There is a jet that is observed at several frequencies. A number of properties suggest that it may be a blazar, but not beamed in the direction of the Earth. It has also been proposed (Baade and Minkowski, 1954) that it is actually two interacting galaxies. The existing low-energy gamma-ray data (See specially Kinzer et al., 1995) seem to support the concept that it is a blazar with the jet not beamed exactly in our direction, since, although there is no high-

energy gamma-ray emission, the low-energy emission extends to higher energies than any of the other AGN listed in Table 8.1. The rather large amount of data on this object is shown in Figure 8.3. As can be seen in the figure, there is not the sharp decrease at a few tenths of 1 MeV seen in Figure 8.1, and evidence exists for emission beyond 1 MeV. However, there is the upper limit to the high-energy gamma-ray flux from the EGRET data.

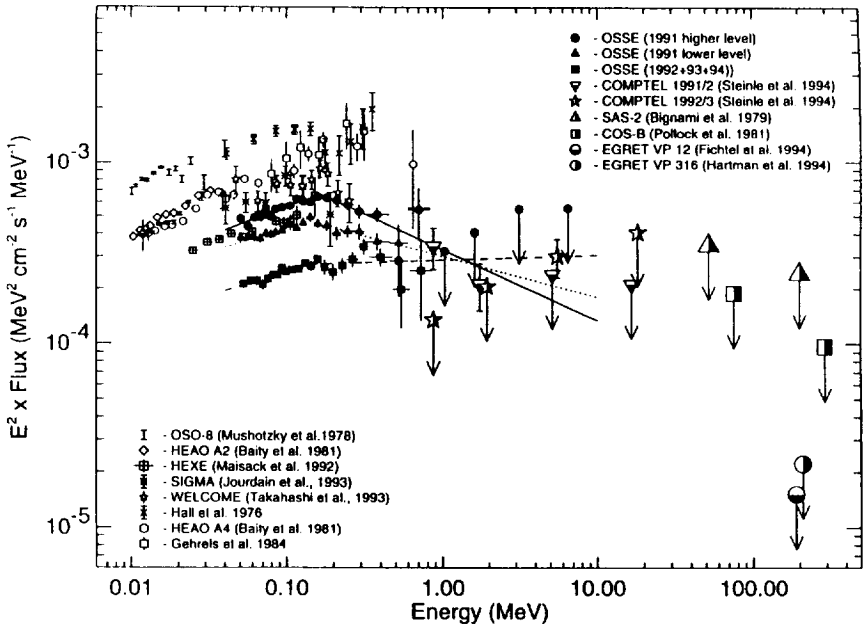

Figure 8.3. Summary plot of selected historical x-ray and gamma-ray spectral measurements for Cen A, together with spectra from the three relatively distinct intensity levels observed by OSSE. Measurements included have accurate spectral measurements above 10 keV; data points below 10 keV were omitted. For clarity, less than 2-σ data points and upper limits have been left off except in the MeV region where few positive detections exist. In this E^2 x dN/dE plot, a power law with $\Gamma = 2$ would appear as a horizontal line. Best-fit broken power-law spectra for the three OSSE observations are shown by a solid line (1991 higher intensity level), a dotted line (1991 intermediate intensity level), and a dashed line (1992-93-94 lower intensity level). The figure is reproduced with the permission of Kinzer et al. (1995).

8.3 Blazars

As mentioned before, one of the most dramatic results in the decade of the 1990s has been the discovery by EGRET of over 50 gamma-ray-emitting quasars (Fichtel et al., 1994; Fichtel, 1994; von Montigny et al., 1995a; Thompson et al., 1995; Mukherjee et al., 1997). Prior to the launch of the Compton Gamma-Ray Observatory only one AGN had been seen in the 100 MeV region. That was 3C 273 (Swanenburg et al., 1978), and its observed intensity was just over the threshold for detection by COS-B. The quasars observed by EGRET are all radio-loud, flat-spectrum radio sources. Many of them are known to be superluminal. They all appear to be members of the blazar class of AGN (containing BL Lac objects, highly polarized quasars, and optically violent variables), which have jets beamed in the general direction of the Earth.

Since one blazar had been seen before the launch of the Compton Gamma-Ray Observatory, it was anticipated that there would be many more seen by EGRET with its over one order of magnitude greater sensitivity, and there was no disappointment. Shortly after the launch in one of the early observations, there was an unexpected result, the first of several. EGRET is a wide-field telescope, and it was looking in the Virgo direction. It was expected that 3C 273 would be seen, and it was, but to everyone's surprise, a very much stronger source was also in the field of view. It was 3C 279 (Hartman et al., 1992). This was the first indication of the highly variable nature of these objects in high-energy gamma rays. Another feature to emerge was that it was not just the closer ones that were being seen, but there was a broad distribution in z. More than half of those now observed have z values greater than 1.0, and a few have z values greater than 2.0.

The implications associated with the tremendous energy being emitted in high-energy gamma rays by these distant objects are many and impressive, but it seems best to begin with a summary of the information that exists at the present. Table 8.2 provides a list of the blazars that have been seen. Notice immediately that, contrary to Table 8.1 for Seyfert and radio galaxies, this table is dominated by observations at high energy by EGRET (e.g., von Montigny et al., 1995a), with only a few of the blazars having even been seen at low energies. This is a clear indication of the difference between these two groups; more differences will become apparent.

Table 8.2.

Flat-Spectrum, Radio-Loud Galaxies Detected Above 50 keV

Source	Type	z	50-150 keV	1-30 MeV	>100 MeV
0202+149	FS,OV				Y: <.18-.24
0208-512	FS,HP	1.003			Y: .38-1.10
0234+285	FS,SL, HP	1.213			Y:.15-.29
0235+164	FS,BL, SL?,HP	0.94			Y: <.31-.83
0420-014	FS,HP, OV	0.915			Y: <.14-.50
0446+112	FS	1.207			Y:<.19-1.14
0454-234	FS,HP	1.009			<.20
0454-463	FS	0.858			<.15
0458-020	FS,OV	2.286			Y:<.10-.31
0521-365	FS,HP	0.055			Y:.19-.21
0522-611	FS	1.4		Y	<0.08
0528+134	FS,OV	2.06	Y	Y	Y:<.46-2.95
0537-441	FS,BL, HP	0.894			Y: <.23-.36
0716+714	FS,BL, SL?,HP				Y: <.18-.24
0804+499	FS,HP	1.43			Y: <.15-.21
0827+243	FS,OV	2.046			Y:.24-.26
0829+046	FS,BL, OV	0.18			Y:.19
0836+710	FS,SL, HP	2.17			Y: <.07-.45
0954+658	FS,BL, HP	0.368			Y: <.09-.11
1101+384	FS,BL	0.031			Y:.12-.23
1156+295	FS,SL, HP,OV	0.729			Y:<.04-2.29
1219+285	FS,BL, HP,OV	0.102			<.06
1222+216	FS	0.435			Y: <.12-.30
1226+023	FS,SL, HP,OV	0.158	Y	Y	Y: <.17-.24

Table 8.2 (continued).

Source	Type	z	50-150 keV	1-30 MeV	>100 MeV
1227+024	QSO	0.57	Y		<0.10
1229-021	FS	1.045			Y: <.11-.14
1253-055	FS,SL, HP,OV	0.538	Y	Y	Y:<.28-2.87
1313-333	FS	1.21			Y:.17-.26
1406-076	FS	1.494			Y:<.10-1.44
1510-089	FS,HP, OV?	0.361			Y: <.18-.48
1517+656	BL		Y		<0.06
1604+159	BL	0.357			Y: <.17-.40
1606+106	FS	1.23			Y: <.28-.60
1611+343	FS,OV	1.401			Y: <.09-.55
1622-253	FS	0.786			Y:<.38-.43
1633+382	FS,SL, OV	1.81			Y:.31-1.05
1739+522	FS	1.375			Y: <.27-.53
1741-038	FS,HP	1.054			<.21
1933-400	FS	0.966			Y: <.08-.97
2005-489	FS,BL, OV	0.071			<0.14
2022-077	FS				Y: <.22-.73
2052-474	FS	1.489			Y: <.11-.25
2155-304	BL	0.116	Y		<.15
2209+236	FS				Y:.13-.14
2230+114	FS,SL?, HP	1.037	Y	Y	Y:.27-.29
2251+158	FS,SL, HP,OV	0.859	Y		Y: .81-.132
2356+196	FS	1.066			Y: <.22-.30

Sy=Seyfert, RG = radio galaxy, Sub = subluminal, SL = superluminal, FS = flat-spectrum radio galaxy, BL = BL Lac, HP = high optical polarization, OV = optically violent variable (OVV).

Note 1: With some exceptions, the colums refer to detections by instruments as follows: 50-150 keV: OSSE and Sigma; 1-3 MeV: COMPTEL; > 100 MeV: EGRET.

Note 2: A "Y" indicates a detection of significance greater than 3.5σ.

Note 3: Numbers in the last column are fluxes or upper limits. Units and references are: 10^{-6} ph $cm^{-2}s^{1}$ (Fichtel et al., 1994; Lin et al., 1993; von Montigny et al., 1995a: Thompson et al., 1995). For clarification, as an example, <.20 - .80 means that, at different times, the flux has varied from less than .20 to a positive detecion of .80.

Note 4: The information in this table comes principally from Thompson et al. (1995), von Montigny et al. (1995a), Dingus et al. (1996), Lin et al. (1996), Dermer and Gehrels (1995), and McNaron-Brown et al. (1995); references to the very large number of individual references are given there.

Note 5: 1101+384 is Mrk 421, 1226+023 is 3C 273, 1253-055 is 3C 279, and 2230+114 is CTA 102..

Consider now the high-energy gamma-ray properties of these objects, starting by looking at the energy spectra, several of which are shown in Figure 8.4. These are representative spectra for those AGN for which there are good photon statistics. The data are well represented by a power law in photon energy. This is typical of the blazars that are observed. There is a range of slopes for the blazars, with the average being -2.1. The distribution in slopes will be discussed later in this section. In most cases, the fit is valid over the entire range in energy for which data exists, although in a few cases there is an indication of a downward break at the highest energies.

Only two blazars are seen at TeV energies with ground-based telescopes; these are Mrk 421 (Punch et al., 1992) and Mrk 501 (Weekes et al., 1996). It should be remembered that, at these energies, the intergalactic photons have a significant effect in absorbing the gamma rays; so one does not expect blazars to be detected beyond a z of about 0.4 to 0.5 (Stecker and De Jager, 1996, and references therein) at energies above about 0.5 TeV. Even the very strong source 3C 279 at $z = 0.54$ can be shown to be much too strongly absorbed to be detectable at 0.5 TeV, assuming the spectral slope seen in the 30 MeV to 10 GeV region continues to higher energies. Hence, only the closer blazars have even a chance of being seen in this very high-energy range.

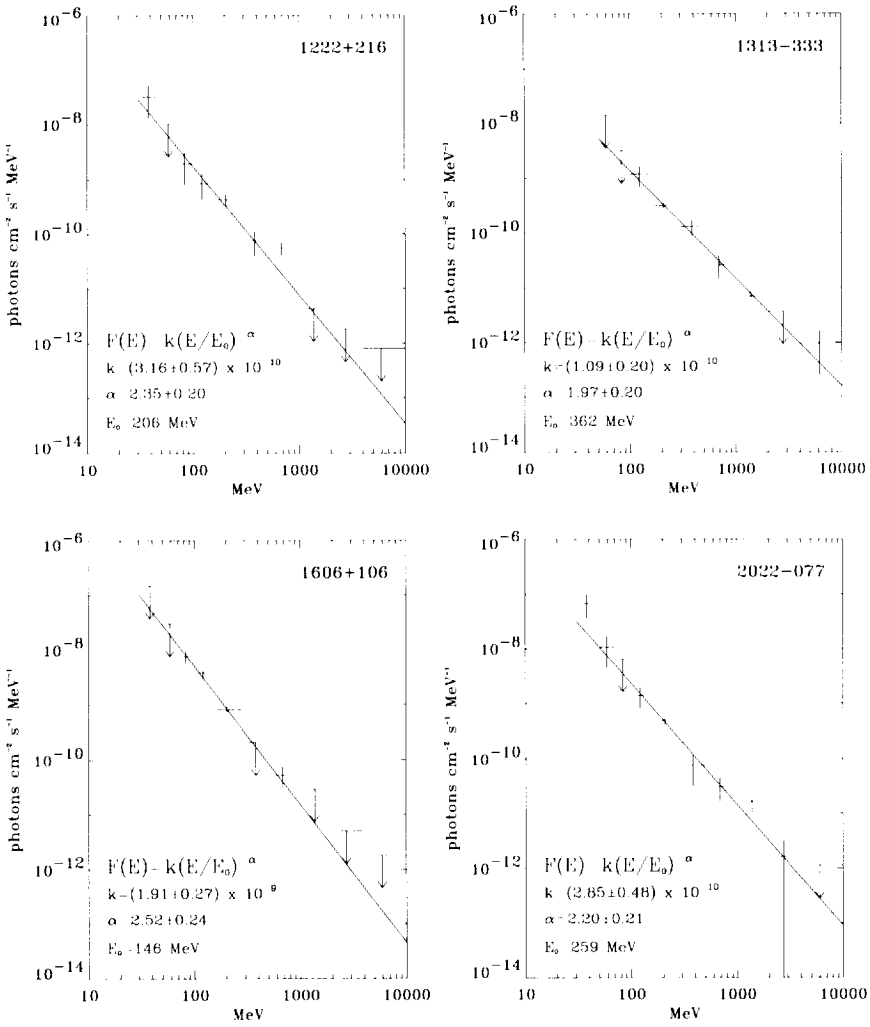

Figure 8.4. The energy spectrum of several blazars as a function of gamma-ray energy. The figure is adapted from von Montigny et al. (1995a) and reproduced with their permission.

In Figure 8.5, the intensity during the several times that EGRET observed PKS 0208-512 (Bertsch et al., 1993; von Montigny et al., 1995b) is shown. As seen, there is substantial variation, and this is typical of the other sources that have been viewed frequently by EGRET. Only a few of those that have been observed several times with adequate fluxes to make a determination have not provided evidence of variation – often quite large. The variation during the order of weeks and months is common in the high-energy gamma-ray observation of blazars that are seen to emit gamma rays. There is some indication that the high-energy gamma-photon spectrum is harder (i.e., the spectrum index is less negative) in the high-luminosity state, but this possible characteristic is not well established. For the few blazars seen in the low-energy gamma-ray range by OSSE, time variations are also seen (McNaron-Brown et al., 1995).

There are a few examples, when the intensity was strong enough in high-energy gamma rays, that variations of the order of days were observed. The best example of a short time variation is that of 3C 279 during the period of June 16 to 28, 1991, shown in Figure 8.6 (Hartman et al., 1992). Here, it is clear that there has been a major change in intensity during a period of 2 days, and a significant change within 1 day. This time variation can be interpreted as a limit for the size of the emitting region, which would be the distance that light travels in about a day, or approximately 3×10^{10} km multiplied by an appropriate relativistic time correction. This distance is very small compared to the dimensions of a galaxy, which are typically of the order of 10^7 times this value in diameter or more.

As another comparison, the Schwarzschild radius for a black hole is given by

$$R_s = \frac{2GM}{c^2}. \qquad (8\text{-}1)$$

For a 10^{10} $M_\odot$ black hole, the value of R_S is about 3×10^{10} km. Hence, it is about the same, except for the relativistic time correction. If the radiation were to be coming from the region of the black hole, it would have to be very close to it. There are a number of difficulties with this interpretation. First of all, the photon density is so high that it is unlikely the gamma rays could escape. Second, it will be difficult to find a mechanism that allows the

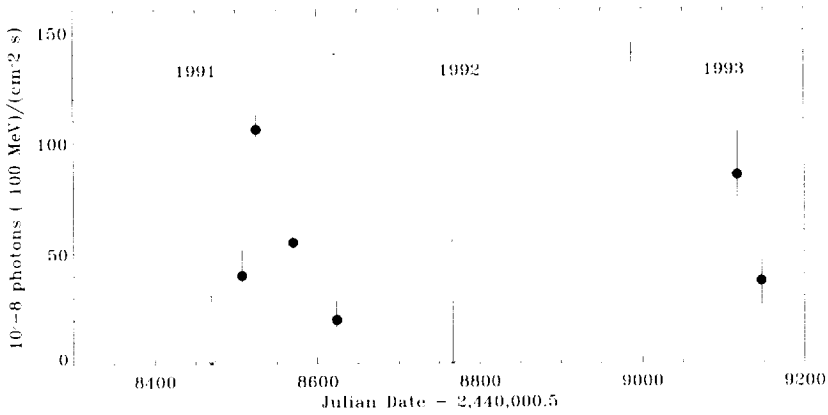

Figure 8.5. The intensity of high-energy gamma rays from PKS 0208-512 as a function of time for the periods when it was observed by EGRET. The figure is reproduced with the permission of von Montigny et al. (1995a).

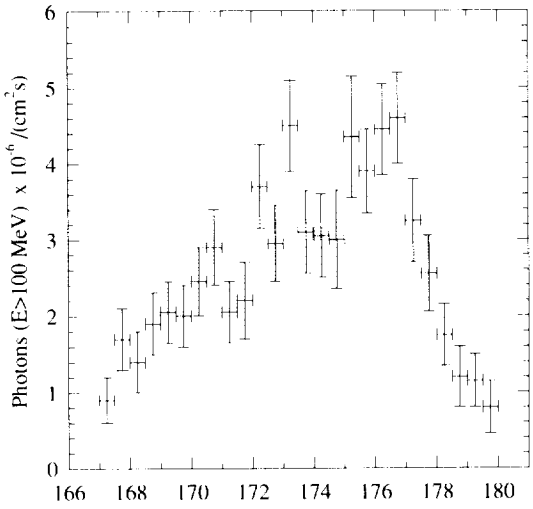

Figure 8.6. The intensity as a funtion of time for 3C 279 during the period of June 16 through 28, 1991. The figure is from Kniffen et al. (1993).

accelerated particles to interact and create the gamma-ray intensity observed. Another problem is that any origin that can be imagined would lead to isotropic emission placing even greater demands on the acceleration mechanism.

The currently more attractive hypothesis is that the gamma-ray emission occurs in the jet, probably near the core or in a shock region, in part, because of the apparent size of the region emitting the gamma rays. The jet concept leads naturally to beaming, which reduces the required energy supply by a very large factor. The subject of the origin of the gamma radiation and the accelerations of the parent relativistic particles will be pursued in Sections 8.6 and 8.7.

If the data are presented in the form of νF_ν, and combined with data from other frequencies, distributions such as those shown in Figure 8.7 are obtained. At least during the high-intensity gamma-ray state when the observations were made, the gamma-ray frequency interval dominates in energy as observed at the Earth. This dominance of the gamma-ray emission must be tempered by several considerations. First, the gamma-ray-emitting blazars vary significantly in intensity, and the average value may be much smaller. Second, only a small fraction of the known high-intensity, flat-spectrum radio AGN are seen in gamma rays. Further, this and other information suggests that the gamma-ray beam may be relatively small in solid angle extent. At the same time, some of the emission at other frequencies may come from a much broader beam, or even may be isotropic. If the emission were isotropic, typical energy emission rates would be 10^{49} ergs s^{-1} for the high-energy gamma rays. Even if the beam is concentrated in a cone that is 10^{-3} of isotropy, the emission is of the order of 10^{46} ergs s^{-1}. This is an extraordinary amount of energy, especially when one considers that there must be an equal or greater amount of energy supplied to the relativistic particles that lead to the creation of the gamma rays through interactions.

In summary, the power-law spectrum in photon energy over the energy range observed and the dominance of energy emission in the gamma-ray region, at least in the high state, appear to be common characteristics of the gamma-ray-emitting blazars that have been observed. Time variability appears to be a frequent, if not universal, characteristic of these objects as well.

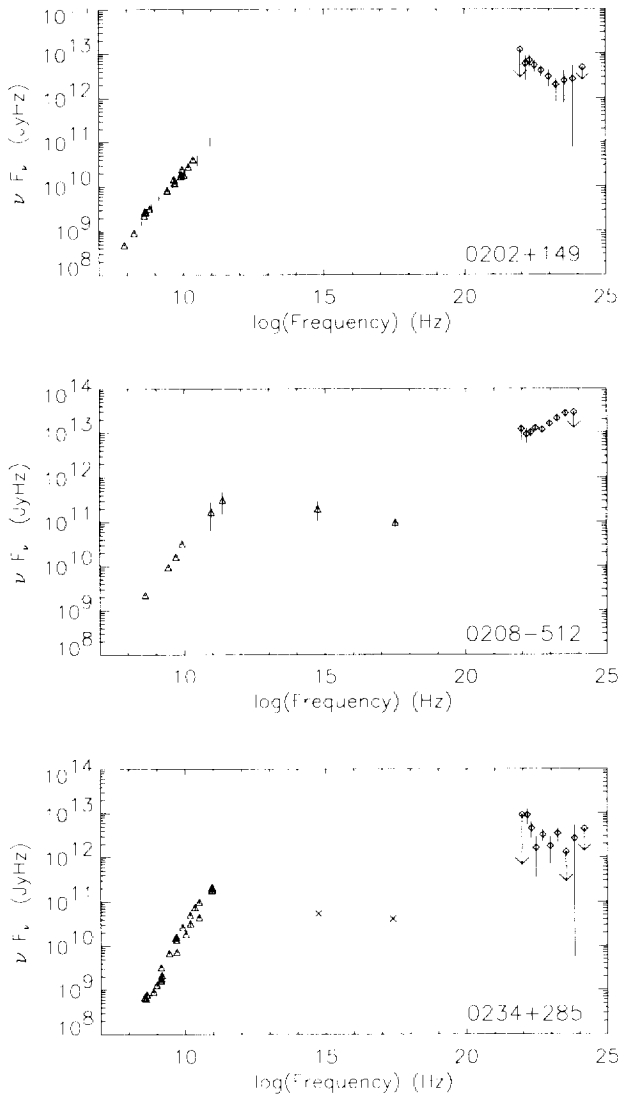

Figure 8.7. The energy observed from several gamma-ray-emitting blazars in Jy Hz as a function of Hz. The figure is reproduced with the permission of von Montigny et al. (1995a).

There are a few cases where there is multifrequency information on gamma-ray-emitting blazars in both the low and the high gamma-ray intensity state. From this limited sample, there is a suggestion that there is a tendency for the gamma-ray intensity to vary more than that at other frequencies. There is even an indication that the amount of the change may increase with photon frequency. The best example of the variation across the electromagnetic spectrum is for 3C 279, as shown in Figures 8.8 and 8.9. Here the increasing difference with frequency for the two times of observation is very clear. In Figure 8.8, although there is almost no perceptable change in the radio region between June and October, 1991, there is about a factor of four change in the high-energy gamma-ray region (Hartman et al., 1996). In Figure 8.8 it is seen that the variation is even more pronounced between the June 1991, and the December 1992 to January 1993 period, with the gamma-ray intensity changing by a factor of about twenty, while the changes at lower frequencies are much smaller (Maraschi et al., 1994). Although caution must be used because of observed short-term fluctuations and the fact that not all of the observations are simultaneous, the increase in the variation as a function of frequency for long-time scales does seem to be a characteristic phenomenon, based, admittedly, on only a few samples.

On the other hand, a study by Mücke et al. (1996) showed no measurable correlation between short-term variations in radio and gamma-ray luminosity, in general. One possible explanation of this phenomenon is that the gamma radiation is almost entirely caused by high-energy particles that are beamed so that, if there is a major change in intensity or even a relatively small change in direction of these particles, there can be a large change in the high-energy gamma radiation. The radio and optical radiation may be broader in angular extent, partially because of secondary processess, or arise in part or primarily from other or secondary origins. Hence, on a longer time scale, there would be some correlation, but the lower frequency changes would be smaller and possibly later. This subject will be considered in more detail in Section 8.6.

Although there is a significant number of radio-loud, flat-spectrum quasars that have been seen in high-energy gamma rays, it is important to recognize that there are a few that might have been expected to have been observed in light of their radio intensity and

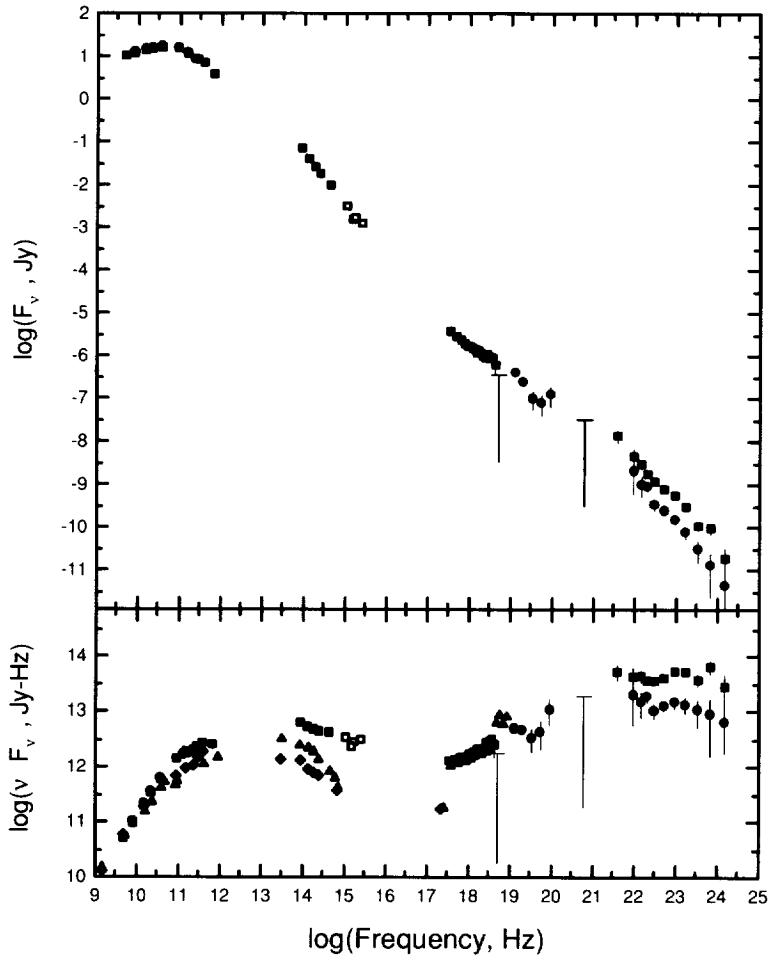

Figure 8.8. The differential intensity of 3C 279 for June 1991 (upper points) and October 1991 (lower points) as a function of frequency. The figure is from Hartman et al. (1996), and references to the several different sets of data are given therein.

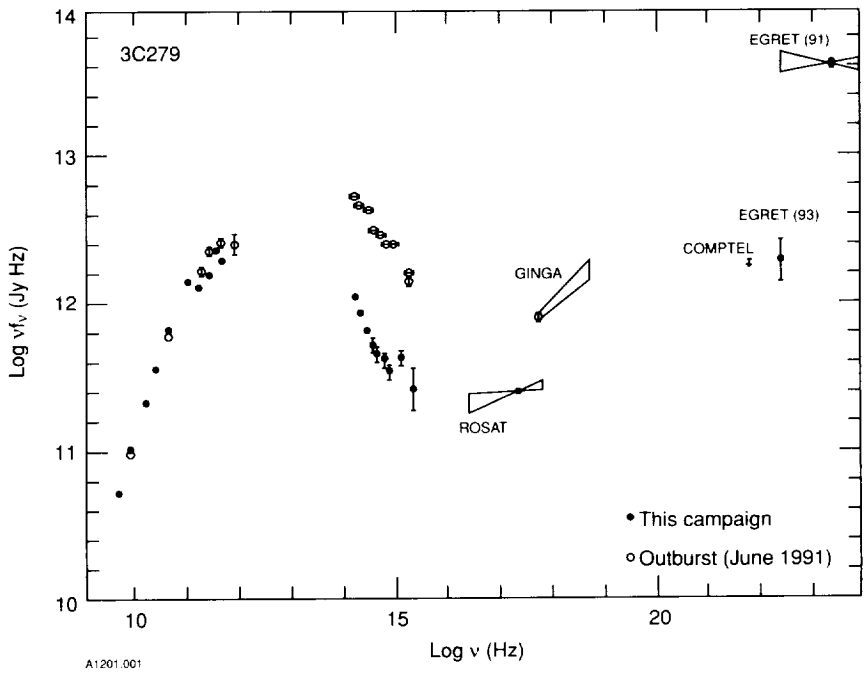

Figure 8.9. The observed intensity in Hz Jy for 3C 279 as a function of frequency for June 1991 (upper points) and the period December 1992 to January 1993 (lower points) as a function of frequency. The figure is reproduced with the permission of Maraschi et al. (1994), and references to the several different sets of data are given therein.

spectral slope, but were not. First, it must be remembered that most or all of the ones that have been seen with any frequency are either variable in their gamma-ray emission or consistent with being variable. Hence, one explanation might be that, at the time of the EGRET observation, the blazar was not in a high state. Alternately, for some reason, the gamma-ray beam may not be directed at the Earth, or it may be that these blazars that have not been seen by EGRET (see von Montigny et al., 1995b, for a list) are just gamma-ray quiet.

Although the observations at the moment point to the gamma-ray emission from blazars being a predominately high-energy gamma-ray phenomenon, with more than 50 having been seen in high-energy gamma rays by EGRET, 6 have been seen in the medium-

energy gamma-ray range (1 to 30 MeV) by COMPTEL, and a similar number in the low-energy range (50 to 150 keV) by OSSE. Not surprisingly, the 5 seen in both the medium- and low-energy range are among the more intense blazars seen in the high-energy range, namely, 0528+134, 1226+023 (3C 273), 1253-055 (3C 279), 2230+114, and 2251+158. (See von Montigny et al., 1993; Johnson et al., 1995; Lichti et al., 1995; McNaron-Brown et al., 1995.) The energy spectrum of simultaneous, or nearly simultaneous, measurements of 3C 273 is shown in Figure 8.10. This figure is similar to the ones of spectra shown earlier in this chapter, with the two exceptions that, although there is a relatively high-energy flux in the gamma-ray range, there are points just as high in the Ultraviolet (IUE) interval, and the maximum in the gamma-ray range is in the low-energy gamma-ray range, not in the high-energy range. The combination of some of the upper limits in the medium-energy range, 1 to 30 MeV, and the observations just mentioned suggest that the energy spectrum does not continue with the same spectral slope, but rather the differential intensity begins to fall below this extension in the medium-energy range. The data on the 5 blazars mentioned above seen in all three energy ranges suggest that, as the gamma-ray energy becomes lower, a flux maximum is reached between 1 and 20 MeV, perhaps typically 5 MeV. Below that energy, the spectral slope is usually -1.5 to -1.9 rather than tyically -2.1 in the high-energy range.

There are seven very-high-energy ground-based gamma-ray instruments whose thresholds range from approximately 0.25 to 3.0 TeV. There have been many attemps by these instruments to detect AGN, but only two have been seen Mrk 421 (Punch et al., 1992) and 501 (Weekes et al., 1996). Mrk 421 is also seen by EGRET, as described here, and the 0.3 TeV flux is consistent with an extension of the spectrum observed by EGRET. Mrk 501 is not, but it is weaker that Mrk 421 in the very-high-energy range. If its spectrum is similar to Mrk 421, its flux would be below the detectable limit in the EGRET range. Of the other blazars seen by EGRET, if their spectra are extended unchanged in slope to the very-high-energy range, most of the upper limits in the very-high-energy range fall above or within the range of the spectral slope extension and its uncertainty. See e.g., Kerrick et al. (1995) and Quinn et al. (1995). Two fall well below the range of the extension of the spectral slope determined by EGRET, namely 1633+382 and 2022-077, and a few others, including 0235+164, 3C 279 (1253-055), and 2251+158 are

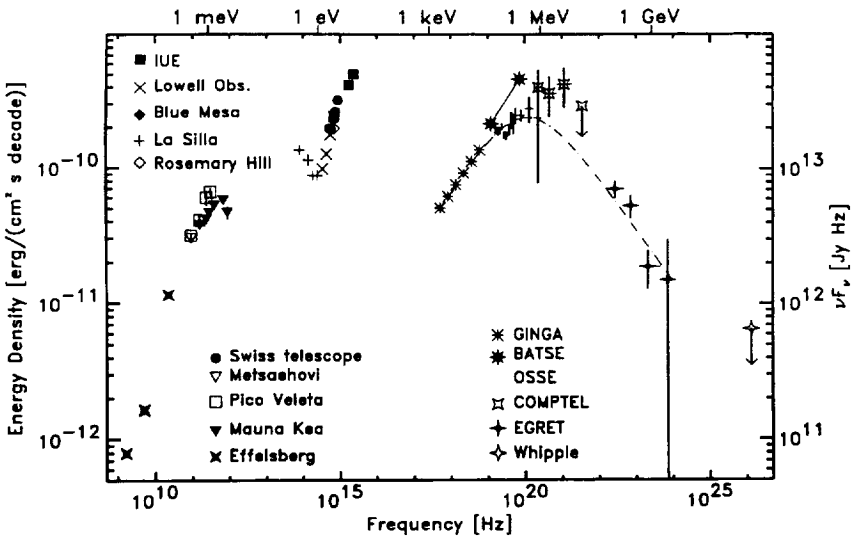

Figure 8.10. Energy per decade spectrum of simultaneous observations of 3C 273. The figure is reproduced with the permission of Lichti et al. (1995).

near the lower bound of the spectral slope extension range (Kerrick et al., 1995).

In considering these results, it must first be remembered that most, and perhaps all, of the high-energy gamma-ray-emitting blazars are time variable, as noted earlier, and the EGRET and very-high-energy measurements were usually not made at the same time. Further, of the five mentioned above for which the z values are known (the z value for 2022-077 is not available), all have a z in excess of 0.5, and 1633+382 has a z of 1.81. At the energies being considered, the dominant process for absorption in intergalactic space is pair production on infrared photons (e.g., Stecker, de Jager, and Salamon, 1992; Dwek and Slavin, 1994). For the energy range involved with the ground-based detectors and the z range just mentioned, the absorption length is estimated to be from 1 to over 10 optical depths; hence, substantial absorption is expected. Even ignoring the time-dependent effect, which could be substantial, none of the results is inconsistent with a power-law spectrum in energy that simply continues into the very-high-energy region at the source, but is effected by absorption in intergalactic space. Hence, for the present, there is no compelling known evidence for a change

in spectral slope at or near the source between the high-energy gamma-ray range of EGRET and that very-high-energy gamma-ray interval of the present ground-based telescopes, although one may, of course, be present. By the end of 1997, a relatively-sensitive, large, ground-based array called CELESTE is expected to be in operation in France. It covers the energy range from approximately 20 to 200 GeV. It should, therefore, provide the spectral information needed to address this matter.

8.4 The Evolution of Gamma-Ray-Emitting Blazars

Although the conclusions that can be drawn are limited at present because of the small sample, there is already a sufficient number of radio-loud, flat-spectrum quasars observed to be emitting high-energy gamma rays that have measured z values to be able to say with reasonable certainty that there has been evolution of these objects. The evolution of flat-spectrum, radio-loud quasars (FSRQs) is well established. Thus, by a simple comparison of the FSRQs' distribution in z to that of the high-energy gamma-ray-emitting blazars, one can see if they are similar. Figure 8.11 clearly shows that they are, although the number of detected gamma-ray emitters with known z values is much smaller than that of the FSR quasars.

There will be a more detailed discussion of cosmology in Chapter 9; here the discussion will be limited to that which is needed for the immediate considerations. The relationship between the observed differential energy flux $S_o(E_o)$, where the subscript "o" denotes the observed, or present value, and Q_e, the power emitted in dE at E, where $E = E_0(1+z)$ in the Friedman universe is given by

$$S_o(E_o) = \frac{Q_e[E_o(1+z)]}{4\pi\Theta D_L^2}, \tag{8-2}$$

where

$$D_L = \frac{c}{H_o q_o^2}\left[1 - q_o + q_o z + (q_o - 1)(2q_o z + 1)^{\frac{1}{2}}\right]. \tag{8-3}$$

H_o is the Hubble parameter, q_o is the deceleration parameter, and Θ is the beaming factor.

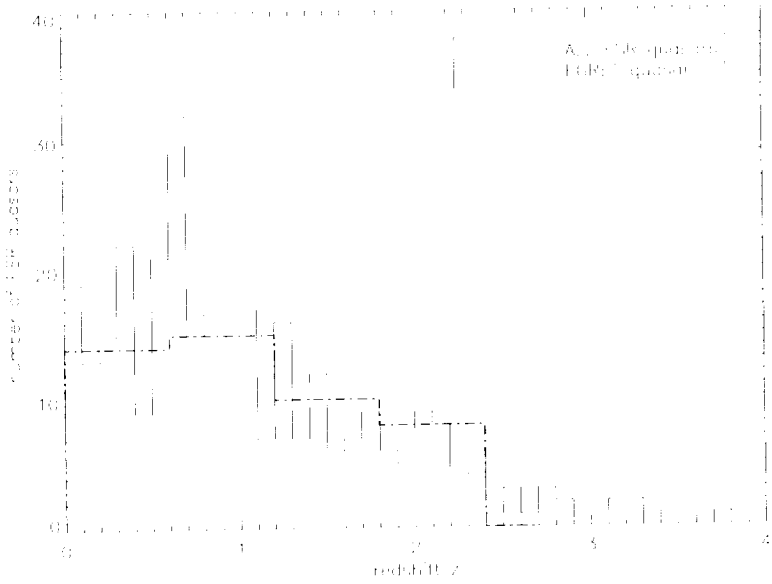

Figure 8.11. The distribution in z for flat-spectrum, radio-loud quasars and high-energy gamma-ray-emitting blazars (EGRET). Note that the number scale for the radio quasars is on the left, and the gamma-ray-emitting blazars is on the right. The figure was developed from the work of Fichtel et al. (1996) and Mukherjee et al., (1997).

For $q_o z << 1$,

$$D_L \approx \frac{cz}{H_o}\left[1 + 0.5z(1 - q_o)\right]. \tag{8-4}$$

In general, it is convenient to write

$$D_L = \frac{cz}{H_o} g(z, q_o). \tag{8-5}$$

where

$$g(z, q_o) = \frac{1}{zq_o^2}\left[1 - q_o + q_o z + (q_o - 1)(2q_o z + 1)^{\frac{1}{2}}\right].$$

For small z, $g(z,q_o)$ approaches 1, and $g(z,1)$ is rigorously 1.

Hence, e.g., (8-2) becomes

$$S_o(E_o) = \frac{Q_e[E_o(1+z)] \cdot H_o^2}{4\pi c^2 z^2 g^2(z,q_o)\Theta}. \tag{8-6}$$

If the photon source spectrum is a power law in E with an exponent −b, so that the energy-flux spectrum is a power law with an exponent −b+1, and, if the evolution function, f(z), is given by the expression

$$f(z) = \frac{1}{(1+z)^a}, \tag{8-7}$$

then Equation (8-6) becomes

$$S_o(E_o) = \frac{Q_o(E_o) \cdot H_o^2 \cdot (1+z)^{(a-b+1)}}{4\pi c^2 z^2 g^2(z,q_o)\Theta}. \tag{8-8}$$

This equation may be rewritten as

$$Q_o[E_o] = \frac{4\pi c^2 S_o(E_o)}{H_o^2} \frac{z^2}{(1+z)^{(a-b+1)}} \Theta g^2(z,q_0). \tag{8-9}$$

The study of the evolution of quasars is complicated by their variability, which should not affect the conclusions, as long as the matter is treated in a systematic fashion. Using the maximum or the best available average value seems to lead to similar results. There is also the complication that z values do not exist for all of the radio and gamma-ray detections. A reasonably unbiased way to handle this matter mathematically seems to be to consider only blazars that are brighter than some level, i.e., that have a visual magnitude less than a given magnitude, and then exclude the relatively few that do not have measured z values. Finally, in the case of the gamma-ray blazars, different exposures and the varying level of the diffuse radiation cause the lower limit for the flux to vary with position in the sky. Chiang et al. (1995) treated this problem by assigning a formal statistical significance to each detection. These authors then applied the standard V/V_{max} test. In this test, the minimum volume that contains the source is calculated based on the value of d(z), the proper distance to the source. V_{max} is the largest volume that can contain an object with the same luminosity and still be detected for the given observation. These authors obtain a value of V/V_{max} for the gamma-ray blazar set of 0.70 ± 0.06; a value of 0.5 corresponds to no evolution.

Assuming an evolution function of the form $f(z) = (1+z)^a$, a value of "a" in the range of 2.4 to 3.2 is obtained for the gamma-ray-emitting blazars. This value compares to 3.2 to 3.7 for the optical range, 2.6 to 2.8 for the x-ray range, and 3.5 to 3.8 for the radio range. Clearly, these values are all similar. Whether they are compatible depends on whether there are remaining systematic errors yet to be determined and eliminated; the exact calculation is not easy for several reasons, including, but not limited to, those listed earlier.

If one takes the value of "a" to be 3.0, and a typical value of "b" is 2.1, then Equation (8-9) becomes

$$Q_o[E_o] = \frac{4\pi c^2 S_o(E_o)}{H_o^2} \frac{z^2}{(1+z)^{1.9}} \Theta g^2(z, q_o). \qquad (8\text{-}10)$$

Although the threshold for detection varies somewhat with exposure, the region of the sky, and the slope, consider for a moment a given threshold for detecion, i.e., $S_o(E_o)_{TH}$, and consider the value of $Q_o[E_o]$ that can be seen. For small z, $Q_o(E_o)_{TH}$ is approximately proportional to z^2; however, for large z, the dependence on z is seen to be very weak. This dependence has the following interesting consequence. For a given detection threshold, if Q_o is relatively weak, one cannot "see" very far in z, and the limiting z does not change very rapidly with Q_o. However, as Q_o increases, a region is reached for which the z, to which objects may be detected, increases relatively rapidly with Q_o. Figure 8.12 illustrates this discussion and also showes the $Q_o(E_o)/\Theta$ values for high-energy gamma-ray-emitting blazars with known z values. H_o was chosen to be 70 and q_o to be 0.5.

Another question of interest is whether there is any spectral evolution. As described earlier in this chapter, the high-energy gamma-ray emitting-blazars are well represented in almost all cases by a power law in energy in the high-energy gamma-ray range. Also, the relativistic transformation of a power law in energy transforms into a power law of the same slope. Figure 8.13 shows the spectral index as a function of z for those blazars for which the z value is known. Whereas the best fit straight line shows a small positive slope, the data are also quite compatible with a flat line, i.e., no variation of the slope with z. Hence, there is no evidence for

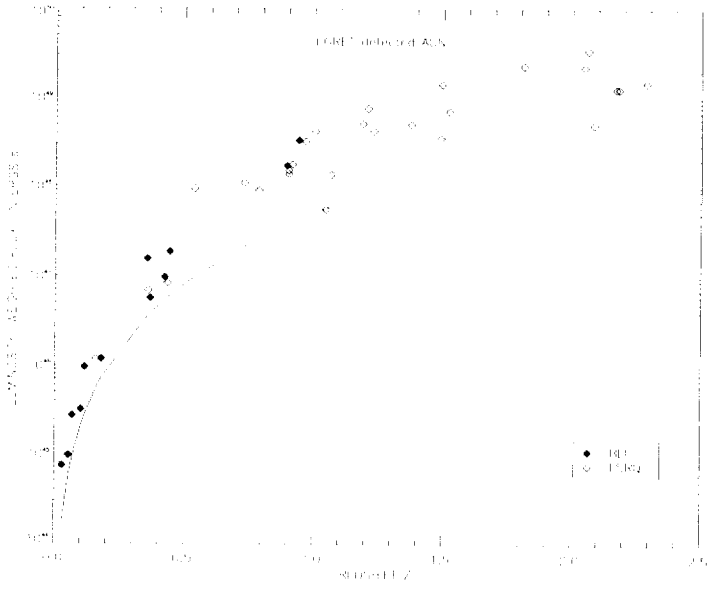

Figure 8.12. The weighted average luminosity divided by the beaming fraction in high-energy gamma rays of blazars detected by EGRET as a function of z. FSRQs are shown as open diamonds and BL Lacs as solid diamonds. The curve is the detection threshold for EGRET as a funcion of z for relative good conditions of exposure and diffuse background. The typical threshold is somewhat higher. The figure is reproduced with the permission of Mukherjee et al. (1997).

evolution of the value of the slope with z. The average value of the spectral slope is 2.12 ± 0.04, and, with the exception of 3C 279, considering the uncertainties of the individual measurements, the distribution in slope is consistent with there being the same (Mukherjee et al., 1997). Small differences are, of course, also possible within the measurement errors.

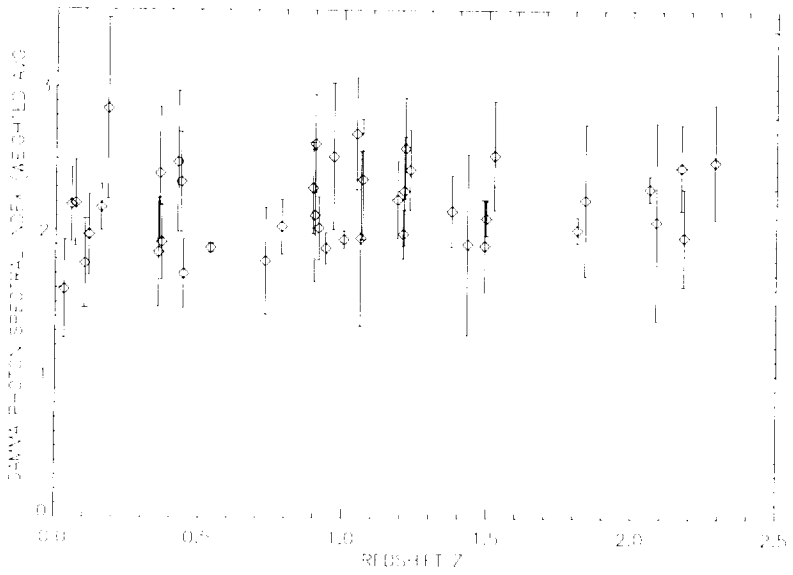

Figure 8.13. The spectral index of individual high-energy gamma-ray blazars detected by EGRET plotted as a function of z. This figure is developed from the work of Fichtel et al.(1996) and Mukherjee et al. (1997).

To continue the spectral slope discussion a bit further, Mukherjee et al. (1997) also showed that the spectral-slope distribution was consistent with there all being the same within uncertainties, except for 3C 279. Hence, the distribution of slopes may be rather tightly grouped; only far better measurements of the slopes can determine what the spread is.

8.5 Two Types of Blazars

Blazars consist of two groups of objects, the BL Lacs and the flat-spectrum radio-loud quasars (FSRQs). They are placed together and called blazars because of their similar continuum properties and their association with jets that are aimed almost directly at the observer, so that the relativistic enhancement effect is especially strong. There are, however, known differences. The FSRQs are generally more distant, more luminous, and have stronger emission lines. One of the principal distinguishing features of the BL Lacs is, in fact, the complete, or nearly complete, lack of emission lines.

Table 8.3 lists the BL Lac objects that have been observed in high-energy gamma-ray emission through March 31, 1995. Only one of these does not have a known value of z. This table is restricted to those which are believed to be statistically significant, using the same criteria defined in the high-energy gamma-ray catalogs referenced below in Table 8.3, note 1. All of the sources have been reanalyzed using the results of the energy-dependent renormalization (Sreekumar et al., 1996). Note that two objects not included in NED are included here; they are included in the four other summaries referenced in the table, including the Hewitt and Burbidge summary (1993) that probably has the most stringent criteria. However, it is appreciated that the process of determining whether an object is a BL Lac is not yet exact.

Although the sample of gamma-ray-emitting blazars is relatively small compared to the radio one, it is possible to look at the distribution of the observed gamma-ray-emitting BL Lacs as a function of z and compare it to that of the radio-loud BL Lacs. This has been done and is shown in Figure 8.14. A similar figure for FSRQs and those that are observed in high-energy gamma rays is shown in Figure 8.15. For the information on the FSRQs and BL Lacs the work of Mukherjee et al. (1997) and the references therein were used. It is obvious from these figures that the distributions of the BL Lac objects emitting high-energy gamma rays is similar to the radio one and quite dissimilar from the FSRQ one. Similarly, the gamma-ray-emitting FSRQ distribution is similar, within the limited statistics, to the radio one, as noted proviously by von Montigny et al. (1995a). A corollary of this result, of course, is that the observed high-energy gamma-ray BL Lacs are less luminous on the average than the high-energy gamma-ray-emitting FSRQs.

Figure 8.16 is the same as Figure 8.12, except that an evolution function of the form $(1+z)^3$ has been added in the form of a series of dashed lines. Note that above the dashed line crossing $z = 0$ at $Q_o[E_o]/Q = 10^{47}$ ergs s^{-1} , there are 27 FSRQs and only 2 BL Lacs emitting high-energy gamma rays. Below this same dashed line, there are only 4 FSRQs emitting high-energy gamma rays, but 10 BL Lacs. This difference is quite significant, despite the relatively small numbers. It is also obvious that the observed high-energy gamma-ray BL Lacs have a much smaller average z value, as noted earlier in this chapter. Similar differences have already been noted frequently in the radio region for FSRQs and BL Lacs. At least two

Table 8.3.

EGRET-Detected BL Lacs in Phases I to IV

Source	Other name	(5)	(6)	(7)	(8)	(9)	Notes
0219+428	3C66A	Y	Y	Y	Y	Y	(1)
0235+164	AO	Y	Y	Y	Y	Y	(1)
0521-365	PKS	Y	Y	Y	Y	No	(1)
0537-441	PKS	No	Y	Y	Y	Y	(1)
0716+714	S5	Y	Y	Y	Y	Y	(1)(2)
0735+178	PKS	Y	Y	Y	Y	Y	(1)
0829+046	OJ 049	Y	Y	Y	Y	Y	(1)
0954+658	S4	Y	Y	No	Y	Y	(1)
1101+384	Mrk 421	Y	Y	Y	Y	No	(1)
1219+285	W Comae	Y	Y	Y	Y	Y	(1)
1604+159	MC 3	Y	Y	No	Y	Y	(1)
2032+107	MC 3	Y	Y	Y	Y	Y	(1)
2155-304	PKS	Y	Y	Y	Y	Y	(3)
2200+420	BL Lac	Y	Y	Y	Y	Y	(4)

Note 1: These sources are listed in the first catalog (Fichtel et al., 1994), the second catalog (Thompson et al., 1995) or the second catalog supplement (Thompson et al., 1996). References to more detailed papers are given in these catalogs.

Note 2: No z is known.

Note 3: Verstrand, Stacy, and Sreekumar (1996).

Note 4: Cantanese et al. (1996).

Note 5: Hewitt and Burbidge (1993).

Note 6: Ciliegi, Bassani and Caroli (1995).

Note 7: Véron-Cetty and Véron (1993).

Note 8: Padovani and Giommi (1995a,b).

Note 9: Information from NASA/IPAC Extragalactic Database (NED).

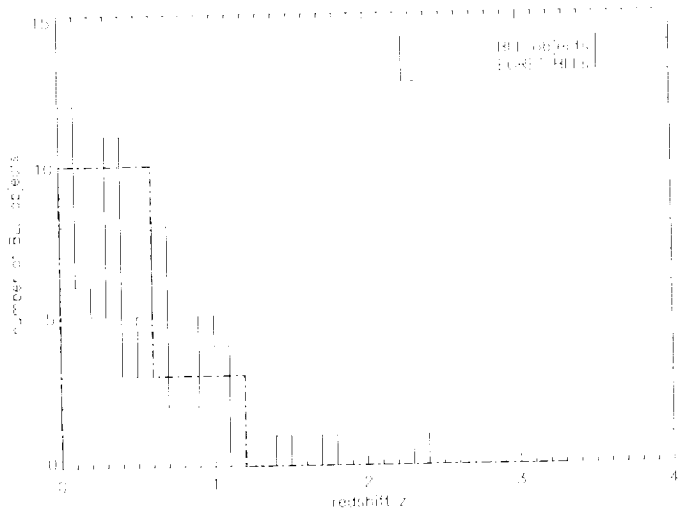

Figure 8.14. The distribution of BL Lacs as a function of z detected in the radio-frequency range, shown as a solid line, and in the high-energy gamma-ray range, shown as a dot-dash line. This figure is reproduced with the permission of Mukherjee et al. (1997).

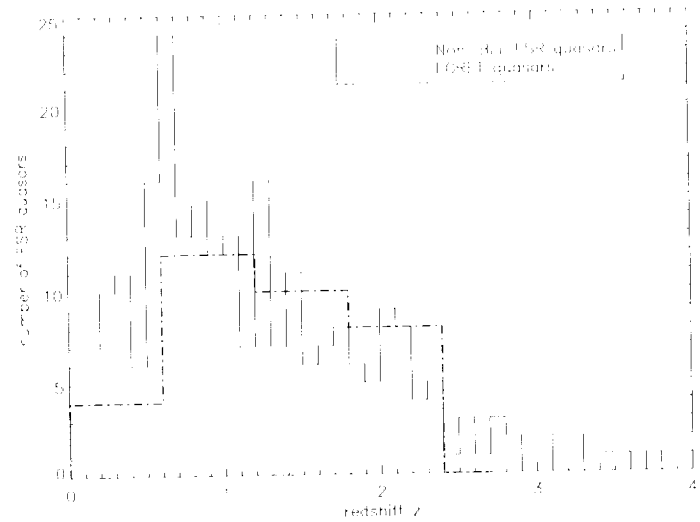

Figure 8.15. The distribution of FSRQ as a function of z detected in the radio-frequency range, shown as a solid line, and in the high-energy gamma-ray range, shown as a dot-dash line. This figure is reproduced with the permission of Mukherjee et al. (1997).

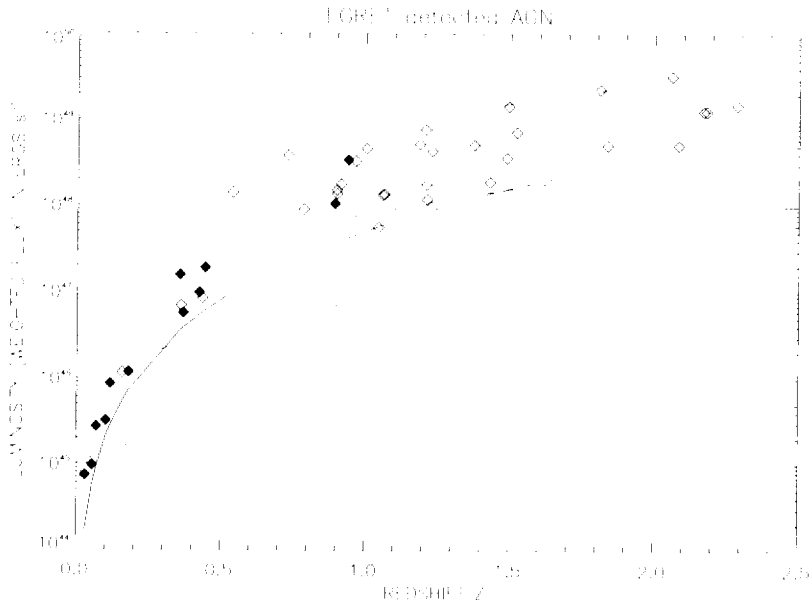

Figure 8.16 The weighted average luminosity in high-energy gamma rays of AGN detected by EGRET as a function of z. FSRQs are shown as open diamonds and BL Lacs as solid diamonds. The solid curve is the detection threshold for EGRET as a function of z for relatively good conditions of exposure and diffuse background. The typical threshold is somewhat higher. The dashed curves are for an evolution function of the form $(1+z)^3$. The figure is from Fichtel et al. (1996).

explanations have been suggested. One is that the difference between the z distribution of FSRQs and BL Lacs is the result of the luminosity values of the BL Lacs being smaller. An alternate possibility is that the BL Lacs are at a later evolutionary state.

The spectra of the two types of high-energy gamma-ray-emitting blazars will now be examined. The average spectral slope of the FSRQs in the high-energy gamma-ray range is -2.20 ± 0.05, and there is no statistically significant variation with z. The average slope of the BL Lacs is -2.03 ± 0.09, and again there is no significant change with z (See Mukherjee et al., 1997). These two average spectral slopes are not significantly different in view of the uncertainties.

For those FSRQs for which adequate observations are available, all seem to be varying in time in high-energy gamma-ray emission or are consistent with their being time variable (von Montigny et al., 1995a; Mukherjee et al., 1997). For BL Lacs, the fluxes are often near the detection threshold, and many have not been examined as often. Hence, the evidence for frequent, significant time variation in high-energy gamma rays appears less strong; however, the time variation may be as strong and just not seen.

It is interesting to consider further the question of what is the relationship between the two types of blazars. Although a number of suggestions have been made, based in large part on the radio data, the discussion here will be limited to two. One is that they are different manifestations of the same phenomenon. The second is that FSRQs evolve into BL Lacs. Regarding the first concept, it has been proposed that the BL Lacs are lower in power and possibly evolve over a longer period of time. It is true that, on the average, the observed luminosity in the radio range is lower for the BL Lacs than the FSRQs. It often has been discussed that BL Lac objects are characterized by blazing at small viewing angles, i.e., the observer is closer to the central axis of the beamed emission than in the case of FSRQs, and the difference in emission may be associated with this feature. The second point of view is that, as the FSRQs evolve, they become weaker, as the experimental evidence suggests, and that, possibly, the beaming increases with the objects evolving into BL Lacs. For a much more complete discussion of this subject, see Urry and Padovani (1995). The high-energy gamma-ray results just discussed certainly do not resolve this issue, but they do emphasize that there are differences between the two types of blazars and reinforce the type of differences already seen in the radio range.

8.6 The Origin of the Gamma-Ray Emission From Blazars and the Requirements for the Parent Relativistic Particles

The tremendous energy supply in blazars is generally believed to originate with the accretion onto a massive black hole (See, e.g., Rees 1984; Begelman, Blandford, and Rees, 1984.). For some of the reasons already mentioned in this chapter and related theoretical considerations, most of the models that appear to be possible explanations for high-energy gamma-ray emission involve beamed emission from a jet of highly relativistic particles, particularly one that is beamed in the direction of the Earth. There is not only the

observed correlation of the gamma-ray emitters with observed beams at other wavelengths, but, as noted earlier in this chapter, the short time scale of days observed for changes in gamma-ray fluxes implies a small volume on the scale of an AGN. The two most natural possibilities are the region near the black hole and a region in a jet, very possibly at the core or base of the jet, but also perhaps in a shock region. The first possibility, as noted earlier, implies isotropic emission and, hence, a much larger energy supply. In addition, it is very doubtful that the gamma rays could even escape from this central region without being absorbed. The high density of the relativistic particles, as inferred from the observations described later in Section 8.7, also strongly supports the jet theory, since it demands that the particles be nearly parallel to avoid major energy losses through interactions.

There has been some recent evidence to suggest that the gamma-ray flux increases may be associated with the core, or base, of the jet, or relatively close to it. For 0528+134 (Pohl et al., 1995), 3C 279 (Wehrle et al., 1996), and 1611+343 (Piner and Kingham, 1996), new superluminal components were observed to be moving outward after a major high-energy gamma-ray increase was observed. These results suggest that these increases are associated with the production of new "knots" or "blobs" that subsequently flow down the jet. Should further observations confirm these preliminary findings to be a common characteristic, the base of the jet would seem to be the most likely location for the acceleration of the relativistic particles responsible for the high-energy gamma radiation, at least for the major high-energy gamma-ray events.

The theories related to the specific mechanisms responsible for collimating and accelerating the jets are still in their formative stages. There is even a major uncertainty regarding whether the the jets are primarily composed of an electron, positron plasma or an electron-proton one (e.g., Celotti and Fabian, 1993). There is even a third possibility that the beam is composed of secondary electrons and protons resulting from decays of ultrarelativistic neutrons emitted by the central engine. The possible particle accelaration mechanisms will be considered in Section 8.7.

If it is accepted that the relativistic particles are present in an adequate flux, there are a number of ways that the gamma rays may be produced. These were mentioned in general terms in the introduction to this chapter. For relativistic electrons, the high-

energy gamma rays may arise from synchrotron radiation in regions of high-magnetic fields. They may come from inverse Compton scattering, either from self-generated synchrotron photons (the synchrotron, self-Compton models) or photons that exist in the region. If there are protons present, there may be nuclear interactions or photomeson production and the subsequent cascade of secondary particles. Hence, the theorists have a large number of choices, especially when several variations are included. Some of the models will be considered, and a more complete discussion may be found in von Montigny et al. (1995a), Marscher (1995), and the references therein.

In perhaps the simplest model conceptually, the gamma rays are produced directly by the synchrotron process (e.g. Ghisellini, 1993). The two principal concerns are that the magnetic-energy density must be large compared to the radiation-energy density, and very large particle energies are required.

The more common models recently are those involving inverse Compton scattering of electrons in jets with low-energy photons producing the gamma rays and a lower-energy electron. There are a number of different proposals for the low-energy photons. They may be low-energy photons in the disk (e.g., Melia and Konigl, 1989; Dermer, Schlickeiser, and Mastichiadis, 1992). Alternately, they may be self generated by the high-energy electrons as synchrotron radiation, the synchrotron, self-Compton models (e.g., Jones, O'Dell, and Stein, 1974; Marscher, 1980). More recently, it has been suggested that the photons are UV or IR thermal radiation from an accretion disk which has been scattered or reprocessed in a halo (Blandford, 1993; Sikora, Begelman, and Rees, 1994). All of these models require relatively high-energy electrons to produce the highest-energy gamma rays that are observed. Marscher (1995) has noted that the ratio of the gamma-ray to x-ray luminosity is higher than the ratio of x-ray to IR luminosity, contrary to the predictions of models in which the gamma-ray emission is second-order self-Compton scattering.

Lovelace and Romanova (1995) have developed a model based on the acceleration and creation of leptons at a propagating discontinuity or front of a Poynting flux jet. The front corresponds to a discrete relativistic jet component as observed by the VLBI. They take into account synchrotoron, synchrotron self-Compton (SSC), and inverse Compton processes, as well as photon-photon pair

production. The apparent synchrotron, SSC, and inverse Compton luminosities as functions of time are determined and seem reasonable, when compared to observations.

In addition to the high-energy electron theories, there are the theories based on a proton-initiated cascade (e.g., Mannheim and Biermann, 1989, 1992, and references therein). There are good reasons for believing that protons are as likely to be accelerated as electrons, and models based on this concept appear to predict rather naturally gamma-ray emission well into the TeV range.

8.7 Particle Acceleration

There appear to be a number of satisfactory ways to create the gamma rays if an adequate number of relativistic particles of the approprate energy distribution exists. The greater theoretical questions indeed appear to be "How do you accelerate the relativistic charged particles in such great numbers and to such high energies in the first place?", and "How is such a high particle density created and maintained?"

Consider first the energy involved. For a typical gamma-ray flux and distance, the energy being emitted by a blazar is determined to be of the order of 10^{48} ergs s^{-1} multiplied by the beaming factor. The energy emitted per unit time may be more or less than this number for a given blazar, but it is typical of those observed, which presumably are the more intense ones. The beaming factor is usually thought to be about one part in a thousand; so the energy emission is 10^{45} ergs s^{-1}. This lever is typically seen for a week or two; so the total energy emitted in that time is 10^{51} ergs. The total energy in relativistic particles, whose energies must exceed that of the gamma rays, is even higher, possibly significantly higher.

There is also the matter of the relativistic particle density. Time variations of the order of about a day have been seen, even in the strongest fluxes seen. This implies that the volume dimensions are of the order of 10^5 seconds times the speed of light, or 3×10^{15} cm times a relativistic factor in one dimension. Assuming that relativistic factor to be of the order of 5 to 10, the volume would be a few times 10^{47} cm^3. Hence, the energy density of particles to provide 10^{51} ergs in high-energy gamma rays exceeds 10^3 ergs cm^{-3}. If the average particle energy is about 1 GeV, the relativistic

particle density is 10^6 particles cm^{-3}. This is a very difficult relativistic particle density to accelerate and maintain without very large losses. It seems to demand a nearly unidirectional velocity of the relativistic particles, and, thus, it is one more factor supporting the concept that there is a beam.

Assuming that there is a sufficient energy supply, ultimately provided by matter falling into the very deep gravitational potential well surrounding a supermassive black hole, the critical question is, how is an adequate fraction converted to relativistic and even extremely relativistic particle energy? It seems that a successful answer to this question must involve the natural evolution of a large, compact, fairly uniform, magnetic field. Two possibilities for the formation of such a magnetic system in association with a supermassive black hole are described below.

First, consider an analogy to a familiar model, the case of isolated pulsars; it seems clear that the dominant role in producing acceleration of charged particles is played by a rotating magnetic field configuration. Extending this conclusion to AGN is not trivial. The formation of a uniform magnetic field is not an immediate consequence of the existence of a supermassive black hole, as it is in the case of a neutron star. Hence, if this approach is to be successful, a reasonable model for the natural evolution of a large, compact, and reasonably uniform magnetic field must be found.

Blandford and Znajek (1977) have proposed a conducting accretion disk around a massive rotating black hole that serves to confine a magnetic field that threads the black hole's event horizon. The horizon acts as a rotating conductor, immersed in an external magnetic field, i.e., a unipolar generator that can deliver power to a magnetospheric load. Apparently, if energy can be extracted in this way, then relativistic beams emerge along the rotation axis of the hole in a natural way. This power-generation process is very similar to a process originally proposed for neutron stars by Goldreich and Julian (1969). A particularly illuminating analysis of the Blandford-Znajek process is given by MacDonald and Thorne (1982).

The key question is the manner in which the intense uniform magnetic field threading the black hole becomes established in the first place. To our knowledge, no one has yet succeeded in proving with mathematical rigor that such a configuration is an inevitable result of the physical situation. However, MacDonald et al. (1986)

have provided qualitative support, based on the natural evolution of the magnetic field near the black hole from a combination of Maxwell pressure and Maxwell tension. They further argue that no matter how chaotic the field threading the disk may be, the field through the hole will be well ordered.

The potential drop available for particle acceleration is approximately

$$\Delta V=(10^{20}V)(a/M_{BH})(B/10^4G)(M_{BH}/10^9M_{\odot}), \qquad (8\text{-}11)$$

where B is the magnetic field threading the horizon, M_{BH} is the mass of the black hole, and "a" is the angular momentum per unit mass of the black hole. For reasonable parameters, as much as 10^{20}–10^{21} V is available. If the charged particle existed in a vacuum, then it could be accelerated up to a maximum energy of $e\Delta V$. But many effects can limit the energy to much lower values. These include synchrotron-radiation losses and Compton cooling. In a plasma, collisions with other particles can limit the maximum attainable energy. Even if the plasma is tenuous and collisions are unimportant, a variety of plasma microinstabilities also may limit the energy. Various stochastic reacceleration mechanisms, such as Fermi shock acceleration, may act to overcome some of these losses.

The total power in the relativistic beams is of the order of

$$L_b=(10^{45}\text{erg s}^{-1})(a/M_{BH})^2x(B/10^4G)^2(M_{BH}/10^9M_{\odot})^2. \qquad (8\text{-}12)$$

For reasonable astrophysical parameters, the beam luminosity can approach 10^{45}–10^{47} erg s^{-1}, consistent with observations.

Lovelace (1976) proposed a different model in which magnetic fields, generated in a differentially rotating disk, can lead to acceleration of charged particles to ultrarelativistic energies and extraction of energy from the disk. The model apparently can produce either a proton beam or an electron-positron beam, depending on whether B is parallel or antiparallel to the angular momentum vector of the disk. The potential drop available for particle acceleration and the beam luminosity in this case are comparable to the above estimates for the Blandford-Znajek effect.

Less powerful, slower jet-acceleration models have also been proposed. They involve acceleration either by hydromagnetic means or by radiation pressure that fills a narrow funnel along the axis of a thick accretion disk (Lynden-Bell, 1978; Paczynski and Wiita, 1980; Camenzind, 1986, 1987). It appears that such models can produce only mildly relativistic beams and that further shock acceleration is then required to produce the energetic particles with adequate energy for the generation of the highest-energy gamma rays.

Both the difficulties of accelerating charged particles near the center of the AGN and the high gamma-ray opacities that are believed to be present are circumvented in a model wherein high-energy neutrons are produced in the central engine (Eichler and Wiita, 1978; Giovanoni and Kazanas, 1990; Mastichiadis and Protheroe, 1990). In these models, the Lorentz factors of the neutrons are high enough that they travel sufficiently far outward before decaying into a proton, electron, and antineutrino. The gamma rays then follow as in the other models from the interactions of the extremely relativistic charged particles.

Still another approach to the accelaration problem has been studied by Dermer, Miller, and Li (1996). They consider the stochastic acceleration of particles that result from resonant interactions with plasma waves in the magnetosphere surrounding an accreting black hole. It is not yet clear which of the several theories just described is closest to the true explanation.

8.8 Summary

With the results from the Compton Gamma-Ray Observatory (CGRO), there has been a truly remarkable advance in our knowledge of the gamma-ray emission from active galaxies, especially in the high-energy gamma-ray emission from blazars. There are two distinctly different classes of objects involved. For the Seyfert and radio galaxies, the gamma-ray emission is in the hard x-ray to low-energy gamma-ray range As noted, the spectrum seems to be consistent with a model including an initial spectrum that is "Comptonized" in a relativistic, optically thin, thermal corona above the surface of an accretion disk and including absorption.

The number of known high-energy gamma-ray emitting blazars has increased dramatically since the launch of CGRO, from 1 to over 50. They have hard spectra that are power laws in photon energy

with an average spectral index of -2.16 ± 0.04. The distributions in z of FSRQs and BL Lacs that emit high-energy gamma rays appear to be similar to those for these objects that are seen in the radio frequency range. Many of the FSRQ blazars that have been observed in gamma rays for which adequate information exists are observed to vary markedly in time, with time scales as short as a day or two in some cases. Based on the correlation with properties measured at other wavelengths and the deduced density of the relativistic particles, the emission seems almost certainly to be associated with a jet directed at least approximately towards the Earth. Recent results suggest that at least the major increases in emission come from relatively near the base, or core, of the jet. The energy released in the form of high-energy gamma rays is extraordinary, typically of the order of 10^{45} ergs s^{-1} or more in the high state. Hence, by implication, the energy in very relativistic particles is at least as great and probably much greater. Since there is now a good knowledge of the basic properties of the high-energy gamma-ray emission and its implication for the parent relativistic particles, a challenge exists to formulate a quantitative theoretical explanation of exactly how these particles are accelerated.

References

Baade, W., and Minkowshi, R., 1954, ApJ, 119, 215.

Baity, W. A., et al., 1981, ApJ, 244, 429.

Baity, W. A., Mushotzky, R. F., Worrall, D. M., Rothschild, R. E., Tennant, A. F., and Primini, F. A., 1984, ApJ, 279, 555.

Begelman, M. C., Blandford, R. D., and Rees, M. J., 1984, RMP, 56, 255.

Bertsch, D. L., et al., 1993, ApJ, 405, L21.

Bignami, G. F., Fichtel, C. E., Hartman, R. C., and Thompson, D. J., 1979, ApJ, 232, 649.

Blandford, R. D., and Znajek, R., 1977, MNRAS, 179, 433.

Blandford, R. D., 1993, in Proc. Compton Symp. 1992, ed. M. Friedlander, N. Gehrels, and D. J. Macomb (New York: AIP), 533.

Camenzind, M., 1986, A&A, 162, 32.

Camenzind, M., 1987, A&A, 184, 341.

Cantanese et al., 1996, to be published.

Celotti, A., and Fabian, A. C., 1993, MNRAS, 264, 228.

Chiang, J., Fichtel, C. E., von Montigny, C., Nolan, P. L., and Petrosian, V., 1995, ApJ 452, 156.

Ciliegi, P., Bassani, L., and Caroli, E., 1995, ApJ, 439, 80.

Dermer, C. D., Miller, J. A., and Li, H., 1996, ApJ, 456, 106.

Dermer, C. D., Schlickeiser, R., and Mastichiadis, A., 1992, A&A, 256, L27.

Dermer, C. D., and Gehrels, N., 1995, ApJ, 447, 103.
Dingus, B., et al., 1996, ApJ, 467, 589.
Dwek, E., and Slavin, J., 1994, ApJS, 94, 551.
Eichler, D., and Wiita, P. J., 1978, Nature, 274, 38.
Esposito, J., et al., 1996, submitted to ApJS.
Fichtel, C. E., 1994, ApJS, 90, 917.
Fichtel, C. E., et al., 1994, ApJS, 94, 551.
Fichtel, C. E., et al. 1996, 188th AAS Meeting, Display Paper 5.09.
Gehrels, N., Cline, T. L., Teegarden, B. J., Paciesas, W. S., Tueller, J., Durouchoux, P., and Hameury, J., 1984, ApJ, 278, 112.
Ghisellini, G., 1993, Adv. Space Res., Vol. 13, No. 12, 587.
Giovanoni, P. M., and Kazanas, D., 1990, Nature, 345, 319.
Goldreich, P., and Julian, W. H., 1969, ApJ, 157, 869.
Griffiths, R. E., et al., 1989, MNRAS, 240, 33.
Grindlay, J. E., 1975, ApJ, 199, 49.
Hall, R. D., Meegan, C. A., Walraven, G. D., Djuth, F. T., and Haymes, R. C., 1976, ApJ, 210, 631.
Hartman, R. C., et al., 1992, ApJ, 385, L1.
Hartman, R. C., et al., 1994, private comunication.
Hartman, R. C., et al., 1996, ApJ, 461, 698.
Herterich, K., 1974, Nature, 250, 311.
Hewitt, A., and Burbidge, G., 1993, ApJS, 87, 451.
Jannuzi, B. T., Smith, P. S., and Elston, R., 1994, ApJ, 428, 130.
Jelley, T. V., 1966, Nature, 211, 472.
Johnson, W. N., et al., 1993, ApJS, 86, 693.
Johnson, W. N., et al., 1994, in The Second Compton Symposium, ed. C. E. Fichtel, N. Gehrels, J. P. Norris, (New York: AIP), 515.
Johnson, W. N., Dermer, C. D., Kinzer, R. L., Kurfess, J. D., Strickman, M. S., 1995, ApJ, 445, 182.
Jones, T. W., O'Dell, S. L., and Stein, W. A., 1974, ApJ, 188, 353.
Jourdain, E., et al., 1993, ApJ, 412, 586.
Kafatos, M., and Leiter, D., 1979, ApJ, 229, 46.
Kanbach, G., et al., 1988, Space Sci. Rev., 49, 69.
Kerrick, A. D., et al., 1995, ApJ, 252, 588.
Kinzer, R. L., et al., 1995, ApJ, 449, 105.
Kniffen, D. A., et al., 1993, ApJ, 411, 133.
Kurfess, J. D., et al., 1995, Advances in Space Research, Vol. 15, No. 5, (5) 103.
Kurfess, J. D., et al., 1996, private communication.
Leiter, D., and Kafatos, M., 1978, ApJ, 226, 32.
Lichti, G. G., et al., 1995, A&A, 298, 711.
Lin, Y. C., et al., 1993, ApJ, 416, L53.
Lin, Y. C., et al., 1996, ApJ, 105, 331.
Lovelace, R. V. E., 1976, Nature, 262, 649.
Lovelace, R. V. E., and Romanova, M. N., 1995, in Proc. of Cygnus A Workshop, ed. Carilli and Harris (Camb. U. Press).
Lynden-Bell, D., 1978, Phys. Scripta, 17, 185.

MacDonald, D., and Thorne, K. P. S., 1982, MNRAS, 198, 345.
MacDonald, D., Thorne, K. P. S., Price, R. H., and Zhang, X-H, 1986, in Black Holes: The Membrane Paradigm, ed. K. P. S. Thorne, R. H. Price, and D. A. MacDonald, (New Haven: Yale Univ. Press), 121.
Madejski, G. M., et al., 1995, ApJ, 438, 672.
Maisack, M., et al., 1992, ApJ, 407, L167.
Mannheim, K., and Biermann, P. L., 1989, A&A, 221, 211.
Mannheim, K., and Biermann, P. L., 1992, A&A, 53, L21.
Maraschi, L., and Treves, A., 1977, ApJ, 218, L113.
Maraschi, L., et al., 1994, ApJ., 298, 711.
Marscher, A. P., 1980, ApJ, 235, 386.
Marscher, A. P., 1995, Proc. of the National Academy of Sciences, 92, 11439.
Mastichiadis, A., and Protheroe, R. J., 1990, MNRAS, 246, 279.
McNaron-Brown, K., et al., 1995, ApJ, 451, 575.
Melia, F., and Konigl, A., 1989, ApJ, 340, 162.
Michelson, P. F., et al., 1992, IAU Circ., 5470.
Montigny, C. von, et al., 1993, A&A, 97, 101.
Montigny, C. von, et al., 1995a, ApJ, 440, 525.
Montigny, C. von, et al., 1995b, A&A, 299, 680.
Morris, S. L., Stocke, J. T., Gioia, I. M., Schild, R. E., Wolter, A., Maccacaro, T., and Deila Deca, R., 1991, ApJ, 380, 49.
Mücke, A., et al., 1996, to be published in A&A.
Mukherjee, R., et al., 1997, submitted to ApJ.
Mushotzky, R. F., 1976, Ph.D. Thesis, University of California, San Diego.
Mushotzky, R. F., Serlemitsos, P., Becker, R., Boldt, E., and Holt, S., 1978, ApJ, 220, 790.
Nolan, P., et al., 1996, ApJ, 459, 100.
Paczynski, B., and Wiita, P. J., 1980, A&A, 88, 23.
Padovani, P., 1992, A&A, 256, 399.
Padovani, P., and Giommi, P., 1995a, ApJ, 444, 567.
Padovani, P., and Giommi, P., 1995b, MNRAS, 277, 1477.
Perlman, E. S., and Stocke, J. T., 1993, ApJ, 406, 430.
Piner, B. G., and Kingman, K. A., 1996, ApJ, submitted.
Pohl, M., et al., 1995, A&A, 303, 383.
Pollock, A., Bignami, G., Hermsen, W., Kanbach, G., Lichti, G., Mansou, J., Swanenberg, B., and Wills, R., 1981, A&A, 94, 116.
Punch, M., et al., 1992, Nature, 358, 477.
Quinn, J., et al., 1995, Proc. of the 24th International Cosmic Ray Conf., 2, 366.
Rees, M. J., 1978, Physica Scripta, 17, 193.
Rees, M. J., 1984, ARAA, 22, 471.
Rothschild, R. E., Mushotzky, R. F., Baity, W. A., Gruber, D. E., Matteson, J. L., & Peterson, L. E., 1983, ApJ, 269, 423.
Schrader, C. R., et al., 1996, in Proc. The Third Compton Symposium, June 12-14 1995, München, to be published.
Shapiro, S. L., and Salpeter, E. E., 1975, ApJ, 198, 671.
Sikora, M., Begelman, M. C., and Rees, M. J., 1994, ApJ, 421, 153.

Sreekumar, P., et al., 1996, ApJ, 464, 628.
Stecker, F. W., de Jager, O. C., and Salamon, M. H., 1993, ApJ, 415, L71.
Stecker, F. W., and De Jager, O. C., 1996, Space and Science Reviews 75, 401.
Steinle, H., et al., 1995, Advances in Space Research, Vol. 15, No. 5, (5) 37.
Stickel, M., Padovani, P., Urry, C. M., Fried, J. W., and Kuehr, H., 1991, ApJ, 374, 431.
Swanenburg, B. N., et al., 1978, Nature, 275, 298.
Takahashi, T., et al., 1993, in Compton Gamma-Ray Observatory, ed. M. Friedlander, N. Gehrels, D. J. Macomb (AIP: New York), 523.
Thompson, D. J., et al., 1993, ApJS, 86, 629.
Thompson, D. J., et al., 1995, ApJS, 101, 259.
Thompson, D. J., et al., 1996, ApJS, 107, in press.
Ubertini, P. Bazzano, A., LaPadula, C., Polcar, V. F., and Manchanda, R. K., 1984, ApJ, 284, 54.
Urry, C. M., and Padovani, P., 1994, in ASP Conf. Ser., 54, The Physics of Active Galaxies, ed. G. Bicknell, M. Dopita, and P. Quinn (San Francisco:ASP), 215.
Urry, C. M., and Padovani, P., 1995, PASP, 107, 803.
Véron-Cetty, M. P., and Véron, P., 1993, A&AS, 100, 521.
Vestrand, W. T., Stacy, J. G., and Sreekumar, P., 1996, ApJ, 454, L93.
Weekes, T. C., et al., 1996, A&AS, submitted [Third Compton Symposium].
Wehrle, A. E., et al., 1996, Proceedings of the Workshop on Blazar Variability, Florida International University, Miami, to be published.
Zdziarski, A. A., Johnson, W. N., Done, C., Smith, D., and McNaron-Brown, K., 1995, ApJ, 438, L63.

CHAPTER 9

DIFFUSE RADIATION

9.1 Introduction

A diffuse celestial radiation that is isotropic, at least on a coarse scale, has been observed from the soft x-ray region all the way through to 10 GeV in the high-energy gamma-ray region. The spectral shape, the intensity, and the established degree of isotropy of this radiation already place significant constraints on the possible explanations. For example, a Galactic halo model, or indeed any Galactic concept, seems very unlikely. The observed diffuse radiation is now generally accepted as being extragalactic in origin and will be referred to as such hereafter.

Numerous theories for the origin of this radiation have been set forth over the years. They include normal galaxies, clusters of galaxies, primordial cosmic rays interacting with intergalactic matter, primordial black holes, cosmic-ray leakage from galaxies, and matter-antimatter annihilation at the boundaries of superclusters of galaxies of matter and antimatter in baryon-symmetric big bang models. As the results have become more complete, extensive, and precise, the constraints on possible models have become more severe. At the moment, active galaxies appear to be the most likely source of this radiation, with Seyfert and radio galaxies dominating in the x-ray and low-energy gamma-ray region and blazars playing the major role in the medium- and high-energy gamma-ray region.

In the next section, the experimental observations will be summarized. This will be followed by a brief summary of some aspects of cosmology that are important in trying to interpret the results, since, at the moment, the more likely explanations seem to be cosmological in nature. In the final section, the experimental findings will be discussed in terms of possible origins, with the reasons why some models seem unlikely being given, along with

the evidence supporting the proposed origins currently thought to be more likely.

9.2 Observations

In the x-ray region, a uniform diffuse radiation has been clearly established for some time with relatively small uncertainty as a result of many measurements, including especially those of HEAO-1 and ASCA. The measurements extend for almost three decades from about 1 keV to nearly 1 MeV. In the high-energy gamma-ray region, after many attempts that led to upper limits or estimated fluxes that later proved to be too high, there was a measurement by SAS-2 giving the intensity in the energy range from 35 to 200 MeV (Thompson and Fichtel, 1982). Although the integral intensity above 100 MeV was reasonably well determined to be (1.3±0.5) x 10^{-5} cm^{-2} s^{-1} sr^{-1}, it was only possible to determine the slope of the spectrum to −2.35 (+ 0.4/ −0.3). The low-energy gamma-ray range remained quite uncertain because of the very great difficulty of subtracting a very high, locally produced, background caused by many effects. Claims of a positive detection were met with doubts about whether the background had really been fully subtracted.

With the results from the Compton Gamma-Ray Observatory, the situation improved considerably. The high-energy gamma-ray diffuse spectrum was determined all the way from 30 MeV to 40 GeV with data from EGRET (Sreekumar et al., 1997); the intensity at 100 MeV and the spectral slope were in agreement with the earlier SAS-2 result. The spectrum was was found by Sreekumar et al. to be well fit by a power law in photon energy with a slope of −2.09 ± 0.02, as shown in Figure 9.1. The intensity above 100 MeV is (1.45 ± 0.05) x 10^{-5} photons cm^{-2} s^{-1} sr^{-1}. This result is consistent with earlier estimates from the EGRET data based on smaller data sets and less detailed analysis (Kniffen et al., 1994; Osborne, Wolfendale, and Zhang, 1994; Fichtel et al., 1994).

In the energy range from just below 1 MeV to 30 MeV, analysis of the COMPTEL results (Kappadeth et al., 1995) showed no evidence of a "bump", as some early results had suggested (Trombka et al., 1977; Schönfelder, Graml and Penningsfeld, 1980). The spectral data points of COMPTEL join smoothly to those above this energy range, as seen in Figure 9.2. The background in the energy range

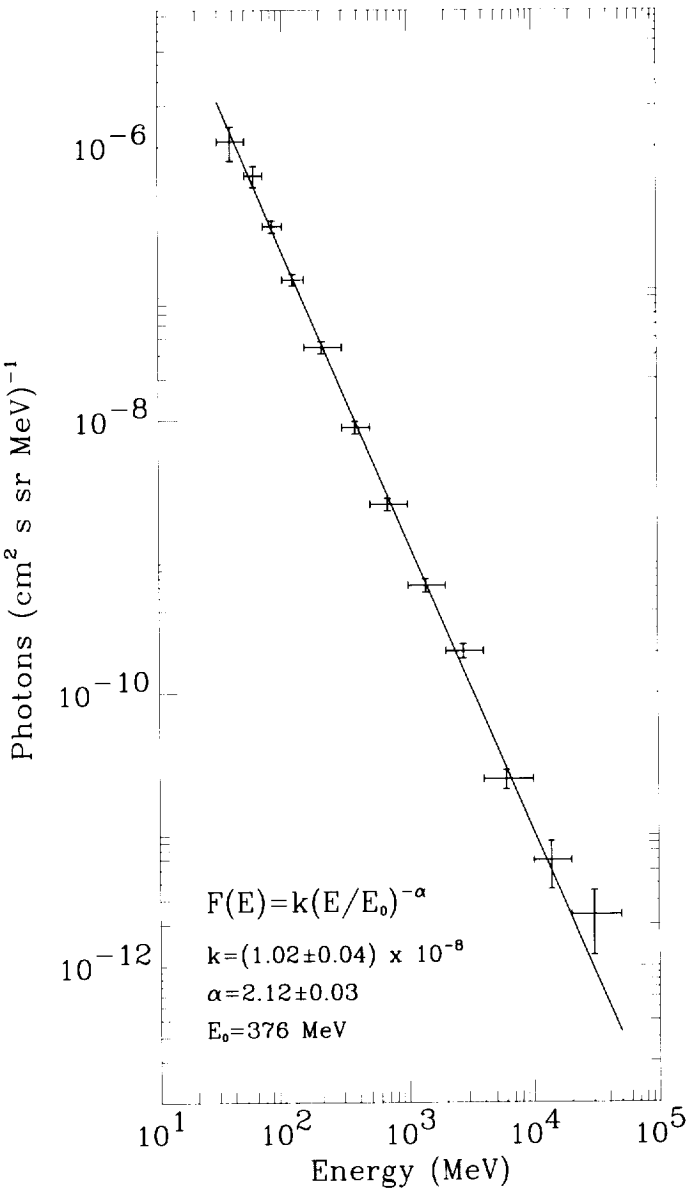

Figure 9.1. The high-energy gamma-ray spectrum of the diffuse, presumably extragalactic, radiation. The figure is reproduced with the permission of Sreekumar et al. (1997).

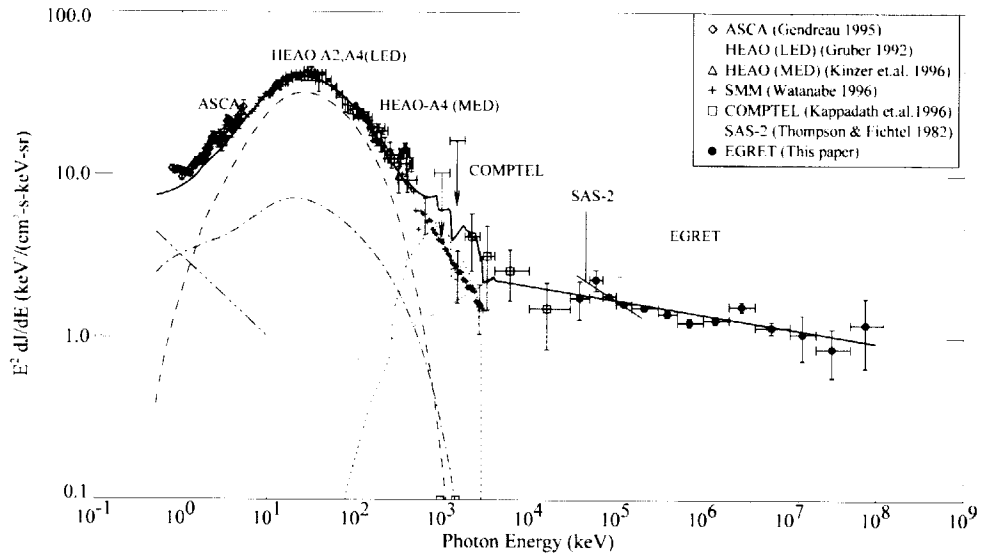

Figure 9.2. Multiwavelength diffuse, presumably extragalactic, spectrum from x-rays through high-energy gamma rays, compared to theoretical expectations. The reference to the EGRET results in the figure is Sreekumar et al. (1997). The estimated contribution from Seyfert 1 (dot-dashed), and Seyfert 2 (dashed) are from the model of Zdziarski (1996); steep-spectrum quasar contribution (dot-dot-dashed) is taken from Chen et al. (1996); Type Ia supernovae (dot) is from The, Leising and Clayton (1993). The blazar contribution below 4 MeV (thin long dashed) is from Sreekumar et al. (1997). The thick solid line indicates the sum of all the components. The figure is reproduced with the permission of Sreekumar et al. (1997).

from approximately 0.2 to 4 MeV is quite high, and there remains significant uncertainty about the corrections; but it is believed now that the intensity is no higher in this range than that shown in Figure 9.2 and likely lower.

These results have been obtained only after a great effort over many years. The difficulties in the two energy ranges are somewhat different. In the low-energy gamma-ray region, the principal problem is the very high background from a large number of orgins. The unwanted background is created both in the surrounding material and in the detector itself. Eliminating this unwanted radiation is a considerable challenge. It was possible in the case of the COMPTEL instrument for several reasons. It was a directional detector, so the direction of the gamma rays was known, at least approximately. There was substantial preflight calibration. A very detailed analysis of the data could be performed as a function of the geomagnetic cutoff rigidity as the Compton Gamma-Ray Observatory circled the Earth. It was possible to separate the effects of the short-lived and long-lived components of the background.

In the high-energy gamma-ray range, as the result of the very careful design of EGRET based on years of study and development, as described in Chapter 13 of this book, the general background was more than an order of magnitude below the level of the radiation to be measured. Hence, that was no longer a problem. However, unlike lower energies, the Galactic diffuse radiation is not negligible. In fact, as seen earlier in Chapter 5, it strongly dominates in the Galactic plane. In order to determine the uniform, presumably extragalactic component, the two must be separated. Using the model discussed in Chapter 5 (Hunter et al., 1996), the expected radiation may be compared to the observed, and, when this is accomplished, there is seen to be a residual that is reasonably uniform over the sky, at least at higher latitudes (Sreekumar et al., 1996). Very close to the plane, within about 10°, the Galactic radiation is so intense that meaningful separation is not possible.

Let us turn next to the question of the uniformity of this radiation. The x-ray emission through about 100 keV at least is known to be isotropic to within about 5% (Schwartz and Gursky, 1973). The degree of uniformity is much less well known in the gamma-ray region. The early SAS-2 results showed that in the high-energy

range, the center-to-anticenter ratio for the radiation with $20° < /b/ < 40°$ was 1.10 ± 0.19, and the ratio of the intensity perpendicular to the Galactic plane intensity to that in the $20° < /b/ < 40°$ region was measured to be 0.87 ± 0.09 (Fichtel, Simpson, and Thompson, 1978). Each of these results, of course, is consistent with isotropy to within uncertainties, and together they showed that there was not a major anisotropy on a broad scale.

In the analysis of the EGRET data Sreekumar et al. (1997) first excluded the region within 10° of the Galactic plane for all longitudes and 30° for an 80° longitude interval above the Galactic center. The remainder of the sky was divided into 36, approximately equal, solid angles. The distribution of intensities in these intervals was found to be consistent with isotropy within an uncertainty of about 20% for the fluxes in the individual solid angle intervals.

9.3 Cosmology

Before discussing the current concepts regarding the origin of the general, presumably extragalactic, diffuse radiation, a very brief summary will be given of the current concept of the universe, the equations used to describe it, and the means of interpreting the diffuse gamma radiation, which might result from phenomena that occurred in the distant past.

The present view of our universe had its beginning about half a century ago, when Hubble first concluded that there was an approximately linear relationship between the distances to galaxies and their red shifts. During the intervening half century, the data have improved remarkably, but the result is still basically the same, that is, galaxies appear to be moving apart at velocities proportional to their separations. Hence, the obvious conclusion is that, at some time in the past, estimated now to be about 15×10^9 years ago, they must all have been close together.

In 1922, Alexandre Friedman found the general homogeneous and isotropic solution to the equations that Einstein had developed in his theory of general relativity for the whole universe. This "Friedman model" forms the basis of what is now considered to be modern standard cosmology. One aspect of the model that has received

much attention is that, depending on the energy density, the universe is either open or closed. If it is open, it will continue to expand forever, and, if it is closed, its present expansion will ultimately stop and a contraction will begin. The detailed theory shows that whether it is open or closed depends on the average energy density. The critical density, ρ_c, is about 10^{-29} g cm^{-3}, or more specifically

$$\rho_c = 0.7 \times 10^{-29}\left(H_o / 60\text{km s}^{-1}\text{ Mpc}^{-1}\right)^2 \text{ g cm}^{-3}, \qquad (9\text{-}1)$$

where H_o is the present best value for the Hubble parameter. Since, at present, matter is the major contributor to the energy density, if the matter density is less than this value, the universe is open, and, if not, it is closed. Currently it appears that ρ for the observed matter is significantly less than ρ_c, but there may be substantial unseen mass.

The other major discovery related to our current view of cosmology was that of the diffuse cosmic 3K microwave radiation by Penzias and Wilson (1965). As Dicke et al. (1965) explained, this radiation is exactly what is expected in a universe that has expanded from a time when it was very hot and dense, and the scattering of free electrons and photons maintained matter and radiation in thermal equilibrium. As the universe expanded and cooled from this early state, a temperature (~3 x 10^3 K) was reached when the electrons combined with nuclei to form atoms. The photons were no longer coupled thermally, and the photon radiation thereafter simply expanded freely. The density and temperature then decreased, but continued to have a spectral shape described by the Planck black-body equation. The microwave radiation mentioned above has been found to have the spectrum appropriate for that of 3K black-body radiation, and, hence, this observation gives strong support to the theory of the expanding universe, or the "big bang".

Knowing the present temperature of the black-body radiation, it is possible to calculate the production of complex nuclei in the early stages of the universe. One result of these calculations is that about 27 percent by mass of the nucleons should have fused into He^4 (Weinberg, 1972). This prediction is in agreement with the range of values allowed by current measurements and, hence, provides additional support for the model.

There is also the question of whether the universe, as a whole, is symmetric with regard to matter and antimatter. Gamma-ray astronomy has already made an important contribution in this area by showing that the steady-state theory, wherein matter and antimatter were continually created (Hoyle, 1948; Bondi and Gold, 1948) throughout the universe, was inconsistent with the level of diffuse radiation by several orders of magnitude.

For a general review of cosmology, see Weinberg (1972) and Peebles (1993).

In order to consider the diffuse emission that is being observed, it is necessary to look as far back in time as necessary to include all the emission that might be reaching the observer at the present time. In proceeding, it is not only necessary to take into account properly spatial considerations, but also to remember that a frequency ν or energy E_o observed now had a larger value in the past. (The subscript "o" will be used in this section to refer to the time of observations). The universe being considered here is isotropic and homogeneous; all the sources, unless otherwise stated, will be assumed to emit radiation isotropically and to be distributed isotropically (at least on a coarse scale or on the average, so that such an assumption is an adequate representation of the real world).

The number density of photons per energy interval at any time t , which at time t_o has an energy between E_o and $E_o + dE_o$, is $N(E_o,t)$. The rate of change of the number of photons in a proper volume $R^3(t)$ and an energy interval dE_o at time t_o is

$$\frac{d}{dt}\left[N(E_o,t)R^3(t)\right] = \Gamma\left(E_o\frac{R(t_o)}{R(t)},t\right)R^3(t)\left(\frac{R(t_o)}{R(t)}\right),$$

$$-A\left(E_o\frac{R(t_o)}{R(t)},t\right)N(E_o,t)R^3(t), \qquad (9\text{-}2)$$

where Γ is the rate of emission per unit volume in the energy interval dE_o, and A is the absorption rate (see Weinberg, 1972).

Consider first the case for which the second term is negligable. This case is appropriate for gamma-ray astronomy for all but the lowest-

energy gamma rays and the earliest times, when the density was much higher. To obtain a rough idea that this statement is correct, consider a gamma ray traveling for 15 x 10^9 years through a medium with a density even as high as that needed to close the universe ~0.7 x 10^{-29}g cm^{-3}. Its total path length is 0.2 g cm^{-2}, which for gamma rays above a few MeV, is only about one-thousandth of an interaction mean free path, and even for gamma rays in the 0.1 to 1 MeV range, it is small. Although the density in the past was larger, the conclusion is still the same. For A = 0, Equation (9-2) becomes

$$d\left[N(E_o,t)R^3(t)\right] = \Gamma\left(E_o\frac{R(t_o)}{R(t)}\right)R^3(t)\left(\frac{R(t_o)}{R(t)}\right)dt. \qquad (9\text{-}3)$$

Using the following expression, which is valid for the standard universe in the matter dominated era (i.e., the matter energy density >> the radiation energy density), which began long before galaxies were formed,

$$dt = \frac{dR}{R} = \frac{dR(t)}{H_oR(t_o)}\left[1 - 2q_o + 2q_o\left(\frac{R(t_o)}{R(t)}\right)\right]^{-\frac{1}{2}}, \qquad (9\text{-}4)$$

where H_o is the present value of the Hubble parameter, and q_o is the deceleration parameter. Since

$$\frac{R(t_o)}{R(t)} = (z+1), \qquad (9\text{-}5),$$

Equation (9-4) may be rewritten as

$$dt = \frac{dz}{H_o(1+z)^2(1+2q_oz)^{\frac{1}{2}}}. \qquad (9\text{-}6)$$

Using Equations (9-5) and (9-6), Equation (9-3) becomes

$$d\left[N(E_o,t)R^3(t)\right] = \Gamma(E_o(1+z),z)\frac{R^3(z=o)dz}{H_o(1+z)^4(1+2q_oz)^{\frac{1}{2}}}. \qquad (9\text{-}7)$$

Hence,

$$N(E_o,t_o)=\frac{1}{H_o}\int_o^{Z\,max}\Gamma(E_o(1+z),z)\frac{dz}{(1+z)^5(1+2q_oz)^{\frac{1}{2}}}, \qquad (9\text{-}8)$$

under the assumption that at the beginning of the time of interest $N(E_o,t_o)R^3(t_i)$ is zero of negligible compared to $N(E_o,t_o)$ x $R^3(t_0)$. The differential intensity associated with this differential density is

$$j\equiv\frac{dJ}{dE}=\frac{c}{4\pi H_o}\int_o^{Z\,max}\Gamma(E_o(1+z),z)\frac{dz}{(1+z)^5(1+2q_oz)^{\frac{1}{2}}}. \qquad (9\text{-}9)$$

To proceed further, it is necessary to make a specific assumption about $(E_o(1+z),z)$. Two specific cases will be considered. First, consider the case wherein the diffuse background is caused by active galaxies. Then, Γ may be written as

$$\Gamma_{1,i}(E_o(1+z),z)=n_iQ_i(E_o(1+z)z), \qquad (9\text{-}10)$$

where n_i is the number of density of galaxies of type i, and $Q_i(z)$ is the average source strength. If it is assumed that the number of galaxies remains constant, then the number density is proportional to $(1+z)^3$. It will also be assumed that $Q_i(z)$ may be written as $Q_i(E_o(1+z))f(z)$, i.e., that intensity evolution is separable and that the spectral form does not change. Equation (9-10) then becomes

$$\Gamma_{1,i}(E_o(1+z),z)=n_i(o)(1+z)^3Q_i(E_o(1+z))f(z), \qquad (9\text{-}11)$$

and Equation (9-9) becomes

$$j_I=\frac{cn_i(o)}{4\pi H_o}\int_o^{Z\,max}\frac{Q_i(E_o(1+z))f(z)dz}{(1+z)^2(1+2q_oz)^{\frac{1}{2}}}. \qquad (9\text{-}12)$$

Another case of interest concerns any photon creating interaction between two species, such as cosmic rays and matter or cosmic-ray electrons and photons. Here, the densities of both species would be

proportional to $(1+z)^3$, and, if neither species were being created during the period of interest and their losses were negligable, then $\Gamma(E_o(1+z),z)$ would be

$$\Gamma_{II,i,j}\left(E_o(1+z),z\right) = n_i(o)n_j(o)(1+z)^6 F_{ij}\left(E_o(1+z)\right), \qquad (9\text{-}13)$$

and Equation (9-9) would become

$$j_{II} = \frac{cn_i(o)n_j(o)}{4\pi H_o}\int_o^{Z\,max} \frac{(1+z)F_{ij}\left(E_o(1+z)\right)dz}{\left(1+2q_oz\right)^{\frac{1}{2}}}. \qquad (9\text{-}14)$$

If one or both of the interacting species changes in number with time (e.g., cosmic rays being created in galaxies and leaking into intergalactic space), then n_i could be written as

$$n_i = (1+z)^3 n_i(o) f_i(z), \qquad (9\text{-}15)$$

and Equation (9-9) would be

$$j_{III} = \frac{cn_i(o)n_j(o)}{4\pi H_o}\int_o^{Z\,max} \frac{(1+z)f_i(z)F_{ij}\left(E_o(1+z)\right)dz}{\left(1+2q_oz\right)^{\frac{1}{2}}}. \qquad (9\text{-}16)$$

Normally, these integrals have to be solved by numerical integration. However, there is an interesting simple case which can be used as an example. Consider the case represented by Equation (9-12) and assume that q_o is near enough to zero to be neglected over the presently estimated age of the existence of galaxies, i.e. z_{max}=(3 to 4). Further assume that Q_i, at the present time, is a power law in energy, so that

$$Q_i\left(E_o(1+z)\right) = C_i E_o^{-b_i}(1+z)^{-b_i}. \qquad (9\text{-}17)$$

Also, assume that

$$f(z) = (1+z)^{a_i}, \qquad (9\text{-}18)$$

then

$$j_I = \frac{cn_i(o)}{4\pi H_o} \int_o^{Z\,max} \frac{C_i dz}{E_o^{b_i}(1+z)^{b_i+1-a_i}}. \qquad (9\text{-}19)$$

For $a_i + 2\text{-}b_i \neq 1$, Equation (9-19) becomes

$$j_I = \left[\frac{cn_i(o)}{4\pi H_o} \left(\frac{C_i}{E_o^{b_i}} \right) \frac{(1+z)^{(a_i-b_i)}}{(a_i - b_i)} \right] \Bigg|_0^{Z\,max}. \qquad (9\text{-}20)$$

Hence, individual sources with power-law spectra of a given slope will yield a power law of the same slope. This aspect of the result (i.e., a power law yields a power law of the same slope) is true even if q_o is not zero; however, the integral over z to obtain the multiplying factor is not as simple.

In the equations just developed, there are the fundamental parameters H_o and q_o. There has been a long effort to determine H_o well, and there is still a fairly wide range of uncertainty. Current estimates, based largely on the results of the Hubble Space Telescope, seem to point to a value of 70 ± 7 km $s^{-1}Mpc^{-1}$. The determination of q_o is a matter of substantial debate at present. Under the assumption that the cosmological constant, Λ, is zero (Friedmann model), q_o equals $\Omega/2$, where Ω is the density parameter defined as p/p_c, and p_c, in turn, is the critical density. There are a number of types of evidence that suggest that the density parameter is small if there is little or no dark matter, possibly of the order of 0.1. One of the stronger arguments in favor of a low density is the density value determined for deuterium. However, if there is substantial dark matter, the value of q_o could be significantly higher.

9.4 The Origin of the Diffuse Radiation

There has been a wide variety of explanations proposed for the diffuse radiation, as well as a number of theories that led to the prediction of a diffuse gamma radiation naturally. As the spectrum of the radiation became better known, the isotropy better established, and other knowledge improved, more and more of these began to appear to be unlikely origins to the gamma radiation that is now observed. It appears now that active galaxies are the most promising explanation, with Seyfert and radio galaxies accounting for most of the x-ray emission (e.g., Zdziarski, 1996; Leiter and

Boldt, 1997) and blazars, including the BL Lacs and the quasars with beams in our direction, supplying the gamma radiation. With this being the situation at present, the blazar explanation for the gamma radiation will be considered first, to be followed by brief descriptions of some of the other alternate ideas, some of which may still play a role.

Blazars

When only one active galaxy had been seen to be a gamma-ray emitter, and even before, there were numerous theoretical proposals suggesting that they might be a possible origin of the diffuse gamma radiation; so this idea is not new. The SAS-2 measurement had given a reasonably accurate measurement of the diffuse intensity at 100 MeV. With the observation now of the numerous blazars seen by EGRET and discussed in Chapter 8, there is a much firmer base for the calculation of what might be expected, and there have been a number of papers written in which estimates have been made. One obvious factor in support of the blazar origin is the spectrum of the diffuse radiation, which is essentially the same in the high-energy region as that of the blazars, i.e., a power-law spectrum in energy with a slope of -2.09 ± 0.02 compared to -2.12 ± 0.04 for the average blazar spectrum. It can be shown relatively quickly that, within the constraints of the present data and the existing energy range for which data exist, the variations in the energy spectra between quasars will not have much effect.

Consider now some of the specific calculations based on the EGRET blazar results. They generally fall into two categories. The first assumes that the gamma-ray evolution is similar to that at other wavelengths and then uses the gamma-ray results to deduce the expected diffuse radiation. The second attempts to deduce the evolution from the gamma-ray data alone. The advantage of the former is that, if the assumption of a common evolution is correct, an estimate with less uncertainty is obtained. The advantage of the latter is that there are no assumptions of this sort, but the uncertainty in the result is relatively large, because the gamma-ray blazar sample is very much smaller than the radio one, for example. In both cases, the result, of course, is dependent on the cosmological model. In the following, the latter will be considered first, since, as will be seen, the result of this work tends to give support to the other approach.

Chiang et al. (1995) used only the gamma-ray data to deduce the evolution function for the gamma-ray emission of these objects and then calculated the diffuse emission using that function. They used the V/V_{max} test in the context of pure luminosity evolution (currently the preferred concept) to show that there was, indeed, evolution of the high-energy gamma-ray-emitting blazars and to deduce the parameters describing the evolution. They found that the evolution was consistent within errors to that seen at other wavelengths and that the observed high-energy gamma-ray diffuse radiation can be explained as the sum of these active galaxies, although the calculated uncertainty is larger, because of the smaller number of the observed gamma-ray-emitting blazars.

Assuming that the evolution is the same as that supported by the work of Chiang et al., described in the last paragraph, several authors have proceeded to accept the radio evolution function and to apply it to the gamma-ray blazar data and estimate the diffuse radiation (e.g., Salamon and Stecker, 1994; Setti and Woltjer, 1994; Erlykin and Wolfendale, 1995). The estimates agree with the observation of the diffuse radiation in intensity, and, of course, the spectra also agree as the observations say they do. The uncertainty estimate varies from 25% to a factor of 2 depending on the author's approach to the uncertainty evaluation.

Since the spectra of the diffuse radiation and the blazars are both power laws in energy and the slope for the diffuse radiation agrees with the average blazar slope to within less than one σ in the high-energy gamma-ray region and the estimated intensity of the diffuse radiation agrees with that which is observed, the blazar origin seems to be the most likely candidate for this radiation, especially in light of the apparent diffculties of other explanations, to be described in the next paragraph. Since this is the case, it is interesting to try to extend the prediction to energies below 30 MeV. McNaron-Brown et al. (1995) describe the data for the five blazars seen by OSSE in the energy range from approximately 0.1 to 10 MeV and compare it to the COMPTEL data (which exist for four of them) and the EGRET fluxes. There is a clear change in spectral slope on the average, although some slopes have relatively large uncertainties. Using these results, although admittedly few in number, Sreekumar et al. (1997) have estimated a spectrum, shown in Figure 9.2 as the long dashed line.

Also shown in Figure 9.2 are the estimated intensities for several x-ray components as deduced by Zdziarsky (1996). The combined result is similar to that deduced by Boldt and Leider (1996), although differing in detail. As one can see in the figure, there is good agreement over eight decades between the observations and the estimated AGN origin, with the exception of the approximately one decade in energy centered on 1 MeV. In this decade, as noted earlier, it is very difficult to obtain good experimental measurements because of the high, time-variable, instrumental and spacecraft background.

One can even take a tentative step further, related to the highest energies shown in Figure 9.2. Based on the agreement with the blazar origin over three decades in energy, there is the suggestion that the average blazar spectrum continues without a major change of slope at least to 100 GeV. In general, the individual blazar spectra extend only from (30 to 50) MeV to (4 to 10) GeV, beyond which there is not a sufficient flux to determine a spectrum. There is no conflict with this proposal, since theoretically the acceleration could easily generate relativistic particles to produce the photons in this energy range and beyond (remember, two close BL Lacs were seen in the TeV range), and intergalactic absorption for the great bulk of the blazars is not expected to be significant until beyond 100 GeV.

Another corollary of the diffuse radiation result, if the blazar origin is accepted, is that there is not a large variation in slopes of the blazars. If there were, there would be a curve to the extragalactic diffuse spectrum, which is not seen. This conclusion is consistent with the analysis of Mukherjee et al. (1997), discussed in Chapter 8, which suggests that there is a relatively narrow distribution of blazar-energy spectral slopes.

Other Possible Origins of the Diffuse Radiation

Many explanations for the diffuse radiation have been proposed over the years. It is not the intention here to give a full, detailed summary of all of them. Rather, a few will be discussed, because they are thought to be of special interest or because it is believed that they might make a modest contribution.

In Chapter 6, the studies of Page and Hawking (1976) and Hawking (1977), attempting to develop a unified theory of general relativity and quantum mechanics, were noted. One aspect of this work is

that the primordial black holes postulated by Page and Hawking can evaporate via a particle tunneling process at an accelerating rate that ends in an explosive gamma-ray burst. The time at which this burst occurs is inversely proportional to the black-hole mass. A 10^{15} gram primordial black hole formed at the time of the big bang would just now be exploding with a burst of gamma rays, whose average energy and duration are about 100 MeV and about 10^{-7}s respectively, although these numbers are somewhat model-dependent. These events are predicted to be significantly harder and faster than the low-energy gamma-ray bursts to be discussed in Chapter 10. The primordial black-hole bursts, should they occur, when integrated over all z, could produce a diffuse background, and the spectral shape would then place constraints on the primordial black hole mass spectrum and the evaporation process. The observed spectral shape of the diffuse radiation is quite different than that predicted in this theory. There is no evidence for the maximum and rapid decrease at high energies that is predicted.

A diffuse radiation is also predicted in a baryon-symmetric universe. The theoretical concept of baryon nonconservation generated a renewed interest in the possibility of a baryon-symmetric universe containing clusters of galaxies of matter and others of antimatter. The nucleons and antinucleons could interact at the boundaries leading to gamma rays. The shape of the spectrum is determined by the interaction process and the cosmological integration. The net result is a spectrum that peaks at about one MeV (Stecker, Morgan, and Bredekamp, 1971). There is now no compelling evidence for such a peak. Also, there is no evidence for an enhancement at the boundaries of superclusters.

It is known that normal galaxies emit gamma rays; hence, it is reasonable to ask what they might contribute to the diffuse radiation. The emission from our galaxy is known, and the density and size distribution of galaxies is reasonably well known. A more difficult question is what should be used for the evolution function, f(z). There is no knowledge of the gamma-ray emission of normal galaxies as a function of z. There can be various arguments put forward, but, mostly because there is no compelling argument to do otherwise, f(z) is generally chosen to be 1. Most authors (e.g., Kraushaar et al., 1972; Strong et al., 1976; Lichti, Bignami and Paul, 1978; Fichtel, Simpson and Thompson, 1978) have come to the same conclusion for the intensity above 100 MeV, namely, that normal galaxies contribute about 3% to 10%. Further, the spectrum

of the Galactic radiation is very different from the observed diffuse radiation, being less steep at low energies and more steep in the 1 GeV to 50 GeV region; therefore, it is unlikely that normal galaxies make a significant contribution.

Two other early suggestions involved either cosmic rays leaking from galaxies or primordial cosmic rays interacting with intergalactic matter or photons. Both the required intensity and the spectral shape that is now known make these possibilities seem unlikely.

9.5 Summary

The intensity level of the diffuse radiation and its spectrum are now well established after a long period of difficult research, except in the region from about 0.3 to 4 MeV, wherein measurements are very difficult and further work is needed. There is no evidence for any major anisotropy. Although there are features of the diffuse radiation about which one would still like to have a better knowledge, the blazar origin seems to be reasonable and consistent with the data in the medium- and high-energy gamma-ray region, both in terms of the spectrum and intensity, and there is no other strong contender. The sum of the fluxes of the Seyfert galaxies seems to explain the x-ray diffuse emission. Thus, it seems that the majority of the diffuse radiation is finally reasonably well described, and there is a satisfactory explanation of its origin.

References

Boldt, E., and Leiter, D., 1996, private communication of work in progress.

Bondi, H., and Gold, T., 1948, Mon. Not. Roy. Ast. Soc., 108, 252.

Bottinelli, L., and Gouguenheim, L., 1976, A&A, 51, 275.

Branch, D., 1979, Mon. Not. Roy. Ast. Soc., 186, 609.

Brown, R. W., and Stecker, F. W., 1979, Phys. Rev. Letters, 43, 315.

Chen, L. W., Fabian, A. C., Gendreau, K. C., 1996, MNRAS, in press.

Chiang, J., Fichtel, C. E., Montigny, C. von, Nolan, P. L., and Petrosian, V., 1995, ApJ, 452, 156.

Dicke, R. H., Peebles, P. J. E., Roll, P. G., and Wilkinson, D. T., 1975, ApJ, 142, 414.

De Vaucouleurs, G., and Bollinger, G., 1979, ApJ, 233, 433.

Erlykin, A. D., and Wolfendale, A. W., 1995, J. of Physics G, 39, 21.

Fichtel, C. E., Simpson, G. A., and Thompson, D. J., 1978, ApJ, 222, 833.

Fichtel, C. E., et al., 1994, AAS High Energy Astrophysics Meeting, Napa Valley.

Gendreau, K. C., 1995, PASJ, 47, L5.
Gott, J. R., III, and Turner, 1977, ApJ, 213, 309.
Gott, J. R., III, Gunn, J. E., Schramm, D. N., and Tinsley, B. M., 1974 ApJ, 194, 543.
Gruber, d. E., 1992, in The X-Ray Background, ed. X. Barcons, and A. C. Fabian, Cambridge University Press, Cambridge, p. 44.
Hawking, S. W., 1977, Scientific American, 236, 34.
Hoyle, F., 1948, Mon. Not. Roy. Ast. Soc., 108, 372.
Hunter, S., et al., 1996, ApJ, submitted.
Kappadath, S. C., et al., 1995, in Proc. 24th Int. Cosmic Ray Conf., Rome, Vol. 2, p. 230.
Kappadeth, S. C., et al., 1995, A&A, {The Third Compton Symposium}, to be published.
Kinzer, R. L., Jung, C. V., Gruber, D. E., Matteson, J. L., Peterson, L. E., 1996, ApJ, submitted.
Kniffen, D. A., et al., 1994, June AAS Meeting.
Kraushaar, W. L., Clark, G. W., Garmire, G. P., Borken, R., Higbie, R., Leong, C., and Thorsos, T., 1972, ApJ, 177, 341.
Leiter, D., and Boldt, E., 1997, in preparation.
Lichti, G. G., Bignami, G. F., and Paul, J. A., 1978, Astrophys. and Space Sci., 56, 403.
McNaron-Brown, K., et al., 1995, ApJ, 451, 575.
Osborne, J. L., Wolfendale, A. W., Zhang, L., 1994, J. of Phys. G.: Nucl. Part. Phys., 20, 1089.
Page, D. N., and Hawking, S. W., 1976, ApJ, 206, 1.
Peebles, P. J., 1993, Principles of Physical Cosmology (Princeton:Princeton University Press, 1993).
Penzias, A. A., and Wilson, R. W., 1965, ApJ, 142, 1149.
Rogerson, J. B., Jr., and York, D. G., 1973, ApJ, 186, L95.
Salamon, M. H., Stecker, F. W. 1994, ApJ, 430, L21.
Sandage, A., and Tammann, G. A., 1976, ApJ, 210, 7.
Schönfelder, V., Graml, F., and Penningsfeld, F. P., 1980, ApJ, 240, 350.
Schramm, D. N., and Wagoner, R. V., 1977, Ann. Rev. Nucl. Sci., 27, 37.
Schwartz, D. A., and Gursky, H., 1973, Gamma Ray Astrophysics, ed. F. W. Stecker and J. I Trombka, NASA SP -339.
Setti, G., and Woltjer, L., 1994, ApJS, 92, 629.
Sreekumar, P., et al., 1997, ApJ, submitted.
Stecker, F. W., 1980, Phys. Rev. Letters, 44, 1237.
Stecker, F. W., Morgan, D. L., Bredekamp, J., 1971, Phys. Rev. Letters, 27, 1469.
Strong, A. W., Wolfendale, A. W., and Worrall, D. M., 1976, J. Phys. A: Math Gen., 9, 1553.
The, L.-S., Leising, M. D., and Clayton, D. D., 1993, ApJ, 403, 32.
Thompson, D. J., and Fichtel, C. E., 1982, A&A, 109, 352.
Tinsley, B. M., 1977, Annals of the New York Academy of Sciences, 302, 423.
Trombka, J. I., et al., 1977, ApJ, 212, 925.
Tulley, R. B., and Fisher, J. R., 1977, A&A, 54, 661.

Weinberg, S., 1972, Gravitation and Cosmology (New York: John Wiley and Sons, 1972), p. 469.
Yang, J., Schramm, D. N., Steigman, G., and Rood, R. T., 1979, ApJ, 277, 697.
Zdziarski, A., 1996, MNRAS, 281, L9.

CHAPTER 10

GAMMA-RAY BURSTS

10.1 Introduction

In spite of the time that has passed since the unexpected discovery in the early 1970s of gamma-ray bursts by Klebesadel, Strong, and Olson (1973), and in spite of the fact that there is now substantial information about them, particularly with the results from the Compton Gamma-Ray Observatorty (CGRO), their origin remains a mystery, although some important constraints have been established now.

One observed feature of these bursts has remained constant since their discovery, namely that they are still a purely gamma-ray phenomenon. In spite of extensive efforts to find counterparts at other wavelengths, including radio, infrared, optical, ultraviolet, and x-ray, none have been found. (It should be noted that counterparts have been found for the separate class of sources called "soft gamma-ray repeaters".) The search for counterparts has been extensive, since it has been hoped that finding them would resolve, or assist in solving, the mystery. The failure to find associated emission at other wavelengths may be because of the character of the phenomenon, specifically that it is a short burst, but, as time passes and searches become more extensive, this explanation becomes less probable. The alternative is that they are basically a gamma-ray phenomenon. Current attempts at a theoretical understanding have been made largely with this concept in mind.

10.2 Distribution of Sources in Direction and Intensity

Even prior to the results from the Compton Observatory, there were strong suggestions that a rather surpising combination of observational results existed. The distribution of directions on the sky was consistent with isotropy, suggesting a uniform distribution

in space for the observed bursts. At the same time, when the number of bursts with intensities greater than a given intensity was compared to the observed intensity, rather than following a simple power law with a −3/2 slope to the minimum detectable intensity, as would be expected if the distribution were uniform, the observed number began to fall increasingly below this line as the observed intensity became increasingly small. Thus, there was the implication that there was a spherical boundary beyond which the number of sources decreased markedly.

The surprise associated with this result was that it was different from either of the two expected cases. One of these was that there would be no deviation from the −3/2 power law down to the threshold, because an isotropic local phenomenon was being detected, and the detector threshold had been reached before the intensity level had gone significantly below that of bursts from beyond one of the boundaries. The other expected alternative was that, if there was a deviation from the −3/2 power law in the number as a function of size, indicating that a boundary had been reached, a nonisotropic distribution would be seen. The latter assumption arises from the fact that the local distributions, namely the solar system and the Galaxy, both naturally would lead to a disk-like distribution of the more distant, or weaker, sources.

The BATSE results from the Compton Observatory have confirmed this experimental picture and have shown it to be true to the degree that it is no longer in doubt. Figure 10.1 shows the intensity distribution of bursts that exceed the 1024 ms trigger threshold of BATSE, based on the BATSE 3B Catalog (Meegan et al., 1996a). The solid line represents the total number of bursts observed, and the dashed line has a slope of −3/2. The deviation from the −3/2 power law is obvious.

The bursts detected by BATSE are generally located in direction to an accuracy of a few degrees. This is quite adequate to determine the distribution on the sky when looking for major characteristics, such as whether there is a disk population or if the distribution is primarily isotropic. Figure 10.2 shows the location of the bursts seen by BATSE in Galactic coordinates. There is clearly no evidence for a Galactic component or solar disk component. The

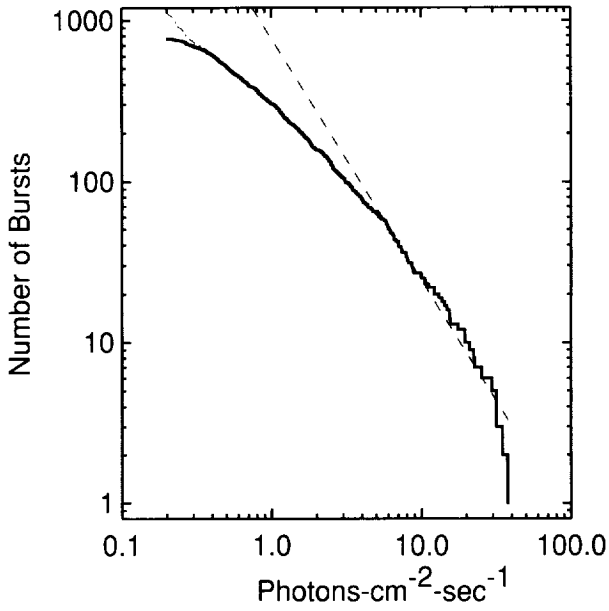

Figure 10.1. The intensity distribution for 772 GRBs that exceed the 1024 ms trigger threshold. This figure is from Meegan et al. (1996a) and is reproduced with their permission.

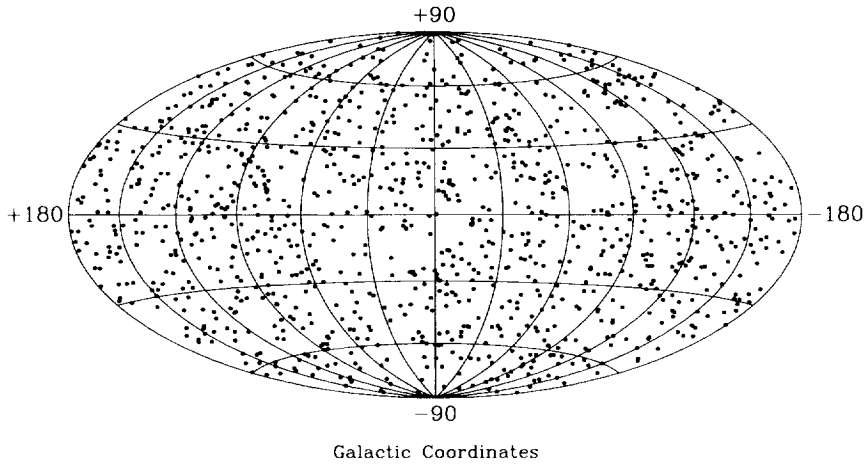

Figure 10.2. The locations of 1122 bursts in Galactic coordinates. This figure is from Meegan et al. (1996a) and is reproduced with their permission.

distribution appears to be consistent with isotropy and, indeed, mathematical tests confirm that it is. Meegan et al. (1996b) found that the Galactic dipole moment corrected for anisotropic sky coverage was $<\cos\Theta> = 0.011 \pm 0.017$, and that the Galactic quadrupole moment corrected for anisotropic sky coverage was $< \sin^2 b - 1/3 > = 0.002 \pm 0.009$. Both of these numbers are seen to be consistent with an isotropic distribution within 1 sigma. The constraint on any nonisotropic component is obviously severe. Several attempts have been made to look for anisotropies in subsets of the data, and none known to the authors has led to a significant result.

For some bursts, the directions are known more precisely. These better locations are achieved in one of two ways. The first is the use of detectors on three or more satellites widely spaced, which measure the arrival time of the burst. The direction can then be determined from a calculation based on these times. The two first reasonably good positions determined in this manner were by Cline et al. (1979a, 1979b); neither was consistent with any known high-energy astronomical source, such as an x-ray emitter. Subsequently, many gamma-ray bursts have been located quite accurately, using the interplanetary network of satellites (See, e.g., Cline et al., 1981, 1984.), some with an accuracy of a fraction of an arcmin square. None of these source locations corresponds with certainly to any known constant, high-energy source, although a few have had subsequent, transient x-ray and optical brightening. The other approach is the use of the high-energy gamma-ray data to be discussed in Section 10.4. It is more precise than the direction determined from BATSE data alone, but generally not as accurate as the several satellite methods just described. However, in conjunction with two other satellites, the small error circle determined by the high-energy gamma-ray data and the arc through it determined from the two satellites gives a small uncertainty in solid angle, a fraction of a degree by about one or a few arc minutes. None of these determinations of direction has led to any certain location of a burst with an obvious constant-intensity counterpart.

It should be noted again, perhaps, that the three soft gamma-ray repeaters, which have been seen many times each, are clearly a different phenomenon, even though they have some properties similar to the bursts being described here. The locations of all three of these are well determined. Two of these three sources have been

clearly identified with supernovae remnants, and the third may be one also. For a further discussion of soft gamma-ray repeaters, see, e.g., Cline (1996) and Hurley (1996).

10.3 Energy and Temporal Distributions in the Low-Energy Gamma-Ray Region

Prior to the launch of the Compton Observatory, the energy spectral measurements were generally limited to a relatively narrow range of energies. For example, the spectra measured by the IMP satellite experiments covered the range from 0.06 to 1 MeV, and those determined by the Konus detectors on Veneras 11 and 12 covered the range from 0.03 to 2 MeV. Detectors on SMM and the Ginga burst detector extended that range somewhat. Finally, the combination of the BATSE, OSSE, COMPTEL, and EGRET instruments on the Compton Observatory covered the range from 10 keV to more than 10 GeV.

If the spectrum is expressed in terms of νf_ν, the spectrum typically has a broad maximum with the peak typically on the order of 150 keV with very few bursts having a maximum above 500 keV. Those bursts observed at high photon energies seem to have a spectrum that continues smoothly, but flattens after a decline from the broad maximum just mentioned. There is substantial variation between events and during an event. Although they are not universal, two trends seem to be characteristic of the variability (Golenetskii, et al., 1983; Norris, et al., 1986; Ford et al., 1995; Band, 1996b). One is that the spectrum tends to be harder when the gamma-ray flux increases. However, it is also true that bursts generally begin with hard spectra, and successive spikes are usually softer, even when the count rate is greater.

Although there had been some evidence for gamma-ray lines in bursts prior to the BATSE observations, particularly from Ginga (e.g., Palmer et al., 1994), no lines have been detected by BATSE at this time (e.g., Band et al., 1996a). Although this result is of concern, there is not yet a very strong conflict in view of detailed sensitivity considerations.

Turning to the time histories, there is a rich variety of types (See, e.g., Fishman et al., 1994.). There are bursts with a single peak, those with two peaks, and those with many peaks. There are those that are jagged and those that are relatively smooth. In a series of

peaks, the largest can come first, last, or in between. The flux can appear to be declining slowly and then have a sharp peak. There are those with a very broad maximum and those with a series of very sharp peaks. Attempts to classify the bursts into various types understandably have been only moderately successful.

If the duration of the bursts is examined, a structure does seem to exist, as shown in Figure 10.3 (Meegan, et al., 1996a). In this figure, the number of bursts is shown as a function of their duration. T_{90} is the time interval in which the integral counts from the burst increases from 5% to 95% of the total. There are clearly two groups of events in the figure, one with a broad maximum between 0.2 and 1 second and another with the maximum between 10 and 100 seconds. Notice also that there are none seen in the low-energy range with a duration in excess of about 700 seconds. The longer events, those with durations more than 1.5 seconds, contain most, but not quite all, of the events with the highest peak fluxes.

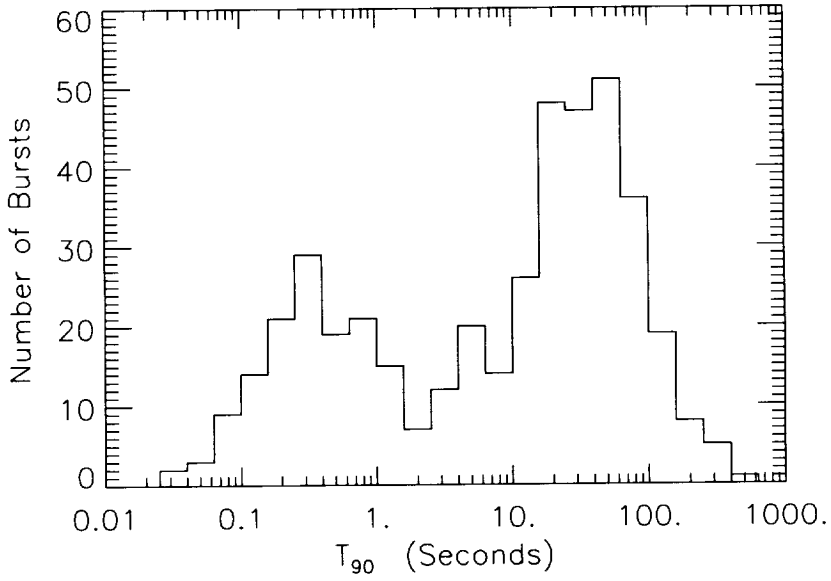

Figure 10.3. Duration distribution for 427 bursts in the 3B BATSE Catalog that do not have significant data gaps and that are above the 64 ms trigger threshold as a function of T_{90}. See the text for a definition of T_{90}. The figure is from Meegan et al. (1966a) and is reproduced with their permission.

However, there are a substantial number of dim bursts with long durations. The short duration bursts (those with durations of less than 1.5 seconds) mostly are relatively dim, i.e., they have relatively small peak fluxes. These latter observations alone would be consistent with the idea that there are a range of intensities of events and types, and that one sees as dim sources the largest peaks of the more distant ones that are jagged. The feature that does not fit well into this picture is the relatively small number of bursts in the 1 to 10 second range. Taking this into account with the other observed characteristics suggests that there are probably two different populations. Whether this, in turn, implies two types of one phenomenon or two distinctly different phenomena is not known at this time.

Attempts have been made to find characteristics that would indicate a cosmological nature of the gamma-ray bursts. Searches for time-dilation effects haven been made, including the study of pulse widths, duration, intervals between peaks in events, and burst intervals. There would also be red-shift effects that might appear in the form of different spectra coupled with any of the time-dilation effects. There are a number of complications associated with these searches, including brightness bias, identifying true pulses, determining true intervals, and width corrections. Nonetheless, there does now seem to be increasing support for a time-dilation effect being seen. For a summary, see, e.g., Norris (1996). The interpretation of the meaning of the observations is, however, not yet certain, and other explanations are possible.

10.4 High-Energy Gamma-Ray Emission

Prior to the launch of the Compton Gamma-Ray Observatory in April 1991 it had been thought that these bursts were a low-energy gamma-ray phenomenon; in fact, they were sometimes referred to as "low-energy gamma-ray bursts". The results from EGRET, with the detection of gamma rays with energies not only of several hundred MeV, but photons with energies well above 1 GeV, have changed this understanding and added a large number of new constraints that any satisfactory explanation of these bursts must now satisfy.

A total of 5 bursts containing photons with energies in excess of 30 MeV have been detected by EGRET during the period from April 1991 through July 1996. They are among the brightest bursts

detected by BATSE that are in the EGRET field of view. They are listed in Table 10.1. Also shown in this table is the highest-energy gamma ray detected in each event. Notice that the energy of one of these photons exceeds 10 GeV, and another is nearly 10 GeV. The high-energy gamma rays detected by EGRET generally arrive later than the low-energy gamma rays detected by BATSE. The total energy per unit area for the whole event is usually somewhat less for the high-energy gamma rays than that of the low-energy gamma rays, but not by more than a factor of 10. The typical time for the low-energy burst observed by BATSE is from a fraction of a second to tens of seconds. The high-energy gamma rays are seen for a longer period of time; in two of the bursts seen at high energies, the high-energy gamma rays have been seen for approximately 5×10^3 seconds. Although the statistics are very limited, there is no indication of a change in distribution of the energies of the high-energy gamma rays with time in the event.

Because of the limited number of high-energy gamma rays that have been detected, it is difficult to construct a meaningful spectral index for any single event. However, if the data for all five events described in Table 10.1 are combined and a spectrum that is a power law in energy is assumed, then a spectral index of -1.95 ± 0.25 is obtained for the energy range from 30 MeV to 10 GeV (Dingus et al., 1996). This slope is slightly less negative than it is in the 1 to 30 MeV range, but consistent with it being the same within uncertainties. It is clearly a shock, acceleration-type spectrum rather than a thermal spectrum.

Table 10.1.

Gamma-Ray Bursts Detected Above 30 MeV by EGRET

Date	Time	Highest-Energy Gamma Ray (GeV)
910503	07:04:13	9.6 ± 2.0 @84 seconds
910601	19:22:14	1.3 ± 0.1 @ 4835 seconds
930131	18:57:12	1.2 ± 0.2 @ 24 seconds
940217	23:02:42	18.4 ± 3.9 @ 4756 seconds
940301	20:10:37	0.16 ± 0.03 @22 seconds

Although high-energy gamma rays have been detected in only a few bursts, they are among the brightest seen by BATSE in EGRET's field of view, as previously noted. Hence, it may very well be that these high-energy photons are a feature of all bursts, but are normally too few in number to be detectable. It is also possible that the low-energy gamma rays persist for a long time at a low flux level, but simply have intensities below the threshold determined by the background in the low-energy detectors.

For more detailed information on the properties of the high-energy gamma rays in bursts, see Schneid (1992), Sommer et al. (1994), Hurley et al. (1994), Dingus et al. (1994), Schneid et al. (1995), and Dingus et al. (1996).

10.5 Possible Sources and Mechanisms

It is useful to begin with a summary of the principal results that have been described in the previous sections, since they provide the constraints that any theory must satisfy. One of the first is the combination of the distribution of the number as a function of size, or intensity, and the observed isotropy. The former implies a boundary or edge of the population, and the latter, a quite uniform distribution. The combination eliminates, for example, a disk-type distribution that might exist for a solar system of Galactic origin, but strongly favors a very large Galactic halo model or a cosmological model wherein the boundary is far back in time.

The bursts appear to be purely gamma ray in nature. The low-energy gamma-ray emission is usually short, from a fraction of a second to seconds, and occasionally hundreds of seconds. The high-energy gamma rays, although seen early in events, also have been observed to last for up to 5×10^3s. The rapid time variations at low energies imply a small region, but the high-energy gamma-ray flux implies a density of energetic particles at the source that is quite high. This combination appears to suggest beaming, if particle-particle interactions that would seriously downgrade the energy of the source particles are to be avoided.

High-energy gamma rays have been seen in five of the stronger events, suggesting that they are a characteristic of these bursts. The energy spectrum of the high-energy gamma rays appears to be a power law in energy with a slope of about -2; it is certainly not a thermal spectrum. Individual photons have energies over 1 GeV,

with one having an energy as high as 18 GeV. These properties, together with the beaming suggested in the last paragraph, point to a shock-type acceleration for the parent high-energy particles.

When the data for the whole spectrum exist, many seem to have a similar general appearance. When expressed in terms of E^2, multiplied by the differential energy spectrum, the curve rises from the lowest observable energy to a broad maximum at about 1 MeV, then decreases a bit, and finally is reasonably flat to as high an energy as is observed, with the latter part being that just noted in the previous paragraph.

The constraints that have just been summarized in the last four paragraphs obviously have seriously restricted the possible origins and mechanisms that are viable explanations of these bursts. The local theories originally had more appeal, primarily because they required much less energy at the source. However, their viability has become of increasing concern as the degree of observed isotropy becomes greater and greater. The large boundary of the halo now required by this constraint is approaching the limit that seems reasonable with the current understanding of our galaxy. The search for absorption features by BATSE in the hard x-ray region was not successful, and hence there was no support for the local theories from this area.

On the other hand, any cosmological theory is immediately faced with the explanation of the large luminosities that are implied, namely on the order of 10^{50}ergs s^{-1}. Additionally, the rapid time variations imply a relatively small size for the emitting region. As in the case of blazars, the combination of the energy and the size implies that the parent particles are beamed. If they were not, the high photon densities would make the source region optically thick by many orders of magnitude, and the escape of gamma rays from the source region would not be possible. From a theorist point of view, beaming also has the advantage of reducing the source luminosity, since the radiation may be only in a small solid angle. On the other hand, the number of sources must be correspondingly greater, since the probability of seeing any one source is smaller by the same amount.

There are two ways that beaming may occur. One is that there is a natural beam associated with the source itself, as is believed to be the case with blazars. Another is the relativistic motion of the source

itself. If the source itself is moving at a velocity relative to an observer on Earth, radiation that is emitted isotropically with respect to the source frame will be beamed in the frame of the observer (Krolik and Pier, 1991; Baring, 1993). An analysis of the observed high-energy gamma-ray spectra by Baring (1995) suggests that, for the major bursts, the minimum bulk Lorentz factor ranges between 36 and over 800. The solid angle is roughly this factor squared; so the number of sources and the intensity are significantly effected, should this be the explanation.

Among the several cosmological models for gamma-ray bursts, a scenario that is currently receiving considerable attention is one wherein two neutron stars or a neutron star and a solar-mass black hole coalesce (Paczynski, 1986; Eichler et al., 1989; Meszaros and Rees, 1992). Calculations of the energy available in the form of gamma rays is at least adequate. The expected number of occurrences per unit time may also be sufficient. A very convincing piece of evidence in favor of the cosmological model would be the observation of a gravitationally-lensed gamma-ray burst, but none has yet been observed.

References

Band, D. L., 1996a, in Gamma-Ray Bursts, 3rd Huntsville Symposium, Huntsville, AL, October 1995, ed. C. Kouveliotaou, M. F. Briggs, and G. J. Fishman, AIP Conference Proceedings 384, 123.

Band, D. L., et al., 1996b, ApJ, 458, 746.

Baring, M. G., 1993, ApJ, 418, 319.

Baring, M. G., 1995, Proc. of the Erice NATO ASI "Currents in High Energy Astrophysics", ed. J. Wefel, R. Silberberg and M. Shapiro (Kluwer Academic, Dordrecht), held in Erice, Sicily, Italy, May 7-17, 1994, p. 21.

Cline, T. L., 1996, Current Perspectives in High Energy Astrophysics, ed. J. F. Ormes, NASA Ref. Pub. 1391, 147.

Cline, T. L., et al., 1979, ApJ. 229, L47.

Cline, T. L., 1981, et al., ApJ, 246, L133.

Cline, T. L., 1984, et al., ApJ, 286, L15.

Cline, T. L., Desai, U. D., Pizzichini, G., Spizzichino, A., Trainor, J. H., Klebesadel, R. W., Helmken, H., 1979, ApJ, 232, L1.

Dingus, B. L., et al., 1994, AIP Conf. Proc 307, 2nd Huntsville Gamma-Ray Burst Workshop, ed. G. J. Fishman, J. J. Brainer, and K. Hurley (New York: AIP), p 22.

Dingus, B. L., et al., 1996, in preparation.

Eichler, D., et al., 1989, Nature, 340, 126.

Fishman, G. J., et al., 1994, ApJS, 92, 229.

Ford, L. A., et al., 1995, ApJ, 439, 307.
Golenetskii, S. V., et al., 1983, Nature, 306, 451.
Hurley, K., et al., 1994, Nature, 372, 652.
Hurley, K., et al., 1996, Symposium, Huntsville, AL, October 1995, ed. C. Kouveliotaou, M. F. Briggs, and G. J. Fishman, AIP Conference Proceedings 384, 889.
Klebesadel, R. W., Strong, I. B., and Olson, R. A., 1973, ApJ, 182, L85.
Krolik, J. H., and Pier, E. A., 1991, ApJ, 373, 277.
Meegan, C. A., et al., 1996a, ApJ, 106, 65.
Meegan, C. A., et al., 1996b, Gamma-Ray Bursts, Third Huntsville Symposium, Huntsville, AL, October 1995, ed. C. Kouveliotaou, M. F. Briggs, and G. J. Fishman, AIP Conference Proceedings 384, 291.
Meszaros, P., and Rees, M. J., 1992, ApJ, 397, 570.
Norris, J. P., 1996, Gamma-Ray Bursts, 3rd Huntsville Symposium, ed. C. Kouveliotou, M. F. Briggs, and G. J. Fishman, AIP Conference Proceedings, 384, 13.
Norris, J. P., et al., 1986, ApJ, 301, 213.
Paczynski, B., 1986, ApJ, 308, L43.
Palmer, D. M., et al., 1994, ApJ, 433, L77.
Schneid, E. J., et al., 1992, A&A, 255, L13.
Schneid, E. J., et al., 1995, ApJ, 453, 95.
Sommer, M., et al., 1994, ApJ, 422, L63.

CHAPTER 11

GAMMA-RAY INTERACTION PROCESSES

11.1 Introduction

Gamma-ray detection in the energy region above 1 keV involves measurements of the energy exchange or energy loss between the gamma rays and the mass of the detector. In most cases of interest, it is the kinetic energy imparted to charged particles by the gamma ray that is lost in the detector and measured in order to obtain spectral knowledge of the incident gamma-ray flux. Furthermore, the angular relationship between the incident gamma-ray photon and the direction of the secondary particles contains important energy information.

We begin by considering the interaction gamma-ray removal processes in matter. This interaction removal process is characterized by the fact that each gamma-ray photon is removed individually from the incident beam. By removal we mean either total absorption or removal from the energy group. The number of photons removed in this manner is proportional to the thickness of matter traversed. In the one dimensional case, then

$$I = \mu I_0 \, \Delta X \, , \tag{11-1}$$

where μ is a constant, ΔX is the distance traversed in matter, I_0 is the number flux for the incident beam, and I is the number flux after the beam has traversed a distance X.

Upon integration, one obtains, of course,

$$I = I_0 \, e^{-\mu x}. \tag{11-2}$$

The product μx is dimensionless. With this in mind, a number of removal cross sections or coefficients can be defined. If the dimension term X is expressed in cm, then μ is expressed in cm^{-1}, and the linear absorption coefficient is numerically equal to the fractional number of photons removed per cm of absorber. Now the dimension unit can also be expressed as $x\rho^{-1}$, where ρ is the density of the material; the dimension of the product is then gm cm^{-2}. The mass absorption coefficient $\mu\rho^{-1}$ has the dimensions cm^2 gm^{-1}. The mass absorption coefficient is equal to the fractional reduction in photon intensity produced by 1 cm of absorber. An electronic absorption coefficient can also be defined as $\mu(\rho N_o)^{-1}$ with units cm^2/electron. N_o is Avagadro's number.

Another term of interest in the attenuation of gamma rays is the range. Regarding the number flux, this term can be considered equal to the inverse of the mass absorption coefficient and is thus expressed in terms of gm cm^{-2}; hence, this term reflects the photon range which reduces the flux by a factor of e^{-1}.

The change in the number flux has been considered above. In the energy region being considered here, there are basically three processes by which gamma rays may interact with matter and be removed from the incident beam: the photoelectric effect, Compton scattering, and pair production. These processes act independently of each other, and thus the total interaction cross section can be separated into three parts: the photoelectric effect cross section, τ; the Compton scattering cross section, σ; and the pair production cross section, χ. Thus

$$\mu = \tau + \sigma + \chi. \tag{11-3}$$

The removal processes have just been considered, but what is important in considering gamma-ray detectors is the amount of gamma-ray energy lost in the detector. Some of the more important loss mechanisms will be discussed in this chapter. These processes will be considered in terms of the three major interaction processes. Detailed considerations of the interaction and processes and cross sections have been given by Bethe and Askin (1953), Heitler

(1954), Rossi (1956), Davisson (1968), Hubbell (1969,and 1977), and Hubbell et al. (1980).

Table 11.1 lists briefly the most important energy loss processes of gamma-ray detectors. In the following sections, outlines of only the mechanisms will be presented in order to provide the background needed to understand the analyses and the operations of the detector systems that will be discussed in the following two chapters.

Table 11.1.

Gamma-Ray Energy-Loss Processes of Importance With Respect to the Detection of Gamma Rays in the > 1 keV Energy Domain

I. Principal energy-loss processes

1. Photoelectric effect
2. Pair production
3. Compton scattering

II. Important secondary energy-loss processes

1. Fluorescence radiation
2. Annihilation radiation
3. Electron leakage or escape
4. Bremsstrahlung

By single or multiple interaction processes mentioned here, either all or part of the gamma-ray energy can be absorbed by matter. One can find methods for calculating this energy-loss mechanism (e.g., Rossi, 1956; Goldstein, 1959). Two further terms, relative to the absorption of gamma-ray energy, need to be presented. The energy absorption coefficient μa is given by

$$\mu_a = \left(\frac{\bar{E}_p}{E_p} \right) \mu \qquad (11\text{-}4)$$

where $\bar{E}_p$ is the average energy loss of photons per collision, and E_p is the initial energy of interacting photons. The radiation length X_0 is defined by

$$\frac{1}{X_0} = 4\alpha\left(\frac{N_0}{A}\right)Z^2 r_e^2 \ln\left(183Z^{-1/3}\right), \qquad (11\text{-}5)$$

where Z and A are the charge and mass number of material traversed, N_0 is Avagadro's number, r_e is the classical radius of the electron, and $\alpha = e^2/\hbar c = 1/137$. This term comes from the quantum mechanical calculation of photon-energy loss. In the description of certain radiation phenomena, there is only a slight dependence of a number of parameters on the atomic number, when thicknesses are measured in radiation lengths. An example of such an effect will be presented in the section on pair production. A more detailed discussion on radiation lengths can be found in Rossi (1956). Detailed compilations of gamma-ray cross sections in the energy region of interest can be found in Hubbell (1969, and 1977), and Hubbell et al. (1980).

11.2 Photoelectric Effect and Fluorescence Radiation

Photoelectric absorption results from the interaction of gamma rays with bound electrons of the detector crystal. All the energy of the gamma ray is lost in this interaction, but not all the energy is imparted to secondary electrons as kinetic energy; some of it is required to overcome the binding energy of the electron. However, after the photoelectric absorption, x-rays are produced with energies almost equal to this binding energy. The absorbtion of these x-rays and their conversion to kinetic energy of secondary electrons will then reclaim, in a sense, the lost energy. In principle, some of the excess photon energy goes into the kinetic energy of the recoiling atom, but this is a negligible fraction. On the other hand, the excess momentum carried off by the recoiling atom is important, for it can be shown that momentum cannot be converted in the photoelectric effect with a free electron. Therefore, the binding of the electron to the atom is all-important for these phenomena (Heitler, 1954).

For photon energies very large compared to the ionization energy, the electron appears to be only lightly bound, and the photoelectric effect becomes relatively improbable. Hence, as the photon energy increases, the cross section for photoelectric emission decreases rapidly from its maximum value at the binding energy. In the low Z-elements, the binding of even the innermost K-electrons is quite weak, and the photoelectric effect is correspondingly small for

photon energies of interest here. As Z increases, the binding energy increases rapidly, and thus the photoelectric effect becomes more prominent. The increase of the cross section with an atomic number goes roughly between Z^4 and Z^5. The K-shell electrons contribute the most to the photoelectric effect, since they are the most tightly bound electrons. For photon energies below the K-ionization energy, however, only the L- and higher shell electrons can be ejected by the photon, and they provide the entire photoelectric effect. Thus, going down in energy, the photoelectric effect drops almost discontinuously at the K-ionization energy to the much lower values characteristic of the L-shell cross sections. Figures 11.1, 11.2, 11.3, 11.4, 11.5, and 11.6 show photoelectric mass-absorption coefficient as a function of energy for plastic, NaI, CsI, Ge, Fe, and Al. The pair-production Compton scattering and total mass-absorption coefficients are also shown.

11.3 Pair Production and Annihilation Radiation

In the pair-production process, the incident gamma ray is annihilated in the field of the nucleus, producing an electron-positron pair. The excess of the gamma ray's energy above that required to produce an electron-positron pair at rest is imparted as kinetic energy to the electron and positron. This kinetic energy is then available to produce scintillations, for example. The positron later annihilates with an electron, usually producing two 0.51 MeV gamma rays. These can lose all or part of their energy by Compton scattering and/or photoelectron absorption in the detector crystal. Thus, the original incident gamma ray interacting through the pair-production process may eventually lose any amount of its energy between its total energy and its total energy minus 1.02 MeV to secondary electrons.

Pair production has a threshold at $2mc^2 = 1.022$ MeV, below which it cannot take place. Like the photoelectric effect, pair production cannot take place in free space; to conserve momentum, it can occur only in the electric field of a particle that can carry away some of the momentum in the recoil. Either the atomic electrons or the

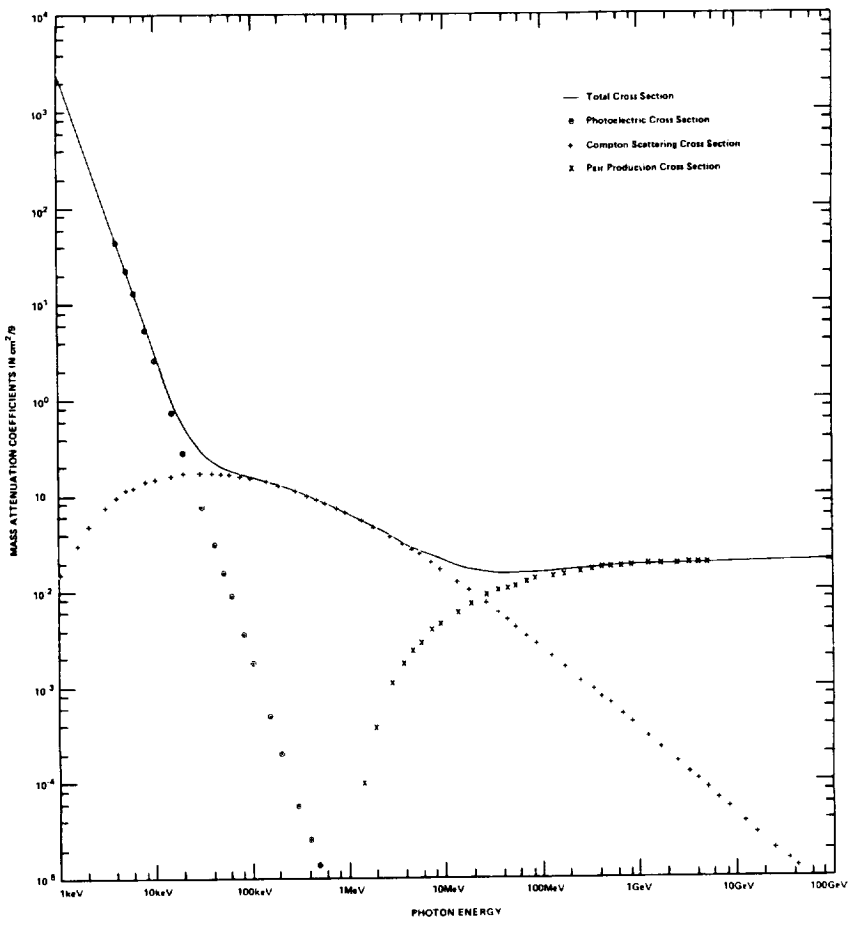

Figure 11.1. The photoelectric, Compton scattering, pair-production, and total mass-absorption coefficient for plastic scintillator plotted flows (from Hubbell, 1969, and 1977, and Hubbell et al., 1980).

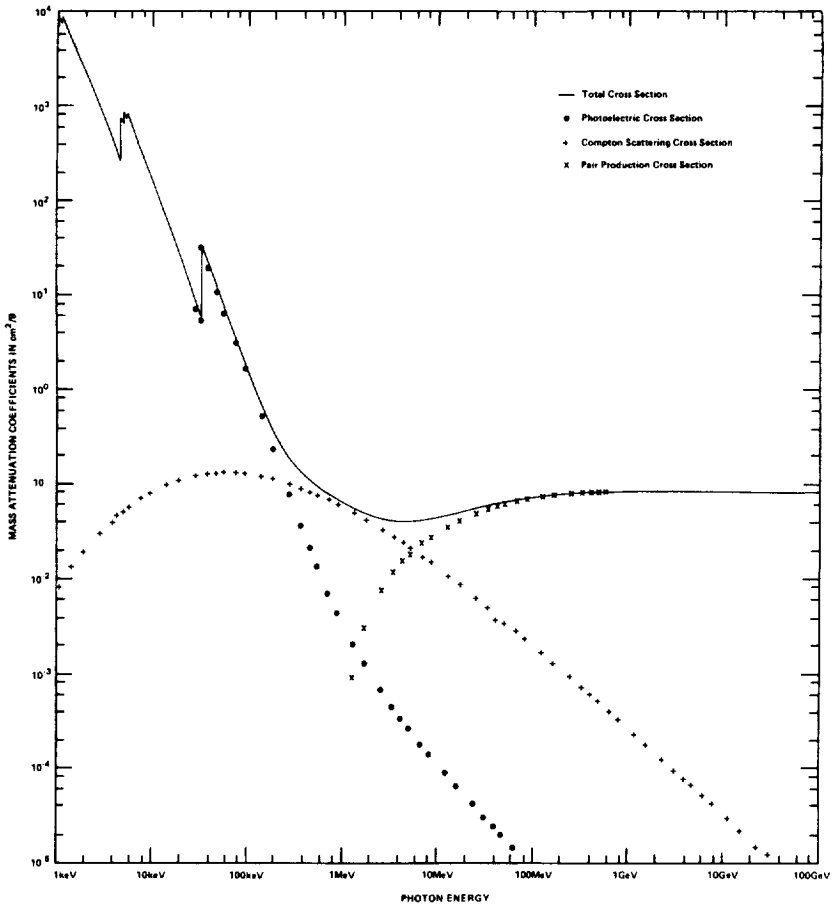

Figure 11.2. The photoelectric, Compton scattering, pair-production, and total mass-absorption coefficient for NaI (from Hubbell, 1969, and 1977, and Hubbell et al., 1980).

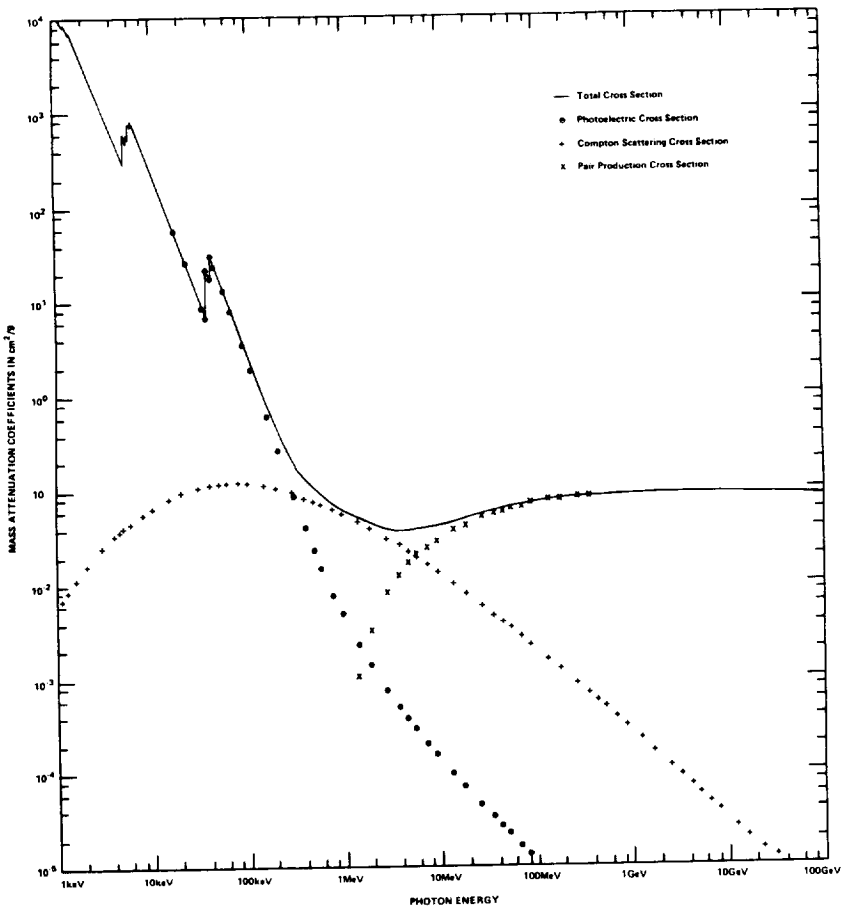

Figure 11.3. The photoelectric, Compton scattering, pair-production, and total mass-absorption coefficient for CsI (from Hubbell, 1969, and 1977, and Hubbell et al., 1980).

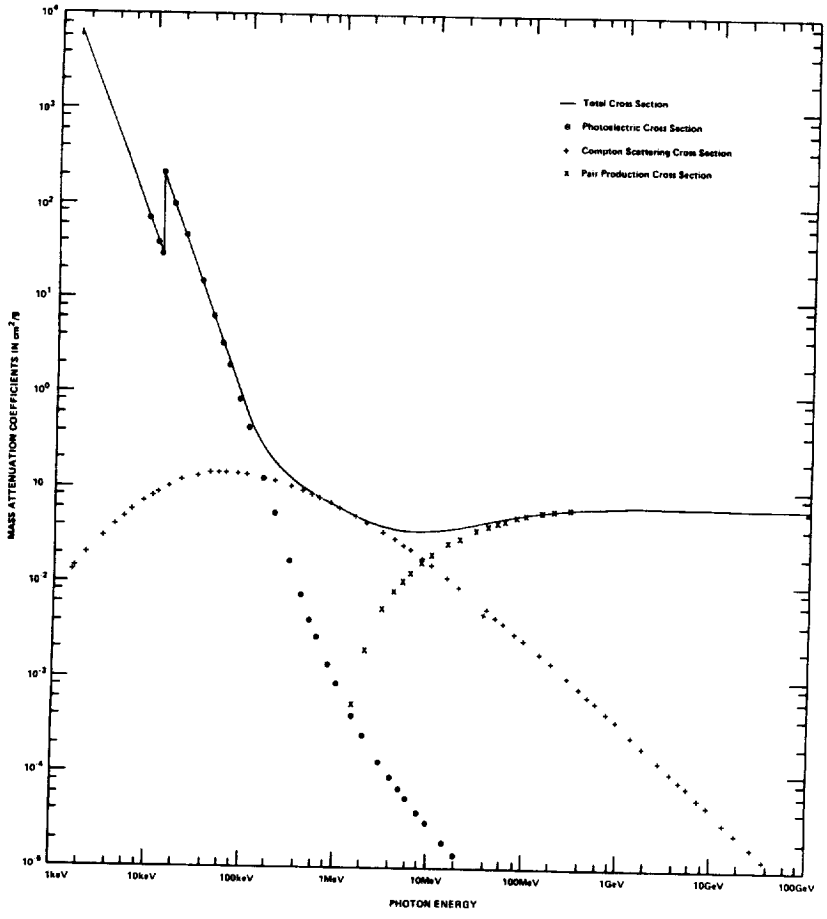

Figure 11.4. The photoelectric, Compton scattering, pair-production, and total mass-absorption coefficient for Ge (from Hubbell, 1969, and 1977, and Hubbell et al., 1980).

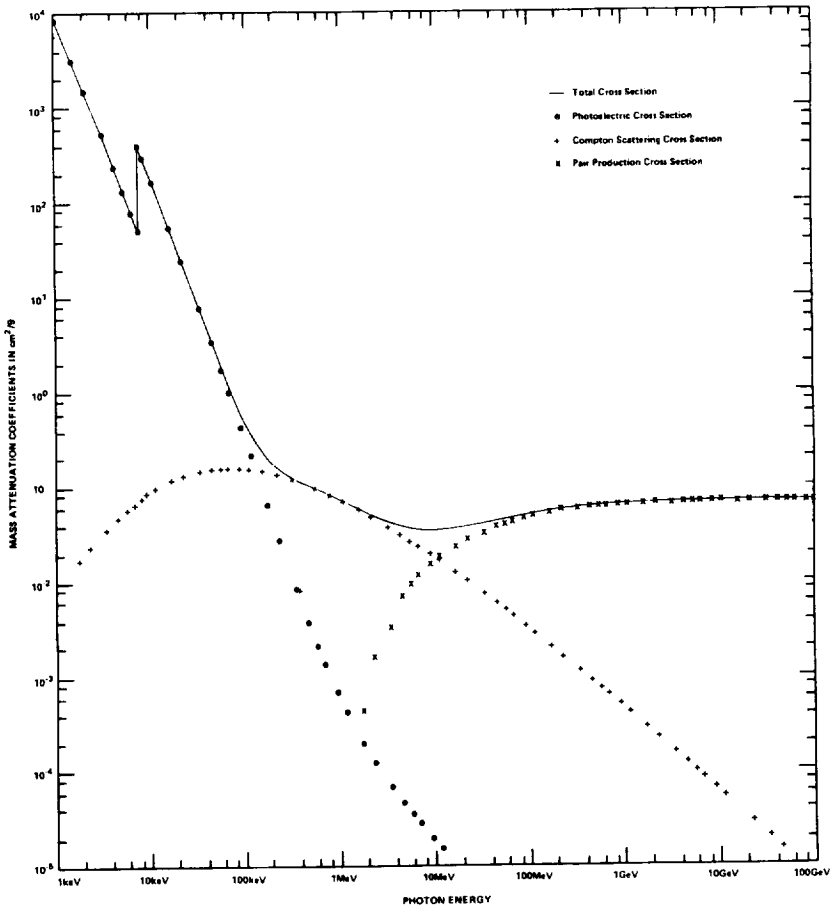

Figure 11.5. The photoelectric, Compton scattering, pair-production, and total mass-absorption coefficient for Fe (from Hubbell, 1969, and 1977, and Hubbell et al., 1980).

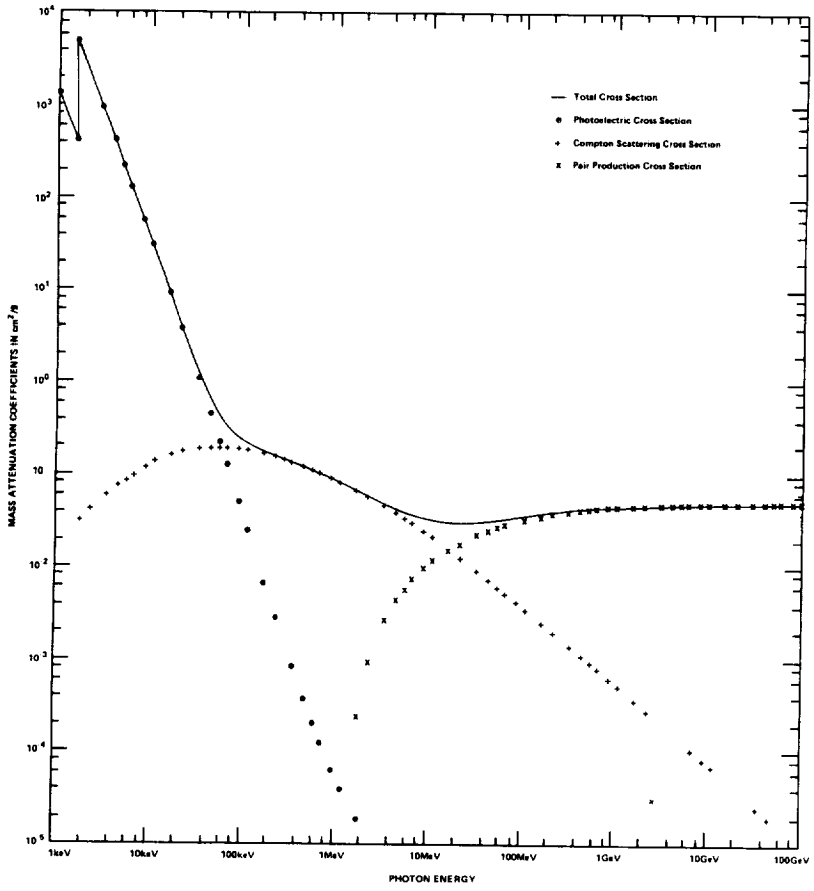

Figure 11.6. The photoelectric, Compton scattering, pair-production, and total mass-absorption coefficient for Al (from Hubbell, 1969, and 1977, and Hubbell et al., 1980).

charge on the nucleus can provide this field. Where the pair production involves an electron, the energy of recoil is quite large, and the threshold for such a process is, in fact, doubled. The amount of pair production caused by the atomic electrons goes as Z; whereas, in the energy range of interest, the pair production in the field of the nucleus goes as Z^2. Since the pair-production cross section is proportional to the square of the particle's charge, for the detectors under consideration most interactions are with the nucleus rather than with one of the atomic electrons. For high-energy photons, a good approximation is simply to add Z to Z^2 in the cross section to include the effect of the atomic electrons. At low energies, near the threshold for this process, the approximation is not valid, because the threshold for pair production in the field of an electron occurs at a higher energy.

The probability per radiation length for pair production is shown in Chapter 13, Figure 13.3. As was mentioned in this chapter for the case of radiation processes, the probability of interaction per radiation length depends only very slightly on the atomic number, or Z, particularly at high energies. As Figure 12.4 shows, the interaction probability rises quickly with energy above a few MeV and then levels to a plateau value of 7/9 at very high energies. The energy at which this process dominates over the Compton scattering process, to be discussed in the next section, depends on the material in which the interaction occurs (see also Figures 11.2 through 11.6). This is an important feature to be kept in mind in the design of a detector.

Since most of the photon's energy goes into the electrons, it is readily understandable that the electrons have a fairly good memory of the arrival direction of the parent gamma ray. The root mean square angle between the trajectory of a secondary electron of energy, E_s, and that of the primary photon of energy, E_p, is given by the equation

$$\left\langle E_p^2 \right\rangle^{1/2} = q\left(E_p, E_s, Z\right)\frac{mc^2}{E_p}\ln\left(\frac{E_p}{mc^2}\right), \qquad (11\text{-}6)$$

where q varies between about 0.6 and 3.5 as $(E_s + mc^2)/E_p$ varies between 0.9 and 0.1 (Stearns, 1949). Averaging over energy, this angle is found to vary from 4° at 30 MeV to 1.5° at 10^2 MeV to 0.2°

at 10^3 MeV, showing the value of higher-energy gamma rays for source location. Unfortunately, the higher-energy gamma rays are much less numerous, as we have seen in the previous chapters.

It is, of course, in general not possible to measure the electrons at the instant they interact or without disturbing them. Hence, it is important to consider what happens to the secondary electrons. An electron undergoes a large number of Coulomb collisions as it passes through matter, most of which produce very small angular deflections. If the layer is very thin, so that energy loss is negligible, and if there is not a very improbable bremsstrahlung interaction in which a significant amount of energy is imparted to a secondary photon, the root mean square angle of scattering is given to a good approximation by

$$\langle\theta\rangle = \left(\frac{E_0}{\beta cp}\right)^2 \frac{\chi}{\chi_0} \quad , \tag{11-7}$$

where χ_0 is the radiation length, χ is the distance traversed in radiation lengths, and E_0 is approximately 21 MeV (Rossi, 1956). In any particle detector, the scattering that occurs before the electron trajectories can be determined increases the uncertainty in the arrival direction of the gamma ray.

In attempting to measure the total energy of the photon by stopping the electron in an absorber, such as a scintillator of some type, it is important to remember that not only the electron and positron must be absorbed. In addition, if an accurate energy estimate is to be obtained, the energy lost by the secondary bremsstrahlung radiation and the energy lost by the two 0.511 MeV photons, formed after the annihilation of the positron, must be taken into account. An alternative method for estimating the energy is to measure the average Coulomb scattering angle of the electrons.

11.4 Compton Scattering

In the Compton scattering process, the electron is treated as unbound or free, and must conserve energy and momentum. The gamma ray can be scattered through an angle with a diminution of energy, and all the energy lost in scattering will be given up to secondary electrons as kinetic energy. Furthermore, the gamma ray

may suffer one or a number of Compton scatterings. As the energy is degraded, the probability of photoelectric absorption increases. Thus, the gamma ray may lose only part of its energy (through Compton scattering only), or it may lose its total energy (through Compton scattering followed by photoelectric absorption) to the crystal.

Since the scattering electrons are taken as free, without interactions among themselves, the effect is strictly additive. The characteristics of the phenomenon can therefore be discussed for a single electron and the cross sections multiplied by Z to obtain the atomic cross section.

The formulas describing the Compton scattering effects take particularly simple form if the photon wavelength, λ, and the energy, E, are expressed in units of the Compton wavelengths, $h/mc = 0.02426$ Å, and the electron rest mass energy, $mc^2 = 0.511$ MeV, respectively. In these units the relation between the change in photon wavelengths and the angle of scattering is simply

$$\lambda' - \lambda = 1 - \cos\theta, \tag{11-8}$$

where λ' is the final wavelength, λ is the initial wavelength, and θ is the angle of scattering. Expressed in terms of energy, Equation (11-8) appears as

$$E_p' = \frac{E_p}{1 + E_p(1 - \cos\theta)}. \tag{11-9}$$

The maximum energy loss occurs for backward scattering (i.e., $\theta = 180°$), at which the scattered energy is

$$E_p' = \frac{E_p}{1 + 2E_p}. \tag{11-9a}$$

For $E >> 1$ (i.e., >> 0.511 MeV) the final energy is never far from 1/2 (i.e., 0.255 MeV), independent of the initial energy. Also, it

should be noted that the photon can never lose all of its energy in any single Compton interaction.

In discussing the total cross section for the Compton effect, it is convenient to introduce the Thomson unit:

$$1 \text{ Thomson unit } = \frac{8\pi}{3}\left(\frac{e^2}{mc^2}\right) = 0.665 \text{ barns}. \qquad (11\text{-}10)$$

In terms of this unit, the total cross section per electron is

$$\sigma_e(E) = \frac{3}{4}\left\{\frac{1+E}{E^2}\left[\frac{2E(1+E)}{1+2E} - \ln(1+2E)\right]\right. \\ \left. + \frac{1}{2E}\ln(1+2E) - \left(\frac{1+3E}{(1+2E)^2}\right)\right. . \qquad (11\text{-}11)$$

For small electron energies, $E << 1$, e is equal to unity in terms of the Thomson unit. In the limit of large energy, $E >> 1$, the formula simplifies to

$$\sigma_e(E) \sim \frac{3}{8E}\left(\ln 2E + \frac{1}{2}\right) E >> 1. \qquad (11\text{-}12)$$

The total Compton scattering cross section thus decreases relatively slowly with increasing energy. Figures 11.1 through 11.6 show the energy-dependent cross sections for a number of elements and materials.

The Compton effect as shown above is independent of Z and is calculated per electron. Thus, when the mass absorption coefficient is considered, this cross section is nearly constant and is almost independent of the Z of the material in which the gamma rays interact.

Another important cross section of interest in detector design is the differential angular cross section, $\sigma(\theta)$, the cross section for scattering into a given angle per unit solid angle. This cross section is given by the Klein-Nishina formula, using Thomson units as

$$\sigma(\theta)d\Omega = \frac{3}{16\pi}\frac{E_p^2}{E_p^2}\left(\frac{E_p}{E_p'} + \frac{E_p'}{E_p} - \sin^2\theta\right)d\Omega, \qquad (11\text{-}13)$$

where d Ω is the differential of solid angle, and remembering that θ and E are related by Equation (11-9).

For very low values of E, that is, as E approaches 0, the Compton scattering cross section is symmetric and varies by a factor of 2 with angle, and has a maximum value at 0 and 180° and a minimum at 90°. As the energy increases, the cross section becomes peaked in the forward direction. By 10 MeV, the cross section falls to half of its initial value at about 12° and drops to about 1 percent of the initial value for 180° or backscattering.

11.5 Electron Leakage and Bremsstrahlung

These effects will be only briefly considered in terms of their effect on energy loss.

Electron leakage or escape complicates our understanding of the amount of energy deposited by the gamma ray in the detector. If enough energy is imparted to secondary electrons so that these electrons can escape the detector mass without losing all their kinetic energy, then the resolution of the detector will be affected and the knowledge of the incident gamma-ray energy will be degraded .

Loss of electron energy by the bremsstrahlung processes can also degrade the knowledge of the incident energy, unless the photoradiation produced is absorbed by the detector mass and produces a measurable energy loss in the detector. These processes will be considered in more detail later when detector response functions are considered.

References

Bethe, H. A., and Askin, J., 1953, in Segre, ed., Experimental Nuclear Physics, I (New York: John Wiley and Sons, Inc., 1953), p. 305-349.

Davisson, C. M., 1968, in Siegbahn, ed., Alpha-, Beta- and Gamma-Ray Spectroscopy, vol. 1, Chapter 2 (Amsterdam: North-Holland Publishing Company, 1968), p. 37-78.

Goldstein, H., 1959, Fundamental Aspects of Reactor Shielding (New York: Addison-Wesley Publishing Co., Inc., 1959).
Heitler, W., 1954, Quantum Theory of Radiation, 3rd ed. (New York: Oxford University Press, 1954).
Hubbell, J. H., 1969, NSRDS-NBS29 (Washington D.C.: U.S. Government Printing Office, 1969).
Hubbell, J. H., 1977, Radiation Research, 70, 58.
Hubbell, J. H., Gimm, H. A., and Overbo, I., 1980, Jour. of Phys. Chem. Ref Data, 9, 1023.
Rossi, B., 1956, High Energy Particles (Englewood Cliffs, N.J.: Prentice Hall, Inc., 1956), p. 10-90.
Stearns, M., 1949, Phys. Rev., 76, 836.

CHAPTER 12

Detectors for Energies Less Than 25 MeV

12.1 Introduction

In the energy domain 100 keV to 10 MeV, both crystal scintillators and semiconductors are widely used for gamma-ray detectors in spectrometer systems. These detectors' operation depends on the fact that gamma rays lose energy by ionization in these materials and electrons and holes are produced. In the case of semiconductors these electrons and holes are collected by an electric field, and they provide an electric signal that is a direct measure of the energy lost by the gamma ray in the material. Scintillation detectors depend on a further conversion of the energy lost in the electron-hole pair to the production of photons. A photomultiplier tube measures the intensity of the photon flux and an electrical pulse, proportional to the photon intensity, is produced at the photomultiplier output. The electrical pulses appearing at the solid state or scintillation detector output can be analyzed as the number of events in terms of the magnitude of the electrical output either as a current or a voltage pulse. The resulting spectrum, called a pulse-height spectrum, is a measure of the energy-loss spectrum in the detector material.

An important characteristic of any such system is that the number of carriers produced and, in the case of scintillation detectors, the number of photons produced is linearly related to the gamma-ray energy lost to the detector. The best material requires only a small amount of energy to produce an electron-hole pair (thus a large number of carriers are produced for a given energy absorption). The statistical fluctuations are small when expressed as a percentage of the total number of electron-hole pairs produced. This fact significantly affects the energy resolution of the spectrometer system.

Gamma-ray energy can be lost to the detector through a variety of processes, such as photoelectric effect, Compton scattering, and pair production. These processes involve both total and partial energy loss. Details of the interaction processes in various types of radiation detectors can be found, for example, in Knoll (1989). High Z materials are used to optimize the absorption process, and anticoincidence systems can reduce the contribution of the partial-absorption events to degrading the shape of the pulse-height spectrum with respect to the incident flux. The measured pulse-height spectrum is further smeared out by the finite resolution of a real gamma-ray spectrometer system and by a number of background and radiation damage effects. The incident-photon spectrum can be used to infer the astrophysical properties of interest in gamma-ray astronomy. Thus, an understanding of the detector interaction processes, of resolution as a function of energy, and of radiation damage and background contributions is essential to infer the true incident-photon spectrum from a measurement of the energy-loss or pulse-height spectrum.

12.2 Properties of Detectors

Important properties of semiconductor and scintillation materials used for the detection of gamma rays are shown in Table 12.1. These properties are used to help select the proper detector, consistent with the requirements imposed by the scientific and space-mission constraints.

Table 12.1

Properties of Semiconductor and Scintillation Materials used for Gamma-Ray Detectors

Properties of Semiconductors.

	Si	Ge	CdTe	CdZnTe	HgI_2
Z	14	32	48, 52	48, 30, 52	80, 53
Density (g/cm^3)	2.33	5.32	6.06	5.8	6.3
Band gap (eV)	1.12	0.74	1.47	>1.47	2.13
Energy per electron-hole (eV)	3.61	2.98	4.43	>4.43	4.22

Properties of Common Scintillators

	CsI(Tl)	CsI(Na)	NaI(Tl)	BGO
Density g/cm^3)	4.51	4.51	3.67	7.13
lmax$^{(nm)}$	540	420	415	505
Index of Refraction	1.80	1.84	1.85	2.15
Decay Constraint (ns)	1000	630	230	300
photons/MeV	64,800	39,000	38,000	8200
Relative Pulse Height*	0.49	1.11	1.00	0.23
Hydroscopic	slightly	yes	yes	no

*relative to NaI

NaI(Tl) Crystal Scintillators

A number of scintillators are of interest in gamma-ray spectroscopy, e.g., NaI(Tl), CsI(Tl), CsI(Na), and BGO, but for the purpose of this discussion we will consider the NaI(Tl) scintillation detector. In Chapter 11 we considered the various mechanisms by which gamma rays interact with matter to produce high-energy electrons. The photoelectric effect, Compton scattering, and pair production are the dominant interaction processes. By any of these processes, or combinations thereof, the gamma ray loses all or part of its energy to the crystal. After having converted this energy to high-energy electrons, the crystal experiences ionization caused by the energy loss dE/dx of the electron in moving through the crystal. This process is similar to the ionization of a gas by a charged particle. In the case of the crystal, however, one deals with a solid insulator in which the band theory is applicable (Sietz, 1940; Kittel, 1953). Band-theory values of energy for bound electrons in a perfect NaI crystal belong to allowed intervals of energy (the valence and conduction bands) which are separated by "unallowed" intervals (the forbidden bands). Passage of a high-energy electron through the crystal excites electron-hole formation in the valence band. The energy gap E_g (i.e., the interval between the valence and conduction band) for NaI is about 30 eV.

The NaI(Tl) crystal is not perfect, and, because of the presence of thallous ions and lattice defects, localized allowed electron levels exist in the normally forbidden interval between the conduction band and the valence band. Some of these levels are in regions called "trapping centers," which are largely caused by lattice imperfections, but partly by the presence of impurity ions. Still other allowed levels that exist between the bands are called "luminescence-center

levels." The mechanisms of light emission (scintillations) are attributed to the existence of these centers. An energy-level diagram is shown in Figure 12.1 (Mott and Sutton, 1958; Price, 1958).

It is believed that the luminescence centers consist of pairs of thallous ions that have some of the properties of ordinary diatomic molecules. The energy, E, of those electrons within the effective radius of a pair of thallous ions as a function of the interior separation, r, is a function of the type shown in Figure 12.1 above the words "luminescence-center levels." The lower curve, A, is for the electronic ground state, and the upper curve, B, is for an excited electronic state. A and B have shapes that allow the thallous ion to vibrate along r. Some of the vibrational levels are excited at room temperatures. By the Frank-Condon principle (Mott, 1952), changes in electronic energy are represented by vertical lines at the

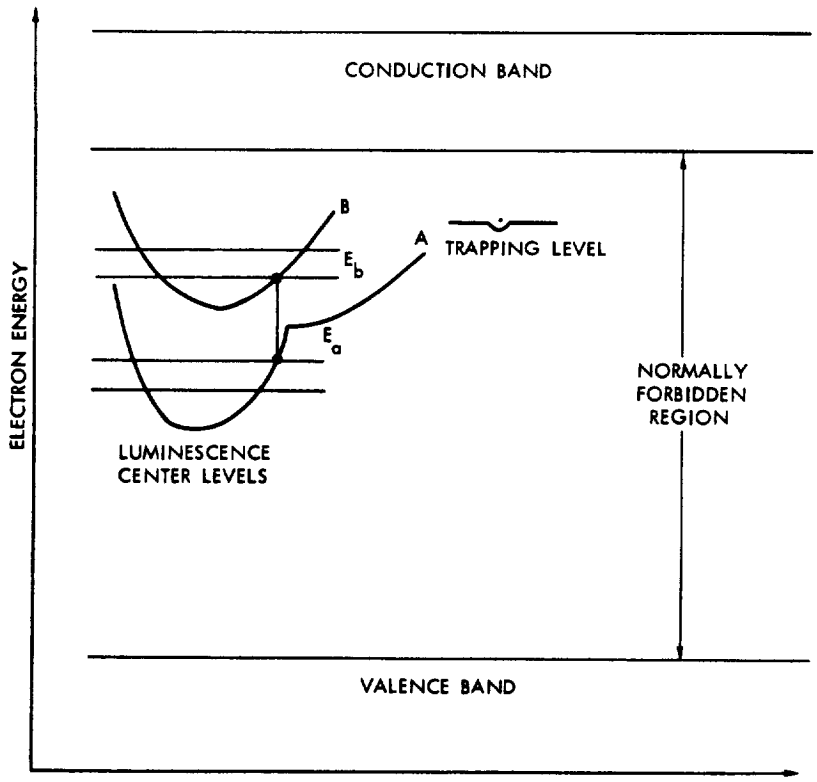

Figure 12.1. Energy-level scheme for NaI(Tl) detector.

extremes of the ion vibration locus on the energy diagrams, whereas in the classical picture the ions are instantaneously stationary. In the transition of the electron from E_b to E_a a photon of energy E_b to E_a is emitted. Since transition may occur at the extremes of these ion vibrations, one obtains photons whose wavelengths lie in a rather broad band with energies corresponding to 3 eV or 4100 Å (Mott and Sutton, 1958).

Electrons excited to the conduction band can fall to the valence band through the trapping level and the luminescence-center levels. Transitions via the luminescence-center levels result in the emission of light: the scintillations of interest in the detection process. Transitions via other levels do not produce detectable scintillations. NaI(Tl) has the property that the intensity of scintillations or total light energy produced is proportional to the energy lost by the gamma rays or fast particles that produced it.

This is the property of NaI(Tl) that makes it useful in gamma-ray spectrometry. Figure 12.2 shows the NaI(Tl) emission spectrum and the absolute efficiency of the conversion of electron excitation energy into emitted light (Bird, 1956). The total light energy produced, ε_l, given an initial electron excitation energy E_e, is

$$\varepsilon_l = E_e \int_0^\infty C_{np}(\lambda)\, d\lambda, \qquad (12\text{-}1)$$

where $C_{np}(\lambda)$ is the conversion efficiency as a function of the wavelength of the photons produced in the scintillation process. $C_{np}(\lambda)$ does not depend upon the initial electron excitation energy (Bird, 1956; Price, 1958). Thus, there is a linear relation between ε_l, and E_e.

Other NaI(Tl) properties of interest in detector design (Heath et al., 1965; Knoll, 1989) are its transparency to its own luminescence, its high light yield relative to other scintillators, the 0.25 s decay constant of its scintillation, and the fact that it is very hygroscopic.

Photomultiplier Tube

The intensity of the scintillations induced in the NaI(Tl) crystal by the absorption and scattering of gamma rays is measured very accurately. The light from the crystal passes through the glass

envelope of the photomultiplier tube and ejects electrons from the photocathode. The photosensitive materials used in the photocathode (e.g., Cs Sb) are chosen because of their good photoelectric yield at the wavelength of the luminescent emission from a scintillation crystal. The photocathode absorption spectrum is shown in Figure 12.2. The photoelectrons are electrically accelerated to the first dynode, where they eject secondary electrons. The process is cascaded in ten stages, so that the charge of the cascade ejected from the tenth dynode and collected on the anode is about 6 x 10^5 times the original charge from the cathode, where the potential difference per stage is 105 V (Mott and Sutton, 1958).

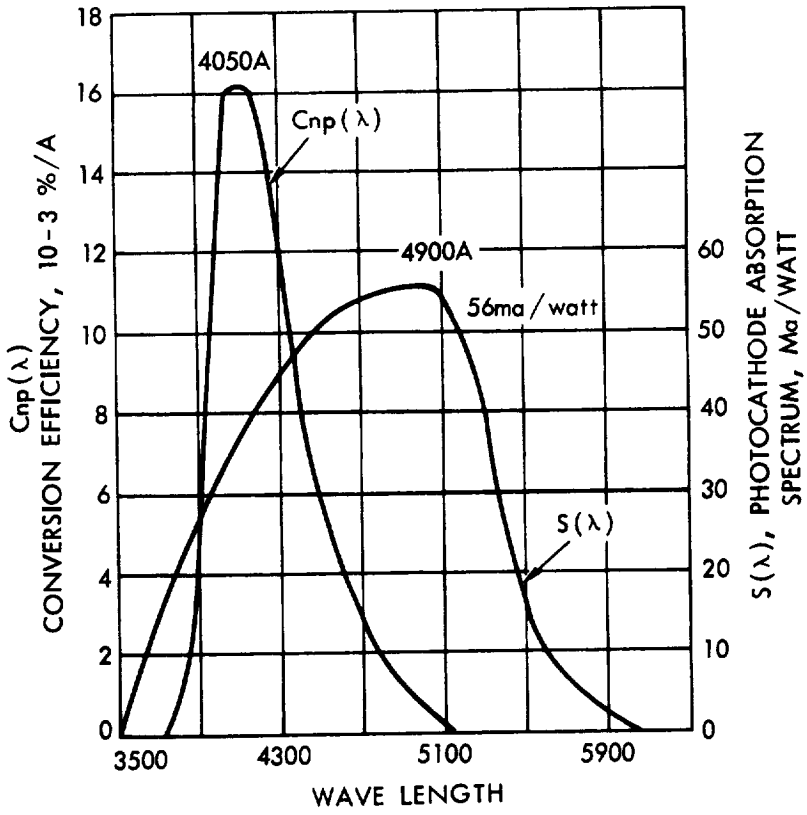

Figure 12.2. Emission spectrum, absolute conversions effieciency spectrum, and photocathode "absorption" spectrum for NaI(Tl) crystal and end window (Sb Cs_3) photocathode multiplier tubes.

The fraction of the photon energy which strikes the photocathode surface is ε_1 (T_p,F_p), where T_p is the transparency both of the crystal and of the phototube optical seal, and F_p is the nonescape probability. For a well-reflected crystal, both T_p and F_p should be nearly one. A fraction of this energy will be absorbed in the photocathode, resulting in the release of a number, n_e, of low-energy photoelectrons inside the phototube. The efficiency of this conversion from photons to electrons depends on the absorption spectrum of Sb Cs_3 and on the electronic stopping power of the photocathode; it is measured by the sensitivity factor $S(\lambda)$, shown also in Figure 12.2.

Now, $dn_e/d\varepsilon = S(\lambda)$ can be written $dn_e = S(\lambda)\, d\varepsilon$, but from Equation (12-1) above

$$d\varepsilon = d\varepsilon_1 = E_e\, C_{np}\, (\lambda)\, d\lambda$$

and

$$dn_e = E_e\, C_{np}\, (\lambda)\, S(\lambda)\, d\lambda,$$

or the number of photoelectrons ejected, n_e, depends on the product of the spectra of Figure 12.2

$$n_e = E_e \int_0^\infty C_{np}(\lambda)\, S(\lambda) d\lambda\,. \qquad (12\text{-}2)$$

If F_c is defined as the fraction of such electrons striking the first dynode and if the subsequent phototube multiplication is M, the total charge q produced at the anode as a result of a single original gamma-ray interaction is

$$q = n_e\, eMF_c = V_0 / C, \qquad (12\text{-}3)$$

where e is the electronic charge and q is the charge collected on the input capacitance C of an amplifier and measured, and a voltage pulse V_0. Thus, it is seen that the pulse height is proportional to the initial electron excitation energy E_e. However, because of the statistical fluctuation in dE/dx, n_e M and F_c, a distribution of voltage pulse heights about some mean value V_0 will be observed, rather than a constant pulse height, as implied by Equation (12-3).

Factors Affecting Resolution. The gamma ray dissipates its energy in the phosphor, thus producing high-energy photoelectrons, which produce pulses with varying magnitudes at the output of a scintillation counter. A number of experimental and theoretical studies (e.g., Wright, 1954; Kelley et al., 1956) have been made of the widths of the generated pulse-height spectra.

Emission of Photons by the Phosphor. There is a statistical fluctuation in the number of photons per scintillation. There are other statistical variations that may be attributed to local variations in luminescent efficiency of the phosphor, caused by possible nonuniform distribution of activity ions by the fact that successive particles lose different amounts of energy to the phosphor through interaction edge and scattering effects, and also by the luminescence process itself (i.e., the ratio of absorbed energy to photon energy fluctuates for successive particles).

Collection of Emitted Photons by the Photocathode. Successive scintillations never occur at exactly the same position in the crystal; thus, the photon-collection efficiency of the photocathode depends on the position where the scintillation is produced. Optical flaws in the crystal and at the various optical seals further spread the distribution.

Emission of Photoelectrons n_e by the Photocathode. There is a statistical fluctuation of the number of photoelectrons released from the photocathode per scintillation. Further, there is a point-to-point variation of photocathode response; also, a random emission of thermal electrons by the photocathode adds to a variation in the pulse-height distribution.

Collection of Photoelectrons by the First Dynode F_e and Multiplication M by the Successive Stages. The variance due to the multiplication process can be shown to be fundamentally statistical in nature (Gamble, 1955). In addition, losses can be attributed to the variations in the fraction of photoelectrons collected by the first dynode, in the collection efficiencies of subsequent dynodes, and in dynode response.

Other processes affecting the pulse-height resolution of scintillation detectors have also been considered, such as the statistical variation in dE/dx; they have been found to be negligible compared with those mentioned above. For further details see Swank and Buch (1952), Wright (1954), and Mott and Sutton (1958). It is also shown in these references that the effects considered in the first two processes are small compared with those of the third and fourth processes. Therefore, the theoretical calculation of energy resolution usually considers only the statistical fluctuation in the third and fourth processes (Swank and Buch, 1952; Gamble, 1955; Kelley et al., 1956; Mott and Sutton, 1958). The statistical variation in the number of photoelectrons n_e produced at the photocathode can be considered to be Gaussian. If $V = V_0$ is the mean value of the pulse height produced by the absorption of the gamma-ray energy E_0 in the phosphor of the anode, then the probability P_1 that a pulse of height V_1 is produced at the phototube anode because of statistical fluctuation in n_e is

$$P_1 = \frac{n_e(V_0)}{n_e(V_1)} = \exp - \left[\frac{(V_1 - V_0)^2}{2\sigma_n^2} \right], \qquad (12\text{-}4)$$

where σ_n^2 is the photoelectron dispersion. Here we have only considered the variance in n_e.

Now consider the effect of the variance in the multiplication M. A Gaussian distribution of statistical fluctuation is assumed (Gamble, 1955; Kelley et al., 1956). Therefore, the probability that the phototube multiplication of $n_e(V_1)$ produces a pulse corresponding to a height of V_2 is

$$P_2 = \frac{M(V_2)}{M(V_1)} = \exp - \left[\frac{(V_2 - V_1)^2}{2\sigma_m^2} \right], \qquad (12\text{-}4a)$$

where σ_m^2 is the dispersion caused by the statistical variation in the multiplication. The probability P_3 of obtaining a pulse of height V_2 from the complete absorption of energy E_0 is the product of P_1 and P_2 integrated over all V_1

$$P_2 = \int_{-0}^{+\infty} \exp -\left[\frac{-(V_1 - V_0)^2}{2\sigma_n^2}\right] \exp\left[\frac{-(V_2 - V_1)^2}{2\sigma_m^2}\right] dV_1. \quad (12\text{-}4b)$$

Letting $\alpha = (V_2 - V_0)$ and $\sigma^2 = \sigma_n^2 + \sigma_m^2$, P_3 reduces to

$$P_3 = K \exp -\left[\frac{\alpha^2}{2\sigma^2}\right], \quad (12\text{-}5)$$

where σ^2 is the dispersion of the photoelectron rate and subsequent multiplication rate, and K is a constant.

Analysis (Swank and Buch, 1952; Gamble, 1955) has shown that

$$\sigma_n^2 \sim \bar{n}_e = K_1 E_0 = K_1' V_0,$$

and

$$\sigma_m^2 \sim \bar{n}_c f(M) = K_2 E_0 = K_2' V_0,$$

where f(M) is a function of the multiplication, and is constant for a fixed M. Also,

$$\sigma^2 = K_3 E_0 = K_3' V_0. \quad (12\text{-}6)$$

It has been found experimentally that Equation (12-6) is true for limited energy ranges (Kelley et al., 1956).

The Gaussian form given in Equation (12-5) describes the pulse-height distribution generally; however, one cannot use Equation (12-6) to determine σ^2 in general. It has been found that over a limited energy range the relation

$$\sigma = K E_0^n = K' V_0^n, \quad (12\text{-}7)$$

where n is determined experimentally, and can be used more generally.

The detector resolution can be defined as the ratio of $V_{1/2}$, the full width at half maximum (FWHM) in pulse-height units, or $E_{1/2}$ the FWHM in energy units to the position V_0 or E_0. That is,

$$R = \frac{\Delta V_{1/2}}{V_0} = \frac{\Delta E_{1/2}}{E_0} \tag{12-8}$$

and $\Delta V_{1/2} = 2(\ln 2)^{1/2}\,\sigma_v$ (in pulse units), and similarly $\Delta E_{1/2} = 2(\ln 2)^{1/2}\,\sigma_e$ (in energy units). The standard measure of the operation of a scintillation detector is the resolution obtained for the photopeak or total absorption of the 0.662 MeV gamma-ray line (the monoenergetic gamma-ray line emitted by ^{137}Cs). An extremely good detector would have a resolution of about 7 percent. This would yield a $\Delta E_{1/2}$ of 46 keV.

Details of the properties of other scintillators of interest in gamma-ray spectroscopy can be found in Neiler and Bell (1968), in Adams and Dams (1970), and in Knoll (1989).

Low Temperature Semiconductor Detectors

Low temperature semiconductor gamma-ray detectors materials such as Si(Li), Ge(Li), and high-purity Ge [Ge(HP)] are extremely attractive for use in x-ray and gamma-ray spectroscopy because of their better energy resolution, when compared with gas proportional counters and scintillation counters in the energy domain of ~6 keV to ~20 MeV. As described above for the case of scintillation detectors, charged particles or gamma rays lose energy to the detector material, and, in doing so, lift electrons from the valence band or deeper conduction band to higher-energy bands. An electron-hole pair is produced. The high-energy electron produced, for example, by photoelectric absorption, Compton scattering, and pair productions rapidly loses energy to other electrons and after about 10^{-12} s a stationary situation is achieved, where the electrons reside in the conduction band. Single gamma-ray interactions are then detected by applying an electric field to the detector and sweeping the electron-hole pair through the material so that an electrical pulse is measured at the detector output. This differs from the scintillation detector technique in that further signal-amplification devices such as photomultiplier tubes are not used.

At this point, one of the reasons for the improved energy resolution can be indicated. About 3 eV is required to produce an electron-hole pair which is finally detected at the output of a Ge semiconductor detector as compared with about 300 eV needed to create one photoelectron at the photocathode in the case of a NaI(Tl) scintillation detector.

Band Theory.

A number of compound semiconductor materials have been studied as possible detectors for x-ray and gamma-ray spectrometers. These include such materials as CdZnTe, GaAs, HgI_2, Ge(Li), Si(Li), and Ge(HP) and Si(HP). In the following discussions we will consider the Ge-type detectors and next the room-temperature solid-state detectors, which require cryogenic cooling.

The same band theory used to describe the properties of scintillators can also be used to describe the properties of semi-conductor detectors. The energy gap separating the valence and conduction bands in Ge is 0.764 keV at absolute zero. In a perfect crystal at absolute zero, the conduction band is vacant and the valence band filled. Electron-hole pairs can be produced as the temperature increases by thermal excitation of electrons, from the top of the valence band to the bottom of the conduction band. The inverse process can also occur, when carriers of the opposite sign collide and annihilate. The creation of one electron-hole pair in the energy-band model corresponds to thermal breaking of one covalent bond between the atoms of the material. The thermally generated electrons and holes give rise to electrical conduction.

The charge carriers are influenced by electric fields present in the lattice, and these fields give a random motion to these charges when an external field is applied. The electrons and holes drift. In pure materials the mobility of the electrons and holes is governed by lattice scattering, and increases rapidly as the temperature is reduced. In pure semiconductor materials, such as Ge, the mobility of both carriers is nearly the same. A maximum mobility for a given material is achieved as a result of the decrease in the mean free path between successive collisions as the average carrier energy is increased by an electric field. The production of electron-hole pairs and their behavior as described above is called intrinsic semiconduction and is characterized by an equilibrium concentration of free electrons and holes in the crystal.

The electrical properties of semiconductors can be modified by the presence of various types of imperfections. These imperfections can be attributed to chemical impurities and structural defects. Chemical impurities in substitutional or interstitial sites introduce transition levels in the forbidden energy band (cf. the consideration of scintillation materials with the introduction of thallous atoms).

Of course, all real crystals will contain various imperfections that change their electrical properties. The semiconductors, in such cases, are said to be extrinsic. Ge is a group-IV element. A diamond lattice characterizes the structural form: each atom of one sublattice is tetrahedrally surrounded by four atoms of the other sublattice, and gives one electron to each of the four atoms to form covalent bonds. If a small concentration of an element from group V is present, for example, this element might replace a Ge atom at a number of sites in the crystal lattice. In this case, four of the valence electrons are used for the covalent bands in the lattice and the fifth electron is loosely bound (a level in the forbidden band) to the impurity atom. This loosely bound electron can easily be thermally excited to the conduction band, and a fixed positive charge results in the lattice. This type of impurity behaves as an electron donor and is denoted as an n-type material. On the other hand, if a group III element is a replacement in the Ge lattice position, it constitutes a fixed negative charge and a positive hole occurs in the valence band. An electron acceptor denoted as a retype material is produced. A group-I type material, such as lithium, can occur interstitially in the semiconductor material and behave as an electron donor.

In all the cases mentioned above, the energy required to ionize these extrinsic materials is less than the energy gap, because of the localized energy levels introduced in the forbidden energy region. Further, in an extrinsic semiconductor the equality does not hold between electron density and hole density. Also, the mobility of electrons and holes in an extrinsic detector differs from that in intrinsic material. This is especially important at low temperatures. In extrinsic materials, the mobilities are controlled by Coulomb attraction at the impurities sites rather than by lattice scattering, as in intrinsic materials.

With the above factors in mind, let us consider these effects with respect to ionization produced by irradiation incident upon such detectors. At a given temperature, the equilibrium distribution is disturbed by the introduction of a number of excess carriers caused

by gamma rays being absorbed in the crystal. In intrinsic semiconductor material, the recombination time is very long (i.e., seconds, because the direct recombination of an electron with positive holes is not a highly probable process). In extrinsic materials the recombination is much shorter (i.e., one millisecond or less). Localized energy levels, caused by impurities or structural defects have significant capture cross sections for the drifting electrons or holes. The electrons or holes can be either lost or re-excited (recombination), or trapped for a time and then restored to the conduction band, if an electron, or to the valence band, if a hole. These properties significantly affect the use of these materials in gamma-ray detectors. Details of the properties discussed above can be found in a number of texts. A rather complete description is given in Bertolini and Coche (1968), and Knoll (1989).

Total Volume Sensitive Detector. The best resolution semiconductor for use in the gamma-ray energy region (about 10 keV to about 20 MeV) is Ge. Such detectors must operate at low temperature (e.g., cooled by liquid nitrogen) to avoid the large leakage current produced at room temperature. Thermal generation of electron-hole pairs at room temperature is large because of the 0.7 eV band gap in this type of material.

Now consider the problem of a volumetric detector with electrodes connected to opposite sides of the detector. The utilization of such contract electrodes may cause large currents to flow through the detector, because there may be a free exchange of holes and electrons between the semiconductor and the electrodes. If this is the case, a current will flow upon the application of high voltage across the detector. It has been found that "blocking contacts" can be used to overcome this problem (Pehl et al., 1968; Pehl and Goulding, 1970; Pehl, 1977). Blocking contacts have the property that the positive contact does not inject holes and the negative contact does not inject electrons. The two approaches that have been used to achieve these properties are as follows: careful choice of metals used for the contacts, or forming contacts with heavily doped n and p regions by either high-temperature diffusion or by ion implantation into the detector surface.

Surface-leakage currents can also cause difficulties in signal detection. This surface-leakage problem can be attributed to, among other things, the absorption of various chemical species on the detector surface, which, by modifying its electrical properties,

results in severe leakage and breakdowns. Thus, during the production of the detector, proper chemical treatment of the surface is required and the surface must be properly proteced during operation.

Detector Resolution. Gamma-ray interactions in the sensitive volume of the detector create highly energetic electrons which can lose energy by the creation of electron-hole pairs or by interactions with the crystal lattice. Considering these loss processes, k is defined as the mean energy per ion pair (about 3 eV for Ge); then the mean number, N, of electron-hole pairs produced is

$$\overline{N} = \left(\frac{E}{k}\right), \tag{12-9}$$

where E is the energy absorbed. There is a statistical function in the production of the electron-hole pairs given by the standard deviation σ_n:

$$\sigma_n = \left(\overline{N}\right)^{1/2} = \left(\frac{E}{k}\right)^{1/2}, \tag{12-10}$$

The actual experimentally determined σ_n has been found to be smaller than that shown in Equation (12-10). A Fano factor, F, similar to an equivalent situation in gaseous detectors, has been used to explain the difference (Pehl et al., 1968) when calculating the variance σ_n,

$$\sigma_n^2 = F\overline{N} = F\left(\frac{E}{k}\right). \tag{12-11}$$

One can consider the Fano factor, F, to be approximately equal to the ratio of the energy imparted to the detector lattice by non-electron-hole production processes, to the total energy imparted to the crystal by the slowing down of the gamma-ray-produced high-energy electrons. If almost all the energy loss produced electron-hole pairs, F would approach zero and there would be "no" statistical variance, while, if the probability of electron-hole creation was very low, F would approach unity. For large germanium detectors the Fano factor seems to be near 0.13.

The measured detector resolution can be understood in terms of the statistical variance related to electron-hole pair production σ_n^2 and the increase in the electronic noise produced variance σ_e^2. The total variance σ_T^2 is then

$$\sigma_T^2 = \sigma_e^2 + \sigma_n^2 = \sigma_e^2 + F\left(\frac{E}{k}\right). \qquad (12\text{-}12)$$

For a given spectrometer/detector system, σ_e^2 should be constant. For scintillation detectors σ_e^2 is small compared to σ_n^2 and was therefore ignored; for semiconductor devices such as Ge detectors these factors are comparable and, thus, σ_e^2 cannot be ignored. With the formulation shown in Equations (12-7) and (12-8), the resolution becomes

$$R = \frac{\Delta E_{1/2}}{E} = \frac{2\sqrt{\ln 2}\ \sigma_T}{E} = \frac{2\sqrt{\ln 2}}{E}\sqrt{\sigma_e^2 + \frac{FE}{k}} = \sqrt{\frac{A}{E^2} + \frac{B}{E}}, \qquad (12\text{-}13)$$

where A and B are constants for a given detector-spectrometer system. The resolution of the 1.33 MeV (^{60}Co) line with a Ge detector of volume about 120 cm^3 can be about 0.2 percent as compared with a good NaI(Tl) detector with a resolution of about 4.7 percent.

Ge(Li) and Ge(HP) Detectors. A major problem with Ge semiconductors is the presence of the trapping centers discussed earlier. A carrier "trapped" by one of these centers can be released or "detrapped" by thermal excitation. The "detrapping time" is defined as the average time a carrier remains trapped. Trapped carriers affect energy resolution. These carriers can be regained, in a sense, if the detrapping time in the detector is very short compared with the pulse-collection time of the external electronics. Because the location of the initial carrier charge production changes and because of the variations in the concentration of traps at different points in the detector, the loss of signal can be different from one event to another, and, thus, there is a further fluctuation in the output signal.

Early in the development of Ge detectors, the high-impurity concentration in the available Ge required that "Lithium drifting" be used to reduce the problems introduced by the presence of acceptor impurities. Lithium acts as an interstitial donor. In this way, Ge detectors could be produced with rather large volumes of material exhibiting low net impurity properties. This process required that the detector be kept constantly at cryogenic temperature: at elevated temperatures the lithium becomes mobile and the impurity properties are destroyed. The lithium-drifted detectors are made by evaporating lithium onto a p-type germanium block and drifting the lithium into the block at higher temperatures. Ge lithium-drifted (Ge(Li)) detectors with large volumes have been produced by means of coaxial drift techniques. The lithium is drifted from the outer surface of a cylindrical block of Ge toward the center, and the drift process is stopped when a small core of p-type Ge remains. This core is the p-contact for the detector.

Higher-purity germanium Ge(HP) detectors are now available. These detectors do not require lithium compensation; the low-impurity concentration does not significantly affect the resolution. These Ge properties are still extrinsic in nature and both p-type and n-type material can be produced. This property is extremely important in radiation-damage considerations. These detectors will not be ruined by elevated temperatures, although they must be operated at low temperatures.

Both planar and coaxial geometries can be used for Ge(HP) and Ge(Li) detectors. However, in terms of the coaxial Ge(HP) detector, the equivalent of the undrifted central core as in Ge(Li) material does not exist. Consequently, a central core of material must be removed and a contact put on the inside surface of the Ge. Detectors of both the n-type and p-type have been manufactured in this coaxial configuration.

Cryogenic Coolers

In Table 12.1 it was seen that the bandgap in Ge is narrower than in the other semiconductor materials. Because of this narrow bandgap, Ge must be operated at cryogenic temperatures, while the other materials can be operated at or near room temperature. This is a major drawback to high-resolution Ge detectors requiring cryogenic temperatures (usually 130 K or less) for operation. Thus, cryogenic cooling systems that can be used in the space environment are

required. Three methods of cooling have been employed: (1) liquid or solid cryogens, (2) mechanical coolers, and (3) passive-radiative cooling systems.

Liquid and solid cryogens have been used on a number of high-altitude balloon flights (Crannell et al. (1984); Teegarden et al. (1985); Rester et al. (1986); Sandie et al. (1988); Mahoney et al. (1980), and on HEAO-3, the third High Energy Astronomy Observatory (Mahoney et al., 1980). For cooling over long time periods large masses of cryogens are needed. The weight and volume requirements for these cryogens could not be accommodated on long-duration space-flight missions.

Passive radiation coolers have been used widely for space-borne instruments that require cryogenic temperatures in the 90-120 K range (Sherman, 1983). The first radiative cooler used to cool a Ge detector in space was flown successfully on the ISEE-3 Mission (Teegarden and Cline, 1980). A "V-groove" passive cooler was developed for the Mars Observer mission (Boynton et al., 1993) and flown successfully. The system operated for eleven months during the cruise phase of the mission. A passive Ge cooling system is in operation on the WIND mission (Owens et al., 1995) and has operated for over a year with excellent results. In terms of cooling on planetary missions, passive systems consume the least power, but require orbits that allow the cooler to point away from the Sun. For Mars this can be achieved in sun-synchronous orbits, but for the Moon the problem of passive cooling is more complicated, since there are no sun-synchronous orbits. For much of the time on some extended planetary missions during the cruise phase and for limited times during the orbital phases, the detector system cannot be cooled. This aspect of operation greatly limits the ability to check the health of the experiment and the calibration during the cruise phase, and may limit the work cycle during orbital operation. Furthermore, leaving the Ge detectors at higher temperatures may exacerbate the radiation-damage problem.

Mechanical coolers can potentially relieve the problem, but current units designed for space are expensive, have large power and volume requirements, and thus are not applicable to small type missions such as Discovery. However, mechanical coolers are being developed for flight in the early years of the 21st century for the European Space Agency (ESA) and the International Gamma-

Ray Astrophysical Laboratory (INTEGRAL) mission (Winkler, 1994).

Room-Temperature Solid-State Detectors

Wide-bandgap, semiconductor materials now under development can potentially overcome the major limitations inherent in cryogenic-based spectrometer systems. Because thermally-generated leakage currents are minimized by the gamma-ray detectors fabricated from these materials, one can operate these spectrometers at room temperature or cooled to temperature around −20° C. The most advanced of such materials are cadmium zinc telluride, mercuric iodide, and cadmium telluride, While the development of new materials is normally a gradual process, progress in the case of CZT has been very rapid over the past 3 years, with detecor efficiencies increasing by a factor of 100 over this period. In 5 years the technology has progressed from counters incapable of energy resolution to spectrometers with resolution far exceeding any system based on scintillators.

In their present state of development, commercial detectors are limited to thicknesses of approximately 2 mm and volumes of less than 1 cm^3. However, experimental gamma-ray spectrometers have been fabricated and tested (James, 1996), having volumes of over 2 cm^3 and thicknesses up to 25 mm. Due to the high effective atomic number of the materials, this thickness provides good absorption efficiency to over 1 MeV (James, 1996). Energy resolution at 662 keV is about 1%, which is an order of magnitude worse than cryogenic Ge(HP) (~0.1%), but at least a factor of 5 better than scintillator-based instruments. In addition to enhanced resolution (relative to scintillators), they offer substantial reductions in weight, volume, and power requirements compared to both scintillators and cryogenic systems. Hole trapping, which degrades energy resolution and is common to wide-bandgap materials, has been greatly reduced through the application of novel electronic pulse-processing techniques and electron-transport-only devices (James and Schlesinger, 1995; and James et al., 1993). A major challenge for future development of these materials is achieving both high-energy resolution and high effeciency throughout the energy region of interest in astrophysical applications (100 keV to 10 MeV). In the short term, efforts are underway to achieve the necessary thickness for high-energy efficiency through large volume CZT and HgI_2

detectors (volumes of 5 cm^3) and stacked detector configurations. In addition to enhanced high-energy efficiency, the multiple detectors can be used to reject internally-produced gamma-ray activity. The achievement of high efficiency and high resolution in monolithic detectors will be the focus of longer-term development efforts.

Semiconductor Photocells

Semiconductor photocells suitable for use with scintillators using both silicon-based and mercuric-iodide-based materials are also underway. These photocells could potentially offer not only performance advantages over photomultiplier tubes, but also operate at a lower voltage and are nearly immune to voltage and magnetic-field fluctuations. Their unusually high quantum efficiency is particularly well matched to CsI, where energy resolutions of better than 5% have been reported at 662 keV (Squillante, 1996; Wang, Iwanczyk, and Patt, 1994; and Iwanczyk and Patt, 1997)

12.3 Detector Response to Monoenergetic Radiation

The pulse-height spectrum obtained when monoenergetic gamma rays are detected by a scintillation or semiconductor detector is never a line, but is of a shape determined by gamma-ray energy and the source-detector configuration. These shapes are determined by (1) the relative magnitude of the photoelectric, Compton, and pair-production cross sections and (2) the losses and statistical fluctuation that characterize the detector system. The latter consideration was discussed in a previous section, and it was indicated that one could describe to a first approximation this spreading for a given energy ε_0 by Gaussian

$$y = A_m \exp -\left[\frac{(E - E_0)^2}{\alpha^2 E_0^n}\right], \qquad (12\text{-}14)$$

where y is the count rate at energy ε or pulse height V, and A_m is a constant.

We first consider the shape of the response of a detector to monoenergetic gamma rays at lower energies, when photoelectric absorption predominates and Compton scattering and pair

production can be considered negligible. The kinetic energy imparted to a secondary electron is equal to the energy of the gamma ray minus the electron-binding energy. This binding energy can be reclaimed in a sense by the absorption of the x-rays produced by photoelectric absorption. There is also the possibility that the x-rays may escape the crystal without being absorbed. The pulse-height distribution caused by photoelectric absorption is characterized by two regions: (1) the region of total absorption (the photopeak) and (2) the region of total absorption minus x-ray escape energy (the escape peak). This distribution spreading, plus the Gaussian spreading discussed above, yields a pulse-height spectrum similar to that shown in Figure 12.3.

When Compton scattering becomes an important energy-loss mechanism, another region is observed in the pulse-height spectrum, the so-called Compton continuum. All the energy lost in scattering will be given up to the electron as kinetic energy. The gamma ray may lose part of its energy to the detector; furthermore, after suffering one or more Compton collisions, it may then undergo a photoelectric absorption, losing its remaining energy. Thus, the gamma ray either loses all or part of its energy in the crystal, or escapes the crystal at a diminished energy. See Figure 12.4 for a NaI(Tl) pulse-height spectrum and Figure 12.5 for a pulse-height spectrum measured with a Ge(HP) detector. The improvement in energy resolution can be easily noted.

At gamma-ray energies higher than 2 MeV pair production becomes appreciable. Two false photopeaks are then observed. Figure 12.6 is the pulse-height spectrum obtained using ^{24}Na source. The gamma-ray energies emitted by ^{24}Na are 2.76 MeV and 1.38 MeV. The three peaks of greatest height are caused, in order of increasing pulse heights, by (1) pair production with escape of both annihilation quanta, (2) by pair production with the absorption of one annihilation quantum, and (3) by pair production with absorption of both annihilation quanta and total absorption by the photoelectric effect, or any combination of other effects leading to total absorption.

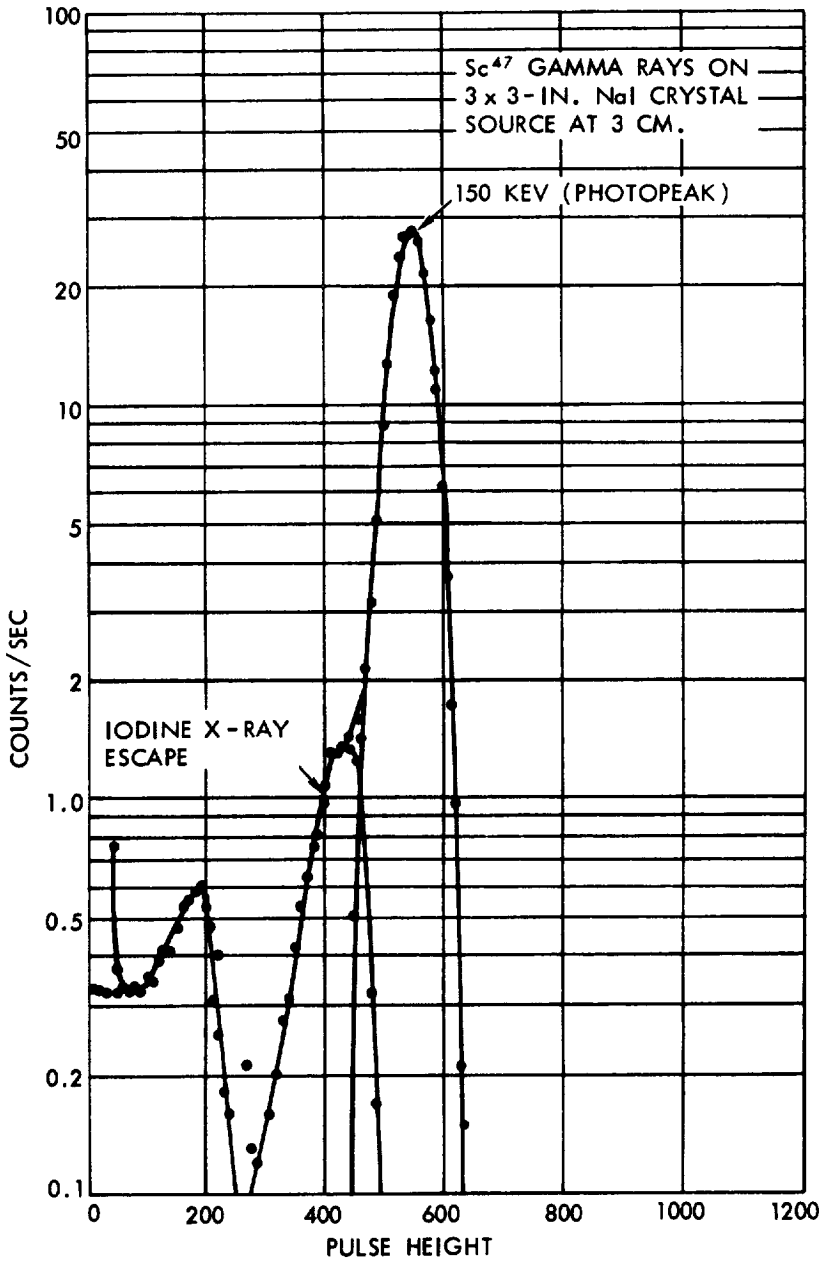

Figure 12.3. ^{47}Sc gamma rays on 3 x 3 in, NaI(Tl) crystal. Source at 3 cm; energy scale 0.42 keV/PHU.

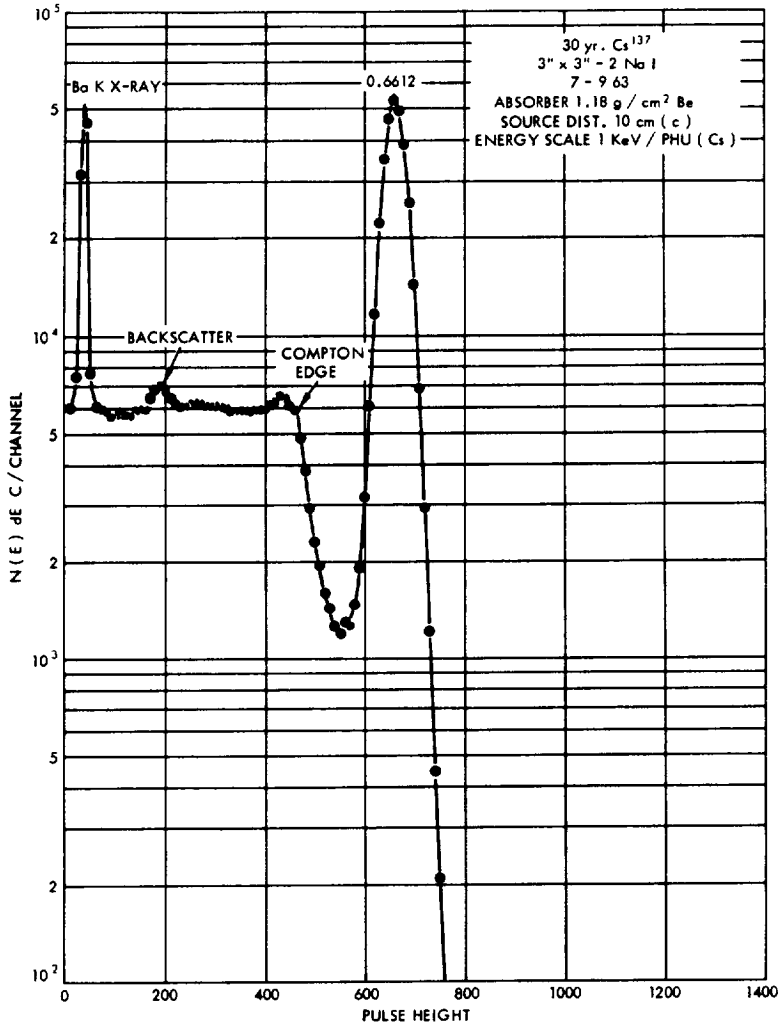

Figure 12.4. ^{137}Cs gamma rays on 3 x 3 in. NaI(Tl) crystal. Source at 10 cm; energy scale, 1 keV/PHU.

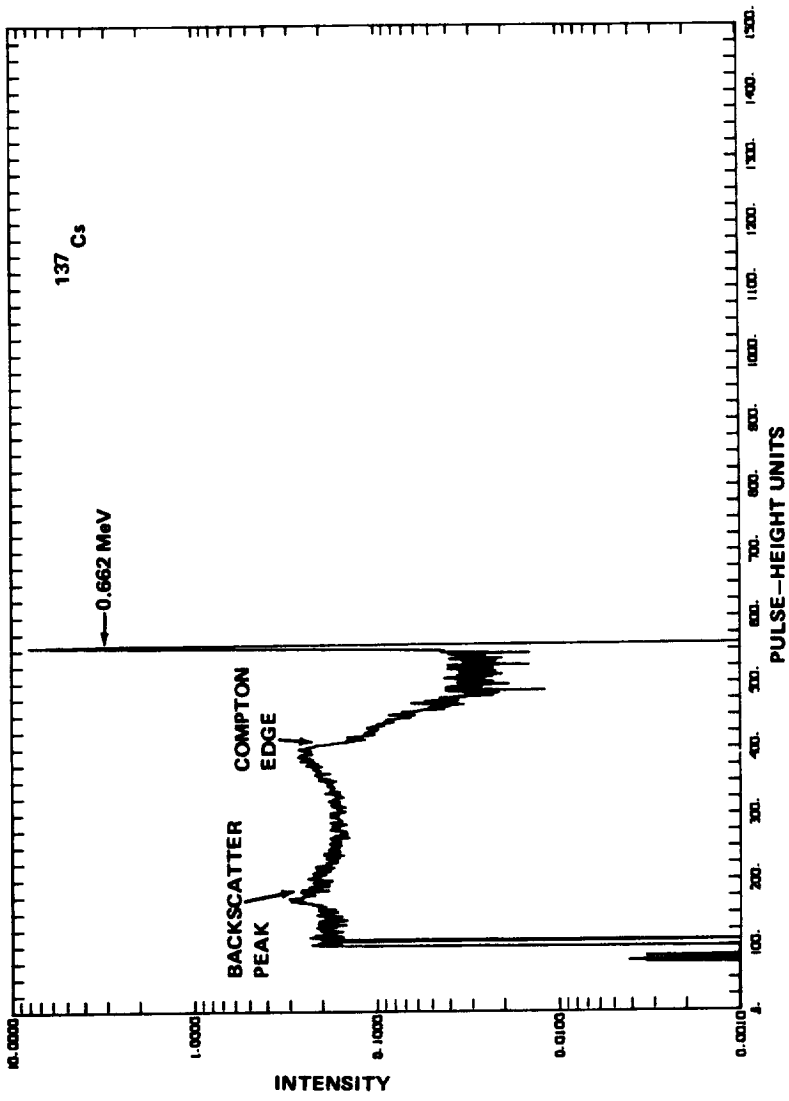

Figure 12.5. ^{137}Cs gamma rays on 120 cc Ge(HP) detector. Source at 25 cm.

In addition to the photopeak, the iodine x-ray escape peak, the Compton continuum, and the pair-escape peaks, there are several other regions characteristic of experimentally determined pulse-height spectra of monoenergetic gamma rays:

1. Multiple Compton scattering region. Because of multiple Compton scattering from materials surrounding the source and crystal, which degrades the primary gamma-ray energy, there is a continuous distribution of gamma-ray incidents upon the crystal with energies less than the initial gamma-ray energy. This tends to spread out the true Compton continuum produced by gamma rays of undegraded energy scattering in the crystal .

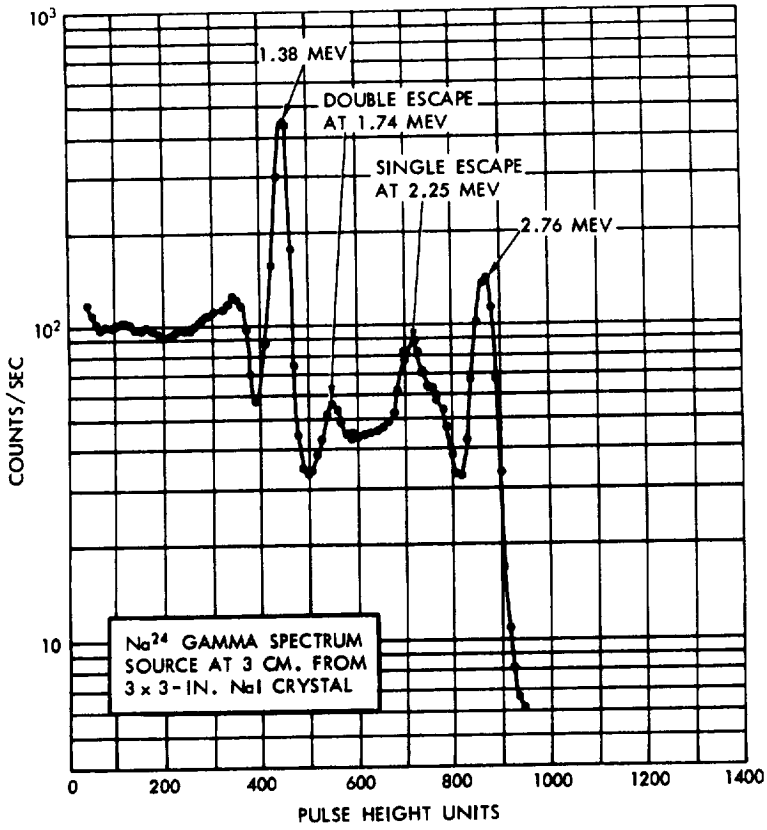

Figure 12.6. ^{24}Na gamma rays on 3 x 3 in. NaI(Tl) crystal. Source at 3 cm; scale about 3.1 keV/PHU.

2. Annihilation radiation from the surroundings. If there is position emission, they may annihilate in surrounding material. Some of the 0.511 MeV gamma rays thus produced reach the crystal, and a pulse-height spectrum characteristic of 0.511 MeV gamma rays is superimposed on the monoenergetic pulse-height spectrum (see Figure 12.7).

3. Coincidence distribution. If two gamma rays interact with the detector during a time that is shorter than the decay time of the light in a scintillator or the electron-hole/collection time in a semiconductor, a pulse will appear whose height is proportional to the sum of the energies lost to the crystal by both interacting gamma rays (see Figure 12.8 for NaI(Tl)). Figure 12.9 shows the

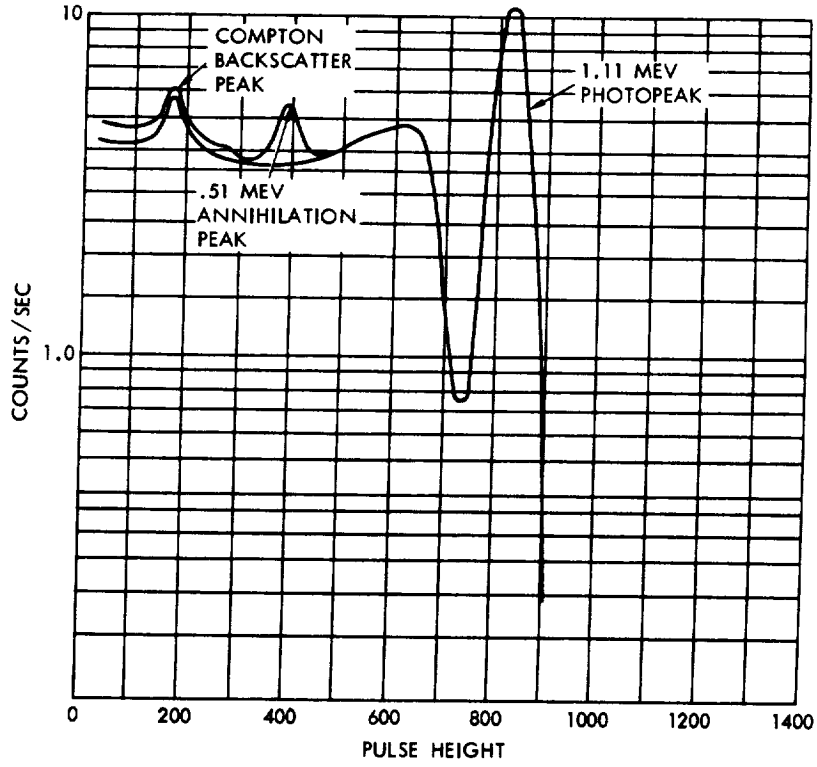

Figure 12.7. ^{65}Zn at 0.24 cm from 3 x 3 in. NaI(Tl) crystal. Scale about 1.3 keV/PHU.

pulse-height spectrum using Ge(HP). The sum peak is not shown in this later figure. A single gamma interacting with the crystal produces only one pulse, whose height is affected by the type and number of interactions by the given gamma ray. If these coincidence effects are negligible, the measured monoenergetic pulse-height spectrum can be considered as a distribution of the probability of energy loss as a function of energy for the given gamma-ray energy and source-detector geometry.

4. Escape of high-energy electrons. When very-high-energy (greater than about 10 MeV) gamma rays interact, the kinetic energy of the electrons may be great enough that some of the electrons escape the detector without losing their total energy to the detector material. This tends to smear out the pulse-height spectrum and at high energies (greater than 15 MeV) the photopeak and the first and second escape peak merge.

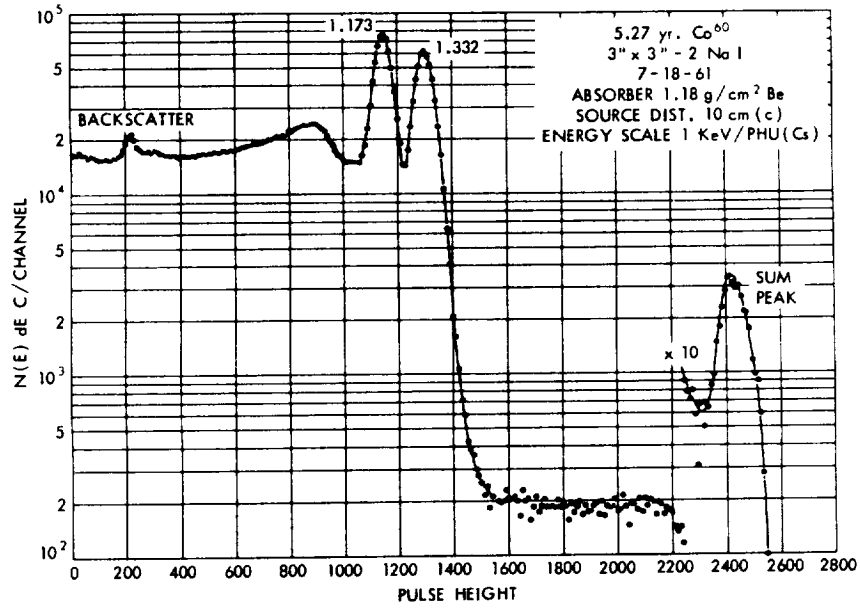

Figure 12.8. ^{60}Co gamma rays on 3 x 3 in. NaI(Tl) crystal. Source at 10 cm; energy scale 1 keV/PHU(Cs).

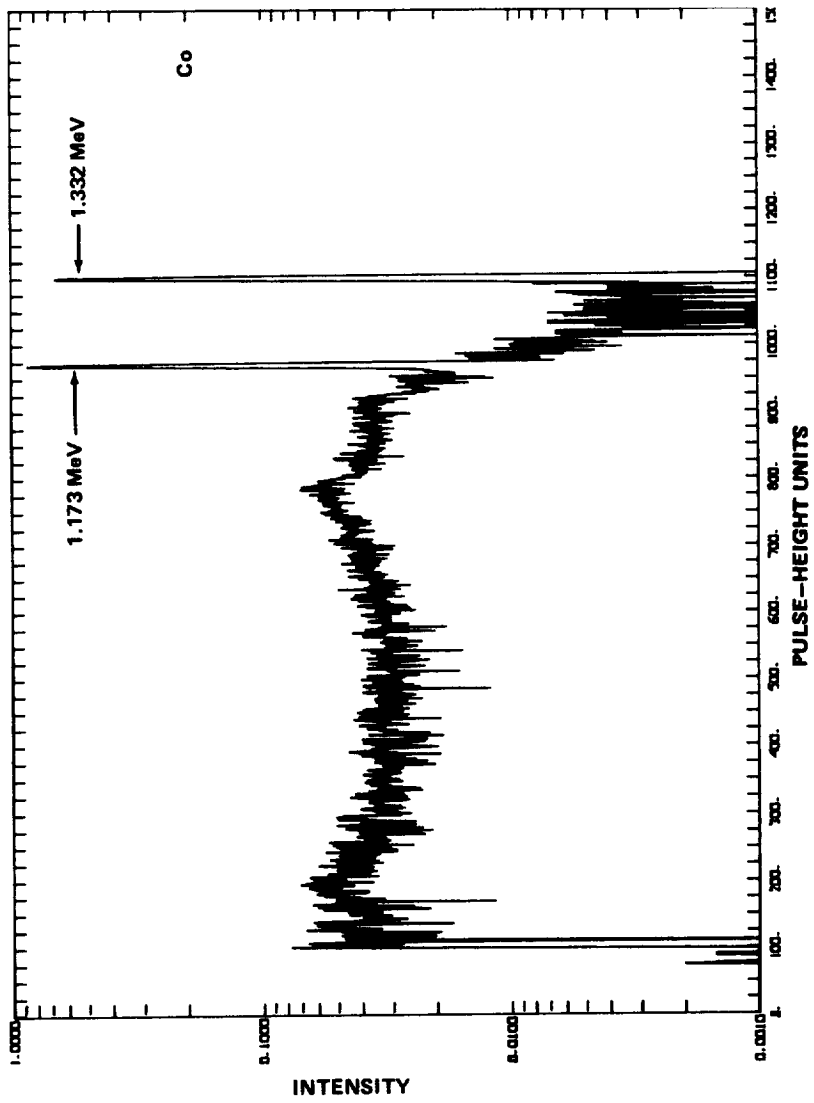

Figure12.9 $^{60}C_o$ on 120 cc Ge(HP) detector. Source at 25 cm above a Ge(HP) detector.

If coincidence losses are negligible, the pulse-height distribution caused by a polyenergetic gamma flux will be a summation of the pulse spectra associated with the various monoenergetic components in the polyenergetic gamma flux. A simple example is shown in Figure 12.10. The pulse-height spectrum caused by a source that

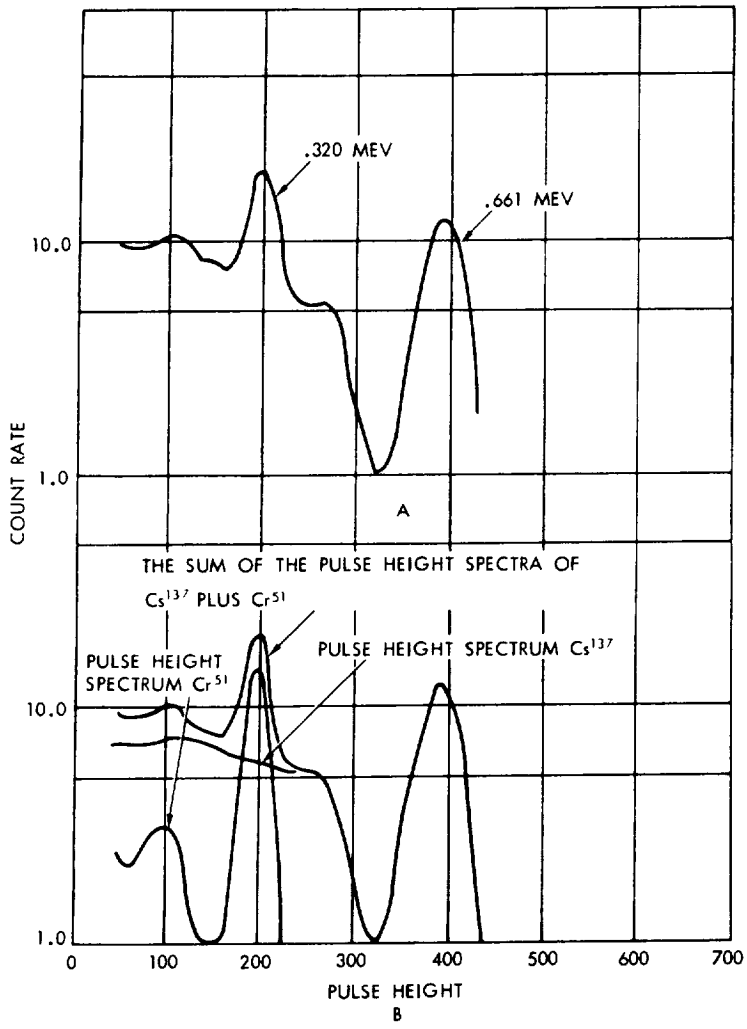

Figure 12.10. Pulse-height spectra of ^{137}Cs and ^{51}Cr, using a 3 x 3 cm NaI(Tl) detector.

was a mixture of ^{137}Cs (0.661 MeV) and ^{51}Cr (0.320 MeV) was measured (Figure 12.10A). The pulse-height spectra of ^{137}Cs and ^{53}Cr were also measured separately, obtaining the pulse-height spectra in Figure 12.10B. The sum of the monoenergetic spectra is also shown in Figure 12.10B and is the same as the spectrum obtained in Figure 12.10A. The source strength of ^{137}Cs and ^{51}Cr was the same in both measurements. Theoretical calculation of detector response functions for NaI(Tl) crystals can be found in Seltzer (1975).

12.4 Detector Efficiency

The response of a particular detecor system can be evaluated in terms of the detection efficiency which depends on the detector and the incident-gamma-ray flux geometry, and the energy distributon of this incident flux. A number of theoretical methods have been developed to calculate these efficiencies (e.g., Trombka et al., 1974; Halbleib, 1992). There are a number of different detector-effeciencies values used in the literature in the following, a number of detector efficiencies are defined.

Total Efficiencies

The total efficiency ε_T is defined for a given energy E_0 as

$$\varepsilon_T(E_0) = \int_\Omega \left(1 - e^{\mu(E_0)\rho(\Omega)}\right)\frac{d\Omega}{4\pi}, \qquad (12\text{-}15)$$

where $\mu(E_0)$ is the total interaction cross section for E_0 defined in (11-12)

$\rho(\Omega)$ is the gamma-ray path through the detector and is a function of Ω, and

Ω is the source/detector geometry,

and the integration is performed over Ω.

A more useful efficiency is the total intrinsic efficiency ε_{Ti}, where

$$\varepsilon_{Ti}(E_0) = \frac{\varepsilon_T(E_0)}{\int_\Omega \frac{d\Omega}{4\pi}}. \qquad (12\text{-}15a)$$

This efficiency now relates only to the angular distribution of the incident flux and thus does not include, in a sense, the source-detector geometry.

Peak Efficiency

The efficiencies as defined above relate to the total interaction. In the analysis of pulse-height spectra it may be more appropriate to study the area under the photopeak only. This area can be determined much more precisely than the total area for a given gamma-ray energy. There are two major reasons for the difficulty in obtaining the total area. First, it is rather difficult to eliminate all scattering effects caused by the surrounding materials. These will appear as pulses in the Compton continuum. Second, pulse-height analyzers cannot direct all pulses down to zero pulse height, for below certain pulse heights the electronic noise dominates the spectrum.

For a monoenergetic gamma-ray incident upon a detector, let

A_p = area under the photopeak
A_t = area under the total pulse-height distribution
ε_{pi} = Intrinsic peak efficiency
$P_t = A_p/A_t$ and is called the peak-to-total ratio,

then

$$\varepsilon_{pi} = P_t \varepsilon_{Ti} . \tag{12-16}$$

For right-cylindrical crystals the peak-to-total ratio P_t does not change significantly as a function of distance along the axis of the cylinder (Francis, Harris, and Trombka, 1957; Miller, Reynolds, and Snow, 1958). Figure 12.11 is a typical plot of P_t as a function of energy for two different geometries. When quoting the efficiency of a Ge detector, the percent efficiency relative to a 3 x 3 in. NaI(Tl) scintillation crystal for the 1.33 MeV line (^{60}CO) is sometimes used. Figure 12.12 is a plot of the intrinsic peak efficiency for a 24 percent Ge detector. The efficiency has been determined for a parallel beam of gamma rays incident on the surface of the detector and perpendicular to the cylindrical axis of the detector (Gordon et al., 1979). At low energies (less than 0.1 MeV) the efficiency drops because of absorption in the dead layer of the detector. The intrinsic efficiencies for the detection of the first and second escape peak are also shown.

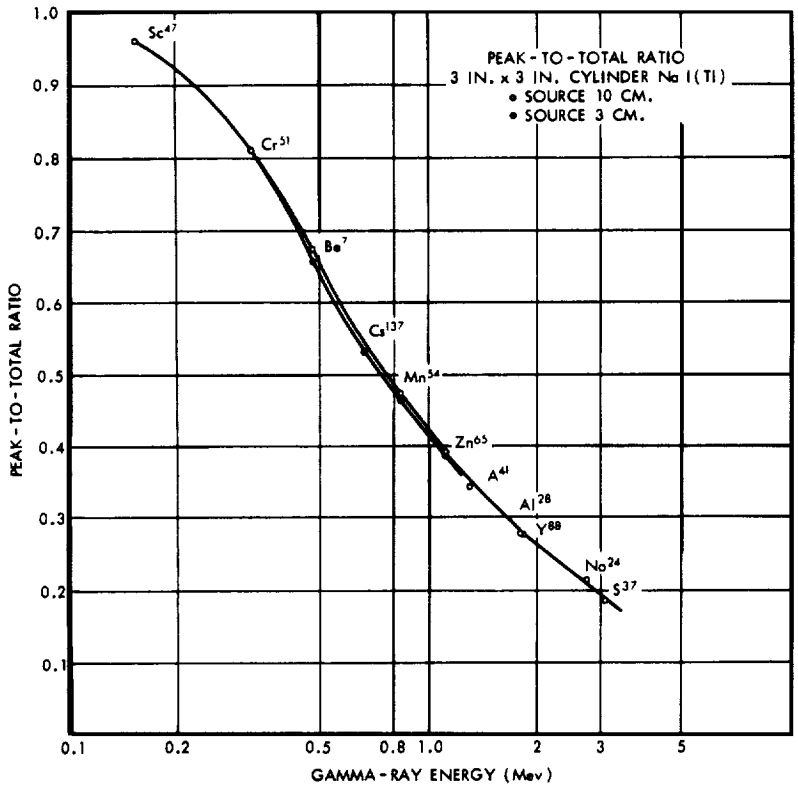

Figure 12.11. Peak-to-total ratio of a 3 x 3 in. cylinder of NaI(Tl) for source distances of 3 cm (closed points) and 10 cm (open points).

The intrinsic peak efficiency can be determined experimentally, using radioactive sources with known source strengths and in a well-defined source-detector geometry. The results can be compared with the theoretical calculations. In the case of scintillation counters, the detector active volume is well-defined. Of course, corrections must be made for absorption in materials housing the detector. The active volume for scintillation is just the volume of the detector. In solid-state detectors such as Ge the active volume depends on the field collection of the electron-hole pairs and the so-called dead layers. There the active volume is not necessarily the total crystal volume. The comparison between theoretical calculation and experiment determination is therefore more difficult.

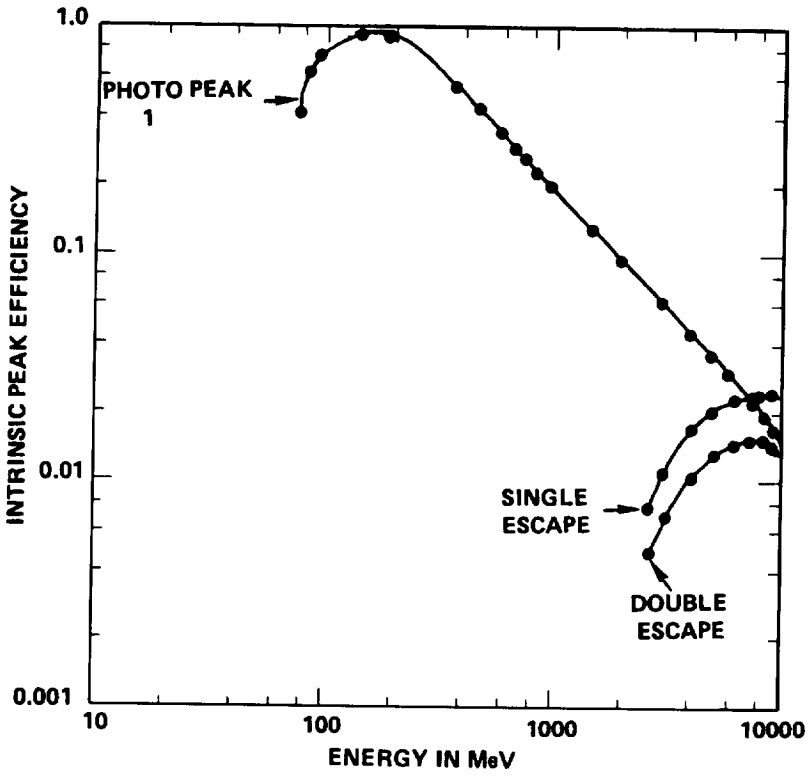

Figure 12.12. Intrinsic peak efficiency for a 100 Ge(HP) detector.

The reason comparison between experimental determinations and theoretical calculations is important is that the experimental determination can only be carried out for certain restricted detector/source geometries. These geometries are not necessarily those geometries characteristic of the space-flight environment. Since the efficiencies can depend on the detector/source geometry, theoretical calculations are needed to obtain the proper intrinsic peak efficiencies.

12.5 Radiation Damage, Annealing

Radiation damage in detectors becomes a significant problem when they are used for extended space-flight missions. Both high-energy protons and neutrons contribute to producing interstitial vacancy pairs in crystals by knocking atoms in the lattice out of their normal positions. Regions of very extensive damage can be produced.

These regions may contain significant concentrations of donors and/or acceptor sites and thus provide trapping sites for holes and electrons. Radiation can also cause degradation of the exposed surfaces, which may lead to an increase in the detector-leakage current. Radiation damage caused by electrons and gamma rays is negligible compared with that caused by energetic neutrons and protons. The radiation damage is reflected in the degradation of the detector resolution. For example, this degradation can become observable when Ge detectors are exposed to fluences of about 10^9 neutrons/cm^2 for fast neutrons and about 10^8 protons/cm^2 for fast protons, which is about a 1-year's dose of Galactic Cosmic-Ray (GCR) protons (Pehl et al., 1978b; and Bruckner et al., 1991). This may be the integral dose during solar-minimum years, but during solar-maximum years the dose may be increased by two orders of magnitude by solar energetic particles. Fast neutron irradiation of Ge detectors produces defects that act predominantly as hole traps (Kraner, Chasman, and Jones, 1968). In conventional high-purity germanium, HpGe, coaxial detectors with p-type electrode configurations, the hole-collection process dominates the signal, whereas when the electrode configuration is reversed (n-type), the electron-collection process dominates the signal. Thus, the effects of radiation damage caused by cosmic-ray primaries and secondaries can be reduced significantly by using the n-type material. In fact, it has been shown that the flux required to produce significant radiation damage can be increased by an order of magnitude in n-type detectors (Pehl, Varnell and Metzger, 1978a; Pehl et al., 1978b). These effects are well known for Ge detectors, and verification of these problems has been obtained as a result of background measurements obtained with the Mars Observer GRS and the WIND TGRS (Seifert, 1996). However, very little experience is available for room-temperature detectors. These detectors must be studied in order to determine their viability for space-flight application during both solar-maximum and solar-minimum periods.

Annealing Ge(HP) at roughly 100° C for about 1 day produces a significant recovery of energy resolution in radiation-damaged detectors. Somewhat higher-temperature annealing (about 150° C) has been used and complete recovery has been achieved. Recent studies in Ge(HP) detectors have shown that radiation damage caused by minimum-ionizing protons is nearly independent of the dose rate and energy (Pehl, personal communication, 1987). The

Mars Observer Gamma-Ray Spectrometer included an annealing capability, since the expected integral particle flux over the three years of the mission (plus possible high particle dose from a major solar flare) would have produced significant degradation in the detector resolution (Bruckner et al., 1991).

The annealing process may cause great mechanical and electrical stress on the detector system. Also, including the anneal capability leads to higher detector temperatures because of thermal paths introduced by the heater wires. This is an important factor because of the strong dependence of the effects of radiation damage on detector resolution as a function of temperature (Bruckner et al., 1991). Again, there is very little experience on the recovery of room-temperature solid-state detectors after annealing. A little experience was gained during the testing of the NEAR PIN solar monitor detector (Golsten, 1995). With the detector HV operating, the detector stopped operating after exposure to a fluence of 10^9 protons/cm^2. Annealing at 100° C recovered operation of the detector. Because of this experience, an anneal mode was included in the NEAR solar monitor. It is important that details of the damage and anneal processes be understood for future missions.

12.6 Sources of Gamma-Ray Background in the Space Environment

A number of sources of background interferences have been identified: photons caused by cosmic-ray and trapped-particles primary and secondary interactions in the spacecraft, in the material surrounding the detector, and in the detector itself; natural radio-activity in the spacecraft; bremsstrahlung produced by high-energy electrons; the diffuse gamma-ray background; atmospheric gamma rays; and scattered gamma rays produced in planetary surfaces and atmospheres.

In the following section a brief description of the major background components is included. The magnitude of these various components will be indicated relative to the measurements of the diffuse gamma-ray spectra carried out during Apollo 15 and Apollo 16 (Trombka et al., 1976). Further details of the identification and computation of the background components can be found in Dyer et al. (1975a and 1975b), Trombka et al. (1976) and Dyer et al. (1980). Figure 12.13 shows energy-loss spectra of the various

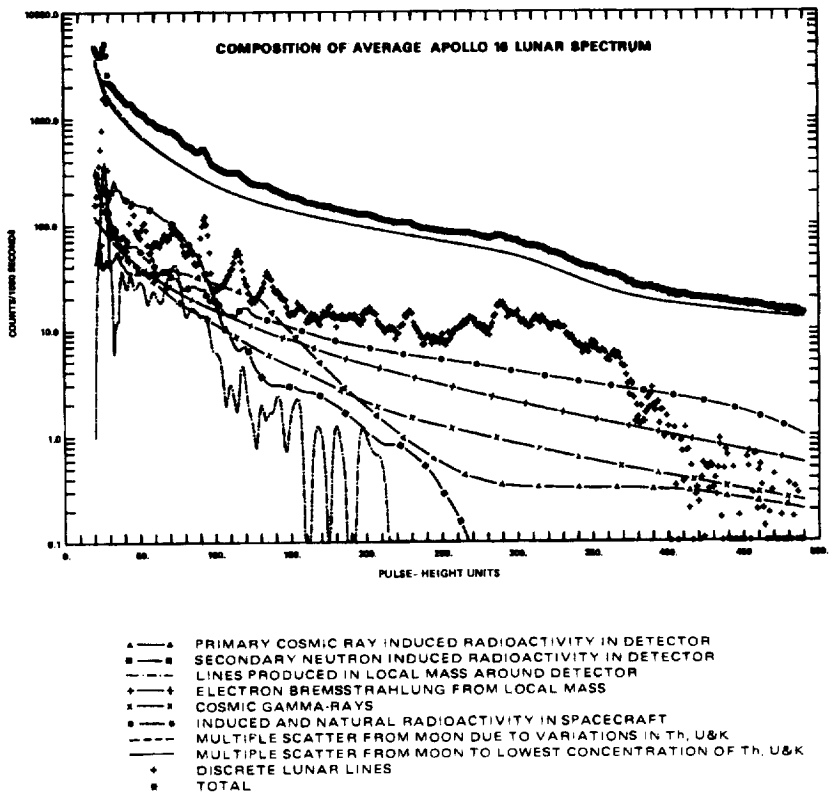

Figure 12.13. The total pulse-height spectra observed at full (7.6 m) boom extension in Apollo 15 and Apollo 16 are compared with the calculated levels of various backgrounds discussed in the text. The latter are shown for the Apollo 16 measurement. The only component to change significantly is the electron bremsstrahlung, which, at the time of Apollo 15, was approximately half the value shown.

background components measured with the Apollo 15 and Apollo 16 gamma-ray spectrometer during translunar flight.

Direct Charged-Particle Detector Counts

Charged particles can interact in the mass of the detector portion of the gamma-ray spectrometer producing a pulse similar to gamma-ray pulses. Active charged-particle shields can be used to decrease significantly the magnitude of this background component. The

Apollo gamma-ray spectrometer with an active charged-particle shield (Harrington et al., 1974) is described below. This system worked well over the energy range up to about 10 MeV (Trombka et al., 1973). Above 10 MeV, the cosmic-ray flux is orders of magnitude higher than the gamma-ray flux for the Apollo spectrometer system, and, therefore, small changes in the rejection efficiency of the active shield changed the observed pulse-height spectrum significantly (Trombka et al., 1976). In near-Earth orbit, the primary cosmic-ray flux will be decreased due to the Earth's magnetic field. Trapped radiation can cause similar problems in active shields and possibly prevent counting during passage through such regions.

Natural Radioactivity

Many sources of natural radioactivity can be found aboard spacecraft, for example, thoriated magnesium, radioactive sources used for space-charge dissipation, radioactive tracers used in quality control of electronic components, and thoriated optical lenses. A survey of sources found aboard the Apollo spacecraft was carried out (Metzger and Trombka, 1972), and may be useful when considering potential radiation-background problems in space-flight experiments and spacecraft design. In terms of the Apollo mission, these sources of background were found to contribute a much lower gamma-ray flux than the cosmic-ray and trapped-particle induced radiation. The Apollo gamma-ray spectrometer pulse-height spectrum of the background contribution caused by natural radioactivity is included in the spacecraft component shown in Figure 12.13.

Spacecraft Component

Cosmic-ray primary and secondary interactions in the spacecraft provide a major background component. The interactions in a planetary body and discussed in Chapter 3 are the same as those that produce gamma-ray emission from the spacecraft material. This component can be significantly reduced by placing the detector on the end of a boom extended away from the mass of the spacecraft. The gamma-ray detector on Apollo was extended 7.6 m from the spacecraft surface. The spacecraft component detected with the detector at this distance is shown in Figure 12.13. In those cases where the detector cannot be extended away from the spacecraft, active BGO, CsI or NaI shields can be used to reduce the

background from the spacecraft. These systems were used on a High Energy Astronomy Observatory (HEAO), Solar Maximum Mission (SMM), and Near Earth Asteroid Rendezvous (NEAR) spacecraft. The SMM and NEAR gamma-ray spectrometer system will be discussed below. However, the materials used in the active shields can contribute to the induced background; thus, great care must be taken in selecting materials for such active shields.

Electron Bremsstrahlung

A change in the translunar pulse-height spectrum measured during the Apollo 15 and Apollo 16 missions was observed. This change is about 15 percent over-all of the energy scale. At the time of the Apollo 16 missions (April 1972) there was a quiet time, a low-energy electron flux increase, possibly of Jovian origin (Teegarden et al., 1974). The electron flux measured during Apollo 16 in the energy domain from 3 to 12 MeV was found to have increased by a factor of about two over that measured during Apollo 15. These higher-energy electrons can produce bremsstrahlung radiation caused by interactions in materials surrounding the detector. It is believed that the difference in the magnitude of the pulse-height spectra measured by Apollo 15 and Apollo 16 can be attributed to the change in low-energy electron flux (Trombka et al., 1976). The pulse-height spectrum caused by the bremsstrahlung component is shown in Figure 12.13.

Cosmic-Ray and Trapped-Particle Induced Radioactivity in the Gamma-Ray Detector

Cosmic-ray primaries and trapped-particle interactions within the gamma-ray detector can produce prompt or delayed gamma-ray and particle emissions. The prompt interactions discussed in Chapters 3 and 4, relative to the Sun and planets, are the important processes producing detector activation. The delayed emissions are a major source of background. They are a result of the build-up of radio-active states which decay with a variety of delays, depending on the lifetimes of the numerous individual excited levels. Since the induced radioactivity is within the detector, both gamma rays and charged particles produce interactions. For high proton energies (in the GeV range), light fragments and their decayed products are produced that deposit most of their energy in the detector: Tables 12-2, 12-3, 12-4, and 12-5 list the sources of the major radioactive

species produced in NaI, CsI, and Ge by cosmic rays and inner-belt protons.

Secondary neutrons are produced as a result of primary charged-particle interactions with the detector and surrounding materials. Both fast and thermal neutron interactions can produce significant induced radioactivity in gamma-ray detectors. The relative effects of cosmic-ray primary and secondary neutron interactions for Apollo gamma-ray measurements are shown in Figure 12.13. These effects dominated over the induced activity produced by passage through the Earth radiation belts (Trombka et al., 1976; Dyer et al., 1975b). The background produced by a given nuclear species depends on the decay mode. In those cases where the gamma-ray emission is preceded by a charged-particle decay, the energy-loss spectra will be observed as the sum of the charged-particle and gamma-ray energy loss, since, in terms of the detection system, these decays are simultaneous. Theoretical calculations of the pulse-height spectra observed as a result of the decay of ^{124}I and ^{24}Na in a NaI detector are shown in Figures 12.14 and 12.15. ^{124}I is a neutron-deficient nuclide. The line features can be attributed to the electron-capture decay branch. Positron emission followed by gamma-ray decay produces the underlying continuum. In the case of a neutron-rich nuclide such as ^{24}Na, electrons' emission always precedes the gamma-ray decay; thus, only a continuum is observed (see Figure 12.15).

A method of calculation for assessing the magnitude and spectral shape of induced activity in a gamma-ray detector has been developed. The method is described in detail in Dyer et al. (1980). Induced activity in the Apollo gamma-ray spectrometer and the OSO-7 gamma-ray detector are considered in detail. The contribution of primary cosmic-ray particles, trapped particles, and their secondaries depends on the flight parameters, the detector type and geometry, and the size of the spacecraft. Comparisons between predictions and observations of induced radioactive background in interplanetary missions can be found in Dyer et al., 1997.

Table 12-2
Isotope List for Activation of Na I/Cs I Crystals by Inner Belt Protons

Isotope	Half-Life	Major Features		Quantity Produced by 1000 Interactions			
				Inner belt		155 MeV	
		β^+ continuum end point (MeV)	γ-ray peaks (keV)	Na I	Cs I	Na I	Cs I
^{133m}Ba	38.9 h		283		2.9		0.8
^{133}Ba	7.2 y		473		2.9		0.8
^{131}Ba	12.0 d		656, 409, 252		36.4		8.4
^{129}Ba	2.2 h	2.45	256, 171		16.2		8.7
^{128}Ba	2.4 d		308, EC		6.5		8.9
^{127}Ba	18.0 m	4.1			3.6		6.3
^{126}Ba	97.0 m	1.8			2.6		4.7
^{132}Cs	6.5 d		702		39.2		41.6
^{131}Cs	9.7 d		EC		33.8		43.6
^{130}Cs	30.0 m	2.99	EC		22.2		29.7
^{129}Cs	32.0 h		451		15.0		21.4
^{128}Cs	3.8 m	3.93	476		10.4		15.7
^{127}Cs	6.2 h	2.1			7.2		11.6
^{126}Cs	1.6 m	4.8			7.0		12.1
^{125}Cs	45.0 m	3.07	146		6.7		11.6
^{124}Cs	26.5 s	5.9			5.9		10.5

Table 12-2 (continued)
Isotope List for Activation of Na I/Cs I Crystals by Inner Belt Protons

Isotope	Half-Life	Major Features		Quantity Produced by 1000 Interactions			
				Inner belt		155 MeV	
		β^+ continuum end point (MeV)	γ-ray peaks (keV)	Na I	Cs I	Na I	Cs I
^{123}Cs	5.9 m	3.1			4.1		7.2
^{122}Cs	4.5 m	6.6			2.2		3.0
^{131m}Xe	11.8 d		164		9.7		12.4
^{129m}Xe	8.0 d		236		8.1		11.7
^{127m}Xe	75.0 s		300, 175, 124	4.5	9.4	1.9	11.6
^{127}Xe	36.4 d		408, 236	4.5	9.4	1.9	11.6
^{125m}Xe	57.0 s		252, 141, 111	26.5	21.7	6.1	12.0
^{125}Xe	17.0 h		277, 222, 87.92	26.5	21.7	6.1	12.0
^{123}Xe	2.08 h	2.7	211,182	20.1	20.8	11.1	17.9
^{122}Xe	20.1 h		272, EC	8.7	9.9	7.3	10.7
^{121}Xe	38.8 m	3.8		3.6	3.6	3.9	4.1
^{126}I	13.02 d	β^- 1.25	1452, 698	61.0	45.4	64.7	51.9
^{125}I	60.14 d		673	51.1	42.1	65.3	58.3
^{124}I	4.18 d	3.16	1357, 635	32.7	27.1	43.6	38.4
^{123}I	13.2 h		191	21.5	18.6	30.6	27.0
^{122}I	3.62 m	3.1	596, EC	14.4	13.4	21.9	19.9

Table 12-2 (continued)
Isotope List for Activation of Na I/Cs I Crystals by Inner Belt Protons

Isotope	Half-Life	Major Features		Quantity Produced by 1000 Interactions			
				Inner belt		155 MeV	
		β^+ continuum end point (MeV)	γ-ray peaks (keV)	Na I	Cs I	Na I	Cs I
^{121}I	2.12 h	2.4	843, 244	10.8	10.9	14.2	14.1
^{120}I	1.30 h	5.6	594	8.2	7.3	13.7	11.1
^{119}I	19.0 m	3.5	289	9.3	7.0	18.4	12.8
^{125m}Te	58.0 d		145	15.1	9.7	19.4	12.4
^{123m}Te	117.0 d		248, 88.5	13.4	8.7	20.0	12.9
^{121m}Te	154.0 d		294, 81.8	10.1	6.8	16.3	10.8
^{121}Te	17.0 d		605, 538	10.1	6.8	16.3	10.8
^{119}Te	15.9 h		1790, 730, 575	7.3	5.3	12.4	8.7
^{119m}Te	4.7 d		2350, 1400, 464, 301, 183	7.3	5.3	12.4	8.7
^{118}Te	6.0 d		EC	11.9	9.7	20.5	15.0
^{117}Te	61.0 m	3.5	750	8.9	6.6	12.0	8.4
^{116}Te	2.50 h		124	4.3	3.3	5.8	4.0
^{120}Sb	15.9 m	2.7	EC	6.1	4.1	10.2	6.6
^{120}Sb	5.8 d		2520, 1500, 1380, 319, 119	6.1	4.1	10.2	6.6

Table 12-2 (continued)
Isotope List for Activation of Na I/Cs I Crystals by Inner Belt Protons

Isotope	Half-Life	Major Features		Quantity Produced by 1000 Interactions			
		β^+ continuum end point (MeV)	γ-ray peaks (keV)	Inner belt		155 MeV	
				Na I	Cs I	Na I	Cs I
^{119}Sb	38.0 h		53	16.0	10.5	24.9	16.2
^{118}Sb	3.50 m	3.7	EC	6.1	4.1	9.1	5.9
^{118m}Sb	5.10 h		2600, 1510, 1370, 1120, 323, 70	6.1	4.1	9.1	5.9
^{117}Sb	2.8 h		187	8.1	5.9	14.2	9.6
^{116}Sb	15.0 m	4.6	1322, 989	3.9	2.9	4.8	3.2
^{116m}Sb	60.0 m	4.6	2929, 1867, 1634, 1089, 673, 574, 268, 128	3.9	2.9	4.8	3.2
^{115}Sb	31.0 m	3.0	2750, 528	5.6	4.3	7.0	4.8
^{114}Sb	3.3 m	6.3		2.1	1.7	2.5	1.7
^{22}Na	2.6 y	2.8	1275	24.6	0.0	26.5	0.0

Note: y = years; d = days; h = hours; m = minutes; s = seconds; EC = electron capture.

Table 12-3
Example Isotopes Produced in Apollo Na I Crystals by the Cosmic Ray Flux

Isotope	Half-Life	Major Features		Cross Section at 1.6 GeV (mbarn)	Decay Rates (Min^{-1}) After 9 d in Flight	
		β^+ continuum endpoint (MeV)	γ-ray peaks (keV)		from cosmic rays	from secondary neutrons
^{126}I	13.02 d	β^+ 2.14 β^- 1.25	1452, 698, EC	57.2	14.4	146.0
^{124}I	4.18 d	3.16	1357, 635	24.4	12.5	129.5
^{123}I	13.2 h		191	17.7	12.5	
^{119}Te	15.9 h		1790, 730, 675	10.5	13.9	
^{117}Te	61.0 m	3.5	750	8.4	11.1	
^{119}Sb	38.0 h		53	15.7	27.9	
^{117}Sb	2.8 h		187	17.5	20.3	
^{116}Sb	15.0 m	4.6	1322, 989	9.2	13.5	
^{116m}Sb	60.0 m	4.6	2929, 1867, 1634, 1089, 673, 574, 268, 128	9.2	6.1	
^{115}Sb	31.0 m	3.0	2750, 528	19.4	12.9	
^{111}In	2.83 d		443, 273, 198	15.2	13.0	
^{110}In	69.1 m	3.93	684	7.1	4.7	

Table 12-3 (continued)
Example Isotopes Produced in Apollo Na I Crystals by the Cosmic Ray Flux

Isotope	Half-Life	Major Features: β^+ continuum endpoint (MeV)	Major Features: γ-ray peaks (keV)	Cross Section at 1.6 GeV (mbarn)	Decay Rates (Mo^{-1}) After 9 d in Flight: from cosmic rays	Decay Rates (Mo^{-1}) After 9 d in Flight: from secondary neutrons
^{110}In	4.9 h		3210, 2510, 2230, 1570, 911, 684	7.1	4.7	
^{24}Na	15.0 h	β^- 4.17		0.0	0.0	30.2
Example light fragments						
^{20}F	11.03 s	β^- 7.03		7.6	6.6	
^{18}F	109.8 m	1.66		9.1	8.5	
^{15}O	122.2 s	2.76		7.5	6.0	
^{16}N	7.1 s	β^- 10.42		4.9	3.9	
^{12}B	0.02 s	β^- 13.37		3.7	3.3	
^{8}Li	0.844 s	β^-, α 16.1		2.7	1.9	
^{6}He	0.802 s	β^- 3.51		9.3	14.5	

Note: y = years; d = days; h = hours; m = minutes; s = seconds; EC = electron capture.

Table 12-4
Major Radioactive Isotopes Produced in Germanium by Trapped Radiation

Isotope	Half-Life[a]	End Point of β^- Contimuum (MeV)	Prominent Lines (keV)	Number Produced per 1000 Interactions
^{76}As	26.5 h	β^- 2.97		2.07
^{74}As	17.9 d	β^- 1.36, β^+ 2.56	607	14.9
^{73}As	80.3 d		EC (11.1)	37.3
^{72}As	26.0 h	β^+ 4.36	641, 797, 846, 905 1061, 1476, 2091, 2212, 2251, 2519, 2921, 3011, 3751	35.0
^{71}As	62.0 h	β^+ 2.01	186	40.9
^{70}As	52.0 m	β^+ 6.24	611, 681, 761, 921, 1051, 1131, 1261, 2041	15.7
^{69}As	15.0 m	β^+ 3.9		
^{75}Ge	82.0 m	β^- 1.20		12.0
^{71}Ge	11.4 d		EC (10.4)	53.1
^{69}Ge	38.0 h	β^- 2.225	EC, 330, 583, 798, 882, 1117	24.6
^{68}Ge	275.0 d		EC(10,4)	11.4
^{67}Ge	19.0 m	β^+ 4.4		3.6

Table 12-4 (continued)
Major Radioactive Isotopes Produced in Germanium by Trapped Radiation

Isotope	Half-Life[a]	End Point of β^- Contimuum (MeV)	Prominent Lines (keV)	Number Produced per 1000 Interactions
^{75}Ga	2.0 m	β^- 3.30		0.97
^{74}Ga	7.9 m	β^- 5.60		3.29
^{73}Ga	4.9 h	β^- 1.55		8.98
^{72}Ga	14.1 h	β^- 4.00		21.3
^{70}Ga	21.1 m	β^- 1.66		32.5
^{68}Ga	68.3 m	β^+ 2.92	EC(9.7), 1088	40.9
^{67}Ga	78.0 h		EC, 101, 195, 216, 306, 398	31.4
^{66}Ga	9.4 h	β^+ 5.175	1049, 2758	19.2
^{65}Ga	15.2 m	β^+ 3.26	125, 162, 216, 760	10.0
^{64}Ga	2.6 m	β^+ 7.07		1.66
^{72}Zn	46.5 h	β^- 0.45		0.51
^{71m}Zn	4.0 h	β^- 2.91		0.93
^{71}Zn	2.4 m	β^- 2.61		0.93
^{69m}Zn	13.8 h		439	3.65
^{69}Zn	57.0 m	β^- 0.90		3.65
^{65}Zn	245.0 d	β^+ 1.35	EC (9.0), 1124	21.6

Table 12-4 (continued)
Major Radioactive Isotopes Produced in Germanium by Trapped Radiation

Isotope	Half-Life[a]	End Point of β^- Continuum (MeV)	Prominent Lines (keV)	Number Produced per 1000 Interactions
^{63}Zn	38.4 m	β^+ 3.366	EC 678, 971	4.2
^{62}Zn	9.3 h	β^+ 1.69	EC 51, 519, 599	1.28
^{64}Cu	12.8 h	β^- 0.57, β^+ 1.68	EC(8.3)	6.0
^{62}Cu	9.8 m	β^+ 3.94	EC	6.7
^{61}Cu	3.32 h	β^+ 2.24	EC, 76, 292, 388, 538, 663, 948, 1158, 1198	2.80
^{58m}Co	9.0 h		25	1.3
^{58}Co	71.3 d	β^+ 2.309	818, 872, 1682	0.65

[a]Isotopes with half-lives of less than 1 minute are not included.

Note: y = years; d = days; h = hours; m = minutes; s = seconds; EC = electron capture.

Table 12-5
Major Radioactive Isotopes Produced in Germanium Under Cosmic Ray Bombardment

Isotopes	Half-Life	Endpoint of β^- Continuum (MeV)	Prominent Lines (keV)	Cross Section at 1.6 GeV (mbarn)	Decay Rate Min^{-1} at Saturation
^{75m}Ge	48.0 s		139	2.7	2.0
^{75}Ge	82.0 m	β^- 1.2		2.7	4.0
^{73m}Ge	0.53 s		54, 13.5, 67.5	11.9	15.60
^{71m}Ge	20.0 m		23, 175, 198	12.3	9.2
^{71}Ge	11.4 d		EC (10.4)	12.3	18.4
^{69m}Ge	5.1 μs		85	7.1	5.3
^{69}Ge	38.0 h	β^+ 2.225	EC, 330, 583, 798, 882, 1117	7.1	10.6
^{68}Ge	275. d		EC (10.4)	6.4	4.8
^{73}Ga	4.9 h	β^- 1.55		9.0	6.7
^{72m}Ga	0.036 s		99	5.4	4.0
^{72}Ga	14.1 h	β^- 4.0		5.4	8.1
^{70}Ga	21.1 m	β^- 1.66		16.8	12.6
^{68}Ga	68.3 m	β^+ 2.92	EC (9.7), 1088	23.1	22.1
^{67}Ga	78.0 h		EC, 101, 195, 216, 306, 398	20.7	15.5
^{66}Ga	9.4 h	β^+ 5.175	1049, 2758	13.3	9.9

Table 12-5 (continued)
Major Radioactive Isotopes Produced in Germanium Under Cosmic Ray Bombardment

Isotopes	Half-Life	Endpoint of β^- Continuum (MeV)	Prominent Lines (keV)	Cross Section at 1.6 GeV (mbarn)	Decay Rate Min^{-1} at Saturation
^{65}Ga	15.2 m	β^+ 3.26	125, 162, 216, 760	7.4	5.5
^{69m}Zn	13.8 h		439	5.7	4.3
^{69}Zn	57.0 m	β^+ 0.901		5.7	8.6
^{67m}Zn	9.3 μs		93	4.3	18.7
^{65m}Zn	1.6 μs		54	13.4	15.5
^{65}Zn	245. d	β^+ 1.349	EC (9.0), 1124	13.4	26.5
^{63}Zn	38.4 m	β^+ 3.366	EC 678, 971	7.4	5.5
^{64}Cu	12.8 h	β^- 0.573, β^+ 1.678	EC (8.3)	9.6	7.2
^{62}Cu	9.8 m	β^+ 3.94	EC	12.3	9.2
^{61}Cu	3.32 h	β^+ 2.236	EC, 76, 292, 388, 588, 663, 948, 1158, 1198	7.6	5.7
^{58m}Co	10.6 μs		80	4.7	3.5
^{58m}Co	9.0 h		25	4.7	7.0
^{58}Co	71.3 d	β^+ 2.300	818, 872, 1682	4.7	10.5
^{57}Co	270. d		22, 129, 143	9.9	7.4
^{55}Fe	2.6 y		EC(6.5)	9.8	7.3
^{54}Mn	303. d		841	11.2	8.4

Table 12-5 (continued)
Major Radioactive Isotopes Produced in Germanium Under Cosmic Ray Bombardment

Isotopes	Half-Life	Endpoint of β^- Continuum (MeV)	Prominent Lines (keV)	Cross Section at 1.6 GeV (mbarn)	Decay Rate Min^{-1} at Saturation
^{51}Cr	27.8 d		EC (5.5), 325	10.0	7.5
^{49}V	330. d		EC(5.0)	10.3	7.7
^{48}V	16.00 d	β^+ 4.013	950, 988, 1317, 2246, 2300, 3229, 3245	4.6	3.4
^{44m}Sc	2.44 d		271	2.4	1.8
^{44}Sc	3.92 h	β^+ 3.647	1160	2.4	3.6
^{28}Al	2.31 m	β^- 4.635		5.4	4.0
^{7}Bc	53.0 d		478	12.2	9.1
^{6}He	0.8 s	β^- 3.510		8.1	6.1
				Σ = 343.5	Σ = 347.5

Note: y = years; d = days; h = hours; m = minutes; s = seconds; EC = electron capture.

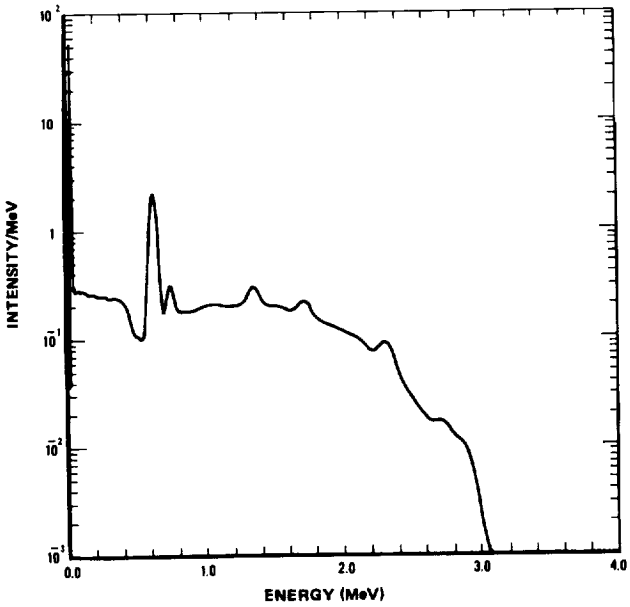

Figure 12.14. Computed Response function for ^{124}I inside the 7 cm x 7 cm NaI(Tl) crystal. The underlying continuum results from several positron branches, while the line features result from electron-capture modes.

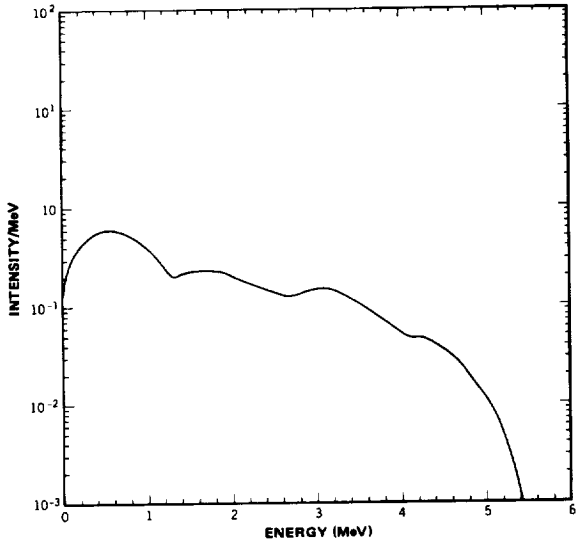

Figure 12.15 Computed response function for ^{24}Na decay inside the Apollo 7 cm x 7 cm NaI(Tl) crystal. The continuum results from b^- spectra shifted by coincidence with one or more emitted gamma rays.

A detailed pulse-height spectrum obtained with the MARS Observatory during the cruise phase of the mission is shown in Figure 12.16 (Evans, L.G., 1997). A number of induced lines are identified. The identification of these lines is discussed in Boynton et al., 1997.

Atmospheric and Planetary Surface Gamma-ray Interactions

Multiple scattered gamma-ray radiation contributed about 80 percent of the background for the measurement made near the lunar surface. The magnitude of this effect is shown in Figure 12.13.

For experiments onboard spacecraft in Earth orbit or near planetary bodies with atmospheres, charged- and neutral-particle interactions will produce discrete-line gamma-ray emission. Natural radioactive elements in the atmosphere are a further source of discrete line gamma-ray emission. Multiple scattering of the gamma-ray and annihilation radiation produces a continuum and a 0.511 MeV line.

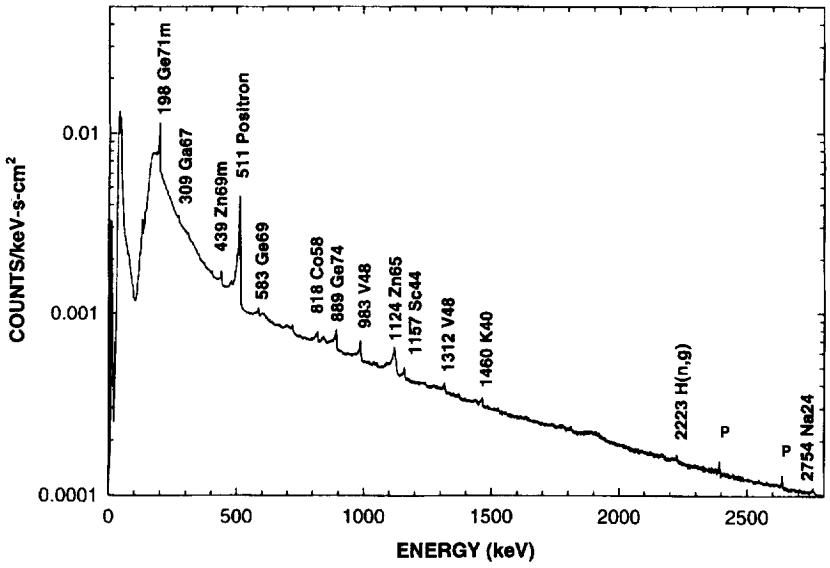

Figure 12.16. Pulse-height spectrum obtained with the MARS Observer Gamma-Ray Spectrometer during the cruise phase of the mission. The dominant induced activation lines are noted.

In Earth orbit the background will be modulated by the longitudinal and latitudinal variations in the magnetic fields and the variations in the trapped-radiation belts, which change the exciting particle flux. Details of studies of terrestrial atmospheric gamma rays can be found in Daniel and Stephens (1974), Kinzer, Share and Seeman (1974), and Ryan et al. (1977).

Diffuse or Cosmic Gamma Rays

The diffuse gamma-ray spectrum was discused in Chapter 5. The pulse-height spectrum of this component as measured by the Apollo gamma-ray spectrometer is shown in Figure 12.13.

Table 12.6 summarizes the contributions of various background components measured during a lunar flight by the Apollo gamma-ray spectrometer.

Gamma-Ray Shields and Collimators

As can been seen from the preceding discussion, a major source of background can be attributed to cosmic-ray-induced activation of the spacecraft and materials surrounding the gamma-ray detector. The spacecraft background can be reduced in two ways. First, by using a boom and extending the detector away from the spacecraft, thus reducing the background relative to the primary gamma-ray emission from planetary surfaces or solar, Galactic or extragalactic sources. The background also can be reduced by the use of shields. Gamma-ray shielding can be achieved using either active or passive shields. Passive shields are high-density shields that absorb gamma rays completely and thus reduce the unwanted flux. The problem with such shields in the space environment is that the cosmic rays interacting with the passive shield will produce more secondary gamma radiation that escapes the shield than will be absorbed by the shield. Usually, active shielding must be used. Active shields utilize high-density scintillation detectors that detect both gamma-ray and particle radiation. The detected radiation is then electronically detected and used to reject any events that occur simultaneously in the shield and central gamma-ray detector. In this way, both the spacecraft background and the background from much of the material surrounding the central detector can be significantly reduced. Three examples of such systems are: the Gamma-Ray Advanced Detector (GRAD) system used on an Antarctic balloon flight in 1987 (Rester et al., 1990) and on the Long-

Table 12-6
Composition of Apollo Spectrum (percentage contribution)

Energy (MeV)	Spacecraft		Bremsstrahlung		Neutron		Cosmic Ray		Discrete Lines		Cosmic	
	A15	*A16*	*A15*	*A16*	*A15*	*A16*	*A15*	*A16*	*A15*	*A16*	*A15*	*A16*
0.4-0.8	18	15	12	19	12	11	12	11	15	16	31	27
0.8-2.0	15	13	14	23	9	8	28	24	11	12	22	20
2.0-5.0	20	17	19	30	6	5	33	29	3	3	19	17
5.0-10.0.	40	33	26	42	0	0	11	10	0	0	23	16

Note: *A15* is *Apollo* 15; *A16* is *Apollo* 16.

duration Antarctic Mars-calibration Balloon Flight (LAMB) (Floyd et al., 1991); and the NEAR Gamma-Ray Spectrometer (GRS) (Trombka et al., 1997). The NEAR design has proven extremely effective as demonstrated during the cruise phase of the mission. Studies to obtain optimum shield design with minimum noise must be carried out when designing a system to be consistent with the particular mission constraints.

Presently BGO material has been used for the shield, but materials such as CsI can also be used. New materials such as gadolinium oxyorthosilicate (GSO) and lutecium oxyorthosilicate (LSO) (Melcher et al., 1990) show great promise. These shields also can be used as a collimator to better define the region on the planetary surfaces or other astrophysical sources, as is done on the NEAR GRS (Trombka et al., 1997). The shields also act to reduce Compton scattering and pair-production effects in the central crystal, thus simplifying data interpretation. For a discussion of these effects, see Knoll, 1989.

Active shield/collimator designs strongly depend on the nature of the central detector (e.g., semiconductor or scintillator) and spacecraft configuration. If an active collimator for both gamma-ray and charged-particle rejection is not used, then a charged-particle detector active shield will be needed. In the past, for the Apollo and Mars Observer GRS, thin scintillation detectors were used. It is believed that thin room-temperature solid-state detectors could be used in place of these scintillation systems. These room-temperature devices should be lighter, smaller, and take less power than the equivalent scintillation system.

12.7 Detector Flight Systems

The design of a flight gamma-ray detector system will strongly depend on: the application in the field of astrophysics, space physics, planetary physics, and geochemistry; spaceflight mission contraints; and allowable costs for the mission. It will be impossible to review all the gamma-ray spectrometer systems which have been flown. Therefore, a number of space-flight systems will be described in the following sections, which will illustrate the various design approaches.

The Apollo Gamma-Ray Spectrometer and Near Earth Asteroid Rendezvous (NEAR) Gamma-Ray Spectrometer

The gamma-ray detector assembly flown on Apollo 15 and Apollo 16 was composed of three major subassemblies: the electronics, the scintillation detector, and the thermal shield (Harrington et al., 1974). A partial view in cross section is shown in Figure 12.17. In flight this system was deployed at the end of a 7.5-meter boom in order to remove it far from both the spacecraft natural radioactivity and the spacecraft activity induced by the ambient cosmic-ray flux.

The detector subassembly consisted of a right cylindrical NaI(Tl) crystal, about 7 x 7 cm. A thin mantle of a scintillating plastic crystal, which was optically isolated from the primary NaI crystal, and both detectors were used in an anticoincidence mode. The plastic scintillator is transparent for all practical purposes to gamma rays with energies greater than 200 keV. Cosmic rays produce significant interactions in both detectors over the energy region of interest. Thus, the plastic shield acts as an active shield reducing the background produced by charged cosmic rays. In the case of Apollo, during translunar flight the cosmic-ray background, relative to gamma rays with energies greater than about 12 MeV, is about two orders of magnitude greater. Rejection efficiencies of the order of 99 percent for the anticoincidence systems reduce the cosmic-ray

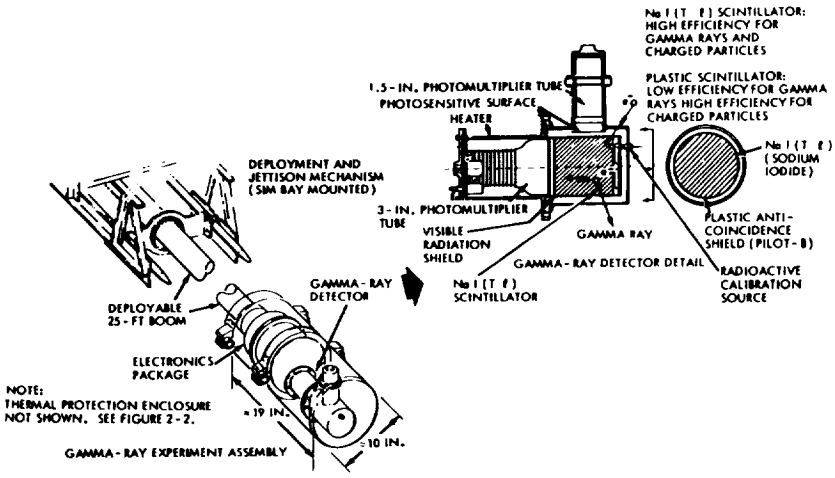

Figure 12.17. Apollo gamma-ray detector detail.

background as detected in the spectrometer to the same magnitude as the gamma-ray flux. It was, therefore, difficult to detect the high-energy (greater than 12 MeV) gamma-ray flux on Apollo 15 and Apollo 16 during the translunar flight (Trombka et al., 1977).

Detector systems similar to the Apollo system were used by the United States on the Ranger Mission (Metzger et al., 1964) and by the Soviets during the Luna program (Vinogradov et al., 1968). The major difference in detector-system design was in the method of observing the scintillations in the active plastic shield and the inorganic central scintillation detector. The two systems were optically separated on the Apollo detector system and separate photomultipliers were used to observe the interactions. In both the Ranger and the Luna detector systems a single photomultiplier was used to observe interaction in both the plastic and the inorganic scintillators. The fluorescence-decay time in the plastic and the inorganic detectors are significantly different. This is reflected as a difference in rise time observed at the output of the photomultiplier tube. The rise time of pulses produced in the plastic scintillator are much shorter than those produced in the inorganic scintillator, and that fact can then be used to discriminate between interactions occurring in the plastic detector and inorganic scintillator. This latter technique is called a "phoswitch."

A further modification of the basic Apollo system had to be made in the design of the gamma-ray spectrometer designed for the Near Earth Asteroid Rendezvous (NEAR) Gamma-Ray Spectrometer (GRS) (Trombka et al., 1997; Golsten et al., 1997). Because of the strict weight, power, and volume constraints, a scintillation system had to be used and the detector had to be mounted on the spacecraft, since a boom mechanism would add too much weight . The NEAR GRS utilized a NaI(Tl) center detector and the outer or shield detector was BGO instead of the plastic scintillator used on the Apollo GRS. The outside active shield of BGO then acts as both an active particle and a gamma-ray shield. Furthermore, the shield acts as a collimator improving the spatial resolution of the detector system. There are further benefits which can be attributed to the Compton suppression achieved in the anti-coincidence mode. Compton suppression will be discussed in the next section.

Some significant information can be lost when a gamma-ray/cosmic-ray active shield is used. The first and second escape peaks, produced as a result of pair-production interactions in the central

detector, are greatly reduced in the anti-coincidence spectrum, For a small central detector like the NEAR design, this loss is unacceptable, since the first and second escape peaks will be larger than the photopeak, and for planetary remote sensing the high-energy lines above 3 MeV are very important. This information can be recovered by collecting two additional NaI spectra in coincidence with, respectively, the 0.511 MeV and the 1.022 MeV gamma rays detected in the BGO shield detector. The first and second escape peaks are easily observed in these coincidence spectra. Thus the NEAR GRS system collects five pulse-height spectra: the raw spectrum; the anti-coincidence spectrum; the 0.511 MeV coincidence spectrum; and the 1.022 MeV coincidence spectrum.

The Solar Maximum Mission (SMM) Gamma-Ray Spectrometer, the Compton Gamma-Ray Observatory (CGRO), and the Oriented Scintillation Spectroscopy Experiment (OSSE)

The gamma-ray detector system carried aboard the SMM spacecraft is shown schematically in Figure 12.18 (Forrest et al., 1980). The central (NaI(Tl)) detector system was designed with a large surface area in order to increase the sensitivity for the detection of low-intensity gamma-ray flux. Furthermore, fine temporal resolution is required for studying the dynamics of gamma-ray emission during solar flares. Good energy resolution is also of utmost importance in determining any line-broadening phenomena, as discussed in Chapter 4. Seven high-resolution (7 percent FWHM for the 0.661 MeV line) detectors constitute the main portion of the gamma-ray spectrometer system. Each of these components consists of a 7.6 x 7.6 cm integral line NaI(Tl) detector. The gain and zero of each of the seven detectors are automatically controlled so that all of the pulse-height scales are the same. An electronic servo loop is used to correct each output continuously to a common position. Three ^{60}Co calibration sources (emitting 1.17 MeV and 1.33 MeV gamma-ray energies) are used to achieve the automatic control (Forrest et al., 1972). A common pulse-height analyzer could be used because all seven detectors were forced to a common gain. The presence of the ^{60}Co did not significantly interfere with the measurements, since transient events were being measured.

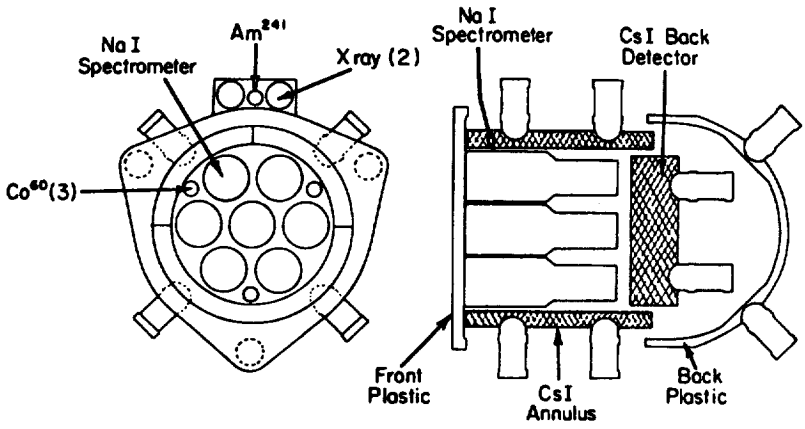

Figure 12.18. Schematic drawing of the SMM gamma-ray spectrometer experiment showing the major subsystem components in top and crosssectional side views (reprinted with permission from David J. Forrest, Forrest et al., 1980).

Detailed background information can be obtained before and after the occurrence of a gamma-ray burst associated, for example, with a solar flare. The calibration lines are subtracted by using these background measurements.

The gain of these detectors is set so that measurements are made up to 9 MeV (relative to the highest channel in the analog-to-digital converter). The lower-level discriminator is set at 0.3 MeV. The digitized signals (out of the common ADC) are stored in two memories. One of these accepts only events associated with any single detector (singles mode), while the other memory accepts only events associated with two or more of the seven detectors (multiple mode). The two modes can be processed separately or summed.

The singles mode adds a further improvement to a discrete-line gamma-ray spectrometer. If the total energy of the gamma ray is lost in a given detector, it will appear as a signal in the photopeak. On the other hand, if a gamma ray is partially absorbed (in terms of energy loss to the detector), some of the energy, either by Compton scatter or annihilation radiation following pair production, may escape from the detector. The escaping gamma-ray energy can interact in the other detectors or in the active shields and thus be

rejected in the singles mode. The Compton continuum and first and second escape peaks can then be suppressed in the pulse-height spectra. The discrete lines (photopeaks) will be enhanced, in a sense, relative to the continuum. In the sum modes, the total events occurring in the seven-detector system can be obtained (multiple mode). This may be important for the detection of rather weak gamma-ray fluxes, where the total pulse-height spectrum may be needed to obtain statistically significant results. This method has also been used successfully with the NEAR GRS.

The basic accumulation time for the singles mode and multiple mode pulse-height spectra is 16.38 seconds. Integral counts can be accumulated for 2-second intervals for three energy regions: 4.2 MeV to 4.6 MeV, 4.6 MeV to 5.6 MeV, and 5.6 MeV to 6.4 MeV. These regions can yield temporal information during flares concerning prompt gamma-ray line emission from ^{12}C and ^{16}O at 4.4 MeV and 6.1 MeV, respectively, on a much finer time scale than differential energy spectral systems. There is one further integral window covering the energy range 300 keV to 350 keV that has an even finer time resolution (0.654 second) for fast-time resolution study of hard x-ray bursts.

The energy range for spectral analysis can be increased up to 15 MeV by command. Evidence of the 15.11 MeV ^{12}C line in very large flares can be investigated (Crannell, Ramaty, and Crannell, 1977). The thick CsI(Na) crystal at the back of the detector is also part of the high-energy detector system. Events in the 10 to 100 MeV range occurring in the NaI and/or the CsI crystals are analyzed by separate pulse-height analysis. A matrix accumulation system is used in order to determine which of the observed interactions are associated with a single high-energy detector, and which elements are associated with gamma-ray showers and interaction in more than one detector.

Two 8 x 0.6 cm thick NaI detectors are present. One has an Al filter used to measure x-rays in the 10 to 80 keV range and the other has an Al-Fe filter, used to measure the x-ray flux in the 95 to 140 keV region. The basic purpose of the x-ray detectors is to provide corrected data for identification of the times when electron acceleration is occurring at the Sun. Photons and ions producing nuclear excitations are most likely to occur during these times.

Active plastic anticoincidence shields completely surround the gamma-ray detector system. These shield elements are used to suppress the background produced by direct charged-particle interactions as well as the background gamma-ray flux coming from all directions, except from the Sun. The CsI and plastic elements shown in Figure 12.18 constitute the active shield. Care must be taken in using such active shields because they themselves can be a major source of induced gamma-ray emission.

A similar type of shield-detector design was used for the HEAO-C gamma-ray spectrometer, except that the NaI(Tl) crystals were replaced with Ge(HP) detectors (Jacobson, 1977; Mahoney et al., 1980). This system is capable of much better energy resolution than the one described above. Many of the background lines produced in the CsI shield have been observed during flight with this detector system.

An active gamma-ray shield can be used in lieu of extending the detector system away from the spacecraft mass on a boom (the Apollo systems, as compared with the SMM system). The active shields are also used for obtaining directional resolution. The shield material contributes to the background. This background varies with time. For short times, the background can be relatively constant. In those cases where transient events of short duration are observed, the background subtraction obtained before and after an event can be used. For those cases where a weak constant source is to be observed, the background problem becomes quite serious. This is especially true in certain Earth orbits where the spacecraft passes through the South American Anomaly and where there is a varying geomagnetic field that will induce rather fast-time variation in the induced background. The use of active blocking shields has been considered. In this design multiple detectors are used looking in the same direction. Half of these detectors are covered by an active shield, thereby allowing simultaneous source and background determination. The major difficulties with this approach are (1) the active blocking crystals become activated in space and contribute to the background and (2) the active shields are not completely opaque to the incident gamma-ray flux. Furthermore, the gamma-ray count rate can increase in the so-called "blocking crystal" configuration, compared with the unblocked detector. Great care must be taken in using such systems.

Successful observations of astrophysical sources in the 0.1-10 MeV range were carried out aboard the Compton Gamma-Ray Observatory (CGRO) by the Oriented Scintillation Spectroscopy Experiment (OSSE). (Johnson et al., 1993) A NaI(Tl)-CsI(Tl) phoswich system was used in which the CsI crystals are placed behind the NaI crystals. In this arrangement, rejection of events entering the rear of teh OSSE system can be rejected by pulse shape discrimination between the CsI and NaI. An active background shield both particle and gamma ray is obtained by combining an annular shield assembly with the CsI portion of the phoswich system. In the OSSE design a passive tungsten alloy collimeter is used. The collimeter is located within the annular shield and defines a field of view of 3.8° x 11.4° throughout the 0.1-10 MeV energy domain. The central NaI detectors have achieved an energy resolution of ~8% at .661 MeV. The OSSE detectors have been operated in coaxial pairs. These detectors can be operated to modulate between pointing on the source and pointing off the source. Thus, while one detector in a pair points on the source, the other monitors the background. High-time resolution studies of transient gamma-ray events by monitoring the time variation profiles in the active shield elements.

Some other examples of such detector systems can be found in Womack and Overbeck (1970), Nakano et al. (1973), Metzger (1973), and Rester et al. (1990).

Gamma-Ray Burst Detection

Gamma-ray burst-detection systems have been included aboard many Earth orbiting and planetary spacecraft. Since, as has been discussed in Chapter 10, gamma-ray-burst source distribution is isotropic, triangulation methods are used to determine source direction. The time of arrival of the gamma-ray burst on a number of widely separated detectors is used. Gamma-ray burst detectors must be designed to obtain detailed measurement of gamma-ray-burst temporal distribution, and for the more complete burst-detection systems energy spectral distributions are measured during the evolution of the burst. The most successful detector system flown up to this time is the Burst and Transient Source Experiment (BATSE), flown aboard CGRO (Fishman et al., 1989). During the 693 days of the mission, 588 cosmic bust have been observed.

BATSE utilizes eight gamma-ray detectors of identical configuration. These modules are placed on the spacecraft so as to provide all-sky coverage. Each module contains two scintillation detectors. One compound is a large-area cylindrical NaI(Tl) detector, 50 cm in diameter and 1.3 cm thick. Three 1.3 cm diameter photomultiplier tubes are optically coupled to this crystal. A charged-particle active shield is used on the front of this detector to reject charged-particle background. A thin lead-and-tin shield inside the housing reduces the amount of background and scattered radiation entering the back side. The second module utilizes a smaller spectroscopy detector with higher energy resolution. This second detector is a NaI(Tl) cylinder, 13 x8 cm thick. These modules are designed to measure gamma-ray temporal variations on timescales down to several microseconds and energy spectra in the energy domain of 30 keV to 1.9 Mev.

High-Energy Resolution Systems

A high-resolution Ge(HP) detector system was flown on HEO-C, as mentioned earlier. A few more systems have been flown successfully on both planetary and astrophysics missions and large imaging systems have been flown on balloon flights and are being planned for future space-flights. Non-imaging systems will be considered in this section, while imaging systems will be considered in the section on coded apertures.

The Mars Observer Gamma-ray (MO GRS) spectrometer will be first be considered. The MO GRS (Boynton et al., 1993) was launched in October, 1992, and operated successfully for eleven months during the cruise phase. The spacecraft was lost prior to insertion into Mars orbit. Much has been learned about the operation of such detectors systems from the data obtained during the cruise period. A cross section of the detector system is shown in Figure 12.19. The major components are: the Ge-detector assembly; the passive cooling system, consisting of the radiator plate and the V-groove thermal shield; and the anticoincidence shield/neutron detector, consisting of boron-loaded, plastic scintillator panels and photomultiplier tubes. Not shown are the sunshade and sunshade door used to keep thermal emission reflected from the Mars surface from striking the radiator.

The anticoincidence/neutron shield serves two purposes. The first is to act as an active charged-particle shield. The plastic scintillator is

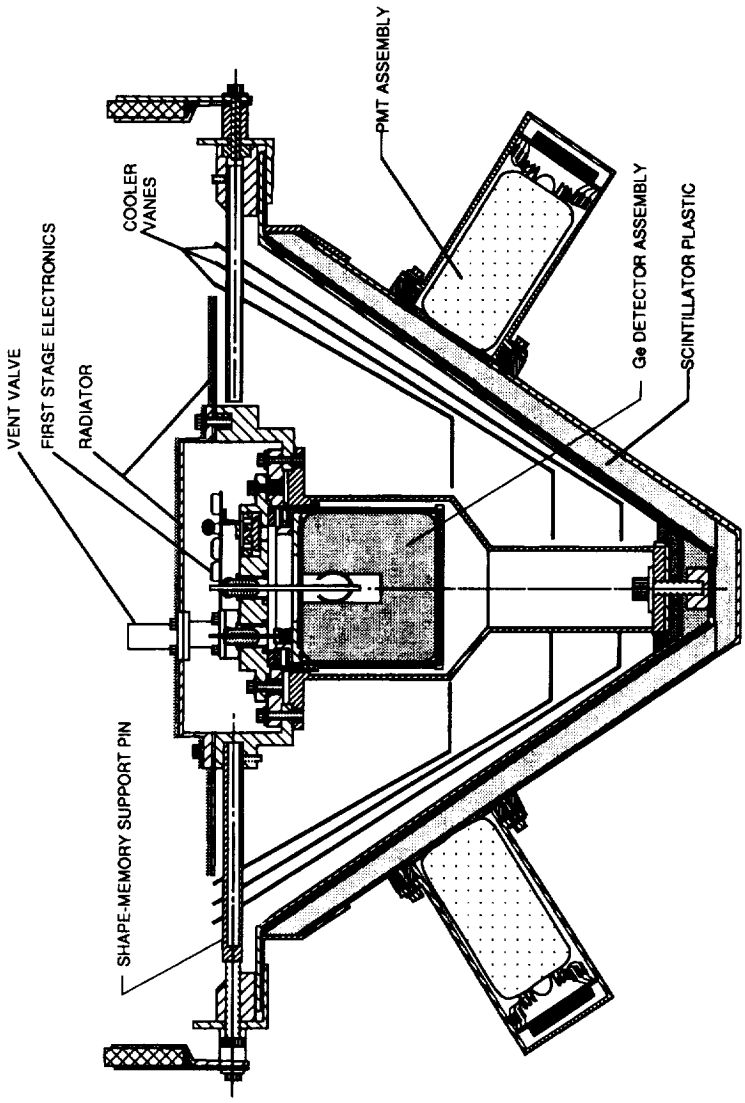

Figure 12.19. The MO GRS sensor head contains the Ge(HP) detector, the front stage electronics, the surrounding boron-loaded plastic scintillator panels and photomultiplier tubes, making up the anticoincidence/neutron detector shield and the radiator and V-groove passive cooling system.

insensitive to gamma rays in this energy domain, that is above 100 keV, but does detect charged particles passing through the shield. Simultaneous events between those detected in the shield and the Ge detector are placed in anticoincidence and thus these events will be, for the most part, rejected and eliminated from the measured gamma-ray pulse-height spectrum.

The second purpose of this shield was to determine the neutron flux at Mars. The front and rear faces of the shield see different count rates for low-energy neutrons, because the spacecraft is moving faster than the velocity of a thermal neutron, but not as fast as an epithermal neutron. The neutron signals can be used to infer the fluxes of thermal and epithermal neutrons (Feldman et al., 1993). The plastic scintillator emits light when struck by a charged particle. It is also doped with Boron, an element with a high cross section for the $^{10}B(n,\alpha)^{7}Li$ reaction, so that it can also generate a light pulse following the capture of a thermal neutron. The pulse from a neutron-capture event has a characteristic energy of about 80 keV that allows it to be distinguished from a charged-particle event, which is typically 500 MeV or more.

The system also included a heater system to allow for annealing if the radiation degradation of the detector became too bad. It was found that annealing at about 100° C for about seventy hours would be sufficient to recover the detector, so that the energy resolution would be almost the same as the undamaged detector. One problem was introduced by including the anneal heater and that was that the parisitic-heat losses introduced by the heater wires caused a degradation in the operation of the passive cooler. The system could not reach temperatures below 100° K. The energy degradation caused by damage is very temperature dependent, and at elevated temperatures the energy-resolution degradation for a given exposure dose is greater. Over the eleven-months period of exposure the energy reslution was degraded.

The design of the passive cooling system for the Transient Gamma-Ray Spectrometer (TGRS) for the WIND mission (Owen et al., 1995) was different. No annealing capability was included in order to achieve the lowest-possible temperature. The TGRS detector is a single, unshielded n-type Ge crystal. The design of the passive cooler is a two-stage radiative Be cooler similar to that flown on ISEE-3 (ISEE-3 reference). The detector operated at 85° K. Over

the first year of operation the resolution degraded, but showed less degradation than that observed on the MOGRS for roughly the same radiation dose.

The Compton Telescope

A rather interesting detector design has been described by Herzo et al. (1975), based on the detection of a Compton-scattering interaction of a gamma ray in a first detector followed by the detection of the Compton-scattered photon in a second detector. This method allows for the simultaneous determination of photon direction and energy. The method can be described using the two-dimensional representation of a so-called "double scatter" Compton telescope shown in Figure 12.20. Neutrons can also be detected in such a system.

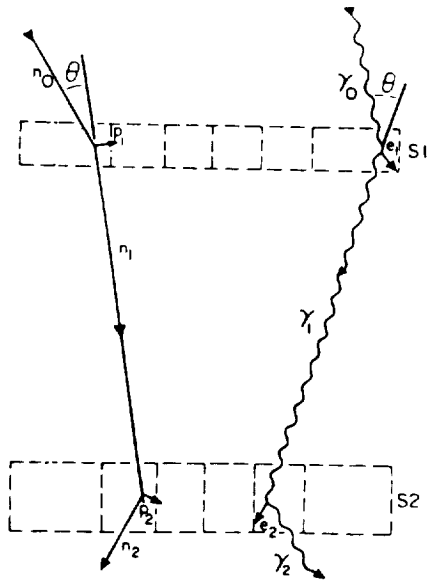

Figure 12.20. Schematic diagram of a double-scatter telescope. Reprinted with permission from Herzo et al. (1975) and Nuclear Instruments and Methods.

The case for gamma-ray detection is first considered. A gamma ray with energy E_{g0} can interact in detector S_1; if the interaction is of the Compton-scattering type, the scattered gamma-ray γ_1 emerges from S_1 at some angle θ with respect to the incident direction. The energy of the scattered gamma ray is E_{g1}. The scattering occurs with an electron e_1 at rest and E_{g1} kinetic energy is imparted to this electron. Thus

$$E_{g0} = E_{g1} + E_{e1}. \tag{12-17}$$

The Compton scattering process was discussed in Chapter 11 and the relationship between θ, E_{g1} and e_2 was given in Equation (11-9), remembering that energy is expressed in electron rest-mass energy units (i.e., 1.022 MeV is expressed as E = 2 electron rest-mass energy units). The scattering angle θ can be written as

$$\theta = \cos^{-1}\left[1 - \left(\frac{1}{E_{g1}} - \frac{1}{E_{g0}}\right)\right]. \tag{12-18}$$

The gamma ray, γ_1, strikes the second detector, S_2, and can be totally absorbed or partially absorbed. If γ_1 Compton scatters once in S_2 and the scattered gamma ray, γ_2, escapes the detector, E_2, kinetic energy is imparted to electron e_2 as a result of the scattering interaction and

$$Eg_1 = Eg_2 + E_{e2}. \tag{12-19}$$

The pulse height measured is S_1, caused by the interaction of γ_0 is E_{g1}. Similarly, E_{g2} is the pulse height measured in detector S_2. If, for example, total absorption of γ_1 occurs in detector S_2, then $e_2 = 0$, $E_{g1} = E_{g2}$ and Equation (12-17) can be written as

$$\theta_m = \cos^{-1}\left\{1 - \left[\frac{1}{E_{g2}} - \frac{1}{E_{g1} - E_{g2}}\right]\right\}. \tag{12-20}$$

E_{g1} and E_{g2} are measured. If total absorption does not take place in the second scintillator, the $E_{g2} < e$ and true scattering angle θ will be less than the values determined in Equation (12-18). Schönfelder et al. (1977) and Herzo et al. (1975) have developed a method for correcting the value E_{g2} by a factor α; that is, αE_{g2} is used in Equation (12-17) for determining θ. The factor α is characteristic of the actual detector system and can be determined using Monte Carlo calculations.

The detector scintillators chosen for S_1 and S_2, their thickness and separation are important in detector design. The detector S_1 should have a large Compton cross section over the energy region of interest. If one studies the interaction cross section plate shown in Chapter 11, and Figures 11.1 to 11.6, it can be seen that lower Z materials (plastic scintillators) will be dominated by the Compton scattering over the energy domain 1 to 10 MeV. The detector, S_1, therefore, should be liquid or plastic and thin to allow for single Compton-scattering interactions. In order to achieve total absorption in S_2 on the other hand, total absorption is desirable. Thus, thicker and/or higher Z detector material should be used.

Higher-energy neutron interaction processes are also shown schematically in Figure 12.20. The incident neutron n_0 scatters from a proton p_1 in scintillator S_1. After scattering, the neutron with a lower energy can continue to move and strike detector S_2. The neutron can scatter again from P_2 and the neutron n_2 leaves the detector with a lower energy. The kinetic energy lost to the protons E_{p1} and E_{p2} is the measured pulse-height output signal from S_1 and S_2. Over a large energy range (i.e., less than 100 MeV) neutron interaction can be differentiated from gamma rays by measuring the time of flight between the two detectors. The neutron velocities are less than the velocity of light for these energies. Above 100 MeV, the velocity of the neutron approaches the velocity of light and is, therefore, difficult to be differentiated. Below about 2 MeV, the proton recoil energy becomes too low to detect in the scintillation counters. This sets the lower limit to the energy sensitivity of the system.

The energy of the incident neutron E_{no} is given simply as

$$E_{n_0} = E_{p_1} + E_{n_1}. \qquad (12\text{-}21)$$

The scattering angle is given by

$$\tan\theta = \left(\frac{E_{p_1}}{E_{n_1}} \right)^{1/2} . \qquad (12\text{-}22)$$

Time-of-flight methods are used to reduce the background effects of high-energy neutrons as well as obtaining some information about the neutron energy spectrum in this energy domain. Gamma rays entering in any direction except through S_1 can be rejected either by coincidence for the case of interaction in one detector or time-of-flight methods (when the gamma ray strikes detector S_2 before S_1). Active and passive shields surrounding the telescope are used to reduce the background produced by cosmic rays and the gamma-ray emission from surrounding materials.

A most successful application of this method was the imaging Compton Telescope (COMPTEL), included as part of the Compton Gamma-Ray Observatory (CGRO) (Schönfelder et al., 1993). COMTEL detects gamma rays by the occurrence of two successive interactions in the telescope: first a Compton scattering collision occurs in a detector of low-Z material (liquid scintillator NE213A) in one of the upper planar modules, then a second interaction takes place in a lower plane of detecor-module high-Z material (NaI(Tl), in which, ideally, the scattered gamma ray is totally absorbed. The two detector modules are separated by a distance of 1.5 m. Each detector is entirely surrounded by a thin plastic scintillator anticoincidence shield for charged-particle rejection. Off to the side, between the two detectors, are two small plastic scintillation detectors, containing weak ^{60}Co electronically gated calibration sources. ^{60}Co decays with the emission of two gamma rays in coincidence. By measuring the energy deposit and the positions in both detector planes, the incident direction of the gamma ray is determined to within a ring on the sky.

A number of Compton telescope designs have been proposed, tested, and flown. The following references describe a number of such systems: Alverez et al. (1973); Herzo et al. (1975); Schönfelder and Lichti (1975); Lichti, Moyano, and Schönfelder (1975); Schönfelder et al. (1977); Ryan et al. (1977); and White et al. (1977).

Coded Aperture

Pinhole systems for imaging gamma-ray-emitting objects have been considered for a number of years. Because the gamma-ray fluxes are so low, a single pinhole would reduce the intensity far below other backgrounds, thus making such fluxes virtually undetectable. The concept of using a coded aperture was first introduced by Dicke (1968) and Ables (1968). The approach discussed in these papers involved replacing a single pinhole opening by many pinholes arranged randomly. The pinhole array is called the aperture. Each point on a self-luminous object deposits a shadow of the aperture on a focal plane. Subsequent processing of the picture yields the reconstructed image, which should resemble the original object. The hole patterns of such an aperture significantly determine the signal-to-noise ratio obtained after image reconstruction. Many such patterns have been considered; but in this section, the Uniformly Redundant Array (URA), discussed by Fenimore and Cannon (1978), Fenimore (1978) and Cannon and Fenimore (1979), will be briefly described.

Essentially, a URA is a special kind of multiple-pinhole mask in which the number of times a particular separation occurs between any pair of pinholes is the same for all separations. The separations are, thus, uniformly redundant. Ideally, URA's have the following properties: A point source is imaged through the URA mask onto a detector and casts a magnified shadow of the URA mask on the detector. Subsequently, the same URA pattern, appropriately magnified, is used to correlate with the image to decode it and the reconstructed point source will be a spike superimposed on a constant background. This essentially means that the point-spread function (autocorrelation function in this case) of a URA is a delta function with perfectly flat, but non-zero, side lobes. Furthermore, this flat and constant background can be removed if digital decoding is used. This is done by representing the holes in the URA as +1, and other locations as −1 instead of zero. This so-called balanced-correlation method imposes the constraint that the point-spread function be a delta function with zero side lobes. Because an extended object represents a superposition of point sources, the delta function point-spread function is extremely important for the artifact-free reconstruction of extended objects.

Under development for the International Gamma Ray Astrophysical Laboratory (INTEGRAL) mission, to be flown in the early years of

the 21st century, a coded-mask system for mapping gamma-ray sources will be incorporated. The INTEGRAL consists of two main instruments: a germanium SPECTROMETER and a cesium Iodide IMAGER (Winkler 1994). The SPECTROMETER consists of an array of Ge(HP) detectors cooled by a Stirling cooler to an operating temperature of 85° K. The energy domain to be measured is 20 keV to 8 MeV with an energy resolution of ~2 keV. The IMAGER is the high-spatial-resolution instrument with an array of Cesium Iodide with a detector area of 2500 cm^2. Also planned is an X-Ray Monitor (XRM), which will extend the continuous spectral coverage to a few keV. The SPECTROMETER, IMAGER and XRM will share a common principle of operation, that is, they are all coded-mask telescopes. The hexagonal coded aperatures for the gamma-ray instruments will be located about 2 meters above the detection planes. Background detection and separation is being implemented by providing split but intermingled detector pixels, capable of viewing the source by one set and the other set being blocked by opaque mask elements, thus determining the background.

There are many more gamma-ray spectrometer systems that have been flown, but only a few systems have been described in this section . These systems were chosen to demonstrate some general principles of system design. A number of other systems are described in the Proceedings of the INTEGRAL Workshop On the Multiwave Approach to Gamma-Ray Astronomy, published as special conference proceedings in The Astrophysical Journal Supplement Series, 92, June 1994, and PCGO proceedings

References

Ables, J. G., 1968, Proc. Astr. Soc. Aust., 4, 172.

Adams, F., and Dams, R., 1970, Applied Gamma-Ray Spectroscopy, Chap. 2 (New York: Pergamon Press,1970), p. 21.

Alverez, L. W., Dauder, P. M., Smith, L. H., Buffington, A., Derenzo, S.E., Muller, R. A., Orth C., and Smoot, G., 1973, The Liquid Xenonl Compton Telescope, A New Technique for Gamma-Ray Astronony, serie. 14, issue 17 (Berkeley, Calif: Space Sciences Lab., 1973).

Bertolini, G., and Coche, A., 1968, eds., Semiconductor Detectors (New York: John Wiley and Sons, Inc., 1968).

Bird, L. L., 1956, Tracerlog, 78, 12.

Boynton, W. V., Evans, L. G., Reedy, R. C., and Trombka, J. I., 1993, "Remote Geochemical Analysis: Elemental and Mineralogical Composition," ed. C. Pieters and P. Englert, Cambridge University Press, P. 395-411.

Boynton, W. V., Evans. L. G., Starr, R., Bruckner, J., Bailey, S. H., and Trombka, J. I., 1997, Proc. of the Intern. Conf. on High-Energy Background in Space, NASA Conf. Pub., 3353, 19-22.

Brückner, J., et al., 1991, Proton-induced Radiation Damage in Germanium Detecotrs, IEEE Trans. Nucl Sci., 38, 209-217.

Cannon, T. M., and Fenimore, E. E., 1979, Appl. Opt, 18, 1052.

Crannell, C. J., Ramaty, R., and Crannell, H., 1977, Proc. of the 12th ESLAB Symp., ESA SP-124, 213.

Crannell, C. J., Starr, R., Stottlemeyer, A. R., and Trombka, J. I., 1984, A Balloon-borne High-Resolution Spectrometer for Observations of Gamma-Ray Emission from Solar Flares, Nucl. Instrum. & Methods, A225, 195-208.

Daniel, R. R., and Stephens, S. A., 1974, Rev. Geophys. Space Phys., 12, 233.

Dicke, R. H., 1968, ApJ, 153, L10.

Dyer, C. S., Tromka, J. I., Metzger, A. E., Seltzer, S. M., Bielefeld, M. J., and Evans, L. G., 1975a, Proc, of the 14th Int. Cosmic Ray Conf., 1,2.

Dyer, C. S., et al., 1975b, Space Sci. Instr., 1, 279.

Dyer, C. S., Tromka, J. I., Seltzer, S. M., and Evans, L. G., 1980, Nucl. Instr. and Meth., 173, 585.

Dyer, C. S., Truscott, P., Evans, H., Evans, L. G., and Trombka, J. I., 1997, the Proc. of the Intern. Conf. on High-Energy Background in Space, NASA Conf. Pub., 3353, 43-46.

Evans, L. G., 1997, private communication.

Feldman, W. C., Boynton, W. V., and Drake, D. M., 1993, "Remote Geochemical Analysis: Elemental and Mineralogical Composition," ed. C. Pieters and P. Englert, Cambridge University Press, p. 213-243.

Fenimore, E. E., 1978 , Appl. Opt., 17, 3562.

Fenimore, E. E., and Cannon, T. M., 1978, Appl. Opt., 17, 337.

Fishman, G. J., et al., 1989, in Proc. GRO Science Workshop, ed. W. N. Johnson (Washington, DC: NASA, 2-39.

Floyd, S. R., et al., 1991, Joint US-USSR Long-Duration Antarctic Mars Calibration Balloon (LAMB) Mission, AIAA Int. Balloon Tech. Conf., 20-24.

Forrest, D. J., Higbie, P. R., Orwig, L. E., and Chupp, E. L., 1972, Nucl. Instr. and Meth, 101, 567.

Forrest, D. J., et al., 1980, Solar Phys., 65, 15.

Francis, J. E., Harris, C. C., and Trombka, J. I., 1957, Variation of NaI(Tl)) Detection Efficiencies with Crystal Size and Geometry for Medical Research, ORNL 2204 (Oak Ridge, Tenn.: Oak Ridge National Laboratory, 1979, College Park, Md.: Univ. of Md., Dept. of Chem., 1979), p. 225.

Gamble, R. L., 1955, Ph. D., Diss., Univ. of Texas.

Goldsten, J., private communication, 1995.

Goldsten, J., et al., 1997, in preparation for publication in Space Science Review.

Gordon, G. E., Walters, W. B., Zoller, W. H., Anderson, D. L., and Failey, M. P., Non-Destructive Determination of Trace Element Concentrations, Technical Report ORO-5173-008.

Halbleib, J. A., Kensek, R. P., Mehlhorn, T. A., Valdez, G. D., Seltzer, S. M., and Berger, M. J., 1992, "ITS Version 3.0 The Intergrated Tiger Series of Coupled Electron/Photon Monte Carlo Transport Codes", Sandia National Laboratory Publication SAND 91-1634, Albuquerque, NM.

Harrington, T. M., Marshall, J. H., Arnold, J. R., Peterson, L. E., Trombka, J. I., and Metzger, A. E., 1974, Nucl. Instr. and Meth., 118, 401.

Heath, R. L., Helmer, R. G., Schmittroth, L. A., and Cazier, G. A., 1965, The Calculation of Gamma-Ray Shapes for Sodium Iodide Scintillation Spectrometers, 39th ed., IDO 17017, AEC Research and Development Report, Physics, TID 4500.

Herzo, D., Koga, R., Millard, W. A., Moon, S., Ryan, J., Wilson, R., Zych, A. D., and White, R. S., 1975, Nucl. Instr. and Meth., 123, 583.

Iwanczyk, J. S., and Patt, B. E, 1997, private communication.

Jacobson, A. S., 1977, High Resolution Gamma-Ray Spectroscopy, Paper presented at the 1977 Spring Meeting (Washington, D.C.: A.P.S., 1977).

James, R. B., Schlesinger, T. E., Sifferet, P., and Frank, L., 1993, ed. "Semiconductors for Room Temperature Radiation Detectors", Material Science Symposium Proceedings, Vol. 302.

James, R. B. and Schlesinger, T. E., 1995, "Semiconductors and Semimetals", Vol.43, Academic Press, NewYork.

James, R. B., 1997, private communication.

Johnson, W. N., et al., 1993, A&A, 86, 693.

Kelley, G. G., Bell, P. R., Davis, R. C., and Lazar, N. H., 1956, Nucleonics, 14, 53.

Kinzer, R. L., Share, G. H., and Seeman, N., 1974, J. Geophys. Res., 79, 4567.

Kittel, C., 1953, Introduction to Solid State Physics (New York: Wiley, 1953).

Knoll, G. F., 1989, "Radiation Detectors and Measurements" 2nd edition, J. Wiley and Sons, Inc.

Kraner, H. W., Chasman, C., and Jones, 1968,.Nucl. Instr. and Meth., 62, 173.

Lichti, G., Moyano, C., Schönfelder, V., 1975, Proc. of the 14th Int. Cosmic Ray Conf., 4, 1423.

Mahoney, W. A., Ling, J. C., Jacobson, A. S., Tapphorn, R. M., 1980, Nucl. Instr. and Meth., 178, 363.

Melcher, C. L., Schweitzer, J. S., Utsu, R., and Akuyama, S., 1990, Scintillations Properties of GSO, ISEE Nucl. Sci., 37, 161-164.

Metzger, A. E., 1973, in Gamma-Ray Astrophysics, ed. F. W. Stecker and J. I. Trombka, NASA SP-339 (Washington, D.C.: Government Printing Office, 1973) p. 97.

Metzger, A. E., Anderson, E. C., VanDilla, M. A., and Arnold, J. R., 1964, Nature, 204, 766.

Metzger, A. E., and Trombka, J. I., 1972, Proc. Nat. Symp. on Natural and Manmade Radiation in Space, ed. E. A. Warman, NASA TMX-2440, 32.

Miller, W. F., Reynolds, J., and Snow, W. J., 1958, Efficiencies and Photofractions for Gamma Radiation on NaI(Tl) Activated Crystals, ANL 5902, (Washingtoll, D.C.: ANL, 1958).
Mott, N. F., 1952, Elements of Wave Mechanics (Cambridge, Eng.: Cambridge University Press, 1952).
Mott, M. E., and Sutton, R. B., 1958, in Handbuch der Physik, Nucl. Instr. II., vol. XLV (Berlin: Springer-Verlag, 1958).
Nakano, G. H., Imhof, W. I., Reagan, J. F., and Johnson, R. G., 1973, in Gamma-Ray Astrophysics, ed. F. W. Stecker and J. I., Trombka, NASA SP-339, (Washington, D.C.: Government Printing Office, 1973) p. 71.
Neiler, J. H., and Bell, P. R., 1968, "The Scintillation Method" in α, β, and γ Spectroscopy, vol. I, ch. 5, ed. K. Siegbaum (Amsterdam: North-Holland Publishing Co., 1968), p. 245.
Owens, A., et al., 1995, Space Science Review, 71, 273-296.
Pehl, R. H., 1977, Physics Today, 30, 50.
Pehl, R. H., 1987, personal communication.
Pehl, R. H., and Goulding, F. S., 1970, Nucl. Instr. and Meth., 81, 329.
Pehl, R. H., Goulding, F. S., Landis, D. A., and Lenzlinger, M., 1968, Nucl. Instr. and Meth., 59, 45.
Pehl, R. H., Varnell, L. S., Metzger, A. E., 1978a, IEEE Trans. Nucl. Sci. NS-25, no. 1, 409.
Pehl, R. H., Madden, N. W., Elliott, J. H., Raudorf, T. W., Trammell, R. D., and Darken, Jr., L. S., 1978b, Radiation Damage Resistance Reverse Electrode of Coaxial Detectors, LBL-8307 Lawrence Berkeley Lab. Tech. Report.
Price, W. J., 1958, Nuclear Radiation Detector (New York: McGraw-Hill,1958).
Rester, A. C., Eichorn, G., Coldwell, R. L., McKisson, J., Ely, D., Mann, H. M., and Jenkins, D. A., 1986, The GRAD Gamma-Ray Spectrometer, IEEE Trans. Nucl. Sci., NS-33, 732-734.
Rester, A. C., Coldwell, R., Trombka, J. I., Starr, R., Eichhorn, G., and Lasche, G., 1990, Performance of Bismuth Germanate Active Shielding on a Balloon Flight over Antarctica, IEEE Trans. Nucl. Sci., 37, 559-65
Ryan, J. M., Dayton, B., Moon, S. H., Wilson, R. B., Zych, A. D., and White, R. S., 1977, J. Geophys. Res., 82, 3593.
Sandie, W. G., Nakano, G. H., Chase, L. F. Jr., Fishman, G. J., Meegan, C. A., Wilson, R. B., Paciesas, W. S., Lasche, G. P., 1988, ApJ, 334, L91.
Schönfelder, V., and Lichti, G., 1975, J. Geophys. Res. Space Physics, 80, 3681.
Schönfelder, V., Graser, U., and Daugherty, J., 1977, ApJ, 217, 306.
Schönfelder, V., et al., 1993, ApJS, 86: 657-692.
Seltzer, S. M., 1975, Nucl. Instr. and Meth., 127, 293.
Sherman, A., 1983, NASA Needs and Trends in Cryiogenic Cooling, in "Refrigeration for Cryogenic Sensors" , ed. M. Glasser, NASA Conf. Publ. 2287, NASA/GSFC.
Sietz, F., 1940, The Modern Theory of Solids (New York: McGraw-Hill, 1940).
Squillante, M., 1996, private communication.
Swank, R. K., and Buch, W. L., 1952, Nucleonics, 10, 51.

Teegarden, B. J., McDonald, F. B., Tramos. J. H., Webber, W. R., and Roelof, E. C., 1974, J. Geophys. Res., 79, 3615.
Teegarden, B. J., Cline, T. L., 1980, High-Resolution Spectroscopy of two Gamma-Ray Bursts in 1978 November, ApJ, L67-L70.
Teegarden, B. J., et al., 1985. The Gamma-Ray Imaging Spectrometer (GRIS), 19th Intern. Cosmic Ray Conf. (La Jolla), 3, 307-310.
Trombka, J. I., Metzger, A. E., Arnold, J. R., Matteson, J. L., Reedy, R. C., and Peterson, L. E., 1973, ApJ, 181, 737.
Trombka, J. I., Vette, J., Stecker, F. W., Eller, E. L., and Wildes, W. T., 1974, Nucl. Instr. and Meth., 117, 99.
Trombka, J. I., et al., 1976, ASTP Preliminary Science Report, NASA TMX-58173.
Trombka, J. I., Dyer, C. S., Evans, L. G., Bielefeld, M. J., Seltzer, S. M., and Metzger, A. E., 1977, ApJ, 212, 925.
Trombka, J. I., er al., 1997, acepter for publication in the J. Geophys. Res.
Vinogradov, A. P., Surkov, Y. A., Chernov, G. M., Kirnozov, F. F., and Nazarkina, G. B., 1968, Moon and Planets, vol. 2, ed. A. Dolifus, (Amsterdalll: North-Holland Publishillg Company, 1968),19. 77.
Wang, W. J., Iwanczk, J. S., and Patt, B. E, 1994, "New Concepts for Scintillator/HgI_2 Gamma Ray Spectroscopy", *I*EEE Trans. on Nuc. Sci., Vol. 41, No. 4.
White, R. S., Dayton, B., Moon, S. H., Ryan, J. M., Wilson, R. B., and Zych, A. D., 1977, ApJ, 218, 920.
Winkler, C., 1994, ApJS, 92, 327.
Womack, E. A., and Overbeck, J. W., 1970, J. Geophys. Res., 75, 1811.
Wright, G. T., 1954, J. Sci. Instr., 31, 462.

CHAPTER 13

DETECTORS FOR ENERGIES GREATER THAN 25 MeV

13.1 Introduction and History

As noted in the last chapter, the detector telescopes used in gamma-ray astronomy are more similar to particle detectors than to optical devices, since the high frequency of the radiation precludes the use of reflection or diffraction techniques, but the high-energy content of each photon does enable them to be detected with scintillators, track imaging chambers, and solid-state detectors. Within the gamma-ray range, the basic design of the instrument changes as the energy of the gamma ray exceeds 10 to 30 MeV, since in that energy range, the region where the Compton effect predominates, the absorption of the gamma ray ends, and the electron pair production becomes most important. The specific energy of the transition depends on the material in which the interaction occurs. The electron pair process is relatively attractive from a gamma-ray detector point of view, since essentially all of the gamma-ray energy goes into two charged particles in one interaction, and a clear signature of the gamma ray can be recorded. Thus, in the energy range from about 25 MeV to several times 10^4 MeV, gamma-ray telescopes are now built in such a way that the electron pair may be seen and the properties of the electrons measured.

As the gamma-ray energy approaches 10^5 MeV, the intensities of celestial gamma rays become too low to be seen with space telescopes. At about the same energy, and above, individual photons may be observed by telescopes on the ground at night. This detection is accomplished by recording the atmospheric Cerenkov radiation, emitted by relativistic charged particles that are produced as secondaries in the electromagnetic cascade resulting from the interaction of the energetic photon with the atmosphere.

This technique is only useful for examining source regions that are quite small in size, since the observations must be made against a high background of Cerenkov-radiation light pulses, produced by the isotropic, charged-particle, cosmic radiation. The photon component from a source region will then appear as a spatial anisotropy in the light pulse distribution.

A number of factors have combined to cause the gamma-ray realm to be the last of the major regions of the celestial electromagnetic spectrum to be examined. The interaction length for a gamma ray in the atmosphere is a small fraction of the total atmospheric depth. At first thought, this may seem strange, since gamma rays can travel through the entire disk of the Galaxy with only about a 1 percent chance of being absorbed, while light can penetrate only about one-tenth of the distance from the Sun to the center of the Galaxy. The explanation lies in the form of the matter. Light is strongly absorbed by the interstellar dust, but is little affected by the gas of our atmosphere. For gamma rays, it is primarily the total amount of matter that is important, although the nuclear charge of the matter is also a factor, and there is almost 10^4 times as much matter along a path from the top of the atmosphere to ground level as there is along a typical path through the disk of the Galaxy. Because the primary gamma rays interact in the atmosphere long before they reach the ground, it was not until the advent of scientific balloons and satellites that observations of the primary radiation could be made directly. Further, within the atmosphere, another problem exists, namely, the background of secondary gamma rays, produced by cosmic rays interacting in the atmosphere. Even on very high-altitude balloons carrying gamma-ray telescopes to residual atmospheric pressures corresponding to only 1.5 to 4 g cm^{-2}, this background makes observations of the very small celestial fluxes extremely difficult. In fact, at balloon altitudes, only a few of the strongest celestial sources are visible above the atmospheric background. Finally, the very low fluxes have forced the development of rather large, sophisticated detectors and have required relatively large observing times. It is the combination of the atmospheric background and the long observing times required that has forced high-energy gamma-ray astronomy to depend on satelite experiments for most of its observations.

It was recognized almost from the beginning that a high-energy gamma-ray telescope had to identify the electron pair unambiguously, if it were to be successful, because the relatively

rare gamma ray has to be clearly separated from the high background of charged-particle cosmic radiation and the charged and neutral secondaries. Further, as was just suggested, the detecting instrument must be designed to record the gamma-ray interactions selectively and have an extremely high efficiency for the rejection of other events. In addition, the detector must be able to provide information on the arrival direction of the gamma ray and a measurement of its energy. Finally, the low flux demands that the detection efficiency must be high.

The earliest experiments developed in the late 1950s and early 1960s were of two basic types: first, oriented nuclear-emulsion stacks and, second, counter telescopes with anticoincidence shields. It was soon realized that celestial gamma-ray intensities were quite low, making simple nuclear-emulsion experiments impractical. However, nuclear emulsions did have the property of being track-imaging detectors, which, in the form of spark chambers, were to become the heart of the gamma-ray telescopes, and they led to upper limits approaching 10^{-3} photons ($E > 100$ MeV) $cm^{-2}s^{-1}$ (Fichtel and Kniffen, 1965). The first of the counter telescopes to be flown in space was a scintillator-Cerenkov counter detector flown on Explorer 11 (Kraushaar and Clark, 1962; and Kraushaar et al., 1965). The detector system was surrounded by a large, very efficient anticoincidence dome to reject charged particles. An improved version of the detector was flown on OSO-3 and led to the first certain measurements of celestial gamma rays. Other scintillator-Cerenkov gamma-ray telescopes were flown on Proton 2 and Cosmos 208 by Bratolyubova-Tsulukidze et al. (1969), on OSO-1 by Fazio and Hafner (1967), and on OSO-3 by Valentine, Kaplon and Badhwar (1969).

It became apparent, however, that more complicated techniques using picture-type, particle-tracking instruments and large, sensitive areas were required to see gamma-ray point sources. Several investigators in the first part of the 1960s turned to spark chambers as the heart of the high-energy gamma-ray detector system. A set of spark chambers provides a high-discrimination, picture-type device, which allows the experimenter to separate the desired electron-positron pairs from the other events which might satisfy the trigger conditions of the telescope. The spark-chamber assembly in a gamma-ray telescope is surrounded by an anticoincidence system, and is triggered by a directional-telescope coincidence signal in the

absence of a signal from the surrounding anticoincidence detector.

Several different types of gamma-ray telescopes were developed and flown on high-altitude balloons. The high-altitude, thin (less than one-thousandth of an inch) polyethylene balloon of large dimensions (10 to 30 million cubic feet in volume when fully inflated at ceiling, although some of even larger dimensions have been flown) proved to be a great asset to gamma-ray astronomy in providing a means to place instruments above most of the atmosphere and to test detectors in an environment similar to space.

Five distinctly different types of image chambers were developed to the point of being included in a gamma-ray telescope. They are the conventional, optical spark chamber, the vidicon system, the sonic spark chamber, the proportional counter, and the multiwire, magnetic-core, digitized spark chamber. Optical spark chambers using cameras and film were built and flown by several groups (Cobb, Duthie and Stewart, 1965; Frye and Smith, 1966; Board, Dean and Ramsden, 1968; and Niel et al., 1969). Vidicon systems were developed for the TD-1 satellite (Voges et al., 1973) and also for balloons (Helmken and Fazio, 1966; Fazio et al., 1968). A vidicon system was also used on GAMMA-1. In this detector, the film is eliminated by using a vidicon tube, which records the spark picture electronically. A somewhat different technique is used in the sonic chamber, which employs microphones to record the position of the spark through accurate timing signals. A chamber of this type was developed for a gamma-ray telescope by Ogelman, Delvaille and Greisen (1966) and flown sucessfully on a balloon. A small sonic chamber was the first spark chamber in space, being flown on OGO-E by Hutchinson et al. (1969). Another gamma-ray telescope including a spark chamber was flown on Cosmos 164 by Volobuev et al. (1969). A 4-gap spark chamber was placed under a one-radiation-length converter, with the main objectives of the spark chamber being to see the shower development. A gamma-ray telescope using proportional counters was developed by Albats et al. (1977) and flown on balloons.

Magnetic-core spark chambers for gamma-ray telescopes were developed at both the Goddard Space Flight Center (Ehrmann et al., 1967; Ross et al., 1969) and the Max Planck Institut für Extraterrestrische Physik (Mayer-Hasselwander et al., 1972), proved on balloon flights, and then flown successfully in the SAS-2 (Derdeyn et al., 1972) and COS-B (Bignami et al., 1975)

gamma-ray telescopes respectively. A magnetic-core spark chamber system was also used in the Energetic Gamma Ray Experiment Telescope (EGRET) on the Compton Gamma-Ray Observatory (CGRO). The SAS-2 satellite provided the first detailed knowledge of the gamma-ray sky and an indication of the ultimate promise of gamma-ray astronomy, while COS-B, following about 3 years later, extended this knowledge. Finally, EGRET provided the great expansion of knowledge described in many of the earlier chapters of this book. These instruments will be described in more detail later in this chapter.

During the same period that gamma-ray experiments were being developed to detect radiation in the 20 to 10^4 MeV range, work was proceeding on the ground to search for gamma rays above 0.1 TeV. The emission of Cerenkov light radiated by electrons in electromagnetic cascades in the atomosphere was discovered by Galbraith and Jelley (1953, 1955), who showed that light pulses from an air shower were detectable above the night-sky background. Attempts were made in the 1960s and 1970s to detect point sources, which could then justifiably be considered gamma-ray sources, since the distribution of arrival directions of primary charged particles in the same energy range that can initiate similar showers, is diffuse. The basic concept of this type of detector is that a large steerable mirror focuses the light, which is then recorded, normally by a set of photomultiplier tubes. Several variations to the basic technique were developed, and these are summarized by Porter and Weekes (1978). Further improvements have been made in detector and analytic techniques in the 1980s and 1990s, and these will also be described later in this chapter. In spite of the very low flux levels in this energy range, it is now accepted that a few very high-energy gamma-ray sources have been detected by this general technique, as was described in earlier chapters.

At energies above 10 TeV, it is possible for the secondary charged particles from a gamma-ray-induced shower to reach sea level. Electromagnetic cascades are expected to have a relatively small number of muons compared to cosmic-ray-nucleon-induced showers; therefore, it was hoped that, in spite of the very low gamma-ray fluxes, a point source might be detectable above the charged particle background. Thus far, no clear evidence for a gamma-ray source detected by this technique exists.

13.2 The High-Energy Gamma-Ray Interaction

The basic interaction of the high-energy gamma ray leading to an electron pair was discussed in Chapter 11. This treatment will be extended somewhat here to emphasize a few points which are particularly relevant to the design of a high-energy gamma-ray detector.

As in the case of radiation processes, the probability of a pair-production interaction per radiation length depends only very slightly on atomic number or Z, particularly at high energies. However, as shown in Figure 13.1, the importance of the Compton

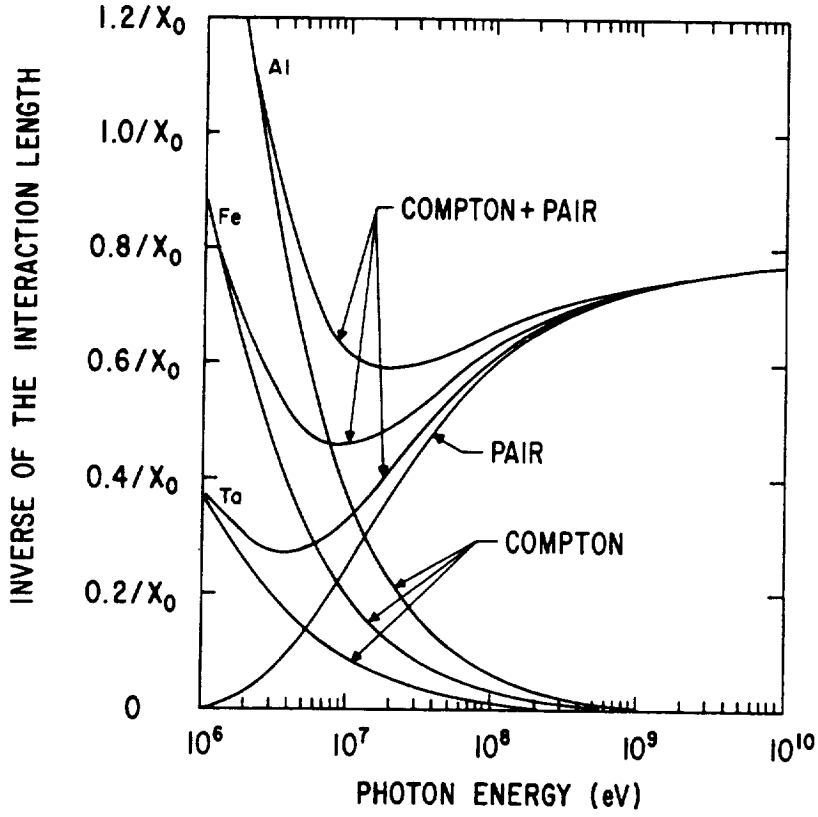

Figure 13.1. The inverse of the interaction length in terms of radiation lengths X_0 for the Compton and pair-production processes for three different materials.

interaction in terms of the interaction probability per radiation length varies strongly with the nuclear charge of the material in which the interaction occurs. This feature results from the Compton interaction being basically one between a photon and an electron, and therefore, the Compton process is proportional to Z rather than to Z(Z + 1). Thus, as the figure shows, the pair-production process can be enhanced relative to the Compton process by using a high Z material in the detector, and this is normally done in high-energy gamma-ray detectors, since the pair-production process is a more desirable one on which to make measurements.

The properties of the gamma ray must, in practice, be determined by measurements on the secondary electrons. (At present, it does not seem practical to measure the very small fraction of the energy and momentum given to the target nucleus.) Of course, it is not possible to measure the electrons at the instant they interact in the detectors or without disturbing them. Hence, it is important to consider what happens to the secondary electrons. As it passes through matter, an electron undergoes a large number of Coulomb collisions, most of which produce very small angle deflections. Both the energy loss and the electron scattering are functions of the radiation length; thus, for a fixed distance measured in radiation lengths, they are unaffected by the selection of material. If the relevant layers in the detector are kept very thin so that energy loss is negligible over the length of the measurement of the electron's direction, and if there has not been a very improbable bremsstrahlung interaction in which a significant amount of energy has gone to a secondary photon, the mean square angle of scattering for an electron is given to a good approximation by the equation

$$< \theta^2 >= \left(\frac{E_o}{\beta c P_e} \right)^2 \frac{x}{X_o}, \qquad (13\text{-}1)$$

where X_o is the radiation length, x is the distance transversed in radiation lengths, and E_o is approximately 21 MeV (Rossi, 1956). The scattering which occurs before the electron trajectories can be determined increases the uncertainty with which the arrival direction of the gamma ray is determined. Therefore, it is desirable to make the individual plates thin.

In attempting to measure the total energy of the photon by stopping the electron in an absorber such as a scintillator of some type, it is

important to remember that not only the electron and the positron must be absorbed, if an accurate energy estimate is to be obtained, but also the secondary bremsstrahlung radiation, and particularly at low energies, the two 1/2 MeV photons formed when the positron comes to rest and annihilates with an electron in the material. An alternate method for estimating the energy is to measure the average Coulomb-scattering angle of each electron, but this is only practical over a very limited energy range.

At very high energies, the emission from the cascade shower, produced by the gamma ray in the atmosphere, is used to study the properties of the gamma-ray radiation. When a very high-energy photon interacts in the atmosphere and produces an electron pair, the secondary very high-energy electrons interact, in turn losing most of their energy through radiation. The photons in the gamma-ray region interact through either the pair formation or the Compton process. This sequence continues, and an electron-photon cascade develops. At each step, the number of electrons and photons increases, and their average energy decreases until the electrons' energies are sufficiently low that collision losses become more important than radiation. Their energy then goes largely into ionization and excitation of atoms. The secondary particles are heavily absorbed, and detection of the shower particles is difficult. Although in principle it should be possible to detect gamma rays above about 10 TeV, the separation of the gamma ray from the enormously greater number of primary charged particles is very complex. Attempts have been made to do so (O'Sullivan et al., 1978).

The Cerenkov light radiated by the electrons in the shower is not absorbed and may be detected on the ground on moonless, clear nights at a dark site. Whereas the gamma-ray-shower Cerenkov light still has to compete on essentially equal terms with the much more numerous cosmic-ray-electron-shower Cerenkov-light events, the cosmic-ray nucleons are somewhat less efficient in the production of Cerenkov light. The flux of optical Cerenkov photons from an air shower of even a 0.1 TeV photon is modest, being only about 10^3 photons m^{-2} within 100 to 200 m of the shower axis (Porter and Weekes, 1978); nonetheless, as will be described in the next section of this chapter, it is possible to obtain useful results from this process in the TeV region.

Thus, there are two quite different approaches to the detection of gamma rays above 25 MeV, depending on the energy range, and these will be treated separately in the next two sections.

13.3 The 20 MeV to 50 GeV Region

The SAS-2 gamma-ray telescope was the first satellite to incorporate all of the basic concepts of the space telescopes which are now looked upon as the essentials of a successful, low-background, high-energy gamma-ray instrument. The COS-B satellite high-energy gamma-ray telescope was quite similar in size and concept, although there were some differences that will be described here, including a calorimeter to estimate the energy. The most recent space high-energy gamma-ray telescope, EGRET, is also based on the SAS-2 concept with the major differences being its much larger size and its capability to measure the energy of the gamma ray with reasonable accuracy over the entire energy of the telescope.

The first of the satellites just mentioned, SAS-2 (Derdeyn et al., 1972), was launched on November 15, 1972. A schematic diagram of that gamma-ray instrument is shown in Figure 13.2. The spark chamber assembly consists of 16 spark chamber modules above a set of four central plastic scintillators and another 16 modules below these scintillators. Thin tungsten plates, averaging 0.010 cm thick –corresponding to 0.03 radiation lengths– are interleaved between the spark chamber modules. In the upper half of the spark-chamber assembly, these plates provide material in which the gamma ray may interact and produce an electron pair.

The combination of the plates and spark chambers in the upper and lower halves provides a means of determining the energy of the electrons in the pair, by measuring the average Coulomb scattering as they pass through the plates. The plates in the lower half of the spark chamber assembly also provide additional pictorial information used to identify the gamma ray.

The spark chamber assembly is triggered if a charged particle passes through one of the four square plastic scintillator tiles and the corresponding directional lucite Cerenkov counter immediately below, while, at the same time, there is no pulse in the surrounding plastic-scintillator-anticoincidence dome. Each of the four scintillator-Cerenkov counter telescopes acts independently of the others and has a full-width half-maximum (FWHM) opening angle

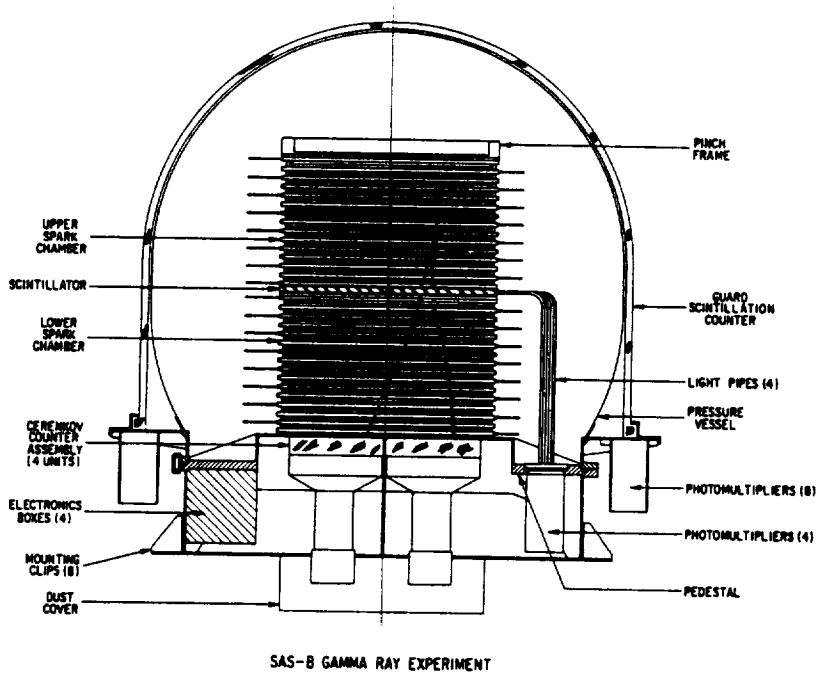

Figure 13.2. A diagram of the gamma-ray telescope flown on SAS-2 (Derdeyn et al., 1972).

of about 30°. The anticoincidence dome prevents the spark chamber system from being triggered by charged particles, and the directional feature of the Cerenkov counter prevents the telescope from being triggered by upcoming particles that might stop above the central scintillator before reaching the anticoincidence dome.

Of the detector subsystems used in this gamma-ray telescope, the spark chamber is one that was not described in Chapter 11; so it deserves attention here. In general, a spark chamber consists of two planes to which a high-voltage differential may be applied, causing an electrical breakdown along the path or paths of charged particles that have passed through very recently and left an ionized path. Normally, the voltage is applied within a few tenths of a microsecond, and the rise time of the voltage pulse is as short as possible. The high voltages cause a free electron to produce an avalanche, which, after reaching a critical size, leads to a

breakdown. The voltage pulse height is chosen to be high enough to give a high efficiency (typically 96 percent) and good multiple-track efficiency, but not so high as to produce undesired track spreading. It is also advantageous to keep the pulse profile short.

As noted earlier, there are several types of readout systems for spark chambers besides the magnetic core one that was used on SAS-2, COS-B, and EGRET. The magnetic core system, however, has proved to be desirable for gamma-ray space telescopes for several reasons, including its inherently rugged nature, the capability for recording multiple tracks with high efficiency, its good single-track position accuracy, the capability for resolving two close tracks, its high reliability, its low power consumption, the reasonable cost, and its two-dimensional readout. In the magnetic-core spark chamber there are two planes of wires replacing the plates, with the wires in one plane being orthogonal to those in the other. The individual wires are typically 1 mm apart, and each threads a magnetic core. When a spark occurs, the current flows down one wire, through the spark, and onto an orthogonal wire in the other plane, setting the core on each wire, thereby giving the "x,y" position of each track at a given level. The third dimension in the picture is provided by the vertical assembly of spark chambers. The information contained in the set cores can be read out in the usual way by threading each core with two additional wires: one, a drive wire, which can be used to switch the core back to its "unset" state, and the other, a sense wire, which sees an induced voltage when a core is switched from one state to another. The cores are read out and reset in one operation that can occur very quickly, so that the spark chamber is ready to be triggered again as soon as the capacitor in the high-voltage pulser is recharged, and the satellite system is ready to receive data.

The information on spark locations is stored in a separate memory until it can be read out and transmitted by the satellite data and telemetry system. Since, typically, only one percent or less of the cores are set, it is more efficient to store the location of each set core rather than to know whether each core is set or not. There are also additional ways of compressing the data further. In a gamma-ray event, there are, of course, two tracks coming from a common vertex. It is desirable to know which track in one view corresponds to which track in the other. It is usually possible to determine the association on the basis of the track characteristics, such as Coulomb scattering or an occasional missing spark on a specific deck; however, to assist in correlating the two views, often a grid

with the wires at 45° with respect to those in the other wire planes is placed at the bottom of the spark chamber stack assembly.

The main constituent of the spark-chamber gas is normally neon. It provides a very good multiple track efficiency as compared, for example, to argon, which has a large breakdown time and, hence, larger fluctuations in spark-formation time, or oxygen, in which the first fired electron from an oxygen molecule normally leads immediately to the only spark seen. A quenching agent is usually added to the gas to absorb ultraviolet light, thereby reducing spurious streamers and keeping the track narrow. Excessive amounts of the quenching agent will reduce efficiency; so, typically, one-half to one percent of a gas such as alcohol or ethane is used. Often a small amount (again one-half to one percent) of argon is added to take advantage of the Penning effect. This concentration allows enough collisions to occur to take advantage of the fact that the ionized states of argon lie below those of neon, so that there is an enhancement of the number of electrons available during an avalanche. Gas purity is not as critical as in proportional chambers, but even very small quantities of electronegative gases must be rigorously avoided. For more detailed information on the theory of spark chambers and proportional counters, see the book by Peter Rice-Evans (1974).

In the case of SAS-2, each spark chamber module had a 25 by 25 cm active area, with 200 wires in each plane. The two planes were separated by 3.6 mm. The material chosen for construction of the module was glass-bonded mica, selected for its low outgassing property, high dielectric strength, yield strength, and relative ease of machining as compared to most ceramics. (A ceramic called Macor was used in EGRET). The spark wires were made of 0.10 mm diameter beryllium-copper.

The effective area of the SAS-2 gamma-ray telescope, which was limited by the size of the Cerenkov counters, was 540 cm^2. The opening angle for detection of gamma rays was approximately 1/4 sr. The efficiency for detection of gamma rays was a function of energy, with a value at very high energies of 0.22, and at 100 MeV of about 0.12. The efficiency and solid angle as a function of energy were determined by calibration at the National Bureau of Standards Synchrotron and the Deutsches Elektronen-Synchrotron (DESY). Timing accuracy for the arrival of each gamma ray to

better than 2 milliseconds was obtained for the study of pulsed gamma-ray sources.

By combining the energy and directional information for each electron, the direction and energy of the primary gamma ray was obtained. The gamma-ray threshold for useful information was not sharp, but was about 35 MeV. The energy of the gamma ray could be measured to about 200 MeV; above that energy, the integral flux could be determined.

SAS-2 was launched from San Marco off the coast of Kenya into an approximately circular orbit with a 2° inclination and a 550 km altitude. The satellite was capable of being pointed in any direction, and viewed a particular region of the sky for about 1 week. Hence, typically, for approximately 0.36 of the orbit, the detector pointed at the Earth, and for another approximately 0.08 of the orbit, the Earth albedo gamma-ray flux was high, leaving about 0.56 of the orbit for collection of celestial gamma-ray data. Combined with a live time (the period when cores are not being read out and the gamma-ray telescope is ready to accept another event) of about 90 percent, the portion of an orbit during which celestial data were collected was about 0.5.

The COS-B gamma-ray telescope (Bignami et al., 1975), launched in August 1975, is similar in concept to the one flown on SAS-2. A schematic diagram of the COS-B instrument is shown in Figure 13.3. Its area and sensitivity were about the same as those of SAS-2. Having additional weight capability, the COS-B satellite was able to carry an energy-measuring crystal that allowed a measurement of the gamma-ray energy to high energies, thereby permitting an extension of the measured energy spectrum. It did not contain the equivalent of the SAS-2 lower chamber, which proved to be an advantage in the data analysis. Another major difference was that the COS-B satellite had a very high apogee, and, consequently, almost all of the orbit was available for viewing, because occultation by the Earth prevailed during only a small fraction of the orbit. The disadvantage was that the instrument saw the full cosmic-ray intensity, which created a significantly higher gamma-ray background from interactions in the thin thermal covering layer, which, for measurement of the diffuse background, was not entirely negligible, as it had been for SAS-2. Finally, the very long life of the COS-B experiment permitted much longer exposures and, hence, markedly improved statistics.

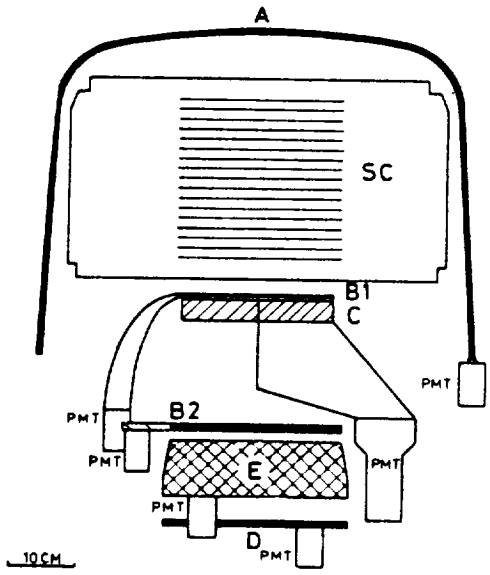

Figure 13.3. A schematic diagram of the gamma-ray telescope flown on COS-B. The figure was supplied by the Caravan collaboration for COS-B, consisting of Space Science Department, ESTEC, Noordwijk, The Netherlands; University of Leiden, The Netherlands; Centre d'Études Nucléaires de Saclay, France; University of Milan, Italy; University of Palermo, Italy; Max-Planck Institut für Extraterrestrische Physik, Garching, FRG; and is reproduced with the permission of Mayer-Hasselwander on behalf of this collaboration.

The GAMMA-1 experiment was a joint effort of four Soviet and two French laboratories. It was also similar in basic concept to the SAS-2 instrument. The sensitive area was about 26 times greater; the area solid-angle factor was about the same, because the viewing angle is smaller. It had an energy-measuring calorimeter which should be able to measure energies with significantly better accuracy than the energy-measuring element of COS-B. The upper spark chamber system was a 12-level, wide-gap Vidicon system. With the aid of a mirror system, the picture of the sparks is recorded and digitized using a Vidicon tube. The directionality of the electron was determined by a time-of-flight system rather than a directional Cerenkov counter. Regrettably, the high-voltage system was found not to be functional after launch, and no pictorial gamma-ray information could be obtained.

The time-of-flight system approach to the directional measurement was also used in the high-energy gamma-ray telescope EGRET, flown on the CGRO. Because of the space required, it could not have been incorporated in the earlier missions, even if the low-power space-flight quality electronics had been developed. The time-of-flight system provides both the coincidence telescope information and the directionality information. The basic unit consists of two charged-particle detectors, which, in this application, are sheets of plastic scintillator separated by a distance large enough to measure the time of flight of a relativistic electron with sufficient accuracy to determine clearly whether it is traveling upward or downward. The time-of-flight technique has proved to be an order of magnitude more efficient in rejecting undesired events than the previously used directional Cerenkov systems. The active area is approximately 80 x 80 cm, and the solid angle is about 1/2 steradians. EGRET, launched on CGRO in April of 1991, has 10 times the area, more than 20 times the area solid-angle factor of SAS-2 or COS-B, and good energy resolution to the highest energies observable. A schematic diagram of this instrument is shown in Figure 13.4 (Fichtel et al., 1978).

The upper chamber consists of 28 magnetic-core spark chamber modules interleaved with 0.02 radiation length plates. The directional time-of-flight coincidence system employs two 16-element layers of plastic scintillators, separated by 60 cm, and has a total of 96 possible coincidence combinations. The energy of the gamma ray is usually determined from measurements made in an 8-radiation length thick, 76 x 76 cm square NaI(Tl) scintillator crystal below the time-of-flight scintillator plane. The energy resolution of the proposed experiment is about 15 percent FWHM. In the lower-energy range, the energy information from the crystal may be supplemented by multiple Coulomb-scattering observations in the spark chamber assembly. The accuracy with which a source can be located depends on the energy-spectrum shape, the intensity, and the location of the sources; for a strong source with a typical spectrum, source locations of 5 to 10 arcmin are achieved. Weak, but detectable, sources are typically determined to 15 to 45 arcmin. For detailed information on the characteristics of EGRET and its calibration see Kanbach et al. (1988) and Thompson et al. (1993).

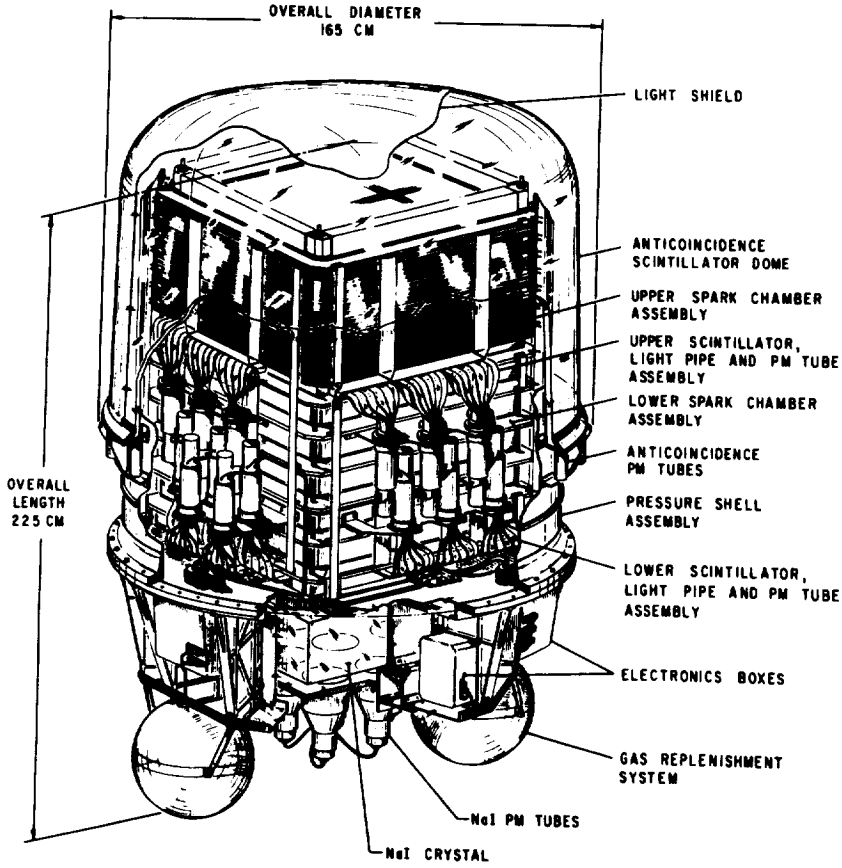

Figure 13.4. A schematic diagram of the gamma-ray telescope flown on CGRO (Fichtel et al., 1978).

Several types of track imaging detectors are now being developed with improved spatial resolution. It is hoped that, when one of these is developed for space flight and can be produced at a reasonable cost, these track-imaging planes can be incorporated into a multi-layer, very-thin-plate telescope of the type described here. This telescope would improve the positions of sources and permit better spectra in the lower part of the high-energy gamma-ray region. However, as discussed in Section 13.6 and Chapter 11, the

lack of knowledge of the recoil nucleus in the initial interaction still severely limits the angular accuracy that can be obtained in high-energy gamma-ray astronomy.

13.4 The Greater Than 20 GeV Region

As mentioned earlier in this chapter, at very high energies photons can be detected by instruments at sea level that record the Cerenkov light produced in the atmosphere from a series of interactions initiated by a single incident gamma ray. Figure 13.5 shows the original 10-meter optical reflector of the Smithsonian Astrophysical Observatory, Mt. Hopkins, Arizona. The reflector is a Davies-Cotton design, consisting of 248 individual, identical mirrors, each with a 14.6 radius of curvature, positioned in a hexagonal pattern covering a spherical bowl of radius 7.3 m with an opening diameter of 10 m. This instrument originally had a threshold of about 0.5 TeV; this was later improved to 0.3 TeV. This design has several advantages including the following: (1) the alignment procedure is straightforward, (2) the overall structure is rigid, and (3) off-axis aberrations are reduced. The principal disadvantage is that the light-travel times from the different portions of the mirror to the focal plane are different. The variations in the light-travel time are as much as six ns from the shortest to the longest path. Although this feature does not directly affect the optical quality, the integration time must be larger, and, hence, there is more noise. For further details, see, for example, Cawley et al. (1990) and Lewis (1990).

Besides the one just described, there are several other ground-based atmospheric Cerenkov gamma-ray telescopes. Others in existence by the summer of 1996 include Mark 6 and CANGAROO in Australia, CAT, ASGAT, and Themistocle in Themis, Shalon-Alatoo in Kazahkstan, and GEGRA in Canaries. Their thresholds vary from 0.2 to 3 TeV.

These instruments examine a region of the sky where a source is suspected, and an attempt is made to detect a directional anisotropy that is statistically significant. Several techniques have been used historically to enhance the sensitivity, primarily by improving the signal-to-noise ratio. Recently, advantage has been taken of the

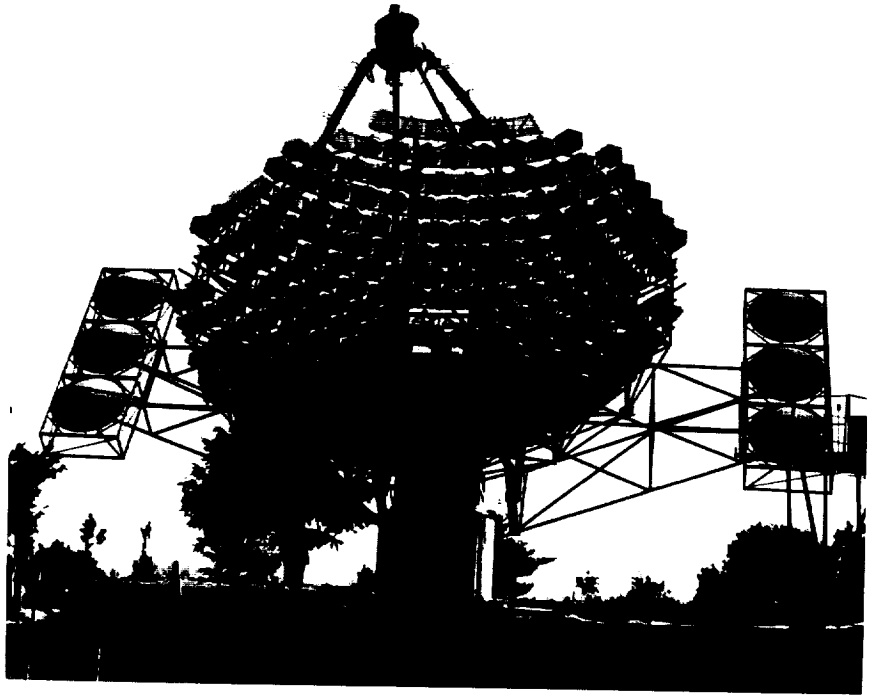

Figure 13.5. Ten-meter optical reflector of the Smithsonian Astrophysical Observatory at Mt. Hopkins (reproduced with permission from Weekes and the Smithsonian Astrophysical Observatory).

characteristically narrow transverse dimensions of the gamma-ray showers relative to their length, as compared to the hadronic showers. Thus, the selection by the nature of the image adds to the selection by direction, i.e., the gamma-ray sources are pointlike as compared to the nearly isotropic hadronic events. For a further discussion of this analysis technique, see Weekes et al. (1989) and Akerlof et al. (1989). This technique did lead to the convincing detections in this energy range discussed in earlier chapters, as well as casting serious doubt on many previously claimed marginal sources.

A major step forward on the ground which will lower the threshold to 20 GeV now seems within reach. The way to achieve this large reduction in the energy threshold is to dramatically increase the collecting area. To increase the effective sensitive area by the required amount using conventional approaches would be very complicated and expensive. Fortunately, facilities exist that can be adopted to achieve the desired goal; these are the large-collecting-area solar-power plants built some time ago. In these large arrays, sunlight, collected with hundreds of mirrors tracking the sun (called heliostats), is reflected unto a central receiver at the top of a tall tower, where the heat is used to generate electrical power. Since this system has tracking capability, a central collecting system, and a large light-collecting area, it has the basic elements needed for a 20 GeV atmosperic Cerenkov gamma-ray telescope. At least two groups (SOLAR-2 AND CELESTE) have been studying an approach to adapting these arrays to very-high-energy gamma-ray astrophysics. It appears now that the CELESTE collaboration (Dumora et al., 1996) will have the system at Themis in operation sometime during the autumn of 1997.

To convert the solar array to a gamma-ray telescope for CELESTE, the heat receiver in the top of the tower is removed and replaced by photodetectors. In particular, a large collection of phototubes is used and placed in such a way that each tube views one and only one heliostat. In this way, the differences in light travel time can be compensated for electonically, thereby maximizing the signal-to-noise ratio. The shape and sharpness of the signal may be used to discriminate against hadronic showers. The expected sensitivity in the 20 to 50 GeV region will substantially exceed that of EGRET. Although the region of the sky that can be examined is limited at the threshold, it rapidly expands, so that by 30 GeV it should include all declinations from about 0° to 55°, and, by 50 GeV, it should cover approximately –5° to 75°.

13.5 Angular Resolution

In the 20 MeV to 50 GeV energy range, where imaging of the electron-positron pair is used in the detection process, the angular uncertainty within which a gamma-ray arrival direction can be determined depends on four factors. The first involves the angular error introduced in the initial interaction between the gamma ray and

the detector material and was discussed in Chapter 10. The second factor, discussed earlier in this chapter, is related to the subsequent Coulomb scattering of the electrons that occurs in the material in the distance over which the electron directions are measured. The third related consideration is the accuracy with which the positions of the electrons can be measured and their relative energies determined. Finally, there is the matter of the accuracy of alignment of the telescope, its alignment to the aspect sensors, and the inherent accuracy of the aspect sensor. This last item is not a determining one, since accuracies well beyond those of the first three factors are readily achievable. The second and third terms are rather complex in general; however, for a multiple thin-plate, track-imaging chamber, a reasonable approximation may be introduced. With these approximations, the angular uncertainty, θ_u , is given by the expression

$$\theta_u \cong \left\{ \theta_{\gamma,ee}^2 + \left[A_1 \frac{M_l^{\frac{1}{3}} L^{\frac{1}{3}}}{E} \right]^2 + \left[A_2 \frac{B}{L} \right]^2 \right\}^{\frac{1}{2}}, \qquad (13\text{-}2)$$

where $\theta_{\gamma,ee}$ is the angle of uncertainty introduced by the lack of knowledge of the direction and energy of the recoil nucleus in the initial interaction, M_l is the mass per unit length, L is the length over which the measurement is made, the A_i's are constants, and B is a measure of the position accuracy. B is best determined experimentally, since it depends on track-location accuracy and the ability to separate two tracks. Another consideration is the ability to measure the relative energy of the electrons, since, for those events where the best choice of length occurs when two tracks exist, the best value of "θ_u " is the weighted average of two separate determinations, one for each electron.

If the expression on the right side of Equation (13-2) is minimized by selecting the optimum choice of L, which can be shown to be proportional to $B^{2/3}E^{2/3}M_l^{-1/3}$, the following form is obtained:

$$\theta_u \cong \left\{ \theta_{\gamma,ee}^2 + \left[A_4 \frac{M_l^{\frac{1}{3}} B^{\frac{1}{3}}}{E^{\frac{2}{3}}} \right]^2 \right\}^{\frac{1}{2}}, \qquad (13\text{-}3)$$

where A_4 is a constant. It is important to remember that this equation is an approximation and as one approaches small values, the exact form must be used. Note that the second term varies most strongly with energy, and remember, from Chapter 10, that

$$< \theta^2_{\gamma,ee} >^{\frac{1}{2}} = q \frac{mc^2}{E_\gamma} \ln\left(\frac{E_\gamma}{mc^2}\right), \qquad (13\text{-}4)$$

where q is a weakly varying function. Hence, the higher-energy gamma rays play an important role in obtaining the best possible angular accuracy. It is, therefore, important to identify these high-energy gamma rays through a good energy measurement. Unfortunately, from the point of view of angular accuracy, the number of high-energy gamma rays is much smaller.

With new detector developments it now seems possible to reduce greatly the error in the position measurement. Coupling this improvement with a modest decrease in the mass per unit length, it should be possible to come within a factor of about $1\frac{1}{2}$ of the interaction limit determined by the first term of Equation (13-2).

Two other considerations should be mentioned. First, a major improvement could be made if information could be obtained on the initial interaction. In practice, accomplishing this involves detecting and measuring the direction of the recoil nucleus, something that, as yet, no one knows how to achieve for a practical gamma-ray telescope. Second, an active collimator or coded-aperture technique, as used at low energies, has been suggested. However, this approach seems to have a major difficulty. High-energy gamma rays are very penetrating; so the mask would have to be very thick. This heavy collimator would almost certainly be the source of a great background that would be very difficult to discriminate against. Remember that, in a high-energy gamma-ray instrument, one is trying to measure a very weak intensity in a very high background of charged particles and other radiation, and the major reason for the advances of the SAS-2 and subsequent high-energy gamma-ray satellite telescopes has been the suppression of the background.

Thus far, the discussion has been limited primarily to single photons. A source can be located more accurately by measuring the distribution of estimated arrival directions of a large number of gamma rays, and determining the error to which the centroid may be

determined. The additional considerations that enter are the source strength, the instrument sensitivity, the diffuse celestial gamma-ray intensity in the region of interest, the source energy spectrum, and the viewing time. With EGRET, 5 to 10 arcmin was possible for the stronger sources, as noted earlier.

In the region above 0.1 TeV, the accuracy with which the direction of a gamma ray can be determined depends on the nature of the development of the shower and how well the light distribution can be determined. Initially, the source could be determined to about 1°. With the use of the technique involving the shape of the image, this was reduced to about 0.1°, and hence similar to EGRET. CAT has an angular resolution of 0.1°. CELESTE should achieve 0.2° in spite of the challenge of proceeding to a 20 GeV threshold.

13.6 Energy Measurement and Resolution

The earliest experiments in the lower part of the high-energy range depended primarily on measuring the Coulomb scattering of the electrons formed in the gamma-ray interaction to obtain an estimate of the gamma-ray energy, or even just on the opening angle of the pair alone. Considerations related to a practical detector having a fixed size and having to meet basic requirements led to a range of about 5 to 10 over which a fair estimate of the energy can be made from the Coulomb scattering measurements. The SAS-2 experiment, for example, was able to estimate the gamma-ray energy to about 30 to 40 percent Root Mean Square (RMS) from 35 MeV to 200 MeV. Below that energy range, the secondary electrons were increasingly lost to scattering from the sides or to absorption. At the high-energy end, a point is reached where a meaningful measurement is increasingly difficult to make, as the noise of the measurement begins to dominate. If the plates are made thicker to obtain meaningful scattering measurements at higher energy, the effective low-energy threshold rises accordingly.

As payload weight restrictions lessened, detectors were added at the bottom of the telescopes to estimate the energy. As was noted earlier, even a relatively simple shallow-energy detector, such as that flown on COS-B, enhanced the experiment significantly, because, although the accuracy in the 35 to 200 MeV energy range was about the same, the energy-absorbing crystal allowed measurements of similar accuracy to be made into the GeV range. The error results

principally from escape of a variable fraction of the energy from the bottom and sides of the energy detector, and, to a lesser degree, from the inaccuracies associated with the measurements of the energy deposited in the energy-measuring device.

A significant improvement was achieved by using a larger and thicker, but still fully active, energy-absorption counter. For EGRET there is an 8-radiation length thick NaI(Tl) crystal covering the entire area. The energy resolution provided by this crystal is about 8% to 10% RMS (or 20% to 25% FWHM) for gamma-ray energies less than a few GeV and increases with energy above that (Thompson et al., 1993). In addition to energy-leakage fluctuations, which decrease with decreasing gamma-ray energy, the realizable energy resolution in the telescope itself is influenced by two other instrumental factors in the range below 1 GeV. These are (1) the presence of about 0.6 radiation length of inert material above the crystal, mostly in the form of the thin plates between the spark chambers, in which the electron-positron pair deposits a variable amount of energy, and (2) the possiblity that one of the secondary particles either is scattered in this inert material and strikes the crystal close to its edge or misses it entirely. In the event analysis, however, partial corrections can be made for both of these effects. The spark chamber reveals the gamma-ray conversion point and the subsequent track lengths for the secondary particle pair through the inert material for each event, and a correction can be made for the mean energy lost in this material below the conversion point. For a particle that fails to hit the crystal, an energy estimate can be made by observing its scattering in the upper and lower spark chamber arrays. The influence of both these effects on the energy resolution is determined by calibration at an accelerator facility.

In the very-high-energy range, the determination of the energy estimates is more uncertain. Initially, only the threshold energy was estimated. More recently, estimates of the energy of the individual events have been made. The energy estimate is ususally made from the total number of counts registered in the analog-to-digital converters. The relationship between the number of counts and the estimated energy is determined by Monte Carlo simulated showers. See Lewis et al. (1993). Given a specific input energy to the Monte Carlo program, both an energy and a deviaton can be determined for each trial, and, with a large number of trials, a mathematical function for the total number of counts as a function of energy and the RMS deviation can be determined. Whereas the estimated RMS error is

only of the order of 10%, it must be remembered that this is based on the Monte Carlo program and not on an actual calibration.

References

Akerlof, C. W., et al., 1989, in Prod. of the Gamma Ray Observatory Workshop (Greenbelt, NASA), 4, 49.

Albats, P., Denchy, B., Frye, G. M., Jenkins, T. L., Koga, R., and Schindler, S. M., 1977, Proc. of the 12th ESLAB Symp., ed. R. Wills and B. Battrich, WSA SP-124, 287.

Bethe, H. A., and Heitler, W., 1934, Proc. Roy. Soc., A146, 83.

Bignami, G. E., Boella, G., Burger, J. J., Keirle, P., Mayer-Hasselwander, H. A., Paul, J. A., Pfeffermann, E., Scarsi, L., Swanenburg, B. N., Taylor, B. G., Voges, W., and Wills, R. D., 1975, Space Sci. Inst., 1, 245.

Board, S. J., Dean, A. J., and Ramsden, D., 1968, Nucl. Instr. and Meth. 65, 141.

Bratolyubova-Tsulukidze, L. I., Grigorov, N. L., Kalinkin, L. F., Melioransky, A. S., Pryakhin, E. A., Savenko, I. A., and Yufrakin, V. Y., 1969, Proc. of the 11th Int. Cosmic Ray Conf., 1, 123.

Cawley, M. F., et al., 1990, Experimental Astrophysics, 1, 173.

Clark, G. W., Farmire, G. P., and Kraushaar, W. L., 1968, Can. J. Phys., 46, S414.

Cobb, R., Duthie, J. G., and Stewart, J., 1965, Phys. Rev.Letters, 15, 507.

Derdeyn, S. M., Ehrmann, D. H., Fichtel, C. E., Kniffen, D. A., and Ross, R. W., 1972, Nucl. Instr. and Methods, 98, 557.

Dumora, D., et al., 1996, Cerenkov Low Energy Sampling and Timing Experiment, CELESTE, Experiment proposal.

Ehrmann, C. H., Fichtel, C. E., Kniffen, D. A., and Ross, R. W., 1967, Nuc. Instr. and Methods, 56, 109.

Fazio, G. G., and Hafner, E. M., 1967, J. Geophys. Res., 72, 2452.

Fazio, G. G., Helmken, H. F., Cavrak, Jr., S. J., and Hearn, D. R., 1968, Can. J. Phys., 46, S427.

Fichtel, C. E., and Kniffen, D. A., 1965, J. Geophys. Res. 70, 4227.

Fichtel, C. E., Hofstadter, R., Pinkau, K., Bertsch, D. L., Favale, N., Hartman, R. C., Hughes, E. B., Kniffen, D. A., Mayer-Hasselwander, H. A., Rothermel, H., Schneid, E., Sommer, M., Thompson, D. J., 1978, Proposal for a High Energy Gamma Ray Telescope on the Gamma Ray Observatory.

Frye, G. M., and Smith, L. H., 1966, Phys. Rev. Letters, 17, 733.

Galbraith, W., and Jelley, J. V., 1953, Nature, 171, 349.

Galbraith, W., and Jelley, J. V., 1955, J. Atmos. Terr. Phys., 6, 250.

Helmken, H. F., and Fazio, G. G., 1966, IEEE Trans. Nucl. Sci., NS-13, 486.

Hughes, E. B., et al., 1980, IEEE, Trans. Nucl. Sci., NS-27, 364.

Hutchinson, G. W., Pearce, A. J., Ramsden, D., and Wills, R. D., 1969, IAU Symp. 37.

Kanbach, B., et al., 1988, Space Sci. Rev., 49, 69.

Kraushaar, W. L., and Clark, G. W., 1962, Phys. Rev. Letters, 8, 106.
Kraushaar, W., Clark, G. W., Garmire, G., Helmken, H., Higbie, P., and Agogino, M., 1965, ApJ, 141, 845.
Lewis, D. A., 1990, Experimental Astronomy, 1, 213.
Lewis, D. A., et al., 1993, Proceedings of the Int. Cosmic Ray Conference (Calgary), 1, 279.
Mayer-Hasselwander, H. A., Pfeffermann, E., Pinkau, K., Rothermel, H., and Sommer, M., 1972, ApJ, 175, L23.
Niel, M., Cassignol, M., Vedrenne, G., and Bouigue, R., 1969, Nucl. Instr. and Meth., 69, 309.
Ogelman, H. B., Delvaille, J. P., and Greisen, K. I., 1966, Phys. Rev. Letters, 16, 491.
O'Sullivan, C., Fegan, D. J., McBreen, B., O'Brien, D., 1978, Proc. of the 15th Int. Cosmic Ray Conf., 1, 188.
Porter, N. A., and Weekes, T. C., 1978, "Gamma Ray Astronomy from 10^{11} to 10^{14} eV Using the Atmospheric Cerenkov Technique," Smithsonian Astrophysical Observatory, Special Report, 381.
Rice-Evans, P., 1974, Spark, Streamer, Proportional, and Drift Chambers (London: The Richelieu Press, 1974).
Ross, R. W., Ehrmann, C. H., Fichtel, C. E., Kniffen, D. A., and Ogelman, H. B., 1969, IEEE Trans. Nucl. Sci., NS-16, 304.
Rossi, B., 1956, High-Energy Particles (Englewood Cliffs: Prentice-Hall, Inc., 1956), Chaps. 2 and 5.
Stearns, M., 1949, Phys. Rev., 76, 836.
Thompson, D. J., et al., 1993, ApJS, 86, 629.
Valentine, D., Kaplon, M. F., and Badhwar, G., 1969, Proc. of the 11th Int. Cosmic Ray Conf., 1, 101.
Voges, W., Pinkau, K., Koechlin, Y., Leray, J. P., Boella, G., Sironi, G., and Turner, M., 1973, Proc. of the 13th Int. Cosmic Ray Conf., 1, 293.
Volobuev, S. A., Galper, A. M., Kirillov-Ugryumov, V. G., Luchkov, B. I., Ozerov, Y. V., Rozenthal, I. L., Shermanzon, E. M., Grigorov, N. L., Kalinkin, L. F., Malioransky, A. S., Savenko, I. A., and Shashko, G. A., 1969, Proc. of the 11th Int. Cosmic Ray Conf., 1, 127.
Weekes, T. C., and Turver, K. E., 1977, Proc. of the ESLAB Symp. on Recent Advances in Gamma Ray Astronomy, 279.
Weekes, T. C., et al., 1989, ApJ, 342, 379.

SUBJECT INDEX

Acceleration 2, 8, 64, 75, 82, 84-87, 191, 209-214, 245, 327

Activation 34, 40, 41, 45, 56, 304, 319, 320
cosmogenic (by cosmic ray), *see* Nuclear gamma-ray lines
measurements of OSO, Apollo 17, Apollo-Soyuz, *see* Background components
nuclear activation by neutrons, *see* Nuclear gamma-ray lines
of CsI, NaI, and Ge, *see* Background components
of spacecraft, *see* Background components

Active Galactic Nuclei (AGN) 176-178, 183, 184, 187, 191, 196, 207, 209, 212, 214, 233
see also Blazars

Annihilation 2, 3, 8 102, 117, 155, 219, 252, 254, 287, 292, 319, 327
matter, antimatter 2
positronium 77, 78
see also Cross sections

Anticoincidence shield 329, 331, 337, 345

Apollo 15 and 16 301-305, 319, 321, 323, 324, 325, 329
diffuse spectrum measurement 220

Background components 301-305, 319, 320, 323, 324, 326, 328, 330, 336-338, 344, 345, 363
atmospheric and planetary surface gamma-ray interactions 3, 4, 250, 344, 345, 348, 364
Compton interaction 113, 264, 349
cosmic-ray and trapped-particle interactions 26, 65, 66, 246, 251, 252, 261-264
cosmic-ray and trapped-particle induced activation in detector 304, 305, 318
electron bremsstrahlung 252, 262, 265

high-energy gamma-ray detector 344, 345, 348, 349, 351, 357, 363
natural radioactivity 4, 8, 19, 55, 57, 58, 63, 75, 76, 80, 82, 85, 87
spacecraft component 303

Balloon(s) (large scientific) 4, 5, 90, 91, 93, 170, 284, 320, 322, 330, 344, 346

Batse 7, 128, 146, 151, 152, 239, 241-247, 329, 330

Big Bang model 1, 154, 219, 234

Blackbody radiation 4, 97, 98, 101, 102

Black holes 103, 118, 119, 154-156, 158, 173, 219, 234

BL lacertae objects 160, 165, 176, 177, 184, 186, 202-204, 206-209, 216, 232, 234

Blazars 7, 160, 177, 184, 187, 189, 191, 193, 195-201, 203, 204, 207, 208, 214, 215, 219, 231-233, 247

Bremsstrahlung 4, 8, 63, 75, 76, 80, 82, 85, 87, 98-102, 111, 116, 252, 262, 265, 301, 302, 304, 349, 350
background source, *see* Background components

Broadening (of spectral lines), *see* Doppler shift

Carbon-12
excited states and gamma-ray emissions (4.4 MeV line) 73, 75, 81, 82, 120, 327

Cerenkov 8, 142, 157, 343-345, 347, 350-352, 354, 356, 357, 359, 361

Coded aperture 330, 337, 363

Comptel 7, 79, 118, 119, 128, 131, 133, 137, 138, 140, 146, 154, 172, 173, 186, 196, 223, 232, 242, 336

Compton
cooling 213
detector efficiencies 297, 345, 353, 354
diffuse emission 106
effect 4, 103, 326, 343, 350
gamma-ray absorption 343, 364, 365
interaction with blackbody radiation 99, 102
inverse compton scattering 126, 137, 210, 211
radiation 348, 349, 350
scattering 63, 75, 251-265, 268, 269, 277, 286, 287, 291, 333, 334, 336, 349; *see also* Cross sections, and Response function
suppression 322, 324, 327
telescope 333

Compton Gamma-Ray Observatory 5, 7, 84, 85, 107, 118, 119, 124, 128, 131, 151, 172, 176, 178, 184, 214, 220, 238, 239, 242, 244, 325, 329, 336, 347

COS-B 106, 126, 132, 137, 138, 139, 157, 184, 346, 347, 351, 353, 355, 356, 357, 364

Cosmic rays 2, 90, 94, 95, 97, 98, 100, 102, 104, 120, 121, 150, 152, 159, 164-167, 219, 228, 229, 235, 323, 344, 347, 350, 355
activation by, cosmogenic 304, 305
active shields 320, 322, 323, 336
distribution in Galaxy 91, 105
electrons 102
gamma-ray background, produced by 98
interactions 2, 3, 4, 30, 34, 53, 100, 110, 114
nuclear abundances 93, 94
see also Nuclear gamma-ray lines and Neutrons

Cosmological red shift 3, 224

Cosmology 8, 198, 219, 224-226

Coulomb scattering 262, 349-351, 353, 357, 362, 364

Crab nebula 5, 123-126, 132-138, 143, 147, 152

Cross sections 19, 96, 99-104, 296
bremsstrahlung 101, 102, 265
Compton 102, 103, 251, 255-260, 262-265, 286, 287, 333, 335, 336, 348-350
definitions (mass, linear, electronic, energy loss, radiation length, and range) 252, 255-260
(e^+, e^-) annihilation 104
neutrons 33, 66, 67
pair production 254-262, 286, 287
π^0-meson production 99-101
photoelectric 27, 251, 253, 255-260, 286, 287
see also Detector efficiency

Doppler shift 65, 73, 95

Egret 107, 111, 112, 128, 131-140, 142, 144, 149-151, 155, 159, 160, 169, 171, 173, 174, 177, 178, 183, 184, 186, 189, 190, 195-199, 202, 203, 206, 208, 221, 224, 225, 232, 233, 243, 245, 347, 351, 353, 354, 356, 357, 361, 364, 365

Electrons 2, 4, 8, 27, 64, 73, 75, 76, 80, 82, 85, 87, 90, 91, 94, 97, 98, 100-102, 108, 116, 124, 152, 154, 209, 210, 211, 225, 228, 251-254, 261-265, 267-282, 285, 287, 292, 293, 297, 298, 300-305, 319, 321, 324, 328, 335, 337, 343-352, 354-357, 361, 362, 364, 365

Electron-positron 125, 213, 254, 345, 361, 365

Electron-hole 267, 268, 298

Energy loss 30, 54, 108, 124, 141, 144, 250, 252, 253, 262, 263, 265, 269, 281, 293, 305, 326, 349

Flat spectrum Radio Quasars (FSRQs), *see* Galaxies
3C 273, *see* Galaxies
3C 279, *see* Galaxies

Galaxies 1, 3, 8
active 176, 214, 219, 228, 231, 232; *see also* AGN
BL lacertae objects 160, 165, 176, 177, 184, 186, 202-204 206, 207-209, 216, 232, 234
Flat Spectrum Radio Quasars (FSRQs) 159, 177, 198, 202, 203, 204, 206-208, 215
Magellanic Clouds 90, 105, 167, 173
Large Magellanic Cloud 167-169
Small Magellanic Cloud 167, 169-171
radio galaxies 167, 168, 176-178, 184, 214, 219, 231
Seyfert 167, 176-178, 181, 184, 214, 219, 231, 235
spiral 96, 97, 167, 177
3C 273 184, 187, 196
3C 279 184, 187, 189, 193, 196, 197, 202, 203, 210

Galaxy, the 1, 2, 4, 5, 8, 90-92, 95-98, 102, 105, 106, 110, 112, 118, 120, 121, 155, 239, 344

Gamma-ray burst 7, 158, 159, 163, 234, 238-245, 248, 326, 329

Gamma-ray lines, *see* Nuclear gamma-ray lines, Annihilation, and Natural radioactivity

Gamma Ray Observatory (GRO), *see* Compton Gamma Ray Observatory

Helium nuclei 92, 120

High Energy Astronomical Observatory (HEAO) 118, 178, 220, 284, 304, 328

Hubble Parameter (H_0) 198-201, 225, 227-230

Hydrogen 1, 32, 33, 80, 94, 96, 97, 100, 111, 120, 170, 171
capture line (2.223 MeV) 73, 80, 81, 84

INTEGRAL 285, 337, 338

Interstellar matter 2, 91, 92, 95-97, 99, 104, 121, 123, 344

Interstellar magnetic field, *see* Magnetic fields

Jupiter 14, 26, 27

Lithium drift and intrinsic germanium 277, 278, 282, 283

Luna 10 and 11 6, 47

Lunakhod 6, 47

Magellanic Clouds, *see* Galaxies

Magnetic bremsstrahlung, *see* Synchroton radiation

Magnetic fields 2, 13, 58, 77, 95, 98, 210, 212, 213, 286, 303, 320, 328
galactic 95, 98, 105, 106
pulsars 125-127, 149

Mars 5, 6, 19, 31, 48, 59-61, 284, 300, 301, 319, 322, 330, 332

Mars 4 and 5 (Russian) 6

Moon 5, 6, 14, 19, 27, 31, 45-49, 51, 54, 56-59, 61, 62, 284

NEAR 19, 59, 301, 304, 322-325, 327

Neutrons 5, 30-34, 65, 72, 73, 84, 209, 214, 299, 300, 305, 332, 333, 336
capture, *see* Nuclear gamma-ray lines
inelastic scatter, *see* Nuclear gamma-ray lines
neatron production 74

Neutron stars 124, 125, 147, 212, 248

Nuclear gamma-ray lines 5, 8, 64, 153
activation (prompt and delayed) 33-44, 68-73, 80-83
inelastic scatter (charged and neutral particles)
natural radioactivity (π^0-mesons) 27-30, 76
spallation 44

Nucleosynthesis 4, 84, 104, 105, 120, 153, 154, 171, 173

OSO-1 345

OSO-3 4, 107, 345

OSO-7 79

OSSE 7, 79, 119, 120, 128, 131, 133, 137, 138, 141, 146, 151, 157, 172, 173, 178, 181-183, 186, 189, 196, 232, 242, 325, 329

Oxygen 6, 47, 65, 354
excited states of (6.12 MeV and 15.11 MeV lines)..73, 81, 82

Pair production 127, 197, 210, 251-261, 268, 269, 277, 282, 286, 287, 326, 343, 348, 349; *see also* Cross sections

Photoelectric effect, *see* Cross sections and Detectors

π^+-mesons 76, 77

π^0-mesons 2, 8, 65, 76, 99

Positron(s) 63, 65, 73, 77, 78, 81, 82, 88, 99, 104, 111, 119, 120, 127, 157, 209, 213, 254, 262, 305, 318, 345, 350, 361, 365

Positronium 77

Pulsars 5, 8, 96, 121, 124-128, 131, 136, 141-142, 144, 146-152, 160, 162, 164-166

Pulse-height and digital spectra 54, 83, 274, 287, 295, 302, 305, 319, 320, 325, 332, 353
relationship to photon spectrum 268

Quasars 167, 177, 184, 193, 198, 200, 231

Radio galaxies, *see* Galaxies

Radioactive capture, *see* Nuclear gamma-ray lines

Ranger 5, 47, 324

Response function 297-296

SAS-2 5, 220, 223, 231, 346, 347, 351-357, 363, 364

Seyfert galaxies, *see* Galaxies

Solar flares 5, 16, 45, 63-65, 67, 73, 75-77, 79, 84, 88, 325
gamma-ray productions, *see* Nuclear gamma-ray lines,
Neutron production, *see* Neutrons

Solar Maximum Mission (SMM) 79, 82-84, 304, 325, 326, 328

Solar-system evolution 3, 5, 8, 11-16, 31, 40, 46

Spallation 44: *see also* Cross sections and Nuclear gamma-ray lines

Supernovae 90, 96, 104, 106, 120, 121, 125, 153-155, 171, 173, 241

Surveyor 1, 6, 47, 59

Synchroton radiation 95, 106

21-cm line 96, 97, 121

Vela pulsar 118, 127, 128, 131, 136, 141, 144, 146, 151, 154

Venera 6, 19, 48, 242

Venus; *see* Venera

Viking 6, 48

WIND 284, 300, 332

X-rays 19, 83, 131, 135, 136, 141, 142, 147, 151, 156, 174, 183, 201, 253, 277, 287, 338
general 3, 8, 20-23, 91, 178, 210, 214, 219, 220, 223, 230, 233, 235, 238, 241, 247
emission from the moon, planets, comets, and asteroids 6, 26, 27, 46-49, 59,