5 / SCIENCE ON THE MOON

THE EXPEDITION OF LEWIS AND CLARK, the voyages of Captain Cook, and the lunar landings of Project Apollo are but a few examples of the symbiosis between scientific investigation and historic human probings of frontiers. Obviously, exploration and study of the Moon itself will be part of lunar surface activities, but what other scientific opportunities arise from permanent human presence? The Lunar Base Working Group, which met in Los Alamos in the spring of 1984, addressed this question in a general way. Consideration was given only to experiments uniquely enhanced in the lunar environment. Gradual expansion of facilities and capability will lead to various experiments that become easy to perform, but special installations or laboratories should be planned only when the research cannot be readily performed on the Earth or in easily accessible locations in space.

Broadly speaking, candidate experiments will be those capitalizing on the unique elements of the lunar environment. These include low gravity, absence of a planetary magnetic field, access to the plasma environment of the solar wind and the Earth's geomagnetic tail, no atmosphere, absence of water and other volatiles, isolation from the terrestrial biosphere, easily created low temperature radiative environments, availability of laboratory volumes with very high thermal and seismic stability, and the easily achievable pointing stability for observations of all kinds. As a rule, the lunar surface is a candidate location for any experiment that suffers interference from noise of geologic, biologic, or human origin.

Astronomical observations from the Moon are both promoted and criticized. Observers appreciate lack of atmospheric absorption but are quick to counter that satellite astronomy offers the same advantage. However, the slow rotation rate of the Moon permits very long integration times (literally days) from a very stable pointing platform. When a lunar base becomes a going concern, ease of maintenance and changeout of equipment will increase the productivity of an observatory.

The papers on astronomical concepts predict significant increases in resolution and sensitivity from lunar-based observations. Bernard Burke's optical interferometer represents a technical challenge that would be impossible without the positional stability of the elements and the freedom from atmospheric absorption and dispersion. Burns takes advantage of the lunar distance from the Earth to create an extremely long baseline interferometer of unprecedented resolving power. Douglas and Smith describe a simply constructed radio interferometer capable of probing the universe at long wavelengths unobservable from the surface of the Earth. Haymes can address a number of important astronomical questions with observations at the high energy end of the spectrum.

Cosmic-ray astronomy, discussed by Adams and Shapiro, not only samples products of physical processes in the galaxy but also provides data on the radiation background that limits the time astronauts can spend unshielded on the lunar surface. Shapiro and Silberberg review the types of neutrino sources that might be observed to advantage from the lunar surface. Cherry and Lande expand on the subject of the background interferences in lunar-based and terrestrial-based neutrino detection and describe what a lunar detector might look like. Petschek evaluates the possibilities of detecting proton decay or neutrino oscillations with a lunar based experiment.

Astronomy and astrophysics dominates the discussion in this section but in no way exhausts the possibilities. For example, Hörz speculates that dating of lunar craters might give perspectives on the proposed periodic bombardments of the Earth associated with mass extinction events. The paper by Anderson *et al.*, in the section on Space Transportation mentions techniques to measure the space plasma environment around the Moon with small, lunar-based sounding rockets.

A large body of science on the Moon can be viewed as applied research. Agriculture will be part of advanced life support systems. Medical and physiological data collection will be part of the study of the adaptation of humans to the lunar environment and will complement work done in space on the zero gravity environment. Microbial engineering will be part of the closed lunar ecosystem and, in addition, may have application to material processing as discussed by White and Hirsch in the section on Lunar Materials. Civil engineering and construction technology will receive attention in the course of lunar development.

In the final analysis, the scientific discoveries enabled by a lunar base are not predictable. Even a reasonable list of possibilities will be assembled only when a large number of minds with a broad spectrum of knowledge are made aware of the research opportunities. One purpose of this book is to stimulate the imagination of the scientific and technical communities.

ASTRONOMICAL INTERFEROMETRY ON THE MOON

Bernard F. Burke

Mail Code 105-24, California Institute of Technology, Pasadena, CA 91125. Permanent address: Room 26-335, Massachusetts Institute of Technology, Cambridge, MA 02139

Optical interferometric arrays are particularly attractive candidates for a manned lunar base. The radio model already exists: the Very Large Array (VLA) of the National Radio Astronomy Observatory, situated on the plains of St. Augustine near Socorro, New Mexico. A Y-shaped array of 27 antennas, each arm being 20 km long, operates as a coherent array, giving 0.1 arcsecond resolution at 2 cm wavelength. An array of similar concept, but with optical elements, would therefore give angular resolution of nearly one microarcsecond resolution at optical wavelengths and would give an absolutely revolutionary new view of objects in the universe. It would not be built on the Earth's surface because the atmosphere damages the phase coherence too severely at optical wavelengths. It could be constructed in Earth orbit as an assemblage of station-keeping free flyers (proposals to do so have been put forward) but the technical problems are not simple, *e.g.*, controlling element position and orientation to 100 A in 20 km. If a permanent lunar base were available, an optical analog of the VLA would, in contrast, be a relatively straightforward project.

THE CASE FOR HIGH ANGULAR RESOLUTION

Galileo's telescope was the first step in improving the angular resolving power of the human eye; this thrust in astronomy continues in our own time. The atmosphere of the Earth has posed a barrier at about one arcsecond (perhaps one-third of an arcsecond at the best sites), but if optical instruments can be mounted in space there seem to be few fundamental difficulties in extending to the microarcsecond range. Most of the problems are of a practical nature, centered on structural stability, satellite station-keeping, instrument adjustment and control, and related technical questions; these problems are solvable in principle, but may turn out to be costly if conventional orbital concepts are followed. Although the surface of the Moon has not been seriously considered in the past, it appears that astronomical instruments of great power could make good use of a lunar location. A permanently occupied lunar base could play a key role in such a program.

Angular resolution can never be better than the diffraction limit λ/D, the wavelength divided by the aperture diameter, and at 5000 Å, a one-meter aperture gives one-tenth arcsecond resolution. Milliarcsecond and microarcsecond resolution will require interferometers of large size, but much wider classes of problems, all of great current interest, become accessible. These are illustrated in Fig. 1, which shows the approximate optical fluxes and angular sizes of a variety of stellar and extragalactic objects. Since it is the maximum flux and largest angular size that is indicated, objects in each class will generally fall along the locus indicated by the upward sloping arrows. An object ten times more distant than the closest member of its class lies at the tip of the arrow,

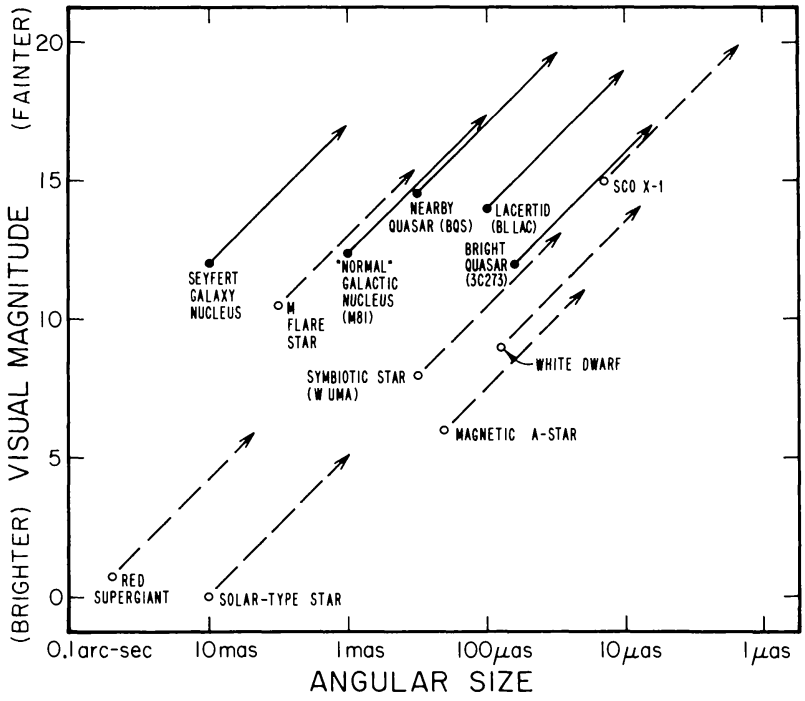

Figure 1. Magnitude and maximum angular sizes for a selection of stellar and extragalactic objects. The scales are chosen so that an object of a given class moves in the direction of the arrow, whose length corresponds to a factor of ten in distance.

for the given scale. The figure, therefore, gives the *largest* expected scale for each class of object.

For the various classes of stars, Dupree *et al.* (1984) have commented that measuring the size of a star is not enough, a conclusion that is generally valid for nearly all astronomical objects. Most interesting objects tend to be complex, and understanding the physical processes requires some detailed knowledge of the phenomena. For most stars, at least a factor of thirty resolution beyond the gross size is certainly needed (*i.e.*, about 100 pixels). Phenomena such as starspots, flares, and other analogs of solar processes will be interesting and, indeed, should be surprising. One is driven to the conclusion that every class of stellar object (except for the closest red supergiants) will demand an angular resolution of a milliarcsecond or better.

The extragalactic phenomena are still more demanding. The complexity of the processes is not known, since we do not have close analogs (such as the sun, for the stellar case) to guide us. The subject matter is of extraordinary interest, however: the physics of quasars, blacertids, and "ordinary" galactic nuclei press close to (or perhaps beyond) the limits of fundamental principles. It is clear that enormous energies are generated, both from radio and x-ray observations of these objects, and the indications are very strong that the energy source must be gravitational.

"Black holes," though not yet demonstrated in nature, may play a key role in these energetic processes. The optical study of the accretion processes and instabilities near the cores of the active extragalactic objects, with high angular resolution, should be as astounding as it has been in the radio case, where milliarcsecond resolution reveals velocities that appear to surpass the speed of light. Reference to Fig. 1 shows that only the broad-

line regions at the nuclei of the closest Seyfert galaxies are accessible to an instrument of milliarcsecond resolution. The rest are smaller in angular size, and it is clear that an optical instrument having angular resolution in the 1–10 microarcsecond range would have truly extraordinary impact. None of the objects are brighter than the twelfth magnitude, and most are substantially fainter; an instrument having at least the collecting area of the Palomar 5-meter telescope is indicated. This challenge of obtaining angular resolution in the milliarcsecond to microarcsecond range, with a net collecting area of at least twenty to thirty square meters, is fully justified by the scientific rewards that would surely be gained.

APERTURE SYNTHESIS

Radio astronomers have, for the past several decades, circumvented the problem of obtaining high angular resolution by using interferometry, culminating in the concept that is called *aperture synthesis*. The methods were, ironically, developed by Michelson (1920) for measuring the diameters of stars at optical wavelengths, but the Earth's atmosphere hindered its quantitative use. The radio version of Michelson's stellar interferometer is illustrated in Fig. 2, which shows a pair of radio telescopes simultaneously receiving radiation from a distant source. There is a difference in arrival time, the geometrical time delay $\Delta\tau_g$, determined by the orientation of the source direction relative to the interferometer baseline. There is obviously no chance of interference if $\Delta\tau_g$ is larger than the coherence time t_c of the radiation, so a time delay must be inserted to compensate for this difference. Then, if the antennas are fixed and the source drifts through the reception

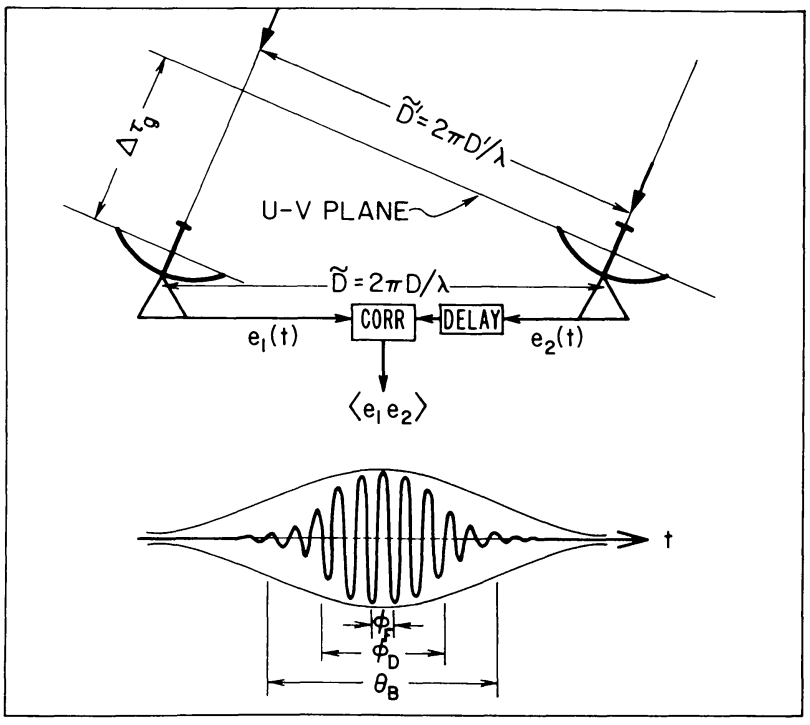

Figure 2. The Michelson stellar interferometer in its radio form. The output of the correlator, with the d.c. term removed, is shown for fixed apertures as a function of time; this is equivalent to variation with angle off-axis.

pattern, the product of the received signal amplitudes varies sinusoidally as the signals alternately interfere, constructively and destructively. These characteristic angular scales are important: the primary reception pattern of half-width θ_B, the fringe spacing ϕ_F, and the delay beam ϕ_D. The analysis is most straightforward if the antennas track the source, with the source itself being small compared to the primary beamwidth θ_B. The fringe spacing is determined by the projected baseline D', which is the projection normal to the incoming radiation.

For the interferometer description, there is a third angle, the delay beam ϕ_D, that is determined by the receiving bandwidth or, equivalently, by the coherence time. If the time delay is set to match $\Delta\tau_g$ perfectly, the central fringe will have full amplitude, but as the time delay error grows, the interference conditions will be different at the upper and lower ends of the band. The interference effects cancel, and the fringe amplitude diminishes over an angle $\phi_D \sim 1/B\tau_B$, where B is the bandwidth and τ_B is the baseline length measured in light travel time. The number of fringes observed as a consequence is of the order of the inverse of the fractional bandwidth, an effect that has strong consequences for optical interferometry.

Given a two-element Michelson interferometer as illustrated in Fig. 2, the output is well-specified if the following conditions are met: the source under study must be small compared to both the primary resolution θ_B and the delay beam θ_D, and the delay compensation must approximate $\Delta\tau_g$ with an accuracy corresponding to a fraction of the fringe angle θ_F, or at least the error must be calibrated to that accuracy. The interferometer output is the convolution of its sinusoidal fringe pattern with the source brightness B(x,y) where x,y are angular coordinates on the sky. This means that the interferometer output is equal to the *Fourier transform* B(u,v) of the brightness distribution. The conjugate coordinates (u,v) are defined by the baseline and the source location as shown in Fig. 2: on a plane normal to the source direction, coordinates (u,v) are defined (North and East, for example) and the interferometer baseline D, measured in wave numbers ($2 \times D/\lambda$), is projected onto that plane with the reference antenna (which can be chosen arbitrarily) at the coordinate origin. The plane is called the *u-v plane*, and the projected vector D'(u,v) defines the conjugate coordinates at which the Fourier transform B(u,v) is defined by the fringe amplitude and phase. If all interferometer baseline lengths and orientations are taken, the complete Fourier transform is determined, and performing Fourier inversion gives a true map B(x,y) of the source. In practice, of course, there is noise introduced by the apparatus; the coverage of the u-v plane is not complete; and due caution and knowledge must be exercised.

The process by which the Fourier transform is developed is known as *aperture synthesis*, and substantial literature has been developed for the radio case. The first complete description, in which the rotation of the Earth was used to move the interferometer baseline, was conceived by Ryle and Hewish (1960), and an authoritative summary of the two-element interferometer has been given by Rogers (1976). The most powerful aperture synthesis instrument is the radio array known as the VLA (the Very Large Array, operated by the National Radio Astronomy Observatory); it is described by Napier *et al.* (1983).
The VLA probably provides the best model for a desirable optical instrument. Its 27 elements

give 351 simultaneous baselines; this means that "snapshots" of fairly complex objects are nevertheless faithful representations if the target is not too complex, or if a dynamic range of a few hundred to one is sufficient. At the same time, for large fields of view and complex targets, its variable configuration and ability to use the rotation of the Earth to obtain more complete u-v plane coverage is vital. The size of the array, 20 km per arm of 35 km equivalent overall size, was set by the original operating requirement that it should equal conventional optical telescope resolution (1″ at 20 cm, 0.3″ at 6 cm). The same considerations will apply to an equivalent optical instrument. The discussion in the beginning of this paper, illustrated by Fig. 1, indicates that a mapping capability of ten microarcseconds would give a rich scientific return. At this angular scale, significant changes can be expected both for stars and active extragalactic objects within brief time spans. The system must therefore have a large number of elements, as in the case of the VLA. This gives two further advantages: a large number of objects can be studied in a short time because of the "snapshot" capability, and the more complete u-v plane coverage can yield maps of high dynamic range. If the optical array contains 27 elements, each element would have to have at least one m diameter to give a total collecting area comparable to the five m Palomar telescope. The instrument should cover the wavelength range 1216 Å (lyman-alpha) to 5 microns; for the mean wavelength of 5000 Å, this implies that an optical aperture-synthesis array should have a diameter of about 10 km.

One of the major considerations of any concept has to be the phase stability of the system. Incoherent and semicoherent interferometers (the Brown-Twiss interferometer is a brilliant example) suffer in signal-to-noise ratio and loss of phase information, and so must be rejected. For the complex objects of greatest interest, phase information is essential. This requirement exacts a price: control (or measurement) of the optical paths to $\lambda/20$ means that 250 Å precision is needed at λ5000, and proportionally tighter specifications are required as one goes to shorter wavelengths. The radio astronomers, in developing VLBI, have formulated a powerful algorithm, phase, and amplitude closure that eases the problem if there are enough receiving apertures. The technique has been applied to VLBI mapping problems with great success (Readhead and Wilkinson, 1978). If one has three elements, and hence three baselines, the instrumental phase shifts sum to zero; similarly, if there are four elements in an array, the instrumental perturbations to the amplitudes cancel. As the number of elements increase, the information recovery becomes more and more complete. For N antennas, a fraction (N-2)/N of the phase information and (N-3)/N-1 of the amplitude information can be recovered. If N is 10 or more, the procedure appears to be thoroughly reliable. The phases must still be stable over the integration period; this means that the precision requirement on the optical paths must be held, but the *time* for which it is held is reduced. The desired sensitivity and the total collecting area therefore set the final stability specifications.

Up to the present time, two general classes of optical space interferometers have been proposed: station-keeping, independently orbiting interferometers, and structually mounted arrays. Examples of the first class are SAMSI (Stachnik *et al.*, 1984), in which pairs of telescopes are placed in near-Earth orbit, and TRIO (Labeyrie *et al.*, 1984), in

which a set of telescopes are maneuvered about the fifth Lagrangian point in the Earth-Moon system. Among the structural arrays that have been proposed are COSMIC (Traub and Carleton, 1984), OASIS, a concept proposed by Noordam, Atherton, and Greenaway (unpublished data), and a variety of follow-on concepts to the Space Telescope being examined by Bunner (unpublished data). At the present time, all of these concepts hold promise for giving useful results in the milliarcsecond class, but when the number of elements grows to the order of 27 (or more) and when the spacings extend to 10 km (or even 100 km, for one microarcsecond resolution at λ5000) the solutions may prove to be expensive, perhaps prohibitively so.

A third class of optical array becomes feasible, however, if there is a permanently occupied lunar base. The Moon turns out to be a most attractive possible location for an optical equivalent of the VLA, capable of microarcsecond resolution.

A LUNAR VLA

Assuming that a lunar base has been established, the general outlines of a large optical array following the pattern of the VLA can be visualized with some confidence.

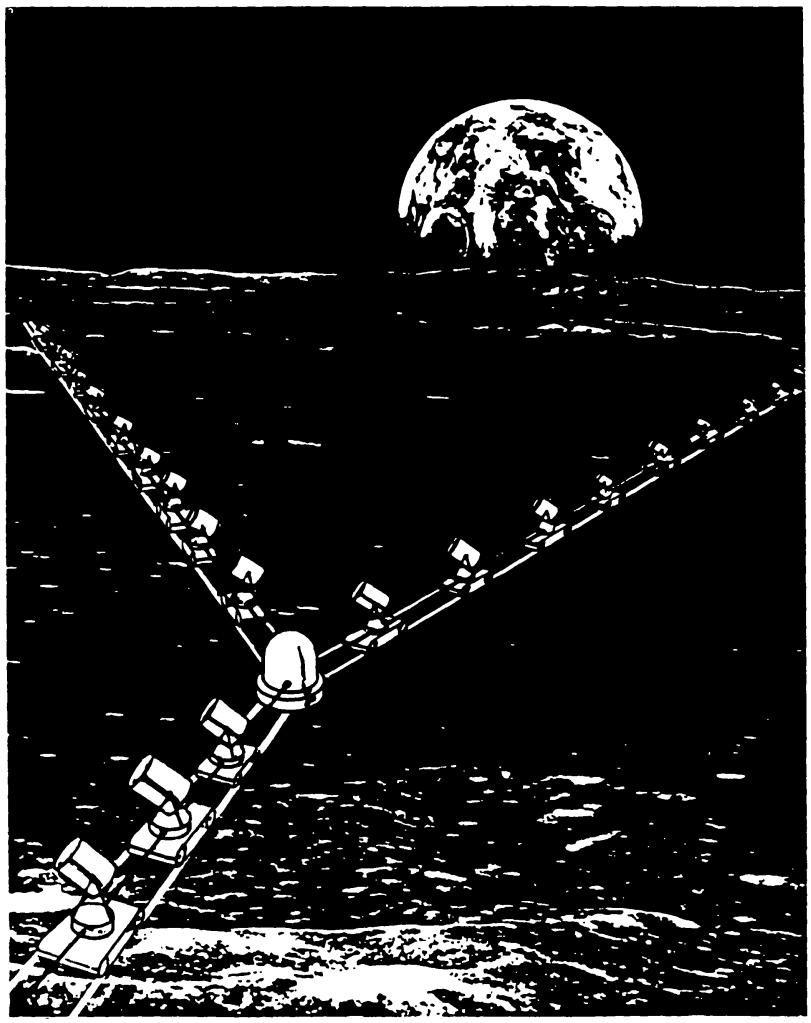

Figure 3. A schematic view of an optical aperture synthesis array on the Moon. The individual elements could assume forms very different from the versions shown.

A schematic form is shown in Fig. 3; a set of telescopes, suitably shielded, are deployed at fixed stations along a Y, each arm being 6 km long, for a maximum baseline length of 10 km. There is a fixed station that monitors the telescope locations by laser interferometers. The telescopes must be movable, but whether they are self-propelled (as shown in Fig. 3) or are moved by special transporters (as in the case of the VLA) is a technical detail. The received light signals are also transmitted to the central correlation station, but time delays must be inserted to equalize the geometrical time delays ($\Delta\tau_g$) illustrated in Fig. 2. These are not shown; a number of configurations are possible, probably in the form of laser-monitored moving mirrors.

The individual telescopes might well be approximately one m in diameter. The telescopes could be transported in disassembled form and, hence, they need not be extremely expensive since launch stress would not be a problem. A simple conceptual design indicates that each telescope might have a mass of 250 kg or less; the total telescope mass would then be about 7 tonnes for 27 telescopes plus a spare. The packing volume could be relatively small, since the parts would nest efficiently. The sketch in Fig. 3 shows each telescope being self-propelled, but if mass transportation to the Moon is a key consideration, one or two special-purpose transporters seems much more likely. Each might have a mass of about 200 kg.

The shielding of the telescopes is an interesting design problem. The simplest scheme would be to adopt the systems used on past telescopes in space, such as the International Ultraviolet Explorer (IUE), but the construction possibilities on the lunar surface may allow concepts that give dramatic improvements. Instead of mounting the shields on the telescopes themselves, the shields could be constructed as independent structures that sit on the lunar surface, free of the telescope itself. The shields might be very simple, low-tolerance, foil and foam baffles, keeping the telescope forever in the shade, radiatively cooled to a very low temperature, or perhaps kept at the average 200 K temperature of the lunar subsurface. It would appear that the thermal stresses might be kept very low by adapting the design to the lunar surface conditions.

Transmission of the received light from the telescopes to the central correlation station must proceed through a set of variable time delays as indicated earlier, and here there is a need for technical studies. For the 10 km maximum baselines proposed here, the maximum time delay rate would be 2.6 cm/s, which is not excessively high. The requirement of $\lambda/20$ phase stability is challenging: the motion should not have a jitter much greater than 100 Å/s rms, so a smoothness of something better than a part per million is needed; not an easy goal, but not beyond reason. The curvature of the lunar surface has to be taken into account unless a convenient crater can be found whose floor is suitably shaped. The height of the lunar bulge along a 6-km chord is 1.5 m and, hence, is not a serious obstacle. For the larger concept (60-km baseline, microarcsecond resolution at λ5000) the intervening rise of 150 m would be more serious, and suitable refraction wedges or equivalent devices would have to be arrayed along the optical path. The transmitted signal should probably be a quasi-plane wave; this translates to the requirement that the receiving aperture at the central correlator station should still be in the near field of the transmitting aperture of the most distant telescope. This specifies the diameter of the transmitted beam, which must have a diameter greater than 10 cm at λ5000,

and 30 cm for 5 μ. If there were a desire to carry out aperture synthesis at 50 m (which there might well be), the transmitted beam would have to be at least a meter in diameter, a requirement that would still be easy to meet, since the tolerances would be relaxed.

The characteristics of the central correlator will depend on the results of detailed studies. Two general classes of optical systems can be projected: the "image plane" correlation geometry developed by Labeyrie (1984) for TRIO (a continuation of the traditional technique of Michelson), and the "pupil plane" correlation scheme generally used by radio astronomers, but realized in the optical regime by the astrometric interferometer of Shao *et al.* (1984).

One interesting advantage generally enjoyed by optical interferometry as compared to radio interferometry is the ease with which multi-banding circumvents the "delay beam" problem described earlier. Labeyrie (1980) has devised an ingenious dispersive system that efficiently eliminates the problem for most cases. The fringes are displayed in delay space and frequency space, but modern two-dimensional detectors such as CCDs (Charged Coupled Devices) handled the increased data rate easily.

The data rates are not excessive, being completely comparable to the data rates now being handled by the VLA. The 351 cross-correlations needed for a 27-element system (or 1404 if all Stokes parameters are derived) requires an average data rate of about 100 kilobaud for a 10-second integration period; future systems always require larger data rates, but even a projection of an order-of-magnitude increase does not seem to present formidable data transmission problems.

Finally, a word is in order concerning the use of heterodyne systems to convert the optical signals to lower frequencies. The technique is in general use in the radio spectrum, extending to wavelengths as short as a millimeter. Unhappily, the laws of physics offer no hope for astronomical use of heterodyne techniques at optical and ultraviolet frequencies. Every amplifier produces quantum noise, and the laws of quantum mechanics are inexorable: approximately one spurious photon per second per Hertz of bandwidth is produced by every amplifier. At radio frequencies, the quantum noise is swamped by the incoming signals since there is so little energy per quantum. Optical systems, with bandwidths of 10^{13} or 10^{14} Hertz, can afford no such luxury. The crossover in technology occurs somewhere between 100 and 10μ. As infrared detectors improve, the shortest wavelength at which heterodyne detectors are practicable will be perhaps 50μ.

Except for these quantum limitations, the concepts developed for radio techniques carry over to the optical domain. The signal-to-noise analysis differs somewhat. The noise limits are determined by the Rayleigh noise of the system in the radio case, while the quantum shot noise of the signal itself determines the signal-to-noise ratio in an optical system. Otherwise, the extensive software armory developed for radio synthesis systems should be directly applicable to optical interferometers.

ARE THERE SERIOUS OBSTACLES?

Relatively little thought appears to have been given thus far to the advantages of the Moon as a base for astronomical instruments. There are a number of current misconceptions that seem to hold little substance.

1. *Does lunar gravity cause problems?* On the whole, the effects of lunar gravity appear to be beneficial. The relatively small (1/6 g) acceleration helps to seat bearings, locate contact points, and generally should provide a reference vector for mechanical systems. The lunar gravity removes dust from above the surface, keeping the density of light-scattering particles low.

Gravitational deflection for telescopes in the one-meter size range is completely negligible. Gravitational deflection does not depend upon the weight of a structure; elementary physics shows that the structural deflection s of a structure depends on the length l of the beam, Young's modulus Y, the density ρ, the gravitational acceleration g_m, and a dimensionless geometrical factor γ that decreases as the depth of the beam increases:

$$S \approx \gamma \left(\frac{\rho}{Y}\right) g_m l^2 \qquad (1)$$

On Earth, 4- and 5-meter telescopes have been built with mirror support systems that limit mirror deflection to a fraction of a wavelength of light under full gravity. A 1-meter mirror, located on the Moon but otherwise similar, would be stiffer than a terrestrial 4-meter mirror by a factor of about 100!

Deflection of the telescope structure can be controlled to high tolerances. Not only are superior materials like carbon-epoxy now available, but there are improved design methods such as the concept of homologous design (introduced by von Hoerner in 1978), in which a structure is designed that always deforms to a similar shape. In summary, gravitational deflection poses no problem.

2. *What about the thermal environment?* The Moon is an approximately 200 K blackbody subtending 2 π steradians on the underside of a lunar-based instrument. For a conventional satellite in low-Earth orbit (LEO), the Earth is an approximately 200 K blackbody subtending nearly 2π beneath the spacecraft; however, if the spacecraft is tracking a celestial object, the aspect is changing rapidly—on the order of 4° per minute. The telescope tracking a celestial source in the lunar environment is changing its aspect at about 0.01° per minute. When one considers the additional advantage of the natural lunar terrain for better thermal shielding to start with, and the ability to upgrade its quality at a permanent base, the lunar environment is almost certainly more favorable than LEO from the point of view of thermal stresses. The L5 case is different, since the elements would always be exposed to direct solar radiation.

3. *Is scattered light a problem?* Again, equipment in LEO has the Earth subtending nearly a hemisphere, but the Earth has high albedo and the Moon has low albedo. The lunar environment is strongly favored, and, as in the thermal case, one should be able to provide superior light shielding on the Moon.

4. *Is direct sunlight a problem?* The sun shines only half the time, and its direction changes slowly. Given the superior light baffling of the lunar-based telescopes, the lunar environment will probably turn out to be far superior to either LEO or L5, but thermal studies of real designs should be made.

5. *What about lunar dust?* The laser retro-reflectors have been in service for over a decade, with little performance degradation reported. Dust seems to be no problem,

probably because the Moon's gravity scavenges it rapidly. A very rare meteorite impact nearby might take one or two telescopes out of service, and the choice would have to be made to clean or replace the instruments.

6. *Is seismic activity a problem?* The Moon is far quieter than the Earth, with a low Q. At good seismic stations on the Earth, the seismic noise is less than one Å rms; the poor locations have high noise because of the effects of wind and surf. Lunar seismic activity is not a concern.

7. *Do the solid body tides of the Moon move the baselines too much*? Earth tides are routinely accommodated by geodesy groups conducting VLBI studies on Earth, where the motions amount to several wavelengths every 12 hours. The lunar tides are larger in amplitude, but they proceed so slowly that they can be compensated for. The 10 km maximum baseline of a lunar VLA is a smaller fraction of the lunar diameter than the 10,000 km VLB baselines are of the Earth's diameter, which diminishes the amplitude of baseline motion. The net tidal motion of the maximum baseline vector should be of the order of a few tenths of a millimeter. This is not a negligible motion, measured in wavelengths of light, but the slow lunar rotation leads to a manageable correction rate of the order of a few wavelengths per hour. The usual interferometric calibration routines should keep this error source under control.

8. *Can the baseline reference system be well defined?* The analogy with terrestrial VLBI is so close that the answer has to be affirmative. The errors can be controlled; the lunar soil is sufficiently competent to stably bear the load of a telescope; and, if necessary, hard points can be established to check on vertical motions. Interferometers are largely self-calibrating: there are enough quasi-stable reference points in the sky to allow the observations themselves to bootstrap the instrumental constants.

SUMMARY

A permanent lunar base can provide support for a variety of astronomical investigations. An optical interferometric array, perhaps of the general form of the VLA but designed for optical instead of radio wavelengths, would lead to a qualitative advance in our understanding of the universe. The Y configuration is well suited to expansion, and the VLA has demonstrated that it can make maps both rapidly (in its snapshot mode) and with high dynamic range (when multiple array configurations are used). Other configurations, such as maximum-entropy-derived circles, should certainly be examined.

A wide variety of scientific problems could be addressed by such an instrument. The stellar analogs of the solar cycle, the behavior of sunspots on other stars, the magnetic field configurations of other stars, and the behavior of dynamic plasma phenomena such as flares and winds, are all examples of star-related problems that ultimately would lead to both fundamental knowledge of how stars formed and evolve, and increase understanding of our own sun. A wide variety of extragalactic problems could be studied, including the fundamental processes associated with black holes and massive condensed objects as they are manifest in quasars, galactic nuclei, and other optically violent variables. There would surely be a number of dramatic surprises, both in stellar and extragalactic studies,

and the instrument would certainly be at the forefront of astronomy from the time of its first use.

There seem to be no fundamental problems in building such an instrument. The total mass to be delivered to the lunar surface for the instrument would be 10–30 tonnes, roughly equivalent to one space station habitat module. The detailed system studies have not yet been made, but even a preliminary conceptual investigation indicates that the elements of the system are relatively straightforward. The presence of man is highly desirable for this particular instrument, in marked contrast to the free-flyer case in which the instruments are too easily perturbed by human presence.

How long would it take to build the instrument? The answer depends upon the timescale of development for a lunar base. Once a clear consensus exists to establish a base on the Moon, development of the components of a lunar VLA could be started and would be ready to be among the first large shipments of non-life-support systems to the Moon. Assembly and development time at the lunar base would depend on the details of the design and on the philosophy of lunar base operations.

Finally, it is clear that a large astronomical community would use the instrument. All the major astronomical facilities on Earth are heavily subscribed, and the VLA probably supports more users than any other astronomical instrument today. An interferometric array has many possible modes of operation: it can take brief snapshots; it can be broken into subarrays to serve multiple user groups simultaneously for specialized projects; and it can interweave long observing sequences with short projects in an efficient fashion. The VLA supports the observing programs of over a thousand scientists per year, and a lunar-based optical equivalent could be expected to do the same.

REFERENCES

Dupree A. K., Baliunas S. L., and Guinan E. F. (1984) Stars, atmospheres, and shells: Potential for high resolution imaging. *Bull. Am. Astron. Soc., 16*, 797.

Labeyrie A. (1980) Interferometry with arrays of large-aperture ground-based telescopes. *Proc. Optical and Infrared Telescopes for the 1990s* (A. Hewitt, ed.), p. 786. Kitt Peak National Observatory, Tucson.

Labeyrie A., Authier B., Boit J. L., DeGraauw T., Kibblewhite E., Koechlin L., Rabout P., and Weigelt G. (1984) TRIO: A kilometric optical array controlled by solar sails. *Bull. Am. Astron. Soc., 16*, 828.

Michelson A. A. (1920) An interferometer for measuring stellar diameter. *Astrophys. J., 51*, 257.

Napier P. J., Thompson A. R., and Ekers R. D. (1983) The VLA: A large aperture synthesis interferometer. *IEEE Proc., 71*, 1295.

Readhead A. C. S. and Wilkinson P. N. (1978) Phase closure in VLBI. *Astrophys. J., 223*, 25.

Rogers A. E. E. (1976) The two-element interferometer. *Methods of Experimental Physics, Vol. 12C: Astrophysics, Radio Observations* (M. L. Meed, ed.), p. 139.

Ryle M. and Hewish A. (1960) The synthesis of a large radio telescope. *MNRAS, 120*, 220–230.

Shao M., Colavita M., Staelin D., and Johnston K. (1984) The technology requirements for a small space-based astrometric interferometer. *Bull. Am. Astron. Soc., 16*, 750.

Stachnik R. V., Ashlin K., and Hamilton S. (1984) Space station SAMSI: A spacecraft array for Michelson spatial interferometry. *Bull. Am. Astron. Soc., 16*, 818.

Traub W. A. and Carleton N. P. (1984) COSMIC: A high-resolution large collecting area telescope. *Bull Am. Astron. Soc., 16*, 810.

A MOON-EARTH RADIO INTERFEROMETER

Jack O. Burns

Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131

In this paper, the logistical considerations and astronomical applications of placing a radio antenna(s) on the Moon as one element of a Moon-Earth Radio Interferometer (MERI) are described. Diffraction, interstellar scintillation, and Compton scattering are considered as processes that will influence the resolution of the interferometer. Compact sources with flux density less than 30 milliJanskys (mJy) can be observed at the optimum resolution of <30 microarcsec for wavelengths <6 cm. With such a resolution, one could perform fundamental astrometry experiments leading to a much improved value for the Hubble constant, or possibly map active regions on other stars and investigate with unprecedented linear resolution the nature of the "engine" at the center of the Milky Way and in active galaxies.

INTRODUCTION

The technique of radio interferometry and, in particular, Earth-Rotation Aperture Synthesis has proven to be enormously successful for ground-based radio astronomy. Operating radio interferometers include the MERLIN and 5-km Cambridge arrays in England, and the Westerbork Synthesis Radio telescope in the Netherlands. The most sophisticated aperture synthesis telescope is the Very Large Array (VLA) located in west-central New Mexico. It is composed of 27 individual antennas arranged in a Y-configuration (*e.g.*, Napier *et al.*, 1983). Each pair of radio antennas samples a particular Fourier component of the radio source brightness distribution at a given instant in time. As illustrated in Fig. 1, the turning of the Earth on its axis effectively synthesizes an aperture with resolution comparable to a single antenna with diameter equal to the maximum baseline between the outermost dishes. This is obviously a cheaper and more practical method for achieving high resolution mapping of extraterrestrial radio sources. The Fourier components gathered during a typical 12-hour integration are Fourier inverted with a computer algorithm (typically a gridded FFT) to produce a map of the sky brightness distribution. The accuracy of this map will depend upon the density of points in the Fourier transform plane (*i.e.*, how well the aperture was synthesized with the available antennas and the length of the integration). Typical maximum resolution for a high declination radio source using the VLA is 0.3 arcsec at 6 cm (35-km baseline) and the dynamic range (peak signal to RMS noise) can be thousands to one.

This technique has been extended to even longer baselines, termed Very Long Baseline Interferometry (VLBI), with individual antennas separated by entire continents. Data are recorded at each antenna on high speed videotape with accurate time markers as determined by hydrogen maser clocks. All the tapes are later brought together at a central computer processor and the data are correlated. Problems arise with the stability of the correlated phase due to differences in tropospheric and ionospheric refraction over the

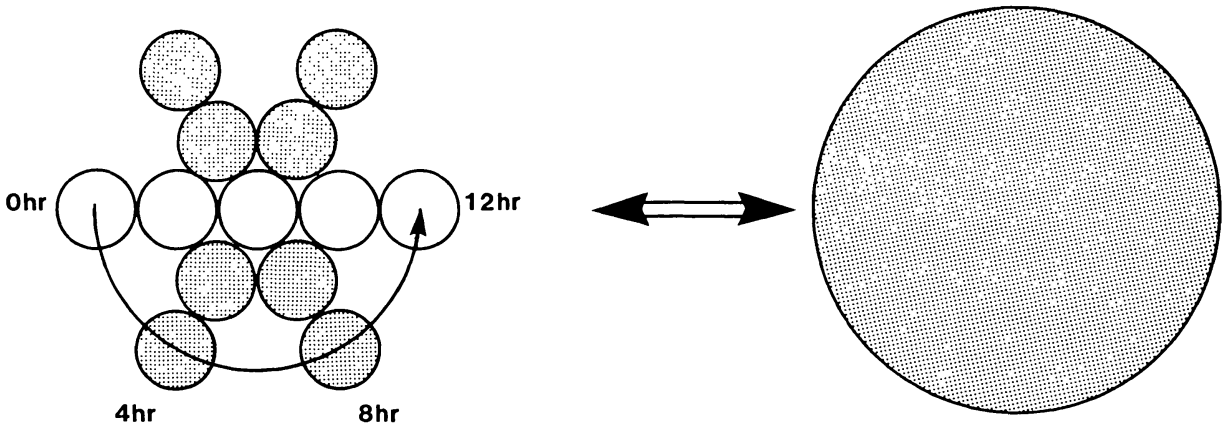

Figure 1. The principle behind Earth rotation aperture synthesis in radio interferometry. Imagine that an observer is stationed above the North Pole of the Earth looking down upon a linear alignment of five antennas. As the Earth rotates, the line sweeps out portions of a filled aperture. In 12 hours, the line has synthesized a circular aperture with diameter equal to the maximum baseline between the outermost antennas. This is equivalent to observing a radio source with a single very large antenna.

different telescope sites, and due to fluctuations in the local oscillator clocks. However, using closure phase and hybrid mapping techniques, it is now possible to recover an accurate map of the source brightness distribution and the relative positioning of radio features (*e.g.*, Pearson and Readhead, 1984). The accuracy and dynamic range depends upon the number and uniformity of antennas in the VLBI network. The proposed Very Long Baseline Array (VLBA), to be operated by the National Radio Astronomy Observatory, will consist of ten identical 25-m dishes located between Hawaii and Puerto Rico, and centered in New Mexico at the VLA. The resolution is expected to be at the submilliarcsecond level at centimeter wavelengths.

The maximum baseline for a ground-based VLBI is obviously limited to the diameter of the Earth. Also, both the European VLBI network and the American VLBA are oriented in a predominantly East-West direction, producing poor sampling of Fourier components in a North-South direction and poor sampling for southerly declination sources. An additional VLBI antenna in space could both increase the resolution and the image restoration accuracy when linked to a ground-based array. Recently, a European-American collaboration has developed a proposal for such a space-based antenna, called QUASAT (quasar satellite), which would have an elliptical orbit with a semi-major axis of 10,000 km and inclined 45° with respect to the equator (Schilizzi *et al.*, 1984). The resolution is expected to be about 350 microarcsec at 6 cm wavelength.

In principle, there is no reason why this VLBI technique could not be applied to baselines ranging between the Earth and the Moon. The major difference would be the observational technique used to synthesize the aperture. For a Moon-Earth Radio

Interferometer (MERI), the Moon's revolution around the Earth would provide the mechanism for the synthesis. Observations would be scattered over two-week intervals rather than the typical 12-hour continuous integrations that are now performed with the VLA and for VLBI. A first-generation antenna on the Moon could be quite simple and inexpensive, unfolding like a flower or an umbrella from the cargo bay of a Moon shuttle craft. Such an antenna, operating at centimeter wavelengths and possibly 10–15 m in diameter, will be needed almost immediately on the Moon for communications and data telemetry. Some time on this antenna could be initially "bootlegged" for tests of the MERI concept. This is an exciting prospect that is relatively inexpensive and promises to yield important new science. This paper examines the advantages of a telescope on the Moon as one element of a MERI, the practical considerations of wavelengths and receivers for setting up such an interferometer, and the scientific merits of this project.

ADVANTAGES OF A MERI

The most obvious advance over previously existing interferometers would be the improvement in resolution. At 6-cm wavelength, the resolution is 30 microarcsec, a factor of ten better than that proposed for QUASAT, 30 times better than the VLBA, and 10,000 times better than the VLA. More discussion on this point is given in the next section.

The thermal stability for an antenna on the Moon would be far better than that of an Earth-orbiting satellite dish. The MERI antenna would experience constant illumination from the sun in two-week intervals. Such thermal stability would be desirable to maintain constant pointing accuracy during a synthesis observation.

A larger antenna (say, 100 m) or a subarray of antennas on the Moon could be built out of relatively simple materials mined on the Moon. This is in keeping with the spirit of self-sufficiency for an advanced Moon colony. Once the mining and manufacturing techniques are developed, fabrication of dish antennas on the Moon will be far cheaper than transporting them from Earth.

Finally, one expects a long lifetime for the antennas on the Moon in comparison to that for a space satellite dish. The main advantage will be easy access to the Moon antenna(s) for repair and, particularly, for cryogen resupply. The lack of weather and low gravity on the Moon should minimize maintenance of the structure of the antennas.

CONSTRAINTS ON MERI

In considering the optimum wavelength at which the receivers will operate and the sensitivity required for interesting science, there are three constraints placed upon MERI. The first is the diffraction limit of the interferometer, which is simply given by

$$\theta_D = 5.47\lambda \quad (1)$$

where θ_D is the FWHM point response function of the instrument (microarcsec) and λ is the wavelength (centimeters). An average Earth-Moon baseline of 3.8×10^{10} cm is assumed.

The second constraint, involving scintillation of the interstellar medium (ISM), will also limit the resolution of MERI. Turbulence in the ISM (a plasma) within our galaxy effectively scatters radio radiation from distant sources, broadening the radio "seeing" disk in a manner somewhat analogous to the "twinkling" of stars produced by the passage of optical light through the turbulent atmosphere of the Earth. The predicted amount of scattering depends critically upon the galactic latitude of the radio source and the assumed model for electron density fluctuations. For sources high above the galactic plane and assuming a simple power-law spectrum for the turbulence following Rickett (1977) and Cordes *et al.* (1984), the scattering angle is given by

$$\theta_{ISM} = 0.60\lambda^{11/5} \quad (2)$$

where θ_{ISM} is again measured in microarcsec and λ is the wavelength in centimeters.

The final constraint is a theoretical limit placed upon the intrinsic sizes of compact radio sources, sometimes referred to as the Compton catastrophe or 10^{12} K brightness temperature limitation. The emission mechanism for active galaxies and quasars is believed to be incoherent synchrotron radiation at radio wavelengths. For compact sources, the densities of relativistic electrons and synchrotron photons are high. Inverse Compton scattering of the photons by electrons is plausible in this environment. Within a limiting angular distance of the core, the optical depth for this Compton process is so high that few, if any, synchrotron photons will escape from this region. Therefore, no information on the structure within this scattering disk can be retrieved. This angle depends upon both the wavelength and the flux density, S (measured in milliJanskys where 1 mJy = 10^{-29} watts/Hz/m^2), such that brighter sources will have larger scattering disks. This limit is given by

$$\theta_{cc} \geq S^{1/2}\lambda \quad (3)$$

This limit will not apply to extended, optically thin sources.

A plot of these three effects is shown in Figure 2. One can see from the above equations that the Compton catastrophe limit is equal to the diffraction limit for radio sources with flux densities of 30 mJy. Furthermore, the ISM scintillation effects are negligible below a wavelength of about 6 cm. Therefore, to achieve the ideal resolution (*i.e.*, diffraction limit), one would like to observe compact sources (for astrometry) of <30 mJy at a wavelength of 6 cm or below.

The fortuitous combination of baselines and wavelengths makes the 6-cm band ideal for MERI (although one should not overlook the potential for operating at much shorter wavelengths). The 6-cm receivers currently available are some of the most sensitive used in radio astronomy, with system temeratures of <50 K. Such sensitivity will be very useful since we desire to observe relatively weak radio sources in the range of a few tens of mJy.

As noted above, we can achieve the diffraction limit for sources at high galactic latitudes with flux densities <30 mJy at 6 cm. This angular resolution of 30 microarcsec corresponds to the following linear dimensions:

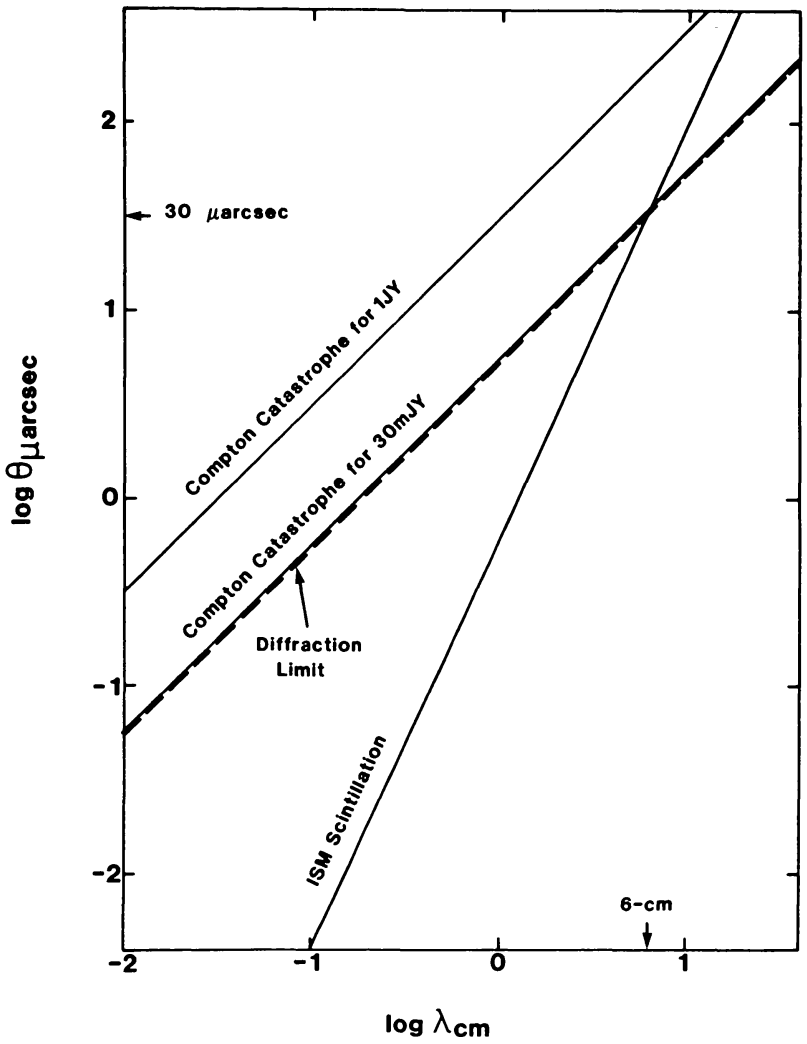

Figure 2. A plot of the minimum FWHM resolution of MERI vs. wavelength. Note that the optimum resolution of 30 microarcseconds occurs in the 6-cm band for compact sources of 30 mJy. ISM scattering effects are negligible for high galactic latitude sources observed at wavelengths <6 cm.

1. At a distance of 300 pc (~240 times the distance to the nearest star), the disk of a solar-type star can be resolved. Studying active regions on such stars will be invaluable in understanding the solar-stellar connection.

2. At the distance of the Galactic center, the linear resolution will be 0.15 astronomical units (1 au = average Earth-sun distance = 1.5×10^{13} cm). With such a resolution, one will be able to "look down the throat of the beast" that is responsible for prodigious amounts of electromagnetic radiation at all wavelengths. This linear dimension corresponds to 760 Schwarzschild radii for a 10^4 solar mass black hole.

3. At the distance of the nearest radio galaxy, Centaurus A, the linear resolution is 450 au. Again one can seriously investigate the nature of the engine at the cores of active galaxies with such a resolution.

Can we achieve the sensitivity necessary to observe sources of a few tens of mJy? The integration time needed to obtain a signal-to-noise ratio, S/N, with radio antennas of diameter D, antenna temperatures T, antenna efficiencies ϵ, and bandwidths $\Delta\nu$ for a source with flux density S on a single Moon-Earth baseline is given by

$$t(s) = 3.1\times 10^{7}(S/N)^{2}(T_{m}T_{e})\ (\epsilon_{m}^{2}\epsilon_{e}^{2}D_{m}^{2}D_{e}^{2}S^{2}\Delta\upsilon)^{-1} \quad (4)$$

Here, the subscript m refers to a Moon antenna and e to an Earth antenna. If we assume equal radio dishes on the Moon and the Earth, and S/N = 10, D = 25 meters (VLA and VLBA size dish), T = 50 K, ϵ = 0.65 (at 6 cm), $\Delta\nu$ = 30 MHz, and S = 30 mJy, the integration time is only 30 minutes. If the Moon radio telescope is linked to one of the ground-based VLBI arrays, the integration time can be further reduced by a factor of N(N–1)/2, where N is the total number of antennas in the network. In either case, the integration time is quite reasonable.

To review, then, our analysis shows that the diffraction-limited resolution can be achieved for wavelengths less than 6 cm for compact sources at high galactic latitudes with flux densities <30 mJy. This requires a moderate aperture radio antenna (~25 meters) linked to a VLBI array on Earth using a wide bandwidth system. It is important to note that *no new technology* is required for MERI as outlined above. However, if aperture synthesis mapping of radio sources is to be seriously attempted at these resolutions, then additional antennas should be placed in orbits between the Earth and Moon.

SCIENTIFIC GOALS

There is a wealth of scientific data that could be collected with a MERI telescope that would significantly add to our knowledge of the local environment and the cosmos. These goals are divided into two parts that will depend upon the number of elements in MERI: astrometry and synthesis mapping.

Astrometry

First, the unprecedented relative position accuracy of MERI could be used to improve the celestial coordinate system, thereby improving celestial navigation and astronomical timekeeping.

Second, observations of point sources could potentially be used to accurately measure distances between the Earth and the Moon. In principle, baseline determinations with millimeter accuracy could improve by an order of magnitude those measured by laser ranging. Improved-stability maser clocks would be required, however.

Third, it may be possible to search for dark companion stars (black holes and neutron stars) or even planets around radio stars. One could simply look for perturbations of the radio star proper motions produced by the gravitational pull of a dark binary companion(s). The few tens of microarcsecond accuracy of MERI would allow very small perturbations (produced by planets of mass less than that of Jupiter) to be detected.

Fourth, one of the most exciting aspects of science with MERI is the fundamental cosmological experiments that could be performed. At present, the Hubble constant, which measures the rate of expansion of the universe at the current epoch, is not known to within a factor of two. Moran (1984) has shown that H_2O masers in our galaxy can be used as independent distance measures using classical proper motion and statistical parallax techniques. The impressive power output of these radio sources at discrete

wavelengths makes them ideal for this purpose. The angular resolution of MERI could enable astronomers to extend this technique to other galaxies and to accurately determine their distances independent of the other less reliable assumptions currently invoked (Reid, 1984). A measure of the Hubble constant would follow from a well-defined statistical sample. Similarly, proper motions of radio galaxies in clusters combined with redshift measurements could allow statistical determinations of cluster distances and, therefore, the Hubble constant.

Synthesis Mapping

If a few radio antennas between the Earth and the Moon could be added to the initial single antenna on the Moon, then synthesis mapping at <30 microarcsec resolution becomes feasible. The following is a short list of the possible mapping projects in order of increasing distance and decreasing linear resolution.

First, one could potentially locate and map radio burst regions on other stars. The primary limitation here is sensitivity. One would probably need a 100-m antenna on the Moon to make this feasible.

Second, resolving regions in and around star formation nebulae would become possible with MERI. This could make significant impacts on our understanding of the early stages of star (and possible planetary) formation. The origins of the bipolar flows recently discovered in star formation zones could be explored.

Third, the center of the Milky Way contains a powerful source of electromagnetic energy. This energy, $\sim 10^{42}$ ergs/sec, emanates from a region of only a few light years across. Mapping this region at the resolution of MERI would almost certainly add new insights if not a definitive answer to our questions concerning the "engine."

Fourth, active galaxies and quasars are now known to possess "jets" of radio emission that appear to illuminate channels by which matter and energy are transported between the engine at the core of the galaxy (or quasar) and extended structures hundreds of thousands of light years out in intergalactic space. We do not understand how or why the radio jets are colliminated as they are in two thin streams. The resolutions of current interferometers are simply too low to explore the nuclear regions where the initial collimation occurs. MERI will allow us to explore these regions in unprecedented detail.

Fifth, in a related vein, the engines themselves are not understood. Are galactic sources such as SS 433 and the Milky Way center simply scaled-down versions of those in the more powerful active galaxies? With MERI, we will be able to map the core regions at high enough resolution to address the nature of the engine.

Sixth, we could potentially test the fundamental physics of compact extragalactic sources. In particular, the Compton catastrophe predicts a minimum size for compact sources of a given flux density at a particular wavelength. We need to test this prediction. Observing compact sources at larger wavelengths and/or higher flux densities will provide the definitive test of this model.

MERI offers us an opportunity for a major leap forward in radio astronomy, both in terms of technique and science. As I have noted, the costs of a single lunar antenna are minimal but the science could be potentially quite promising. Therefore, I would hope

that a single radio antenna linked as an interferometer to Earth-based radio telescopes would be one of the first "flowers" planted on mankind's return to the Moon.

Acknowledgments. *I wish to thank Pat Crane, Ed Fomalont, Steve Gregory, Frazer Owen, Marc Price, Peter Wilkinson, and Stan Zisk for useful input into the preparation of this paper. I would also like to thank the Lunar and Planetary Institute for travel support to the Symposium.*

REFERENCES

Cordes J. M., Ananthakrishnan S., and Dennison B. (1984) Radio wave scattering in the galactic disk. *Nature, 309*, 689.

Moran J. M. (1984) Masers in the nuclei of galaxies. *Nature, 310*, 270.

Napier P. J., Thompson A. R., and Ekers R. D. (1983) The Very Large Array: Design and performance of a modern synthesis radio telescope. *Proc. I.E.E.E., 71*, 1295.

Pearson T. J. and Readhead A. C. S. (1984) Image formation by self-calibration in radio astronomy. *Ann. Rev. Astron. Astrophys., 22*, 97.

Reid M. J. (1984) H_2O masers and distance measurements: The impact of QUASAT. In *QUASAT—a VLBI Observatory in Space*, p. 181. ESA Scientific Publication, The Netherlands.

Rickett B. J. A. (1977) Interstellar scattering and scintillation of radio sources. *Ann. Rev. Astron. Astrophys., 15*, 479.

Schilizzi R. T., Burke B. F., Booth R. S., Preston R. A., Wilkinson R. N., Jordan J. F., Preuss E., and Roberts D. (1984) The QUASAT project. In *VLBI and Compact Radio Sources* (R. Fanti, K. Kellerman, and G. Setti, eds.), p. 407. D. Reidel, Boston.

A VERY LOW FREQUENCY RADIO ASTRONOMY OBSERVATORY ON THE MOON

James N. Douglas and Harlan J. Smith

Astronomy Department, University of Texas, Austin, TX 78712

Because of terrestrial ionospheric absorption, very little is known of the radio sky beyond 10 m wavelength. We propose an extremely simple, low-cost Very Low Frequency (VLF) radio telescope, consisting of a large (approximately 15 × 30 km) array of short wires laid on the lunar surface, each equipped with an amplifier and digitizer, and connected to a common computer. The telescope could do simultaneous multifrequency observations of much of the visible sky with high resolution in the 10- to 100-m wavelength range, and with lower resolution in the 100- toward 1000-m range. It would explore structure and spectra of galactic and extragalatic point sources, objects, and clouds, and would produce a detailed quasi-three-dimensional mapping of interstellar matter within several thousand parsecs of the sun.

INTRODUCTION

The spectral window through which ground-based radio astronomers can make observations spans about five decades of wavelength, from a bit less than a millimeter to something more than ten meters. The millimeter cutoff produced by molecular absorption in the Earth's atmosphere is fairly stable, but the long wavelength cutoff caused by the terrestrial ionosphere is highly variable with sunspot-cycle, annual, and diurnal effects; scintillation on much shorter time scales is also present. Radio frequency interference imposes further limits, making observations at wavelengths longer than 10 m normally frustrating and frequently impossible.

Consequently, the radio sky at wavelengths longer than 10 m is poorly observed and is virtually unknown for wavelengths longer than 30 m, except for a few observations with extremely poor resolution made from satellites. Exploration of the radio sky at wavelengths longer than 30 m must be done from beyond the Earth's ionosphere, preferably from the farside of the Moon, where physical shielding completes the removal of natural and manmade terrestrial interference that the inverse square law has already greatly weakened.

THE LONG WAVELENGTH RADIO SKY

What may we expect in the long-wavelength radio sky (apart from the unexpected, which experience often shows to be more important)?

First, non-thermal radiation from plasma instabilities in solar system objects is present in rich variety, especially from the sun, Jupiter, Saturn, and Earth itself (which was unexpectedly discovered by telescopes flown for other purposes).

Second, the synchrotron radiation from the galaxy reaches a peak of intensity near 4 MHz (75 m), then drops off as absorption by ionized hydrogen becomes important, and possibly for other reasons, as well. This behavior has been seen by the low-resolution (20° beam) telescopes already flown.

Third, the plane of the Milky Way—already dimming at 10 m—becomes even more absorbed by ionized hydrogen, and many black blots of HII regions are seen in absorption against the bright radiation background. Such clouds, whose emission measure is too small to be noticed optically, would be obvious using a moderate to high resolution (1°–0.1°) VLF telescope. At longer wavelengths, our distance penetration becomes increasingly limited, decreasing with the square of the wavelength until, by 300 m, unit optical depth corresponds to only a few hundred parsecs.

Fourth, extragalactic discrete sources continue to be visible as wavelength increases (outside the gradually expanding zone of avoidance at low galactic latitudes), and their spectra can be measured, although their angular structure will be increasingly distorted by interstellar and interplanetary scattering. At these wavelengths, one is looking at the expanded halo parts of such objects, and turn-overs will be noted in the spectra of many. For wavelengths longer than about 300 m, HII absorption in our own galaxy will effectively prevent extragalactic observations, even at the galactic poles, and, at wavelengths longer than a kilometer or so, we will be limited to studying objects within a few tens of parsecs of the sun.

Finally, there are possible new features of the sky that can be studied only by a high resolution and high sensitivity telescope at long wavelengths. These include non-thermal emission from stars and planets or other such sources within a few parsecs (if any of these are significantly more powerful than the sun and Earth); radio emission of very steep spectra from new classes of galactic or extragalactic discrete sources that may have gone undetected to date in even the faintest surveys at short wavelengths, yet be detectably strong at 100 m; nearby and compact gas clouds, visible in absorption, whose presence has hitherto been unsuspected; and fine-scale structure in the galactic emission, which—given data at high-resolution and multiple low-frequencies—can be studied in depth as well as direction. In this connection the proposed telescope should provide a uniquely detailed and effectively three-dimensional map of interstellar matter in the galaxy out to distances of thousands of parsecs.

THE LUNAR VLF OBSERVATORY

As noted above, the low-frequency telescopes flown to date have had very poor resolution, although valuable for some studies on very bright sources, *e.g.*, dynamic spectra of the sun, Earth, Jupiter, and the cosmic noise spectrum. Significant advances, however, will require high resolution (say 1°, corresponding to 15 km aperture at 300 m wavelength) and high sensitivity (many elements). A lunar base offers probably the best location in the solar system for constructing an efficient low-cost VLF radio telescope.

In contemplating any lunar-based experiment, the question must first be asked whether it is preferable to carry out the work in free space. For the proposed VLF observatory, the Moon offers a number of advantages:

1. It is a fine platform, able to hold very large numbers of antenna elements in perfectly stable relative positions over tens or even hundreds of kilometers separation (this would be excessively difficult and expensive to try to do in orbital configuration);

2. The telescope can begin modestly, though still usefully, and can continue to grow to include thousands of antenna elements added in the course of traverses of lunar terrain undertaken at least in part for other purposes;

3. The dry dielectric lunar regolith permits simply laying the short thin-wire antenna elements on the surface. No structures, difficult to build and maintain, are required;

4. Lunar rotation provides a monthly scan of the sky;

5. The lunar farside is shielded from terrestrial interference, although even the nearside offers orders-of-magnitude improvement over Earth orbit because of the inverse square law, and the much smaller solid angle in the sky presented by the Earth.

Limiting Factors

Various natural factors limit the performance of a lunar VLF observatory.

Long wavelength limits. (1) Interplanetary plasma at 1 AU has about 5 electrons/cm^3 corresponding to a plasma frequency (f_p) of 20 kHz, or a wavelength of 15 km. (2) The Moon may have an ionosphere of much higher density than the solar wind; 10^{-12} torr corresponds to about 40,000 particles/cm^3, if the mean molecular weight is 20. If such an atmosphere were fully singly ionized, f_p would be around 1.8 MHz, usefully but not vastly better than the typical values for the Earth of around 9 MHz. However, ground-based observations of lunar occultations suggest that N_e is actually less than 100. In this case, f_p would be less than 90 KHz (wavelength 4 km), and would set no practical limit to very low frequency lunar radio astronomy. It will clearly be very important for detailed planning of the lunar VLF observatory to have good measures of the lunar mean electron density and its diurnal variations.

Scattering. (1) The interstellar medium produces scattering and scintillation, and thus angular broadening of sources—e.g., the angular size of an extragalactic point source would be about 8 arcseconds if observed at 30 m wavelength. The size grows with wavelength to the 2.2 power, becoming 1/3° at 300 m. (2) Interplanetary scintillation is more important. Obeying essentially the same wavelength dependence as interstellar scintillation, it ranges from about 50 arcsec at 30 m to a few degrees at 300 m (1 MHz). However, it is still worthwhile designing the telescope with higher resolution than 2° at 1 MHz, since techniques analogous to speckle interferometry may recover resolution down to the limits set by interstellar scintillation, which will be relatively small especially for nearby sources in our galaxy.

Interference. (1) Solar: the intensity of the cosmic background radiation is on the order of 10^{-15} w m^{-2} Hz^{-1} $ster^{-1}$. The sun is already known to emit bursts stronger than this by an order of magnitude in the VLF range, so the most sensitive observations may have to be carried out during lunar night. (2) Terrestrial: Nearside location will always expose the telescope to terrestrial radiations. Consider two known types: auroral kilometric radiation is strong between 100 and 600 kHz; an extremely strong burst would produce flux density at the Moon of about $2(10)^{-15}$ w m^{-2} Hz^{-1}—far stronger than the cosmic

noise we are trying to study. Fortunately it is sporadic, and limited to low frequencies. Also, it probably comes from fairly small areas in the auroral zones, so that its angular size as seen from the Moon will be small. Lunar VLF observations below 1 MHz will therefore be limited unless the telescope is highly directive with very low sidelobes, or built on the lunar farside.

Terrestrial radio transmitters may leak through the ionosphere in the short wavelength portions of the spectrum of interest. If we assume a 1-Mw transmitter on Earth with a 10 kHz bandwidth, the flux density at the Moon would be about $5(10)^{-17}$ w m^{-2} Hz^{-1} without allowing for ionospheric shielding. This would be a serious problem; much weaker transmitters with some ionospheric shielding would merely be an occasional nuisance. Again, this is an argument in favor of a farside location, particularly for frequencies above 4 MHz or so.

Considerations of Telescope Design

It would be futile to carry out a detailed telescope design at this point; however some general considerations can be addressed:

Frequency range. The telescope should be broadband, but capable of observing in very narrow bands over the broad range to deal with narrow-band interference. The upper limit of frequency should be around 10 MHz or 30 m. Even though this wavelength can be observed from the ground, it is extraordinarily difficult to do so. The initial normal lower limit should be about 1 MHz or 300 m, although the capability for extending observations with reduced resolution to substantially longer wavelengths should be retained.

Resolution. It is probably useless to attempt resolution at any given frequency better than the limit imposed by interstellar scintillation, e.g., about 1/3° at 1 MHz. A reasonable initial target resolution for the observatory might be 1° at 1 MHz. Although this is somewhat better resolution than the limit normally set by interplanetary scintillation, it is probably attainable using restoration procedures. This choice of target resolution implies antenna dimensions of 15×15 km for a square filled array, or of 30×15 km for a T configuration.

Filling factor. A 1° beam may be synthesized from a completely filled aperture (100×100 elements, for a total of 10^4) or by a T, one arm of which has 200 elements, the other 100, for a total of 300 elements—far less. Many other ways of filling a dilute aperture also exist, including a purely random scattering of elements over the aperture. The filled array has far greater sensitivity, but, what is also important in this context, it has much better dynamic range and a cleaner main beam. This will be of great benefit in mapping the galactic background, particularly in looking at the regions of absorption, which will be of such interest at these frequencies. The sensitivity of the filled array is also decidely better: a 1° beam produced by a filled aperture at 1 MHz with a bandwidth of 1 kHz and an integration of 1 min has an rms sensitivity of 1 Jy; the same sensitivity would require an integration of 1 day with the dilute array of 300 elements. The most sensible approach is probably to begin with a dilute aperture and work toward the filled one, the power of the system increasing as more antenna elements are set out.

Telescope construction. The telescope would be an array of many elements. Each element should be thought of as a field sensor rather than as an ordinary beam-forming

antenna—in other words as a very short dipole. The inefficiency of such devices can be great before noise of the succeeding electronics becomes a factor, in view of the high brightness temperature of the cosmic background radiation. An A/D converter at each element would put the telescope on a digital footing immediately. The exceedingly low power requirement at each antenna element could be met with a tiny solar-powered battery large enough to carry its element through the lunar night.

Communication with the telescope computer at lunar base via radio or perhaps by individual optical-fiber links would bring all elements together for correlation. Bandwidth of the links need only be about 1 kHz per element if only one frequency is to be observed at a time, although maximum bandwidth consistent with economics will produce maximal simultaneous frequency coverage. In any event the central computer will produce instant images of a large part of the visible hemisphere with the 1° resolution, at one or many frequencies, which can be processed for removal of radio frequency interference (rfi) and bursts prior to long integrations for sky maps at various frequencies in the sensitivity range of the system.

Short wavelength operation. Operation at the short-wavelength boundary of the telescope range will be a different proposition. Element spacing for 1 MHz is very dilute indeed for 10 MHz; some portion will have to be more densely filled, and operated against the rest of the system as a dilute aperture. At 10 MHz the system would have a resolution of about 0.1°. In this way, an extremely powerful telescope for work both on extragalactic sources and on galactic structure would result.

ESTABLISHING THE VLF OBSERVATORY

The individual antenna elements—short wires—will probably weigh about 50 gm each. Their associated microminiaturized amplifiers, digitizers, transmitters, and solar batteries can all be on several tiny chips in a package of similar weight. Allowing for packaging for shipment to the Moon, the initial array should still weigh less than 50 kilograms! Materials for the entire filled array would only need about a ton of payload. If individual optical fiber couplings to the central computer are used, each of these should add only a few tens of grams to the total, not appreciably affecting the extraordinarily small cost of transporting the system to the Moon.

A powerful computer is of course required, to process continuously the full stream of digital information. Some on-base short-term storage of processed data is probably also desirable, but at frequent intervals this would presumably be dumped back to Earth. Again, with the increasing miniaturization yet steady growth in power of computer hardware over the next twenty years, the required computer facilities may also be expected to weigh less than a hundred kilograms. It thus seems clear that at least the initial, and quite possibly the ultimate, VLF observatory system could be carried to the Moon as a rather modest part of the very first scientific payload.

Laying out the initial system of several hundred antenna elements on the lunar regolith should require only a few days of work with the aid of an upgraded lunar rover having appropriate speed and range. (Such vehicles will be an essential adjunct of any lunar

base for exploration, geological and other studies, and general service activities). The elements need not be placed in accurately predetermined positions, but their actual relative positions need to be known to a precision of about a meter. This can easily be done, as the layout proceeds, by surveying with a laser geodometer. The conspicuous tire marks produced by the rover vehicle will delineate the sites of each of the antenna elements for future maintenance or expansion of the system. A concentrated month using two vehicles each carrying teams of perhaps three workers would probably suffice to lay out the full proposed field of 100×100 elements.

These estimates, while necessarily rough at this preliminary stage of planning, strongly suggest that because of its extreme simplicity and economy, its almost unique suitability for lunar deployment, and its high scientific promise, the VLF observatory is a major contender for being the initial lunar observatory—perhaps even the first substantial scientific project that should be undertaken from Lunar Base.

LUNAR BASED GAMMA RAY ASTRONOMY

Robert C. Haymes

Space Physics and Astronomy Department, Rice University

INTRODUCTION

Gamma ray astronomy is the study of the universe through analysis of the information carried by the highest energy electromagnetic radiation. The energy of the photons it analyzes ranges upwards from about 0.1 MeV, the upper energy end of the x-ray band, through the MeV energies of nuclear transitions and radioactivity, up through the GeV energies of cosmic ray-matter interactions, to beyond the 10^3TeV energies radiated by the highest energy galactic cosmic rays. The celestial sources where gamma ray emission is a major fraction of the energy release are some of the most bizarre, energetic objects in the universe, including supernovae, neutron stars, black holes, galactic cores, and quasars.

SIGNIFICANCE

If the gamma ray photon fluxes are above the sensitivity threshold, then a variety of phenomena may be studied through their measurement. We currently believe that nucleosynthesis of the heavier elements takes place in *supernovae*, stellar catastrophic explosions that each rival a whole galaxy of stars in brightness. Supernovae are expected to emit specific gamma ray energy spectra from which the mode(s) of nucleosynthesis may be deduced (Ramaty and Lingenfelter, 1982).

Black holes are places where present-day physics is, at best, on shaky ground. Gamma ray astronomy permits us to study matter as it falls into a black hole, because the matter becomes heated to gamma ray temperatures as it does so.

Active galaxies generate energies comparable with their relativistic self energy, but the nature of their energy source(s) is unknown. The various theoretical possibilities suggested thus far all predict different gamma ray spectra.

Quasars may each generate as much energy as do ten billion stars, but they appear somehow to do it in a volume not much larger than that of only one star. Much of the quasar's radiation is in the gamma ray band, and our most important clues to the phenomenon may therefore come from such astronomy.

The sources of *cosmic rays* have long been mysterious. Whatever and wherever they are, the sources are likely to generate high-energy gamma rays because of the acceleration of the charged-particle cosmic rays. Because photons are uncharged and therefore travel in straight lines that are unaffected by magnetic fields en route, gamma ray astronomy

uniquely offers the opportunity of at long last locating and studying the sources of this ubiquitous, extreme-energy particle radiation.

Quantized cyclotron emission may already have been detected from the presumed direction of one *neutron star*. If confirmed, this evidence would provide for the existence of magnetic fields seven orders of magnitude more intense than any generated in the laboratory. *Gamma ray bursters* are possibly associated with neutron stars and have been known for over a decade, but their nature is as mysterious as ever. Also not understood is the nature of *transient sources* of cosmic gamma radiation.

Over thirty steady sources of high energy gamma radiation have already been detected, sources that do not seem to have counterparts in other spectral bands (*e.g.*, the optical, radio, and x-ray). These sources therefore seem capable of somehow accelerating charged particles to the extreme energies required for gamma ray production, while also suppressing the usually copious radiation of lower energy photons.

In addition to acceleration of ultrarelativistic charged particles, the gamma radiation is produced in several ways. These include radioactivity and nuclear de-excitation, matter-antimatter annihilation, decays of elementary particles, and some effects of relativity. Through this branch of astronomy, qualitatively different and often uniquely available information will be acquired on the mechanisms, history, and sites of cosmic nucleosynthesis, the structure and dynamics of the Milky Way, the nature of pulsars, the sites and properties of intensely magnetized regions, the isotopic composition of the matter in the space surrounding black holes, the nature of the huge energy sources powering active galactic nuclei, and the universality, composition, and sources of very high-energy matter throughout the universe. We already know that, at least in some of these phenomena, the emissions are most luminous in the gamma ray part of the spectrum. Full understanding of these sources will require detailed study of their gamma radiation, in combination with study of the emissions in the other spectral bands.

SENSITIVITY LIMITATIONS FOR GAMMA RAY ASTRONOMY

In most situations, if a source radiates high-energy photons, it will radiate smaller fluxes of them than it will of lower energy photons. The sensitivity of observational gamma ray astronomy is limited by the small fluxes of gamma ray photons from celestial sources. It is also limited by the background.

Gamma ray astronomical background appears to have two components. One component is a sky (possibly cosmic) background. The other component has two sources. One arises from both ambient gamma radiation, and the other from radioactivity induced in the observing instrument by the particle radiation environment that exists in space.

Gamma ray astronomy is best conducted far from Earth, because the atmosphere is a source of gamma rays. Energetic particles continually bombard the atmosphere. Examples of such particles include cosmic rays and the high-energy protons that compose the Inner Van Allen belt; others are solar-flare accelerated ions. These particles produce gamma ray photons when they interact with the atmosphere. Some of the produced photons head out into space, forming a "gamma ray albedo." The ambient gamma radiation

comprising the albedo has an intensity that exceeds, by several orders of magnitude, the fluxes of gamma ray photons from even the brightest cosmic sources.

Bombardment by high-energy particles also activates the materials composing a gamma ray detector, making the instrument itself a source of the very radiation one is attempting to measure from the cosmos (*e.g.*, Paciesas *et al.*, 1983). The brightness of this background component is dependent on the mass of target matter and the magnitude of the bombarding fluxes. It may be reduced to cosmic ray levels by observing from sites that are outside the geomagnetically trapped radiation and that are shielded from solar protons.

GAMMA RAY ASTRONOMY OBSERVATIONS

Earth's atmosphere is an absorber of cosmic gamma radiation. To avoid unacceptable attenuation of the small photon fluxes from celestial sources, gamma ray astronomy must therefore be conducted from outside most, if not all, of the atmosphere. Fragmentary data have been acquired with balloon flights near the top of our atmosphere. Almost all of the balloon flights have been restricted to durations of a day or less, and all were sporadic. Even though most of the major advances in understanding the phenomena encountered in other spectral bands were made when long-term measurements were undertaken, there has been little systematic, long-term observational gamma ray astronomy.

The HEAO-1 and HEAO-3 satellites scanned the sky and undertook some short-duration observations of selected discrete gamma ray sources. The major long-term efforts thus far have been the European COS-B mission and the groundbased monitoring of Cerenkov pulses in Earth's atmosphere. For about six years in the 1970s, the low altitude COS-B (Mayer-Hasselwander *et al.*, 1982) satellite carried a relatively small spark chamber. The chamber measured the arrival directions and the energies of gamma rays in approximately the 0.1–5 GeV energy band. That mission has produced almost all of our present information on astronomical sources of gamma rays in that energy band.

All of our information on emissions in the 1–1000 TeV energy band, which may be providing our first direct looks at the sources of galactic cosmic rays and which could tell us how high in energy particle acceleration goes in the nuclear regions of active galaxies, has come from groundbased monitoring of nanosecond Cerenkov light pulses in the atmosphere. Some of these pulses are due to the interaction of very high energy gamma ray photons with the atmosphere (Samorski and Stamm, 1983). The relative frequency of such pulses increases when the gamma ray source transits the observer's meridian. These extreme-energy events are relatively rare in occurrence; several-year integration times are necessary for statistical validity. For absolute flux information, it is necessary to distinguish those showers produced by γ-ray photons from those due to cosmic ray nuclei. Such a distinction is now made on the basis of relative μ-meson richness in the showers.

In 1988, NASA plans to launch the Gamma Ray Observatory (GRO), the first full-fledged systematic investigation into gamma ray astronomy, into a low altitude orbit. The three-axis stabilized spacecraft will operate for two years. Data will be collected

from selected targets for one to two weeks at a time. The GRO will carry four gamma ray astronomy experiments in the one-square-meter class. Described by Kniffen (written communication, 1981), each experiment has different scientific objectives.

One GRO experiment is a spark chamber. Its collecting area for photons is about ten times greater than the chamber that flew on COS-B, and it will have about 1° angular resolution. Since the observing times will be comparable with the COS-B times, the spark chamber experiment, called EGRET, will therefore extend the sensitivity of the COS-B observations of 100 MeV–5 GeV cosmic sources and the gamma ray background.

The second GRO experiment, called COMPTEL, is an imaging double Compton telescope for 1–30 MeV studies. Very little is presently known about this spectral region, most of which lies above the energies of nuclear transitions and radioactivity but below the energy where decay of neutral π-mesons into gamma rays is an important source of photons. Compton scattering is a major photon-matter interaction mechanism at these energies.

Double Compton telescopes consist of two layers of scintillation counters. A gamma ray that interacts in the first layer generates a pulse. The recoil Compton photon, if its direction is suitable, interacts with the second layer. Delayed coincidence between the two layers is required for a given event to be accepted as due to a gamma ray in the direction of the instrumental cone of acceptance. The time between these two pulses is given by the ratio of the separation distance to the speed of the recoil, which is the speed of light. COMPTEL's spacing is about 3 m, so the delay time is about 10 nanoseconds. This is an excellent way of rejecting background; particles and photons from other directions will not give the correct time signature. Double Compton experiments, however, are inefficient (the photon-detection efficiency is of order 10^{-5}), since the recoils have to be directed only in the direction of the second layer for an event to be counted. They also do not require total energy deposition, and reliance must therefore be placed on calculations of most probable energy loss as a function of energy, in order to convert the observed pulse-height spectrum to an energy spectrum. COMPTEL's imaging will be crude at best, since its angular resolution is several degrees.

Third is OSSE, the Oriented Scintillation Spectrometer Experiment, which consists of four independently pointable, equal-area actively collimated scintillation counters. Its goal is to conduct astronomical spectroscopy in the 0.1–10 MeV energy band. This band is characteristic of radioactivity and transitions of excited atomic nuclei.

Actively collimated counters almost completely surround the photon counter with a thick collimator that consists of an efficient scintillation counter whose output is connected in anticoincidence with the photon counter. Absence of coincidence between the scintillation and photon counters is required for a photon-counter event to be accepted. Total deposition of energy in the photon counter is imposed by this requirement. The coincidence requirement also rejects counts due to charged particles; a particle that caused a count in the photon counter most likely had to traverse the scintillation counter-collimator in order to reach the photon counter, and it would have generated a pulse from the scintillation counter in order to do so. The magnitude of the field of view of such experiments is defined by the size of the opening in the "active collimator" (*i.e.*, the surrounding scintillation

counter). Although total energy-deposition is required and the photon-detection efficiency may be near 100%, activation of the crystals themselves by particle bombardment is a serious source of background. For example, the "collimator" is itself induced by the bombardment to become a radioactive source of gamma rays, which are not distinguished by the system from external cosmic gamma rays coming through the aperture. At high energies in the band, photon leakage through the finite thickness collimator is also a serious background source. To date, the noise (*i.e.*, background)-to-signal ratios of actively collimated astronomical experiments have typically been 10 or more, even for the brightest cosmic sources. The two brightest cosmic sources, the central region of the Galaxy and the Crab Nebula supernova remnant, have photon fluxes at 1 MeV that are of order 10^{-3} photons/cm^2-s. OSSE will attempt to reduce background effects by simultaneously measuring source and background, using its independently targetable 2000 cm^2 modules. OSSE has an angular resolution of about one degree and an energy-dependent energy resolution that is about 0.05 MeV at 1 MeV.

The fourth experiment, called BATSE, is primarily intended to measure cosmic gamma ray bursts with a lower fluence threshold than heretofore available. Its very large photon collection area also makes it a very sensitive detector of rapidly varying existing gamma ray sources, such as pulsars. BATSE consists of six one-square-meter scintillation counters that are each pointed in different directions from the stabilized spacecraft. They cover the entire hemisphere of sky. So far as is known, bursts seem to originate from all sky locations with equal probability; a burst that occurs anywhere on the hemisphere will illuminate all six counters differently. Each counter uses sodium iodide as the scintillator, for maximum light output and best energy resolution from a large-area detector. The ratios of their count rates will locate the burst on the sky to an accuracy of a degree or so. The time history of their count rates will measure the light curve of the event, and pulse height data from the six will provide some information on the energy spectrum of the emitted gamma radiation. Pulses from existing sources are part of BATSE's data stream. Sources that have known "signatures," such as pulsars that have known periods, may be sorted out from the other data and their energy dependence measured with good precision out to higher energies than previously done.

There can be little doubt that, if successful, the GRO will add greatly to our meager knowledge of the universe at gamma ray wavelengths. But, the two-year overall lifetime limits the time that may be devoted to a given source; variability information, which may be crucial to a correct understanding, will suffer. The low altitude orbit means that sensitivity will also suffer, since the materials composing the spacecraft and the instruments will be subjected to a continuous irregular bombardment by high energy particles, resulting in relatively high, time-varying backgrounds.

THE FUTURE AND THE NEEDS OF THE SCIENCE

It is vital, for progress in gamma ray astronomy, to establish small error boxes on the sky for the locations of the different sources. Good locations will make deep searches practical in other wavelength bands, such as the radio and the optical, for counterparts

of the gamma ray sources. With optical and other identifications made, progress in understanding the sources is likely. Good positional accuracy for small-flux sources may be obtained with increased photon collecting area, or by increased observing time, or by a combination of the two. The importance of long observing times for variability studies has already been noted. As the next step in the post-GRO era, instruments in the 10–100 m^2 class appear to be in order.

For missions beyond the GRO, there is the low Earth orbit, long-duration space station. The space station offers the opportunity to conduct first-time measurements of the variability of gamma ray sources. The low altitude orbit, however, also means that a sensitivity limitation will again be imposed, because of the comparatively intense bombardment by Van Allen and South Atlantic Anomaly particles.

What is ideally needed for support of gamma ray astronomy conducted from within Earth's gravitational sphere of influence is an indefinite-duration observatory that is (a) capable of orienting large instruments; (b) located far from large masses; (c) at worst, bombarded by high energy particle fluxes no greater than cosmic ray fluxes; and (d) operated such that its instrumentation may be updated as technology advances.

An observatory located at one of the Earth-Moon Lagrangian Points best fits this ideal. If other considerations rule out such a location, it appears that the surface of the Moon itself would be an acceptable alternative site.

A LUNAR OBSERVATORY

From Apollo data, the Moon's surface is already known to be a low radioactivity environment, compared with Earth's surface or atmosphere. Background radiation from the surroundings will be lower on the lunar surface than it will be in a satellite in orbit about Earth.

The Moon is a satellite orbiting at 60 Earth radii. Instruments on the Moon therefore are in orbits well beyond the regions where the geomagnetically trapped particles exist. There will be no activation by the intense particle fluxes encountered in the South Atlantic Anomaly.

The monthly passage of the Moon through the plasmas in the geomagnetic tail is unlikely to present a problem for gamma ray astronomy, because the energies of the plasma particles in the far tail are all too low to generate gamma ray photons. Should a solar flare occur while the Moon crosses the tail and while solar-flare ions with high energies travel back "upstream" along the tail, the background will be increased. But the tail crossings are only about five days long, and the rest of each month should be free of such problems.

With certain modifications, gamma ray astronomy instrumentation resembling that of a scaled-up GRO payload is envisaged for the lunar observatory. In the context of gamma ray astronomy, a lunar observatory would have much the same goal as does GRO: increased sensitivity measurements over as wide an energy range as feasible.

One observatory instrument that would not require pointing is a spark chamber. In a spark chamber, the incident gamma ray is converted into a positron-electron pair

of particles; the tracks of these two particles may be measured by the pattern of little sparks they cause as they move through high electric fields in the chamber's fill gas. The length of the tracks yields energy information, and their direction yields directional information. Post-GRO progress would require a spark chamber whose total sensitive volume would be 15 m (diameter) × 3 m thick, filled to one-atmosphere pressure with an inert gas such as argon. Its overall weight is likely to be nearly 100 tons on Earth. Given a GRO-like observing time for a given source, the sensitivity would be improved by a factor of 10 over that of EGRET. Data processing with so large a chamber would become a serious problem, because the hundred-fold increase in gamma ray detection rate would correspondingly increase the rates of non-background (*i.e.*, non-single-track-events in the chamber's volume) events. Each such non-background event requires measurement of the length and direction of the two visible tracks in the chamber. This may be done in principle by humans with photographs; it is more likely to be automatically done with digitized TV pictures and a computer, in the observatory, or done non-photographically, as in the GRO, but with a hundred-fold increase in data rate. Lunar observatory-large spark chamber data rates are therefore likely to be continuously over one Mbit/s.

Continuous observations mean that gas leaks must be compensated for. A lunar observatory would require some ability to store replacement filling gas for the chamber.

Double Compton instruments also do not require active pointing; all sources on the visible sky would be simultaneously measured by a double Compton on the lunar surface. A scaled-up GRO double Compton with a ten-meter diameter and three-meter height would weigh perhaps 30 tons on Earth. It would have a sensitivity of 10^{-5} photons cm^{-2} sec^{-1}, given one week of observing.

Actively collimated astronomical instruments do require pointing, usually in an on-source, off-source sequence. Future instruments of this type are likely to need pointing accuracy and stability of one arcminute. A hypothetical instrument would have several square meters (total collecting area) for the photon counter, and the photon counter would be constructed of segmented hyperpure germanium solid-state radiation detectors for maximum energy resolution and lowest induced background. The active collimator used to define the field of view and impose a total energy deposition requirement and the active coded aperture used to define the angular resolution would most likely be comprised of thick scintillation counters. On Earth, the instrument would weigh perhaps 10 tons.

Scintillation counters operate well at room temperature. Their photomultiplier tubes or photodiodes require temperature stabilization. Germanium gamma ray counters require an operating temperature of below 100 K. If radiative cooling is not practical, cryostats will be necessary, requiring replenishment of the cryogen. The observatory must be able to resupply the cryogen (*e.g.*, liquid nitrogen or solid carbon dioxide) as needed.

Finally, measurements of gamma ray bursts must be extended. The principal objective here is to attempt to identify the source(s) of the bursts. This means determining their positions on the celestial sphere with arcsecond accuracy, so that other wavelength identifications may be confidently made. Probably the most progress would not be made

with a simple increase in the size of a BATSE derivative, but by significantly increasing the length of the baseline available for measurements of position. A network of burst detectors, each not much larger than the BATSE instrument and that have good (*i.e.*, microsecond or better) event timing accuracy over long times, should be operated simultaneously on the Moon and throughout as much of the solar system as feasible.

Each instrument would be sensitive to photons from all directions. Therefore, none of the members of the network would require pointing. The relative timing of the detection at the various sites of a given burst would provide high accuracy data on the angular coordinates of the burst site. Detectors located on the Moon and in low orbit about Earth could form the beginning of such a network of detectors; these would give a baseline 60 Earth radii long. Such a two-station network could locate bursts with one arcsecond accuracy (Chupp, 1976). Each detector should additionally have spectroscopic capabilities, at least in the 0.1–1.0 MeV spectral range and preferably beyond.

Large area and long observing time will place severe demands on the dependability of all the instruments. Large instruments tend to be more complex; they are likely to composed of more modules. Long times without failure are difficult to achieve, technically. The availability of a lunar observatory staffed with trained maintenance personnel makes this more practical. A not insignificant contribution to the increased practicality arises from the penetrating power of gamma radiation. Providing the walls of the observatory are at most a small fraction of a gamma ray mean free path in thickness, gamma ray astronomy instruments may be located *inside* a shirtsleeve environment, which facilitates maintenance and calibration. At an energy of 1 MeV, the gamma ray mean path in aluminum is 16 gm/cm^2, and it is 40 gm/cm^2 at 100 MeV photon energy.

Long-term measurements could run the risk of instrument obsolescence, if the instruments were not upgraded. The availability of a well equipped observatory staffed with trained scientific personnel and in good communication with Earth obviates this problem. Because of the observatory, investments in large instruments may be cost effective; modifications and improvements may be made as they develop.

REFERENCES

Chupp E. L. (1976) *Gamma-ray Astronomy.* D. Reidel, Boston. 195 pp.

Mayer-Hasselwander H. A., Bennett K., Bignami G. F., Buccheri R., Caraveo P. A., Hermsen W., Kanbach G., Lebrun F., Lichti G. G., Masnou J. L., Paul J. A., Pinkau K., Sacco B., Scarsi L., Swanenburg B. N., and Wills R. D. (1982) Large-scale distribution of galactic gamma radiation observed by COS-B. *Astron. Astrophys., 105,* 164.

Paciesas W., Baker R., Bodet D., Brown S., Cline T., Costlow H., Durouchoux P., Ehrmann C., Gehrels N., Hameury J., Haymes R., Teegarden B., and Tueller J. (1983) A balloon-borne instrument for high-resolution astrophysical spectroscopy in the 20–8000 keV energy range. *Nucl. Instrum. Methods, 215,* 261–276.

Ramaty R. and Lingenfelter R. E. (1982) Gamma-ray astronomy. *Annu. Rev. Nucl. Part. Sci., 32,* 235–269.

Samorski M. and Stamm W. (1983) Detection of 2×10^{15}–2×10^{16} eV gamma-rays from Cygnus X-3. *Astrophys. J. Lett., 268,* L17.

IRRADIATION OF THE MOON BY GALACTIC COSMIC RAYS AND OTHER PARTICLES

James H. Adams, Jr.

E. O. Hulburt Center, Naval Research Laboratory, Washington, DC 20375

Maurice M. Shapiro

Max Planck Institüt für Astrophysik, 8046 Garching bei München, Federal Republic of Germany

Men and sensitive instruments on a lunar base can be profoundly affected by the radiation environment of the Moon. The ionizing radiation incident upon the lunar surface is comprised of the galactic cosmic rays (GCR) and energetic particles accelerated in the solar neighborhood. The latter consist mainly of solar energetic particles (SEP) from flares and of other particles energized in the heliosphere. The cosmic radiation bombarding the Moon consists overwhelmingly of relativistic and near-relativistic atomic nuclei ranging in energy from 10^8–10^{20} eV, approximately 98.6% of which consists of hydrogen and helium. The remainder spans the rest of the periodic table, with conspicuous peaks in abundance at C, O, Ne, Mg, Si, and Fe. The GCR composition is roughly similar to that of the sun, with some notable differences. Differential energy spectra and composition of cosmic rays as well as the intensities, composition, and the spectra of SEP and particles accelerated in the heliosphere are reviewed. We also summarize the analytic models developed by the Naval Research Laboratory (NRL) group to describe the energy spectra and elemental compositions of the various components.

THE LUNAR RADIATION ENVIRONMENT

The Moon is constantly bombarded by galactic cosmic rays (GCR). Figure 1 (from Simpson, 1983) shows a sampling of the data available on the differential energy spectra of the most prominent particle types in galactic cosmic rays. The intensity of this highly penetrating particle radiation varies in response to solar activity. In a way that is not yet fully understood (Fillius and Axford, 1985), the out-flowing solar wind modulates the cosmic ray intensity so that it is anti-correlated with the general level of solar activity. This causes the average intensity of cosmic rays with energies greater than 10 MeV/amu to increase 2.5 times from the maximum to the minimum of the 11-year solar activity cycle. Low energy cosmic rays are affected more strongly than the higher energy ones. Figure 2 (from Simpson, 1983) compares the elemental compositions of galactic cosmic rays and the solar system, normalized at silicon. The two compositions are comparable for the most abundant elements. The odd elements, in general, and Li, Be, B, F, Sc, Ti, V, Cr, and Mn, in particular, are overabundant in galactic cosmic rays. This difference is the result of the propagation of galactic cosmic rays through approximately 7 g/cm^2 of interstellar gas, on average, before reaching Earth (Shapiro and Silberberg, 1970).

During periods of minimum solar activity, additional components can be observed at low energy. One component is constantly present. This component, discovered by Garcia-

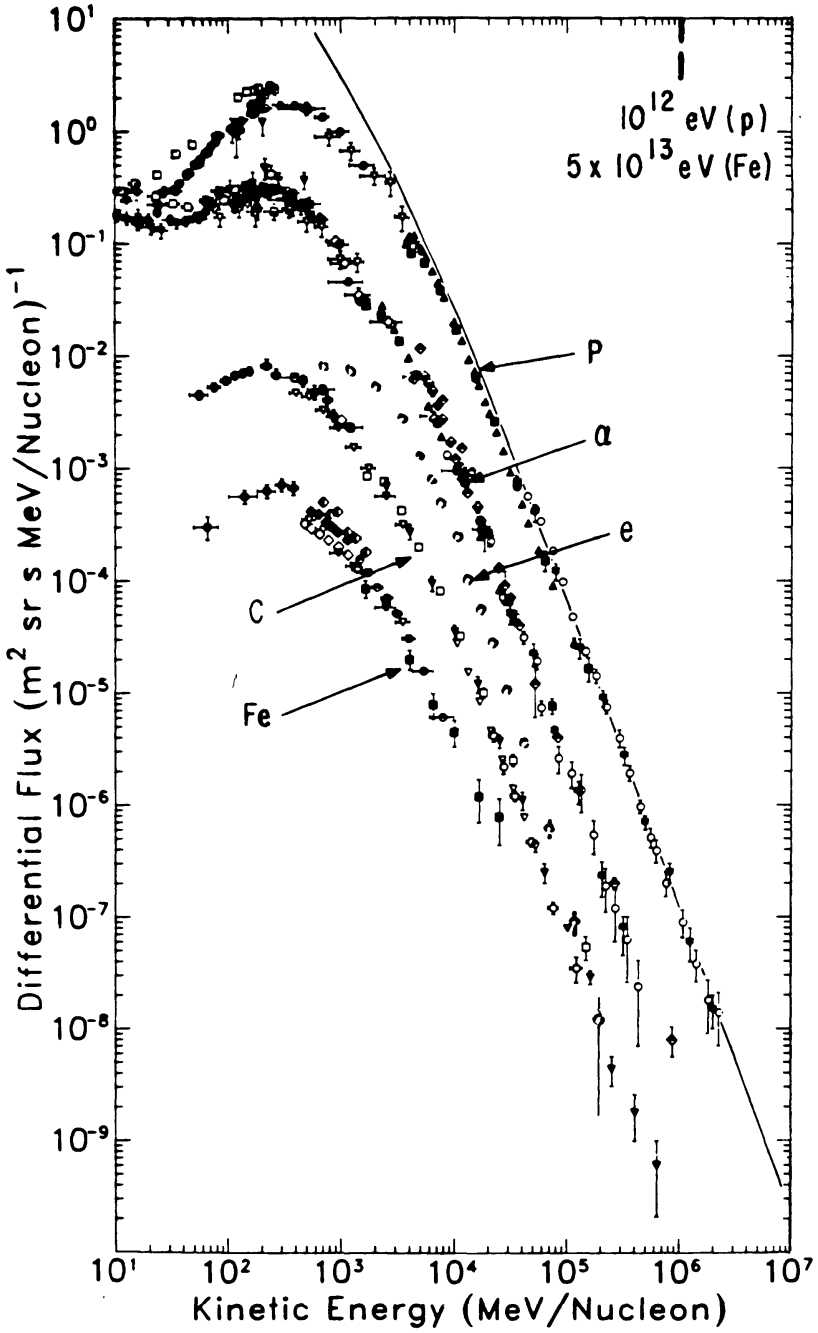

Figure 1. The differential energy spectra for the elements (from the top) hydrogen (P), helium (α), carbon (C), and iron (Fe). Also shown is the electron spectrum (labeled e). The solid curve shows the hydrogen spectrum extrapolated to interstellar space by unfolding the effects of modulation. The turn-up of the helium spectrum below about 60 MeV/nucleon is due to the contribution of the anomalous component of helium. This figure was taken from Simpson (1983).

Munoz *et al.* (1973), is called the anomalous component because of its unusual nature. Figure 3 shows the spectra of H, He, C, and O in the interplanetary medium. The anomalous component is the broad peak in the low energy oxygen spectrum and the absence of a dip in the helium spectrum at 10 MeV/amu that causes the helium flux to exceed the hydrogen flux in this energy range. A second component is sometimes accelerated in regions where fast and slow moving solar wind streams collide (Gloeckler, 1979). These co-rotating energetic particles will sometimes cause modest increases in the 1–10 MeV/

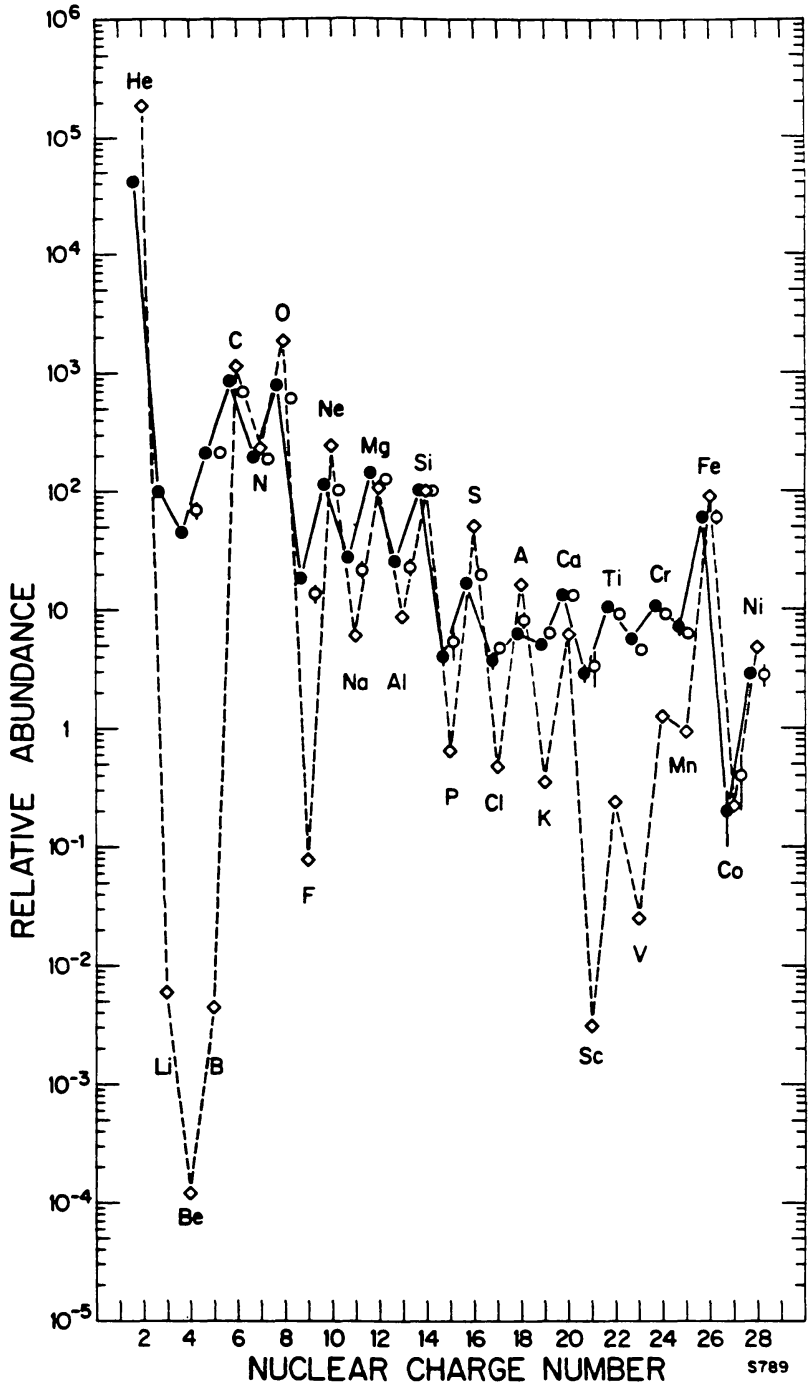

Figure 2. The cosmic ray element abundances (He-Ni) measured at Earth compared to the solar system abundances. The two abundances are normalized at silicon. The diamonds represent the solar system abundances, while the open circles are cosmic ray measurements at high energies in the 1000–2000 MeV/nucleon range. Hydrogen, not shown, is about 20 times more abundant in the solar system than in cosmic rays, using silicon as the normalization. This figure was taken from Simpson (1983).

amu hydrogen and helium fluxes bombarding the Moon. Because these components exceed the cosmic ray background only at low energies, their contribution to the total particle intensity bombarding the Moon is small.

Occasionally there are major increases in the radiation intensity at the Moon due to solar energetic particle (SEP) events. These events last from hours to days and range in size from the limit of detection to an intensity more than 70,000 times that of galactic

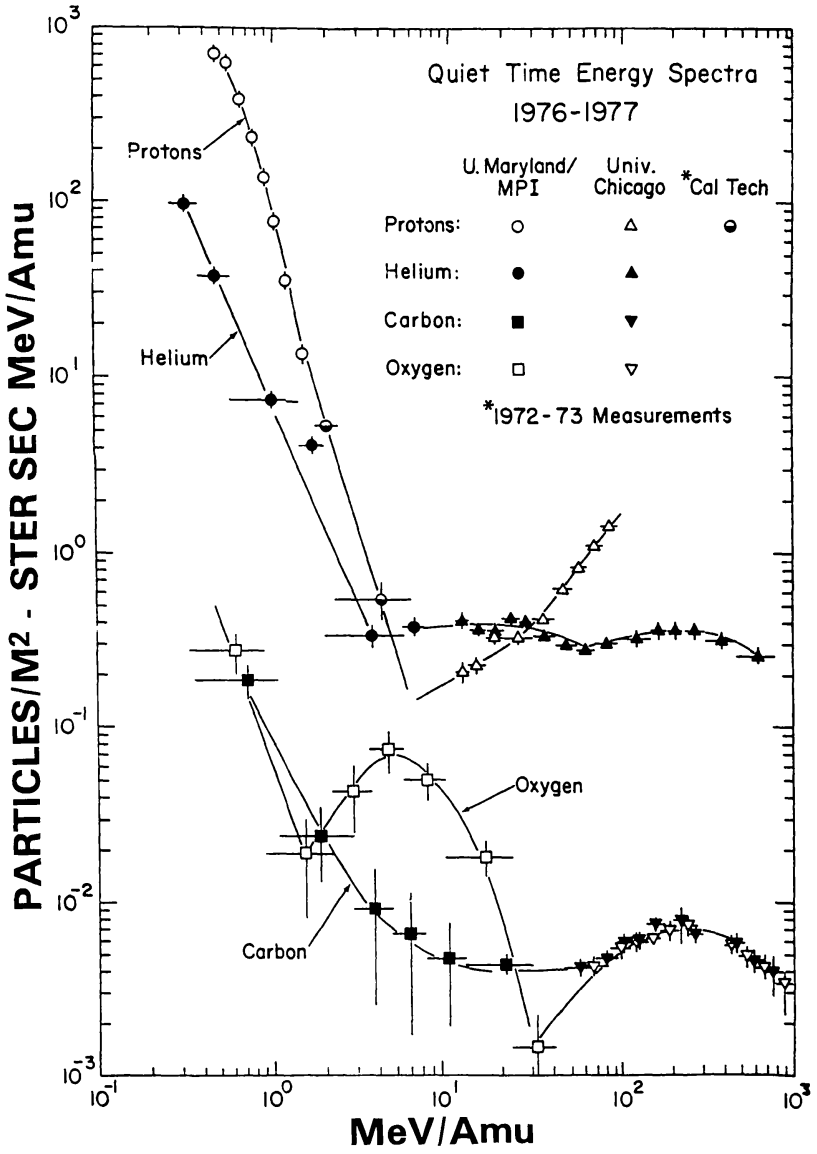

Figure 3. Differential energy spectra of hydrogen, helium, carbon, and oxygen observed in the interplanetary medium near the Earth during the solar minimum in 1976–77 during quiet times. The anomalous cosmic ray component appears between 2 and 30 MeV/ nucleon and is characterized by the large overabundance of helium and oxygen compared to hydrogen (protons) and carbon, respectively. This figure was taken from Gloeckler (1979).

cosmic rays. So large are the largest of these events that they determine the particle fluence at the lunar surface over a solar cycle. It is usually true that half the particles to strike the Moon in an 11-year solar cycle arrive in less than a day and are the result of one, or at most a few, large SEP events. This striking feature will make a flare watch an important part of any future lunar expedition, as it was during the Apollo program.

McGuire *et al.* (1983) show the record of SEP events for the last three solar cycles. Their results are reproduced in Fig. 4. As this figure shows, the frequency of SEP events varies with the overall level of solar activity as gauged by the smoothed Zurich sunspot number. McGuire *et al.* find the solar-cycle-averaged hydrogen fluxes above 10 MeV for cycles 19, 20, and 21 are 378, 93, and 65 particles/cm^2 s. These fluxes are 132, 33, and 23 times larger than the solar-cycle-averaged galactic cosmic ray hydrogen flux, respectively.

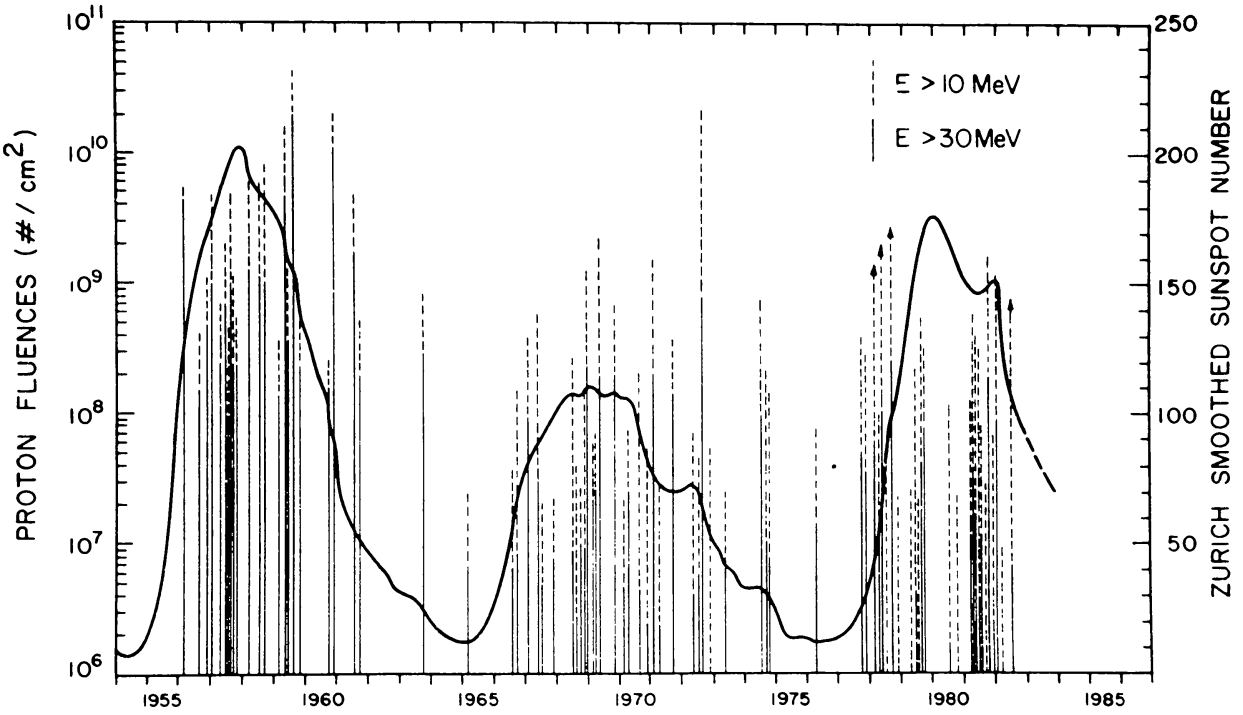

Figure 4. Hydrogen fluences above 10 and 30 MeV in Solar Energetic Particle Events during solar cycles 19, 20, and 21. The solid curve represents Zurich smoothed sunspot numbers. This figure was taken from McGuire et al. (1983).

The spectra of SEP events are much softer than the galactic cosmic ray spectrum. Even during the peak intensity of most SEP events, galactic cosmic rays are the principal source of particles above a few hundred MeV/amu. The largest observed flares have, at their peak, dominated the flux up to 10,000 MeV/amu.

The elemental composition of SEP events is very similar to the solar system composition (shown in Fig. 2) on average but can be highly variable from one event to the next. The composition even varies with particle energy in individual events (Chenette and Dietrich, 1984). SEP events that are enriched in one heavy ion tend to be enriched in the others as well. Dietrich and Simpson (1978) have shown that this systematic enrichment increases strongly with atomic number.

THE NRL CREME MODEL

A procedure has been developed at Naval Research Laboratory to characterize cosmic ray effects on microelectronics (CREME) used in spacecraft and aircraft (see Adams *et al.*, 1981; Adams *et al.*, 1983; and Tsao *et al.*, 1984). This procedure relies on a detailed numerical model of the near-Earth particle environment (Adams *et al.*, 1981), which is directly applicable to characterizing the radiation environment on the Moon. A set of formulas describes the differential energy spectra of each of the elements in galactic cosmic rays and how these spectra are modified by the contributions from the anomalous

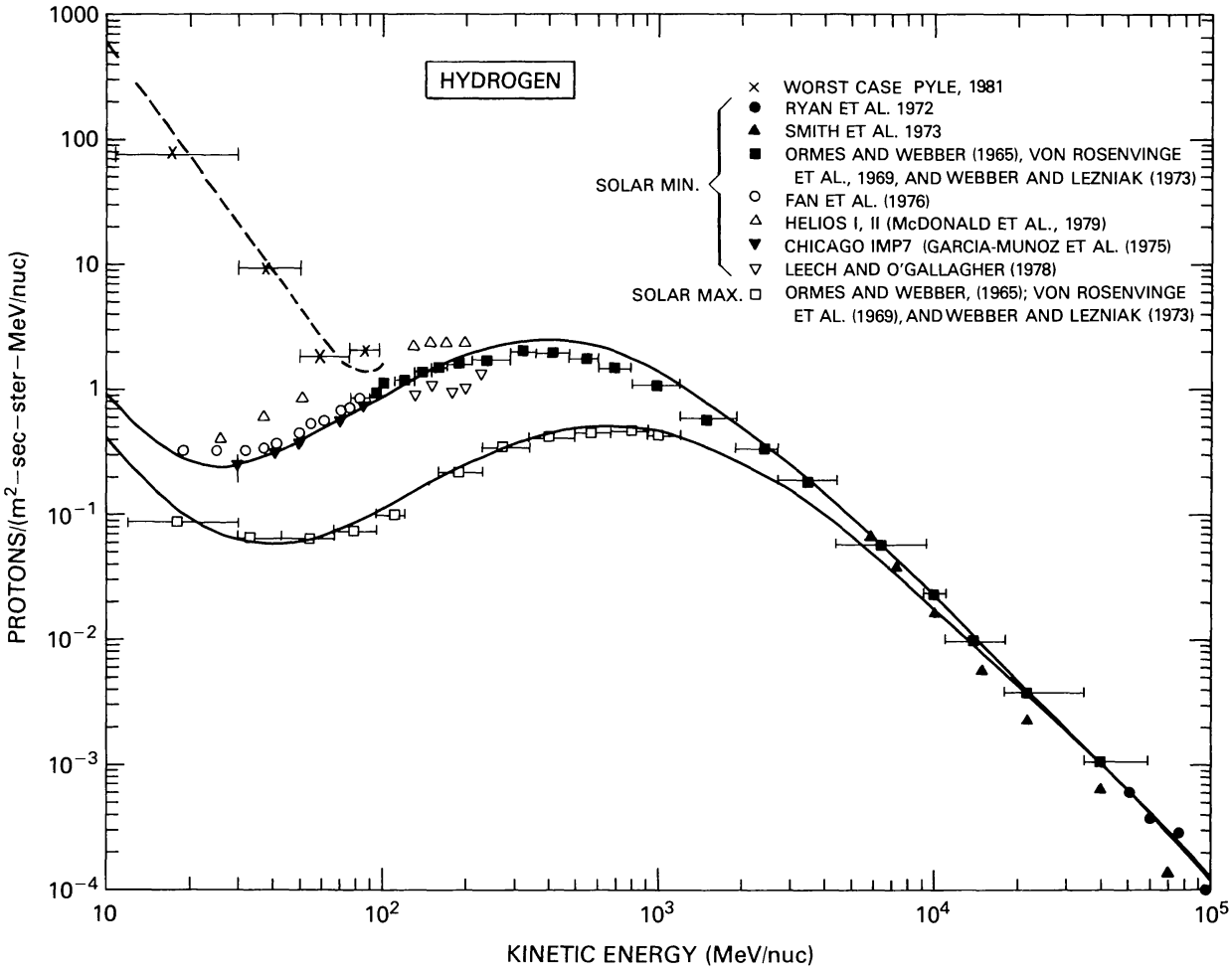

Figure 5. Hydrogen differential energy spectra (taken from Adams et al., 1981). The data are selected for the extremes of solar maximum and solar minimum. The solid curves are from the formulas to fit the cosmic ray spectra for solar minimum (upper curve) and solar maximum (lower curve). The dashed curve is from a formula constructed to give instantaneous flux levels so high at each energy that they are exceeded only 10% of the time.

component, co-rotating energetic particle streams, and from small flares. The model also contains formulas for the differential energy spectra of each of the elements in SEP events and formulas for calculating the probability of occurrence of such events.

Galactic cosmic rays were modeled by using all the available data to determine the shapes of the differential energy spectra of hydrogen, helium, and iron at the extremes of solar maximum and solar minimum. Figures 5, 6, and 7 show how the model (solid lines) fit the data for hydrogen, helium, and iron, respectively. The model spectra are for the extremes of solar maximum and minimum. Intermediate cases are interpolated with a sinusoidal solar modulation factor of sin[2 $(t-t_0)/10.9$ years)], where $t_0 = 1950.6$. The other elemental spectra are obtained by multiplying either the helium or the iron spectrum by a constant or, in some cases, energy-dependent scale factor. By comparing

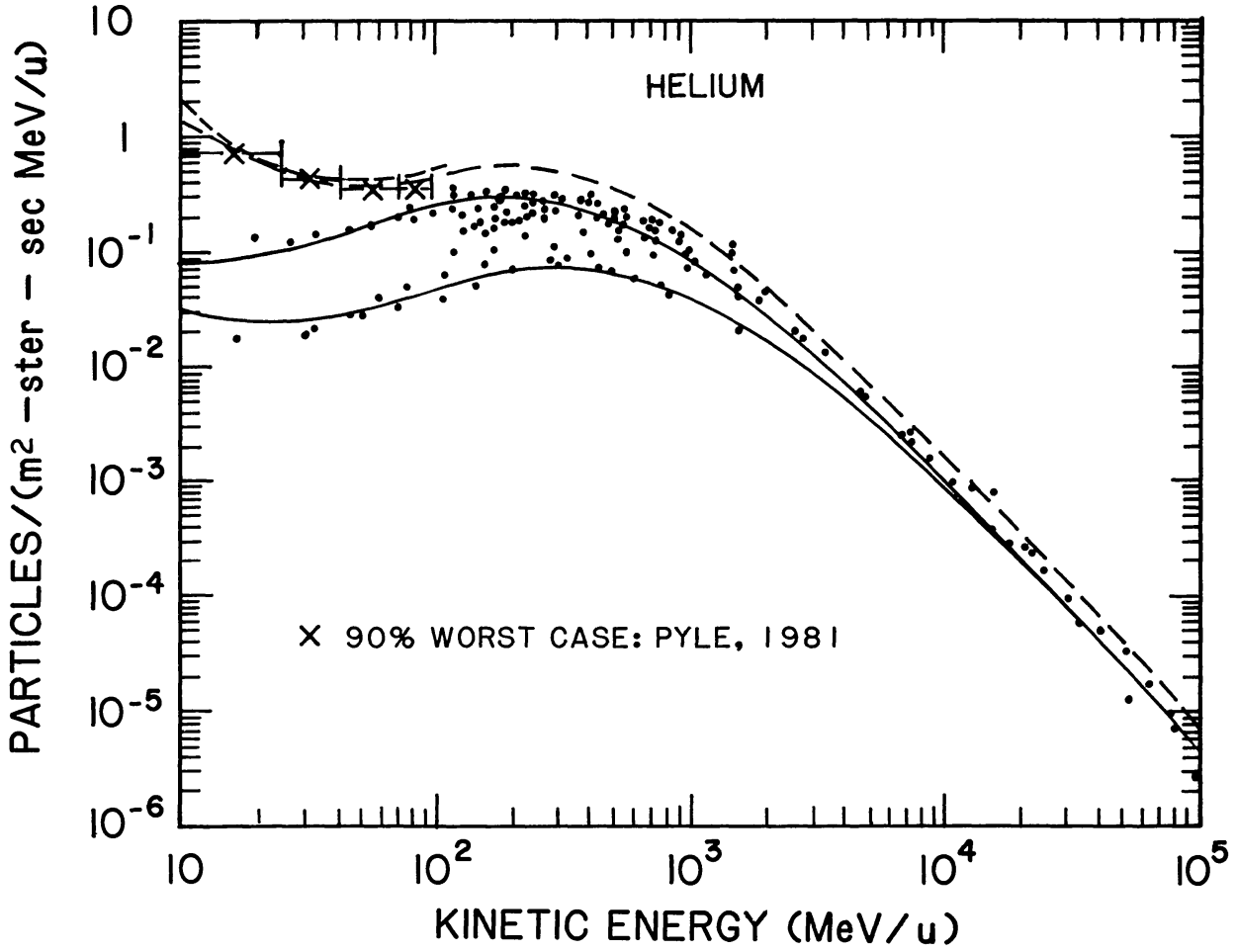

Figure 6. Helium differential energy spectra (taken from Adams et al., 1981). The solid curves are from the formulas to fit the cosmic ray spectra for solar minimum (upper curve) and solar maximum (lower curve). The dashed curve is from a formula constructed to give instantaneous flux levels so high at each energy that they are exceeded only 10% of the time.

this model with recent data, we find that it seems to predict the absolute cosmic ray flux to within a factor of two. The relative abundances are accurate to about 20%.

The contributions at low energies from co-rotating particle streams and small SEP events were accounted for along with the overall uncertainty by the 90% worst-case model. This model is shown as the dashed curves in Figs. 5–7. The contributions of the anomalous component to the helium, nitrogen, and oxygen spectra are modeled by Adams *et al.* (1981), who show how these may be combined with the cosmic ray model spectra to account for the contributions of the anomalous component at low energies.

Following the scheme of King (1974), Adams *et al.* (1981) divided all the large SEP events into ordinary large flares and anomalously large flares. A formula was fitted to the means of the log-normal distributions of the integral SEP flux above several energy thresholds and then differentiated to obtain the mean hydrogen differential energy spectrum for ordinary large SEP events. This procedure was repeated using values 1.28 standard

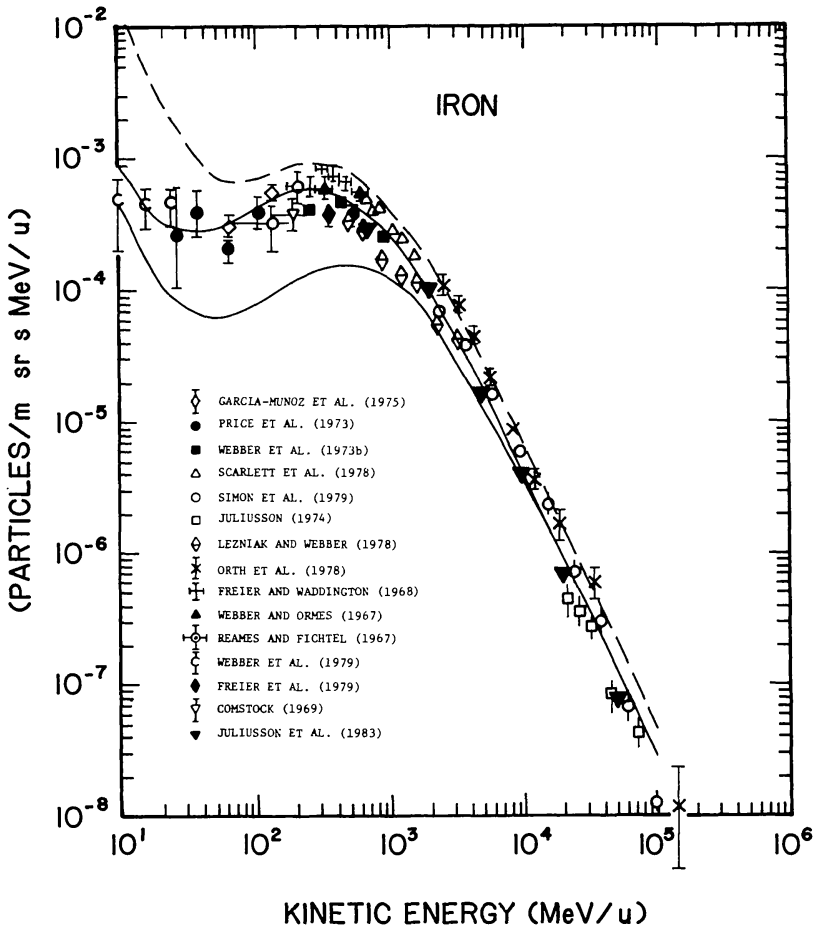

Figure 7. Iron differential energy spectra (taken from Adams et al., 1981). The solid curves are from the formulas to fit the cosmic ray spectra for solar minimum (upper curve) and solar maximum (lower curve). The dashed curve is from a formula constructed to give instantaneous flux levels so high at each energy that they are exceeded only 10% of the time. This dashed curve has been constructed by comparison with helium, since iron data to establish this curve directly are lacking.

deviations above the means of the log-normal distributions to obtain a spectrum for the 90% worst-case SEP event. Again, following King (1974), the SEP event of August 4, 1972 was used as the model for anomalously large events. These three model hydrogen spectra are shown in Fig. 8.

The composition of the SEP events is given by Adams *et al.* (1981) as elemental abundances relative to hydrogen for both the mean heavy ion composition and a 90% worst-case enrichment in the heavy elements. These two compositions, shown in Fig. 9, indicate the degree of variability in the SEP composition. Burrell distribution formulas are also provided by Adams *et al.* (1981) to calculate the probability of an SEP event during any time period.

RADIATION EFFECTS ON THE MOON

If a large permanent base is established on the Moon, the 5 rem/y exposure limit for Earth-based radiation workers might be a more appropriate standard for radiation protection. As Silberberg *et al.* (1985) show, adherence to this standard would make it necessary to bury a lunar habitat beneath several meters of lunar regolith and limit human activity on the lunar surface to "regular working hours," *i.e.*, about 1800 hours

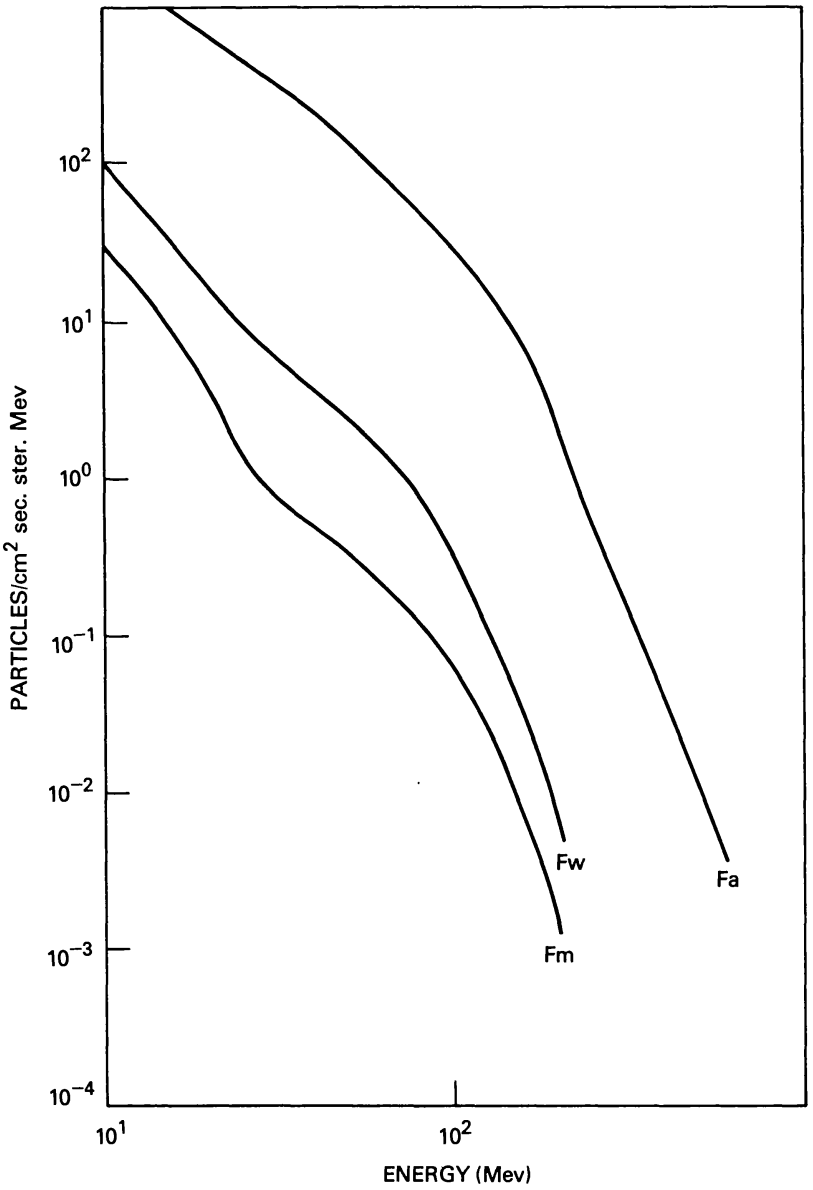

Figure 8. Hydrogen differential energy spectra (taken from Adams et al., 1981). These spectra are for the peak intensities of three model solar energetic particle (SEP) events. The curve labeled Fm is for the mean large SEP event (using the definition of King, 1974). A second model SEP event (curve Fw) has been constructed such that only one SEP event in 10 will have a peak intensity, at any energy, that is greater than predicted by this event. These two curves may be compared to get a feel for the range of flare sizes. The Fa curve is modeled after the peak of the SEP event of August 4, 1972. This is one of the most severe SEP events ever observed.

per year inside an enclosed vehicle. Extravehicular activity would have to be restricted. Even then, people working on the surface would have to remain near the habitat so that they could rush to shelter in case of a large SEP event. Long expeditions across the lunar surface would be risky unless shelters could be constructed a few hours travel time apart, or a means could be provided to quickly rescue the expedition and return its members to the safety of the buried lunar habitat.

The lunar radiation environment affects not only people, but electronic systems as well. It has long been known that electronic components are affected by the total radiation dose they have accumulated. This radiation damage produces changes in conductivity or shifts in device thresholds that cause a malfunction of the electronic circuit. Electronic components have been developed that can tolerate very large total doses, so that it is

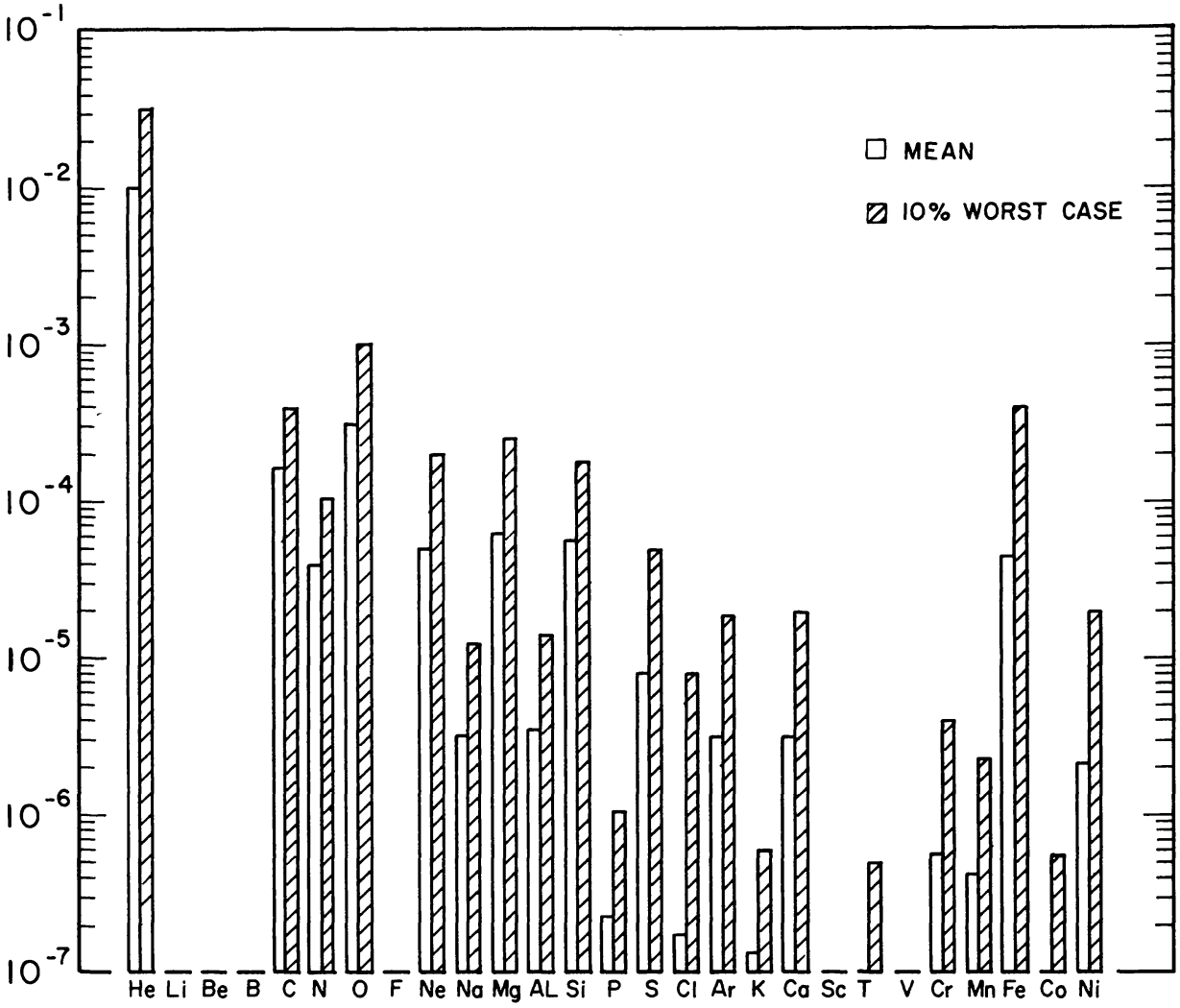

Figure 9. Elemental composition of Solar Energetic Particle (SEP) events (taken from Adams and Gelman, 1984). The mean SEP composition, normalized to hydrogen, is compared with a composition (10% worst case) constructed so that only one SEP event in 10 will be richer in any heavy ion. Comparing the two compositions, we can see that the iron to hydrogen ratio in a heavy ion-rich event may exceed the oxygen to hydrogen ratio for a typical event.

possible to design electronic systems for use on the Moon with operational lifetimes in excess of 10 years.

Recently, it has been discovered that single, intensely ionizing particles can produce a burst of hole-electron pairs so large that the resulting charge or current can change the logic state of a modern digital microcircuit (Adams *et al.*, 1981, 1983; Tsao *et al.*, 1984). This change of state damages not the electronic circuitry but the information stored in it. These events are therefore called "soft upsets."

The operational impact of a soft upset depends on the microcircuit affected. If the microcircuit is in the program memory of a computer, the program will no longer be

operable and must be reloaded. If it is in the microprocessor's address registers or program counter, the actions the computer takes will be unpredictable. Soft upsets in control circuitry can also result in unplanned events such as thruster firings. It is clear that a single soft upset could cause the loss of equipment and personnel. Unlike total dose sensitivity, soft upset susceptibility is a fundamental feature of modern large-scale integrated circuits. It appears unlikely that such compact circuits can be made immune to soft upsets. The problem has been attacked at the system level instead, with redundancy, fault tolerance, and software checking. These methods reduce, but do not eliminate, the risks posed by soft upsets.

COSMIC RAY EXPERIMENTS FOR A LUNAR BASE

The Moon offers the possibility of doing cosmic ray experiments that would be difficult to carry out in Earth orbit. On the Moon, lunar regolith can be used for the massive absorbers needed in some large detector systems. The Moon also offers a site for the construction of large detector arrays beyond the protection of the Earth's magnetic field. Two possible experiments are discussed here.

The energy spectrum of cosmic rays has been measured directly up to 1 TeV/nuc (Watson, 1975). The only direct measurement above this energy is due to Grigorov *et al.* (1971). Therefore, most of what is known about cosmic rays above 1 TeV/nuc is based on indirect measurements that provide only total particle energy as determined from the shower of secondary particles produced in the atmosphere by an incident cosmic ray. Direct measurements at these higher energies will make it possible to establish the particle's charge and, hence, its velocity and magnetic rigidity. The latter quantity can be compared with the available data on galactic magnetic fields to determine whether the particles at these high energies could have come from our galaxy or must be extra-galactic.

The best device for measuring these high energies directly is an ionization calorimeter of the type developed at Goddard Space Flight Center (Balasubrahmanyan and Ormes, 1973). One square meter calorimeters could be constructed on the lunar surface, using regolith to replace the heavy iron plates in the Goddard design. A single calorimeter of this size would detect events up to 10,000 TeV/nuc in the first year of operation, and 100 such units could extend the spectrum to the interesting region above 100,000 TeV/nuc in a few years. A secondary benefit of such an experiment might be the opportunity to study elementary particle interactions at energies well above those achieved at accelerators. Such investigations could lead to new discoveries pointing the way for new particle physics experiments on Earth.

Experiments employing NASA's Long Duration Exposure Facility, presently underway and planned for the near future, are expected to establish the flux of actinide nuclei in galactic cosmic rays. This will tell us whether cosmic ray source material resembles the interstellar medium or is enriched in nuclei synthesized by the rapid neutron capture (r-) process. Whatever the nucleosynthetic origin of cosmic rays may turn out to be, these near-term experiments will not tell us how much time has elapsed since cosmic

ray material was synthesized. To answer this question, it will be necessary to measure the abundances of the individual actinide nuclei. The relative abundances of Th, U, Np, Pu, and Cm tell us the elapsed time since the nucleosynthesis of cosmic rays in the range of 10^7 to 10^9 years. To measure them would require 2000 m^2 ster. years of collecting power. As suggested by Waddington (personal communication, 1984), this could be provided by a cylindrical array of scintillators 10 m in diameter and 10 m high. Such an apparatus could be placed on the lunar surface, and it would collect a suitable sample of events in less than 5 years. The array would use time of flight across the cylinder to measure velocity, so that the scintillator signals could be corrected for velocity to obtain the particle's charge.

CONCLUSIONS

The ionizing radiation environment on the lunar surface poses a hazard to men and sensitive instruments. Measures to protect crews from this environment can be expected to influence the design of lunar bases and the planning of lunar surface activities.

The lunar surface offers a site for large cosmic ray experiments to measure the abundances of rare elements and extremely high energy particles. The experiments that are possible on the Moon will provide new information on the origin of cosmic rays and possibly on the interaction of ultra-high energy particles with matter.

REFERENCES

Adams J. H. and Gelman A. (1984) The effects of solar flares on single event upset rates. *IEEE Trans. Nucl. Sci., NS-31*, 1212–1216.

Adams J. H., Silberberg R., and Tsao C. H. (1981) Cosmic ray effects on microelectronics, Part I: *The Near-earth Particle Environment.* NRL Memorandum Report 4506, Naval Research Laboratory, Washington, DC. 92 pp.

Adams J. H., Letaw J. R., and Smart D. F. (1983) *Cosmic Ray Effects on Microelectronics, Part II: The Geomagnetic Cutoff Effects.* NRL Memorandum Report 5099, Naval Research Laboratory, Washington, DC. 45 pp.

Balasubrahmanyan V. K. and Ormes J. F. (1973) Results on the energy dependence of cosmic-ray charge composition. *Ap. J., 186*, 109–122.

Chenette D. L. and Dietrich W. F. (1984) The solar flare heavy ion environment for single event upsets: A summary of observations over the last solar cycle, 1973–1983. *IEEE Trans. Nucl. Sci., NS-31*, 1217–1222.

Dietrich W. F. and Simpson J. A. (1978) Preferential enhancements of the solar flare-accelerated nuclei carbon to zinc from ~20–300 MeV nucleon. *Ap. J., 225*, L41–L45.

Fillius W. and Axford I. (1985) Large scale solar modulation of $\lesssim$500 MeV/Nucleon galactic cosmic rays seen from 1 to 30 AU. *J. Geophys. Res., 90*, 517–520.

Garcia-Munoz M., Mason G. M., and Simpson J. A. (1973) A new test for solar modulation theory: The 1972 May-July low-energy galactic cosmic-ray proton and helium spectra. *Ap. J., 182*, L81–L84.

Gloeckler G. (1979) Compositions of energetic particles populations in interplanetary space. *Rev. Geophys. Space Phys., 17*, 569–582.

Grigorov N. L., Gubin Yu V., Rapaport I. D., Savenko I. A., Akimov V. V., Nesterov V. E., and Yakovlev B. M. (1971) Energy spectrum of primary cosmic rays in the 10^{11}–10^{15} eV energy range according to the data of proton-IV measurements. *Proc. 12th Intl. Cosmic Ray Conf., 5*, pp. 1746–1751.

King J. H. (1974) Solar proton fluences for 1977–1983 space missions. *J. Spacecraft Rockets, 11*, 401–408.

McGuire R. E., Goswami J. N., Jha R., Lal D., and Reedy R. C. (1983) Solar flare fluences during solar cycles 19, 20, and 21. *Proc. 18th Intl. Cosmic Ray Conf., 44*, pp. 66–69.

Shapiro M. M. and Silberberg R. (1970) Heavy cosmic ray nuclei. *Ann. Rev. Nucl. Sci., 20*, 323.

Silberberg R., Tsao C. H., Adams J. H. Jr., and Letaw J. R. (1984) Radiation doses and LET distributions of cosmic rays. *Rad. Res., 98*, 209–226.

Silberberg R., Tsao C. H., Adams J. H. Jr., Hulburt E. O., and Letaw J. R. (1985) Radiation transport of cosmic ray nuclei in lunar material and radiation doses (abstract). This volume.

Simpson J. A. (1983) Elemental and isotopic compositions of the galactic cosmic rays. *Ann. Rev. Nucl. Part. Sci., 33*, 323–381.

Tsao C. H., Silberberg R., Adams J. H. Jr., and Letaw J. R. (1984) Cosmic Ray Effects on Microelectronics, Part III: Propagation of Cosmic Rays in the Atmosphere. NRL Memorandum Report 5402, Naval Research Laboratory, Washington, D.C. 87 pp.

Watson A. A. (1975) Energy spectrum and mass composition of cosmic ray nuclei from 10^{12} to 10^{20} eV. In *Origin of Cosmic Rays* (J. L. Osborne and A. W. Wolfendale, eds.), pp. 61–96. Reidel, Dordrecht.

CELESTIAL SOURCES OF HIGH-ENERGY NEUTRINOS AS VIEWED FROM A LUNAR OBSERVATORY

Maurice M. Shapiro

Max-Planck Institut für Astrophysik, 8046 Garching bei München, West Germany

Rein Silberberg

Hulburt Center, Naval Research Laboratory, Washington, DC 20375

The detection of high-energy (HE) cosmic and solar flare neutrinos near the lunar surface would be feasible at energies much lower than for a terrestrial observatory. At these lower energies ($\geq 10^9$ eV) the neutrino background is drastically reduced below that generated by cosmic rays in the Earth's atmosphere. Because of the short mean free path (< 1m) of the progenitor pi and K mesons against nuclear interaction in lunar rocks, the neutrino background would be quite low. At 1 GeV, less than 1% of the pions would decay; at 10 GeV, 0.1% would decay. Thus, if the neutrino flux to be observed is intense enough and its spectrum is steep enough, then the signal-to-noise ratio is very favorable. The reduction in cross section at lower energies would not cancel the advantage of enhanced flux. The observation of HE neutrinos from solar flares would be dramatically enhanced, especially at lower energies, since the flare spectra are very steep. Detection of these neutrinos on Earth does not appear to be feasible. Moreover, higher-energy neutrinos ($> 10^{12}$ eV) that could, in principle, be detected are virtually absent from solar flares. A remarkable feature of solar flares as viewed in HE neutrinos from a lunar base is that the entire surface of the sun would be "visible." Indeed, flares on the *far* side of the sun would be producing more neutrinos moving toward the detector than those on the near side. Diffuse sources of HE neutrinos, such as the galactic disc (especially from the galactic center), would be detectable at energies between, say, 10^9 and 10^{11} eV. On Earth, they are swamped by the overwhelming atmospheric background.

INTRODUCTION

The advantages of a lunar observatory for neutrino astronomy were discussed some years ago by F. Reines (1965). In the present paper, we suggest that the investigation of neutrinos from astrophysical sites at energies between 1 and 10^3 GeV can be better carried out on the Moon than on the Earth. In the dense lunar materials, competition between nuclear interactions of pions and their decay suppresses the frequency of decay. In the tenuous upper atmosphere of the Earth, on the other hand, the decay of pions (and of their muon progeny) does generate neutrinos. Hence, the flux of neutrinos near the surface of the Moon is about 10^{-3} of that on the Earth at energies between 1 and 10^2 GeV, and about 10^{-2} at 10^3 GeV. Only the background due to prompt neutrinos from the decay of charmed particles in the atmosphere is not suppressed.

At energies below 1 GeV, however, the path length of pions against decay diminishes as the Lorentz factor approaches unity, and pion decay is no longer suppressed, even on the Moon. Furthermore, due to the absence of magnetic shielding on the Moon, the

flux of low-energy cosmic rays incident on the lunar surface is much higher than the average flux at the top of the Earth's atmosphere. This further enhances the low-energy neutrino intensity ($E < 1$ GeV) on the Moon. [The suppression of neutrino background was quantitatively explored by Cherry and Lande (1984) in a paper presented at this conference.]

Accordingly, a lunar base is probably an unsuitable site for observing the low-energy neutrinos (~10 MeV) from stellar gravitational collapse. Moreover, it is not competitive for recording neutrinos at very high energies ($E > 10^3$ GeV); this can be done more readily with Cerenkov light detectors in a large volume of sea water (some 10^8 m^3) near the bottom of the ocean. Such an array—DUMAND (a Deep Underwater Muon and Neutrino Detector)—will be emplaced in the waters near Hawaii in the near future (Peterson, 1983).

CRITERIA FOR CANDIDATE NEUTRINO SOURCES TO BE EXPLORED ON THE MOON

What types of neutrino sources are likely to be observable between 1 and 10^3 GeV? This is the energy interval for optimum detection by a neutrino observatory under the lunar surface (about 100 m below). The sources should emit neutrinos much more copiously above 1 GeV than above 1 TeV, so as to permit the construction of a neutrino observatory significantly smaller than DUMAND. An important constraint is imposed by the interaction cross section of neutrinos, which increases linearly with energy between 1 and 10^3 GeV. As a result, the observation of lower-energy neutrinos becomes more difficult. This cross section is given by

$$\sigma_{\nu N} = (0.7 \text{ or } 0.8) \times 10^{-38} E_\nu \text{ cm}^2 \tag{1}$$

and

$$\sigma_{\bar{\nu}N} = 0.3 \times 10^{-38} E_{\bar{\nu}} \text{ cm}^2 \tag{2}$$

for neutrinos and anti-neutrinos, respectively. Let the energy spectrum of the neutrinos be

$$\frac{dJ}{dE_\nu} = KE_\nu^{-\alpha} \tag{3}$$

Then the event rate is proportional to

$$\int_{E_0}^{E_{max}} \sigma(E_0) K E_\nu^{-\alpha} \, dE_\nu \tag{4}$$

i.e., it is proportional to

$$E_0^{-(\alpha-2)} - E_{max}^{-(\alpha-2)} \tag{5}$$

Thus, one criterion for significant source strength in the energy interval between 1 and 10^3 GeV is a steep neutrino spectrum, with the exponent α appreciably greater than 2.

SOME PROMISING CANDIDATE SOURCES

Solar flares generate particles having steep energy spectra, with $\alpha = 4$–7 at proton energies above 1 GeV. Erofeeva *et al.* (1983) explored the use of a deep underwater detector of 10^6 tons for observing neutrinos from solar flares. They did not investigate the neutrino background in their paper. We estimate that the background rate is about 10^3 per day. If the neutrinos are emitted in about 20 minutes, as are the gamma rays from a flare, then the background rate is down to 10 for the duration of the flare. If, moreover, an angular resolution of 1 steradian is obtained, then the background is down to ~1 event for the duration of the flare.

For observation of neutrinos from very large flares, such as occur about once per solar cycle, a *terrestrial* underwater observatory of 10^6 tons seems adequate. However, for larger observatories, $>10^6$ tons, the neutrino background on Earth becomes prohibitive. Thus, for observing fine-time structure or neutrino energy spectra of very large flares, or for recording somewhat smaller flares, a lunar observatory of $>10^6$ tons provides an opportunity to carry out studies of flares that are not possible on the Earth. Even flares on the remote side of the sun become observable, since neutrinos with energies $<10^{11}$ eV can traverse the solar diameter. In fact, for a given size of flare, neutrinos should reach the detector in greater numbers from the far side than from flares on the near side. This is due to the favorable rate of production of pions (hence, of daughter neutrinos) that move toward the observer, when the progenitor protons or other energetic nuclei—on the far side—are directed toward the solar surface.

Another, more diffuse source of neutrinos with a fairly steep energy spectrum $\alpha = 2.7$ is that from the central annulus of the galactic disk, ± 60° in longitude and ± 5° in latitude about the galactic center. Stecker *et al.* (1979) explored the detectability of these neutrinos at 10^3 GeV with a DUMAND array of 10^9 tons (having an effective detection volume of some 10^{10} tons). The estimated rate of neutrino events to be expected was 130 per year, swamped by 1.8×10^4 background events per year. At $E > 1$ GeV, the event rate is about 100 times higher, so that even in a smaller detector of ~10^7 tons, the event rate is about 10 per year, with the signal exceeding the background in a lunar observatory.

In addition, there are many interesting discrete candidate sources of neutrinos: accreting neutron stars (including pulsars) in binary systems, active galactic nuclei with accretion disks from which matter drifts into ultra-massive black holes (Silberberg and Shapiro, 1979), and the expanding shells around young pulsars (Berezinsky, 1976; Shapiro and Silberberg, 1979). However, the energy spectra of neutrinos from these sources are as yet unknown.

Presented here are the results of a sample calculation for SS433, which appears to be one of the most promising candidate sources in our galaxy, at a distance of about

3 Kpc. This object is probably an accreting black hole in a binary system; it has two relativistic jets and other remarkable features. Its estimated power output is 3×10^{39} ergs/s (Grindlay *et al.*, 1984), but values that are higher by an order of magnitude have also been proposed (Eichler, 1980). If we assume that a power input of 3×10^{39} ergs/s yields protons of energy ≥ 10 GeV and that these protons suffer nuclear collisions, a detector of 10^6 tons would permit the observation of about 30 neutrino events per year. With 10^7 tons, several different sources of neutrinos become detectable.

CONCLUSIONS

We conclude that a neutrino detector of $\geq 10^6$ tons on the Moon—*i.e.*, one considerably more compact than the proposed DUMAND array—would open up a new window of neutrino astronomy, making possible the study of neutrinos at $1\text{–}10^3$ GeV*. The effort must probably await the establishment of a substantial lunar colony; because of its large size, the detector would probably have to be locally constructed, perhaps of glass fabricated from lunar materials.

Acknowledgments. *One of the authors (MMS) expresses his appreciation to Professors R. Kippenhahn and W. Hillebrandt for their hospitality at the Max-Planck Institut für Astrophysik in Garching. He thanks Professor F. Reines for stimulating his interest in this problem.*

REFERENCES

Berezinsky V. S. (1976) Ultra HE neutrinos and detection possibilities by DUMAND. In *Proc. 1976 DUMAND Workshop* (A. Roberts, ed.), pp. 229–255. Univ. Hawaii, Honolulu.

Cherry J. L. and Lande K. (1985) Proposal for a neutrino telescope on the Moon (abstract). In *Papers Presented to the Symposium on Lunar Bases and Space Activities of the 21st Century*, p. 50. NASA/Johnson Space Center, Houston.

Eichler D. (1980) SS433: A possible neutrino source. In *Proc. of 1980 DUMAND Symposium, Vol. 2,* pp. 266–271. Hawaii Dumand Center, Honolulu.

Erofeeva I. N., Lyotov S. I., Murzin V. S., Kolomeets E. V., Albers J., and Kotzer P. (1983) Detection of solar flare neutrinos. In *Proc. 18th Intl. Cosmic Ray Conf.*, pp. 104–107. Tata Institute for Fundamental Research, Bombay.

Grindlay J. E., Band D., Seward F., Leahy D., Weisskopf M. C., and Marshall F. E. (1984) The central x-ray source in SS433. *Ap. J., 277,* 286.

Peterson V. Z. (1983) Deep underwater muon and neutrino detection. In *Composition and Origin of Cosmic Rays* (M. M. Shapiro, ed.), pp. 251–268. Reidel, Dordrecht.

Reines F. (1965), as reported in Shapiro M. M. (1965) Galactic cosmic rays. In *NASA 1965 Summer Conference on Lunar Exploration and Science*, pp. 317–330, NASA SP-88. NASA, Washington.

Shapiro M. M. and Silberberg R. (1979) Neutrinos from young supernova remnants. In *Proc. 16th Intl. Cosmic Ray Conf., Vol. 10,* pp. 363–366. Univ. Tokyo, Kyoto.

Silberberg R. and Shapiro M. M. (1979) Neutrinos as a probe for the nature of and process in active galactic neutrinos. In *Proc. 16th Intl. Cosmic Ray Conf., Vol. 10,* pp. 357–362. Univ. Tokyo, Kyoto.

**Even a smaller detector of 10^4 tons could detect a giant solar flare like that of Feb. 23, 1956. The pulse is likely to be of such short duration (< 20 min) that the atmospheric background would not degrade terrestrial observation.*

Stecker F. W., Shapiro M. M., and Silberberg R. (1979) Galactic and extra-galactic HE neutrinos. In *Proc. 16th Intl. Cosmic Ray Conf., Vol. 10*, pp. 346–351. Univ. Tokyo, Kyoto.

A LUNAR NEUTRINO DETECTOR

M. Cherry and K. Lande

Department of Physics, University of Pennsylvania, Philadelpha, PA 19104

The major experimental difficulty in neutrino astronomy lies in the fact that expected event rates are exceedingly small (typically 10^{-4} or fewer neutrinos per year per sr per ton of detector). A detector must therefore be extremely massive and must be located in a very low background environment. Over the energy range 1 GeV–10 TeV, the neutrino background on the Moon is lower than on the Earth (at some energies by as much as 10^{-3}–10^{-4}). At both lower and higher energies, the lunar background is just as high as that on Earth, but in the proper energy range, the Moon may be the only possible site for neutrino astronomy. We review the properties of terrestrial neutrino detectors located deep underground or underwater, discuss the calculated and measured backgrounds, and demonstrate the improvement to be obtained with a lunar location. In addition, we briefly discuss a possible design for a 10^6 ton lunar neutrino detector.

INTRODUCTION

Neutrinos are produced as a result of collisions of cosmic rays with the ambient material in astronomical sources, the Earth's atmosphere and surface, and the lunar surface. Neutrinos are produced directly when cosmic rays interact to produce secondary charmed mesons, when π and K mesons decay to μ mesons and electrons, and when the μ mesons decay to electrons. Gamma rays are produced from neutral π° meson decays resulting from the same cosmic ray interactions, and the numerous satellite, balloon, and ground-based measurements of astronomical gamma ray sources make it clear that such cosmic ray interactions occur in a great variety of sources and that large numbers of neutrinos must also be emitted. Neutrinos have a much greater interaction mean free path than do gamma rays, however ($\lambda_{\nu_\mu}/\lambda_\gamma \sim 10^{11}$ at 100 GeV), so that if the neutrinos can be detected, they can be used to probe to far greater depths in dense media. The significance of astronomical neutrino observations and the relationship to gamma ray astronomy have been discussed previously by, among others, Berezinsky and Zatsepin (1970), Lande (1979), Fichtel (1979), Stecker (1979), Shapiro and Silberberg (1983), Silberberg and Shapiro (1983), and Lee and Bludman (1985).

Due to the low detection rates and high terrestrial background, the Moon may provide the best (and perhaps the only) location for observations of astronomical neutrinos at energies 1 GeV–10 TeV. In the following section, we briefly review the estimates of neutrino detection rates from astronomical sources and the Earth's atmosphere. We list the existing large underground detectors and discuss some of the experimental difficulties associated with the low anticipated astronomical event rates and the high atmospheric background. In a later section, we describe the lower neutrino background expected for a detector on the Moon; in the last section, we discuss how a lunar neutrino detector might look.

PRESENT EXPERIMENTAL STATUS OF NEUTRINO ASTRONOMY

If observed high energy gamma rays are primarily nucleonic in origin (*i.e.*, they result from cosmic ray interactions and the resulting π^0 meson decays) and if the bulk of the emission originates in source regions whose thickness (t) is small compared to the gamma ray attenuation length (*i.e.*, $t \lesssim 10^2 g\ cm^{-2}$), then the gamma ray and neutrino emission must be comparable. Unfortunately, since the neutrino interacts so weakly, it can be detected only very inefficiently. Fichtel (1979) has estimated neutrino detection rates based on observed gamma ray fluxes; for the quasar 3C273, the active galaxy Cen A, the central region of our own galaxy, γ-ray pulsars, and the galactic γ-ray source Cyg X-3, he finds expected interaction rates of only 3×10^{-9} detected neutrinos per yr per kiloton of detector above 1 TeV and 3×10^{-10} neutrinos per yr per kiloton above 10 TeV, corresponding to neutrino fluxes of 10^{-12} and $10^{-13} cm^{-2}\ s^{-1}$, respectively. At lower energies, Cherry and Lande (unpublished data, 1985) expect rates of 10^{-5} neutrinos $ton^{-1}\ yr^{-1}\ sr^{-1}$ from the Crab and 10^{-4} $ton^{-1}\ yr^{-1}\ sr^{-1}$ from the galactic center above 100 MeV.

These small rates of neutrinos interacting in a detector must be visible in the presence of a nearly isotropic background of neutrinos produced locally by cosmic rays interacting in the Earth's atmosphere. The fluxes of atmospheric neutrinos have been calculated. The recent analytic calculation of Dar (1983) and the detailed Monte Carlo calculation of Gaisser *et al.* (1983) are in substantial agreement with each other and with the measured interaction rates in the underground IMB proton decay detector (Bionta *et al.*, 1983). In this and other large underground detectors, it is impossible to measure energies as high as 1 TeV; rather, one measures the integral flux of all neutrinos with energies above a relatively low threshold (~200 MeV for the IMB detector). The rates measured deep underground ($10^{-2}\ T^{-1}yr^{-1}$) are in substantial agreement with the calculated interaction rate of atmospheric neutrinos (Gaisser and Stanev, 1984).

Gaisser and Stanev (1984) have also considered the case of neutrinos traveling upward through the Earth, interacting in the rock beneath the detector, and producing muons that then continue upward through the detector. If the detector is located sufficiently deep underground that the flux of penetrating downward-moving cosmic ray muons is sharply reduced, then the small upward flux can be measured. The resulting measured upward fluxes ($\sim 2 - 7 \times 10^{-13} cm^{-2}\ s^{-1}\ sr^{-1}$) are also in substantial agreement with the calculations of atmospheric background.

It appears, then, that the atmospheric neutrino background is reasonably well understood, and one can compare the calculated atmospheric neutrino spectrum to predicted astronomical spectra. An example of such a comparison is shown in Fig. 1. The neutrino flux from the Earth's atmosphere is based on the analytic approach described by Dar (1983). The line labeled "Galactic" gives an estimate of the diffuse neutrino flux expected from the region of the galactic center. This estimate agrees substantially with the detailed calculation of Stecker (1979). One can see that the atmospheric spectrum is steeper than the galactic spectrum and falls below the galactic spectrum for neutrino energies $E \lesssim 10$ TeV. It is for this reason that many previous discussions of neutrino astronomy have typically emphasized neutrino energies above several TeV.

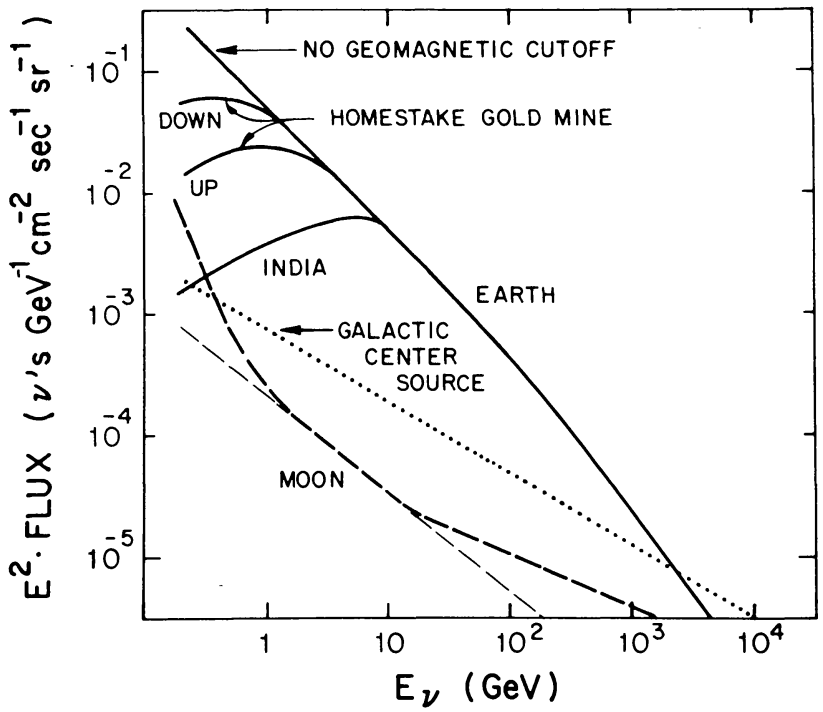

Figure 1. Flux of muon neutrinos and anti-neutrinos. The terrestrial flux is shown for the case of no geomagnetic field, for vertically downward and upward neutrinos at the Homestake and IMB sites, and for downward neutrinos in India in the direction of maximum geomagnetic shielding. The lunar flux is calculated (light line) for π and K meson decays; the heavy line takes into account the effect of low-energy K^+ and high-energy charm decays.

The detectors envisioned for these observations have typically been massive instruments located deep underground or underwater in order to reduce the background. For example, the proposed DUMAND detector (Stenger, 1984) involves instrumenting a 50 Mt 5×10^7 m^3 volume of the Pacific Ocean at a depth of 4.7 km; the IMB detector is an 8 kT water detector at a depth of 600 m in the Morton Salt Mine near Cleveland, Ohio; and a 1–5 kT liquid scintillation detector (Cherry *et al.*, 1983) is eventually planned at a depth of 1500 m in the Homestake Gold Mine in Lead, South Dakota. The largest underground detector so far proposed is approximately 30 kT of water (A. Mann and

Table 1. Operating Underground Detectors

Detector	Location	Mass (tons)	Type
< 100 Tons			
Soudan I	Minnesota, U.S.	30	Fe Calorimeter
100–1000 Tons			
Kolar	India	140	Fe Calorimeter
Homestake	South Dakota, U.S.	140	Liquid Scintillator
NUSEX	Italy-France	150	Fe Calorimeter
Baksan	USSR	330	Liquid Scintillator
Frejus	Italy-France	800	Fe Calorimeter
HPW	Utah, U.S.	800	Water Cerenkov
> 1000 Tons			
Kamioka	Japan	3000	Water Cerenkov
IMB	Ohio, U.S.	8000	Water Cerenkov

B. Cortez, private communications, 1985). A list of existing large underground detectors is given in Table 1. A number of detection techniques have been utilized and shown to be effective. In each case, though, the cost of the experiment is typically $1–5 million per kiloton.

Unfortunately, even for massive detectors located deep underground where the remnant cosmic ray background fluxes are low, neutrino and neutrino-induced muon rates are still meager—both in terms of the absolute number of events per year and compared to the atmospheric background. As an example, the calculated rates of high-energy muons penetrating from the Earth's surface to a depth of 1500 m and the rate of lower-energy neutrino-induced muons at the same depth are shown in Fig. 2 for the case of the Homestake Large Area Scintillation Detector (Cherry *et al.*, 1985). The penetrating muons must have a minimum energy of $E_\mu \sim 2.6$ TeV at the Earth's surface in order to penetrate to the depth of the detector; the minimum energy detectable at the detector is $E_\mu \sim 2$ GeV. By contrast with the rates of Fig. 2, the galactic center spectrum of Fig. 1 would give 3 neutrino-induced muon events per year per sr above 2 GeV and 2 $yr^{-1}sr^{-1}$ above 100 GeV. With such extremely low event rates, it is absolutely essential to minimize the background as much as possible.

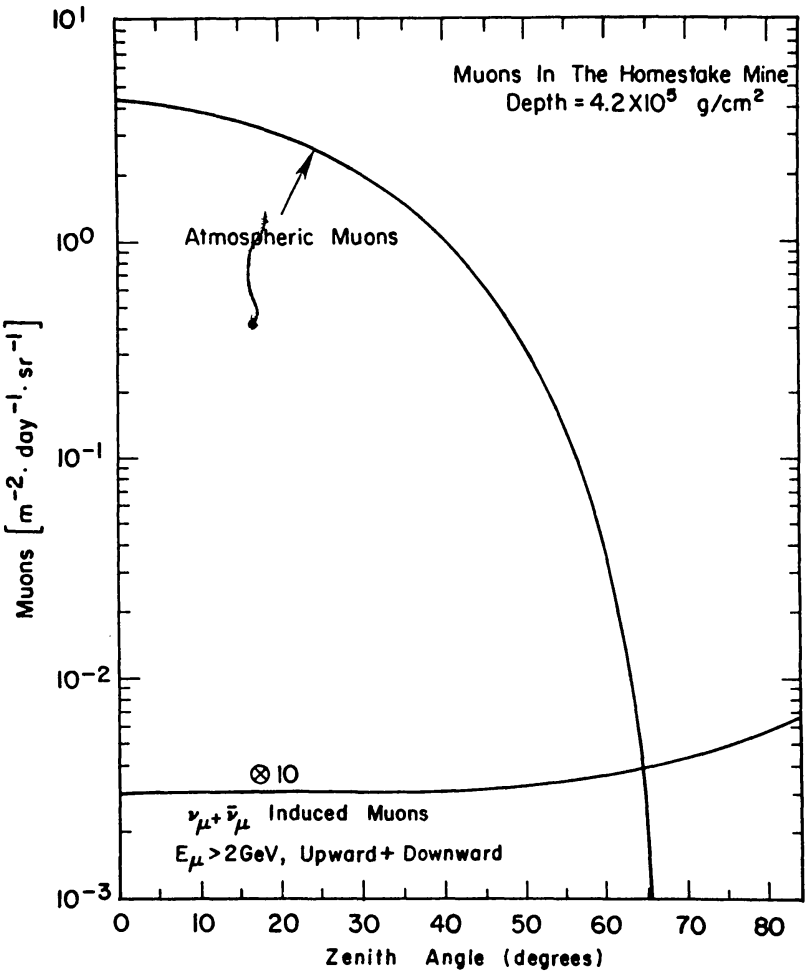

Figure 2. Predicted muon fluxes in the Homestake Gold Mine (Dar, written communication, 1984).

LUNAR AND TERRESTRIAL BACKGROUNDS

Over a large range of energies, the background problem can be alleviated by placing the detector above the atmosphere—for example, on the Moon. In the Earth's tenuous upper atmosphere, cosmic ray primaries have interaction lengths of 80 g cm^{-2}; the secondary pion interaction lengths are 120 g cm^{-2}. The total thickness of the atmosphere (1020 g cm^{-2} from the top of the atmosphere down to sea level) is many interaction lengths, so that even the Earth's surface is exceptionally well shielded from the primary cosmic rays and the hadronic components of the cosmic ray showers. In order to obtain the same hadron shielding, a lunar detector should be buried at a depth of at least 10^3g cm^{-2}, or 3 m.

Once a π (or K) meson is produced in the Earth's upper atmosphere, it can either interact or decay. Neutrinos are produced from the decays in Table 2. If the initial meson interacts, then neutrinos are produced by the decays of later-generation mesons formed lower in the atmosphere. The contribution from these secondary interactions is relatively small, however (20%). The meson decay length is a function of the meson energy E: $\lambda_\pi^{decay} = \beta\gamma c\tau = 7.8\beta\gamma$ m, where $\beta = v/c$ is the pion velocity in units of the speed of light c, $\gamma = E/mc^2$ is the energy in units of the mass, and τ is the pion lifetime at rest. By contrast, the pion interaction length is nearly independent of energy: $\lambda_\pi^{inter} = 120$ g cm^{-2} ($\lambda_\pi^{inter} \sim 6$ km in the Earth's upper atmosphere, $\lambda_\pi^{inter} \sim 0.4$ m in lunar rock). The probability of decay depends on the relative values of λ_π^{decay} and λ_π^{inter}.

$$\frac{\lambda_\pi^{inter}}{\lambda_\pi^{decay}} = \begin{cases} \dfrac{6000}{7.8\beta\gamma} \sim \dfrac{100}{E_\pi(\text{GeV})} & \text{for the Earth} \\ \dfrac{0.4}{7.8\beta\gamma} \sim \dfrac{1}{110E_\pi(\text{GeV})} & \text{for the Moon} \end{cases} \quad (1)$$

As long as this ratio is large, mesons will decay before they have a chance to interact, and the neutrino flux will be high. In the Earth's tenuous upper atmosphere, a large fraction of the mesons below 100 GeV decay; above 100 GeV, the decreasing decay probability suppresses neutrino production. In the dense lunar (or terrestrial) rock, the interactions occur long before the mesons have a chance to decay, and at all energies above 10 MeV, neutrino production is highly suppressed. The ratio of lunar and terrestrial neutrino fluxes from decay is roughly

$$\frac{\phi_\nu^{\pi}(\text{Moon})}{\phi_\nu^{\pi}(\text{Earth})} \sim \frac{\lambda_\pi^{inter}/\lambda_\pi^{decay}(\text{Moon})}{\lambda_\pi^{inter}/\lambda_\pi^{decay}(\text{Earth})} \sim \begin{cases} 1/110\ E_\pi \text{ for } E_\pi \lesssim 100 \text{ GeV} \\ 10^{-4} \text{ for } E_\pi \gtrsim 100 \text{ GeV} \end{cases} \quad (2)$$

We assume here that the production is the same in a lunar or terrestrial target of thickness 10^3 g cm^{-2}; we ignore the extra neutrinos produced in the terrestrial rock, since few mesons penetrate through the entire atmosphere; and we let the terrestrial ratio λ_π^{inter}/

Table 2. Most Probable Neutrino-Producing Decay Modes of Non-charmed Mesons

$\pi^\circ, \eta^\circ \rightarrow$	2γ	
$\pi^-, K_1 \rightarrow$	$\mu^- \bar{\nu}_\mu$	
$\pi^+, K^+ \rightarrow$	$\mu^+ \nu_\mu$	
		$\mu^+ \rightarrow e^+ \nu_e \bar{\nu}$
		$\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$
	$\pi^+ e^- \bar{\nu}_e$	
	$\pi^- e^+ \nu_e$	
$K^\circ \rightarrow$		
	$\pi^+ \mu^- \bar{\nu}_\mu$	
	$\pi^- \mu^+ \nu_\mu$	

λ_π^{decay} saturate at 1 for $E_\pi \lesssim 100$ GeV, where essentially all terrestrial pions decay. The neutrino flux from K mesons is similarly suppressed on the Moon.

The ratio ϕ_ν (Moon)/ϕ_ν (Earth) due to the competition between π and K meson interactions and decays is shown as the dashed curve in Fig. 3. At energies above 10 GeV the lunar suppression is reduced, however, by the effect of charmed meson production. The energy dependence of charm production is not yet well understood, but at Fermilab energies Ball *et al.* (1984) found cross sections for $pN \rightarrow D\bar{D}X$ and $pN \rightarrow \Lambda DX$ of 10–20μb/nucleon, about 10^3 times smaller than the corresponding $pN \rightarrow \pi X$ cross-sections. The lifetime of the charmed mesons is exceedingly short, however ($\tau_{D^\circ,D^\pm} \sim 4 - 9 \times 10^{-13}$ s, $\tau_{F\pm} \sim 2 \times 10^{-13}$ s), so that all charmed mesons decay promptly either on Earth or on the Moon. The effect of the Moon is therefore to suppress only the neutrinos from π and K mesons. The maximum suppression is given by the ratio (R) of prompt neutrinos from charm decay to neutrinos from π and K meson decay: $R \sim 10^{-3} - 10^{-4}$ (Elbert *et al.*, 1981; Inazawa and Kobayakawa, 1983). Above 10 GeV, the relative neutrino flux from charmed and π meson decays is given by the production ratio times the ratio of charmed meson to π meson decay probabilities:

$$\frac{\phi_\nu^c(\text{Moon})}{\phi_\nu^\pi(\text{Moon})} \sim 10^{-3} \times \frac{1}{1/110\, E_\pi} \sim \frac{E_\pi(\text{GeV})}{10} \qquad (3)$$

The solid curve above 10 GeV shows this behavior.

Below about 1 GeV, the terrestrial neutrino flux is reduced by the effects of the Earth's geomagnetic field. At geomagnetic latitude (λ), zenith angle (θ), and azimuthal angle (φ)

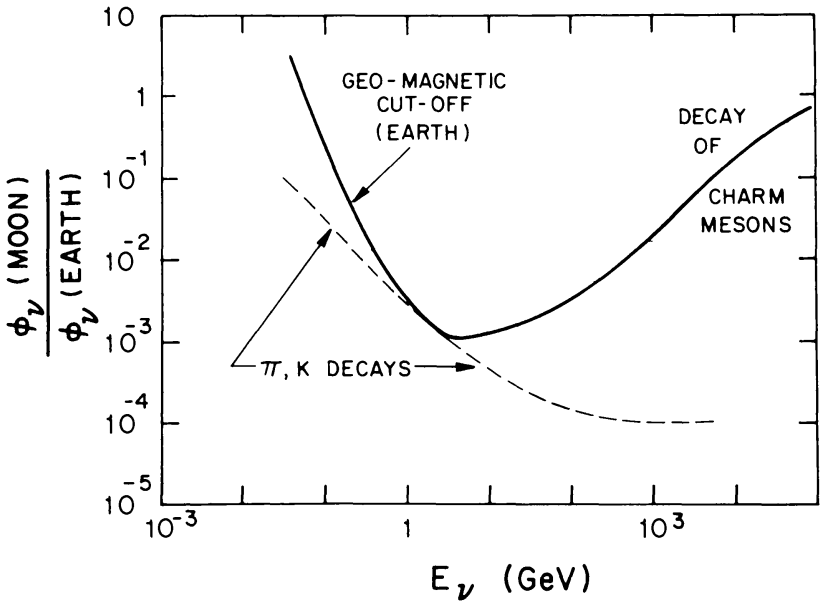

Figure 3. Ratio of lunar to terrestrial neutrino fluxes versus neutrino energy.

from magnetic North, the Earth's magnetic field effectively prevents primary cosmic rays with rigidity

$$p < 60 \cos^4\lambda\ (1 + \sqrt{1 + \cos^3\lambda \sin\theta \sin\varphi}\ )^2 GV \qquad (4)$$

from reaching the atmosphere. For downward-moving protons at the IMB and Homestake detectors, this cutoff is near 2 GV; at the Kolar Gold Fields in India, the cutoff is at 16 GV. This geomagnetic suppression does not apply to the Moon, so ϕ_ν(Moon)/ϕ_ν(Earth) increases at low energies; the details depend on the particular location and direction of cosmic ray incidence on the Earth, but the qualitative effect is shown in Fig 3.

Figure 1 shows the flux of neutrinos we calculate for the Earth without any geomagnetic field, for downward and upward trajectories at the Homestake detector, and for the direction of maximum cutoff rigidity (~60 GV) at the Kolar detector in India. The dashed line is the lunar flux under 4 m of rock. The dotted line is the flux expected from cosmic ray-matter interactions near the galactic center. The Moon offers a major suppression of the neutrino background between about 1 GeV and 1 TeV and, in particular, may make it possible to see the very interesting galactic center source as well as numerous other potential astronomical sources.

A LUNAR NEUTRINO DETECTOR

Since neutrinos can escape from dense sources that are opaque to gamma rays, it is quite possible that actual neutrino fluxes may turn out to be much larger than those predicted on the basis of the observed γ-rays. If one is to embark on a project as complex and costly as a lunar neutrino detector, however, one must probably adopt extremely

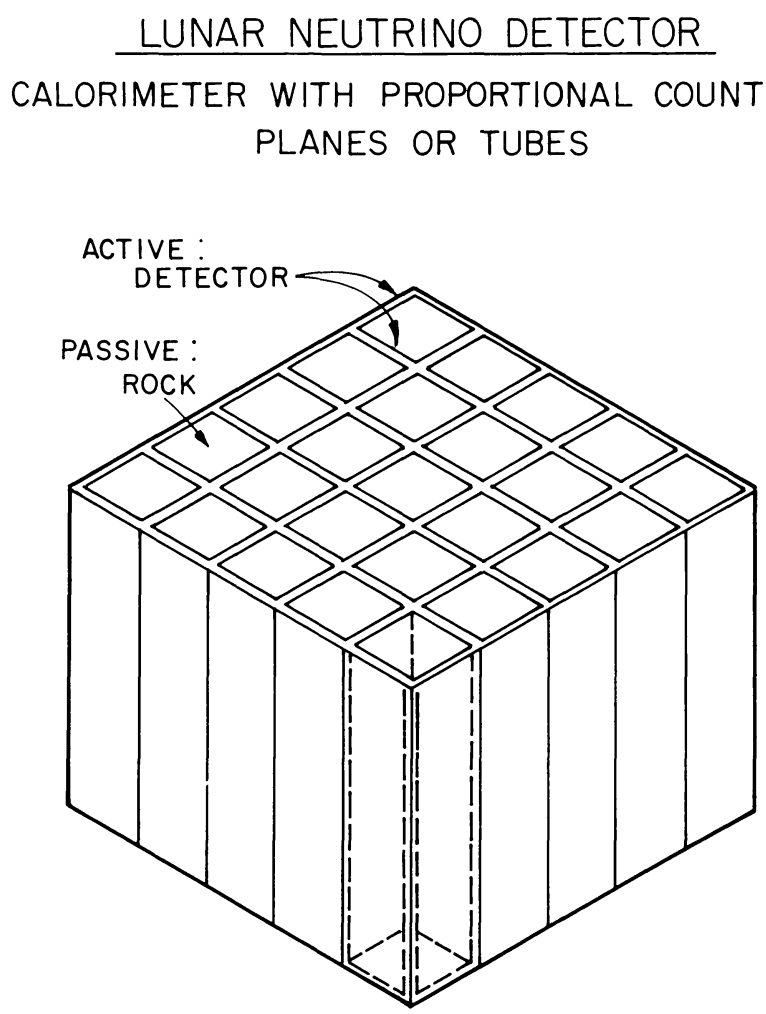

Figure 4. Schematic diagram of a possible lunar neutrino detector, consisting of active detector planes distributed through a large mass (10^6 tons) of lunar rock.

conservative design criteria. In particular, one must probably start from the very low neutrino fluxes estimated from the γ-ray measurements. For a flux level 10^{-5} $ton^{-1}yr^{-1}sr^{-1}$, one therefore needs a detector mass on the order of 10^6 tons. Since it is presumably unreasonable to transport 10^6 tons of water or liquid scintillator to the Moon, one must rather use local lunar material for the main detector mass. For example, a block of lunar surface material 30 m high $\times$ 100 m wide $\times$ 100 m long buried beneath several meters of soil would provide a well-shielded 900 kT neutrino target.

The detector might be instrumented as a calorimeter with planes of gas-filled drift chambers set out through the detector volume, as in Fig. 4. If the detector planes are separated by 1 m, then the detector threshold is given by the minimum muon energy required to penetrate 2 detector layers—about 1 GeV. The 30–100 m dimensions of the detector make muon energy measurements possible up to 100 GeV for some cases.

The detector would presumably have to be fabricated in a lunar laboratory. Wire, filling gas, and electronics could be supplied from Earth, but the main detector elements might be locally constructed gas-tight glass structures. Trenching, drilling, and digging of the detector volume would need to be done on site. Individual drift chamber dimensions

might be 3 m × 3 m × 3 cm, requiring 3×10^4 elements. Although the size of such a detector is certainly mammoth by terrestrial standards, the electronic complexity is comparable to other high energy experiments.

The logistical, engineering, and financial problems associated with a lunar neutrino detector would be enormous. From the observational point of view, however, the low neutrino backgrounds would make lunar viewing conditions significantly better than anything possible on the Earth in the energy range 1–1000 GeV.

Acknowledgments. *Funding for the University of Pennsylvania/Homestake program is provided by the United States Department of Energy. We have benefitted from many discussions with Drs. S. Bludman, A. Dar, T. Gaisser, H. Lee, and T. Stanev.*

REFERENCES

Ball R. C. *et al.*, (1984) Prompt muon-neutrino production in a 400-GeV proton beam-dump experiment. *Phys. Rev. Lett., 51*, 743–746.

Berezinsky V. S. and Zatsepin G. T. (1970) Cosmic neutrinos of ultrahigh energy. *Sov. J. Nucl. Phys., 11*, 111–114.

Bionta R. M. *et al.* (1983) Results from the IMB detector. In *Fourth Workshop on Grand Unification* (H. A. Weldon, P. Langacker, and P. J. Steinhardt, eds.) pp. 46–68. Birkhauser, Boston.

Cherry M. L., Davidson I., Lande K., Lee C. K., Marshall E., Steinberg R. I., Cleveland B., Davis R., and Lowenstein D. (1983) Physics opportunities with the Homestake large area scintillation detector. *ICOMAN '83: International Colloquium on Matter Non-Conservation*, p. 133–143. INFN, Rome.

Cherry M. L., Corbato S., Kieda D., Lande K., Lee C. K., and Steinberg R. I. (1985) The Homestake large area scintillation detector and cosmic ray telescope. In *Solar Neutrinos and Neutrino Astronomy* (M. L. Cherry, K. Lande, and W. A. Fowler, eds.) pp. 32–49. American Institute of Physics, New York.

Dar A. (1983) Atmospheric neutrinos, astrophysical neutrinos, and proton decay experiments. *Phys. Rev. Lett., 51*, 227–231.

Elbert J. W., Gaisser T. K., and Stanev T. (1981) Possible studies with DUMAND and a surface air shower detector. In *DUMAND-80* (V. J. Stenger, ed.), pp. 222–228. Hawaii DUMAND Center, Honolulu.

Fichtel C. E. (1979) The significance of gamma-ray observations for neutrino astronomy. *Proc. 1978 DUMAND Summer Workshop* (A. Roberts, ed.) p. 289–312. DUMAND Scripps Institute of Oceanography, La Jolla, CA.

Gaisser T. K. and Stanev T. S. (1984) Interaction rates of atmospheric neutrinos. In *ICOBAN '84: International Colloquium on Baryon Non-Conservation* (D. Cline, ed.), Park City, UT. In Press.

Gaisser T. K., Stanev T., Bludman S. A., and Lee H. (1983) Flux of atmospheric neutrinos. *Phys. Rev. Lett., 51*, 223–226.

Inazawa H. and Kobayakawa K. (1983) The production of prompt cosmic ray muons and neutrinos. *Prog. Theor. Physics (Japan), 69*, 1195–1206.

Lande K. (1979) Experimental neutrino astrophysics. *Ann. Rev. Nucl. and Part. Sci., 29* (J. D. Jackson, H. E. Gove, and R. F. Schwitters, eds.), pp. 395–410. Annual Review, Palo Alto, CA.

Lee H. and Bludman S. A. (1985) Neutrino production from discrete high-energy gamma ray sources. *Astrophys. J., 290*, 28–32.

Shapiro M. M. and Silberberg R. (1983) Gamma rays and neutrinos as complementary probes in astrophysics. *Space Sci. Revs., 36*, 51–56.

Silberberg R. and Shapiro M. M. (1983) Sources of extragalactic cosmic rays: Photons and neutrinos as probes. In *Composition and Origin of Cosmic Rays* (M. M. Shapiro, ed.) pp. 231–244. Reidel, Dordrecht.

Stecker F. W. (1979) Diffuse fluxes of cosmic high-energy neutrinos. *Astrophys. J., 228*, 919–927.

Stenger V. J. (1984) The production of very high energy photons and neutrinos from cosmic proton sources. *Astrophys. J., 284*, 810–816.

NEUTRINO MEASUREMENTS ON THE MOON

Albert G. Petschek

Los Alamos National Laboratory, Los Alamos, NM 87545 and
New Mexico Institute of Mining and Technology, Socorro, NM 87801

Several possible neutrino experiments on the Moon are discussed in the light of the expected background rates. Observations of neutrino oscillations may be feasible.

Grand unified theories suggest a variety of novel physical processes, including nucleon decay, that is, the breakup of neutrons or protons into lighter particles, and neutrino oscillations, a process in which neutrinos with one type of weak interaction change back and forth into neutrinos with another interaction. In order to test these theories, a number of very large proton decay detectors with masses between 0.14 and 3 k tons have been built on Earth.

Protons are presumed to decay into various leptons and mesons, possibly $e^+ + \pi^0$. For a more complete list, see the review of particle properties (Particle Data Group, 1984). The detectors attempt to observe these decay products, or their daughters, by observing Cherenkov radiation in a large volume of water (Bionta *et al.*, 1983) or ionization in gas-filled proportional tubes (Peterson, 1983; Krishnaswamy *et al.*, 1983). A principal source of background in these detectors, none of which has observed a proton decay, is neutrino interactions. The neutrinos in question are generated by the interaction of cosmic rays with the upper atmosphere, producing relativistic pions and other particles that decay to produce neutrinos of high (Doppler shifted) energy before they interact with another nucleus in the rarefied upper atmosphere. Naively, one might expect a substantially lower background on the Moon, since interactions with the solid lunar material would take place before decay. Accordingly, the Lunar Base Working Group (Duke *et al.*, 1984) suggested that the Moon would be a suitable location for a proton decay detector sensitive to 10^{34} or 10^{35} years proton lifetime. Such a lifetime corresponds to between 0.06 and 0.006 decays per kiloton-year if all the nucleons, including those bound in nuclei, are active and fewer decays otherwise. Thus, exposures of 100 to more than 1000 k ton-years would be required, a massive undertaking on Earth, let alone on the Moon.

As is detailed by Cherry and Lande (1985) the neutrino background on the Moon is less than that on the Earth only in a limited energy band, one that just begins at the 1 GeV proton decay energy. The total energy of proton decay, including the rest energy of the decay products is, of course, the proton rest energy of 938 MeV, less than that at which the background is reduced. Higher backgrounds mean poor experiments; poor experiments do not get funded. Hence, it is neither worthwhile nor possible to move proton decay experiments to the Moon, and it will not be possible to piggyback neutrino detection on the huge decay detectors that would be required. Nevertheless, it is of interest

to discuss what neutrino experiments might usefully be done on the Moon. Further discussion may be found in the paper by Shapiro and Silberberg (1985).

Another source of background, of interest in its own right, is astrophysical neutrinos. Their spectrum dominates atmospheric neutrinos even on Earth above 1000 GeV. These neutrinos originate in a variety of places. The ones with highest energy originate in interactions between cosmic rays and interstellar matter (Dar, 1983) or cosmic rays and the 3 K background radiation (Stecker, 1979). Lower energy neutrinos arise from stellar collapse (Burrows, 1984). This background should be the same on Earth as on the Moon, and the lower local background above 1 GeV will allow it to be studied to somewhat smaller energies on the Moon than on Earth.

Neutrinos produced in the Earth's atmosphere will also reach the Moon. Since the Earth subtends a solid angle of 5×10^{-5} of 4π at the Moon, the rate of interaction of atmospheric neutrinos with a lunar detector will be reduced from the terrestrial value of 100 per k ton year (Gaisser and Stanev, 1983) to 5×10^{-3} k ton^{-1} year^{-1}. This is a few during the 100–1000 k ton years of exposure required in the proton-decay experiment. If a really large detector were to be built, it might be possible to detect neutrino oscillations in a new range of Δm^2, the difference in the squares of the masses of the two neutrinos. The probability of transition to a second neutrino, for example, from the electron neutrino to the μ or τ neutrino, depends on the parameter Δm^2 L/E (Boehm, 1983) where L is the flight path and E the neutrino energy. Experiments explore values of this parameter ~1 if Δm^2 is in eV^2, L is m, and E is MeV. Thus reactor experiments explore $\Delta m^2 =$ 1 eV^2 or a little less (L is a few m, E a few MeV). Accelerator experiments use much higher energies and larger flight paths, but still explore a similar range of Δm^2. Neutrinos have been observed from the sun at well below the expected rate (Bahcall *et al.*, 1982). Since the neutrinos emitted by the sun are electron neutrinos and the detection method (by the transmutation of Cl to Ar) detects only these same neutrinos, a possible explanation of the discrepancy between theory and experiment is that an oscillation into another neutrino has taken place. With this assumption the solar neutrino experiment can be viewed as an oscillation experiment with E a few MeV and $L = 1.5 \times 10^{11}$ m, corresponding to very small Δm^2.

Gaisser and Stanev (1984) have put limits on neutrino oscillations by observing the dependence of the intensity of upward-going neutrinos produced in the Earth's atmosphere on angle, that is, on path length to the detector from the source (the atmosphere). In this experiment the characteristic energy is 1000 MeV and the characteristic distance is an Earth diameter so that $\Delta m^2 \sim 10^{-4}$ eV^2 is explored, less than for accelerator or reactor experiments but much greater than for the solar experiment. If, as is suggested at the beginning of the preceding paragraph, these same neutrinos can be detected on the Moon, then another region of Δm^2 can be explored. This region corresponds to mass differences much larger than those of the solar experiments but smaller than those of Gaisser and Stanev. In contrast to the experiment with solar neutrinos, which must rely on a calculation of the source, the experiment on the Moon would derive its neutrino source from terrestrial measurements such as those used by Gaisser and Stanev.

In summary, a few interesting neutrino experiments can be contemplated on the Moon, but they require massive detectors. A body with no atmosphere but with a magnetic moment comparable to the Earth's to reduce the low energy neutrino background would be much more suitable.

REFERENCES

Bahcall J. N., Huebner W. F., Lubow S. H., Parker T. D., and Ulrich R. K. (1982) Standard solar models and the uncertainties in predicted capture rates of solar neutrinos. *Rev. Mod. Phys., 54*, 767–800.

Bionta R. M., *et al.* (1983) Results from IMB detector. In *Fourth Workshop on Grand Unification* (H. A. Weldon, P. Langacker, and P. J. Steinhardt, eds.), pp. 46–68. Birkhäuser, Boston.

Boehm F. (1983) Neutrino mass and neutrino oscillations. In *Fourth Workshop on Grand Unification* (H. A. Weldon, P. Langacker, and P. J. Steinhardt, eds.), pp. 163–173. Birkhäuser, Boston.

Burrows A. (1984) On detecting stellar collapse with neutrinos. *Astrophys. J., 283*, 848–852.

Cherry M. and Lande K. (1985) Proposal for a neutrino telescope on the Moon. This volume.

Dar A. (1983) Atmospheric neutrinos and astrophysical neutrinos in proton decay experiments. In *Fourth Workshop on Grand Unification* (H. A. Weldon, P. Langacker, and P. J. Steinhardt, eds.), pp. 101–114. Birkhäuser, Boston.

Duke M. B., Mendell W. W., and Keaton P. W. (1984) *Report of the Lunar Base Working Group.* LALP-84-43. Los Alamos National Laboratory, Los Alamos. 41 pp.

Gaisser T. K. and Stanev T. (1983) Calibration with cosmic ray neutrinos. *AIP Conference Proceedings #114* (M. Blecher and K. Gotow, eds.), pp. 89–97. American Institute of Physics, New York.

Gaisser T. K. and Stanev T. (1984) Neutrino induced muon flux deep underground and search for neutrino oscillations. *Phys. Rev. D 30*, 985–990.

Krishnaswamy M. R., Menon M. G. K., Mondal N. K., Narasimham V. S., Sreekantan B. V., Hayashi Y., Ito N., Kawakami S., and Miyake S. (1983) The K.G.F. nuclear decay experiment. In *Fourth Workshop on Grand Unification* (H. A. Weldon, P. Langacker, and P. J. Steinhardt, eds.), pp. 25–34. Birkhäuser, Boston.

Particle Data Group (1984) Review of Particle Properties, *Rev. Mod. Phys., 56*, S1–S304.

Peterson E. (1983) New results from the Soudan 1 detector. In *Fourth Workshop on Grand Unification* (H. A. Weldon, P. Langacker, and P. J. Steinhardt, eds.), pp. 35–45. Birkhäuser, Boston.

Shapiro M. M. and Silberberg R. (1985) High-energy neutrino astronomy from a lunar observatory. This volume.

Stecker F. W. (1979) Diffuse fluxes of cosmic high-energy neutrinos. *Astrophys. J., 228*, 919–927.

MASS EXTINCTIONS AND COSMIC COLLISIONS: A LUNAR TEST

Friedrich Hörz

Experimental Planetology Branch, SN4, NASA/Johnson Space Center, Houston, TX 77058

Chemical and physical evidence strongly indicates synchroneity between the Cretaceous/Tertiary mass extinctions and the collision of a large cosmic object with Earth some 65 m.y. ago. Statistical time series analysis of the marine extinction record for the past 200 m.y. reveals periodicity on the order of 30 m.y. Time series analyses on the formation ages of terrestrial impact craters may yield similar periodicities; extinction and cratering cycles may even be in phase. However, the crater analyses are somewhat ambiguous because of the small number of terrestrial craters that can be dated precisely. It is therefore suggested that additional formation ages of lunar craters be obtained such that statistically improved and sound time series analysis for collision events in the Earth-Moon system can be performed. If synchroneity between cratering rate and mass extinctions were confirmed, far-reaching implications for the evolution of life would result.

INTRODUCTION

Paleontologists, geochemists, planetologists, astrophysicists, and others are currently debating whether some or all major mass extinctions in the geologic record of Earth are caused by the collision of massive, cosmic objects (*e.g.*, Silver and Schultz, 1982; Holland and Trendall, 1984).

The fossil record abounds with evidence that major mass extinctions are followed by an explosive "blooming" of new life forms that rapidly occupy apparently empty habitats. Thus, the sudden disappearance of old life forms and the almost equally sudden radiation of new forms (or of a few survivors) are closely intertwined. The brevity of the time scales involved can be termed "sudden" only in a geologic context. Generally, an enormous diversity of both marine and continental life is affected. This argues very strongly for some "catastrophic" deterioration of the environment on global scales. As a consequence, modern paleontological views allow room for both catastrophism and orderly evolution; the term "spasmodic" evolution is frequently used (Raup, 1984).

As detailed below, a strong case can be made for the correlation between one major mass extinction and the impact of a massive cosmic object; tentative evidence exists for a second case. Speculative suggestions postulate a general, causal link between most (all?) mass extinctions and hypervelocity collisions. Large-scale collisions must have taken place throughout geologic time within the entire solar system because most planetary surfaces are pockmarked by hypervelocity craters. Many of these craters are 100 km in diameter and represent global catastrophies.

THE CRETACEOUS/TERTIARY MASS EXTINCTION

The entire debate started with a watershed discovery by Alvarez *et al.* (1980). They found that siderophile trace elements such as Ir and Pt are greatly enriched and sharply concentrated in exactly the same strata in which palenotologists had placed the boundary between the Cretaceous (K) and Tertiary (T) geologic time periods. The K-T boundary can be identified in some localities with centimeter precision by a sharp decline, if not total disappearance, of many Cretaceous life forms and the subsequent blooming of new, Tertiary species. It is estimated that at least 85% of all Cretaceous biomass became extinct and no land animal larger than 25 kg survived this catastrophe some 65 m.y. ago (Russell, 1979). Alvarez *et al.* (1980) proposed that the unusual concentration of siderophile elements in these strata must be remnants of a gigantic meteorite. Subsequent work by many geochemists demonstrated that this siderophile-rich layer is of global extent; importantly, it may be found in marine as well as continental sediments (Fig. 1). In addition, Bohor *et al.* (1984) discovered "shocked" quartz grains in the K-T layer. These grains display deformation features that are diagnostic for the passage of transient, high pressure shock waves, which in turn can only be generated in nature by high speed impact. Thus, there

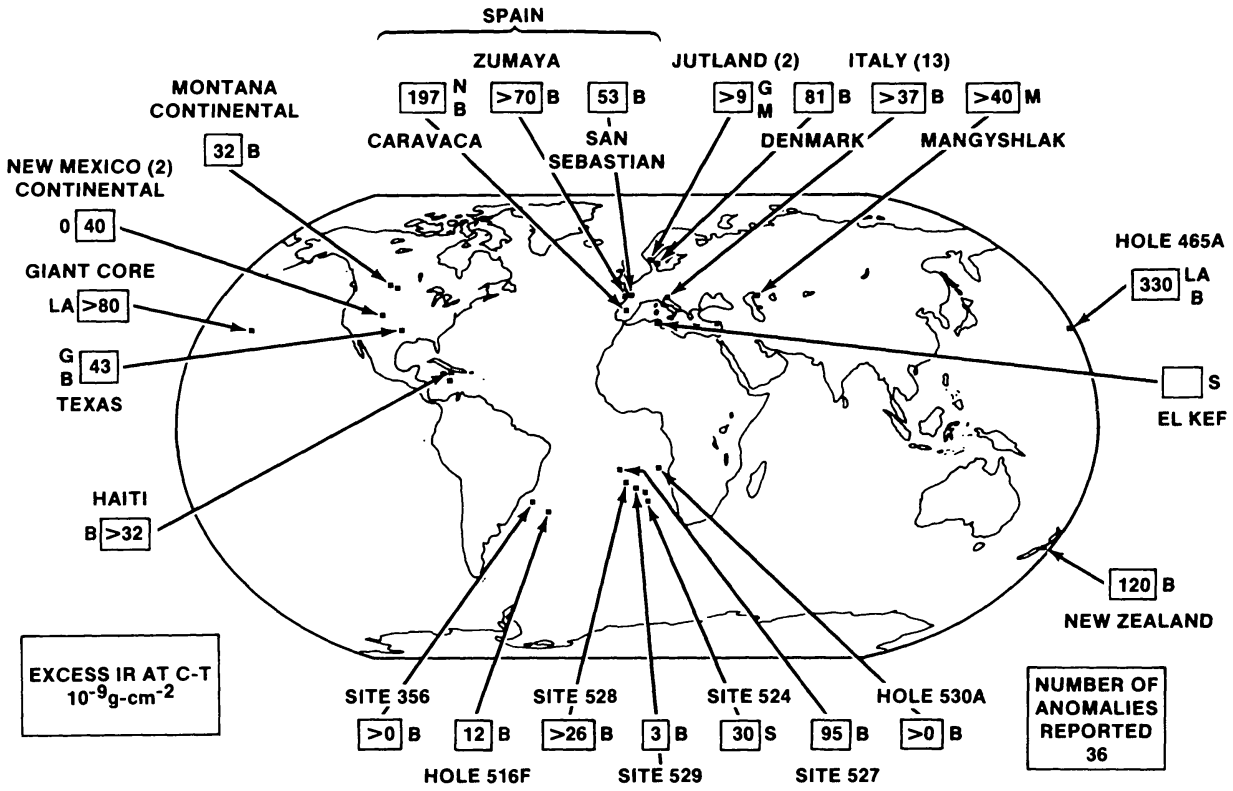

Figure 1. Worldwide distribution of locations where 65-m.y.-old strata of anomalously rich siderophile element concentrations have been observed. The letters next to the boxes refer to major analytical groups specializing in such analyses, and the numbers in the box refer to the absolute concentration of the key element Ir in units of 10^{-9} g per cm^2 (from Alvarez et al., 1982).

is chemical and physical evidence that the mass extinctions at the K–T boundary are indeed synchronous with a large collisional event.

The quantity of meteoritic material in the K–T layer requires an impacting body some 5–10 km across. Upon impact, such a body releases energies measured in millions of megatons of TNT (billions of Hiroshima bombs!), and it produces a crater some 50–100 km in diameter. These are minimum estimates; the crater could have been larger. Nevertheless, such a large, 65-m.y.-old crater is not presently known on Earth. This is a somewhat disturbing aspect of the collisional scenario, but by no means a fatal one, considering the dynamic nature of Earth. Also, very little is known about the actual kill mechanism(s), *i.e.*, what kind of environmental changes and specific stresses did the impact generate? Extreme temperature excursions can readily be envisioned as can changes of the pH of oceans and surface water. Direct poisoning by heavy metals and catastrophic floodings, if the impact occurred in the ocean, were also suggested (Toon *et al.*, 1982; O'Keefe and Ahrens, 1982; Hsu *et al.*, 1982; Gault and Sonett, 1982; and others). In all likelihood, a combination of these (and other?) effects resulted in short- and long-term imbalances in the environment and associated breakdown(s) of the food chain.

What is the probability of collision between a cosmic object some few kilometers across and the surface of Earth? Fortunately, there are three independent ways to estimate this probability: (1) The total number of impact structures on Earth is about 100 (see Fig. 2). This crater population may be used to calculate a minimum crater production rate for Earth (Grieve, 1982; Grieve *et al.*, 1985). (2) It was possible to date the ages of lunar volcanic lava flows returned by Apollo with isotopic methods. All craters that punched through these lava flows must therefore be younger than the flows themselves. Combining the observable number of such craters and the age of the target rock, lunar

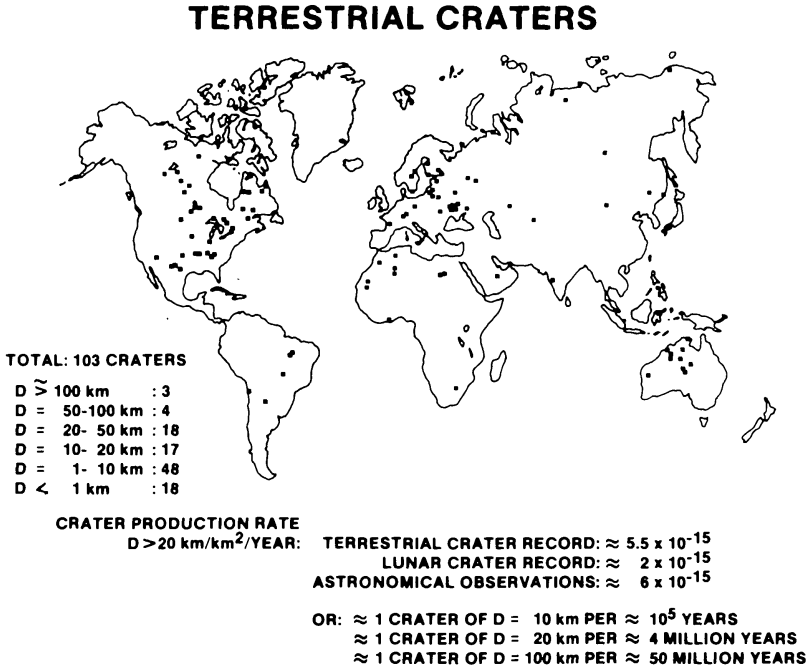

Figure 2. Locations of known terrestrial craters (from Grieve, 1982) and their approximate size distribution. Note that three structures of 100 km or larger in diameter are known; however, none of them is close in age to the K-T boundary at 65 m.y. Approximately 50% of all craters are less than 10 km in diameter. While this inventory of craters contains a few older than 250 m.y., the crater production rates indicated are those calculated for the past 250 m.y.

crater production rates can be calculated (Neukum and Wise, 1976; Shoemaker, 1984). (3) Present-day astronomical telescope observations measure the frequency and orbits of small objects, generally less than 10 km in diameter, in the inner solar system. The collision probability of these objects with Earth may be computed (Wetherill and Shoemaker, 1982; Shoemaker, 1983). While the astronomical observations refer to the present, the terrestrial cratering record is an average over the last few hundred m.y. and the lunar cratering data integrate over the last 3 to 3.5 b.y., the age of lunar basaltic volcanism. Nevertheless, collision rates based on these different approaches agree to better than an order of magnitude, if not better than a factor of 5. Therefore, the probability for large-scale collisions on Earth may be calculated with considerable confidence (see Fig. 2). Large-scale collisions appear unavoidable in Earth's geologic record.

SYNCHRONEITY OF MASS EXTINCTIONS AND CRATERING RATE

From the above, it seems clear that the K-T collision was by no means unique. The question therefore arises whether mass extinctions may be associated with collisions in general. The search for siderophile element enhancement in paleontological boundary layers has just begun. Ganapathy (1982) found such an enhancement for the Eocene/Oligocene extinctions some 34 m.y. ago. This evidence, however, is tentative and lacks confirmation on a worldwide scale. Searches at other paleontological boundaries were unsuccessful to date.

A different perspective, however, provides incentive to consider a general correlation between impact and mass extinctions. Raup and Sepkoski (1984) subjected the marine mass extinction record for the last 250 m.y. to statistical time series tests. They found distinct periodicity of about 27 m.y. in the occurrence of mass extinctions. Depending on the approach and on various weighting factors, others have subsequently arrived at similar periodicities ranging from 26 to 32 m.y. Simultaneously, the terrestrial cratering record was subjected to time series tests by Alvarez and Muller (1984) and by Rampino and Stothers (1984a,b). Surprisingly, maxima in the terrestrial cratering record seem to reveal similar periodicity of 30 m.y. Importantly, both cycles appear to be approximately in phase (see Fig. 3).

If we accept a causal link between mass extinctions, cosmic collisions, and a period of some 30 m.y. as dictated by the biological record, we must identify mechanisms that gravitationally perturb small solar system objects on such time scales to become Earth-crossing objects. The most populous objects are comets. They reside at the outer fringes of the solar system, and their number is estimated to be 10^{13}. The existence of comets in the inner solar system is proof that comets can be gravitationally deflected into orbits that come close to the sun and thus Earth (Weissmann, 1982). Periodic character of this process requires a massive object that disturbs the comet reservoir on a cyclic schedule. Such an object could be a small companion star of the sun itself (Whitmire and Jackson, 1984; Davis *et al.*, 1984), but a more probable cause is the passage of our solar system through the galactic plane, where encounters with gigantic and massive molecular clouds seem unavoidable (Schwartz and James, 1984). In principle, both suggestions appear valid,

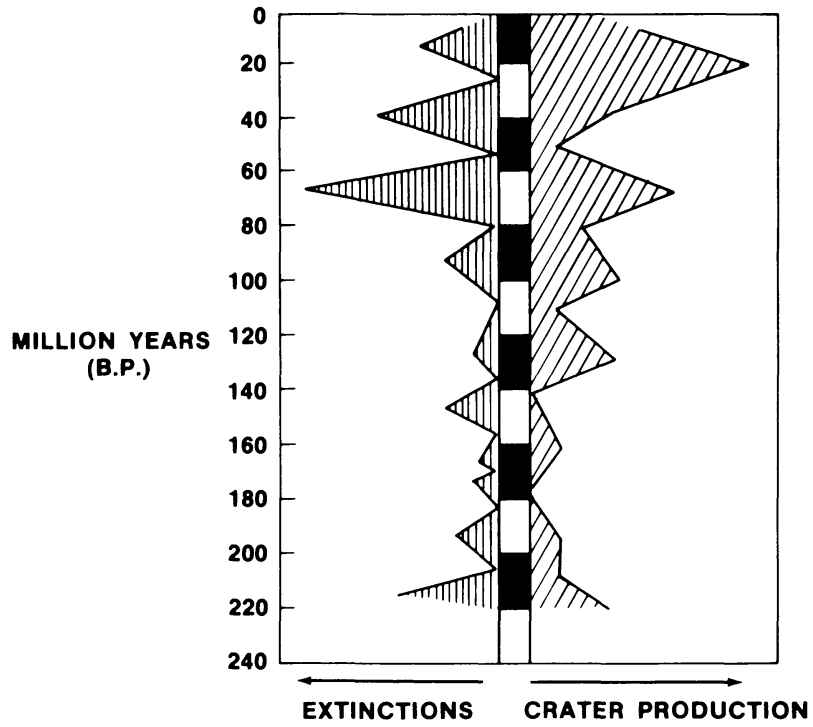

Figure 3. Periods of pronounced mass extinctions for the last 250 m.y. with peak height reflecting their magnitude (left-hand side; after Raup and Sepkoski, 1984). The right-hand side depicts maxima in the crater production rate as assembled by the author using the tabulation of Grieve (1982) for 35 crater ages (isotope ages and a few stratigraphic ages) of craters larger than 5 km.

although passage through the galactic plane is probably the more likely and less *ad hoc* mechanism (Clube and Napier, 1984; Hills, 1984; Torbett and Smoluchowski, 1984).

According to the above, statistical time series tests seem to yield a positive correlation between impact and mass extinctions. Plausible, if not compelling, mechanisms that may drive this 30 m.y. cycle appear to exist. Nevertheless, these generalizations are severely questioned by many.

Major criticism relates to the small number of terrestrial craters that can be subjected to statistical time series. Most craters are badly eroded. In many cases, erosion has totally removed the surficial impact melts. It is such melts, however, that are the choice material in isotopic age dating. Without impact melts, a crater formation time can generally not be determined exactly. Isotope ages (± 1%) exist for some 25 craters only. The formation times of all other craters are bracketed—at best—by stratigraphic evidence. The crater must be younger than its target rocks and older than the first sediments that backfilled the cavity. This stratigraphic bracketing yields on occasion relatively good ages, but it results mostly in age estimates that are useless for incorporation in the statistical time series tests.

Additionally, crater size must be considered. A 1 km diameter crater will certainly not produce a global catastrophy. Nevertheless, relatively small craters could be the sole survivors of impactor swarms, also referred to as "showers." The impact scars of the more massive shower members may have been totally obliterated by erosion or sedimentation, or the "large" impact may have happened in the ocean.

This brings us to a crucial point: The terrestrial cratering record is incompletely preserved as summarized by Grieve (1982) and Grieve *et al.* (1985). Erosion and sedimentation rates on Earth are highly variable on regional if not local scales. Craters

may be obliterated by surface erosion; they may also be buried by sedimentation. Otherwise identical craters are subject to vastly different mechanisms and rates of erosion (or burial) depending on local, geological environment. For example, one of the best preserved large craters on Earth, the Ries Crater in Germany, 26 km in diameter, formed 15 m.y. ago and therefore at approximately the time when the Colorado River started (!) to carve the Grand Canyon. Some 1500 m of rocks were removed at Grand Canyon and yet the uppermost meters of ejecta are still preserved at the Ries. Due to fortuitous circumstances, the Ries was completely buried immediately after formation by 200 m of sediments; erosion and removal of these sediments started a few million years ago and has presently proceeded to a stage where the crater is being exhumed again. Thus, highly variable erosion and sedimentation environments rendered a fair number of craters unrecognizable during the past 200 m.y., and the terrestrial cratering record is incomplete.

From the above it follows that the terrestrial crater population may represent a subset of "survivors" only. For the purposes of statistical treatment, the formation ages of this limited sample may be biased. Second, choices have to be made as to which crater formation ages are appropriate, possibly introducing additional bias. Third, a minimum crater size applicable to the problem must be defined; the cut-off diameters used in the above time series tests are typically 5 or 10 km. Thus, considerable high-grading of the terrestrial crater population is required for time series analysis. Additional bias may be introduced in selecting this limited subset. Depending on personal preference and judgment, only some 20–40 craters may be suitable. This is a small statistical sample; inclusion or exclusion of a few craters may have significant effects as demonstrated by Grieve *et al.* (1985). Within permissible ranges of geological weighting factors, Grieve *et al.* defined a variety of sample sets and subjected them to time series tests. They obtained periodicities ranging from 18–30 m.y. Also, 30 m.y. cycles resulted that may or may not be in phase with the mass extinctions. Indeed, the possibility that the cratering ages are entirely random cannot be positively excluded.

As a consequence, statistical time series tests of the terrestrial cratering record yield ambigous results. Statistical analyses of the small number of craters is simply not robust enough to demonstrate synchroneity with the mass extinctions. By the same token, none of the analyses demonstrate that such a correlation is inconsistent with the (limited) evidence at hand. It is therefore permissible to adopt a general, causal relationship between impact and mass extinction as a viable working hypothesis. We must subject this hypothesis to more rigorous tests.

Kyte (1984) analyzed chemically a deep ocean drill core that represented an essentially complete sedimentary record for the past 70 m.y. Enrichment of siderophile elements was only observed at the K–T boundary. Thus, this conceptually elegant test failed in detecting periodicity in Earth's collision record.

The most obvious solution for this important debate is a statistically sound data base. More craters need to be dated with the prerequisite precision. As stated above, isotope geochronology can only be performed on impact melts. Most terrestrial craters containing such melts are dated and already part of the data set. Some additional craters may be discovered, no doubt, as our abilities to recognize these impact scars improve; many of these new discoveries, however, will identify relatively eroded structures. Most

well preserved structures that yield datable impact melts are probably part of the present data set because they are recognized with relative ease. Thus, in the author's intuitive view, the number of well dated terrestrial craters may be doubled, with luck.

THE LUNAR CRATERING RECORD

Within the solar system, Earth and its Moon occupy essentially one location; both bodies were subject to the same bombardment history. As a consequence, formation

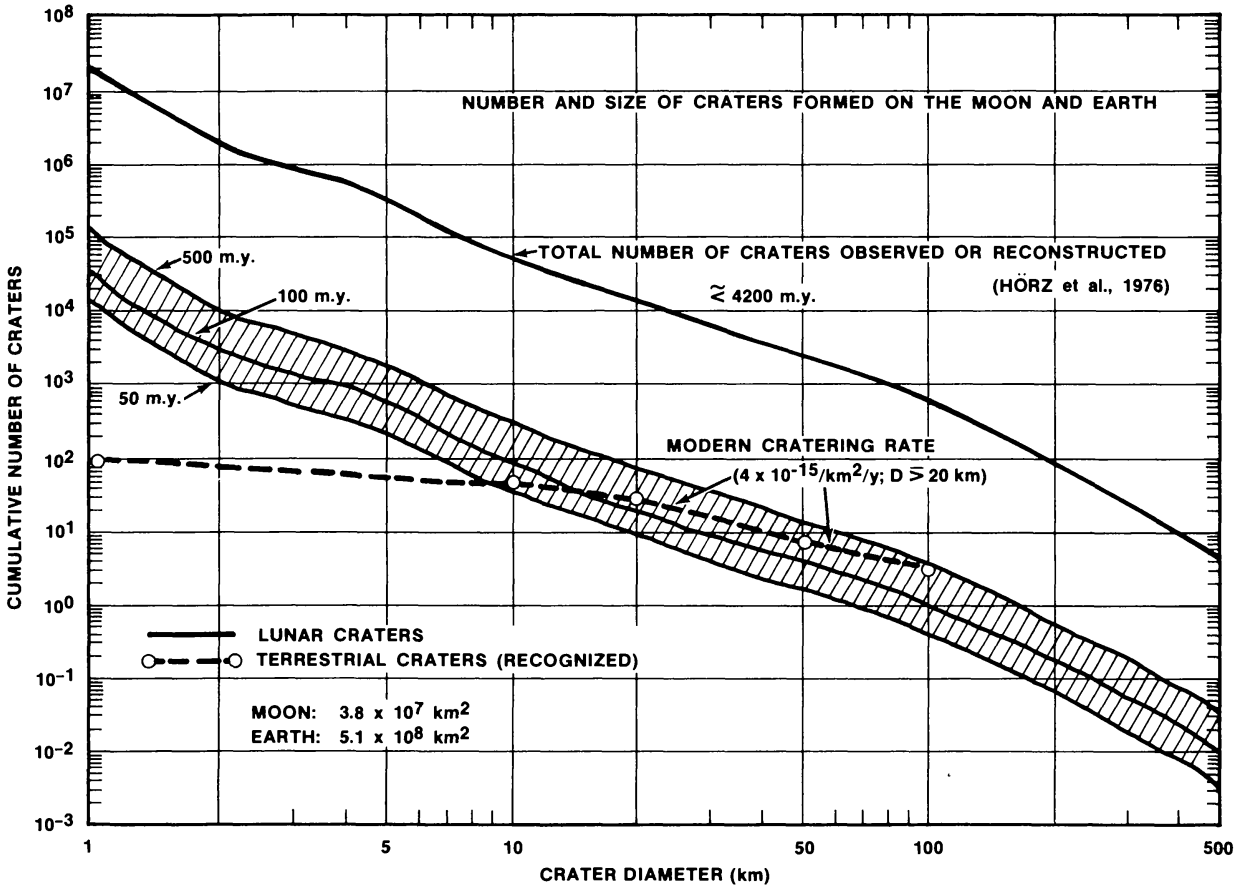

Figure 4. Observed and calculated cumulative size frequencies of lunar craters 1 to 500 km diameter. The uppermost curve depicts the observable frequency for craters >50 km formed during the past 4 b.y.; however, the absolute number of craters <50 km are computed values (Hörz et al., 1976). These calculations are based on the well documented size frequency distribution of small craters on relatively young surfaces. The numbers of craters above a specific size that were produced during the past 500, 100, and 50 m.y.—bracketed by the hatched band—were calculated with Grieve's "modern" crater production rate for craters >20 km and assuming that the size frequency distribution of lunar craters remained constant through geologic time. The latter is a reasonable assumption; if anything, the size distribution of <50 km diameter craters is better known than that of very large craters for recent geological history. The terrestrial crater size distribution is depicted also; it becomes relatively flat at small crater sizes, indicating that a fair number of "small" craters were obliterated or have not yet been recognized on Earth. The absolute crater numbers refer to the entire surface areas of the Moon and Earth. The figure may therefore be used to estimate how many square kilometers need to be covered during manned exploration to sample a specific number of craters above a specific diameter.

times of lunar collisions may be mixed with terrestrial crater ages in the statistical time series tests.

The heavily cratered lunar surface attests to a much better preservation of craters than on Earth. The Moon is geologically less active and lacks efficient erosive agents, such as water. The total number of craters larger than 100 km in diameter is 700 (as illustrated in Fig. 4). However, most large craters are remnants of an early bombardment phase with unusually high impactor fluxes (Wetherill, 1975; Neukum and Wise, 1976; Soderblom, 1977). The number of craters produced during the last 50, 100, and 500 m.y. is indicated in Fig. 4. Accordingly, a 100-km-diameter crater was formed once every 100 m.y., during which time the Moon also retained some 5 craters larger than 50 km or approximately 100 craters larger than 10 km. All of these craters are exceptionally well preserved by terrestrial standards and should yield proper impact melt samples for isotope geochronology. If craters as small as 5 km in diameter were included, the number of datable collisions could reach the hundreds. Clearly, the "modern" lunar cratering record has the potential to place any time series tests on statistically firm footing.

Figure 4 depicts the crater population of the entire Moon to illustrate that exploration, sampling, and ultimate dating of, for instance 100 lunar craters, requires systematic coverage of large fractions of the lunar surface. Long range sampling mobility of about 1000 km radius is necessary. Clearly, this amounts to a major undertaking, although sample acquisition spread over a 10 year period, possibly even longer, appears acceptable. Such time scales may be short compared to the efforts required in expanding the terrestrial cratering record to an adequate degree. Moreover, terrestrial studies include the risk of being only partly successful.

SUMMARY

In conclusion, synchroneity and causal relationship between mass extinctions and collisional events may be accepted as a viable working hypothesis, but not more. The hypothesis must be subjected to additional and more diagnostic tests. The most direct test demands that more craters are dated to a precision measured in small fractions of the proposed period, *i.e.*, to a precision of ± 5 m.y. The number of datable terrestrial craters will always remain small compared to the inventory of relatively recent lunar craters. There is considerable doubt whether the terrestrial cratering record can ever be reconstructed with enough confidence to accept or reject a temporal and causal relationship between mass extinctions and collisions on Earth. A return to the Moon and extensive lunar exploration may therefore hold crucial clues in our quest to better understand how life evolved on Earth.

REFERENCES

Alvarez L. W., Alvarez W., Asaro F., and Michel H. V. (1980) Extraterrestrial cause for the Cretaceous-Tertiary extinction. *Science, 208*, 1095–1108.

Alvarez W., Alvarez L. W., Asaro F., and Michel H. V. (1982) Current status of the impact theory for the terminal Cretaceous extinctions. *Geol. Soc. Am. Spec. Pap. 190*, 305–316.

Alvarez W. and Muller R. A. (1984) Evidence from crater ages for periodic impacts on the Earth. *Nature, 308*, 718–720.

Bohor B. F., Foord E. F., Modreski J., and Triplehorn D. M. (1984) Mineralogic evidence for an impact event at the Cretaceous-Tertiary boundary. *Science, 224*, 867–869.

Clube S. V. M. and Napier V. M. (1984) Terrestrial catastrophism-Nemesis or galaxy. *Nature, 311*, 635–636.

Davis M., Hutt P., and Muller R. A. (1984) Extinction of species by periodic comet showers. *Nature, 308*, 715–717.

Ganapathy R. (1982) Evidence for a major meteorite impact on the Earth 34 million years ago: implications on the origin of North American tektites and Eocene extinctions. *Geol. Soc. Am. Spec. Pap. 190*, 513–516.

Gault D. E. and Sonett C. P. (1982) Laboratory simulation of pelagic asteroidal impact: atmosphere injection, benthic topography, and the surface radiation field. *Geol. Soc. Am. Spec. Pap. 190*, 69–92.

Grieve R. A. F. (1982) The record of impact on Earth: implications for major Cretaceous/Tertiary impact event. *Geol. Soc. Am. Spec. Pap. 190*, 25–38.

Grieve R. A. F., Sharpton V. L., Goodacre A. K., and Garvin J. B. (1985) A perspective on the evidence for periodic cometary impacts on Earth. *Earth Planet. Sci. Lett.* In press.

Hills J. G. (1984) Dynamical constraints on the mass and perihelion distance of Nemesis and the stability of its orbit. *Nature, 311*, 636–638.

Holland H. D. and Trendall A. F. (editors) (1984) *Patterns of Change in Earth Evolution*. Springer Verlag, Berlin. 431 pp.

Hörz F., Gibbons R. V., Hill R. E., and Gault D. E. (1976) Large-scale cratering of the lunar highlands: some Monte Carlo model considerations. *Proc. Lunar Sci. Conf. 7th*, pp. 2932–2945.

Hsu K. J., McKenzie J. A., and He Q. X. (1982) Terminal Cretaceous environmental and evolutionary changes. *Geol. Soc. Am. Spec. Pap. 190*, 317–328.

Kyte F. (1984) Cenozoic iridium sedimentation: Death of Nemesis? (abstract) *Geological Society of America Abstract with Programs, 16*, 567.

Neukum G. and Wise D. U. (1976) Mars: a standard crater curve and possible new time scale. *Science, 194*, 1381–1387.

O'Keefe J. D. and Ahrens T. J. (1982) The interaction of the Cretaceous/Tertiary bolide with the atmosphere, ocean and the solid Earth. *Geol. Soc. Am. Spec. Pap. 190*, 103–120.

Rampino M. R. and Stothers R. B. (1984a) Terrestrial mass extinction, cometary impacts and the Sun's motion perpendicular to the galactic plane. *Nature, 308*, 709–712.

Rampino M. R. and Stothers R. B. (1984b) Geological rhythms and cometary impacts. *Science, 226*, 1427–1431.

Raup D. M. (1984) Evolutionary radiations and extinctions. In *Patterns of Change in Earth Evolution* (H. D. Holland and A. F. Trendall, eds.), pp. 5–15. Springer Verlag, Berlin.

Raup D. M. and Sepkoski J. J., Jr. (1984) Periodicity of extinctions in the geologic past. *Proc. Natl. Acad. Sci. U.S.A., 81*, 801–805.

Russell D. A. (1979) The enigma of the extinction of dinosaurs. *Annu. Rev. Earth Planet. Sci., 7*, 163–182.

Schwartz R. D. and James P. B. (1984) Periodic mass extinctions and the Sun's oscillation about the galactic plane. *Nature, 308*, 712–713.

Shoemaker E. M. (1983) Asteroid and comet bombardment of the Earth. *Annu. Rev. Earth Planet. Sci., 11*, 15–41.

Silver L. T. and Schultz P. H. (editors) (1982) Geological implications of impacts of large asteroids and comets on Earth. *Geol. Soc. Am. Spec. Pap. 190*, 528 pp.

Soderblom L. A. (1977) Historical variations in the density and distribution of impacting debris in the inner solar system: evidence from planetary imaging. In *Impact and Explosion Cratering* (D. J. Roddy, R. O. Pepin, and R. B. Merrill, eds.), pp. 629–633. Pergamon, New York.

Toon O. B. (1984) Sudden changes in atmospheric composition and climate. In *Patterns of Change in Earth Evolution* (H. D. Holland and A. F. Trendall, eds.), pp. 41–61. Springer Verlag, Berlin.

Torbett M. V. and Smoluchowski B. (1984) Orbital stability of the unseen solar companion linked to periodic extinction events. *Nature, 311*, 641–642.

Weissmann P. R. (1982) Terrestrial impact rates for long and short-period comets. *Geol. Soc. Am. Spec. Pap. 190*, 15–24.

Wetherill G. W. (1975) Late heavy bombardment of the moon and the terrestrial planets. *Proc. Lunar Sci. Conf. 6th*, pp. 1539–1561.

Wetherill G. W. and Shoemaker E. M. (1982) Collision of astronomically observable bodies with the Earth. *Geol. Soc. Am. Spec. Pap. 190*, 1–14.

Whitmire D. P. and Jackson A. A. (1984) Are periodic mass extinctions driven by a distant solar companion? *Nature, 308*, 713–715.

(This Page Intentionally Left Blank)